university of groningen

faculty of science and engineering

mathematics and applied mathematics

The Proof of Dirichlet's Theorem on Arithmetic Progressions and its Variations

Bachelor's Project Mathematics

13th of July 2018

Student: M. W. Mintjes

First supervisor: Prof.dr. J. Top

Second assessor: Dr. A. E. Sterk

Abstract

In this bachelor's thesis we study the proof of Dirichlet's Theorem on primes in Arithmetic Progressions. We find out about the various proofs of this theorem and see how they are different, but also very related. We start looking at the idea of the proof that Dirichlet originally came up with it and end up at very recent variations, published in the American Mathematical Monthly. The influences from time are visible in the mathematics, but the general idea of the proof stays intact. We can construct different proofs for different audiences and find one for readers from different backgrounds to enjoy and understand. From classrooms to devoted readers of mathematics, there is something interesting about the proof of Dirichlet's Theorem for all mathematicians.

Contents

1 Introduction

The aim of this bachelor project is to compare and combine different proofs for Dirichlet's Theorem on primes in arithmetic progressions. The theorem, and often a proof, can be found in most textbooks on Analytic Number Theory. We will start by studying the proof as given by Jean-Pierre Serre in *A Course in Arithmetic* [5]. This is one of the more classic proofs that explains all concepts and steps in detail. It relies heavily on the fields of Group Theory, Complex Analysis and Real Analysis. After understanding this proof, we will move on to study the hardest step and we discuss how various authors prove this step. We will find that different proofs rely on different fields of study in mathematics more heavily. Finding our way through multiple proofs, we form our own idea on what is important and simple to understand. We compare variations of the proof in the final chapter and recommend readers which proof might best suit their interests. This could be expanded on in other projects, as there are many more proofs of Dirichlet's Theorem and we only focused on proving one step of the proof in different ways. An even broader manual could be created to provide a better tailored proof to different audiences.

Let us study Dirichlet's Theorem now. The theorem states the following:

Theorem 1. Let a and m be positive integers with $gcd(a, m) = 1$. There exist infinitely many prime numbers p such that $p \equiv a \pmod{m}$.

Equivalently, we can say that there exist infinitely many primes in the so called arithmetic progression a, $a + m$, $a + 2m$, $\ldots$ for two positive coprime integers a and m. The arithmetic progression is the collection of integers $a \pmod{m}$. This is why the theorem is known as Dirichlet's Theorem on Arithmetic Progressions, shortened to Dirichlet's Theorem.

To prove the theorem, Dirichlet used the series

$$\sum_{p \equiv a \pmod{m}} \frac{1}{p^s}, \tag{1.1}$$

where p is a prime number and $s \in \mathbb{C}$ has a real part bigger than 1. He showed that this series diverges for $s = 1$. This is of course only possible if the sum in (1.1) consists of infinitely many terms, hence this divergence implies Dirichlet's Theorem. To prove this divergence, Dirichlet split the series up into two separate series, which were more convenient to work with. Of these two, he showed one to be divergent and one to be convergent. The combination of these two series was then thus shown to be divergent. As we cannot say anything about the combination of two divergent series, proving the one series to be convergent is crucial. To prove this, Dirichlet had to come up with new functions and series. These are still used to prove various results in Number Theory. This original way of working is why this theorem and its proofs are still relevant and interesting to us. The study of Dirichlet's Theorem and its proofs throughout the years teaches us about mathematical styles in different times. The latest variation on a part of the proof we will study was published in the American Mathematical Monthly as recent as 2017 [6]. But even this relies on the same ideas that Dirichlet came up with. To understand the proofs, we first study how Dirichlet himself came up with the mathematics behind them. That is what we will do now in the next chapter.

2 Preliminaries

2.1 History and characteristics

History of the proof

Prime numbers have very useful applications. Most of these are in the field of cryptography.

However, the investigations on primes started far before these applications, according to [1]: Already around 300 BC, Euclid wrote about prime numbers in his *Elements*. He provided the first proof that there are infinitely many prime numbers. In *Elements*, Euclid also laid the foundation for what is now known as the Fundamental Theorem of Arithmetic. This theorem states that any integer greater than 1 either is a prime or can be represented as a product of primes in a unique way.

