Dirichlet's Theorem on Primes in Arithmetic Progressions

By: William Wright 260415362

April 25, 2013

The goal of this paper is to give a proof of Dirichlet's Theorem by showing various facts about L-functions. The Dirichlet Theorem is stated as follows:

Theorem: Let $a, m \in \mathbb{Z}$ be such that $m > 1$, and $gcd(a, m) = 1$. Then there are infinitely many primes congruent to a modulo m.

Before we can prove this however, we need to familiarize ourselves with a few definitions.

Definition 1: Let G be a finite abelian group. Then a character χ is a group homomorphism of G to the group of complex numbers under multiplication $\mathbb{C}^*$.

Propostion 1: The set containing all characters χ of G is a multiplicative group, and is denoted $\mathrm{Hom}(G, \mathbb{C}^*)$, ie. $(\chi_1 \cdot \chi_2)(g) = \chi_1(g) \cdot \chi_2(g)$ for all $g \in G$.

Proof: Consider the group homomorphism $\chi_e : G \to \mathbb{C}^*$ defined by $\chi_e(g) = 1$ for all $g \in G$. Then for any character χ of G, we have:

$$(\chi_e \cdot \chi)(g) = \chi_e(g) \cdot \chi(g) = \chi(g) = \chi(g) \cdot \chi_e(g) = (\chi \cdot \chi_e)(g)$$

Therefore, χ_e is the identity element in $\mathrm{Hom}(G, \mathbb{C}^*)$. From now on, we will denote χ_e by 1.

Since every element in the complex numbers has a multiplicative inverse, if χ is a character of G, there will always exist another character of G, χ^{-1} such that $\chi^{-1}(g) = \frac{1}{\chi(g)}$ for all $g \in G$. Therefore, we have that:

$$(\chi \cdot \chi^{-1})(g) = (\chi^{-1} \cdot \chi)(g) = \frac{\chi(g)}{\chi(g)} = 1 \text{ for all } g \in G$$

Thus, every character of G has an inverse.

If χ_1, χ_2 are two characters of G, $g_1, g_2 \in G$, then

$$\begin{aligned}(\chi_1 \cdot \chi_2)(g_1 g_2) &= \chi_1(g_1 g_2)\chi_2(g_1 g_2) \\ &= \chi_1(g_1)\chi_1(g_2)\chi_2(g_1)\chi_2(g_2) \\ &= (\chi_1 \cdot \chi_2)(g_1)(\chi_1 \cdot \chi_2)(g_2).\end{aligned}$$

This shows that characters of G are closed under multiplication.

If χ_1, χ_2, χ_3 are three characters of G, $g \in G$, then

$$\begin{aligned}((\chi_1 \cdot \chi_2) \cdot \chi_3)(g) &= (\chi_1 \cdot \chi_2)(g)\chi_3(g) \\ &= \chi_1(g)\chi_2(g)\chi_3(g) \\ &= \chi_1(g)(\chi_2 \cdot \chi_3)(g) \\ &= (\chi_1 \cdot (\chi_2 \cdot \chi_3))(g).\end{aligned}$$

Therefore, characters of G are associative, and thus $\mathrm{Hom}(G, \mathbb{C}^*)$ is a group under multiplication.

$\boxdot$

Definition 2: $Hom(G, \mathbb{C}^*)$ is called the dual of G, and will now be denoted $\hat{G}$.

Definition 3: A group G is called divisible if for all pairs (n, g), where $n \in \mathbb{Z}^+$, and $g \in G$, there exists $y \in G$ such that $n \cdot y = g$.

Proposition 2: Let $H \leq G$ be a subgroup of G. Then every character of H can be extended to a character of G.

Proof: To prove this proposition, we will induct on the index of G and H: $[G : H] = \frac{|G|}{|H|}$
If $[G : H] = 1$, then $G = H$, so every character of H is a character of G.
If $[G : H] > 1$, there exists $g \in G$ such that $g \notin H$. Define n as the smallest integer such that $g^n \in H$. This n always exists, because since G is finite abelian, there exists $t \in \mathbb{Z}$ such that $g^t = e \in H$, where e is the identity element in G. Now let $\chi \in \hat{H}$, and $\chi(g^n) = k$. Since $\mathbb{C}$ is a divisible group under multiplication, there exists $m \in \mathbb{C}$ such that $m^n = k$. Now consider the subgroup $H_1 =< H, g >$, the subgroup of G generated by H and g. Then every element of H_1 is of the form: $h_1 = hg^a$ for some $h \in H, a \leq n$ integer. Define $\chi_1(h_1) = \chi(h)g^a$. χ_1 is a character of H_1, which extends the character χ of H. Since $g \notin H, |H_1| > |H|$, it follows that $[G : H_1] < [G : H]$. Thus, by the induction hypothesis, χ_1 can be extended to a character of G, and since χ_1 extends χ, χ can be extended to a character of G.

$\boxdot$

Proposition 3: $\hat{G}$ is a finite abelian group, and $|\hat{G}| = |G|$.

