Dirichlet's Theorem

Martin Orr

August 21, 2009

1 Introduction

The aim of this article is to prove Dirichlet's theorem on primes in arithmetic progressions:

Theorem 1.1. *If $m, a \in \mathbb{N}$ are coprime, then there are infinitely many primes congruent to a mod m.*

This was first proved by Dirichlet in 1837, in a paper which Davenport described as marking the beginning of analytic number theory.

Dirichlet's proof used only real analysis, including some big integrals. This article gives a more modern presentation, using a little complex analysis to simplify one part and using an entirely complex-analytic proof for the last section (which is less work than Dirichlet's methods, although of less independent interest).

A so-called "elementary" proof (i.e. not using infinitary methods of analysis) was found by Selberg in 1949, but is reputed to be long and difficult.

In this article, I have aimed to present the full details of all the analysis. This makes things rather long, but it divides up into several mostly independent sections. The outline of the proof (Section 2) explains how these fit together.

On the other hand I have not bothered to prove the key results on characters of finite abelian groups which are used in Section 3, in order to avoid distraction from the analysis. If you don't know any character theory, you can simply take this section on trust or read almost any other account of the proof of Dirichlet's theorem.

I relied upon several sources for preparing this article, because I couldn't find any which collects together all the details. Probably the canonical reference is Davenport's *Multiplicative Number Theory*. I also used Neukirch's *Algebraic Number Theory*, Jameson's *The Prime Number Theorem*, a project by Daniel Weissman of the University of Washington[1] and Tom Körner's Part III Topics in Fourier Analysis notes.[2] An English translation of Dirichlet's original paper by Ralf Stephan can be found on the arXiv.[3]

2 Outline of proof

Instead of proving Theorem 1.1 directly, we shall prove the following stronger theorem:

Theorem 2.1. *If $m, a \in \mathbb{N}$ are coprime, then $\sum \frac{1}{p}$, summed over primes $p \equiv a \bmod m$, is infinite.*

In fact we work with $\sum_{p\equiv a} p^{-s}$, and show that this tends to $+\infty$ as $s \to 1$ from above. The reason for introducing the variable s is that many of the sums we consider (including of course this one) do not converge or only converge conditionally at $s = 1$, but they converge absolutely for $s > 1$, and absolute convergence is necessary for parts of the argument. (Note: seeing that $\sum_{p\equiv a} p^{-s} \to +\infty$ as $s \to 1+$ implies that $\sum_{p\equiv a} p^{-1} = +\infty$ is not hard, since $\sum_{p\equiv a} p^{-1}$ must be greater than $\sum_{p\equiv a} p^{-s}$ for each $s > 1$.)

The proof depends on considering functions of the form

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

where χ is a *Dirichlet character* (a function $\mathbb{Z} \to \mathbb{C}$ which is periodic mod m, such that $\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in \mathbb{Z}$ and $\chi(a) = 0$ iff $\gcd(a, m) > 1$).

We study these L-functions because the sum over all n has more accessible analytic properties than a sum over primes. They are related to the primes by an Euler product: $L(\chi, s) = \prod_p (1 - \chi(p)p^{-s})^{-1}$.

The algebraic part of the proof is to show that $\sum_{p\equiv a} p^{-s}$ can be written as a linear combination

$$\sum_{\chi} c_\chi \sum_p \chi(p) p^{-s},$$

where the sum is over Dirichlet characters χ mod m. Furthermore, $c_{\mathbb{1}} \neq 0$, where $\mathbb{1}$ is the *principal Dirichlet character*: $\mathbb{1}(n) = 1$ for all n coprime to m (Lemma 3.1).

[1] http://modular.math.washington.edu/129/projects/weissman/project.pdf
[2] http://www.dpmms.cam.ac.uk/~twk/Fou3.pdf
[3] http://arxiv.org/abs/0808.1408v1

Hence it will suffice to show that $\sum_p \mathbb{1}(p)p^{-s} \to +\infty$ as $s \to 1+$ and $\sum_p \chi(p)p^{-s}$ is uniformly bounded on $s > 1$ for each non-principal Dirichlet character χ.

We relate $\sum_p \chi(p)p^{-s}$ to $L(\chi, s)$ using the following steps:

(i) $\prod_p(1 - \chi(p)p^{-s})$ converges and $\prod_p(1 - \chi(p)p^{-s})^{-1} = L(\chi, s)$ for $s > 1$, for all χ (Lemma 4.1);

(ii) $\sum_p \log(1 - \chi(p)p^{-s})$ converges uniformly on $s > 1 + \delta$ for any $\delta > 0$ and $\left| - \sum_p \chi(p)p^{-s} - \sum_p \log(1 - \chi(p)p^{-s}) \right|$ is bounded independent of s for $s > 1$, for all χ (Corollaries 5.2 and 5.3);

(iii) $L(\chi, s) = \exp(\sum_p \log(1 - \chi(p)p^{-s}))$ for $s > 1$ (Lemma 5.4).

It follows that if $L(\chi, s)$ tends to a finite, non-zero limit as $s \to 1+$, then $\sum_p \chi(p)p^{-s}$ tends to a finite limit (Corollary 5.6).