Around 2000 years later, Euler came up with his proof for the infinitude of primes. This proof relies on the divergence of

$$\sum_{p \in \mathcal{P}} \frac{1}{p^s} \tag{2.1}$$

at $s = 1$, where $\mathcal{P}$ is the set of all prime numbers. He introduced the idea of the method that Dirichlet would use later to prove Dirichlet's theorem using the series (1.1). Euler's series (2.1) was also defined in such a way that divergence would require an infinite number of terms. In Euler's method of showing said divergence, he introduces the *zeta function* $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$. We will see later that this function also plays a key role in Dirichlet's proof. Euler also used the Fundamental Theorem of Arithmetic in his proof to connect a series over prime numbers to a series over all natural numbers. This way, he could connect the series (2.1) to the zeta function. The zeta function diverges as $s \to 1$ and this is how Euler proved the required divergence. This is not the whole proof of course, but a few highlights of how Euler proved that the series (2.1) diverges when $s = 1$.

Euler proved that there is an infinitude of primes, but he also showed something about the distribution of them. For example, the sum $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is known to converge, opposed to the series (2.1) that diverges as $s \to 1$. From this, it can be concluded that there are more prime numbers than there are squares. The knowledge of distribution of primes ties in with how Dirichlet's theorem was proven later as he connected the set of primes in arithmetic progressions to the set of all primes. Dirichlet proved that there are infinitely many primes in arithmetic progressions. For this, he showed the divergence of $\sum_{p \equiv a \pmod{m}} \frac{1}{p^s}$ as $s \to 1$, where $p \equiv a \pmod{m}$ is the collection of all prime numbers p congruent to a modulo m. To show this, he looked at the ratio between this series over primes modulo a and the series over all primes.

Approaching the proof

Showing that the series

$$\sum_{p \equiv a \pmod{m}} \frac{1}{p^s} \tag{2.2}$$

diverges as $s \to 1$, is not an easy task. This is a hard problem to approach, harder than a sum over all primes. It is convenient to rewrite (2.2) as

$$\sum_{p \in \mathcal{P}} \frac{f(p)}{p^s}, \tag{2.3}$$

where the domain of f is $\mathbb{N}$. We define f so that

$$f(p) = \begin{cases} 1 \text{ when } p \equiv a \pmod{m} \\ 0 \text{ otherwise.} \end{cases}$$

The rewritten series is very closely related to the series that Euler used in his proof that there are infinitely many primes. It is therefore appealing to express the series (2.3) in terms of the series (2.1), of which we know it diverges. For this, the function $f(p)$ first needs to be rewritten as a combination of so called multiplicative functions:

Definition 1. A function $g : \mathbb{N} \to \mathbb{C}$ is called **multiplicative** if $g(1) = 1$ and it has the property that $g(ab) = g(a)g(b)$ for all $a, b \in \mathbb{N}$.

The function $f(p)$ does not have this property. It is possible, though, to write $f(p)$ as a combination of multiplicative functions from the natural numbers to the complex plane. Using this, we rewrite the series (2.3) as a sum of two separate series that only contain multiplicative functions. Then, we show that one of these series is divergent and the other is convergent. The combination of these series, must then be divergent. This shows that the series (2.3) is divergent, and so the series (2.2) must diverge, which concludes the proof.

This rewriting makes the mathematics a lot more complicated. Finding the functions and writing $f(p)$ as a combination of them is the first part. After this, we have to show that one of our series diverges and the other converges. Showing this convergence especially is a very difficult step which we do not show here yet. It is also a step that mathematicians have not been able to avoid. The fundamental idea of modern proofs of this theorem is still the same as has been introduced by Dirichlet. This is why Dirichlet's Theorem on Arithmetic Progressions is intriguing, all proofs follow the same general line but are elegant in their own way. To gain more of an understanding, we study an example of the proof for a specific case.