Proof: To see that $\hat{G}$ is abelian, if χ_1, χ_2 are two characters of G, by the fact that the complex numbers are abelian we have that

$$(\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g) = \chi_2(g)\chi_1(g) = (\chi_2\chi_1)(g) \text{ for all } g \in G.$$

As for$|G| = |\hat{G}|$, we go by induction on $|G|$.
If $|G| = 1$, there is only one character of G, since if χ a character of G, $g \in G$ the only element of G, then

$$\chi(g) = \chi(g \cdot g) = \chi(g)^2 \Rightarrow \chi(g) = 1.$$

Therefore, there is only one character of the trivial group.
Now, let $H < G$ be a strict subgroup of G. Consider the map $\alpha : \hat{G} \to \hat{H}$ defined by the restriction of characters of G to the subgroup H. By propostion 2, since every character of H can be extended to a character of G, the map α is surjective. Also,

$$\begin{aligned} ker(\alpha) &= \{\text{characters } \chi \text{ of G such that } \chi(h) = 1 \text{ for all } h \in H\} \\ &= \{\text{characters of the group } G/H\} \\ &= \widehat{G/H}. \end{aligned}$$

By the first isomorphism theorem, $\frac{\hat{G}}{\widehat{G/H}} \simeq \hat{H}$, and therefore $|\hat{G}| = |\widehat{G/H}||\hat{H}|$.

Since $|G/H|, |H| < |G|$, by induction hypothesis, $|G/H| = |\widehat{G/H}|$, and $|H| = |\hat{H}|$. Also, by LaGrange's Theorem, $|G/H| = \frac{|G|}{|H|}$. Putting all this together yields

$$\begin{aligned} |\hat{G}| &= |\widehat{G/H}||\hat{H}| \\ &= |G/H||H| \\ &= \frac{|G|}{|H|}|H| \\ &= |G|. \end{aligned}$$

$\boxdot$

Proposition 4: Let $g \in G$. Then the function $X_g : \hat{G} \to \mathbb{C}^*$ defined by $X_g(\chi) \mapsto \chi(g)$ defines a character of $\hat{G}$.

Proof: Let $g \in G$, and $\chi_1, \chi_2 \in \hat{G}$. Then

$$X_g(\chi_1\chi_2) = (\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g) = X_g(\chi_1)X_g(\chi_2).$$

Therefore, X is a homomorphism, and thus a character of $\hat{G}$.

$\boxdot$

Proposition 5: The map $\varepsilon : G \to \hat{\hat{G}}$ defined by $\varepsilon(g) = X_g$, where $g \in G$ is an isomorphism.

Proof: Since $|G| = |\hat{G}| = |\hat{\hat{G}}|$, the map ε is injective if and only if it is surjective. To show that ε is injective, we need to show that if $g \in G, g \neq 1$, then there exists a character χ of G such that $\chi(g) \neq 1$, or else the kernel of ε would not be trivial. Consider the subgroup $H =< g >, |H| = n$, subgroup generated by g. Now consider any character χ of H. Since $g^n = 1$, if $\chi(g) = \omega$, then

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

Therefore, ω is an n^{th} root of unity. By proposition 3, we know that $|H| = |\hat{H}|$, which means there must exist a non-trivial n^{th} root of unity ω such that $\chi(g) = \omega$ for some character χ of H, since otherwise there would only be one character of H. By proposition 1, the character χ of H with the property that $\chi(g) = \omega$ can be extended to a character of G, and therefore if $g \neq 1$, there exists a character χ of G such that $\chi(g) \neq 1$. Therefore, the map ε is injective, and thus surjective. To see that ε is a homomorphism, if $g_1, g_2 \in G$, then for all $\chi \in \hat{G}$, we have

$$\varepsilon(g_1g_2) = X_{g_1g_2} = \chi(g_1g_2) = \chi(g_1)\chi(g_2) = X_{g_1}X_{g_2} = \varepsilon(g_1)\varepsilon(g_2).$$

Therefore, ε is an isomorphism.

$\boxdot$

Proposition 6: Let $|G| = n$, χ be a character of G, then

i)

$$\sum_{g\in G} \chi(g) = \begin{cases} n & : \chi = 1 \\ 0 & : \chi \neq 1, \end{cases}$$

and

ii)

$$\sum_{\chi\in \hat{G}} \chi(g) = \begin{cases} n & : g = 1 \\ 0 & : g \neq 1. \end{cases}$$

Proof:
i) If $\chi = 1$, then

$$\sum_{g\in G} \chi(g) = 1 + 1 + \cdots + 1 \text{ (n times)} = n.$$

If $\chi \neq 1$, let $h \in G$ be such that $\chi(h) \neq 1$. Then

$$\chi(h) \sum_{g\in G} \chi(g) = \sum_{g\in G} \chi(h)\chi(g) = \sum_{g\in G} \chi(hg).$$

Since G is a group, if $h \cdot g = h \cdot k$, then $g = k$ by cancellation.
Therefore, the map: $\beta : G \longrightarrow G$ defined by $\beta(g) = h \cdot g$ is injective, and thus surjective (since $|G| < \infty$).
This shows that $\sum_{g\in G} \chi(hg) = \sum_{k\in G} \chi(k)$, where $k = xy$.

Therefore, we have

$$(\chi(h)-1)\cdot\sum_{g\in g}\chi(g)=\chi(h)\sum_{g\in g}\chi(g)-\sum_{g\in g}\chi(g)=\sum_{g\in g}\chi(g)-\sum_{g\in g}\chi(g)=0.$$

This implies that either $(\chi(h)-1)=0$, or $\sum_{g\in g}\chi(g)=0$, but since $\chi(h)\neq 1$ by assumption, $\sum_{g\in g}\chi(g)=0$.

