THE PRIME NUMBER THEOREM FOR ARITHMETIC SEQUENCES

NADIR HAJOUJI

The two classic results of 19th century analytic number theory are the Prime Number Theorem, and Dirichlet's Theorem for Arithmetic Progressions. The first theorem describes the asymptotic distribution of primes, and the second asserts that there are infinitely many primes in arithmetic sequences of the form $a, a+q, a+2q, \cdots$, (with $\gcd(a,q) = 1$, obviously). In this paper, I will prove both theorems simultaneously, following the method of Ivan Soprounov ([4]).

1. The Analytic Theorem

The first result we prove is purely analytic. It was introduced by Newman in [2], and it serves a similar purpose in [4]. The key to proving it is the novel integral representation shown below.

Lemma 1.1. *Let $\Omega \subset \mathbb{C}$ be a simply connected open set containing 0 and suppose f is holomorphic on $\overline{\Omega}$.*

$$f(0) = \frac{1}{2\pi i} \int_{\partial\Omega} (f(\zeta)) e^{\zeta T} \left(1 + \frac{\zeta^2}{R^2}\right) \frac{d\zeta}{\zeta}$$

Proof. Using the residue theorem, one computes that:

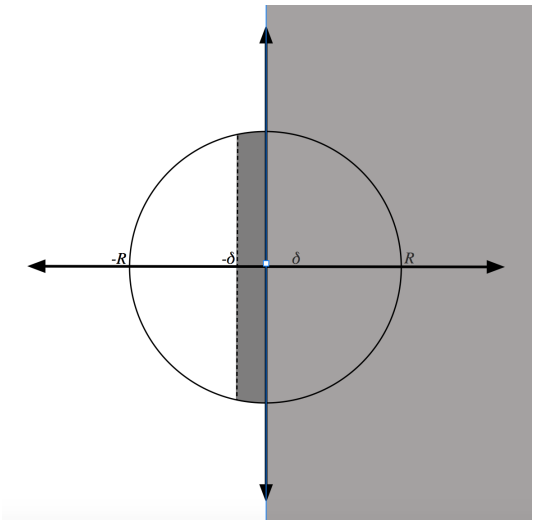
$$\begin{aligned}
\frac{1}{2\pi i} \int_\gamma f(\zeta) e^{\zeta T} \left(1 + \frac{\zeta^2}{R^2}\right) \frac{d\zeta}{\zeta} &= \mathrm{Res}_0 \left(f(\zeta) \frac{e^{\zeta T}}{\zeta} \left(1 + \frac{\zeta^2}{R^2}\right)\right) \\
&= \lim_{\zeta \to 0} \zeta f(\zeta) \frac{e^{\zeta T}}{\zeta} \left(1 + \frac{\zeta^2}{R^2}\right) \\
&= f(0) e^{0 \cdot T} \left(1 + \frac{0}{R^2}\right) \\
&= f(0)
\end{aligned}$$

□

Theorem 1.2. *Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be a bounded and locally integrable function, with $|f(t)| \leq M$ for all $t \in \mathbb{R}_{\geq 0}$. Suppose that $g(z) = \int_0^\infty f(t) e^{-zt} dt$, defined on $\mathrm{Re}(z) > 0$, extends holomorphically to $\mathrm{Re}(z) \geq 0$. Then $\int_0^\infty f(t) dt$ exists and equals $g(0)$.*

Proof. For $T > 0$, define $g_T(z) = \int_0^T f(t)e^{-zt}dt$. Choose $R > 0$. We know that $g(z)$ is analytic on the imaginary axis. For every point iy on the imaginary axis, with $|y| < R$, we can find $\epsilon_y > 0$ such that g is analytic on $B(iy, \epsilon_y)$. These open balls cover the interval $[-R, R]$, which is compact, so we can find $\delta > 0$ such that g is holomorphic on $B(iy, \delta)$ for all $|y| < R$.

Let $\gamma = \{z \in \mathbb{C} | |z| < R, \mathrm{Re}(z) > -\delta\}$. Then g is analytic inside γ and on γ. The picture below shows the domain we are working on. The mapping g was assumed to be analytic on the light grey area, and we extended the analyticity to the light grey strip to the left of the imaginary axis. On the picture, γ is the boundary of the portion of the circle that is shaded.



By the preceding lemma:

$$g(0) - g_T(0) = \frac{1}{2\pi i} \int_\gamma (g(\zeta) - g_T(\zeta)) e^{\zeta T} \left(1 + \frac{\zeta^2}{R^2}\right) \frac{d\zeta}{\zeta}$$

Now, let $\gamma_+ = \{z \in \gamma | \mathrm{Re}(z) \geq 0\}$ and $\gamma_- = \{z \in \gamma | \mathrm{Re}(z) \leq 0\}$. On γ_+:

$$|g(\zeta) - g_T(\zeta)| = \left| \int_T^\infty f(t)e^{-\zeta t}dt \right| \leq \int_T^\infty |f(t)||e^{-\zeta t}|dt \leq M \int_T^\infty e^{-t\mathrm{Re}(\zeta)}dt$$

Since $\mathrm{Re}(\zeta) > 0$, we can evaluate the integral to obtain the bound:

$$|g(\zeta) - g_T(\zeta)| \leq M \left(\frac{e^{-t\mathrm{Re}(\zeta)}}{-\mathrm{Re}(\zeta)} \right|_T^\infty = \frac{Me^{-T\mathrm{Re}(\zeta)}}{\mathrm{Re}(\zeta)}$$

Furthermore, on the circle $|\zeta| = R$, $\frac{1}{\zeta} = \frac{\overline{\zeta}}{R^2}$ so:

$$\left| e^{\zeta T} \left(1 + \frac{\zeta^2}{R^2}\right) \frac{1}{\zeta} \right| = e^{\mathrm{Re}(\zeta)T} \cdot \left| \frac{1}{\zeta} + \frac{\zeta}{R^2} \right| = e^{\mathrm{Re}(\zeta)T} \left| \frac{\overline{\zeta}}{R^2} + \frac{\zeta}{R^2} \right| = \frac{2e^{\mathrm{Re}(\zeta)T}\mathrm{Re}(\zeta)}{R^2}$$

Combining results, we obtain:

$$\left|\int_{\gamma_+}(g(\zeta)-g_T(\zeta))e^{\zeta T}\left(1+\frac{\zeta^2}{R^2}\right)\frac{d\zeta}{\zeta}\right| \leq \frac{Me^{-T\mathrm{Re}(\zeta)}}{\mathrm{Re}(\zeta)}\cdot\frac{2e^{\mathrm{Re}(\zeta)T}\mathrm{Re}(\zeta)}{R^2}=\frac{2M}{R^2}$$

Consequently, we see the contribution of $\int_{\gamma_+}$ is bounded above by $\frac{2M}{R^2}$, and since R is arbitrary, this contribution must be 0.

Next, we look at the contribution of $\int_{\gamma_-}$. We will consider the representations of g_T and g separately this time. Since g_T is entire, we can replace γ_- by $\gamma'_- = \{z \in \mathbb{C} | \mathrm{Re}(z) \leq 0, |z| = R\}$, in order to derive similar bounds as earlier:

$$\begin{aligned}\left|\int_0^T f(t)e^{-\zeta t}dt\right| &\leq \int_0^T |f(t)||e^{-\zeta t}|dt \leq M\int_0^T e^{-t\mathrm{Re}(\zeta)} \leq M\int_{-\infty}^T e^{-t\mathrm{Re}(\zeta)} \\ &= M\left(\frac{e^{-t\mathrm{Re}(\zeta)}}{-\mathrm{Re}(\zeta)}\right|_{-\infty}^T dt = M\frac{e^{-T\mathrm{Re}(\zeta)}}{-\mathrm{Re}(\zeta)}\end{aligned}$$

Note that $\mathrm{Re}(\zeta) < 0$ since $\zeta \in \gamma'_-$, which is why we took the limit to $-\infty$. It is also clear why we extended our integral - we have exactly the same inequality as earlier, and we get the second inequality from the fact that we are still on the circle $|\zeta| = R$. Thus this integral also contributes nothing.