2.2 An example

Example 1. Let $a = 3$ and let $m = 8$. Dirichlet's theorem states that there are infinitely many prime numbers in the arithmetic progression $3 \pmod 8 = \{3, 11, 19, 27, 35, 43, \dots\}$. To prove this, we use the series

$$\sum_{p \equiv 3 \pmod 8} \frac{1}{p^s},$$

for p prime and $s \in \mathbb{C}$ such that the real part of s is bigger than 1. In the end of the proof, we will take the limit as $s \to 1$.

As has been noted before, the restrictions on p for this series are a problem. We rewrite as

$$\sum_{p \in \mathcal{P}} \frac{f(p)}{p^s},$$

where

$$f(p) = \begin{cases} 1 \text{ when } p \equiv 3 \pmod 8 \\ 0 \text{ otherwise.} \end{cases}$$

Again, the set of all prime numbers is denoted $\mathcal{P}$. To use the divergence of $\sum_{p \in \mathcal{P}} 1/p$, we first write $f(p)$ as a combination of multiplicative functions from $\mathbb{N}$ to $\mathbb{C}$. To find these functions, we look at the domain that they must have. These functions should be able to take any prime number as an input value, as the set of prime numbers is the set we sum over. We know that all primes except for 2 are odd. If we divide an odd number by 8, the remainder will also be odd. All primes, except for 2, are thus of the form $p \equiv 1 \pmod 8 \equiv \overline{1}$, $p \equiv \overline{3}$, $p \equiv \overline{5}$ or $p \equiv \overline{7}$. Here we introduce a new bit of notation. The overline implies that we are considering the element modulo 8 in this example.

This narrows down our functions nicely to only give a non-zero output for these four input values. This does mean that we leave the prime number 2 out of consideration, but that is not a problem. The set of all prime numbers except for 2 is still an infinite set of primes as we take away merely one element. More than this infinitude is not necessary for the argument later on and so we set $f(2) = 0$.

We look for functions that take these four elements to the real or complex plane, as we want to end up with a value of 1 or 0 in the end. Let

$$\{g : \mathbb{N} \to \mathbb{C}\}$$

denote all multiplicative functions g from the set of natural numbers to the complex plane. This is a vector space over the complex plane with infinite dimension. A linear subspace of this is the set of multiplicative functions g that are periodic with period 8.

$$\{g : \mathbb{N} \to \mathbb{C} \mid g \text{ has period } 8\} = \{\psi \circ \pi \mid \psi : \mathbb{Z}/8\mathbb{Z} \to \mathbb{C}\}$$

Here π denotes the canonical projection $\pi : \mathbb{N} \to \mathbb{Z}/8\mathbb{Z}$ that is surjective. This linear subspace has dimension 8. Even more specific is the linear subspace of this that only takes the units modulo 8 as inputs:

$$\{g : \mathbb{N} \to \mathbb{C} \mid g \text{ has period 8 and } g(n) = 0 \text{ if } gcd(n, 8) \neq 1\}$$

These functions only map elements in $(\mathbb{Z}/8\mathbb{Z})^*$ to non-negative values. This subspace has dimension 4 as there are four natural numbers $n < 8$ such that $gcd(n, 8) \neq 1$: the numbers $1, 3, 5$ and 7.

A basis of this linear subspace can be defined using four functions. We use the functions χ_0, χ_1, χ_2 and χ_3 as a basis:

	$\overline{1}$	$\overline{3}$	$\overline{5}$	$\overline{7}$
χ_0	1	1	1	1
χ_1	1	-1	-1	1
χ_2	1	-1	1	-1
χ_3	1	1	-1	-1