It also follows that $\sum_p \mathbb{1}(p)p^{-s} \to +\infty$ as $s \to 1+$ (Corollary 6.3).

All that is left is to show that $L(\chi, s)$ tends to a finite, non-zero limit as $s \to 1+$, for non-principal Dirichlet characters χ. In fact $L(\chi, s)$ is defined and continuous at $s = 1$ so $L(\chi, s) \to L(\chi, 1)$ as $s \to 1+$ (Corollary 7.5).

Showing that $L(\chi, 1) \neq 0$ is surprisingly hard and falls into two cases, depending on whether χ takes only real values or not.

We will use complex analysis to show this, so need first to establish that $L(\chi, s)$ defines a meromorphic function on $\Re s > 0$, and is analytic except for $\chi = \mathbb{1}$ at $s = 1$ (Sections 7 and 8).

Considering $\prod_\chi L(\chi, s)$ shows that $L(\chi, 1)$ can be zero for at most one character χ mod m, and it follows that $L(\chi, 1) \neq 0$ if χ takes non-real values (Lemma 9.2).

The hard case is when χ takes only real values. The proof I have given here makes significant use of complex analysis, as well as the previously developed theory of Euler products and L-series. It is based on the Taylor expansion of the carefully chosen function $\psi(s) = L(\chi, s)L(\mathbb{1}, s)/L(\mathbb{1}, 2s)$ (Lemma 10.4).

3 Dirichlet characters

The aim of this section, which is algebraic rather than analytic in flavour, is to establish the following lemma:

Lemma 3.1. *There are constants $c_\chi \in \mathbb{C}$, with $c_{\mathbb{1}} \neq 0$, such that*

$$\sum_{p \equiv a \bmod m} p^{-s} = \sum_\chi c_\chi \sum_p \chi(p) p^{-s},$$

where the sum on the right is over Dirichlet characters χ mod m.

Recalling the definition from the previous section, a *Dirichlet character mod m* is a function $\chi : \mathbb{Z} \to \mathbb{C}$ such that

(i) $\chi(a) = \chi(b)$ whenever $a \equiv b \bmod m$;
(ii) $\chi(a) = 0$ iff $(a, m) > 1$;
(iii) $\chi(ab) = \chi(a)\chi(b)$ for all a, b.

For convenience, define a function $\delta_{m,a} : \mathbb{Z} \to \mathbb{C}$ by

$$\delta_{m,a}(n) = \begin{cases} 1 & \text{if } n \equiv a \bmod m, \\ 0 & \text{if } n \not\equiv a \bmod m. \end{cases}$$

We will show that there are constants c_χ such that $\delta_{m,a} = \sum_\chi c_\chi \chi$. In fact the following much more general statement is true:

Lemma 3.2. *The Dirichlet characters mod m are a basis for the vector space of functions $f : \mathbb{Z} \to \mathbb{C}$ satisfying*

(i) $\chi(a) = \chi(b)$ *whenever* $a \equiv b \bmod m$*;*
(ii) $\chi(a) = 0$ *whenever* $\gcd(a, m) > 1$.

Proof. There is an obvious bijection between the set of functions described in the lemma and the set of functions $(\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}$.

Under this bijection, Dirichlet characters correspond to group homomorphisms $(\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times$.

Since $(\mathbb{Z}/m\mathbb{Z})^\times$ is an abelian group, all its irreducible characters are one-dimensional, so homomorphisms $(\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times$ are the same as irreducible characters of $(\mathbb{Z}/m\mathbb{Z})^\times$.

It is a standard fact from character theory that the characters of a finite group are a basis for the complex-valued functions on its set of conjugacy classes (and for an abelian group, conjugacy classes are the same as elements of the group). □

The function $\delta_{m,a}$ satisfies the conditions of the lemma if a and m are coprime, so is a linear combination of Dirichlet characters mod m.

To compute the coefficient $c_{\mathbb{1}}$, we work again with the corresponding functions on $(\mathbb{Z}/m\mathbb{Z})^\times$.

Character theory tells us that the irreducible characters of G are orthonormal with respect to the inner product $\langle f, g\rangle = \frac{1}{|G|}\sum_{x\in G} f(x)\overline{g(x)}$, so we can compute $c_{\mathbb{1}}$ as

$$\langle \delta_{m,a}, \mathbb{1}\rangle = \frac{1}{\phi(m)} \sum_{x\in(\mathbb{Z}/m\mathbb{Z})^\times} \delta_{m,a}(x).1 = \frac{1}{\phi(m)} \neq 0.$$

4 Euler products

Euler product representations are a key property of L-functions. The lemma shows that a Dirichlet L-function, defined as a sum over all integers, can also be described as a product over the primes. This can be seen as a version of the Fundamental Theorem of Arithmetic, from which it is clear that the identity holds formally. A little analysis is required to show that the identity holds numerically for $s > 1$ (where the series converges absolutely).

Lemma 4.1. *For $s > 1$ and for any Dirichlet character χ,*

$$L(\chi, s) = \prod_{p \text{ prime}} (1 - \chi(p)p^{-s})^{-1}.$$

(This includes the fact that the right hand side converges.)