Finally, we have the integral with g on γ_-. We can't work on a circle using g as we did in the earlier cases, but we now have the advantage that g is independent of T. As $T \to \infty$, $e^{\zeta T} \to 0$ since $\mathrm{Re}(\zeta) < 0$, and furthermore, the convergence is uniform on compact sets. Thus, we can pass the limit through the integral to show that:

$$\lim_{T\to\infty}\int_{\gamma_-} g(t)e^{\zeta T}\left(1+\frac{\zeta^2}{R^2}\right)\frac{d\zeta}{\zeta} = \int_{\gamma_-}\left(\lim_{T\to\infty}e^{\zeta T}\right)g(t)\left(1+\frac{\zeta^2}{R^2}\right)\frac{d\zeta}{\zeta} = 0$$

Thus $\lim_{T\to\infty} g_T(0) = g(0)$, so we are done, since:

$$\lim_{T\to\infty} g_T(0) = \lim_{T\to\infty}\int_0^T f(t)e^0 dt = \int_0^\infty f(t)dt.$$

□

2. Dirichlet Series

Before we start introducing algebraic machinery, we prove some more purely analytic results about Dirichlet series. Recall the definition of a Dirichlet series:

Definition 2.1. *Let $\{\lambda_n\}_{n\geq 1}$ be an increasing sequence of positive real numbers tending to $+\infty$. A Dirichlet series with exponents $\{\lambda_n\}$ is a series of the form:*

$$\sum_{n=1}^{\infty} a_n e^{-\lambda_n z} \qquad (a_n \in \mathbb{C}, z \in \mathbb{C})$$

The classic example of a Dirichlet series is the Riemann zeta function:

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

Since:

$$|n^s| = |e^{\ln ns}| = |e^s|^{\ln n} = e^{\mathrm{Re}(s)\ln n} = n^{\mathrm{Re}(s)}$$

for all s in the right half-plane $\mathrm{Re}(s) > 1$:

$$\begin{aligned} |\zeta(s)| = \left|\sum_{n=1}^{\infty} \frac{1}{n^s}\right| &\leq \sum_{n=1}^{\infty} \frac{1}{|n^s|} = \sum_{n=1}^{\infty} \frac{1}{n^{\mathrm{Re}(s)}} \approx \int_1^{\infty} \frac{1}{x^{\mathrm{Re}(s)}} dx \\ &= \left(\frac{1}{x^{\mathrm{Re}(s)-1}(1-\mathrm{Re}(s))}\right|_1^{\infty} = \frac{1}{1-\mathrm{Re}(s)} < \infty \end{aligned}$$

That is, ζ converges absolutely on the right half-plane $\mathrm{Re}(s) > 1$. At $s = 1$, $\zeta(s)$ is the harmonic series, which also famously diverges. However, this is the only pole on the right half-plane $\mathrm{Re}(s) > 0$, and the pole is simple.

Theorem 2.2. *Let $f(z) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n z}$ be a Dirichlet series, and suppose $f(z_0)$ converges for some $z_0 \in \mathbb{C}$. Then f converges uniformly in every domain of the form $\mathrm{Re}(z - z_0) \geq 0$, $-\alpha \leq \arg(z - z_0) \leq \alpha$ for some $0 < \alpha < \frac{\pi}{2}$.*

Corollary 2.3. *$\sum a_n e^{-\lambda_n z}$ is analytic in $\mathrm{Re}(z - z_0) > 0$.*

These result was proven in class, so I will not reprove it. We do need the following result, however, which we have not proven:

Proposition 2.4. *Let $f = \sum a_n e^{-\lambda_n z}$ be a Dirichlet series such that $a_n \in \mathbb{R}_{\geq 0}$ for all n. Suppose that f converges for $\mathrm{Re}(z) > \rho$ for some $\rho \in \mathbb{R}$, and that f can be extended analytically to a function holomorphic in a neighborhood of the point $z = \rho$. Then there exists a real number $\epsilon > 0$ such that f converges for $\mathrm{Re}(z) > \rho - \epsilon$.*

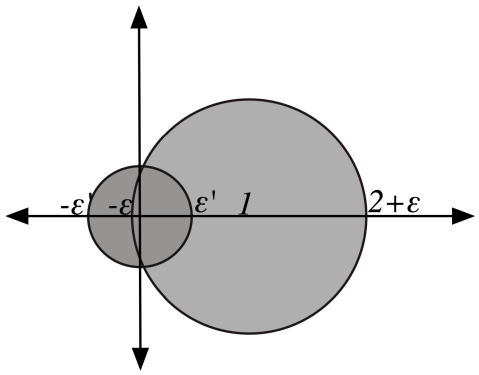
Proof. After an affine transformation, we can assume that $\rho = 0$. Our hypothesis states that f is holomorphic on a neighborhood of the form $|z| < \epsilon'$. Consequently, there exists a neighborhood of radius $1+\epsilon$ about 1 on which f is holomorphic, and f converges absolutely to its Taylor series in that disk:

$$f(z) = \sum_{\ell=0}^{\infty} \frac{f^{(\ell)}(1)(z-1)^\ell}{\ell!} \qquad (|z-1| \leq 1+\epsilon)$$

where the ℓ^{th} derivative is readily seen to be:

$$f^{(\ell)}(z) = \sum_{n=1}^{\infty} a_n(-\lambda_n)^\ell e^{-\lambda_n z} \Longrightarrow f^{(\ell)}(1) = (-1)^\ell \sum_{n=1}^{\infty} a_n(\lambda_n)^\ell e^{-\lambda_n}$$

The summation above converges, and if we ignore the $(-1)^\ell$ term in the front, we see that the summands are all nonnegative, so the summation converges *absolutely*.



Consequently, we can write $f(z)$ as a an absolutely convergent double summation:

$$f(z) = \sum_{\ell=0}^{\infty}\sum_{n=1}^{\infty} a_n(-1)^\ell(\lambda_n)^\ell e^{-\lambda_n}\frac{(z-1)^\ell}{\ell!} = \sum_{\ell=0}^{\infty}\sum_{n=1}^{\infty} a_n(\lambda_n)^\ell e^{-\lambda_n}\frac{(1-z)^\ell}{\ell!} \qquad (|z-1| \leq 1+\epsilon)$$

In particular:

$$f(-\epsilon) = \sum_{\ell=0}^{\infty}\sum_{n=1}^{\infty} \frac{a_n(1+\epsilon)^\ell(\lambda_n)^\ell e^{-\lambda_n}}{\ell!}$$

Now since every summation in sight is absolutely convergent, we can switch the order of summation, so we do:

$$\begin{aligned} f(-\epsilon) &= \sum_{n=1}^{\infty}\left(\sum_{\ell=0}^{\infty}\frac{(1+\epsilon)^\ell(\lambda_n)^\ell}{\ell!}\right)a_n e^{-\lambda_n} = \sum_{n=1}^{\infty} a_n e^{-\lambda_n}e^{(1+\epsilon)\lambda_n} \\ &= \sum_{n=1}^{\infty} a_n e^{\epsilon\lambda_n} \end{aligned}$$

We have just shown that the Dirichlet series $f(z) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n z}$ converges absolutely at $z = -\epsilon$, so by the previous corollary, $f(z)$ converges on $\mathrm{Re}(z+\epsilon) > 0$, i.e. on $\mathrm{Re}(z) > -\epsilon$. $\square$

Let us now prove a fun result about a special type of Dirichlet series that shows how they can be used for algebraic purposes.

Definition 2.5. *Let $f : \mathbb{N} \to \mathbb{C}$ be a function. We say f is* multiplicative *if $f(1) = 1$ and:*

$$\gcd(m,n) = 1 \implies f(mn) = f(m)f(n)$$

The most boring multiplicative function one can think of is the function $n \mapsto 1$ for all n; that function will produce the Riemann zeta function in the upcoming construction.

Proposition 2.6. *Let $\lambda_n = \log n$, $f : \mathbb{N} \to \mathbb{C}$ a multiplicative function satisfying $|f(n)| \leq 1$ for all n, and set $a_n = \chi(n)$. The corresponding Dirichlet series converges absolutely on the half-plane $\mathrm{Re}(s) > 1$, and has the following product representation on that half-plane:*

$$\prod_p (1 - \chi(p)p^{-s})^{-1}$$

where the product is taken over all primes p.

