1 Elementary Results on Primes in Arithmetic Progressions

Dirichlet considered a question very similar to the one which inspired Euler’s introduction of the \(\zeta\)-function: namely, how the primes are distributed modulo \(m\). The simplest question of this type is whether there are infinitely many primes congruent to \(a\) modulo \(m\). Obviously there can only be infinitely many primes of this form if \(a\) and \(m\) are relatively prime. Unfortunately, this problem turned out to be much more difficult than proving that there are infinitely many primes. Elementary results along the lines of Euclid’s proof only sufficed to show very special cases. For example,

**Proposition 1.1.** There are infinitely many primes \(p \equiv 3 \pmod{4}\).

**Proof.** Suppose there were only finitely many such primes, \(p_1, p_2, \ldots, p_n\). Consider the number

\[
Q = 4p_1p_2\cdots p_n - 1.
\]

Clearly this number is not divisible by any of the primes which are 3 modulo 4. Thus, \(Q \equiv 3 \pmod{4}\) is a product of primes all of which are 1 modulo 4. This is clearly a contradiction.

**Proposition 1.2.** There are infinitely many primes \(p \equiv 1 \pmod{4}\).

**Proof.** Here we use the fact from basic number theory that \(-1\) is a square modulo \(p\) exactly when \(p \equiv 1 \pmod{4}\). Again, suppose there were only finitely many primes \(p_1, p_2, \ldots, p_n \equiv 1 \pmod{4}\). Let

\[
Q = (2p_1p_2\cdots p_n)^2 + 1.
\]

Clearly \(Q\) is not divisible by any of the primes which are 1 modulo 4. Since \(-1 \equiv (p_1p_2\cdots p_n)^2 \pmod{Q}\), any prime which divides \(Q\) must be 1 modulo 4. Again, this is a contradiction.

Although similar methods will work for \(m = 3\) or \(m = 6\), they are doomed to failure in general. The above arguments generalize only to the following two results:

**Proposition 1.3.** Suppose \(H\) is a proper subgroup of the group of units \((\mathbb{Z}/m\mathbb{Z})^\times\). Then there exist infinitely many primes which are not in \(H\) when reduced modulo \(m\).

**Proof.** Suppose there were only finitely many such primes, \(p_1, p_2, \ldots, p_n\). Consider the number

\[
Q = mp_1p_2\cdots p_n + p_1.
\]

Clearly this number is not divisible by any of the primes which are not in \(H\). Thus, \(Q \equiv x \pmod{m}\) is a product of primes all of which are in \(H\). This is clearly a contradiction.

**Proposition 1.4.** There are infinitely many primes congruent to 1 modulo \(m\) for any \(m\).

**Proof.** Let \(\Phi_m(x)\) be the \(m\)th cyclotomic polynomial (that is the minimal polynomial of a primitive \(m\)th root of unity). Suppose there are finitely many primes \(p_1, p_2, \ldots, p_n \equiv 1 \pmod{m}\). Let \(N = m \cdot p_1 \cdot p_2 \cdots p_n\). Consider \(\Phi_m(N)\). Suppose that a prime \(q\) divides \(\Phi_m(N)\). Then, modulo \(q\), we must have a primitive \(m\)th root of unity. Therefore, \(q \equiv 1 \pmod{m}\). Thus \(q = p_i\) for some \(i\). However, none of the \(p_i\) can divide \(\Phi_m(N)\). This is a contradiction.
In fact, using class field theory, one can show that it is impossible to find a polynomial which outputs only numbers divisible by primes congruent to \( a \pmod{m} \) for a general relatively prime pair \( a \) and \( m \).

Rather than trying to use these sorts of elementary proofs, Dirichlet instead tried to adapt Euler’s analytic methods to this situation.

2 A Special Case of Dirichlet’s Theorem

In particular, Dirichlet wanted to prove

**Theorem 2.1.** For any relatively prime positive integers \( a \) and \( m \), the series

\[
\sum_{p \equiv a \pmod{m}} \frac{1}{p}
\]

diverges.

Let us first consider the particular case \( m = 4 \). Since we will consider things more rigorously in the general case, here we shall be a bit lax.