Proof. For any $N \in \mathbb{N}$,

$$\prod_{p<N} (1 - \chi(p)p^{-s})^{-1} = \prod_{p<N} \sum_{r=0}^{\infty} \chi(p)^r p^{-sr}.$$

Now let A_N be the set of $n \in \mathbb{N}$ such that all prime factors of n are less than N.

By the Fundamental Theorem of Arithmetic (and using that χ is totally multiplicative), if we multiply out the product on the RHS above, then we get a term $\chi(n)n^{-s}$ exactly once for each $n \in A_N$.

Since $\sum \chi(n)n^{-s}$ converges absolutely for $s > 1$, we can reorder the terms to get

$$\prod_{p<N} \sum_{r=0}^{\infty} \chi(p^r)p^{-sr} = \sum_{n\in A_N} \chi(n)n^{-s}.$$

Every $n < N$ is in A_N, so

$$\left| \prod_{p<N} (1 - \chi(p)p^{-s})^{-1} - \sum_{n=1}^{\infty} \chi(n)n^{-s} \right| = \left| \sum_{n\notin A_N} \chi(n)n^{-s} \right| \leq \sum_{n\geq N} \left|\chi(n)n^{-s}\right|$$

and this tends to 0 as $N \to \infty$. □

5 Sums over primes and L-functions

In this section we relate $L(\chi, s)$ to $\sum_p \chi(p)p^{-s}$, using $\sum_p \log(1-\chi(p)p^{-s})$ as an intermediate step. The aim is to show that if $L(\chi, s)$ tends to a finite, non-zero limit as $s \to 1+$, then $\sum_p \chi(p)p^{-s}$ tends to a finite limit.

Note that $-\sum_p \log(1-\chi(p)p^{-s})$ looks like the expression we could expect to obtain by taking the logarithm of the Euler product for $L(\chi, s)$. Talking about $\log L(\chi, s)$ is awkward, because it is not clear that a consistent choice of branch of logarithm is possible. But exp is unambiguously defined and continuous, so we shall establish that

$$L(\chi, s) = \exp(-\sum_p \log(1-\chi(p)p^{-s})).$$

Expanding $-\sum_p \log(1-\chi(p)p^{-s})$ using Taylor series, the dominant terms give $\sum_p \chi(p)p^{-s}$. We shall show that the other terms of the Taylor series are small, and in particular that their sum is bounded as $s \to 1+$.

We shall do the second of the steps mentioned above first, i.e. relating $-\sum_p \log(1-\chi(p)p^{-s})$ to $\sum_p \chi(p)p^{-s}$. This will show that the former sum converges, so that talking about $\exp(-\sum_p \log(1-\chi(p)p^{-s}))$ makes sense.

Unless otherwise specified, log shall denote the principal branch of the logarithm throughout this section, and χ will denote any Dirichlet character.

Lemma 5.1. *If $|a_n| \leq \frac{1}{2}$ for all n, $\sum_n |a_n|^2$ converges and $\sum_n \log(1+a_n)$ converges, then $\sum_n a_n$ converges and $|\sum_n \log(1+a_n) - \sum_n a_n| \leq \sum_n |a_n|^2$.*

Proof. Since $|a_n| \leq 1$, $\log(1+a_n)$ is defined by its Taylor expansion at 1 and (using $|a_n| \leq \frac{1}{2}$ for the last inequality),

$$|\log(1+a_n) - a_n| \leq \sum_{r=2}^{\infty} \frac{|a_n|^r}{r} \leq \frac{|a_n|^2}{2} \sum_{r=0}^{\infty} |a_n|^r = \frac{|a_n|^2}{2(1-|a_n|)} \leq |a_n|^2$$

Hence $\sum |\log(1+a_n) - a_n| \leq \sum |a_n|^2$; now apply simple facts about convergence of sums. □

Corollary 5.2. *$\sum_p \log(1-\chi(p)p^{-s})$ converges uniformly on $s > 1+\delta$, for any $\delta > 0$ (and so defines a continuous function on $s > 1$).*

Proof. Note that $|\chi(p)p^{-s}| \leq \frac{1}{2}$ for all primes p, Dirichlet characters χ and $s > 1$. So we can apply Lemma 5.1 and get that

$$\left|\sum_{p>N} \log(1-\chi(p)p^{-s}) - \sum_{p>N} (-\chi(p)p^{-s})\right| \leq \sum_{p>N} \left|\chi(p)p^{-s}\right|^2$$

so that

$$\left|\sum_{p>N} \log(1-\chi(p)p^{-s})\right| \leq \left|\sum_{p>N}(-\chi(p)p^{-s})\right| + \sum_{p>N} \left|\chi(p)p^{-s}\right|^2 \leq 2\sum_{p>N} p^{-(1+\delta)}$$

which tends to 0 as $N \to \infty$ and is independent of s. □

Corollary 5.3. $\left|-\sum_p \chi(p)p^{-s} - \sum_p \log(1-\chi(p)p^{-s})\right|$ *is bounded independent of s on $s > 1$.*

Proof.

$$\left|-\sum_p \chi(p)p^{-s} - \sum_p \log(1-\chi(p)p^{-s})\right| \leq \sum_p \left|\chi(p)p^{-s}\right|^2 \leq \sum_p p^{-2} < +\infty.$$

□

Now relating this to $L(\chi, s)$ is not hard.

