Dirichlet's Theorem on Arithmetic Progressions

Anthony Várilly

Harvard University, Cambridge, MA 02138

1 Introduction

Dirichlet's theorem on arithmetic progressions is a gem of number theory. A great part of its beauty lies in the simplicity of its statement.

Theorem 1.1 (Dirichlet). *Let a, $m \in \mathbb{Z}$, with $(a, m) = 1$. Then there are infinitely many prime numbers in the sequence of integers a, $a + m$, $a + 2m, \ldots, a + km, \ldots$ for $k \in \mathbb{N}$.*

A sixth grader knows enough mathematics to understand this particular formulation of the theorem. However, many deep ideas of algebra and analysis are required to prove it.

In order to motivate some of the ideas we will introduce, we will sketch how to show there are infinitely many primes of the form $4k + 1$, the special case $a = 1, m = 4$ of Theorem 1.1. We shall follow Knapp's exposition in our sketch [2].

Define the (real valued) *Riemann zeta function* as

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \qquad s > 1. \tag{1}$$

Throughout this paper, p shall denote a prime number, unless otherwise indicated. It is possible to write the zeta function as the infinite product

$$\zeta(s) = \prod_{p} \frac{1}{1 - p^{-s}}. \tag{2}$$

To see why this is true, notice that for finite N,

$$\prod_{p \leq N} \frac{1}{1 - p^{-s}} = \sum_{n \in \mathbf{S}} \frac{1}{n^s}$$

where $\mathbf{S}$ is the set of natural numbers whose prime factors do not exceed N. Letting $N \to \infty$ we obtain the result. With this product formula for $\zeta(s)$, it is possible to show (and we will do so in the proof of Theorem 1.1) that

$$\log \zeta(s) = \sum_{p} \frac{1}{p^s} + g(s) \tag{3}$$

where $g(s)$ is bounded as $s \to 1$.

Define a function $\chi : \mathbb{Z} \to \{-1, 0, 1\}$ by

$$\chi(a) = \begin{cases} 0 & \text{if } a \text{ is even,} \\ 1 & \text{if } a \equiv 1 \bmod 4, \\ -1 & \text{if } a \equiv 3 \bmod 4. \end{cases}$$

This function will allow us to distinguish primes of the form $4k+1$ and $4k+3$ from one another. Notice that $\chi(mn) = \chi(m)\chi(n)$ for all integers m and n. Now let

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}. \tag{4}$$

Since χ is multiplicative for all integers (we say χ is *strictly multiplicative* in this case), one can write, just like in the case of the zeta function,

$$L(s, \chi) = \prod_{p} \frac{1}{1 - \chi(p)p^{-s}}. \tag{5}$$

This product formula can be used to show that

$$\log L(s, \chi) = \sum_{p} \frac{\chi(p)}{p^s} + g_1(s, \chi) \tag{6}$$

where $g_1(s, \chi)$ is a function that remains bounded as $s \to 1$.

Combining (3) and (6) we get

$$\log \zeta(s) + \log L(s, \chi) = 2 \sum_{p \equiv 1(4)} \frac{1}{p^s} + \left(\frac{1}{2^s} + g(s) + g_1(s, \chi) \right), \tag{7}$$

$$\log \zeta(s) - \log L(s, \chi) = 2 \sum_{p \equiv 3(4)} \frac{1}{p^s} + \left(\frac{1}{2^s} + g(s) - g_1(s, \chi) \right). \tag{8}$$

From the Taylor expansion of $\arctan x$, $L(1, \chi) = \pi/4 > 0$. But $\zeta(s)$ diverges as $s \to 1$, so the left hand sides of (7) and (8) tend to infinity as $s \to 1$. Since both $1/2^s + g(s) + g_1(s, \chi)$ and $1/2^s + g(s) - g_1(s, \chi)$ remain bounded, it follows that both sums $\sum_{p \equiv 1(4)} 1/p^s$ and $\sum_{p \equiv 3(4)} 1/p^s$ diverge as $s \to 1$. This proves there are infinitely many primes of the form $4k+1$. As a bonus, we obtained the existence of infinitely many primes of the form $4k+3$.

There were two crucial ideas that made this last proof possible. First, it was imperative that $L(1, \chi)$ was finite and non-zero, so that its logarithm remain bounded in (8). The other key idea was the use of the function χ to 'filter out' the primes of the form $4k+1$ from all other primes. To prove Dirichlet's theorem, we'll need functions like χ that will filter out primes of the form $a + km$. We thus direct our attention to such functions: group characters.