Δ

ii) Since the function $X:\hat{G}\longrightarrow\mathbb{C}^*$ given by $X(\chi)=\chi(g)$ is a character of $\hat{G}$, the second part of the proposition follows immediately from the first part, when applied to $\hat{G}$.

$\boxdot$

We will now focus our discussion to characters with certain properties which will later be used to define L-functions. These characters are generally called "Dirichlet Characters of modulus m", where $m\in\mathbb{Z}$.

Definition 4: Consider $(\mathbb{Z}/m\mathbb{Z})^*$, the multiplicative group of invertible elements of $\mathbb{Z}/m\mathbb{Z}$. Then, if $\chi:(\mathbb{Z}/m\mathbb{Z})^*\longrightarrow\mathbb{C}^*$ is such that $\chi(ab)=\chi(a)\chi(b)$ for all $a,b\in(\mathbb{Z}/m\mathbb{Z})^*$, then χ is called a character modulo m.

Note that we can extend χ to a character of $\mathbb{Z}$ in the following way.

Definition 5: A function $\chi:\mathbb{Z}\longrightarrow\mathbb{C}^*$ is called a Dirichlet character of modulus m if it satifies the following three properties:
i) $\chi(ab)=\chi(a)\chi(b)$ for all $a,b\in\mathbb{Z}$.
ii) If $gcd(a,m)>1$, then $\chi(a)=0$.
iii) $\chi(a+m)=\chi(a)$ for all $a\in\mathbb{Z}$.

Example: $m=5$.
Note that the values that any Dirichlet character of modulus 5 takes for 1, 2, 3, and 4 completely determine the character for all $a\in\mathbb{Z}$ by condition *iii*). Since $\chi(1)=\chi(1^2)=\chi(1)\chi(1)$, it follows that $\chi(1)=1$. Also, $\chi(4^2)=\chi(16)=\chi(1)=1$, which implies that $\chi(4)=\pm 1$. $\chi(2^2)=\chi(4)=\chi(9)=\chi(3^2)$, and $\chi(3)\chi(2)=\chi(6)=\chi(1)=1$. Therefore:
If $\chi(4)=1$, then either $\chi(3)=\chi(2)=1$, or $\chi(3)=\chi(2)=-1$
If $\chi(4)=-1$, then either $\chi(2)=i,\chi(3)=-i$, or $\chi(2)=-i,\chi(3)=i$.
Therefore, the only four Dirichlet characters of modulus 5 are

mod 5	1	2	3	4
1	1	1	1	1
χ_1	1	-1	-1	1
χ_2	1	i	$-i$	-1
χ_3	1	$-i$	i	-1.

Note that χ_1 is simply $\left(\frac{a}{5}\right)$, where $\left(\frac{a}{5}\right)$ is the Legendre symbol which is defined in the following way.

Definition 6: The Legendre symbol $\left(\frac{a}{p}\right)$, $a \in \mathbb{Z}$, p prime is a function defined as

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if there exists } b \in \mathbb{Z}/p\mathbb{Z} \text{ such that } a = b^2(\text{mod}p) \\ 0 & \text{if } p \text{ divides } a \\ -1 & \text{otherwise.} \end{cases}$$

Proposition 7: $a^{(p-1)/2} \equiv 1(modp)$ if and only if there exists $b \in \mathbb{Z}/p\mathbb{Z}$ such that $a \equiv b^2(modp)$, or equivalently, $\left(\frac{a}{p}\right) \equiv a^{(p-1)/2}(modp)$.

Proof: Assume $a^{(p-1)/2} \equiv 1(modp)$. Let $g \in (\mathbb{Z}/p\mathbb{Z})^*$ be a primitive root (ie, $p-1$ is the smallest power of g such that $g^{(p-1)} \equiv 1(modp)$). Since g a primitive root, there exists $x < p-1$ such that

$$g^x \equiv a(modp).$$

This implies that

$$g^{(x(p-1)/2)} \equiv a^{(p-1)/2} \equiv 1(modp).$$

Therefore, 2 divides x as $x(p-1)/2$ must be a multiple of $p-1$, or else $g^{(x(p-1)/2)} \not\equiv 1(modp)$, by the fact that g is a primitive root. Let $x = 2 \cdot y$. Then

$$g^x = (g^y)^2 = b^2, \text{ where } b = g^y.$$

Therefore

$$b^2 = g^x \equiv a(modp).$$

Δ

Conversly, assume that there exists b such that $b^2 \equiv a(modp)$. Then

$$a^{(p-1)/2} = (b^2)^{(p-1)/2} = b^{p-1} \equiv 1(modp)$$

by Fermat's Little Theorem.

$\square$

Proposition 8: The function $\left(\frac{a}{p}\right)$ is a Dirichlet character of modulus p.

Proof: By the definition of the Legendre symbol, if $gcd(a,p) > 1$, then $\left(\frac{a}{p}\right) = 0$. Also, we have that

$$\left(\frac{a+p}{p}\right) = \left(\frac{a}{p}\right) \text{ for all } a \in \mathbb{Z},$$

since $a + p = a(modp)$.
To show multiplicativity of the Legendre symbol, we note that

$$\left(\frac{ab}{p}\right) = (ab)^{(p-1)/2} = a^{(p-1)/2} \cdot b^{(p-1)/2} = \left(\frac{a}{p}\right) \cdot \left(\frac{b}{p}\right).$$

Therefore, The Legendre symbol $\left(\frac{a}{p}\right)$ is a Dirichlet character of modulus m.