The most obvious attempt at modifying Euler’s method is to consider the function

\[
f_1(s) = \prod_{p \equiv 1 \pmod{4}} \frac{1}{1 - p^{-s}},
\]

By the same arguments as used the last section we can conclude that

\[
\sum_{p \equiv 1 \pmod{4}} p^{-s} = \log f_1(s) + O(1).
\]

Unfortunately, \( f_1(1) \) does not obviously diverge. If we multiply out the Euler product, we see that

\[
f_1(s) = \sum_{n \in S} n^{-s},
\]

where \( S \) is the set of all numbers which are products of primes which are 1 modulo 4. This is entirely unhelpful. We need to find some functions whose Euler factorizations depend only on what the prime is modulo 4, and where the terms in the series do not depend on the prime factorization of \( n \).

Dirichlet’s insight was to look at the functions

\[
L_1(s) = \sum_{n \text{ odd}} (-1)^{\frac{n-1}{2}} n^{-s} = \prod_{p \equiv 1 \pmod{4}} \frac{1}{1 - p^{-s}} \prod_{p \equiv 3 \pmod{4}} \frac{1}{1 + p^{-s}}
\]

\[
L_0(s) = \sum_{n \text{ odd}} n^{-s} = \prod_{p \equiv 1 \pmod{4}} \frac{1}{1 - p^{-s}} \prod_{p \equiv 3 \pmod{4}} \frac{1}{1 - p^{-s}} = (1 - 2^{-s})\zeta(s).
\]

Just as in the last section, we can take logarithms and use the Taylor series expansion. As in the last section the contribution from quadratic and higher terms in the Taylor series are bounded. Thus,

\[
\log L_1(s) = \sum_{p \equiv 1 \pmod{4}} p^{-s} - \sum_{p \equiv 3 \pmod{4}} p^{-s} + O(1)
\]

\[
\log L_0(s) = \sum_{p \equiv 1 \pmod{4}} p^{-s} + \sum_{p \equiv 3 \pmod{4}} p^{-s} + O(1)
\]

Therefore,

\[
\frac{1}{2} (\log L_0(s) + \log L_1(s)) = \sum_{p \equiv 1 \pmod{4}} p^{-s} + O(1)
\]

\[
\frac{1}{2} (\log L_0(s) - \log L_1(s)) = \sum_{p \equiv 3 \pmod{4}} p^{-s} + O(1)
\]
Thus in order to prove this special case of Theorem 2.1, we need only show that \( \log L_0(s) + \log L_1(s) \) and \( \log L_0(s) - \log L_1(s) \) are both unbounded as \( s \to 1^+ \). Obviously \( (1 - 2^{-s}) \zeta(s) \) blows up at \( s = 1 \). Hence we’ve reduced this problem to showing that \( \log L_1(1) \) is finite.

But, \( L_1(s) \) is an alternating series; thus, we can bound it by the first two partial sums, i.e. \( \frac{2}{3} 3^{-s} < L_1(s) < 1 \). Thus \( \log \frac{2}{3} < \log L_1(1) < 0 \), so we have proved Theorem 2.1 for the case of \( m = 4 \).

## 3 Some General Results on Dirichlet Series

Before we attack the general proof of Theorem 2.1, it will be useful to have a few technical definitions and results.

**Definition 3.1.** A Dirichlet series is a series of the form,

\[
    f(s, a_n) = \sum_{n=1}^{\infty} \frac{a_n}{n^s},
\]

where \( a_n \) is some sequence of complex numbers.

We will want to know when a Dirichlet series actually converges. In order to prove this, we will need the following theorem which is the analog for sums of integration by parts.

**Theorem 3.2 (Abel Summation Formula or Summation by Parts).** Suppose \( (a_n) \) and \( (b_n) \) are two sequences. Then \( \sum_{n=1}^{N} n a_n(b_n - b_{n-1}) = a_N b_N - a_1 b_1 - \sum_{n=1}^{N} b_n(a_n - a_{n-1}) \).

**Proof.** To verify this theorem, we simply check that each term \( a_i b_j \) occurs with the same multiplicity on both sides.

Using the notion of a Stieltjes integral we can rewrite the Abel Summation Formula as a generalization of integration by parts.

**Definition 3.3.** Let \( \alpha \) be a non-decreasing (not necessarily continuous function) on \([a, b]\) and \( f \) a function bounded on \([a, b]\). Let \( P \) be some partition of \([a, b]\), let \( \Delta \alpha_i = \alpha(x_i) - \alpha(x_{i-1}) \). Define the upper and lower sums, \( U(P, f, \alpha) = \sum_{i=1}^{\ell} M_i \Delta \alpha_i \) and \( L(P, f, \alpha) = \sum_{i=1}^{\ell} m_i \Delta \alpha_i \), where \( M_i \) and \( m_i \) are the maximum and minimum respectively of \( f \) on the \( i \)th interval in the partition. We define the upper and lower Stieltjes integrals to be \( \int_a^b f \, d\alpha = \lim_{\ell \to \infty} U(P, f, \alpha) \) and \( \int_a^b f \, d\alpha = \lim_{\ell \to \infty} L(P, f, \alpha) \). If the upper and lower integrals are equal we denote their common value by \( \int_a^b f \, d\alpha \) and say that \( f \) is integral with respect to \( \alpha \).