2 Group Characters

Let G be a finite abelian group. A group character is a homomorphism $\chi : G \to \mathbb{C}^*$. The characters of a group form themselves a group under pointwise multiplication. We call this group the *dual* of G and denote it $\widehat{G}$.

If G is a cyclic group of order n, then it is easy to describe $\widehat{G}$. Let g be a generator of G. Then $\chi(g) = w$ for some $w \in \mathbb{C}^*$. Since χ is a homomorphism,

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

Hence w is an n^{th} root of unity. Conversely, let w be an n^{th} root of unity. Then we can define a character χ of $\widehat{G}$ by setting $\chi(g) = w$. Notice that $\chi^{-1}(a) = \overline{\chi(a)}$ for all $a \in G$. We have a bijective correspondence between the group of n^{th} roots of unity μ_n and $\widehat{G}$. In fact, it is easy to see that this correspondence gives an isomorphism. Since $\mu_n \cong \widehat{G}$, it follows that $G \cong \widehat{G}$.

Now let G be any finite abelian group. The structure theorem for finite abelian groups tells us G can be written as a direct product of cyclic groups, $G \cong C_{n_1} \times \cdots \times C_{n_k}$. Let g_i be a generator of C_{n_i}. Every element of G can be written as a product of g_i's to the appropriate powers, so a character of G is completely determined by the images of the g_i's. These images must again be roots of unity.

Conversely, we can define a character χ_i of $\widehat{G}$ by sending g_i to an n_i^{th} root of unity w_{n_i} and all other generators g_j to the identity element. It is easy to see that in this case $G \cong \widehat{G}$ as well. In particular, a group and its dual have the same order.

2.1 Examples of Groups Characters

Example 2.1. The *trivial character* $\chi : G \to \mathbb{C}^*$ defined by $\chi(a) = 1$ for all $a \in G$.

Example 2.2. *Dirichlet Characters modulo m*: Let $G = (\mathbb{Z}/m\mathbb{Z})^*$. Then G is a finite abelian group with $\phi(m)$ elements (here ϕ is the Euler totient function). The Dirichlet characters can be extended to all of $\mathbb{Z}$ by setting $\chi(a) = 0$ if $(a, m) > 1$ and letting $\chi(a+m) = \chi(a)$ for all integers a. These extensions are not themselves group characters (a character can't take the value 0), but they are multiplicative functions on $\mathbb{Z}$. Through an abuse of language, we will often times refer to these extensions as Dirichlet characters modulo m.

Example 2.3. The principal Dirichlet character modulo m is the extension to $\mathbb{Z}$ (as a multiplicative function) of the trivial character of $(\mathbb{Z}/m\mathbb{Z})^*$:

$$\chi_0(a) = \begin{cases} 1 & \text{if } (a, m) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

We shall often write 1 for the principal character instead of χ_0

Example 2.4. Let $m = 4$ in Example 2.2. Then $(\mathbb{Z}/4\mathbb{Z})^* \cong \mathbb{Z}/2\mathbb{Z}$, so the dual of $(\mathbb{Z}/4\mathbb{Z})^*$ has one non-trivial character; it is given by

$$\chi(a) = \begin{cases} 1 & \text{if } a \equiv 1 \bmod 4, \\ -1 & \text{if } a \equiv 3 \bmod 4, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2.5. Let $m = p$ in Example 2.2; here p is an odd prime number. The dual of $(\mathbb{Z}/p\mathbb{Z})^*$ will be cyclic of order $p-1$. Hence there will be a character χ of order 2, that is $\chi^2 = \chi_0$. If a is a quadratic residue modulo p, then $\chi(a)$ is forced to be 1. If a is a quadratic non-residue, then $\chi(a)$ is forced to be -1. Thus we can identify χ with the familiar Legendre symbol, $\chi(a) = \left(\frac{a}{p}\right)$.

Remark. The characters χ of $\widehat{G}$ are strictly multiplicative, that is, $\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in G$. This follows from the definition of group homomorphism.

2.2 Orthogonality Relations

As we said earlier, group characters are essential to the proof of Dirichlet's theorem because they let us 'filter out' the primes of the form $a + km$. But how is this the case? It turns out the characters of a group satisfy certain orthogonality relations. These relations hold the key to the process of 'filtering primes'.