Lemma 5.4. *For $s > 1$, $L(\chi, s) = \exp(-\sum_p \log(1-\chi(p)p^{-s}))$.*

Proof. Since we are relating a finite sum and product, for any N,

$$\prod_{p<N}(1-\chi(p)p^{-s})^{-1} = \exp(-\sum_{p<N} \log(1-\chi(p)p^{-s})).$$

exp is continuous, so taking the limit of both sides as $N \to \infty$ gives the result. □

Now we tie together all the results of this section (valid on $s > 1$) to establish what happens as $s \to 1+$.

Lemma 5.5. *If $L(\chi, s)$ tends to a finite, non-zero limit as $s \to 1+$, then $\sum_p \log(1-\chi(p)p^{-s})$ tends to a finite limit as $s \to 1+$.*

Proof. Let the limit of $L(\chi, s)$ as $s \to 1+$ be a.

Let D be the open disc around a of radius $|a|$.

Since D is simply connected and does not contain 0, we can choose a branch of the logarithm which is continuous on D (this branch will be meant whenever we talk about $\log L(\chi, s)$ or $\log a$ for the rest of this proof).

Since $L(\chi, s) \to a$ as $s \to 1+$, there exists $\delta > 0$ such that $L(\chi, s) \in D$ for all s with $1 < s < 1 + \delta$.

Then $\log L(\chi, s)$ is a continuous function on $(1, 1 + \delta)$.

From Lemma 5.4, $-\sum_p \log(1-\chi(p)p^{-s}) = \log L(\chi, s) + 2\pi i h$ for some integer h, for each $s \in (1, 1+\delta)$. At first sight h might depend on s, but since

$\sum_p \log(1-\chi(p)p^{-s})$ and $\log L(\chi, s)$ are both continuous and h is always an integer, h must in fact be independent of s.

Since log is continuous on D, $\log L(\chi, s) \to \log a$ as $s \to 1+$.

So $\sum_p \log(1-\chi(p)p^{-s}) \to -\log a - 2\pi i h$ as $s \to 1+$. $\square$

Corollary 5.6. *If $L(\chi, s)$ tends to a finite, non-zero limit as $s \to 1+$, then $\sum_p \chi(p)p^{-s}$ tends to a finite limit as $s \to 1+$.*

Proof. Combine Lemma 5.5 and Corollary 5.3. $\square$

6 The principal Dirichlet character

This section takes the results of the last section and applies them to the principal character, to establish that $\sum_p \mathbb{1}(p)p^{-s} \to +\infty$ as $s \to 1+$.

This is much easier than dealing with non-principal characters, because we are only working with sums of positive real numbers.

It is well-known that $\sum_n n^{-1} = +\infty$. We show that $\zeta(s) = \sum_n n^{-s}$ tends to $+\infty$ as $s \to 1+$, then that $L(\mathbb{1}, s) \to +\infty$. (Note that $\zeta(s)$ is not the same as $L(\mathbb{1}, s)$, because the sum defining $L(\mathbb{1}, s)$ does not include integers n which share a factor with the modulus m.)

Lemma 6.1. *$\zeta(s) \to +\infty$ as $s \to 1+$.*

Proof. It is well-known that $\zeta(1) = \sum_n n^{-1} = +\infty$.

So given $K > 0$, there exists N such that $\sum_{n=1}^N \frac{1}{n} > K$.

Now $\sum_{n=1}^N n^{-s}$ is continuous at $s = 1$ (since it is a finite sum of continuous functions).

So there exists $\delta > 0$ such that $s \in (1, 1+\delta) \implies \sum_{n=1}^N n^{-s} > K - 1$.

But then $s \in (1, 1+\delta) \implies \zeta(s) > K - 1$. $\square$

Lemma 6.2. *$L(\mathbb{1}, s) \to +\infty$ as $s \to 1+$.*

Proof. $\zeta(s)$ is the L-function of the constant function $\mathbb{Z} \to \mathbb{C}$ with value 1, so Lemma 4.1 gives that for $s > 1$,

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

and $\prod_{p|m}(1-p^{-s})^{-1}$ (which is a finite product of continuous functions) is continuous at 1. $\square$

Corollary 6.3. *$\sum_p \mathbb{1}(p)p^{-s} \to +\infty$ as $s \to 1+$.*

Proof. Lemma 5.4 gives that $L(\mathbb{1}, s) = \exp(-\sum_p \log(1 - \mathbb{1}(p)p^{-s}))$ for $s > 1$.

Since we are working only with positive real numbers, there is a canonical choice of logarithm and we can say that $\log L(\mathbb{1}, s) = -\sum_p \log(1 - \mathbb{1}(p)p^{-s})$.

From Lemma 6.2, we get that $\log L(\mathbb{1}, s) \to +\infty$ as $s \to 1+$.

So $-\sum_p \log(1 - \mathbb{1}(p)p^{-s}) \to +\infty$ as $s \to 1+$.

