

An introduction to analytic number theory

This expository article aims to convey a taste of analytic number theory via the study of two problems centering around the Riemann zeta function, namely the prime number theorem and the Dirichlet theorem on primes in arithmetic progressions. We shall assume some knowledge on the reader of complex analysis (for instance, infinite products, contour integrals and analytic continuation) and algebra (limited to the language of group theory).

1 Riemann zeta function

The Riemann zeta function is defined by

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

for all complex numbers s with $\operatorname{Re} s > 1$. Its relation to number theory is already seen in the following **Euler's product formula**, which relates it to one of the most primitive notions in number theory, namely prime numbers:

Lemma 1. *For all real numbers $s > 1$,*

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}}$$

where in the product p runs through all prime numbers.

Proof. Every positive integer n can be written uniquely as the product of prime numbers. Hence if p_1, p_2, \dots is a listing of all prime numbers, then $p_1^{e_1} p_2^{e_2} \dots$ runs through all positive integers when the exponents e_1, e_2, \dots independently run through all non-negative integers

with all but finitely many being non-zero. As a result,

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{n^s} &= \lim_{k \rightarrow \infty} \sum_{e_1 \geq 0} \sum_{e_2 \geq 0} \cdots \sum_{e_k \geq 0} \frac{1}{(p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k})^s} \\
&= \lim_{k \rightarrow \infty} \left(\sum_{e_1 \geq 0} \frac{1}{p_1^{s e_1}} \right) \left(\sum_{e_2 \geq 0} \frac{1}{p_2^{s e_2}} \right) \cdots \left(\sum_{e_k \geq 0} \frac{1}{p_k^{s e_k}} \right) \\
&= \lim_{k \rightarrow \infty} \left(\frac{1}{1 - p_1^{-s}} \right) \left(\frac{1}{1 - p_2^{-s}} \right) \cdots \left(\frac{1}{1 - p_k^{-s}} \right) \\
&= \prod_p \frac{1}{1 - p^{-s}}
\end{aligned}$$

as desired, at least formally. It is easy to justify the manipulations of the infinite series and products here, which we leave to the reader. \square

Such Euler product formula are very useful in general. There are various generalizations of the above product formula to other zeta functions and their generalizations, called L -functions (see also Section 2.3 below). They relate the analytic object, namely the zeta functions, to arithmetic information by means of a product involving one term at each prime. By a similar token, they are very useful in the study of geometric objects (like elliptic curves): this belongs to another huge area of mathematics that we shall not take up here.

As a warm-up we shall use the Euler product formula to prove the well-known fact that there are infinitely many primes. Indeed we shall prove a stronger quantitative assertion, namely

Theorem 2.

$$\sum_p \frac{1}{p} = \infty$$

where the sum runs over all prime numbers p .

Proof. By the Euler product formula, at least for real values of $s > 1$,

$$\log \zeta(s) = - \sum_p \log(1 - p^{-s})$$

where again the sum runs through all prime numbers p . The Taylor series expansion of the logarithm says $-\log(1 - x) = x + O(x^2)$ for x close to zero, so for $s > 1$ but close to 1,

$$\log \zeta(s) = \sum_p (p^{-s} + O(p^{-2s})) = \sum_p \frac{1}{p^s} + O(1).$$

Here $O(x^2)$ is an error term that upon division by x^2 remains bounded as $x \rightarrow 0$, and $O(1)$ is a term that remains bounded as $s \rightarrow 1^+$. The last equality in the above equation holds because for every prime p , we can choose the same constant $C > 0$ so that each of term $O(p^{-2s})$ is bounded by Cp^{-2} for s close to 1^+ , and $\sum p^{-2} \leq \sum n^{-2}$ converges. We will use this fact again implicitly later. Letting $s \rightarrow 1^+$, we get

$$\sum_p \frac{1}{p} = \infty$$

if we know $\lim_{s \rightarrow 1^+} \zeta(s) = \infty$. This is indeed the case, since the series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. \square

In fact it is readily seen from the definition that ζ is an analytic function of s for $\text{Re } s > 1$, and that the Euler product formula continues to hold in that range of s . We shall see later that ζ can be analytically continued so that it has a simple pole at $s = 1$. In other words, we shall see that

$$\lim_{s \rightarrow 1} \zeta(s) = \infty,$$

which is stronger than the assertion $\lim_{s \rightarrow 1^+} \zeta(s) = \infty$ that we have just used above.

The above proof of the infinitude of primes can be modified to prove the following **Dirichlet's theorem** on the infinitude of primes in any (non-trivial) arithmetic progressions:

Theorem 3. *In any arithmetic progression $\{qk + l : k \in \mathbb{N}\}$, where $q, l \in \mathbb{N}$ are relatively prime, there exists infinitely many primes.*

Clearly for the arithmetic progression $\{qk + l\}$ to contain infinitely many primes, q and l have to be relatively prime. The remarkable discovery of Dirichlet is that this is also sufficient. The goal of section 2 is to prove this theorem; note that the case $q = 4$, $l = 3$ says that there are infinitely many primes of the form $4k + 3$, which is well-known and elementary.

The study of the zeta function shall also lead to the famous **prime number theorem**:

Theorem 4. *Let $\pi(x)$ be the number of prime numbers not exceeding x . Then*

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

as $x \rightarrow \infty$, i.e. $\lim_{x \rightarrow \infty} \left(\frac{\pi(x)}{\frac{x}{\log x}} \right) = 1$.

We shall prove this in section 3 using complex analysis.

The two sections are basically logically independent and can be read in any order. We hope they can give the reader some idea how analysis helps one solve outstanding questions in number theory.

2 Dirichlet's theorem on primes in arithmetic progressions

In this section we fix two positive integers q and l that are relatively prime. To prove Dirichlet's theorem, following the proof of the infinitude of primes given above, we shall prove

$$(1) \quad \sum_{p \equiv l \pmod{q}} \frac{1}{p} = \infty,$$

where the sum is only over those *primes* p that are congruent to $l \pmod{q}$. (Hereafter the symbol p is reserved for *prime* numbers.) To do so we shall need some Fourier analysis on finite abelian groups, to which we now turn.

2.1 Fourier analysis on finite abelian groups

Let G be a *finite abelian* group. In application we shall take $G = (\mathbb{Z}/q\mathbb{Z})^\times$, the multiplicative group of units in the ring $\mathbb{Z}/q\mathbb{Z}$. In other words, G will be the multiplicative group of integers modulo q that are relatively prime to q . Since the results below are fairly general, we shall state it for an arbitrary finite abelian group G .

Definition. A *character* χ of G is a group homomorphism $\chi: G \rightarrow U(1)$, where $U(1)$ is the multiplicative group of complex numbers whose moduli are 1. The group of all characters of G (under pointwise multiplication) will be denoted by \hat{G} .

For example, there is the trivial character χ_0 , defined by $\chi_0(a) = 1$ for all $a \in G$. It is often useful to consider this trivial character separately, as we will do in the proof of the Dirichlet's theorem.

For readers who know some representation theory, the χ 's in the above definition are precisely the characters of the complex irreducible representations of G . The key fact that we shall need is the following version of Peter-Weyl theorem, which says that the characters of a finite abelian group G can be used to obtain a series expansion of any complex-valued functions on G that resembles the ordinary Fourier series of L^2 functions on the unit circle:

Theorem 5. Let H be the space of complex-valued functions on G , equipped with the inner product

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{a \in G} f(a) \overline{g(a)}$$

for any $f, g \in H$. Then \hat{G} is an orthonormal basis of H . In particular, for $f \in H$ we define for all $\chi \in \hat{G}$

$$\hat{f}(\chi) = \langle f, \chi \rangle = \frac{1}{|G|} \sum_{a \in G} f(a) \overline{\chi(a)};$$

then

$$f(a) = \sum_{\chi \in \hat{G}} \hat{f}(\chi) \chi(a)$$

for all $a \in G$.

Again, the readers familiar with group representations will recognize $\frac{1}{|G|} \sum_{a \in G} \delta_a$ as the Haar measure on G , and H as the Hilbert space $L^2(G)$ with respect to this measure. Hence the above theorem is a special case of the Peter-Weyl theorem in the setting of finite abelian groups. Since for finite abelian groups G the above theorem admits a much simpler proof, for the convenience of readers who do not know representation theory, we shall give the simple proof in this case below (see section 2.5).