Theorem 2.1. *Let $\chi \in \widehat{G}$. Then*

$$\frac{1}{|G|} \sum_{a \in G} \chi(a) = \begin{cases} 1 & \text{if } \chi = 1, \\ 0 & \text{if } \chi \neq 1. \end{cases} \tag{9}$$

Proof. If $\chi = 1$, the sum adds up to the number of elements in G. Otherwise, choose $b \in G$ such that $\chi(b) \neq 1$. Then

$$\chi(b) \cdot \frac{1}{|G|} \sum_{a \in G} \chi(a) = \frac{1}{|G|} \sum_{a \in G} \chi(a)\chi(b) = \frac{1}{|G|} \sum_{a \in G} \chi(ab) = \frac{1}{|G|} \sum_{a \in G} \chi(a).$$

The last equality follows from the fact that as a ranges through the elements of G, so does ab. Hence we have

$$(\chi(b) - 1) \cdot \frac{1}{|G|} \sum_{a \in G} \chi(a) = 0.$$

Since $\chi(b) \neq 1$, (9) follows. □

By applying Theorem 2.1 to the dual group $\widehat{G}$ and since $G \cong \widehat{\widehat{G}}$, we get the following result.

Corollary 2.2. *Let $a \in G$. Then*

$$\frac{1}{|\widehat{G}|}\sum_{\chi\in\widehat{G}}\chi(a) = \begin{cases} 1 & \text{if } a = 1, \\ 0 & \text{if } a \neq 1. \end{cases} \tag{10}$$

Consider now two characters $\chi, \psi \in \widehat{G}$. Since $\widehat{G}$ is a group, $\chi\psi^{-1} \in \widehat{G}$, and $\sum_a \chi\psi^{-1}(a) = \sum_a \chi(a)\psi^{-1}(a) = \sum_a \chi(a)\overline{\psi(a)}$. By Theorem 2.1, we have

$$\frac{1}{|G|}\sum_{a\in G}\chi(a)\overline{\psi(a)} = \begin{cases} 1 & \text{if } \chi = \psi, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

Similarly, $\sum_\chi \chi(ab^{-1}) = \sum_\chi \chi(a)\chi(b^{-1}) = \sum_\chi \chi(a)\overline{\chi(b)}$, so that by Corollary 2.2, we get

$$\frac{1}{|\widehat{G}|}\sum_{\chi\in\widehat{G}}\chi(a)\overline{\chi(b)} = \begin{cases} 1 & \text{if } a = b, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Equations (11) and (12) are refered to as the orthogonality relations for group characters. A special case of these relations, which is of interest to us, occurs when $G = (\mathbb{Z}/m\mathbb{Z})^*$.

Corollary 2.3 (Orthogonality relations for Dirichlet Characters). *Let χ and ψ be Dirichlet characters modulo m, and let a, b be integers. Then*

$$\frac{1}{\phi(m)}\sum_{a=0}^{m-1}\chi(a)\overline{\psi(a)} = \begin{cases} 1 & \text{if } \chi = \psi, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

$$\frac{1}{\phi(m)}\sum_{\chi}\chi(a)\overline{\chi(b)} = \begin{cases} 1 & \text{if } a \equiv b \mod m, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

These last two relations shall do us a great service when we try to 'filter out' primes of the form $a + km$ from the zeta function. This is all the character theory we will need. If the reader is interested in a more thorough treatment of it, we recommend Serre's book [4]. For a treatment closer to ours, Ireland and Rosen [1] would be a good book to look at.

We now turn our attention to series like (4). A careful study of them, together with our knowledge of group characters is enough to prove Theorem 1.1.

3 Dirichlet Series

A series

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

with a_n and s complex is called a *Dirichlet series*. We will be primarily concerned with series where a_n is a Dirichlet character modulo m. First, we must know something about a Dirichlet series' region of convergence. We follow Knapp's [2] treatment on Dirichlet series for the following theorems.

Theorem 3.1. *Let $\sum_{n=1}^{\infty} \frac{a_n}{n^s}$ be a Dirichlet series. If the series converges for a particular $s = s_0$, then it converges uniformly on the open half-plane $\operatorname{Re} s > \operatorname{Re} s_0$. Furthermore, the sum is analytic in this region.*

We will need Abel's summation formula to prove the theorem. Suppose $\{u_n\}$ and $\{v_n\}$ are sequences of complex numbers such that $\sum_{n=1}^{\infty} u_n v_n$ converges. Let $U_n = \sum_1^n u_i$; if $U_n v_n \to 0$ as $n \to \infty$, then

$$\sum_{n=1}^{\infty} u_n v_n = \sum_{n=1}^{\infty} U_n (v_n - v_{n+1})$$

Proof of Theorem 3.1. We have $\dfrac{a_n}{n^s} = \dfrac{a_n}{n^{s_0}} \dfrac{1}{n^{s-s_0}}$. Let $u_n = \dfrac{a_n}{n^{s_0}}$ and $v_n = \dfrac{1}{n^{s-s_0}}$. We know $\{U_n\}$ is convergent by hypothesis, and $v_n \to 0$ uniformly on the half-plane $\operatorname{Re} s > \operatorname{Re} s_0$. Thus $U_n v_n \to 0$ as $n \to \infty$ in this region. Say $U_n \to U$ as $n \to \infty$. Then

$$\begin{aligned} \left| \sum u_n v_n \right| = \left| \sum U_n (v_n - v_{n+1}) \right| &\le \sum |U_n| |v_n - v_{n+1}| \\ &\le U \sum |v_n - v_{n+1}|. \end{aligned}$$