$\square$

Going back now the the example of the dirichlet characters modulo 5, we can see that

$$\left(\tfrac{1}{5}\right) = 1^{(5-1)/2} = 1^2 \equiv 1(mod5)$$
$$\left(\tfrac{2}{5}\right) = 2^2 = 4 \equiv -1(mod5)$$
$$\left(\tfrac{3}{5}\right) = 3^2 = 9 \equiv -1(mod5), \text{ and}$$
$$\left(\tfrac{4}{5}\right) = 4^2 = 16 \equiv 1(mod5) \quad .$$

Therefore, $\chi_1(a)$ is precisely the Legendre symbol $\left(\frac{a}{5}\right)$.

Now that we have familiarized ourselves with the definition of a character, and some of its properties, we will move on to defining a special kind of series known as Dirichlet

series. This will prove to be very useful later, as both the Riemann zeta function and the L-function are both Dirichlet series.

Definition 7: Let $(\lambda_n)_{n=1}^{\infty}$ be a sequence of real numbers such that
$$0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq \cdots.$$
Then, a Dirichlet series with exponents $(\lambda_n)_{n=1}^{\infty}$ is defined as
$$f(z) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n z}, \text{ where } a_n \in \mathbb{C}, z \in \mathbb{C}.$$

Proposition 9: If the Dirichlet series $f(z) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n}$ converges for some $z = z_0$, then for any domain in the first quadrant with $\Re(z) \geq \Re(z_0)$, $f(z)$ converges uniformly, where $\Re(z)$ denotes the real part of the complex number z.

In order to prove the above proposition, we need to make use of the following two lemmas.

Lemma 1: Let (a_n), (b_n) be two sequences. Define
$$A_{m,t} = \sum_{n=m}^{t} a_n, \text{ and } S_{m,t} = \sum_{n=m}^{t} a_n b_n.$$
Then
$$S_{m,t} = \sum_{n=m}^{t-1} A_{m,n}(b_n - b_{n+1}) + A_{m,t} b_t.$$
This is known as Abel's Lemma.

Proof: first note that $a_n = A_{m,n} + A_{m,n-1}$. Then

$$\begin{aligned} S_{m,t} &= \sum_{n=m}^{t} (A_{m,n} - A_{m,n-1}) b_n \\ &= (A_{m,m}(b_m - b_{m+1} + A_{m,m+1}(b_{m+1} - b_{m+2}) + \cdots + A_{m,t-1}(b_{t-1} - b_t) + A_{m,t} b_t \\ &= \sum_{n=m}^{t-1} A_{m,n}(b_n - b_{n+1}) + A_{m,t} b_t. \end{aligned}$$

⊡

Lemma 2: Let $a, b \in \mathbb{R}$, $0 < a < b$, $z = x + iy$, $x > 0$. Then
$$|e^{-az} - e^{-bz}| \leq |\tfrac{z}{x}|(e^{-ax} - e^{-bx}).$$

Proof: We start by noticing that
$$z \int_a^b e^{-tz} dt = (-e^{-tz})_a^b = e^{-az} - e^{-bz}.$$
Therefore
$$|e^{-az} - e^{-bz}| = |z \int_a^b e^{-tz} dt| \leq |z| \int_a^b e^{-tx} dt = |\tfrac{z}{x}|(-e^{-tz})_a^b = |\tfrac{z}{x}|(e^{-ax} - e^{-bx}).$$

⊡

We now return to the proof of Proposition 9.

Proof: Without loss of generality, we can assume that $z_0 = 0$ (by a translation argument). Therefore, if $\sum_{n=1}^{\infty} a_n e^{-\lambda_n z_0}$ converges, then so does $\sum_{n=1}^{\infty} a_n$ (as $z_0 = 0$). Therefore, we are reduced to proving that any domain with $|z|/\Re(z) \leq (\pi/2)$, and $\Re(z) \geq 0$, $f(z)$ converges uniformly. Let

$$A_{m,t} = \sum_{n=1}^{\infty} a_n, \text{ and } S_{m,t} = \sum_{n=1}^{\infty} a_n e^{-\lambda_n z}, \text{ where } \Re(z) > 0.$$

Then for all $\varepsilon > 0$, since $\sum_{n=1}^{\infty} a_n$ converges, there exists $N \in \mathbb{Z}^+$ such that for all $m, t \geq N$,

$$|A_{m,t}| < \varepsilon.$$

By Abel's Lemma, we have

$$S_{m,t} = \sum_{n=m}^{t-1} A_{m,n}(e^{-\lambda_n z} - e^{-\lambda_{n+1} z}) + A_{m,t} e^{-\lambda_t z}.$$

By Lemma 2, we have

$$|e^{-\lambda_n z} - e^{-\lambda_{n+1} z}| \leq |\tfrac{z}{x}|(e^{-\lambda_n x} - e^{-\lambda_{n+1} x}).$$

Therefore,

$$\begin{aligned} |S_{m,t}| &\leq \sum_{n=m}^{t-1} |A_{m,n}||e^{-\lambda_n z} - e^{-\lambda_{n+1} z}| + |A_{m,t}||e^{-\lambda_t z}| \\ &\leq \sum_{n=m}^{t-1} \varepsilon |\frac{z}{x}|(e^{-\lambda_n x} - e^{-\lambda_{n+1} x}) + \varepsilon e^{-\lambda_t x} \\ &\leq \varepsilon(1 + (\pi/2)(e^{-\lambda_m x} - e^{-\lambda_t x}) \\ &\leq \varepsilon(1 + \pi/2). \end{aligned}$$

Therefore, $f(z)$ converges uniformly.