Proof. First, let us show the product on the right hand side converges absolutely. We have:

$$\prod_p \frac{1}{1 - f(p)p^{-s}} = \prod_p 1 + \left(\frac{1}{1 - f(p)p^{-s}} - 1\right) = \prod_p 1 + \frac{f(p)p^{-s}}{1 - f(p)p^{-s}} = \prod_p 1 + \frac{f(p)}{p^s - 1}$$

Now:

$$\implies \left|\frac{f(p)}{p^s - 1}\right| = \frac{|f(p)|}{|p^s - 1|} \leq \frac{1}{p^{\mathrm{Re}(s)} - 1} < \frac{1}{(p-1)^{\mathrm{Re}(s)}} = \frac{1}{|(p-1)^s|}$$

$$\implies \sum_p \left|\frac{f(p)}{p^s - 1}\right| < \sum_p \frac{1}{|(p-1)^s|} < \sum_{n=1}^{\infty} \frac{1}{|n^s|} < \infty$$

so the product converges absolutely, and we may rearrange the terms. Observe that $|f(p)p^{-s}| = p^{-\mathrm{Re}(s)} < p^{-1} < 1$ so the factors in the product are geometric series:

$$\prod_p \frac{1}{1 - f(p)p^{-s}} = \prod_p \sum_{r \geq 0} f(p)^r p^{-rs}$$

If we set $p_1 = 2$, $p_2 = 3$, etc., we can expand the infinite product of infinite sums like so:

$$\begin{aligned}\prod_p \sum_{r\geq 0} f(p)^r p^{-rs} &= \sum_{(r_1,r_2,\cdots)\in\mathbb{N}^{\oplus\mathbb{N}}} f(p_1)^{r_1} p_1^{-r_1 s} f(p_2)^{r_2} p_2^{-r_2 s}\cdots \\ &= \sum_{(r_1,r_2,\cdots)\in\mathbb{N}^{\oplus\mathbb{N}}} f(p_1^{r_1}\cdot p_2^{r_2}\cdots)(p_1^{r_1}\cdot p_2^{r_2}\cdots)^{-s} \\ &= \sum_{n=1}^{\infty}\frac{f(n)}{n^s} = \sum_{n=1}^{\infty} f(n)e^{-\log(n)s} = \sum_{n=1}^{\infty} a_n e^{-\lambda_n s}\end{aligned}$$

Two remarks should be made regarding the final display. The notation $\mathbb{N}^{\oplus\mathbb{N}}$ refers to a countable direct sum of $\mathbb{N}$, so its elements take the form $(r_1, r_2, \cdots)$, with only finitely many of the r_i nonzero. The fundamental theorem of arithmetic states that every $n \geq 1$ can be written as $p_1^{r_1} p_2^{r_2}\cdots$ for some $(r_1, r_2, \cdots) \in \mathbb{N}^{\mathbb{N}}$, which justifies the penultimate equality, so we are done.

□

3. Characters

In order to start making meaningful progress, we need to address a very significant obstacle: how do we study something which we only understand algebraically using analytic tools? Well, we want algebraic objects to live in the world of complex analysis, while retaining their algebraic properties, so we send our algebraic objects to the complex plane in a package that preserves structure. In other words, we use *group characters*.

A character of a group G is is simply a group homomorphism from the group into the multiplicative group of $\mathbb{C}$, i.e. $\chi : G \to \mathbb{C}^\times$. Every group has at least one character, called the principal, or trivial, character, defined by $g \mapsto 1$ for all $g \in G$, and often denoted χ_0. The set of characters of G form the dual group of G, and denoted $\widehat{G}$. The group operation is pointwise multiplication: $\chi \cdot \chi' : g \mapsto \chi(g)\chi'(g)$, which inherits associativity from that of multiplication in $\mathbb{C}$, and the identity is χ_0. We have closure under inverses since $g \mapsto \frac{1}{\chi(g)}$ is also a character, so $\widehat{G}$ is, indeed, a group.

Furthermore, note that if $\chi : G \to \mathbb{C}$ is a homomorphism, then so is $\overline{\chi}$, where $\overline{\chi}(g) = \overline{\chi(g)}$, since complex conjugation is multiplicative. The conjugate of a character has the following interesting property:

$$\chi(g)\chi(g^{-1}) = \chi(1) = 1 = |\chi(g)|^2 = \overline{\chi(g)}\chi(g) \Longrightarrow \overline{\chi(g)} = \chi(g^{-1}) = \frac{1}{\chi(g)}$$

The theory of group characters is very interesting in its own right, and has applications in Galois theory and algebraic number theory. We will not discuss those applications, but we will need some general results in character theory to proceed. In particular, we would like to know how big $\widehat{G}$ can get. If G is finite abelian, then the result is quite nice - $|G| = |\widehat{G}|$. Let us prove this.

Lemma 3.1. *Let G be a group of order n and $\chi : G \to \mathbb{C}^\times$ a character. Then $\chi(g)$ is an n-th root of unity for all $g \in G$.*

Proof. We have:

$$\chi(g)^n = \chi(g^n) = \chi(1) = 1$$

□

Consequently, if G is a cyclic group, then we can easily classify all of the characters of G: every character is determined by where it maps the generator and the generator must go to a $|G|$-th root of unity. Thus there are $|G|$ characters in all. Furthermore, it seems like we might be able to prove the result constructively by dealing with one generator at a time, but

to make this rigorous, we will need to understand how the dual group of G is related to the dual group of its subgroups.

On the one hand, given a group G, a subgroup H and a character $\chi : G \to \mathbb{C}^\times$, it is easy to produce a character for H: we restrict the character of G. In other words, restriction is a mapping from the dual of G to the dual of H, i.e. $\rho : \widehat{G} \to \widehat{H}$, and it is clearly a homomorphism since it doesn't actually change the characters in any way. We can also go the other way.

Lemma 3.2. *Let G be a finite abelian group, H a subgroup of G and $\chi : H \to \mathbb{C}^\times$. There exists $\tilde{\chi} : G \to \mathbb{C}^\times$ such that $\tilde{\chi}|_H = \chi$.*

Proof. We argue by induction on $[G : H]$. If $[G : H] = 1$, then $G = H$ so we may take $\tilde{\chi} = \chi$. Otherwise, choose some $x \in G$ such that $x \notin H$. Let n be the smallest integer such that $x^n \in H$. Since H is finite, $\chi(x^n)$ is an m-th root of unity, i.e. $\chi(x^n) = e^{2\pi ik/m}$, where $m = |H|$.

Now let $K = (H, x)$, the subgroup generated by H and x, and let $\chi_K : K \to \mathbb{C}^\times$ be the homomorphism determined by:

$$\chi_K(x) = e^{2\pi ik/mn}, \qquad \chi_K|_H = \chi$$

Since x, H generate K, χ_K is defined on all of K, and the definition is consistent throughout since:

$$\chi_K(x^n) = \chi_K(x)^n = e^{2\pi ik/m} = \chi(x^n)$$

so χ_K is well-defined. Since $[G : K] < [G : H]$, we can inductively extend $\tilde{\chi_K} : G \to \mathbb{C}^\times$ so that $\tilde{\chi_K}|_K = \chi_K$, and thus $\tilde{\chi_K}|_H = \chi_K|_H = \chi$.

□

The previous lemma basically says that $\rho : \widehat{G} \to \widehat{H}$ is surjective. If $\chi \in \ker(\rho)$, then $\chi|_H = \chi_0$, so χ descends to a homomorphism $\chi_{G/H} : G/H \to \mathbb{C}^\times$, where $\chi_{G/H}(gH) = \chi(g)$. This is well-defined because if $gH = g'H$, then $\chi_{G/H}(gH)\chi_{G/H}(g'H)^{-1} = \chi(g)\chi(g')^{-1} = \chi(g(g')^{-1})$, and $g(g')^{-1} \in H$ since $gH = g'H$ so $\chi(g(g')^{-1}) = 1$. Thus $\chi_{G/H}(gH) = \chi_{G/H}g'H$.

Consequently, we have a short exact sequence:

$$0 \longrightarrow \widehat{G/H} \longrightarrow \widehat{G} \xrightarrow{\rho} \widehat{H} \longrightarrow 0$$

With that, we are good to go.

Proposition 3.3. *Let G be a finite abelian group. Then $|G| = \widehat{G}$.*

Proof. We induct on $|G|$. If $|G| = 1$, then $\widehat{G} = \{e \mapsto 1\}$ so the result holds.

Otherwise let $h \in G$ be a nonidentity element and let H be the subgroup generated by h. Then $|\hat{H}| = |H|$ since H is cyclic and $\widehat{G/H}$ has order strictly less than G, so $|\widehat{G/H}| = |G/H|$ by the inductive hypothesis.

From the short exact sequence above, we have that $|\widehat{G}| = |\widehat{G/H}||\widehat{H}|| = |G/H||H| = |G|$ so the result follows.

□

We are particularly interested in characters of $(\mathbb{Z}/q\mathbb{Z})^\times$, where q is some positive integer.

Definition 3.4. *A homomorphism $\chi : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}^\times$ is called a character* mod q.