If we can show that $\sum |v_n - v_{n+1}| = \sum \left| \dfrac{1}{n^{s-s_0}} - \dfrac{1}{(n+1)^{s-s_0}} \right|$ converges uniformly on the half-plane $\operatorname{Re} s > \operatorname{Re} s_0$, we will be done. For $n \le t \le n+1$, we have

$$\begin{aligned} \frac{n^{-(s-s_0)} - t^{-(s-s_0)}}{1} &\le \sup_{n \le t \le n+1} \left| \frac{d}{dt} \left(n^{-(s-s_0)} - t^{-(s-s_0)} \right) \right| \\ &= \sup_{n \le t \le n+1} \left| \frac{s - s_0}{t^{s-s_0+1}} \right| \le \frac{|s - s_0|}{n^{1+\operatorname{Re}(s-s_0)}}, \end{aligned} \tag{15}$$

and so $|v_n - v_{n+1}| \le \dfrac{|s - s_0|}{n^{1+\operatorname{Re}(s-s_0)}}$. Hence

$$\sum_n |v_n - v_{n+1}| \le |s - s_0| \sum_n \frac{1}{n^{1+\operatorname{Re}(s-s_0)}},$$

and this last expression converges uniformly when $\operatorname{Re}(s - s_0) > 0$. The analyticity of the sum follows from the analyticity of each term in the half-plane. □

Corollary 3.2. *If the Dirichlet series $\sum_{n=1}^{\infty} a_n/n^s$ converges* absolutely *at $s = s_0$, then it converges uniformly and absolutely in the half-plane $\operatorname{Re} s \ge \operatorname{Re} s_0$.*

Proof. With absolute convergence, we deduce $\sum \left| \dfrac{a_n}{n^s} \right| = \sum \left| \dfrac{a_n}{n^{s_0}} \right| \dfrac{1}{n^{s-s_0}} \le \sum \left| \dfrac{a_n}{n^{s_0}} \right|$ and since the sum on the right hand side of these relations converges, we get the result by a simple application of the Weierstrass M-test. □

3.1 Dirichlet L-series and Euler Products

Dirichlet series that have Dirichlet characters modulo m (extended to $\mathbb{Z}$) as their coefficients are called L-functions.

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}. \tag{16}$$

When we studied primes of the form $4k+1$, we came across an example of L-function. Notice though that a general L-function can have s and $\chi(n)$ take complex values.

Remark. $L(s,1)$ looks like a zeta function with complex s that is missing all integers n that are divisible by m, since $\chi(n) = 0$ for such n.

Lemma 3.3. *The zeta function $\zeta(s)$ is meromorphic in the half-plane $\operatorname{Re} s > 0$. Its only pole is $s = 1$ and it is simple.*

Proof. We have

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

Notice that $\int_n^{n+1} (n^{-s} - t^{-s})\, dt$ is an analytic function for $\operatorname{Re} s > 0$. To show the sum of such integrals (as n ranges from 1 to ∞) is analytic, all we need is convergence on compact sets for which $\operatorname{Re} s > 0$. Now,

$$\Big| \int_n^{n+1} n^{-s} - t^{-s}\, dt \Big| \leq \int_n^{n+1} |n^{-s} - t^{-s}|\, dt \leq \sup_{n \leq t \leq n+1} |\left(n^{-s} - t^{-s}\right)|,$$

and this last expression is at most $\dfrac{|s|}{n^{1+\operatorname{Re} s}}$ by (15). The series $\displaystyle\sum_n \frac{1}{n^{1+\operatorname{Re} s}}$ converges for $\operatorname{Re} s > 0$. Hence the desired series of integrals converges in this region as well. □

Our next goal is to obtain a product expansion for $L(s,\chi)$ like that of the zeta function. We use the crucial fact that Dirichlet characters are strictly multiplicative.