$\boxdot$

Definition 8: The derivative of a complex valued function at a point z_0 is defined as

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}.$$

If the limit approaches the same value for every sequence of complex numbers approaching z_0, $f(z)$ is said to be complex-differentiable. A complex-valued function $f(z)$ is said to be holomorphic in a domain D of the complex plane, if $f(z)$ is complex-differentiable for all $z_0 \in D$.[2]

Definition 9: If a Dirichlet series $f(z)$ converges for $\Re(z) > x_0$, we call x_0 the abscissa of convergence, and the open set of $\mathbb{C}$ defined by $\Re(z) > x_0$ the half plane of convergence.

Corollary: The Dirichlet series $f(z) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n z}$ is holomorphic in the half plane of convergence.

To see this, we use the result from Complex Analysis that states:
If there is a sequence of holomorphic functions (f_n) on an open set $\mathbb{C}$, such that (f_n) converges uniformly to f, then f is holomorphic on U.[2]
By proposition 9, we saw that $f(z) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n z}$ converges uniformly in the half plane of convergence, and thus by the above result, $f(z)$ is holomorphic.

We now turn our attention to a specific type of Dirichlet series known as the ordinary Dirichlet Series.

Definition 10: The ordinary Dirichlet series is a Dirichlet series with $(\lambda_n)_{n=1}^{\infty} = (\log(n))_{n=1}^{\infty}$ ie

$$f(s) = \sum_{n=1}^{\infty} a_n n^{-s}.$$

Proposition 10: If there exists $M \in \mathbb{Z}^+$ such that $|a_n| < M$ for all n, where $f(s) = \sum_{n=1}^{\infty} a_n n^{-s}$, then $f(s)$ is absolutely convergent for $\Re(s) > 1$.

Proof: We have that

$$\sum_{n=1}^{\infty} |a_n| n^{-s} \leq \sum_{n=1}^{\infty} M n^{-s} = M \sum_{n=1}^{\infty} n^{-s}.$$

Since $\sum_{n=1}^{\infty} n^{-s}$ converges for all $s > 1$, so does $f(s)$.

$\boxdot$

Proposition 11: Let $A_{m,t} = \sum_{n=m}^{t} a_n$, where $f(s) = \sum_{n=1}^{\infty} a_n n^{-s}$. If there exists $M \in \mathbb{Z}^+$ such that $|A_{m,t}| \leq M$ for all m,t, then $f(s)$ converges for $\Re(s) > 0$.

Proof: Let $S_{m,t} = \sum_{n=m}^{t} a_n n^{-s}$. Then, by Abel's Lemma, we get

$$\begin{aligned} |S_{m,t}| &= |\sum_{n=m}^{t-1} A_{m,n}(n^{-s} - (n+1)^{-s}) + A_{m,t} t^{-s}| \\ &\leq M(\sum_{n=m}^{t-1} |n^{-s} - (n+1)^{-s}| + |t^{-s}|) \\ &\leq M(m^{-s}). \end{aligned}$$

Therefore, $f(s)$ converges for $\Re(s) > 0$, since m^{-s} goes to zero as m goes to infinity for all $s > 0$.

$\boxdot$

With the properties of Dirichlet series that we just discussed, we are now in good shape to define the Riemann zeta function, one of the most important functions in Number Theory, and L-functions, which play a critical role in Dirichlet's proof.

Definition 11: The Riemann zeta function is an ordinary Dirichlet series with $a_n = 1$ for all n, ie

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

Definition 12: Let $\chi(n)$ be a Dirichlet character of modulus m, for some $m \geq 1$, then the L-function with respect to χ is an ordinary Dirichlet series with $a_n = \chi(n)$ for all n ie

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

Proposition 12: If $g(x)$ is a bounded, strictly multiplicative function ($g(xy) = g(x)g(y)$ for all pairs (x, y)), then

$$\sum_{n=1}^{\infty} g(s) n^{-s} = \prod_{p \in P} (1 - g(p)p^{-s})^{-1},$$

where P is the set of all primes. This is known as Euler product factorization.