Now we specialize to the case where $G = (\mathbb{Z}/q\mathbb{Z})^\times$: we shall take f to be the function on $(\mathbb{Z}/q\mathbb{Z})^\times$ defined by

$$f(a) = \delta_l(a) = \begin{cases} 1 & \text{if } a = l \pmod{q} \\ 0 & \text{for all other } a \in (\mathbb{Z}/q\mathbb{Z})^\times \end{cases},$$

where l is the integer that is fixed at the beginning of this section that is relatively prime to q . Then since $|G| = \phi(q)$, where ϕ is the Euler phi function,

$$\hat{f}(\chi) = \frac{1}{\phi(q)} \overline{\chi(l)}$$

for all characters χ of $(\mathbb{Z}/q\mathbb{Z})^\times$ and thus

$$(2) \quad \delta_l(a) = \frac{1}{\phi(q)} \sum_{\chi} \overline{\chi(l)} \chi(a)$$

for all $a \in (\mathbb{Z}/q\mathbb{Z})^\times$. Note that both $\delta_l(a)$ and $\chi(a)$ above were only defined for $a \in (\mathbb{Z}/q\mathbb{Z})^\times$, and the equality was asserted only for such a 's. Below we shall extend the domain of definitions of the δ_l 's and the χ 's to the whole $\mathbb{Z}/q\mathbb{Z}$ and see that (2) continues to hold. This will lead us to the concept of Dirichlet characters.

2.2 Dirichlet characters

Our goal in this section is to extend (2) so that it holds for all $a \in \mathbb{Z}/q\mathbb{Z}$, and indicate how this is related to the proof of (1).

The extension of (2) is trivial; one simply extend the definitions of δ_l and the χ 's by setting them to vanish for the non-units in $\mathbb{Z}/q\mathbb{Z}$, i.e. we set

$$\delta_l(a) = 0 \quad \text{for } a \notin (\mathbb{Z}/q\mathbb{Z})^\times,$$

and for all characters χ of $(\mathbb{Z}/q\mathbb{Z})^\times$, we set

$$\chi(a) = 0 \quad \text{for } a \notin (\mathbb{Z}/q\mathbb{Z})^\times.$$

The extended χ 's are called the Dirichlet characters modulo q ; there are $\phi(q)$ of them. The extension of the trivial character is still denoted by χ_0 ; so

$$\chi_0(a) = \begin{cases} 1 & \text{if } a \in (\mathbb{Z}/q\mathbb{Z})^\times \\ 0 & \text{if } a \notin (\mathbb{Z}/q\mathbb{Z})^\times. \end{cases}$$

Now (2) obviously continues to hold on all of $\mathbb{Z}/q\mathbb{Z}$, i.e.

$$(3) \quad \delta_l(a) = \frac{1}{\phi(q)} \sum_{\chi} \overline{\chi(l)} \chi(a)$$

for all $a \in \mathbb{Z}/q\mathbb{Z}$, where the sum is still over all characters χ of $(\mathbb{Z}/q\mathbb{Z})^\times$. By abuse of notation, for $a \in \mathbb{Z}$ we shall also denote by $\delta_l(a)$ and $\chi(a)$ the values of δ_l and χ evaluated at the equivalent class a modulo q , and clearly (3) continues to hold for all $a \in \mathbb{Z}$. We shall also call such χ Dirichlet characters modulo q .

With this notation in mind, by (3) we can write the sum in (1) as

$$\begin{aligned} \lim_{s \rightarrow 1^+} \sum_{p \equiv l \pmod{q}} \frac{1}{p^s} &= \lim_{s \rightarrow 1^+} \sum_p \frac{\delta_l(p)}{p^s} \\ &= \lim_{s \rightarrow 1^+} \frac{1}{\phi(q)} \sum_{\chi} \overline{\chi(l)} \sum_p \frac{\chi(p)}{p^s}; \end{aligned}$$

here once again \sum_p denotes sum over all primes p . Note that we have put in the power $s > 1$ to make sure that the sums converge absolutely so that it can be rearranged. Separating the term involving the trivial character χ_0 , we get

$$\sum_{p \equiv l \pmod{q}} \frac{1}{p^s} = \frac{1}{\phi(q)} \sum_{p \not\equiv q} \frac{1}{p^s} + \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0} \overline{\chi(l)} \sum_p \frac{\chi(p)}{p^s}.$$

As $s \rightarrow 1^+$, the first term diverges to infinity by Theorem 2. If we can show that the second term remains bounded as $s \rightarrow 1^+$, then it follows that (1) holds, and Dirichlet's theorem is proved. In fact we shall prove the following:

Theorem 6. *If χ is a non-trivial character of $(\mathbb{Z}/q\mathbb{Z})^\times$, then*

$$\sum_p \frac{\chi(p)}{p^s}$$

remains bounded as $s \rightarrow 1^+$, where the sum is over all primes p .

This is plausible because if χ is a non-trivial Dirichlet character modulo q , then χ takes on more than one value on the unit circle $U(1)$, and thus in the sum $\sum_p \frac{\chi(p)}{p^s}$, it is likely that there are cancellations that were not present in the sum $\sum_p \frac{1}{p^s}$. This extra cancellation will lead to the desired boundedness of the sum $\sum_p \frac{\chi(p)}{p^s}$ as $s \rightarrow 1^+$.

In the next section we prove this theorem, using the Dirichlet L -functions.

2.3 Dirichlet L -functions

To prove Theorem 6, we imitate the proof of Theorem 2. There we made use of the Euler product formula. It turns out that there is a version of the Euler product formula involving the Dirichlet characters, namely

Lemma 7. *For any real numbers $s > 1$ and any Dirichlet characters χ modulo q , we have*

$$(4) \quad \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \frac{1}{1 - \chi(p)p^{-s}}$$

where in the product p runs through all prime numbers.

The interest of the lemma is of course in the case when χ is a non-trivial Dirichlet character; this is the case that we will apply the lemma below. The left side of the above lemma is usually called the Dirichlet L -function. It is usually denoted $L(s, \chi)$, i.e.

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

The sum is convergent for real numbers $s > 1$. The proof of Lemma 7 makes use of the multiplicative property of the Dirichlet characters, namely

$$\chi(ab) = \chi(a)\chi(b)$$

for all $a, b \in \mathbb{Z}$, and is otherwise along the same line as that of the ordinary Euler's product formula. We omit its proof.

In fact both sides of the above product formula have natural analytic continuation to the half-space $\{\operatorname{Re} s > 1\}$ in the complex plane, and the equality continues to hold there. Since we shall not make use of that, we shall be contented with the above 'real' version of the product formula.

Proof of Theorem 6. Following the proof of Theorem 2, we take the logarithm of both sides of (4). Now both sides of (4) involve complex numbers since they involve non-trivial Dirichlet characters. As a result we must take the complex logarithm instead of the real one. The result is, for appropriate branches of the complex logarithm, that

$$(5) \quad \log L(s, \chi) = - \sum_p \log \left(1 - \frac{\chi(p)}{p^s} \right)$$

for all $s > 1$. We shall be more careful about how the branches are chosen, but arguing formally for the moment, assuming that we have the appropriate power series expansion of the right hand side of the above identity, we get

$$(6) \quad \log L(s, \chi) = \sum_p \frac{\chi(p)}{p^s} + \sum_p O(p^{-2s}) = \sum_p \frac{\chi(p)}{p^s} + O(1)$$

for $s > 1$. Hence the desired boundedness of the sum

$$\sum_p \frac{\chi(p)}{p^s}$$

as $s \rightarrow 1^+$ follows from the following theorem:

Theorem 8. *If χ is a non-trivial Dirichlet character modulo q , then $L(s, \chi)$ continues to a continuously differentiable function for $s \in (0, \infty)$, and*

$$L(1, \chi) \neq 0.$$

We shall prove this in the next section.