Lemma 3.4. *The Dirichlet series $\sum_n \frac{\chi(n)}{n^s}$ converges absolutely for $Re s > 1$. Furthermore,*

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \frac{1}{1 - \chi(p)p^{-s}}. \tag{17}$$

Proof. χ is a bounded function. This gives the desired absolute convergence for $\operatorname{Re} s > 1$. To see why the product expansion holds note that for $\operatorname{Re} s > 1$ and a fixed prime number q,

$$\prod_{\substack{p=2,\\ p\,\text{prime}}}^{q} \frac{1}{1-\chi(p)p^{-s}} = \prod_{\substack{p=2,\\ p\,\text{prime}}}^{q} \left(1+\chi(p)p^{-s}+\chi^2(p)p^{-2s}+\cdots\right) \tag{18}$$

$$= \prod_{\substack{p=2,\\ p\,\text{prime}}}^{q} \left(1+\chi(p)p^{-s}+\chi(p^2)p^{-2s}+\cdots\right) = \sum_{n\in\mathbf{S}} \frac{\chi(n)}{n^s} \tag{19}$$

where $\mathbf{S}$ is the set of natural numbers whose prime factors do not exceed q. This means the partial product (18) is equal to a convergent infinite sum. Now fix a natural number N. We have

$$\sum_{n=1}^{N} \frac{\chi(n)}{n^s} = \prod_{\substack{p=2,\\ p\,\text{prime}}}^{r} \frac{1}{1-\chi(p)p^{-s}} - \sum_{\substack{n\in\mathbf{S}\\ n>N}} \frac{\chi(n)}{n^s} \tag{20}$$

where r is the largest prime number less than or equal to N, and now $\mathbf{S}$ is the set of natural numbers whose prime factors do not exceed r. Letting $q \to \infty$ in (18) and $N \to \infty$ in (20) we see that the product expansion and the series converge or diverge together. Since we know the series converges for $\operatorname{Re} s > 1$, the product expansion must also converge in that region. Furthermore, letting $q \to \infty$ in (18), we obtain (17) □

With the above three lemmas in hand, we can extend our remark about $L(s,1)$. Applying Lemma 3.4 to the principal character, we have

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

This last equality follows from the fact that we have extended the zeta function to the region $\operatorname{Re} s > 0$ in Lemma 3.3. Since the product over $p|m$ is finite, it follows that $L(s,1)$ is meromorphic in the region $\operatorname{Re} s > 0$ and its only pole is simple at $s = 1$. Note, however, that the product expansion of $L(s,1)$ is only valid in the region $\operatorname{Re} s > 1$.

The product expression (17) is an example of an *Euler product* of first degree.

If χ is not the principal character, then we can go further and show that the series $L(s,\chi)$ is convergent and analytic in the region $\operatorname{Re} s > 0$.

Theorem 3.5. *Let χ be a Dirchlet character modulo m different from the principal character. Then the* series *$L(s,\chi)$ converges and is analytic in* $\operatorname{Re} s > 0$.

Proof. We extended Dirchlet characters to $\mathbb{Z}$ by setting $\chi(a) = 0$ when $(a,m) > 1$ and by letting $\chi(a+m) = \chi(a)$ for all integers a. Using the extended characters, it follows, by Theorem 2.1, that

$$\sum_{n=1}^{m} \chi(m+a) = 0 \tag{21}$$

for any a.

Let $s > 0$ for now. We use Abel's summation formula with $u_n = \chi(n)$ and $v_n = 1/n^s$. Equation (21) says $\{U_n\}$ is bounded; say $|U_n| \le U$. It is easy to see that $U_n v_n \to 0$ as $n \to \infty$. Hence

$$\Big| \sum_{n=M}^{\infty} \frac{\chi(n)}{n^s} \Big| = \Big| \sum_{n=M}^{\infty} u_n v_n \Big| = \Big| \sum_{n=M}^{\infty} U_n (v_n - v_{n+1}) \Big| \le U \sum_{n=M}^{\infty} |v_n - v_{n+1}| = \frac{U}{M^s}$$

for any finite M. The last equality follows because $|v_n - v_{n+1}| = (v_n - v_{n+1})$ for $s > 0$. As $M \to \infty$, the last expression tends to zero. Therefore the series $\sum_n \chi(n)/n^s$ is convergent for s real and positive. By Theorem 3.1, the series is convergent and analytic in the region $\operatorname{Re} s > 0$. □

As a consequence of Theorem 3.5, we see that when χ is not the principal character, $L(s, \chi)$ is well defined at $s = 1$. We will need to show that in fact it is not zero to prove Dirichlet's theorem.

Theorem 3.6. *For non-principal χ, $L(1, \chi) \neq 0$.*

We will postpone the proof of this theorem until we prove Dirichlet's theorem.