Proof: First, we see that

$$\begin{aligned} \prod_{p<x} (1 - g(p)p^{-s})^{-1} &= \prod_{p<x} \sum_{n=0}^{\infty} g(p)p^{-s} \\ &= \lim_{m\to\infty} \prod_{p<x} (1 + g(p)p^{-s} + \cdots + g(p^m)p^{-ms}) \\ &= \lim_{m\to\infty} \sum_{n \in S_x^m} g(s) n^{-s}, \end{aligned}$$

where $S_x^m = \{n \in \mathbb{Z}^+$such that each prime in the factorization of n is $\leq x$, and occurs with multiplicity $\leq m\}$.
Therefore, we have

$$\prod_{p \in P} (1 - g(p)p^{-s})^{-1} = \lim_{x\to\infty} \prod_{p<x} (1 - g(p)p^{-s})^{-1} = \lim_{x,m\to\infty} \sum_{n \in S_x^m} g(s) n^{-s} = \sum_{n=1}^{\infty} g(n) n^{-s}.$$

$\boxdot$

In particular, we get that

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p \in P} (1 - p^{-s})^{-1},$$

and

$$L(s,\chi) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p \in P} (1 - \chi(p)p^{-s})^{-1}.$$

Definition 13: A deleted-neighbourhood of $z_0 \in \mathbb{C}$ is defined as

$$D = \{z : 0 < |z - z_0| < r\},$$

ie, it is an open disc of radius r centered at z_0, where z_0 is not in the deleted-neighbourhood. Let $f(z)$ be afunction defined in a deleted neighbourhood of z_0. Then, if the function $(1/f)(z)$ is holomorphic in the full neighbourhood of z_0, where $(1/f)(z_0) := 0$, then z_0 is called a pole of $f(z)$ [2].

Proposition 13: The zeta function has a pole at $s = 1$.

Proof: We will prove the proposition by showing that

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

where $H(s)$ is holomorphic in $\Re(s) > 0$. We start by noticing that

$$\tfrac{1}{s-1} = \textstyle\int_1^{\infty} t^{-s} dt.$$

Therefore, we can rewrite the zeta function as

$$\begin{aligned}\zeta(s) &= \sum_{n=1}^{\infty} n^{-s} + \frac{1}{s-1} - \int_1^{\infty} t^{-s}dt \\ &= \frac{1}{s-1} + \sum_{n=1}^{\infty} n^{-s} - \sum_{n=1}^{\infty} \int_n^{n+1} t^{-s}dt \\ &= \frac{1}{s-1} + \sum_{n=1}^{\infty} \int_n^{n+1} (n^{-s} - t^{-s})dt.\end{aligned}$$

Let $H_n(s) = \int_n^{n+1}(n^{-s} - t^{-s})dt$, and $H = \sum_{n=1}^{\infty} H_n(s)$. Now since $H_n(s)$ is holomorphic for all $\Re(s) > 0$, if $\sum_{n=1}^{\infty} H_n(s)$ converges, H will have the desired properties. Now we have that

$$\begin{aligned}|H_n(s)| &= |\int_n^{n+1} (n^{-s} - t^{-s})dt| \\ &\leq \int_n^{n+1} |n^{-s} - t^{-s}|dt \\ &\leq \sup_{n \leq t \leq n+1} |n^{-s} - t^{-s}| \int_n^{n+1} dt \\ &= n^{-s} - (n+1)^{-s}.\end{aligned}$$

By the Mean Value Theorem, we have
$\frac{(n^{-s}-(n+1)^{-s})-(n^{-s}-n^{-s})}{(n+1)-n} = (d/dt)(n^{-s} - t^{-s})(c)$
for some $c \in (n, n+1)$. Therefore,
$|H_n(s)| = |(n^{-s} - (n+1)^{-s}| = |s/c^{s+1}| \leq| s|/|n^{s+1}| = |s|/n^{Re(s)+1}$.
This implies that
$$\sum_{n=1}^{\infty} H_n(s) \leq \sum_{n=1}^{\infty} |s|/n^{Re(s)+1}.$$
Therefore, H will converge for all $\Re(s) > 0$.

$\boxdot$

Proposition 14: $\lim_{s\to\infty} \left(\sum_{p\colon \text{prime}} \frac{1}{p^s} \right) \Big/ \log\left(\frac{1}{s-1}\right) = 1$, i.e. $\sum_p \frac{1}{p^s} \sim \log\left(\frac{1}{s-1}\right)$.

Proof: From the proof of Proposition 13, we know that

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

Therefore,

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

Now,

$$\log \zeta(s) = \log \prod_{p\colon \text{prime}} (1 - p^{-s})^{-1} = \sum_{p\colon \text{prime}} -\log(1 - p^{-s}).$$

By Taylor expanding $-\log(1-p^{-s})$, we have that

$$\begin{aligned}
\sum_p -\log(1-p^{-s}) &= \sum_p \left(p^{-s} + \frac{p^{-2s}}{2} + \frac{p^{-3s}}{3} + \cdots\right) \\
&= \sum_p p^{-s} + \sum_p \left(\frac{p^{-2s}}{2} + \frac{p^{-3s}}{3} + \cdots\right) \\
&\leq \sum_p p^{-s} + \sum_p \left(p^{-2s} + p^{-3s} + \cdots\right) \\
&= \sum_p p^{-s} + \sum_p p^{-2s}(1 + p^{-s} + p^{-2s} + \cdots) \\
&= \sum_p p^{-s} + \sum_p p^{-2s}\left(\frac{1}{1-p^{-s}}\right).
\end{aligned}$$

Since $\frac{1}{1-p^{-s}} \geq 2$, we get that

$$\begin{aligned}
\sum_p p^{-s} + \sum_p p^{-2s}\left(\frac{1}{1-p^{-s}}\right) &\leq \sum_p p^{-s} + 2\sum_p p^{-2s} \\
&< \sum_p p^{-s} + 2\sum_{n=1}^{\infty} n^{-2} \\
&= \sum_p p^{-s} + \frac{\pi^2}{3}.
\end{aligned}$$

Therefore, $\sum_p p^{-s} + A = \log\zeta(s), which implies that \sum_p p^{-s} \sim \log\left(\frac{1}{s-1}\right)$, where $A = \sum_p \left(\frac{p^{-2s}}{2} + \frac{p^{-3s}}{3} + \cdots\right)$ is bounded.