For any prime $p \in \mathbb{Z}$ relatively prime to q, $p + q\mathbb{Z} \in G$. We denote by $f(p)$ the order of $p + q\mathbb{Z}$ in G and by $g(p)$ the index of the subgroup generated by $p + q\mathbb{Z}$ in G. Since $|G| = \varphi(q)$, we have $f(p)g(p) = \varphi(q)$ whenever f, g are defined. Using that notation, we prove the following technical result:

Lemma 3.5. *Let p be a prime and T a variable over $\mathbb{C}$. Then:*

$$\prod_{\chi \in \widehat{G}} 1 - \chi(p)T = (1 - T^{f(p)})^{g(p)}$$

Proof. Let $\omega = e^{2\pi i/f(p)}$ be a primitive $f(p)^{th}$ root of unity. Then we have $f(p)$ distinct characters $\chi_k : (p + q\mathbb{Z}) \to \mathbb{C}^\times$, given by $\chi_k(p + q\mathbb{Z}) = \omega^k$, defined on the (multiplicative) subgroup generated by $p + q\mathbb{Z}$. Each of these characters can be extended to a character of G in $|G/\widehat{(p + q\mathbb{Z})}| = g(p)$ different ways so we have:

$$\prod_{\chi \in \widehat{G}} 1 - \chi(p)T = \left(\prod_{k=0}^{f(p)-1} 1 - \omega^k T \right)^{g(p)} = (1 - T^{f(p)})^{g(p)}$$

□

Lemma 3.6. *(First Orthogonality Theorem) Let p be prime. Then:*

$$\sum_{\chi \in \widehat{G}} \chi(p) = \begin{cases} \varphi(q) & \text{if} \quad p \equiv 1 \bmod q \\ 0 & \text{otherwise} \end{cases}$$

Proof. By the previous lemma we have:

$$\prod_{\chi \in \widehat{G}} 1 - \chi(p)X = (1 - X^{f(p)})^{g(p)}$$

On the one hand:

$$\prod_{\chi \in \widehat{G}} 1 - \chi(p)X = \left(\prod_{\chi \in \widehat{G}} (-\chi(p)) \right) X^{\varphi(q)} + \cdots - \left(\sum_{\chi \in \widehat{G}} \chi(p) \right) X + 1$$

On the other:

$$(1 - X^{f(p)})^{g(p)} = \sum_{i=0}^{g(p)} (-1)^i \binom{g(p)}{i} X^{f(p)i}$$

It is clear from the second display that X has a nonzero coefficient if and only if $f(p) = 1$. From the first display, we see that the coefficient of X is $\sum_{\chi \in \widehat{G}} \chi(p)$, so we have $\sum_{\chi \in \widehat{G}} \chi(p) = 0$ if and only if $f(p) = 1$.

The condition $f(p) = 1$ is equivalent to $p \equiv 1 \bmod q$. In that case, $\chi(p) = 1$ for all $\chi \in \widehat{G}$. Since $|\widehat{G}| = \varphi(q)$, the result follows.

□

Given a character $\bmod q$, χ, we can define a "lift" of χ to a multiplicative function $\tilde{\chi} : \mathbb{Z}^+ \to \mathbb{C}$ by setting:

$$\tilde{\chi}(n) = \begin{cases} \chi(n + q\mathbb{Z}) & \text{if} \quad \gcd(n, q) = 1 \\ 0 & \text{if} \quad \gcd(n, q) > 1 \end{cases}$$

It is easy to verify that $\tilde{\chi}$ is, indeed, multiplicative, i.e. $\tilde{\chi}(nm) = \tilde{\chi}(n)\tilde{\chi}(m)$. If $\gcd(n, q) = \gcd(m, q) = 1$, then $\gcd(nm, q) = 1$ so:

$$\tilde{\chi}(nm) = \chi(nm + q\mathbb{Z}) = \chi(n + q\mathbb{Z})\chi(m + q\mathbb{Z}) = \tilde{\chi}(n)\tilde{\chi}(m)$$

Otherwise, $\gcd(n, q) > 1$ or $\gcd(m, q) > 1$; WLOG assume $\gcd(n, q) > 1$. Then $\gcd(nm, q) > 1$ so:

$$\tilde{\chi}(nm) = 0 = 0 \cdot \chi(m + q\mathbb{Z}) = \tilde{\chi}(n)\tilde{\chi}(m)$$

From here on, we will abuse notation and refer to $\tilde{\chi}$ as χ, with the understanding that $\chi(n) = 0$ if $n + q\mathbb{Z} \notin (\mathbb{Z}/q\mathbb{Z})^\times$. Observe that $\overline{\chi_0}(n) = \begin{cases} 1 & \text{if} \quad \gcd(n, q) = 1 \\ 0 & \text{if} \quad \gcd(n, q) > 1 \end{cases}$. We also allow principal character $\bmod 1$, whose lift is $\chi_0(n) = 1$, for all n.

Definition 3.7. *Let χ be a character* $\bmod q$. *The* Dirichlet L-series for χ *is the formal power series:*

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

Since $\chi(n)$ is multiplicative on $\mathbb{N}$, we immediately get product representations:

$$L(s,\chi) = \prod_p \frac{1}{1-\chi(p)p^{-s}} \qquad (\mathrm{Re}(s) > 1)$$

Furthermore, we see that no L-function $L(s,\chi)$ has a zero in the open half-plane $\mathrm{Re}(s) > 1$: since $L(s,\chi) = \prod_p \frac{1}{1-\chi(p)p^{-s}}$ is an absolutely convergent product, it is nonzero, thus $L(s,\chi) \neq 0$.

If χ_0 is the principal character mod q, then:

$$L(s,\chi_0) = \prod_{p\not| q} \frac{1}{1-p^{-s}} = \zeta(s) \prod_{p|q} 1-p^{-s}$$

Since $\chi(p) = 1$ for all p relatively prime to q, those primes contribute $\frac{1}{1-\chi(p)p^{-s}} = \frac{1}{p^{-s}}$ to the product expansion. The remaining primes divide q, so $\chi(p) = 0$, so they contribute $\frac{1}{1-0\cdot p^{-s}} = 1$ to the product; we may as well ignore them so we have:

$$L(s,\chi_0) = \prod_{p\not| q} \frac{1}{1-p^{-s}} = \left(\prod_p \frac{1}{1-p^{-s}}\right) \cdot \left(\prod_{p|q} 1-p^{-s}\right) = \zeta(s) \prod_{p!q} 1-p^{-s}$$

Proposition 3.8. *The function $\zeta(s) - \frac{1}{s-1}$ is analytic on the half-plane $\mathrm{Re}(s) > 0$.*

Proof. First, observe that:

$$\int_1^\infty \frac{1}{x^s} dx = \left(\frac{1}{x^{s-1}(1-s)}\right|_1^\infty = \lim_{x\to\infty} \frac{1}{x^{s-1}(1-s)} - \frac{1}{1^{s-1}(1-s)} = \frac{1}{s-1}$$

so we may write:

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

The rightmost summation converges absolutely for $\mathrm{Re}(s) > 0$. To prove this, we show $\left|\int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right)\right| \leq \frac{|s|}{n^{\mathrm{Re}(s)+1}}$ by rewriting the integrand as an integral itself:

$$s\int_n^x \frac{du}{u^{s+1}} = s\left(-\frac{1}{su^s}\right|_n^x = \frac{1}{n^s} - \frac{1}{x^s}$$

$$\begin{aligned}
\implies \left|\int_n^{n+1} \frac{1}{n^s} - \frac{1}{x^s} dx\right| &= \left|\int_n^{n+1} s\int_n^x \frac{du}{u^{s+1}} dx\right| = |s| \left|\int_n^{n+1} \int_n^x \frac{du}{u^{s+1}} dx\right| \\
&\leq |s| \int_n^{n+1} \int_n^x \left|\frac{du}{u^{s+1}}\right| dx \leq \max_{n<x\leq n+1} \max_{n\leq u\leq x} \left|\frac{s}{u^{s+1}}\right| \\
&= \max_{n\leq u\leq n+1} \frac{|s|}{u^{\mathrm{Re}(s)+1}} = \frac{1}{n^{\mathrm{Re}(s)+1}}
\end{aligned}$$

The series therefore converges absolutely on $\mathrm{Re}(s) > 0$ as claimed.

□

Corollary 3.9. *Let χ_0 be the principal character mod q for any $q > 0$. Then $L(s, \chi_0) - \frac{\varphi(q)}{q}\frac{1}{s-1}$ extends holomorphically to the right half plane* $\mathrm{Re}(s) > 0$

Proof. We have the representation:

$$L(s, \chi_0) = \zeta(s) \prod_{p|q} 1 - p^{-s}$$

The product $\prod_{p|q} 1 - p^{-s}$ is clearly analytic, being a finite product of analytic functions, and it is clearly nonzero on the right half plane $\mathrm{Re}(s) > 0$, and $\zeta(s)$ has a simple pole at $s = 1$ with residue 1. Consequently, $L(s, \chi_0)$ is a meromorphic function with a simple pole at 1. The residue is $\prod_{p|q} 1 - p^{-1} = \frac{\varphi(q)}{q}$, so $L(s, \chi_0) = f(s) + \frac{\varphi(q)}{q}\frac{1}{1-s}$ for some holomorphic function f, so $L(s, \chi_0) - \frac{\varphi(q)}{q}\frac{1}{1-s} = f(s)$ which is holomorphic on the desired half plane.
□

We can also extend $L(s, \chi)$, for non-principal χ, to the full half-plane using a different method.