Then apply Corollary 5.3, that $\left| -\sum_p \mathbb{1}(p)p^{-s} - \sum_p \log(1 - \mathbb{1}(p)p^{-s}) \right|$ is bounded as $s \to 1+$. □

7 Analyticity of L-functions

At this point we have done all the work relating $\sum_{p \equiv a} p^{-1}$ to L-functions, and all that remains is to prove that for non-principal Dirichlet characters, $L(\chi, s)$ tends to a finite, non-zero limit as $s \to 1+$.

Let χ be a non-principal Dirichlet character of modulus m. In this section, we shall show that $\lim_{s \to 1+} L(\chi, s)$ is defined and finite, by showing that the series $L(\chi, s)$ converges and defines a continuous function on $s > 0$ (so $\lim_{s \to 1+} L(\chi, s) = L(\chi, 1)$).

Establishing the above only requires working with real s, but exactly the same proof will show that the series converges for all complex s with $\Re s > 0$, and defines an analytic function. This will be needed when we come to show that $L(\chi, 1) \neq 0$.

We rely on the following convergence test.

Lemma 7.1. *Let a_n be a sequence of complex numbers.*

Suppose that $\left| \sum_{n=1}^{N} a_n \right| \leq M$ for all $N \in \mathbb{N}$.

Let $s = \sigma + it$ with $\sigma > 0$.

Then $\sum_n a_n n^{-s}$ converges, and $\left| \sum_{n=N+1}^{\infty} a_n n^{-s} \right| \leq M N^{-\sigma} \left(\frac{|s|}{\sigma} + 1 \right)$.

We shall prove this by comparing with an integral.

Lemma 7.2. *Let a_n be a complex sequence and $A_x = \sum_{n \leq x} a_n$ for any $x \in \mathbb{R}$. Then $\sum_{n=1}^{N} a_n n^{-s} = A_N N^{-s} + s \int_1^N A_x x^{-s-1} dx$.*

If $s \neq 0$ and $A_N N^{-s} \to 0$ as $N \to \infty$, then

$$\sum_n a_n n^{-s} \text{ converges} \Leftrightarrow s \int_1^{\infty} A_x x^{-s-1} dx \text{ converges}$$

and both have the same value.

If these converge then for $N \in \mathbb{N}$,

$$\sum_{n=N+1}^{\infty} a_n n^{-s} = -A_N N^{-s} + s \int_N^{\infty} A_x x^{-s-1} dx.$$

Proof. Some algebraic manipulation gives that

$$\sum_{n=1}^{N} a_n n^{-s} = A_N N^{-s} - \sum_{n=1}^{N-1} A_n((n+1)^{-s} - n^{-s})$$

(this is the discrete version of integration by parts).

Now $\int_n^{n+1} A_x x^{-s-1} dx = -\frac{1}{s} A_n((n+1)^{-s} - n^{-s})$, so we get the first part of the lemma.

The rest follows immediately. □

Lemma 7.3. *Let f be a measurable function $\mathbb{R} \to \mathbb{C}$.*

Suppose there exists M such that for all $x \geq 1$, $|f(x)| \leq M$.

Suppose that $s = \sigma + it$ has $\sigma > 0$.

Then $\int_1^\infty f(x)x^{-s-1}dx$ converges and for $N > 1$,

$$\left| \int_N^\infty f(x)x^{-s-1}dx \right| \leq \frac{M}{\sigma N^\sigma}.$$

Proof. We have that $|f(x)x^{-s-1}| \leq Mx^{-\sigma-1}$ for all $x \geq 1$, so

$$\left| \int_N^\infty f(x)x^{-s-1}dx \right| \leq M \int_N^\infty x^{-s-1}dx.$$

Since $-\sigma - 1 < -1$, $\int_N^\infty x^{-\sigma-1}dx = \frac{1}{\sigma N^\sigma}$ which tends to 0 as $N \to \infty$. □

These two lemmas combine (taking $f(x) = A_x$) to give Lemma 7.1.

We can use this convergence test to prove:

Lemma 7.4. *Let a_n be a sequence of complex numbers.*

Suppose that $\left|\sum_{n=1}^N a_n\right| \leq M$ for all $N \in \mathbb{N}$.

Then $\sum_n a_n n^{-s}$ converges locally uniformly on $\Re s > 0$ (and so defines a holomorphic function on the same domain).

Proof. We shall show that $\sum_n a_n n^{-s}$ converges uniformly on the rectangle (extending infinitely to the right) $U = \{\sigma + it \mid \sigma \geq \delta, |t| \leq R\}$ for any $\delta, R > 0$.

By Lemma 7.1, we have that for $s \in U$,

$$\left| \sum_{n=N+1}^{\infty} a_n n^{-s} \right| \leq \frac{M}{N^\sigma}\left(\frac{|s|}{\sigma} + 1\right) \leq \frac{M}{N^\delta}\left(\frac{\sigma + |t|}{\sigma} + 1\right) \leq \frac{M}{N^\delta}\left(\frac{R}{\delta} + 2\right),$$

which tends to 0 as $N \to \infty$ and is independent of s. □

Corollary 7.5. *For a non-principal Dirichlet character* χ*,* $L(\chi, s)$ *is analytic on* $\Re s > 0$*.*

Proof. In order to apply the lemma, we just need to know that $\sum_{n=1}^{N} \chi(n)$ is bounded (this is the step that fails for $\chi = \mathbb{1}$, for which of course $\sum_{n=1}^{N} \chi(n)$ is essentially proportional to N).