To justify the formal argument given above, first we choose an appropriate branch of logarithm for the two sides of (5). For the left hand side of (5), we define, for $s > 1$ and $\chi \neq \chi_0$,

$$\log L(s, \chi) = - \int_s^\infty \frac{L'(t, \chi)}{L(t, \chi)} dt;$$

this is legitimate since $L(t, \chi) = 1 + O(e^{-ct})$ and $L'(t, \chi) = O(e^{-ct})$ as $t \rightarrow \infty$. For the right hand side of (5), note that $\chi(p)p^{-s}$ has absolute value less than 1 for all primes p and all $s > 1$; thus one can simply use the principal branch of logarithm for the terms on the right hand side of (5). In particular, the power series representation is valid:

$$\log \left(\frac{1}{1 - \chi(p)p^{-s}} \right) = - \sum_p \frac{\chi(p)}{p^s} + O(p^{-2s}),$$

which incidentally justifies the equality (6). To show that the equality in (5) holds, note that since we are just taking possibly different branches of logarithm, the two sides of the

equation differ at most by an integral multiple of $2\pi i$ that may depend on s . Write the difference of the two sides as $2\pi i M(s)$ where $M(s)$ is an integer for all s . Then M is a continuous function of s that takes only integer values, and thus M is constant. Letting $s \rightarrow \infty$, M is identically zero. Hence the two sides of (5) are equal.

It thus remains to prove Theorem 8, which is the goal of the next section. □

2.4 Non-vanishing of $L(1, \chi)$

Let χ be a non-trivial Dirichlet character modulo q , fixed throughout this section.

Proposition 9. $L(s, \chi)$ can be analytically continued to $\operatorname{Re} s > 0$.

Proof. We use summation by parts. Let $s_n = \sum_{k=1}^n \chi(k)$ (and $s_0 = 0$). Then for any positive integer N ,

$$(7) \quad \sum_{n=1}^N \frac{\chi(n)}{n^s} = \sum_{n=1}^N \frac{s_n - s_{n-1}}{n^s} = \sum_{n=1}^{N-1} s_n \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right) + \frac{s_N}{N^s}.$$

Now s_n is a bounded sequence, since

$$\sum_{k=1}^q \chi(k) = \phi(q) \langle \chi, \chi_0 \rangle = 0$$

by the orthogonality of χ and χ_0 asserted in Theorem 5; the same holds as long as the sum is over a complete residue class modulo q . It follows that $|s_n| \leq q$ for all $n \in \mathbb{N}$. Next for $n \in \mathbb{N}$ and s with $\operatorname{Re} s > 0$,

$$\left| \frac{1}{n^s} - \frac{1}{(n+1)^s} \right| \leq \frac{|s|}{n^{\operatorname{Re} s + 1}}.$$

Hence

$$\sum_{n=1}^{N-1} s_n \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right)$$

converges absolutely and locally uniformly for $\operatorname{Re} s > 0$ as $N \rightarrow \infty$, and this gives via (7) the desired analytic continuation of $L(s, \chi)$ to $\operatorname{Re} s > 0$. □

Proposition 10. $L(1, \chi) \neq 0$.

Proof. We consider two cases: the case where χ takes on some (non-real) complex values, so that $\bar{\chi}$ is a (non-trivial) Dirichlet character different from χ , and the case where χ takes on only real values, i.e. when $\bar{\chi} = \chi$.

Case 1: $\bar{\chi} \neq \chi$

We need the following lemma:

Lemma 11. *If $s > 1$ then*

$$(8) \quad \prod_{\chi} L(s, \chi) \geq 1$$

where the product extends over all Dirichlet characters χ modulo q . (In particular the product is real.)

Proof of lemma. By the Euler product formula (5) for L -functions, if \log represents the principal branch of the complex logarithm, then for $s > 1$,

$$\begin{aligned} \prod_{\chi} L(s, \chi) &= \prod_{\chi} \exp \left(\sum_p \log \frac{1}{1 - \chi(p)p^{-s}} \right) \\ &= \exp \left(\sum_{\chi} \sum_p \sum_{k=1}^{\infty} \frac{1}{k} \frac{\chi(p)^k}{p^{sk}} \right) \\ &= \exp \left(\sum_p \sum_{k=1}^{\infty} \frac{1}{kp^{sk}} \sum_{\chi} \chi(p^k) \right) \\ &= \exp \left(\sum_p \sum_{k=1}^{\infty} \frac{\phi(q)\delta_1(p^k)}{kp^{sk}} \right) \geq 1 \end{aligned}$$

where the last equality follows from the Fourier series expansion (3) of the function δ_1 . \square

With the lemma we now prove $L(1, \chi) \neq 0$ for our fixed non-trivial complex Dirichlet character χ . In the product in (8), all but the trivial character contribute a factor that is holomorphic near $s = 1$. The only factor that blows up corresponds to the trivial character χ_0 , and since by (4),

$$L(s, \chi_0) = (1 - p_1^{-s})(1 - p_2^{-s}) \dots (1 - p_N^{-s})\zeta(s)$$

where p_1, p_2, \dots, p_N is a listing of all prime factors of q , it follows that the factor $L(s, \chi_0)$ can only blow up like a constant multiple of

$$\zeta(s) \sim \frac{1}{s-1}$$

near $s = 1$. (The fact that ζ has a simple pole at $s = 1$ was mentioned after the proof of Theorem 2 and will be proved in section 3.1 below.) Now if for our fixed non-trivial complex Dirichlet character, we have $L(1, \chi) = 0$, then $L(s, \chi)$ vanishes up to order at

least 1 near $s = 1$; the same happens to $\bar{\chi}$, since then $L(1, \bar{\chi}) = \overline{L(1, \chi)} = 0$. However $\bar{\chi}$ is a different (non-trivial) Dirichlet character modulo q from χ . Hence there are two different terms in the product (8), each vanishing up to order at least 1 near $s = 1$. This kills the only possible singularity of the product, namely the one corresponding to the term $L(s, \chi_0)$, and thus the product $\prod_{\chi} L(s, \chi)$ tends to 0 as $s \rightarrow 1^+$, contradicting (8). This shows that $L(1, \chi)$ cannot be zero, and concludes the proof in this case.

Case 2: $\bar{\chi} = \chi$

Since $\bar{\chi}$ is the same Dirichlet character as χ , the above argument would not work in this case. We shall take another route: we shall consider the sums

$$S_k = \sum_{mn \leq k} \frac{\chi(n)}{(mn)^{1/2}}$$

where k is a positive integer and the sum is over all ordered pairs (m, n) of positive integers whose product does not exceed k . The key is the following:

Lemma 12. *If χ is a real non-trivial Dirichlet character, then*

- (a) $S_k = 2k^{1/2}L(1, \chi) + O(1)$ as $k \rightarrow \infty$;
- (b) $S_k \geq c \log k$ for some constant $c > 0$.

It follows that $L(1, \chi)$ cannot be zero.

Proof of lemma. (a) The proof depends on two different ways of summing S_k . Note that the sum is over the lattice points (m, n) that lie in the first quadrant of the (m, n) plane under the hyperbola $mn = k$. Draw the (m, n) plane such that m represent the horizontal coordinate and n the vertical coordinate. First we sum ‘vertically and horizontally’: we write S_k as

$$S_k = \left(\sum_{1 \leq m < k^{1/2}} \sum_{k^{1/2} < n \leq k/m} + \sum_{1 \leq n \leq k^{1/2}} \sum_{1 \leq m \leq k/n} \right) \frac{\chi(n)}{(mn)^{1/2}} = I + II.$$

Now

$$(9) \quad I = \sum_{1 \leq m < k^{1/2}} \frac{1}{m^{1/2}} \sum_{k^{1/2} < n \leq k/m} \frac{\chi(n)}{n^{1/2}},$$

and to exploit the cancellation due to the factors $\chi(n)$ in the inner sum, we estimate the inner sum by summation by parts in a way similar to how we analytically continued

the L -functions. In fact again denoting $\sum_{k=1}^n \chi(k)$ by s_n , we have by the boundedness of $\{s_n\}$ that

$$\sum_{n=a}^b \frac{\chi(n)}{n^{1/2}} = \sum_{n=a}^{b-1} s_n \left(\frac{1}{n^{1/2}} - \frac{1}{(n+1)^{1/2}} \right) + O(a^{-1/2}) \leq C \sum_{n=a}^{b-1} \frac{1}{n^{3/2}} + O(a^{-1/2}).$$

However, the sum $\sum_{n=a}^{b-1} \frac{1}{n^{3/2}}$ can be approximated by the integral $\int_a^b \frac{1}{x^{3/2}} dx$, and the result is that $\sum_{n=a}^{b-1} \frac{1}{n^{3/2}} = O(a^{-1/2})$. Hence applying this to the inner sum in I , we get

$$I \leq Ck^{-1/4} \sum_{1 \leq m < k^{1/2}} \frac{1}{m^{1/2}}.$$

Once again we can use the estimate $\sum_{1 \leq m < k^{1/2}} \frac{1}{m^{1/2}} = O(k^{1/4})$ that we obtain by approximating the sum by the integral $\int_1^{k^{1/2}} \frac{1}{x^{1/2}} dx$. Hence

$$I = O(1)$$

is a bounded term as $k \rightarrow \infty$.