For our purposes we will only be looking at Stieltjes integrals where either \( \alpha \) is differentiable in which case \( \int_a^b f \, d\alpha = \int_a^b f(x) \alpha'(x) \, dx \) or when \( \alpha \) is a step function. In the latter situation the integral is simply a sum over the jumps of the value of the function at that jump times the size of the jump. For example, \( \int_a^b f \, d[x] = \sum_{n \in \mathbb{Z} \cap [a, b]} f(n) \). Lastly in the case of \( \alpha \) a step function, we shall always take the value of \( \alpha \) at each jump to be halfway between the values on either side. Although this is simply a convention, later on it will hopefully become clear why it is such a useful one.

In the context of a Stieltjes integral, summation by parts is simply a special case of integration by parts for which holds for all Stieltjes integrals. (This is slightly misleading, in reality one *uses* summation by parts to prove this general integration by parts formula.) Generally we will prefer to write things as Stieltjes integrals and use integration by parts, but occasionally we will directly refer to Abel Summation.

**Proposition 3.4.** If \( \sum a_n n^{-s_0} \) converges, then the Dirichlet series \( f(s, a_n) \) converges for all complex numbers \( s \) with \( \Re(s) > s_0 \). In fact, this convergence is uniform in any wedge to the right of the point \( s_0 \): \( \{ s : \Re(s) > \Re(s_0) \text{ and } 0 < \frac{|s-s_0|}{\Re(s)-\Re(s_0)} < M \} \), where \( M \) is an arbitrary positive constant. (Since this convergence is uniform, \( f(s, a_n) \) is analytic on that region.)

**Proof.** Letting \( M \) grow arbitrarily shows that the second assertion implies the first. Without loss of generality, we can assume \( s_0 = 0 \) (since we can look at the Dirichlet series \( \sum_n a_n n^{-s_0} n^{-s} \)). Also, without loss of generality, we can subtract off the first term \( a_1 \) and so assume \( a_1 = 0 \).
Since we are assuming that \( \sum_n a_n \) converges for any \( \varepsilon > 0 \), there exists an integer \( N \) such that for any \( \ell, m > N, |A_{\ell, m}| < \varepsilon \).

We want to get a good bound on \( |\sum_{n=\ell}^m a_n n^{-s}| \). By Abel’s summation formula,

\[
|\sum_{n=\ell}^m a_n n^{-s}| = |A_{\ell, m} b_m - \sum_{n=\ell}^m A_{\ell, m} ((n + 1)^{-s} - n^{-s})| < \varepsilon \left( 1 + \sum_{n=\ell}^m |e^{-s \log n} - e^{-s \log (n+1)}| \right).
\]

To get a bound on that last term, we notice that for any \( \alpha > \beta \geq 0 \),

\[
e^{-\alpha z} - e^{-\beta z} = z \int_{\alpha}^{\beta} e^{-t \Re(z)} dt.
\]

Therefore,

\[
|e^{-\alpha z} - e^{-\beta z}| \leq |z| \int_{\alpha}^{\beta} e^{-t \Re(z)} dt = \frac{|z|}{\Re(z)} \left( e^{-\alpha \Re(z)} - e^{-\beta \Re(z)} \right).
\]

Applying this to our particular case, we see that

\[
|\sum_{n=\ell}^m a_n n^{-s}| < \varepsilon \left( 1 + M \sum_{n=\ell}^m e^{-\Re(s) \log n} - e^{-\Re(s) \log (n+1)} \right)
= \varepsilon \left( 1 + M (e^{-\Re(s) \log \ell}) - e^{-\Re(s) \log (m)} \right) < \varepsilon (1 + M).
\]

Thus for large enough \( N \), this goes to zero independently of \( s \), so the series converges uniformly in this region.

It turns out that if a function can be written as a Dirichlet series then it can be done so in only one way.