But the non-zero values of $\chi(n)$ on $n = 1, \ldots, m$ are the d-th roots of unity, repeated $\phi(m)/d$ times, for some $d \in \mathbb{N}$ dividing $\phi(m)$.

So $\sum_{n=1}^{m} \chi(n) = 0$ and $\left|\sum_{n=1}^{N} \chi(n)\right| \leq m$ for $1 \leq N \leq m$.

Since χ is periodic with period m, in fact $\left|\sum_{n=1}^{N} \chi(n)\right| \leq m$ for all N. □

8 Meromorphic continuation of the Riemann zeta function

At this point, all that is left is to show that $L(\chi, 1) \neq 0$ for non-principal Dirichlet characters χ. This turns out to be surprisingly difficult, and is the only part of the proof where we will need to use complex analysis or consider $L(\chi, s)$ for non-real s.

As a preliminary, we need to know that $L(\mathbb{1}, s)$ extends to a meromorphic function on $\Re s > 0$, with a simple pole at $s = 1$. We prove this first for the Riemann zeta function $\zeta(s) = \sum_n n^{-s}$, then use the fact that the Euler product for $L(\mathbb{1}, s)$ differs from that for $\zeta(s)$ by only finitely many factors.

Lemma 8.1. $\zeta(s)$ *has an analytic continuation to* $\{s \mid \Re s > 0$ *and* $s \neq 1\}$*.*

Proof. In fact ζ has analytic continuation to $\mathbb{C} \setminus \{1\}$, but we shall not need that here. There are many ways of proving this; here we shall use a very simple method.

First note that the series $\sum_n n^{-s}$ converges absolutely and locally uniformly for all complex s with $\Re s > 1$, so defines a holomorphic function there.

Now consider $\eta(s) = \sum_n (-1)^n n^{-s}$. By Lemma 7.4, this converges locally uniformly for $\Re s > 0$ and defines a holomorphic function there.

We have

$$\begin{aligned}\eta(s) &= \sum_{n \text{ odd}} n^{-s} - \sum_{n \text{ even}} n^{-s} \\ &= \zeta(s) - 2\sum_{n \text{ even}} n^{-s} \\ &= \zeta(s) - 2\sum_{n=1}^{\infty}(2n)^{-s} \\ &= \zeta(s) - 2^{1-s}\zeta(s).\end{aligned}$$

This is valid on $\Re s > 1$ since everything converges absolutely there.

Hence $\frac{\eta(s)}{1-2^{1-s}}$ defines a meromorphic continuation of $\zeta(s)$ to all s with $\Re s > 0$, and is analytic whenever $1 - 2^{1-s} \neq 0$ i.e. $s \neq 1$. □

Lemma 8.2. *$\zeta(s)$ has a simple pole at $s = 1$.*

Proof. We know that $\eta(s)$ is holomorphic at 1, so it is sufficient to show that $\frac{s-1}{1-2^{1-s}}$ has a removable singularity at $s = 1$.

By l'Hôpital's rule,

$$\lim_{s\to 1} \frac{s-1}{1-2^{1-s}} = \lim_{s\to 1} \frac{1}{\log 2.2^{1-s}} = \frac{1}{\log 2}$$

and this is finite. □

Corollary 8.3. *$L(\mathbb{1}, s)$ has an analytic continuation to $\{s \mid \Re s > 0$ and $s \neq 1\}$, with a simple pole at $s = 1$.*

Proof. The Euler products for $\zeta(s)$ and $L(\mathbb{1}, s)$ are valid not just for real values of $s > 1$, but for complex s with $\Re s > 1$ (the proof given for Lemma 4.1) still works).

So for $\Re s > 1$,

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

Now $\prod_{p|m}(1 - p^{-s})$ is holomorphic on all of $\mathbb{C}$, so $\zeta(s)\prod_{p|m}(1 - p^{-s})$ defines an analytic continuation of $L(\mathbb{1}, s)$ to $\{s \mid \Re s > 0$ and $s \neq 1\}$, with a simple pole at $s = 1$. □

9 $L(\chi, 1)$ is not zero: non-real characters

Finally we show that $L(\chi, 1) \neq 0$, where χ is a non-principal Dirichlet character mod m. We already know that $L(\chi, s)$ defines a holomorphic function on $\Re s > 0$.

This falls into two cases. This section deals with the easier case, in which $\chi(n)$ does not take only real values.

In this case, $\bar{\chi}$ (the complex conjugate of χ) is a Dirichlet character distinct from χ, and if $L(\chi, 1) = 0$ then also $L(\bar{\chi}, 1) = 0$. We show that this implies that $F(s) = \prod_\chi L(\chi, s)$ vanishes at $s = 1$.

However by writing $F(s)$ as an Euler product, we can show that $F(1) \neq 0$.

Lemma 9.1. *$F(s) = \prod_\chi L(\chi, s)$ does not have a zero at $s = 1$ (the product is over all Dirichlet characters mod m, including the principal character).*

Proof. $F(s)$ is a product of finitely many meromorphic functions on $\Re s > 0$, so is itself meromorphic on this domain, and is holomorphic except perhaps at $s = 1$.