4 Dirichlet's theorem

We are now in a position to prove Theorem 1.1. For the first part of the proof we will loosely follow Knapp's [2] treatment.

Proof of Theorem 1.1. First, we will show that for a Dirichlet character modulo m,

$$\log L(s, \chi) = \sum_p \frac{\chi(p)}{p^s} + g(s, \chi) \tag{22}$$

for real $s > 1$, where $g(s, \chi)$ is a function that remains bounded as $s \to 1$. Even if s is real, $L(s, \chi)$ could still be complex valued, so if we want to take its logarithm, we better choose a branch. For a given p and $s \ge 1$, define the value of the logarithm of the p^{th} factor in the L-function's Euler product by

$$\log \frac{1}{1 - \chi(p) p^{-s}} = \frac{\chi(p)}{p^s} + \sum_{n=2}^{\infty} \frac{\chi(p^n)}{n p^{ns}}$$

and let $g(s, \chi, p) = \displaystyle\sum_{n=2}^{\infty} \frac{\chi(p^n)}{n p^{ns}}$. For this choice of branch, and for $|z| \le 1/2$, we have

$$\begin{aligned} \Big| \log \frac{1}{1-z} - z \Big| = \Big| \sum_{n=2}^{\infty} \frac{z^n}{n} \Big| &\le \sum_{n=2}^{\infty} \frac{|z|^n}{n} \\ &\le |z|^2 \sum_{n=0}^{\infty} \frac{1}{n+2} \Big(\frac{1}{2}\Big)^n \le |z|^2 \sum_{n=0}^{\infty} \Big(\frac{1}{2}\Big)^{n+1} = |z|^2 \end{aligned}$$

Now set $z = \dfrac{\chi(p)}{p^s}$. Since $\left|\dfrac{\chi(p)}{p^s}\right| \leq \dfrac{1}{2}$ we obtain

$$\left|g(s,\chi,p)\right| = \left|\log\frac{1}{1-\chi(p)p^{-s}} - \frac{\chi(p)}{p^s}\right| \leq \left|\frac{\chi(p)}{p^s}\right|^2 \leq \frac{1}{p^2} \qquad \text{for } s \geq 1.$$

Finally, set $g(s,\chi) = \sum_p g(s,\chi,p)$. Now $\sum_p |g(s,\chi,p)| \leq \sum_p \frac{1}{p^2} \leq \sum_n \frac{1}{n^2}$ which converges. Thus, $g(s,\chi)$ remains bounded as $s \to 1$.

By adding all the logarithms of the factors in the L-function's Euler product, we obtain a branch of $\log L(s,\chi)$,

$$\sum_p \log\frac{1}{1-\chi(p)p^{-1}} = \sum_p \frac{\chi(p)}{p^s} + g(s,\chi)$$

This establishes (22). Now we use group characters to 'filter out' the primes of the form $a + km$. Recall from our discussion of Dirichlet characters modulo m the following orthogonality relation

$$\frac{1}{\phi(m)}\sum_\chi \chi(a)\overline{\chi(b)} = \begin{cases} 1 & \text{if } a \equiv b \bmod m, \\ 0 & \text{otherwise.} \end{cases}$$

If we multiply (22) by $\overline{\chi(a)}$ and sum over all χ we get

$$\sum_\chi \overline{\chi(a)}\log L(s,\chi) = \sum_\chi \overline{\chi(a)}\sum_p \frac{\chi(p)}{p^s} + \sum_\chi \overline{\chi(a)}g(s,\chi)$$

Using the orthogonality relation, we rearrange this last equation to give

$$\phi(m)\sum_{p\equiv a(m)}\frac{1}{p^s} = \sum_\chi \overline{\chi(a)}\log L(s,\chi) - \sum_\chi \overline{\chi(a)}g(s,\chi). \tag{23}$$

We know $L(s,1)$ has a pole at $s = 1$; in fact $L(s,1) \to \infty$ as $s \to 1$ for real s. On the other hand, by Theorem 3.6, we know $\log L(s,\chi)$ is bounded at $s = 1$ for χ non-principal. Hence the sum $\sum_\chi \overline{\chi(a)}\log L(s,\chi)$ has *only one* unbounded term as $s \to 1$. This means the sum must itself be unbounded as $s \to 1$. This is why Theorem 3.6 is so important. Had two or more terms of the previous sum been unbounded, they could have cancelled, depending on the sign of $\overline{\chi(a)}$. The term $\sum_\chi \overline{\chi(a)}g(s,\chi)$ is bounded as $s \to 1$ because $g(s,\chi)$ is. Thus, overall, the right hand side of (23) is unbounded as s approaches 1. This means there must be infinitely many primes contributing to the sum on the left hand side. This concludes the proof of the theorem. □