Lemma 3.10. *(Second Orthogonality Theorem) Let χ be a non-principal character mod q, and let $G = (\mathbb{Z}/q\mathbb{Z})^\times$. We have:*

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

Proof. Since χ is not principal, there exists $g_0 \in G$ such that $\chi(g_0) \neq 1$. Compute that:

$$\chi(g_0) \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(g)\chi(g_0) = \sum_{g \in G} \chi(gg_0)$$

But $\sum_{g \in G} \chi(gg_0) = \sum_{g \in G} \chi(g)$, so if $\sum_{g \in G} \chi(g) \neq 0$, we could divide both sides by $\sum_{g \in G} \chi(g)$ and contradict the assumption $\chi(g_0) \neq 1$. Thus $\sum_{g \in G} \chi(g) = 0$. □

Theorem 3.11. *Let χ be a non-principal character mod q. Then $L(s, \chi)$ extends holomorphically to the right half plane* $\mathrm{Re}(s) > 0$.

Proof. Define $A(x) = \sum_{n \leq x} \chi(n)$. Compute:

$$\begin{aligned}
\frac{A(x)}{x^s} + s\int_1^x \frac{A(t)dt}{t^{s+1}} &= \frac{A(x)}{x^s} + s\sum_{i=1}^{\lfloor x \rfloor - 1} A(i) \int_i^{i+1} \frac{dt}{t^{s+1}} + sA(x)\int_{\lfloor x \rfloor}^x \frac{dt}{t^{s+1}} \\
&= \frac{A(x)}{x^s} + s\sum_{i=1}^{\lfloor x \rfloor - 1} A(i)\, \frac{-1}{st^s}\Big|_i^{i+1} + sA(x)\, \frac{-1}{st^s}\Big|_{\lfloor x \rfloor}^x \\
&= \frac{A(x)}{x^s} + \sum_{i=1}^{\lfloor x \rfloor - 1} A(i)\left(\frac{1}{i^s} - \frac{1}{(i+1)^s}\right) + A(x)\left(\frac{1}{\lfloor x \rfloor^s} - \frac{1}{x^s}\right) \\
&= \sum_{i=1}^{\lfloor x \rfloor - 1} A(i)\left(\frac{1}{i^s} - \frac{1}{(i+1)^s}\right) + A(x)\frac{1}{\lfloor x \rfloor^s}
\end{aligned}$$

Now we have something that looks like a telescoping series. To exploit that property, we rewrite A as a sum and then switch the order of summation:

$$\frac{A(x)}{x^s} + s\int_1^x \frac{A(t)dt}{t^{s+1}} = \sum_{i=1}^{\lfloor x \rfloor - 1} A(i)\frac{1}{i^s} - \sum_{j=2}^{\lfloor x \rfloor} A(j)\frac{1}{j^s} \qquad (j = i+1)$$

$$\sum_{n \leq x} \frac{\chi(n)}{n^s} = \frac{A(x)}{x^s} + s\int_1^x \frac{A(t)dt}{t^{s+1}}$$

Then we can write:

$$L(s,\chi) = s\int_1^\infty \frac{A(t)dt}{t^{s+1}}$$

By the preceding lemma, $A(rq) = 0$ for all integer multiples of q, so A is periodic, hence bounded ($|A(x)| \leq \max_{1 \leq \xi \leq n} |A(\xi)| := m$ for all $x \in \mathbb{R}^+$), so:

$$\left|\int_1^\infty \frac{A(t)dt}{t^{s+1}}\right| \leq \int_1^\infty \left|\frac{A(t)dt}{t^{s+1}}\right| \leq m\int_1^\infty t^{-(\mathrm{Re}(s)+1)}dt = \frac{m}{-\mathrm{Re}(s)}\left(t^{-\mathrm{Re}(s)}\Big|_1^\infty\right.$$

which converges on $\mathrm{Re}(s) > 0$.

□

Next, define:

$$\zeta_q(s) = \prod_{\chi \in \widehat{G}} L(s,\chi)$$

Since ζ_q is a finite product of meromorphic functions on $\mathrm{Re}(s) > 0$, it is also meromorphic on $\mathrm{Re}(s) > 0$. We would like to show that $\zeta_q(s)$ has a simple pole at 1.

Proposition 3.12.

$$\zeta_q(s) = \prod_{p \nmid q} \frac{1}{\left(1 - \frac{1}{p^{f(p)s}}\right)^{g(p)}}$$

Furthermore, this is a Dirichlet series with nonnegative coefficients.

Proof. First, let us derive the representation. We substitute the product representation of $L(s, \chi)$ into the definition of $\zeta_q(s)$, and then use the fact that one of our products is finite and the other absolutely convergent to rearrange terms. We use Lemma 3.5, with $X = p^{-s}$to complete the argument:

$$\zeta_q(s) = \prod_{p \nmid q} \prod_{\chi \in \widehat{G}} \frac{1}{1 - \chi(p)p^{-s}} = \prod_{p \nmid q} \frac{1}{\prod_{\chi \in \widehat{G}} 1 - \chi(p)p^{-s}} = \prod_{p \nmid q} \frac{1}{\left(1 - \frac{1}{p^{f(p)s}}\right)^{g(p)}}$$

Thus we have the desired product representation. Now the p-th factor is:

$$\frac{1}{\left(1 - \frac{1}{p^{f(p)s}}\right)^{g(p)}} = \left(\sum_{j=0}^{\infty} p^{-f(p)sj}\right)^{g(p)}$$

so ζ_q is a product of terms of the form $\sum_{j=0}^{\infty} p^{-f(p)sj}$ over all primes not dividing q. Each of these "subfactors" is clearly a Dirichlet series with nonnegative coefficients, and the product of Dirichlet series with nonnegative coefficients is a Dirichlet series with nonnegative coefficients, so the result follows.

□

The following theorem is quite important, and turns out to be the key ingredient in most proofs of Dirichlet's theorem. There seems to be some consensus that the nicest way to prove the result is by studying zeta functions of algebraic number fields [3], but that proof is hard to find. Many people, including [4], simply omit this proof. The proof in here is due to Serre. The key idea in the proof is the fact that the domain of convergence of a Dirichlet series is an open half-plane, and if the boundary of that half-plane does not contain a pole, then it can be extended via Prop. 2.4.

Theorem 3.13. *Let χ be a non-principal character* $\bmod q$ *for some q. Then $L(1, \chi) \neq 0$.*

Proof. Suppose $L(1, \chi) = 0$ for some nontrivial χ. Then the pole that ζ_q would have at 1 becomes a removable singularity. Since ζ_q is a Dirichlet series with nonnegative coefficients, and since it extends holomorphically to $\mathrm{Re}(s) \geq 0$, it converges to its Dirichlet series representation on that entire halfplane. Consequently, the product expansion must be valid on the entire half-plane.

But this cannot be. The p^{th} factor in the product expansion is:

$$\frac{1}{(1-p^{-f(p)s})^{g(p)}} = (1 + p^{-f(p)s} + p^{-2f(p)s} + \cdots)^{g(p)}$$

On the real axis (i.e. $\mathrm{Im}\,(s) = 0$), that factor clearly dominates the series:

$$1 + p^{-\phi(q)s} + p^{-2\phi(q)s} + \cdots$$

because the cross terms, which are the terms we've dropped, are all positive. Consequently, the coefficients of ζ_m are greater than those of:

$$\sum_{n\in\mathbb{N}:\gcd(n,q)=1} n^{-\phi(q)s}$$

If we take $s = \frac{1}{\phi(q)}$, the series above becomes:

$$\sum_{n\in\mathbb{N}:\gcd(n,q)=1} n^{-1} \geq \sum_{k\in\mathbb{N}} (1+kq)^{-1} < \sum_{k\in\mathbb{N}} (kq)^{-1} = q^{-1} \sum_{k\in N} k^{-1} = \infty$$

which diverges. This is absurd - thus, the product expansion of $\zeta_q(s)$ cannot converge at $s = \frac{1}{\phi}$, so the Dirichlet series ζ_q must have a pole at $s = 1$, so $L(s, \chi) \neq 0$.