Next we estimate II : we write II as

$$II = \sum_{1 \leq n \leq k^{1/2}} \frac{\chi(n)}{n^{1/2}} \sum_{1 \leq m \leq k/n} \frac{1}{m^{1/2}}.$$

By comparing it to an integral, the inner sum is $2(k/n)^{1/2} + c + O((k/n)^{-1/2})$ for some constant c ; in fact

$$\begin{aligned} \left| \sum_{1 \leq m \leq b} \frac{1}{m^{1/2}} - \int_1^b \frac{1}{x^{1/2}} dx \right| &\leq \sum_{1 \leq m \leq [b]} \int_m^{m+1} \left(\frac{1}{m^{1/2}} - \frac{1}{x^{1/2}} \right) dx \\ &\leq \sum_{1 \leq m \leq [b]} \frac{1}{2m^{3/2}} \\ &= \sum_{m=1}^{\infty} \frac{1}{2m^{3/2}} - \sum_{m=[b]+1}^{\infty} \frac{1}{2m^{3/2}} \\ &= c + O(b^{-1/2}). \end{aligned}$$

Hence

$$\begin{aligned} II &= 2 \sum_{1 \leq n \leq k^{1/2}} \frac{\chi(n)}{n^{1/2}} \left(\frac{k}{n} \right)^{1/2} + c \sum_{1 \leq n \leq k^{1/2}} \frac{\chi(n)}{n^{1/2}} + O \left(\sum_{1 \leq n \leq k^{1/2}} \frac{\chi(n)}{n^{1/2}} \left(\frac{k}{n} \right)^{-1/2} \right) \\ &= 2k^{1/2} \sum_{1 \leq n \leq k^{1/2}} \frac{\chi(n)}{n} + c \sum_{1 \leq n \leq k^{1/2}} \frac{\chi(n)}{n^{1/2}} + O(1) \\ &= 2k^{1/2} L(1, \chi) + 2k^{1/2} \sum_{n > k^{1/2}} \frac{\chi(n)}{n} + c \sum_{1 \leq n \leq k^{1/2}} \frac{\chi(n)}{n^{1/2}} + O(1) \end{aligned}$$

The two sums can be estimated by first summing by parts and then approximating by integrals, in very much the same way that we tackled the inner sum in (9). The result is that the two terms remain bounded as $k \rightarrow \infty$, so together with the estimate for I we get

$$S_k = 2k^{1/2}L(1, \chi) + O(1)$$

as desired.

(b) This time we sum S_k by summing along the hyperbolas; we write

$$S_k = \sum_{j=1}^k \sum_{mn=j} \frac{\chi(n)}{(mn)^{1/2}} = \sum_{j=1}^k \frac{1}{j^{1/2}} A_j$$

where

$$A_j = \sum_{n|j} \chi(n).$$

However,

$$A_j \geq 0$$

for all j , and

$$A_j \geq 1$$

if j is a perfect square. This can be proved as follows: first, since χ is multiplicative, A_j as a function of j is also multiplicative. It follows that we only have to consider the case where $j = p^N$ is a prime power. Now if $j = p^N$ for some prime p , then since now χ is assumed to take only real values (i.e. 0, 1 and -1 since the values of χ lies on $U(1)$), we have

$$A_j = \sum_{i=0}^N \chi(p)^i = \begin{cases} N+1 & \text{if } \chi(p) = 1 \\ 1 & \text{if } \chi(p) = -1 \text{ and } N \text{ is even} \\ 0 & \text{if } \chi(p) = -1 \text{ and } N \text{ is odd} \\ 1 & \text{if } \chi(p) = 0, \text{ i.e. if } p \text{ divides } q \end{cases}.$$

In any case $A_j \geq 0$, and $A_j \geq 1$ if $j = p^N$ with N even. This proves the assertions about A_j . As a result,

$$S_k = \sum_{j=1}^k \frac{1}{j^{1/2}} A_j \geq \sum_{\{a \in \mathbb{N}: a^2 \leq k\}} \frac{1}{a} = \sum_{a \leq k^{1/2}} \frac{1}{a} \geq c \log k,$$

as desired. This completes the proof of the lemma, and thus completes the proof of the Dirichlet's theorem.

□

□

2.5 Representations of finite abelian groups

Finally, we prove the Peter-Weyl theorem as stated in Theorem 5 for finite abelian groups. There is a more general version of Peter-Weyl theorem for topological groups (not necessarily abelian or finite). Here we shall just give the proof in the simple setting of finite abelian groups.

In this section G will be a finite abelian group (group law written multiplicatively), and \hat{G} will be the group of characters of G . $L^2(G)$ will denote the *complex* vector space of all complex-valued functions on G , with inner product given as in Theorem 5.

First we verify the orthogonality between distinct characters: If $\chi_1 \neq \chi_2$ are both characters of G , then

$$\langle \chi_1, \chi_2 \rangle = \frac{1}{|G|} \sum_{a \in G} \chi_1(a) \overline{\chi_2(a)} = \langle \chi_1 \chi_2^{-1}, \chi_0 \rangle$$

where χ_0 is the trivial character, and $\chi_1 \chi_2^{-1}$ is a non-trivial character. Hence it suffices to verify the orthogonality between the trivial character χ_0 and a non-trivial character χ , i.e.

$$\sum_{a \in G} \chi(a) = 0$$

for any non-trivial $\chi \in \hat{G}$, to which we now turn our attention. Let $S = \sum_{a \in G} \chi(a)$. Then since $\chi \in \hat{G}$ is non-trivial, $\chi(b) \neq 1$ for some $b \in G$. Hence by change of variable,

$$S = \sum_{a \in G} \chi(a) = \sum_{a \in G} \chi(ab) = \chi(b)S$$

which implies that $S = 0$, as desired.

Next we verify that there are enough characters to span $L^2(G)$. Our proof is based on a lemma in linear algebra, concerning the simultaneous diagonalization of a family of commuting linear operators on a finite dimensional complex inner product space.

First note that if $\chi \in \hat{G}$, then for any $b \in G$, we have

$$\chi(ab) = \lambda_b \chi(a)$$

for all $a \in G$, where $\lambda_b = \chi(b)$. In other words, if $T_b: L^2(G) \rightarrow L^2(G)$ is the operator given by translation by b (i.e. given any function f on G , $T_b(f)$ is defined to be the function such that $T_b(f)(a) = f(ab)$ for all $a \in G$), then

$$T_b(\chi) = \lambda_b \chi$$

for all $b \in G$ and all $\chi \in \hat{G}$. Hence given any $b \in G$, any character χ of G is an eigenvector of the operator T_b . The converse is to a large extent true: if $f \in L^2(G)$ is an eigenvector

of the operator T_b for all $b \in G$, then there exists a family of complex constants λ_b indexed by $b \in G$ such that

$$f(ab) = \lambda_b f(a)$$

for all $a, b \in G$. It follows that λ_b has to be $f(b)/f(1)$, i.e.

$$f(ab) = \frac{1}{f(1)} f(a)f(b)$$

for all $a, b \in G$. Hence if in addition $f(1) = 1$, we have then $|f(a)| = 1$ for every $a \in G$ and so f is a character of G , i.e. $f \in \hat{G}$. This says

Lemma 13. *A function f defined on G is a character of G if and only if $f(1) = 1$ and f is an eigenvector of T_b for any $b \in G$.*

Hence our goal reduces to showing that there are enough simultaneous eigenvectors of the family of operators $\{T_b : b \in G\}$ such that they span $L^2(G)$. The key observations are that all these T_b 's commute, i.e.

$$T_a T_b = T_b T_a$$

for all $a, b \in G$, and that the T_b 's are all unitary operators on $L^2(G)$, i.e.

$$\langle T_b(f), T_b(g) \rangle = \langle f, g \rangle$$

for all $f, g \in H$ and all $b \in G$, which allow us to invoke the following lemma.