4.1 The missing link

We have to prove Theorem 3.6. Let χ be a Dirichlet character modulo m. Define $\zeta_m(s)$ by

$$\zeta_m(s) = \prod_\chi L(s,\chi). \tag{24}$$

We know $L(s,1)$ has a simple pole at $s=1$ and all other $L(s,\chi)$ are analytic for $\operatorname{Re} s > 0$. Suppose there is some non-principal χ such that $L(1,\chi)=0$. The function $\zeta_m(s)$ would then be analytic on $\operatorname{Re} s > 0$. We will prove this is not the case. The theorem will follow.

Suppose p is a prime not dividing m. Let $f(p)$ be the order of the image $\overline{p}$ of p in $(\mathbb{Z}/m\mathbb{Z})^*$. Define $g(p) = \phi(m)/f(p)$, that is, the order of the quotient of $(\mathbb{Z}/m\mathbb{Z})^*$ by (p).

Lemma 4.1. *If $p \nmid m$ then*

$$\prod_{\chi} \left(1 - \frac{\chi(p)}{p^s}\right) = \left(1 - \frac{1}{p^{f(p)s}}\right)^{g(p)}. \tag{25}$$

Proof. Let μ_n denote the group of n^{th} roots of unity. Then

$$\begin{aligned}\prod_{w\in\mu_{f(p)}} (1-wx) = \prod_{w\in\mu_{f(p)}} w^{f(p)} \prod_{w\in\mu_{f(p)}} \left(\frac{1}{w} - x\right) &= e^{i\pi(n-1)} \prod_{w\in\mu_{f(p)}} (w-x) \\ &= (-1)^{n-1}(-1)^n \prod_{w\in\mu_{f(p)}} (x-w) = 1 - x^{f(p)}.\end{aligned}$$

For any $w \in \mu_{f(p)}$, there are $\phi(m)/f(p) = g(p)$ Dirichlet characters modulo m such that $\chi(\overline{p}) = w$. Letting $x = \frac{1}{p^s}$ we obtain the desired result. □

Lemma 4.2. *The function $\zeta_m(s)$ has a product expansion for $\operatorname{Re} s > 1$ given by*

$$\zeta_m(s) = \prod_{p\nmid m} \left(\frac{1}{1-p^{-f(p)s}}\right)^{g(p)}. \tag{26}$$

Proof. We have

$$\zeta_m(s) = \prod_{\chi} L(s,\chi) = \prod_{p\nmid m} \left(\prod_{\chi} \frac{1}{1-\chi(p)p^{-s}}\right) = \prod_{p\nmid m} \left(\frac{1}{1-p^{-f(p)s}}\right)^{g(p)},$$

where the last step follows from Lemma 4.1. □

Notice that $\dfrac{1}{1-p^{-f(p)s}}$ is given by a Dirichlet series with positive coefficients. Hence, Lemma 4.2 shows $\zeta_m(s)$ is itself a Dirichlet series with positive coefficients. This is the key fact we shall exploit to obtain our contradiction.

Lemma 4.3. *Let $\sum_n a_n/n^s$ be a Dirichlet series with coefficients $a_n \geq 0$. Suppose the series converges for $\operatorname{Re} s > s_0$ for some real s_0, and that it extends analytically to an analytic function in a neighborhood around the point s_0. Then there is an $\epsilon > 0$ for which the series converges for $\operatorname{Re} s > (s_0 - \epsilon)$.*

Proof. We will follow Serre's [3, p. 67] ideas in this proof, though our proof is not as general as his. We may assume without loss of generality that $s_0 = 0$. Just replace s with $s - s_0$. For convenience, denote the series above by $f(s)$. Because $f(s)$ is analytic in the region $\operatorname{Re} s > 0$ and in a neighborhood around 0, there is an $\epsilon > 0$ such that $f(s)$ is analytic in the disc $|s-1| \leq 1+\epsilon$. This means the Taylor series of $f(s)$ must converge in this disc. The p^{th} derivative of $f(s)$ is given by

$$\begin{aligned} f^{(p)}(s) &= \sum_{n=1}^{\infty} \frac{a_n(-\log n)^p}{n^s} \\ \to f^{(p)}(1) &= \sum_{n=1}^{\infty} (-1)^p \frac{a_n(\log n)^p}{n} \end{aligned}$$