□

4. Major Theorems

We are now in position to make some meaningful progress towards our problem. Let us define some new functions. Fix a positive integer q, let $G = (\mathbb{Z}/q\mathbb{Z})^\times$, let $\overline{G}$ be the group of characters of G and let $a \in \mathbb{Z}$, $\gcd(a, q) = 1$. Set:

$$\Phi(s,\chi) = \sum_p \frac{\chi(p)\log(p)}{p^s}, \qquad \Psi_q(s) = \sum_{\chi\in\widehat{G}} \Phi(s,\chi), \qquad \Psi_{q,a}(s) = \sum_{\chi\in\widehat{G}} \overline{\chi(a)}\Phi(s,\chi)$$

Lemma 4.1. *For all $\epsilon > 0$, $\lim_{n\to\infty} \frac{\log(n)}{n^\epsilon} \to 0$.*

Proof.

$$\lim_{n\to\infty} \frac{\log(n)}{n^\epsilon} = \lim_{n\to\infty} \frac{\frac{1}{\epsilon}\log(n^\epsilon)}{n^\epsilon} = \frac{1}{\epsilon} \lim_{n^\epsilon:=m\to\infty} \frac{\log(m)}{m} = 0$$

□

Proposition 4.2. *$\int_{x=1}^{\infty} \frac{\log(x)}{x^s} dx$ converges for $\mathrm{Re}(s) > 1$.*

Proof. We can evaluate the following integral using the previous lemma:

$$\begin{aligned}\int_{x=1}^{\infty} \frac{\log(x)}{x^s} dx &= \left(-\frac{x^{1-s}((s-1)\log(x)+1)}{(s-1)^2} \right|_1^\infty \\ &= \lim_{t\to\infty} \left(-\frac{t^{1-s}((s-1)\log(t)+1)}{(s-1)^2} \right) + \frac{1}{(s-1)^2}\end{aligned}$$

With $\epsilon = \mathrm{Re}(s) - 1 > 0$, we have $|t^{1-s}\log(t)| = t^{-\epsilon}\log(t)$ so the limit goes to zero and the integral converges to $\frac{1}{(s-1)^2}$. □

Corollary 4.3. *$\Phi(s,\chi)$ converges absolutely for $\mathrm{Re}(s) > 1$.*

Proof.

$$\sum_p \left| \frac{\chi(p)\log(p)}{p^s} \right| = \sum_p \frac{\log(p)}{p^{\mathrm{Re}(s)}} \le \sum_n \frac{\log(n)}{n^{\mathrm{Re}(s)}} < \infty$$

□

Theorem 4.4. *For any χ and s satisfying $\mathrm{Re}(s) > 1$, we have:*

$$-\frac{L'(s,\chi)}{L(s,\chi)} = \Phi(s,\chi) + h(s,\chi),$$

where $h(s,\chi)$ is holomorphic for $\mathrm{Re}(s) > \frac{1}{2}$.

Proof. By routine calculations, one obtains:

$$-\frac{L'(s,\chi)}{L(s,\chi)} = -\frac{d}{ds}\log L(s,\chi) = -\frac{d}{ds}\log\left(\prod_p \frac{1}{1-\chi(p)p^{-s}}\right) = -\sum_p \frac{d}{ds}\log\left(\frac{1}{1-\chi(p)p^{-s}}\right)$$
$$= -\sum_p (1-\chi(p)p^{-s})\frac{(-\chi(p)p^{-s}\log(p))}{(1-\chi(p)p^{-s})^2} = \sum_p \frac{\chi(p)p^{-s}\log(p)}{1-\chi(p)p^{-s}} = \sum_p \frac{\chi(p)\log(p)}{p^s-\chi(p)}$$

Now, we use the identity $\frac{1}{a-b} = \frac{1}{a} + \frac{b}{a(a-b)}$, with $a = p^s$ and $b = \chi(p)$ to rewrite the last summation:

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

It remains to show that $\sum_p \frac{\chi(p)^2\log(p)}{p^s(p^s-\chi(p))}$ is holomorphic for $\mathrm{Re}(s) > \frac{1}{2}$. Compute:

$$\left|\frac{\chi(p)^2\log(p)}{p^s(p^s-\chi(p))}\right| = \frac{\log(p)}{p^{\mathrm{Re}(s)}|p^s-\chi(p)|} \leq \frac{\log(p)}{p^{\mathrm{Re}(s)}(p^{\mathrm{Re}(s)}-1)} \leq \frac{\log(p)}{(p-1)^{2\mathrm{Re}(s)}}$$

By the preceding proposition and the integral test, it is clear that we have absolute convergence on $\mathrm{Re}(s) > \frac{1}{2}$.

□

Corollary 4.5. *$\Phi(s,\chi)$ is meromorphic on $\mathrm{Re}(s) \geq 1$, with a simple pole at $s = 1$, which carries a residue of 1, if and only if χ is principal. If χ is not principal, then $\Phi(s,\chi)$ is holomorphic on $\mathrm{Re}(s) \geq 1$.*

Proof. We have shown that the logarithmic derivative of $L(s,\chi)$ is equal to $\Phi(s,\chi) + h(s,\chi)$, where $h(s,\chi)$ is holomorphic. We know that the residue of the logarithmic derivative at a point is equal to the multiplicity of the zero at that point. For non-principal χ, we know that $L(s,\chi)$ has no zeros on the plane $\mathrm{Re}(s) \geq 1$, so the logarithmic derivative of $L(s,\chi)$ cannot have any poles on that plane. Consequently, $\Phi(s,\chi)$ has no poles on $\mathrm{Re}(s) \geq 1$ if χ is non-principal, and $\Phi(s,\chi_0)$ has a pole at 1 which corresponds to the simple pole of $L(s,\chi_0)$ at 1, and thus carries residue 1.

□

Corollary 4.6.

$$-\frac{\zeta_q'(s)}{\zeta_q(s)} = \Psi_q(s) + h(s)$$

where $h(s)$ is holomorphic on $\mathrm{Re}(s) > \frac{1}{2}$.

Proof. This follows easily from the previous lemma:

$$-\frac{\zeta_q'(s)}{\zeta_q(s)} = \sum_{\chi \in \widehat{G}} -\frac{L'(s,\chi)}{L(s,\chi)} = \sum_{\chi \in \widehat{G}} \Phi(s,\chi) + h(s,\chi) = \Psi_q(s) + h(s)$$

where $h(s) = \sum_{\chi \in \widehat{G}} h(s,\chi)$ is clearly holomorphic on $\mathrm{Re}(s) > \frac{1}{2}$.

□

Theorem 4.7. *For any character mod q, χ, the L-function $L(s,\chi)$ has no zeros in the half-plane* $\mathrm{Re}(s) \geq 1$.

Proof. Since $L(1,\chi_0) = \zeta(1)\prod_{p|q} 1 - p^{-1}$, $\prod_{p|q} 1 - p^{-1}$ is a finite product and ζ has a simple pole at 1, $L(1,\chi_0)$ is a simple pole. Furthermore, $L(1,\chi) \neq 0$ for all non-principal χ, so $\mathcal{L}$ has a simple pole at 1.

We have proven the result for $\mathrm{Re}(s) > 1$ and $s = 1$. Now suppose ζ_q has a zero of order $\mu \geq 0$ at some $t = 1 + i\alpha$ with $\alpha \neq 0$. Denote by ν the order of the zero at $u = 1 + 2i\alpha$.

Observe that, for any $s \in \mathbb{R}$:

$$\overline{\zeta_q(s)} = \overline{\prod_{\chi \in \widehat{G}} L(s,\chi)} = \prod_{\chi \in \widehat{G}} \overline{L(s,\chi)} = \prod_{\chi \in \widehat{G}} L(s,\overline{\chi}) = \prod_{\chi \in \widehat{G}} L(s,\chi) = \zeta_q(s)$$

so $\zeta_q(\mathbb{R}) \subset \mathbb{R}$. As a result, ζ_q must also have zeros of order μ, ν at $\overline{t}, \overline{u}$, respectively.