$\boxdot$

Proposition 15: If $\chi = 1$ is the trivial character modulo m, then

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

Proof: We have by proposition 12 that

$$L(s,1) = \prod_{p\in P}(1-\chi(p)p^{-s})^{-1},$$

with

$$\chi(p) = \begin{cases} 0 & : p|m \\ 1 & : otherwise. \end{cases}$$

Therefore, we get that

$$\begin{aligned}
L(s,1) &= \prod_{p\not|m}(1-p^{-s})^{-1} \\
&= \prod_{p\not|m}(1-p^{-s})^{-1}\prod_{p|m}(1-p^{-s})^{-1}/\prod_{p|m}(1-p^{-s})^{-1} \\
&= \zeta(s)/\prod_{p|m}(1-p^{-s})^{-1} \\
&= \zeta(s)\prod_{p|m}(1-p^{-s}).
\end{aligned}$$

□

Corollary: $L(s,1)$ has a pole at s=1.
Since for all $m \in \mathbb{Z}$, $\prod_{p|m}(1-p^{-s})$ is finite, since $\zeta(s)$ has a pole at $s=1$, it follows immediately from Propostion 15 that $L(s,1)$ has a pole at $s=1$. Also, we note that since the zeta function can be analytically extended to a function defined for all $s \neq 1$, it again follows from proposition 15 that $L(s,1)$ can be extended in the same way.

Proposition 16 For all $\chi \neq 1$, $L(1,\chi) < \infty$.

Proof : To prove this, we will show that $L(s,\chi)$ converges for $\mathrm{Re}(s) > 0$ ($\chi \neq 1$). Since $L(s,\chi)$ is an ordinary Dirichlet series, if we can show that $|A_{r,t}| = \sum_{n=r}^{t}$ are bounded, by Proposition 11 $L(s,\chi)$ will converge for $\mathrm{Re}(s) > 0$. By Proposition 6, we know that $\sum_{n=r}^{r+m-1} \chi(n) = 0$ (where m is the modulus of the character). Therefore, the only case that we care about is $\sum_{n=r}^{t} \chi(n)$, where $t - r < m$. But this turns out to be trivial, as

$$|A_{n,m}| = \left|\sum_{n=r}^{t} \chi(n)\right| \leq \sum_{n=r}^{t} |\chi(n)| \leq \varphi(m).$$

□

We will now introduce some notation that we will use to prove that $L(1,\chi) \neq 0$ for all $\chi \neq 1$.

Fix an $m \geq 1$. Given p a prime, if $gcd(p,m) = 1$, then define $\bar{p} = p(modm)$. Denote the order of $\bar{p}$ by $\theta(p)$ = smallest t such that $\bar{p}^t = 1(modm)$, and let $\psi(p) = \phi(m)/\theta(p)$.

Definition 14: We define a new function $\zeta_m(s)$ as

$$\zeta_m(s) = \prod_{\chi} L(s,\chi),$$

where the product is over all the Dirichlet characters modulo m.

Propostion 17: $\zeta_m(s) = \prod_{p \not| m} (1 - p^{-s\theta(p)})^{-\psi(p)}$.

Proof: To prove this result, we will rely on the following two results. The first result states that if Ω is the set of $\theta(p)$ roots of unity, then the following identity holds:

$$\prod_{\omega \in \Omega} (1 - \omega x) = 1 - x^{\theta(p)}. \text{ [1]}$$

The second result we use is the statement that there are exactly $\psi(p)$ characters modulo m with the property that $\chi(\bar{p}) = \omega$ for every fixed $\omega \in \Omega$. [1].
Combining the above two results gives us that

$$\prod_{\chi} (1 - \chi(p)x) = (1 - x^{\theta(p)})^{\psi(p)}.$$

Plugging in $p^{-s} = x$ into the above formula gives

$$\begin{aligned}\zeta_m(s) &= \prod_{\chi} L(s,\chi) \\ &= \prod_{\chi}\prod_{p\not| n}(1-\chi(p)p^{-s})^{-1} \\ &= \prod_{p\not| n}\prod_{\chi}(1-\chi(p)p^{-s})^{-1} \\ &= \prod_{p\not| n}(1-p^{-s\theta(p)})^{-\psi(p)}.\end{aligned}$$

⊡

We have now made it to the most crucial part of Dirichlet's proof, which is the following theorem.

Theorem 1: Let χ be a Dirichlet character modulus m, $\chi \neq 1$. Then, the associated L-function $L(s,\chi)$ has the property that $L(1,\chi) \neq 0$.