**Proposition 3.5.** cf. [7, Thm. 11.4] Suppose that \( \sum_n \frac{a_n}{n^s} = 0 \) on some right halfplane \( \Re(s) > \sigma_0 \). Then, \( a_n = 0 \) for all \( n \). Therefore, if we have two Dirichlet series with \( \sum_n \frac{b_n}{n^s} = \sum_n \frac{c_n}{n^s} \) for all \( \Re(s) > \sigma_0 \), then \( b_n = c_n \) for all \( n \).

**Proof.** Without loss of generality, by considering \( a_n n^{-\sigma_0} \) we can assume that \( \sigma_0 = 0 \). Thus, in order to have \( \sum_n \frac{a_n}{n^s} \) converge near \( s = 0 \), we must have \( a_n = O(1) \). Now suppose that \( a_N \) is the first non-zero term. Then

\[
0 = a_N N^{-s} \left( 1 + \sum_{n \geq N} \frac{a_n}{a_N} \left( \frac{n}{N} \right)^{-s} \right).
\]

Multiplying by \( N^s \) we get

\[
0 = a_N \left( 1 + \sum_{n \geq N} \frac{a_n}{a_N} \left( \frac{n}{N} \right)^{-s} \right).
\]

Now send \( s \to +\infty + 0i \). Each of the terms in the sum dies exponentially. Therefore, since the coefficients are bounded, the whole sum dies. Therefore, \( 0 = a_N \). This is a contradiction; therefore, \( a_n = 0 \) for all \( n \).

For the second conclusion, we simply consider the Dirichlet series \( \sum_n \frac{b_n - c_n}{n^s} = 0 \), from which it follows that \( b_n - c_n = 0 \) and thus \( b_n = c_n \) for all \( n \).

**Definition 3.6.** A sequence \( a_n \) is called multiplicative if \( a_n a_m = a_{nm} \) for all relatively prime positive integers \( n \) and \( m \). Similarly, a sequence \( a_n \) is called strongly multiplicative if \( a_n a_m = a_{nm} \) for all pairs of positive integers.
Theorem 3.7. If \( a_n \) is multiplicative then the Dirichlet series \( f(s, a_n) \) has the Euler factorization:

\[
f(s, a_n) = \prod_p \sum_{\ell=1}^{\infty} \frac{a_p^{\ell} n^{\ell}}{p^{s\ell}}.
\]

Furthermore if \( a_n \) is strongly multiplicative, summing this geometric series we see that

\[
f(s, a_n) = \prod_p \frac{1}{1 - \frac{a_p}{p^s}}.
\]

Proof. The proof here is identical to the proof of Theorem ??.

Finally we have a formula for the product of two Dirichlet series.

Proposition 3.8. \( f(s, a_n) f(s, b_n) = f(s, \sum_{d|n} a_d b_\frac{n}{d}) \).

Proof. By definition,

\[
f(s, a_n) f(s, b_n) = \sum_{m=1}^{\infty} \sum_{\ell=1}^{\infty} \frac{a_m b_\ell}{(m\ell)^s}.
\]

Make a change of variables \( n = m\ell \) and \( d = m \) to get,

\[
f(s, a_n) f(s, b_n) = \sum_{n=1}^{\infty} \sum_{d|n} \frac{a_d b_\frac{n}{d}}{n^s} = f(s, \sum_{d|n} a_d b_\frac{n}{d}).
\]

4 Dirichlet’s L-series

In order to generalize Dirichlet’s argument from the case of \( m = 4 \) to a general \( m \), we should look at all Dirichlet series with particularly nice Euler factorizations in which \( a_p \) depends only on what \( n \) is modulo \( m \) and vanishes when \( p|m \). From our results above, it is clear that we should be looking at the series corresponding to sequences of the following form:

Definition 4.1. Let a Dirichlet character modulo \( m \) be any function from \( \chi : \mathbb{Z} \to \mathbb{C} \) with the properties:

1. If \( n \) and \( m \) are not relatively prime, then \( \chi(n) = 0 \).
2. If \( n \) and \( m \) are relatively prime, then \( |\chi(n)| = 1 \).
3. If \( n_1 \) and \( n_2 \) are any two positive integers, then \( \chi(n_1 n_2) = \chi(n_1) \chi(n_2) \).

If we restrict a Dirichlet character to \((\mathbb{Z}/m\mathbb{Z})^\times\) we get a homomorphism. (By abuse of notation, we will also call it \( \chi : (\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times \).) Furthermore, if one considers any homomorphism \( \chi : (\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times \) it will be a Dirichlet character modulo \( k \) for any \( m|k \). If \( \chi \) is an injective homomorphism \( (\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times \), then we say that the corresponding Dirichlet character is a primitive character modulo \( m \). The homomorphism sending everything to 1 and the corresponding Dirichlet character will both be called trivial.