By the argument principle, we know the residue of the logarithm derivative at a point is equal to the multiplicity of the zero at that point. Furthermore, by Corollary 4.6, we have that $\frac{\zeta_q'(s)}{\zeta_q(s)} = \Psi_q(s) + h(s)$, where $h(s)$ is analytic, and thereby does not contribute to the residue. We summarize explicitly (while negating the arguments to clean up the notation):

$$\begin{aligned}
\mathrm{Res}_1(\Psi_q) &= \lim_{\epsilon \to 0} \epsilon \Psi_q(1+\epsilon) = 1 \\
\mathrm{Res}_{t,\bar{t}}(\Psi_q) &= \lim_{\epsilon \to 0} \epsilon \Psi_q(1+\epsilon \pm i\alpha) = -\mu \\
\mathrm{Res}_{s,\bar{s}}(\Psi_q) &= \lim_{\epsilon \to 0} \epsilon \Psi_q(1+\epsilon \pm 2i\alpha) = -\nu
\end{aligned}$$

Now compute that:

$$\begin{aligned}\sum_{r=-2}^{2}\binom{2+r}{4}\Psi_q(1+\epsilon+ri\alpha) &= \sum_{r=-2}^{2}\binom{2+r}{4}\left(\sum_{\chi\in\widehat{G}}\sum_{p}\frac{\chi(p)\log(p)}{p^{1+\epsilon+ri\alpha}}\right)\\ &= \sum_{p}\frac{\log p}{p^{1+\epsilon}}\left(\sum_{\chi\in\widehat{G}}\chi(p)\right)\left(\sum_{r=-2}^{2}\binom{2+r}{4}\frac{1}{p^{ri\alpha}}\right)\\ &= \varphi(q)\sum_{p\equiv 1 \bmod q}\frac{\log p}{p^{1+\epsilon}}\left(p^{i\alpha/2}+p^{-i\alpha/2}\right)^4\end{aligned}$$

Thus on the one hand, we have:

$$\lim_{\epsilon\to 0}\epsilon\sum_{r=-2}^{2}\binom{2+r}{4}\Psi_q(1+\epsilon+ri\alpha) = \sum_{r=-2}^{2}\binom{2+r}{4}\mathrm{Res}_{1+ri\alpha}(\Psi_q) = 6-8\mu-2\nu$$

On the other hand:

$$\begin{aligned}\lim_{\epsilon\to 0}\epsilon\varphi(q)\sum_{p\equiv 1 \bmod q}\frac{\log p}{p^{1+\epsilon}}\left(p^{i\alpha/2}+p^{-i\alpha/2}\right)^4 &= \varphi(q)\lim_{\epsilon\to 0}\epsilon\sum_{p\equiv 1 \bmod q}\frac{\log p}{p^{1+\epsilon}}\left(e^{i\log p\alpha/2}+e^{-i\log p\alpha/2}\right)^4\\ &= 2\varphi(q)\cos^4(\ln(p|\alpha|/2))\lim_{\epsilon\to 0}\epsilon\sum_{p\equiv 1 \bmod q}\frac{\log p}{p^{1+\epsilon}} \geq 0\end{aligned}$$

Since $\mu,\nu\geq 0$, $6-8\mu-2\nu\geq 0$ we have $6-8\mu-2\nu\geq 6-8\mu\geq 0$ so $\mu=0$. Thus ζ_q has no zeros on the line $\mathrm{Re}(s)=1$.

□

Theorem 4.8. *Let a be relatively prime to q. Then $\Psi_{q,a}(s)-\frac{1}{s-1}$ is holomorphic for $\mathrm{Re}(s)\geq 1$.*

Proof. By definition we have:

$$\Psi_{q,a}(s) = \overline{\chi(a)}\sum_{\chi\in\widehat{G}}\Phi(s,\chi) = \Phi(s,\chi_0)+\overline{\chi(a)}\sum_{\chi\in\widehat{G}:\chi\neq\chi_0}\Phi(s,\chi)$$

By Theorem 4.4, we have:

$$-\frac{L'(s,\chi)}{L(s,\chi)} = \Phi(s,\chi)+h(s,\chi)$$

so:

$$\Psi_{q,a}(s) = \Phi(s,\chi_0)-\overline{\chi(a)}\sum_{\chi\in\widehat{G}:\chi\neq\chi_0}\frac{L'(s,\chi)}{L(s,\chi)}+h(s)$$

where $h(s)$ is holomorphic on $\mathrm{Re}(s) > \frac{1}{2}$. Since $L(s,\chi)$ is finite and nonzero for all Re$s \geq 1$, $\chi \neq \chi_0$, we have $\frac{L'(s,\chi)}{L(s,\chi)} = 0$ for all such s, χ so the summation disappears and we have:

$$\Psi_{q,a}(s) = \Phi(s,\chi_0) + h(s)$$

Thus it suffices to show $\Phi(s,\chi_0) - \frac{1}{s-1}$ is holomorphic on $\mathrm{Re}(s) \geq 1$. This is Corollary 4.5.

.

□

5. Main Theorem and Corollaries

Let a, q be relatively prime integers and define:

$$\vartheta_{q,a}(x) = \varphi(q) \sum_{\substack{p \leq x \\ p \equiv a \bmod q}} \log(p)$$

Lemma 5.1. $\vartheta_{q,a}(x) = \mathcal{O}(x)$.

Proof. We have:

$$\vartheta_{q,a}(x) = \varphi(q) \sum_{\substack{p \leq x \\ p \equiv a \bmod q}} \log(p) \leq \varphi(q) \sum_{p \leq x} \log p$$

so it suffices to prove the result for $\vartheta(x) := \sum_{p \leq x} \log p$.

Now compute that:

$$2^{2n} = (1+1)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} \geq \binom{2n}{n} = \frac{(2n)!}{(n!)^2}$$

If p is a prime and $n < p < 2n$, then $p | (2n!)$ but $p \not| (n!)^2$ so $p \left| \binom{2n}{n} \right.$. Thus $\prod_{n<p<2n} p \left| \binom{2n}{n} \right.$ so $2^{2n} \geq \prod_{n<p<2n} p = e^{\vartheta(2n)-\vartheta(n)}$, so:

$$2n \ln 2 \geq \vartheta(2n) - \vartheta(n)$$

With $\delta(x) := \vartheta(x) - \vartheta(x/2)$, the previous display shows $\delta(x) = \mathcal{O}(x)$. For any $C > \ln 2$, we can find x_C such that for all $x \geq x_C$, $\delta(x) \leq Cx$.

Fix $x \geq x_C$ and let r be the largest integer for which $\frac{x}{2^r} \geq x_0$. Then we have:

$$\delta\left(\frac{x}{2^k}\right) \leq C\frac{x}{2^k} \quad (k = 0, 1, \cdots, r) \Longrightarrow \sum_{k=0}^{r} \delta\left(\frac{x}{2^k}\right) = \theta(x) - \vartheta(x/2^r) \leq Cx \sum_{k=0}^{r} \frac{1}{2^k} \leq Cx \sum_{k=0}^{\infty} \frac{1}{2^k}$$

By maximality of r and monotonicity of ϑ, we have:

$$\frac{x}{2^{r+1}} < x_0 \leq \frac{x}{2^r} \Longrightarrow \frac{x}{2^r} < 2x_0 \Longrightarrow \vartheta\left(\frac{x}{2^r}\right) < \vartheta(2x_0)$$

so:

$$\vartheta(x) \leq Cx \sum_{k=0}^{\infty} \frac{1}{2^k} + \vartheta\left(\frac{x}{2^r}\right) \leq 2Cx + \vartheta(2x_0) \Longrightarrow \vartheta(x) = \mathcal{O}(x)$$

□

Theorem 5.2. *The integral:*

$$\int_1^\infty \frac{\vartheta_{q,a}(x) - x}{x^2} dx$$

converges.

Proof. Let a^{-1} be a positive integer satisfying $aa^{-1} \equiv 1 \bmod q$, i.e. a^{-1} is the multiplicative inverse of a in G, as suggested by the notation. Recall that $\chi(a^{-1}) = \overline{\chi(a)}$, and compute, using Lemma 3.6:

$$\begin{aligned}\Psi_{q,a}(s) &= \sum_{\chi\in\widehat{G}} \overline{\chi(a)}\Phi(s,\chi) = \sum_{\chi\in\widehat{G}} \chi(a^{-1}) \sum_p \frac{\chi(p)\log p}{p^s} = \sum_{\chi\in\widehat{G}}\sum_p \frac{\chi(a^{-1}p)\log p}{p^s} \\ &= \sum_p \left(\sum_{\chi\in\widehat{G}} \chi(a^{-1}p)\right) \frac{\log p}{p^s} = \sum_{p:a^{-1}p\equiv 1 \bmod q} \varphi(q)\frac{\log p}{p^s} = \sum_{p\equiv a \bmod q} \varphi(q)\frac{\log p}{p^s}\end{aligned}$$

Thus, $\Psi_{q,a}$ looks like $\vartheta_{q,a}$ if we were to let the summation go to infinity while dividing the summands by powers of p. We can exploit that seemingly superficial resemblance to rewrite the summation in the previous display as a Riemann-Stieltjes integral, and then apply integration by parts to obtain yet another representation for Ψ.

$$\Psi_{q,a}(s) = \int_1^\infty \frac{d\vartheta_{q,a}(x)}{x^s} = \frac{\vartheta_{q,a}(x)}{x^s}\bigg|_1^\infty + s\int_1^\infty \frac{\vartheta_{q,a}(x)}{x^{s+1}}dx = \lim_{t\to\infty}\frac{\vartheta_{q,a}(t)}{t^{s+1}} - \frac{\vartheta_{q,a}(1)}{1^s} + s\int_1^\infty \frac{\vartheta_{q,a}(x)}{x^{s+1}}dx$$