We shall show that $F(s) > 1$ for all real $s > 1$, so that $F(s) \not\to 0$ as $s \to 1$. As F is meromorphic, it follows that F does not vanish at $s = 1$.

Since everything converges absolutely on $s > 1$ (and we are taking a product over only finitely many characters χ), we can reorder the product and get

$$F(s) = \prod_\chi \prod_p (1 - \chi(p)p^{-s})^{-1} = \prod_p \prod_\chi (1 - \chi(p)p^{-s})^{-1}.$$

For any fixed prime p not dividing m, the values of $\chi(p)$ are the d-th roots of unity, repeated r times, for some d, r such that $dr = \phi(m)$.

Hence $\prod_\chi (1 - \chi(p)X) = (1 - X^d)^r$ as an identity of polynomials, and so $\prod_\chi (1 - \chi(p)p^{-s})^{-1} = (1 - p^{-sd})^{-r} > 1$.

For a prime p dividing m, of course $\prod_\chi (1 - \chi(p)X) = 1$.

So $F(s) > 1$ for all real $s > 1$, and so $F(1) \neq 0$. □

Lemma 9.2. *If χ takes non-real values, then $L(\chi, 1) \neq 0$.*

Proof. $\bar{\chi}$ (the complex conjugate of χ) is also a Dirichlet character mod m, and since χ takes non-real values, $\bar{\chi} \neq \chi$.

If $L(\chi, 1) = 0$ then also $L(\bar{\chi}, 1) = 0$.

So in the product $F(1) = \prod_\chi L(\chi, 1)$, there are at least two functions with zeros, while only $L(\mathbb{1}, s)$ has a pole at $s = 1$ and it has a simple pole.

So $F(1) = 0$ contradicting the previous lemma. □

Note that complex analysis was only used in this proof to say that the order of a pole or a zero makes sense.

If you wanted to use only real analysis, instead of saying $L(\chi, s)$ holomorphic, it is sufficient to know that $L(\chi, s)$ has a continuous derivative at $s = 1$ (for χ non-principal) and that $(s-1)L(\mathbb{1}, s)$ has a finite limit at $s = 1$.

You can then use l'Hôpital's rule and the mean value theorem to show that if $L(\chi, 1) = 0$ for more than one character χ then $F(s) \to 0$ as $s \to 1$.

10 $L(\chi, 1)$ is not zero: real characters

This section deals with the harder case of showing that $L(\chi, 1) \neq 0$: the case in which $\chi(n)$ is real for all n. This is the hardest part of the whole proof.

There are several ways of proving this, none straightforward. The method I have used here is taken from Prof Körner's Topics in Fourier Analysis notes. Dirichlet proved the result for prime modulus m by using the Poisson summation formula, which requires more computation although perhaps less magic than this proof. For composite m, he relied on a more arithmetic approach, relating $L(\chi, 1)$ for a real character χ to the number of quadratic forms with a certain discriminant.

It is based on the surprising function

$$\psi(s) = \frac{L(\chi, s)L(\mathbb{1}, s)}{L(\mathbb{1}, 2s)}.$$

This function is chosen because it has a simple Euler product expansion, and has useful values at $s = 1$ and $s = \frac{1}{2}$.

We show that $\psi(s)$ is analytic on $\Re s > \frac{1}{2}$, except perhaps at $s = 1$, and also at $s = \frac{1}{2}$; indeed $\psi(\frac{1}{2}) = 0$.

At $s = 1$, $\psi(s)$ has a removable singularity if $L(\chi, 1) = 0$ and a simple pole otherwise. This tells us that if $L(\chi, 1) = 0$, then the Taylor series for ψ about some $s_0 > 1$ has radius of convergence greater than $s_0 - \frac{1}{2}$. In particular the Taylor series converges at $s = \frac{1}{2}$, with value $\psi(\frac{1}{2})$.

In order to estimate the terms of this Taylor series, we write ψ first as an Euler product, then as an L-series. This allows us to compute the derivatives of ψ, and show that $(-1)^m \psi^{(m)}(s) > 0$ for real $s > 1$. It follows that if the Taylor series converges at $s = \frac{1}{2}$, it must have a positive value there, contradicting $\psi(\frac{1}{2}) = 0$.

Lemma 10.1. *$\psi(s)$ is analytic on $\{s \mid \Re s > \frac{1}{2}\} \setminus \{1\} \cup \{\frac{1}{2}\}$, and $\psi(\frac{1}{2}) = 0$.*

Proof. Note that $\psi(s)$ is certainly meromorphic on $\Re s > 0$.

By Lemma 5.4, $L(\mathbb{1}, s) = \exp(-\sum_p \log(1 - \mathbb{1}(p)p^{-s}))$ for $\Re s > 1$. (We only stated this lemma for real values of s, but the proof still works for complex s with $\Re s > 1$.) Hence $L(\mathbb{1}, s) \neq 0$ for $\Re s > 1$.

So the denominator $L(\mathbb{1}, 2s)$ of ψ is non-zero for $\Re s > \frac{1}{2}$.

The numerator is analytic on $\{s \mid \Re s > \frac{1}{2}\} \setminus \{1\}$, so ψ is analytic on $\Re s > \frac{1}{2}$ except perhaps at $s = 1$.