The above notion of homomorphisms \( \chi : (\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times \) generalizes to the more general notion of an abelian group character, which is a homomorphism \( \chi : G \to \mathbb{C}^\times \) where \( G \) is an abelian group.

As far as I can tell, the linguistic history of this term character is quite the opposite of what one might expect. The notion of a Dirichlet character came before the notion of a general character, and the name character seems to come from the fact that it is a generalization of the quadratic character that is the quadratic nature of a number modulo \( m \).
Definition 4.2. If \( \chi \) is a Dirichlet character modulo \( m \), then define the Dirichlet L-series

\[
L(s, \chi) = \sum_{n=1}^{\infty} \chi(n)n^{-s} = \prod_p \frac{1}{1 - \frac{\chi(p)}{p^s}}.
\]

If \( \chi_0 \) is the trivial character modulo \( m \), then obviously

\[
L(s, \chi_0) = \prod_{p|m} \left(1 - \frac{\chi(p)}{p^s}\right) \zeta(s)
\]
converges for any \( \text{Re}(s) > 1 \).

For a nontrivial character, we certainly could get the same region of convergence since \( \sum_n \chi(n)n^{-s} < \sum_n |\chi(n)|n^{-s} \). However, we might hope to find some cancellation and be able to find a larger region of convergence.

Lemma 4.3 (Dirichlet). If \( G \) is an abelian group, and \( \chi \) is a nontrivial character of that group, then

\[
\sum_{g \in G} \chi(g) = 0.
\]

Proof. For any \( h \in G \), notice that

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

But, as \( g \) runs over all of \( G \), so does \( hg \). Hence,

\[
\chi(h) \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(g).
\]

Therefore, either \( \chi(h) = 1 \) or \( \sum_{g \in G} \chi(g) = 0 \). Since \( \chi \) is nontrivial, we can choose \( h \) so that the former condition is not true. Thus the lemma is proved.

Corollary 4.4. If \( \chi \) is a nontrivial Dirichlet character modulo \( m \), then \( L(s, \chi) \) converges for \( \text{Re}(s) > 0 \).

Proof. By Lemma 4.3, we know that \( \sum_n \chi(n) \) is bounded. Thus, \( \sum_n \chi(n)n^{-s} \) converges for any positive \( s \), which by Proposition 3.4 is enough.

5 Reducing Dirichlet’s Theorem to an Analytic Theorem

Now that we have proved enough technical results, we can return to Dirichlet’s original question. Following Euler, we notice that

\[
\log L(s, \chi) = \sum_p \log \frac{1}{1 - \frac{\chi(p)}{p^s}}.
\]

Again the Taylor series expansion is valid, and our estimate on the terms still holds:

\[
\left| \sum_{n=2}^{\infty} \frac{1}{n} \sum_p \frac{\chi(p)}{p^{ns}} \right| < \int_2^{\infty} \int_2^{\infty} x^{-1}y^{-sz}dydx < \frac{1}{32}.
\]

Therefore,

\[
\log L(s, \chi) = \sum_p \chi(p)p^{-s} + O(1).
\]

To get Dirichlet’s result, we need to write \( f_a(s) = \sum_{p \equiv a (m)} p^{-s} \) as a sum of \( \log L(s, \chi) \) for various \( \chi \).

We know that

\[
\log L(s, \chi) = \sum_{a \in (\mathbb{Z}/m\mathbb{Z})^*} \chi(a)f_a(s) + O(1).
\]

In order to get this result, Dirichlet noticed and proved a finite analog of Fourier inversion called the orthogonality of characters.
Theorem 5.1 (Dirichlet). If \( \chi_1 \) and \( \chi_2 \) are distinct characters of a finite abelian group \( G \), then
\[
\sum_{g \in G} \chi_1(g)\chi_2(g)^{-1} = 0.
\]
Furthermore, if we let \( G^* \) denote the dual group of characters of \( G \), then, if \( g \neq h \),
\[
\sum_{\chi \in G^*} \chi(gh^{-1}) = 0.
\]

Proof. The first assertion follows immediately from applying Lemma 4.3 to the character \( \chi_1\chi_2^{-1} \). Together with the obvious fact that \( \sum_{g \in G} \chi(g)\chi(g)^{-1} = |G| \), this implies that the matrix \( (\chi(g)\chi(g)^{-1})_{\chi,g} \) has orthogonal rows. By the structure theorem for finite abelian groups, it is easy to see that \( |G| = |G^*| \) (since it is obvious for cyclic groups). Thus this matrix is a square matrix. By standard linear algebra we know that having orthogonal columns is the same as having orthogonal rows and the second half of the result follows.