Obviously $\vartheta_{q,a}(1) = 0$ because there are no primes $p \le 1$ so that term disappears. Thus:

$$\Psi_{q,a}(s) = s\int_1^\infty \frac{\vartheta_{q,a}(x)}{x^{s+1}}dx$$

Since x is a positive real number, it is perfectly sensible to use the substitution $x = e^t$ where $t = \ln x$. Plugging that into the previous display:

$$\Psi_{q,a}(s) = s\int_1^\infty \vartheta_{q,a}(e^t)e^{-(s+1)t}(e^t dt) = s\int_1^\infty \vartheta_{q,a}(e^t)e^{-st}dt$$

We would like to use the analytic theorem, so define:

$$f(t) = \vartheta_{q,a}(e^t)e^{-t} - 1, \qquad g(s) = \int_0^\infty f(t)e^{-st}dt$$

Let us verify that the conditions of the theorem are met. The function f is bounded, because $\vartheta_{q,a} = \mathcal{O}(x)$ which is no match for e^{-t}. Since $\vartheta_{q,a}$ is a step-function with points of discontinuity only at primes $p \equiv a \bmod q$, it is locally integrable, so f is locally integrable and hence satisfies all of the conditions of the analytic theorem.

On the other hand:

$$g(s) = \int_0^\infty (\vartheta_{q,a}(e^t)e^{-t} - 1)e^{-st}dt = \int_0^\infty \vartheta_{q,a}(e^t)e^{-t(1+s)}dt - \int_0^\infty e^{-st}dt$$
$$= \frac{\Psi_{q,a}(s+1)}{s+1} - \frac{e^{-st}}{-s}\Big|_0^\infty = \frac{\Psi_{q,a}(s+1)}{s+1} - \frac{1}{s} = \frac{\Psi_{q,a}(z)}{z} - \frac{1}{z-1}$$

Now, $\Psi_{q,a}(z) - \frac{1}{z-1}$ is holomorphic for $\mathrm{Re}(z) \geq 1$ so $\frac{\Psi_{q,a}(z)}{z}$ is also meromorphic for $\mathrm{Re}(z) \geq 1$, with a simple pole at $z = 1$ with residue still 1, so $\frac{\Psi_{q,a}(z)}{z} - \frac{1}{z-1}$ is holomorphic for $\mathrm{Re}(z) \geq 1$. Since $z = s+1$, $g(s)$ is holomorphic on $\mathrm{Re}(s) \geq 0$ so g satisfies the conditions of the analytic theorem.

Hence, we conclude that:

$$\int_0^\infty f(t) = g(0) = \int_0^\infty f(t)e^{-0\cdot t}dt = \int_0^\infty \vartheta_{q,a}(e^t)e^{-t} - 1dt = \int_0^\infty \frac{\vartheta_{q,a}(e^t) - e^t}{e^t}dt$$

We reverse our previous substitution: e^t becomes x, 0 becomes 1 and dt becomes $\frac{dx}{x}$.

$$g(0) = \int_0^\infty \frac{\vartheta_{q,a}(e^t) - e^t}{e^t}dt = \int_1^\infty \frac{\vartheta_{q,a}(x) - x}{x}\frac{dx}{x} = \int_1^\infty \frac{\vartheta_{q,a}(x) - x}{x^2}dx$$

Thus the desired integral converges because $g(0)$ converges.

□

Corollary 5.3. $\vartheta_{q,a}(x) \sim x$

Proof. Since $\int_1^\infty \frac{\vartheta_{q,a}(x)-x}{x^2}dx$ converges, $\frac{\vartheta_{q,a}(x)-x}{x^2} \to 0$. Thus $\frac{\vartheta_{q,a}(x)}{x^2} \to \frac{x}{x^2} = \frac{1}{x}$ so $\frac{\vartheta_{q,s}(x)}{x} \to 1$, i.e. $\vartheta_{q,s}(x) \sim x$.

□

Theorem 5.4.

$$\pi(x,q) \sim \frac{1}{\varphi(q)}\frac{x}{\log x}$$

Proof. To prove the theorem, we need to show that:

$$\lim_{x\to\infty} \frac{\pi(x,q,a)}{\frac{x}{\varphi(q)\log(x)}} = \lim_{x\to\infty} \frac{\pi(x,q,a)\varphi(q)\log(x)}{x} = 1$$

First, observe that:

$$\vartheta_{q,a}(x) = \varphi(q) \sum_{\substack{p \leq x \\ p \equiv a \bmod q}} \log(p) \leq \varphi(q) \sum_{\substack{p \leq x \\ p \equiv a \bmod q}} \log(x) = \varphi(q)\pi(x,q,a)\log x$$

Now let $\epsilon > 0$.

$$\begin{aligned}\vartheta_{q,a}(x) &\geq \varphi(q) \sum_{\substack{x^{1-\epsilon} \leq p \leq x \\ p \equiv a \bmod q}} \log(p) \\ &\geq \varphi(q) \sum_{\substack{x^{1-\epsilon} \leq p \leq x \\ p \equiv a \bmod q}} (1-\epsilon)\log(x) \\ &= \varphi(q)(1-\epsilon)\log(x)(\pi(x,q,a) - \pi(x^{1-\epsilon},q,a))\end{aligned}$$

Thus we've shown:

$$\begin{aligned}&\varphi(q)(1-\epsilon)\log(x)(\pi(x,q,a) - \pi(x^{1-\epsilon},q,a)) \leq \vartheta_{q,a}(x) \\ \implies &\varphi(q)\log(x)\pi(x,q,a) \leq \frac{1}{1-\epsilon}\vartheta_{q,a}(x) + \varphi(q)\log(x)(\pi(x^{1-\epsilon},q,a) \\ \implies &\frac{\vartheta_{q,a}(x)}{x} \leq \frac{\varphi(q)\log(x)\pi(x,q,a)}{x} \leq \frac{1}{1-\epsilon}\frac{\vartheta_{q,a}(x)}{x} + \frac{\varphi(q)\log(x)\pi(x^{1-\epsilon},q,a)}{x}\end{aligned}$$

Now $\pi(x,q,a) \leq x$ for all x so $\pi(x^{1-\epsilon},q,a) \leq x^{1-\epsilon}$ and hence $\frac{\pi(x^{1-\epsilon},q,a)}{x} \leq x^{-\epsilon}$, so we can rewrite the previous display as:

$$\frac{\vartheta_{q,a}(x)}{x} \leq \frac{\varphi(q)\log(x)\pi(x,q,a)}{x} \leq \frac{1}{1-\epsilon}\frac{\vartheta_{q,a}(x)}{x} + \frac{\varphi(q)\log(x)}{x^{\epsilon}}$$

When $x \to \infty$, $\frac{\varphi(q)\log(x)}{x^{\epsilon}} \to 0$ and $\frac{\vartheta_{q,a}(x)}{x} \to 1$ by the previous theorem, so we have:

$$1 \leq \lim_{x\to\infty} \frac{\varphi(q)\log(x)\pi(x,q,a)}{x} \leq \frac{1}{1-\epsilon}$$

Let $\epsilon \to 0$ and the result follows.

□

Corollary 5.5.

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

Proof. Let $q = a = 1$.

□

Corollary 5.6. *For any relatively prime a, q, there are infinitely many prime numbers in the arithmetic sequence $a, a+q, a+2q, \cdots$. Furthermore, the primes are evenly distributed amongst all such arithmetic sequences for a fixed q.*

Proof. For all relatively prime a, q, $\pi(x, q, a) \sim \frac{\varphi(q)x}{\log x}$, so in particular, the number of primes congruent to a mod q cannot be finite. Since there are $\varphi(q)$ choices of arithmetic sequences, the number of primes overall tends to $\frac{x}{\log x}$ and the density in each of those sequences is $\frac{1}{\varphi(q)}\frac{x}{\log x}$, the result follows. □

References

[1] L.V. Ahlfors. *Complex Analysis: an Introduction to the Theory of Analytic Functions of One Complex Variable*. International Series in Pure and Applied Mathematics, 1979.

[2] D. J. Newman, *Simple analytic proof of the prime number theorem*, The American Mathematical Monthly, Vol. 87, (1980), 693-696.

[3] J.P. Serre *A Course In Arithmetic*. Springer-Verlag, 1973.

[4] I. Soprounov. *The Prime Number Theorem for Arithmetic Sequences*. 1998.

[5] D. Zagier. *Newman's Short Proof of the Prime Number Theorem*. The American Mathematical Monthly, Vol. 104, No. 8. 1997.