Proof: by Proposition 16, we know that $L(s,\chi)$ converges for all $\Re(s) > 0$, where $\chi \neq 1$, and by the corollary to Proposition 15, we know that $L(s,1)$ can be anayltically extended to $\Re(s) > 0$. Therefore, since $\zeta_m(s) = \prod_{\chi} L(s,\chi)$, $\zeta_m(s)$ can also be analytically extended to $\Re(s) > 0$.
Assume that there exists $\chi \neq 1$ such that $L(1,\chi)$=0. This implies that $\zeta_m(s)$ is holomorphic when $\Re(s) = 1$, and therefore it is holomorphic for all $\Re(s) > 0$, and thus $\zeta_m(s)$ converges for all $\Re(s) > 0$ [1].
Now, by expanding $\zeta_m(s)$, we get that

$$\begin{aligned}\zeta_m(s) &= \prod_{p\not| n}(1-p^{-s\theta(p)})^{-\psi(p)} \\ &= \prod_{p\not| n}(1+p^{-\theta(p)s}+p^{-\theta(p)2s}+\cdots)^{\phi(m)/\theta(m)} \\ &> \prod_{p\not| n}(1+p^{-\phi(m)s}+p^{-\phi(m)2s}+\cdots) \\ &= \sum_{gcd(n,m)=1} n^{-\phi(m)s}.\end{aligned}$$

The last line diverges for $s \leq \phi(m)$, and therefore $\zeta_m(s)$ diverges for $s \leq \phi(m)$. This is a contradiction, and therefore, $L(1,\chi) \neq 0$ for all $\chi \neq 1$.

⊡

Definition 15: Let P be the set of all prime numbers, $K \subseteq P$ be a subset. Then, the Dirichlet Density is defined as

$$\lim_{s\to 1^+}(\sum_{p\in K} p^{-s})/(\log(1/s-1)).$$

Now we are about done, as by using this density, Dirichlet's Theorem becomes a direct result of the following three lemmas.

Definition 16 : Let χ be a character modulo m, then

$$f_\chi(s) = \sum_{p \not| m} \frac{\chi(p)}{p^s}.$$

Lemma 3 : In the case that $\chi = 1$, we have that $f_1(s) \sim \log\left(\frac{1}{s-1}\right)$.

Proof : From Proposition 14, we have that $\sum_p \frac{1}{p^s} \sim \log\left(\frac{1}{s-1}\right)$. Since $\chi = 1$, then

$$f_1(s) = \sum_p \frac{1}{p^s} - \sum_{p \mid m} \frac{1}{p^s}.$$

(since $\chi(p) = 0$ if and only if $p \mid m$). But $\sum_{p \mid m} \frac{1}{p^s}$ is finite, as there are finitely-many primes less than m. Therefore, $f_1(s) \sim \sum_p \frac{1}{p^s} \sim \log\left(\frac{1}{s-1}\right)$.

$\boxdot$

Lemma 4: As $s \to 1^+$, if $\chi \neq 1$, then $f_\chi(s)$ is bounded.

Proof: From the proof of proposition 14, we can see through a similar argument that

$$\log L(s,\chi) = \sum_{p:prime} \chi(p)p^{-s} + A = \sum_{p \not| m} \chi(p)p^{-s} + A = f_\chi(s) + A,$$

where A remains bounded as $s \to 1^+$. Also, from proposition 16, we know that $L(s,\chi)$ remains bounded as $s \to 1^+$, which shows that $\log L(s,\chi)$ remains bounded as $s \to 1^+$. Therefore, we have that

$f_\chi(s) = \log L(s,\chi) - A$, which implies that $f_\chi(s)$ is bounded as $s \to 1^+$.

$\boxdot$

Lemma 5: Let P_a denote the set of primes congruent to a modulo m, and consider

$$f_a(s) = \sum_{p \in P_a} p^{-s}.$$

Then

$$f_a(s) = (1/\varphi(m)) \sum_\chi \chi(a)^{-1} f_\chi(s).$$

Proof: We have that

$$\begin{aligned} \sum_\chi \chi(a)^{-1} f_\chi(s) &= \sum_{p \not| m} \sum_\chi \chi(a)^{-1}\chi(p)p^{-s} \\ &= \sum_{p \not| m} \sum_\chi \chi(a^{-1}p)p^{-s} \\ &= \varphi(m) \sum_{p \in P_a} p^{-s}, \end{aligned}$$

since, by proposition 6, $\sum_\chi \chi(a^{-1}p) = \varphi(m)$ if and only if $a^{-1}p \equiv 1 mod(m)$, and $= 0$ otherwise.

$\boxdot$

We can now reword Dirichlet's Theorem in the following way.

Theorem 2: The set P_a has Dirichlet Density of $1/\varphi(m)$, and is therefore infinite.

Proof: By lemma 3, $f_1(s) \sim \log\left(\frac{1}{s-1}\right)$, and by lemma 4, all the other f_χ's are bounded. Therefore, by lemma 5, $f_a(s) \sim (1/\varphi(m)) \log\left(\frac{1}{s-1}\right)$. This implies that the Dirichlet density of P_a is

$$f_a(s)/\log\left(\frac{1}{s-1}\right) \sim (1/\varphi(m)).$$

Therefore, since the set P_a has a non-zero density, and any finite set has Dirichlet density of zero, the set must be infinite, completing the proof of Dirichlet's Theorem.

$\boxdot$

References

[1] Serre, Jean-Pierre. A Course in Arithmetic. New York: Springer-Verlag, 1973. Print.
[2] Stein, Elias, and Rami Shakarchi. Complex Analysis. New Jersey: Princeton UP, 2003. Print.