Thus by Theorem 5.1, we see that the characters \( \chi \) are all linearly independent and thus form a basis of the space of all complex valued functions on \( G \). In particular,
\[
\frac{1}{|G|} \sum_{\chi \in G^*} \chi(h^{-1})\chi(g) = \begin{cases} 1 & \text{if } g=h \\ 0 & \text{otherwise} \end{cases}.
\]
Therefore, if we let \( G = (\mathbb{Z}/m\mathbb{Z})^\times \),
\[
f_a(s) = \frac{1}{|G|} \sum_{\chi \in G^*} \chi(a^{-1}) \log L(s, \chi) + O(1).
\]
Since \( \log L(1, \chi_0) \) blows up, and \( \log L(1, \chi) < \infty \) for all nontrivial characters, all that remains is to show that \( \log L(1, \chi) > -\infty \). That is to say, we have shown that Theorem 2.1 is equivalent to the following theorem:

Theorem 5.2 (Dirichlet). If \( \chi \) is a nontrivial Dirichlet character, then \( L(1, \chi) \neq 0 \).

Proof. First we claim that for every \( m \) there is at worst one \( \chi \) with \( L(1, \chi) = 0 \). Notice that
\[
\sum_{\chi \in G^*} \log L(s, \chi) = \varphi(m) \sum_{k=1}^{\infty} \frac{1}{k} \sum_{p | m} p^{-ks} > 0.
\]
Now we already know that \( \lim_{s \to 1^+} (s-1)L(s, \chi_0) \) is finite. Hence, \( \log L(s, \chi_0) = \log \frac{1}{s-1} + O(1) \). But we know by Theorem 3.4 that for all the other \( \chi \neq \chi_0 \), \( L(s, \chi) \) are analytic near 1. Therefore, for all of these, \( \log L(s, \chi) \) either goes to \( -\infty \) or is bounded.

Suppose one of these series, for instance \( L(s, \tau) \), had a zero at 1. Since it is analytic, by considering the Taylor expansion \( \frac{L(s, \tau)}{s-1} \) is analytic and bounded at 1. Hence, \( \log L(s, \tau) = -\log(s-1) + O(1) \). So, if we had \( L(1, \tau_1) = L(1, \tau_2) = 0 \), then
\[
\sum_{\chi \in G^*} \log L(s, \chi) = \log(s-1) - 2\log(s-1) + \varepsilon(s),
\]
where \( \varepsilon(s) \) is either bounded or goes to \(-\infty \) as \( s \to 1^+ \). Thus, the right hand side would be negative for small enough \( s \), and we’ve reached a contradiction.

Notice that this proves the theorem for every character whose image is not contained in the reals. In this case there is a distinct character \( \bar{\chi}(n) = \chi(n) \) with \( L(1, \chi) = 0 \iff L(1, \bar{\chi}) = 0 \).

Furthermore, this shows that there is at worst 1 primitive Dirichlet character with \( L(s, \chi) = 0 \). If there were two, say one primitive modulo \( m_1 \) and the other primitive modulo \( m_2 \), then we could consider both of them as Dirichlet characters modulo \( m_1m_2 \). Thus their \( L \)-series modulo \( m_1m_2 \) would differ from.
the originals by only finitely many terms. Thus there would be two different Dirichlet $L$-series modulo $m_1 m_2$ with $L(1, \chi) = 0$, which is a contradiction.

So we have proven this theorem for all but one primitive character which must be real. Any real primitive character modulo $m$ must have, for $g$ a generator of $\mathbb{Z}/m\mathbb{Z}$, $\chi(g^k) = (-1)^k$. Clearly, this means $\chi(a)$ is 1 or $-1$ exactly when $a$ is a square or a non-square respective. Thus, $\chi(a) = \left( \frac{a}{m} \right)$.

Dirichlet spent several years trying to prove that $L(1, \left( \frac{\cdot}{m} \right)) \neq 0$. Eventually he was able to prove this using his famous class number formula, which we will prove (in part) in a later lecture. In next week’s homework we will give Dirichlet’s proof in the special case of $m$ is prime. Finally the next week in homework we will be giving a much faster modern proof of the general case.