Lemma 14. *If $\{T_i\}_{i=1}^k$ is a commuting family of unitary operators on a finite dimensional complex inner product space H , then there is a basis $\{f_1, \dots, f_n\}$ of H such that each f_j is an eigenvector of all the T_i 's, i.e. the T_i 's can be simultaneously diagonalized.*

This lemma applied to the family $\{T_b : b \in G\}$ of operators on $H = L^2(G)$ allows us to assert that there is an eigenbasis $\{f_1, \dots, f_n\}$ such that each f_j is an eigenvector of all the T_b 's. Renormalizing the f_j 's such that

$$\tilde{f}_j := \frac{f_j}{f_j(1)},$$

we obtain a basis $\{\tilde{f}_1, \dots, \tilde{f}_n\}$ of $L^2(G)$ such that each \tilde{f}_j is a character of G , according to Lemma 13. This proves that $L^2(G)$ admits a basis consisting of characters of G and thus concludes the proof of Theorem 5, save that we have to verify that it is possible to renormalize the f_j 's, i.e. we are left to check that $f_j(1) \neq 0$ for all j . This is easy though: if $f_j(1) = 0$ then since f_j is an eigenvector of all the T_b 's, we have

$$f_j(b) = T_b(f_j)(1) = \lambda_{j,b} f_j(1) = 0$$

for all $b \in G$, where $\lambda_{j,b}$ is the eigenvalue of T_b corresponding to the eigenvector f_j ; but this says f_j is identically zero, which is absurd since f_j is a basis vector of $L^2(G)$. This proves that $f_j(1) \neq 0$ for all j , and we are done.

Proof of Lemma 14. The proof is by induction on the number k of commuting unitary operators involved. When $k = 1$, this is the well-known fact that a unitary operator on a finite dimensional complex inner product space is diagonalizable. So now assume that the assertion holds for some $k \in \mathbb{N}$ and take $k + 1$ commuting unitary operator on H . First T_{k+1} is unitary, thus H decomposes into eigenspaces corresponding to distinct eigenvalues, say $H = H_1 \oplus H_2 \oplus \dots \oplus H_m$, with H_j being the eigenspace of T_{k+1} with eigenvalue λ_j . Now T_1, \dots, T_k preserve each of the H_j 's, because if $v \in H_j$, then for $i = 1, \dots, k$,

$$T_{k+1}(T_i v) = T_i T_{k+1} v = T_i(\lambda_j v) = \lambda_j(T_i v).$$

It follows that $\{T_1, \dots, T_k\}$ are commuting unitary operators on each of these H_j 's, thus are simultaneously diagonalizable on these H_j 's by induction hypothesis. This gives a simultaneous diagonalization of the $k + 1$ operators $\{T_1, \dots, T_{k+1}\}$ on the bigger space H , and Lemma 14 is proved. \square

3 Prime number theorem

There are a number of different approaches to the prime number theorem, some more elementary and others more sophisticated. Since our aim is to illustrate how analysis can be used to study number theory, we shall present here a classical proof using complex analysis. We shall derive the prime number theorem from properties of the Riemann zeta function. Before that, we shall need to take a closer look at the zeta function, and in particular investigate the zeros of the zeta function on the line $\{\operatorname{Re} s = 1\}$.

3.1 Analytic continuation of ζ

The zeta function was defined by a series

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

which converges for all complex s with $\operatorname{Re} s > 1$. For a deeper understanding of the zeta function, it is necessary to analytically continue to a larger region in the complex plane. There are two customary ways of doing that, one via a functional equation that arises from the connection of the zeta function with a modular form, namely the Jacobi theta function $\vartheta(t)$, and another via comparing the sum with an integral. The latter approach is more like what we have been doing in proving the non-vanishing of $L(1, \chi)$ for real χ , and despite being more elementary, it is less powerful than the first approach, for we do not get a functional equation out of it. Below we describe both approaches.

3.1.1 Functional equation approach

Here we shall make use of the theta function

$$\vartheta(t) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 t}$$

($t > 0$) and the Poisson summation formula, which says that for a function f defined on the real line that satisfies suitable decay estimates, e.g.

$$|f(x)| \leq \frac{A}{1+x^2} \quad \text{and} \quad |\hat{f}(\xi)| \leq \frac{A}{1+\xi^2},$$

we have

$$(10) \quad \sum_{n=-\infty}^{\infty} f(x+n) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{2\pi i n x}$$

as L^1 periodic functions of x (x taking values in $[0, 1]$). In other words, the two ways of periodizing the function f gives the same result, and if further f is continuous, then setting $x = 0$ we get

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(n).$$

The proof of the Poisson summation formula is not too difficult; one simply checks that both sides of (10) are periodic L^1 functions of x defined on $[0, 1]$, and that their Fourier coefficients agree at every $n \in \mathbb{N}$. The details are left to the reader.

The point here is that the Poisson summation formula applied to $e^{-\pi n^2 t}$ gives an elegant functional equation for ϑ , namely

$$\vartheta(t) = t^{-1/2} \vartheta(1/t)$$

for real $t > 0$. One just has to note that $e^{-\pi t^2}$ is its own Fourier transform, and apply the Poisson summation formula. The details are omitted.

Now what has the zeta function has to do with the theta function? The connection here is given by the principle of subordination via the Gamma function. Recall that

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

for complex s with $\operatorname{Re} s > 0$. We want to study $\zeta(s)$, which is the sum of n^{-s} over n , using the ϑ function. So we squeeze out the term n^{-s} from this definition of Γ , and at the

same time create terms of the form $e^{-\pi n^2 t}$: in fact by the customary change of variable, we have, for $\operatorname{Re} s > 1$, that

$$\Gamma\left(\frac{s}{2}\right) = \int_0^\infty e^{-\pi n^2 t} (\pi n^2 t)^{s/2} \frac{dt}{t} = \pi^{s/2} n^s \int_0^\infty e^{-\pi n^2 t} t^{s/2} \frac{dt}{t},$$

and thus

$$\pi^{-s/2} n^{-s} \Gamma\left(\frac{s}{2}\right) = \int_0^\infty e^{-\pi n^2 t} t^{s/2} \frac{dt}{t}.$$

Summing over $n \in \mathbb{N}$ we get

$$\pi^{-s/2} \zeta(s) \Gamma\left(\frac{s}{2}\right) = \int_0^\infty \frac{\vartheta(t) - 1}{2} t^{s/2} \frac{dt}{t}.$$

Break the integral into two parts, namely $\int_0^1 + \int_1^\infty$, and using the functional equation for ϑ along with a change of variable in the first integral, we get

$$\pi^{-s/2} \zeta(s) \Gamma\left(\frac{s}{2}\right) = \frac{1}{s-1} - \frac{1}{s} + \int_1^\infty (t^{(1-s)/2-1} + t^{s/2-1}) \frac{\vartheta(t) - 1}{2} dt$$

in which the integral converges for all $s \in \mathbb{C}$ and is invariant under the substitution $s \leftrightarrow 1 - s$. Hence if we define $\xi(s)$ to be the right hand side of the above equality and insist that we have

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

then we have meromorphically extended ζ to the whole complex plane, and we have the functional equation

$$\xi(s) = \xi(1 - s),$$

which also serves as a functional equation for ζ upon the substitution $s \leftrightarrow 1 - s$.

Note that ξ is holomorphic except at the simple poles $s = 0$ and $s = 1$, and from

$$(11) \quad \zeta(s) = \pi^{s/2} \xi(s) \frac{1}{\Gamma(s/2)}$$

we see that the simple pole of ξ at $s = 0$ is cancelled by the simple zero of $\frac{1}{\Gamma(s/2)}$ there. Furthermore, at $s = 1$, $\Gamma(1/2) = \pi^{1/2}$. As a result, ζ is now holomorphic on the whole complex plane, with a simple pole at $s = 1$ where ζ behaves like

$$\zeta(s) \sim \frac{1}{s-1}.$$

This settles the questions about the singularities of ζ .