The Taylor series of $f(s)$ around $s = 1$ is given by

$$f(s) = \sum_{p=1}^{\infty} \frac{f^{(p)}(1)}{p\,!}(s-1)^p, \qquad |s-1| \leq 1+\epsilon.$$

Now for $s = -\epsilon$, we obtain

$$f(-\epsilon) = \sum_{p=1}^{\infty} \frac{f^{(p)}(1)}{p\,!}(1+\epsilon)^p(-1)^p.$$

But $(-1)^p f^{(p)}(1) = \displaystyle\sum_{n=1}^{\infty} \frac{a_n(\log n)^p}{n}$ is a convergent series with positive terms. This means the following double sum converges:

$$f(-\epsilon) = \sum_p \sum_n a_n \frac{1}{p\,!} \frac{(1+\epsilon)^p(\log n)^p}{n}.$$

But this sum can be re-expressed as

$$\begin{aligned} f(-\epsilon) &= \sum_n \frac{a_n}{n^s} \sum_{p=1}^{\infty} \frac{1}{p\,!}(1+\epsilon)^p(\log n)^p \\ &= \sum_n \frac{a_n}{n^s} e^{(\log n)(1+\epsilon)} \\ &= \sum_n a_n n^{\epsilon} \end{aligned}$$

This is the Dirichlet series we started with evaluated at $-\epsilon$! Therefore, the series converges at $s = -\epsilon$, and thus, by Theorem 3.1, for all $\operatorname{Re}(s) > -\epsilon$. □

So far we have argued that if $L(1,\chi) = 0$ for some non-principal χ, then $\zeta_m(s)$ must be analytic for $\operatorname{Re} s > 0$. We saw that $\zeta_m(s)$ is a Dirichlet series with positive coefficients. Moreover, since *all* $L(s,\chi)$ are convergent for $\operatorname{Re}(s) > 1$, $\zeta_m(s)$ converges in this region as well. Since $\zeta_m(s)$ is analytic in the region $\operatorname{Re}(s) > 0$, Lemma 4.3 tells us we can push back the region of convergence of $\zeta_m(s)$ to $\operatorname{Re}(s) > 0$. Our contradiction is at hand.

Let s be a real number greater than 1. Consider the p^{th} factor in the product expansion of $\zeta_m(s)$ (which is valid for $\operatorname{Re} s > 1$)

$$\begin{aligned}\left(\frac{1}{1-p^{-f(p)s}}\right)^{g(p)} &= (1 + p^{-f(p)s} + p^{-2f(p)s} + \cdots)^{g(p)} \\ &\geq 1 + p^{-f(p)g(p)s} + p^{-2f(p)g(p)s} + \cdots \\ &= 1 + p^{-\phi(m)s} + p^{-2\phi(m)s} + \cdots \\ &= \frac{1}{1-p^{\phi(m)s}}.\end{aligned}$$

This shows that for $s > 1$,

$$\begin{aligned}\zeta_m(s) = \prod_{\chi} L(s,\chi) = \prod_{\chi}\prod_{p}\Big(\frac{1}{1-\chi(p)p^{-s}}\Big) &= \prod_{p}\Big(\frac{1}{1-p^{-f(p)s}}\Big)^{g(p)} \\ &\geq \prod_{p}\Big(\frac{1}{1-p^{-\phi(m)s}}\Big) = \sum_{(n,m)=1} n^{-\phi(m)s}\end{aligned} \tag{27}$$

Thus, $\zeta_m(s)$ has all its *coefficients* greater than those of the series in (27). These coefficients of $\zeta_m(s)$ remain unchanged if we take s between 0 and 1. But the series (27) diverges for $s = \frac{1}{\phi(m)} > 0$. Hence $\zeta_m(s)$ is unbounded for this value of s, and thus the series $\zeta_m(s)$ diverges for a value of s whose real part is greater than zero. This is a contradiction, since we showed $\zeta_m(s)$ converges for $\operatorname{Re} s > 0$. This completes the proof of Theorem 3.6. □

Remark. Even though the product expansion of $\zeta_m(s)$ is only valid for $\operatorname{Re} s > 1$, we were able to use the expansion to look at the coefficients of the *series* representation for $\zeta_m(s)$, which is valid for $\operatorname{Re} s > 0$

References

[1] K. Ireland, M. Rosen, *A Classical Introduction to Modern Number Theory* Second Edition. Springer, New York, 1990.

[2] A. Knapp, *Elliptic Curves* Princeton UP, New Jersey, 1992.

[3] J-P. Serre, *A course in Arithmetic* Springer, New York, 1973.

[4] J-P. Serre, *Linear Representations of Finite Groups* Springer, New York, 1977.