The numerator is holomorphic at $s = \frac{1}{2}$ and $L(\mathbb{1}, 2s)$ has a pole there, so ψ is holomorphic at $s = \frac{1}{2}$ with $\psi(\frac{1}{2}) = 0$. □

Lemma 10.2. *There are nonnegative integers c_n such that $\psi(s) = \sum_n c_n n^{-s}$ for $\Re s > 1$.*

Proof. First we compute the Euler product for ψ.

At a prime p not dividing m, the Euler factor is

$$\frac{(1 - p^{-2s})}{(1 - \chi(p)p^{-s})(1 - p^{-s})} = \frac{1 + p^{-s}}{1 - \chi(p)p^{-s}} = \begin{cases} 1 & \text{if } \chi(p) = -1, \\ \frac{1+p^{-s}}{1-p^{-s}} & \text{if } \chi(p) = +1. \end{cases}$$

(Because χ is real and $\chi(p)$ must be a root of unity, $\chi(p) = \pm 1$ are the only possibilities.)

Hence for $\Re s > 1$,

$$\psi(s) = \prod_{\chi(p)=1} \frac{1 + p^{-s}}{1 - p^{-s}}.$$

By the same arguments as for Lemma 4.1, we get that for $\Re s > 1$,

$$\begin{aligned} \prod_{\chi(p)=1} (1 + p^{-s}) &= \sum_{n \in S} n^{-s}, \\ \prod_{\chi(p)=1} (1 - p^{-s})^{-1} &= \sum_{n \in T} n^{-s}, \end{aligned}$$

where $S = \{n \in \mathbb{N} \mid \chi(p) = 1 \text{ for all prime factors of } n \text{ and } n \text{ is square-free}\}$ and $T = \{n \in \mathbb{N} \mid \chi(p) = 1 \text{ for all prime factors of } n\}$.

Now $(\sum_n a_n n^{-s})(\sum_n b_n n^{-s}) = \sum_n c_n n^{-s}$ where $c_n = \sum_{d|n} a_d b_{n/d}$. (The reordering required to prove this is valid for real $s > 1$ because we are summing positive real numbers; it follows that $\sum c_n n^{-s}$ converges absolutely for $\Re s > 1$, and so the reordering is valid for all such s.)

Taking a_n, b_n to be 0 or 1 depending on whether or not n is in S, T, we get that $\psi(s) = \sum_n c_n n^{-s}$ for $\Re s > 1$ for nonnegative integers c_n. □

Lemma 10.3. *If $s > 1$ is real, then $(-1)^m \psi^{(m)}(s) > 0$.*

Proof. Differentiating each term of the sum $\psi(s) = \sum_n c_n n^{-s}$ gives

$$\frac{d}{ds} c_n n^{-s} = c_n(-\log n) n^{-s}.$$

Given $\delta > 0$, for large enough n, $(\log n) n^{-(1+\delta)} < n^{-(1+\frac{1}{2}\delta)}$.

But $\sum_n c_n n^{-(1+\frac{1}{2}\delta)}$ converges, so $\sum_n c_n (\log n) n^{-(1+\delta)}$ converges and the partial sums are bounded.

Then by Lemma 7.4 (taking $a_n = c_n(-\log n) n^{-(1+\delta)}$ as the sequence in the lemma), $\sum_n c_n(-\log n) n^{-s}$ converges locally uniformly for $\Re s > 1 + \delta$.

Hence term-by-term differentiation is valid for $\Re s > 1$, and $\psi'(s) = \sum_n c_n(-\log n) n^{-s}$.

Similarly, we get that $\psi^{(m)}(s) = \sum_n c_n(-\log n)^m n^{-s}$ for any $m \in \mathbb{N}$.

So as the c_n are nonnegative (and not all zero), for real $s > 1$ we get that

$$(-1)^m \psi^{(m)}(s) = \sum_n c_n (\log n)^m n^{-s} > 0. \qquad \square$$

Lemma 10.4. *If $L(\chi, s)$ takes only real values, then $L(\chi, 1) \neq 0$.*

Proof. Assume for contradiction that $L(\chi, 1) = 0$.

Then $\psi(s) = L(\chi, s) L(\mathbb{1}, s)/L(\mathbb{1}, 2s)$ has a removable singularity at $s = 1$ (since $L(\mathbb{1}, s)$ has a simple pole and $L(\mathbb{1}, 2s)$ is non-zero).

Fix a real value $s_0 > 1$, and consider the Taylor series for ψ about s_0.

Since ψ has a removable singularity at $s = 1$, it is holomorphic on the closed disc centred at s_0 with radius $s_0 - \frac{1}{2}$. Standard complex analysis tells us that the Taylor series converges on this closed disc, with limit $\psi(s)$.

In particular, at $s = \frac{1}{2}$, we get that

$$\psi(\tfrac{1}{2}) = \sum_m \frac{\psi^{(m)}(s_0)}{m!} (\tfrac{1}{2} - s_0)^m.$$

$\frac{1}{2} - s_0$ is negative, so by Lemma 10.3 each term of this sum is positive.

But this contradicts the fact that $\psi(\frac{1}{2}) = 0$ (Lemma 10.1). $\square$