The zeros of ζ are far more fascinating. From (11) one sees immediately that $\zeta(s) = 0$ whenever s is a negative even integer, for $\frac{1}{\Gamma(s/2)}$ vanishes there. These are called the trivial zeros of ζ , and they are the only zeros of ζ off the *critical strip* $\{0 \leq \operatorname{Re} s \leq 1\}$. This

is because ζ is nowhere zero on $\{\operatorname{Re} s > 1\}$, and thus so is ξ ; then by the functional equation, ξ has no zeros in the region $\{\operatorname{Re} s < 0\}$ as well, so by (11) one sees that the only possible zeros of ζ off the critical strip comes from those of $\frac{1}{\Gamma(s/2)}$, and they are precisely the negative even integers. The famous **Riemann hypothesis** asserts that the only non-trivial zeros of ζ (i.e. those lying in the critical strip) lie on the line $\operatorname{Re} s = \frac{1}{2}$. This has remained open for more than a hundred years, despite the effort of many first rate mathematicians. In fact the zeros of ζ inside the critical strip has been of great interest since Riemann's famous memoir in 1859. The proof of the prime number theorem that we shall give involves showing that ζ has no zeros on the line $\operatorname{Re} s = 1$, and in fact it was known that the prime number theorem is equivalent to the fact that ζ does not vanish on the line $\operatorname{Re} s = 1$.

3.1.2 Integral approach

Next we describe the integral approach of analytically continuing ζ . Here we compare the partial sums of $\zeta(s)$ to an integral, and continue ζ to the region $\{\operatorname{Re} s > 0\}$. More precisely, we write, for $\operatorname{Re} s > 1$ and $N \in \mathbb{N}$, that

$$(12) \quad \sum_{n=1}^{N-1} \frac{1}{n^s} - \int_1^N \frac{1}{x^s} dx = \sum_{n=1}^{N-1} \delta_n(s),$$

where

$$\delta_n(s) = \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) dx.$$

Note that each δ_n is an entire function of s , and satisfies an estimate

$$|\delta_n(s)| \leq \int_n^{n+1} \left| \frac{1}{n^s} - \frac{1}{x^s} \right| dx \leq \frac{|s|}{n^{\operatorname{Re} s + 1}}$$

by the mean value inequality applied to $x \mapsto x^{-s}$. Hence

$$H(s) := \sum_{n=1}^{\infty} \delta_n(s)$$

converges locally uniformly on the region $\{\operatorname{Re} s > 0\}$, and defines a holomorphic function there. Letting $N \rightarrow \infty$ in (12), we have

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

for $\operatorname{Re} s > 1$. However, since the right hand side of this equation is defined and meromorphic on $\{\operatorname{Re} s > 0\}$, it provides an meromorphic continuation of ζ to this larger region, and it is again readily seen that ζ has a simple pole $\frac{1}{s-1}$ at the point $s = 1$, as we have observed using the functional equation approach.

3.2 Estimates of ζ and ζ'

Our goal now will be to establish some estimates of ζ (and its derivative) on the line $\operatorname{Re} s = 1$, which is a key to proving the prime number theorem. This will be based on the integral approach of continuing ζ that we described above.

Lemma 15. *For each $\sigma_0 \in (0, 1]$ and each $\varepsilon > 0$, there exists a constant c_ε such that*

$$(a) \quad |\zeta(s)| \leq c_\varepsilon |\operatorname{Im} s|^{1-\sigma_0+\varepsilon} \text{ for all } s \text{ with } \operatorname{Re} s \geq \sigma_0 \text{ and } |\operatorname{Im} s| \geq 1.$$

$$(b) \quad |\zeta'(s)| \leq c_\varepsilon |\operatorname{Im} s|^\varepsilon \text{ for all } s \text{ with } \operatorname{Re} s \geq 1 \text{ and } |\operatorname{Im} s| \geq 1.$$

Proof. To see that this is true, first observe that the second assertion follows from the first by Cauchy's estimates: if $\operatorname{Re} s \geq 1$ and $|\operatorname{Im} s| \geq 1$, then given $\varepsilon > 0$, Cauchy's estimate for holomorphic functions gives

$$|\zeta'(s)| \leq \frac{1}{\varepsilon/2} \max_{|t-s|=\varepsilon/2} |\zeta(t)|,$$

so applying (a) with $\sigma_0 \geq 1 - \varepsilon/2$ and $\varepsilon/2$ in place of ε we have

$$|\zeta'(s)| \leq \frac{c_{\varepsilon/2} (|\operatorname{Im} s| + \varepsilon/2)^{1-(1-\varepsilon/2)+\varepsilon/2}}{\varepsilon/2} = O(|\operatorname{Im} s|^\varepsilon)$$

with an implicit constant depending on ε . For the first assertion it suffices to consider the case where ε is small, say $1 - \sigma_0 + \varepsilon < 1$. In this case the inequality follows from two estimates of $\delta_n(s)$, one being the estimate

$$|\delta_n(s)| \leq \frac{|s|}{n^{\operatorname{Re} s+1}}$$

which we already alluded to, the other being

$$|\delta_n(s)| \leq \frac{2}{n^{\operatorname{Re} s}}$$

which holds trivially by estimating the integrand of the integral defining δ_n . Now taking a geometric mean of these two estimates, for any $\eta \in [0, 1]$, we have

$$|\delta_n(s)| \leq \left(\frac{|s|}{n^{\operatorname{Re} s+1}} \right)^\eta \left(\frac{2}{n^{\operatorname{Re} s}} \right)^{1-\eta} \leq \frac{2|s|^\eta}{n^{\sigma_0+\eta}}.$$

Taking $\eta = 1 - \sigma_0 + \varepsilon$, we have

$$|\delta_n(s)| \leq C \frac{|\operatorname{Im} s|^{1-\sigma_0+\varepsilon}}{n^{1+\varepsilon}}.$$

This is because if $\operatorname{Re} s \leq 1 + \varepsilon$, then

$$|s| \leq \sqrt{(1 + \varepsilon)^2 + |\operatorname{Im} s|^2} \leq |\operatorname{Im} s| \sqrt{(1 + \varepsilon)^2 + 1}$$

while if $\operatorname{Re} s \geq 1 + \varepsilon$ the estimate is trivial:

$$|\delta_n(s)| \leq \frac{2}{n^{\operatorname{Re} s}} \leq \frac{2}{n^{1+\varepsilon}} \leq \frac{2|\operatorname{Im} s|^{1-\sigma_0+\varepsilon}}{n^{1+\varepsilon}}.$$

Hence by (12),

$$|\zeta(s)| \leq \left| \frac{1}{s-1} \right| + C|\operatorname{Im} s|^{1-\sigma_0+\varepsilon} \sum_{n=1}^{\infty} \frac{1}{n^{1+\varepsilon}} \leq c_\varepsilon |\operatorname{Im} s|^{1-\sigma_0+\varepsilon}$$

as desired. □

We shall need just one more estimate of ζ on the line $\{\operatorname{Re} s = 1\}$, this time from below. This is really a quantitative way of saying that ζ has no zeros on $\operatorname{Re} s = 1$ (note that ζ blows up near the pole $s = 1$), and in fact if one examines the proof that follows, one will see that we have actually shown that $\zeta(s) \neq 0$ on the line $\operatorname{Re} s = 1$.

Lemma 16. *For every $\varepsilon > 0$, there exists c_ε such that*

$$\frac{1}{\zeta(s)} \leq c_\varepsilon |\operatorname{Im} s|^\varepsilon$$

for all s with $\operatorname{Re} s \geq 1$ and $|\operatorname{Im} s| \geq 1$.

Proof. First observe that

$$\log |\zeta^3(\sigma)\zeta^4(\sigma + it)\zeta(\sigma + 2it)| \geq 0$$

for all real $\sigma > 1$ and real t . This is because by the Euler product formula, for $\operatorname{Re} s > 1$,

$$\log \zeta(s) = \sum_p \log \left(\frac{1}{1 - p^{-s}} \right) = \sum_p \sum_m \frac{p^{-ms}}{m} = \sum_{n=1}^{\infty} c_n n^{-s}$$

where $c_n = 1/m$ if n is a prime power with $n = p^m$ and $c_n = 0$ otherwise. It follows that for $\sigma > 1$ and t is real, then

$$\begin{aligned} & \log |\zeta^3(\sigma)\zeta^4(\sigma + it)\zeta(\sigma + 2it)| \\ &= 3\operatorname{Re} \log \zeta(\sigma) + 4\operatorname{Re} \log \zeta(\sigma + it) + \operatorname{Re} \log \zeta(\sigma + 2it) \\ &= \sum_n c_n n^{-\sigma} (3 + 4 \cos \theta_n + \cos 2\theta_n) \end{aligned}$$

where $\theta_n = t \log n$. However, $c_n \geq 0$ for all n , and $3 + 4 \cos \theta + \cos 2\theta = 2(1 + \cos \theta)^2 \geq 0$ for all real θ . Thus

$$\log |\zeta^3(\sigma)\zeta^4(\sigma + it)\zeta(\sigma + 2it)| \geq 0$$

as desired.

From this we prove the lemma: note now for $\sigma > 1$ and $|t| \geq 1$, we have

$$|\zeta^3(\sigma)\zeta^4(\sigma + it)\zeta(\sigma + 2it)| \geq 1.$$

Hence for σ large, using Lemma 15, we have

$$|\zeta(\sigma + it)| \geq c_\varepsilon^{-1} |\zeta(\sigma)|^{-3/4} |2t|^{-\varepsilon/4} \geq c'_\varepsilon |t|^{-\varepsilon/4}.$$

The inequality we want to prove follows easily from this.

For σ close to 1, since zeta function has simple pole at 1, we have

$$(13) \quad |\zeta(\sigma + it)| \geq c_\varepsilon^{-1} |\zeta(\sigma)|^{-3/4} |2t|^{-\varepsilon/4} \geq c'_\varepsilon (\sigma - 1)^{3/4} |t|^{-\varepsilon/4}.$$

If $\sigma - 1 \geq A|t|^{-5\varepsilon}$ for some constant A (the value of A to be specified below), then we are in good shape, and

$$|\zeta(\sigma + it)| \geq c''_\varepsilon |t|^{-4\varepsilon}$$

as desired. Hence it suffices to consider the case where $\sigma - 1 < A|t|^{-5\varepsilon}$. However, if this is the case, then we can first choose an $\tilde{\sigma}$ such that $\tilde{\sigma} > \sigma$, and $\tilde{\sigma} - 1 = A|t|^{-5\varepsilon}$. Then since

$$|\zeta(\sigma + it)| \geq |\zeta(\tilde{\sigma} + it)| - |\zeta(\sigma + it) - \zeta(\tilde{\sigma} + it)|,$$

we have by (13) that

$$|\zeta(\sigma + it)| \geq c'_\varepsilon (\tilde{\sigma} - 1)^{3/4} |t|^{-\varepsilon/4} - |\zeta(\sigma + it) - \zeta(\tilde{\sigma} + it)|.$$

Applying mean value theorem, the last term in absolute value is estimated by

$$|\zeta(\sigma + it) - \zeta(\tilde{\sigma} + it)| \leq c'''_\varepsilon |\tilde{\sigma} - \sigma| |t|^\varepsilon \leq c'''_\varepsilon (\tilde{\sigma} - 1) |t|^\varepsilon$$

using the estimate for ζ' in Lemma 15. Hence

$$|\zeta(\sigma + it)| \geq c'_\varepsilon (\tilde{\sigma} - 1)^{3/4} |t|^{-\varepsilon/4} - c'''_\varepsilon (\tilde{\sigma} - 1) |t|^\varepsilon.$$

Set now $A = (c'_\varepsilon / (2c'''_\varepsilon))^4$ and using the fact that $\tilde{\sigma} - 1 = A|t|^{-5\varepsilon}$, we have the first term on the right hand side exactly equal to $2c'''_\varepsilon (\tilde{\sigma} - 1) |t|^\varepsilon$. Hence

$$|\zeta(\sigma + it)| \geq c'''_\varepsilon (\tilde{\sigma} - 1) |t|^\varepsilon \geq c''''_\varepsilon |t|^{-4\varepsilon}$$

as was to be shown in this case as well. □

3.3 Chebyshev's ψ function

Having understood more the zeta function, now we turn to the proof of the prime number theorem. We shall need two auxiliary functions. The first one is the Chebyshev's ψ function, defined for real x by

$$\psi(x) = \sum_{p^m \leq x} \log p$$

where the sum is over all prime powers p^m not exceeding x . It is readily seen that

$$\psi(x) = \sum_{p \leq x} \left\lfloor \frac{\log x}{\log p} \right\rfloor \log p$$

where here the sum is over all primes p not exceeding x . The second one is the integral of ψ , which we denote by ψ_1 :

$$\psi_1(x) = \int_1^x \psi(t) dt.$$

The following notation will be adopted: we write

$$f(x) \sim g(x)$$

to mean

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

Hence the prime number theorem says that

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

To prove the prime number theorem, we make a series of reductions. First we have

Theorem 17. $\pi(x) \sim \frac{x}{\log x}$ if and only if $\psi(x) \sim x$.

Proof. Note that

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

and that for all $\alpha \in (0, 1)$,

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

the last inequality following from the fact that $\pi(x^\alpha) \leq x^\alpha$. Hence

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

It follows that if $\psi(x) \sim x$, then letting $x \rightarrow \infty$ in the above estimate, we get

$$1 \leq \liminf_{x \rightarrow \infty} \frac{\pi(x)}{\frac{x}{\log x}} \leq \limsup_{x \rightarrow \infty} \frac{\pi(x)}{\frac{x}{\log x}} \leq \frac{1}{\alpha}$$

for all $\alpha \in (0, 1)$, and letting $\alpha \rightarrow 1^-$, we get $\pi(x) \sim \frac{x}{\log x}$. One can prove the converse in a similar fashion. \square

Next we note that

Theorem 18. $\psi(x) \sim x$ if and only if $\psi_1(x) \sim x^2/2$.

Proof. This is because for any $\beta > 1$,

$$\psi(x) \leq \frac{1}{(\beta - 1)x} \int_x^{\beta x} \psi(t) dt = \frac{1}{(\beta - 1)x} (\psi_1(\beta x) - \psi_1(x)),$$

the first inequality following from the fact that ψ is an increasing function of x . It follows that

$$\frac{\psi(x)}{x} \leq \frac{1}{\beta - 1} \left(\frac{\psi_1(\beta x)}{(\beta x)^2} \beta^2 - \frac{\psi_1(x)}{x^2} \right),$$

and if $\psi_1(x) \sim x^2/2$ then letting $x \rightarrow \infty$ we get

$$\limsup_{x \rightarrow \infty} \frac{\psi(x)}{x} \leq \frac{1}{\beta - 1} \left(\frac{1}{2} \beta^2 - \frac{1}{2} \right) \leq \frac{1}{2} (\beta + 1).$$

Letting $\beta \rightarrow 1^-$ we get

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

Similarly, this time using the fact that for $0 < \alpha < 1$,

$$\psi(x) \geq \frac{1}{(1 - \alpha)x} \int_{\alpha x}^x \psi(t) dt = \frac{1}{(1 - \alpha)x} (\psi_1(x) - \psi_1(\alpha x)),$$

we get

$$\liminf_{x \rightarrow \infty} \frac{\psi(x)}{x} \geq 1$$

assuming $\psi_1(x) \sim x^2/2$. It follows that $\psi(x) \sim x$, and the converse can be proved similarly. \square

3.4 Proof of the prime number theorem

According to the results in the last section, to prove the prime number theorem it suffices to prove $\psi_1(x) \sim x^2/2$. This is what we shall aim at now. Here more complex analysis comes in. First we relate ψ_1 to the zeta function, so that we can make use of our knowledge of ζ that we gathered in the previous sections. This is done via a magic identity:

Lemma 19. *For all real $c > 1$,*

$$(14) \quad \psi_1(x) = -\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} ds,$$

the contour integral being upwards along the vertical line $\operatorname{Re} s = c$.

Proof. We have on one hand the ψ_1 function, and on the other the ζ function. To make the two ends meet, we use the Λ function, defined for $n \in \mathbb{N}$ by

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ for some prime } p \text{ and some } k \geq 1 \\ 0 & \text{otherwise} \end{cases}.$$

First,

$$\psi(t) = \sum_{1 \leq n \leq t} \Lambda(n) = \sum_{n=1}^{\infty} \Lambda(n) f_n(t)$$

where f_n is the characteristic function of the interval $[n, \infty)$. Hence

$$\psi_1(x) = \int_1^x \psi(t) dt = \sum_{n=1}^{\infty} \Lambda(n) \int_1^x f_n(t) dt = \sum_{1 \leq n \leq x} \Lambda(n)(x - n).$$

Next, taking the logarithmic derivative of the Euler product formula, we get

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

We would like to plug this into the right hand side of the desired identity and evaluate the contour integrals. The relevant contour integrals are given by the following lemma.

Lemma 20. *If $c > 0$ then*

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{a^s}{s(s+1)} ds = \begin{cases} 0 & \text{if } a \in (0, 1] \\ 1 - 1/a & \text{if } a \geq 1 \end{cases},$$

the integral taken along the vertical line $\operatorname{Re} s = c$.

Assuming this for the moment, we see that

$$\begin{aligned}
-\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} ds &= x \sum_{n=1}^{\infty} \Lambda(n) \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{(x/n)^s}{s(s+1)} ds \\
&= x \sum_{1 \leq n \leq x} \Lambda(n) \left(1 - \frac{n}{x}\right) \\
&= \psi_1(x),
\end{aligned}$$

proving Lemma 19. □

Hence it suffices now to prove the contour integral in Lemma 20.

Proof of Lemma 20. Indeed the contour integral can be evaluated by residue theorem. If $a \geq 1$, then for $T \gg 0$, replace the contour by $\Gamma_T = L_1 + C_2 + L_3$, where L_1 is the vertical line from $c - i\infty$ to $c - iT$, C_2 is the half circle centered at c and of radius T that lies to the left of $\operatorname{Re} s = c$, and L_3 is the vertical line from $c + iT$ to $c + i\infty$. The residue that one gains this way (at the poles $s = 0, -1$) is precisely $1 - 1/a$, and as $T \rightarrow \infty$ the new contour integral tends to zero (here one uses that $a \geq 1$). This proves the second part of the assertion in Lemma 20. The other part is proved similarly, except that this time one has to replace the contour by $\tilde{\Gamma}_T$ which one obtains by reflecting Γ_T along the line $\operatorname{Re} s = c$. That this is necessary is due to the fact that one needs this to argue that the contour integral along the new contour tends to zero as $T \rightarrow \infty$, and when $a \in (0, 1]$ this can only be achieved by using the contour $\tilde{\Gamma}_T$. The details are left to the reader. □

We shall now use the magic identity in Lemma 19 to finish the proof of the prime number theorem. We will verify that $\psi_1(x) \sim x^2/2$ using Lemma 19 and the estimates in section 3.2. The idea is to shift the contour integral to from $\operatorname{Re} s = c$ ($c > 1$) to $\operatorname{Re} s = 1$, so that the power of x in the integrand of (14) gets closer to the desired value of 2.

More precisely, we shall do the following: write the integrand in (14) as $F(s)$. Then for any $T > 0$, if we denote by γ_T the contour $l_1 + l_2 + l_3 + l_4 + l_5$, where

- l_1 is the vertical line from $1 - i\infty$ to $1 - iT$
- l_2 is the horizontal line from $1 - iT$ to $c - iT$
- l_3 is the vertical line from $c - iT$ to $c + iT$
- l_4 is the horizontal line from $c + iT$ to $1 + iT$
- l_5 is the vertical line from $1 + iT$ to $1 + i\infty$,

we see that

$$(15) \quad \int_{c-i\infty}^{c+i\infty} F(s)ds = \int_{\gamma_T} F(s)ds.$$

To see this we make use of the residue theorem. For this we use the estimate

$$\left| \frac{\zeta'(s)}{\zeta(s)} \right| \leq C|\operatorname{Im} s|^{1/2}$$

for $\operatorname{Re} s \geq 1$, $|\operatorname{Im} s| \geq 1$ that we obtain by Lemma 15 and 16, which imply the decay

$$|F(s)| \leq C|\operatorname{Im} s|^{-3/2}$$

for the same s . This shows that $F(s)$ is integrable along the contour $(c-i\infty, c+i\infty) - \gamma_T$, and by residue theorem, since $F(s)$ has no singularity in the region enclosed by this contour, it follows that its integral along the contour is zero, and (15) is established.

Next we change the contour again: let $\gamma_{T,\delta}$ be the contour $l_1 + l_6 + l_7 + l_8 + l_5$, where l_1, l_5 are as above, and

- l_6 is the horizontal line from $1 - iT$ to $1 - \delta - iT$
- l_7 is the vertical line from $1 - \delta - iT$ to $1 - \delta + iT$
- l_8 is the horizontal line from $1 - \delta + iT$ to $1 + iT$.

In doing so we have to be careful, and we need to ensure that $F(s)$ and hence $1/\zeta(s)$ has no singularity on $\gamma_{T,\delta}$ so defined. This we do by choosing $\delta \in (0, 1)$ sufficiently small (how small it is depend on our choice of T) so that ζ has no zero on $l_6 + l_7 + l_8$, which is possible since ζ has no zero on the line $\operatorname{Re} s = 1$ (and hence in an open neighborhood of $(1 - iT, 1 + iT)$ for each finite $T > 0$). Now

$$\int_{\gamma_T} F(s)ds = \int_{\gamma_{T,\delta}} F(s)ds + (2\pi i)\left(-\frac{x^2}{2}\right),$$

since the residue of $F(s)$ at the pole $s = 1$ is $-x^2/2$. This is because

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

near $s = 1$ for some non-zero holomorphic function $h(s)$, and carrying out logarithmic differentiation we see that

$$\frac{\zeta'(s)}{\zeta(s)} = -\frac{1}{s-1} + h_1(s)$$

for some holomorphic function $h_1(s)$ near $s = 1$.

To summarize what we have got so far, we see that

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

where $T > 0$ can be arbitrarily large and $\delta \in (0, 1)$ is sufficiently small relative to T . Hence to verify that $\psi_1(x) \sim x^2/2$, which has been the goal of this section, we shall prove that

$$\lim_{x \rightarrow \infty} \int_{\gamma_{T,\delta}} \frac{x^{s-1}}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} ds = 0.$$

Now fix $\eta > 0$, and we want to see that the integral is smaller than η when x is sufficiently large. Choosing T sufficiently large, we have

$$\int_{l_1+l_5} \frac{x^{s-1}}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} ds \leq \frac{\eta}{2}.$$

This is because $\zeta'(s)/\zeta(s) \leq C|\operatorname{Im} s|^{1/2}$ on $\operatorname{Re} s = 1$, again from the estimates we got in Lemma 15 and 16. Next fixing this T , we observe that

$$\int_{l_7} \frac{x^{s-1}}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} ds \leq C_T x^{\operatorname{Res}-1} = C_T x^{-\delta},$$

and

$$\int_{l_6+l_8} \frac{x^{s-1}}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} ds \leq C \int_{1-\delta}^1 x^{\sigma-1} d\sigma \leq C \frac{1}{\log x}.$$

Hence choosing x large enough, we have

$$\int_{l_6+l_7+l_8} \frac{x^{s-1}}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} ds \leq \frac{\eta}{2}.$$

This proves that

$$\int_{\gamma_{T,\delta}} \frac{x^{s-1}}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} ds \leq \eta,$$

for large x , and hence $\psi_1(x) \sim x^2/2$. This completes the proof of the prime number theorem.

References

- [1] E. Stein and R. Shakarchi, *Fourier analysis*, Princeton University Press, 2003.
- [2] E. Stein and R. Shakarchi, *Complex analysis*, Princeton University Press, 2003.
- [3] T. Apostol, *Introduction to analytic number theory*, New York: Springer-Verlag, 1976.