The functions in the first column take the input values in the first row and map them to the corresponding entry, which is either 1 or -1. These functions map any other element of the natural numbers to 0. This defines four multiplicative functions. We will now construct our function f as a combination of these functions, by a clever use of signs:

$$f(p) = \frac{1}{4}(\chi_0(p) - \chi_1(p) - \chi_2(p) + \chi_3(p)),$$

as

$$\begin{aligned}
f(\overline{1}) &= \frac{1}{4}(\chi_0(\overline{1}) - \chi_1(\overline{1}) - \chi_2(\overline{1}) + \chi_3(\overline{1})) = \frac{1}{4}(1 - 1 - 1 + 1) = 0, \\
f(\overline{3}) &= \frac{1}{4}(\chi_0(\overline{3}) - \chi_1(\overline{3}) - \chi_2(\overline{3}) + \chi_3(\overline{3})) = \frac{1}{4}(1 + 1 + 1 + 1) = 1, \\
f(\overline{5}) &= \frac{1}{4}(\chi_0(\overline{5}) - \chi_1(\overline{5}) - \chi_2(\overline{5}) + \chi_3(\overline{5})) = \frac{1}{4}(1 + 1 - 1 - 1) = 0, \\
f(\overline{7}) &= \frac{1}{4}(\chi_0(\overline{7}) - \chi_1(\overline{7}) - \chi_2(\overline{7}) + \chi_3(\overline{7})) = \frac{1}{4}(1 - 1 + 1 - 1) = 0.
\end{aligned}$$

For any prime p, this gives us indeed the following result for $f(p)$:

$$f(p) = \begin{cases} 1 \text{ when } p \equiv 3 \pmod 8 \\ 0 \text{ otherwise.} \end{cases}$$

It is interesting to note here that different combinations of signs would give us the functions for 1, 5 and 7 modulo 8. It is only this that influences where we get a 1 and where we get a 0.

Now that we have rewritten f as a combination of multiplicative functions, we can rewrite our original series. The new series is the one for which we want to show divergence:

$$\begin{aligned}
\sum_{p \equiv 3 \pmod 8} \frac{1}{p^s} &= \sum_{p \in \mathcal{P}} \frac{f(p)}{p^s} \\
&= \sum_{p \in \mathcal{P}} \frac{\frac{1}{4}(\chi_0(p) - \chi_1(p) - \chi_2(p) + \chi_3(p))}{p^s} \\
&= \frac{1}{4} \sum_{p \in \mathcal{P}} \frac{\chi_0(p) - \chi_1(p) - \chi_2(p) + \chi_3(p)}{p^s}
\end{aligned}$$

This series can be split up into two parts. We separate χ_0, which will have a value of 1 for every input in the series:

$$\begin{aligned}\sum_{p\equiv 3 \pmod 8} \frac{1}{p^s} &= \frac{1}{4}\left(\sum_{p\in\mathcal{P}} \frac{\chi_0(p)}{p^s} + \sum_{p\in\mathcal{P}} \frac{-\chi_1(p)-\chi_2(p)+\chi_3(p)}{p^s}\right) \\ &= \frac{1}{4}\left(\sum_{p\in\mathcal{P}} \frac{1}{p^s} + \sum_{p\in\mathcal{P}} \frac{-\chi_1(p)-\chi_2(p)+\chi_3(p)}{p^s}\right)\end{aligned}$$

Looking at the first part of the right-hand side, we recognize this as almost exactly the same as the series that Euler had shown to be divergent when $s = 1$. There are only two differences. The first one is the factor $1/4$, but this does not influence divergence. The second one is that we have a series over all prime numbers, except for 2. In Euler's case, 2 was included as he summed over all primes. This does not change our divergence though. Even without 2, we have an infinite set to sum over and the series does still diverge for $s = 1$.

Analysis teaches us that we cannot say anything about the sum of two divergent series. They could diverge in completely different directions and this would leave us wondering what happens when we add them up. When we add a divergent series and a convergent series however, we get a divergent series. So we want to show that the second part of our right-hand side,

$$\sum_{p\in\mathcal{P}\setminus\{2\}} \frac{-\chi_1(p)-\chi_2(p)+\chi_3(p)}{p^s},$$

converges. Then we have shown that our sum over all prime numbers congruent to 3 modulo 8 is divergent, because it is the sum of a divergent and a convergent series. This is exactly what we want to achieve, because this series being divergent proves that there are infinitely many prime numbers in this specific arithmetic progression.

This last step is a lot more involved. Even for our specific example, we cannot give the proof for this in short. Therefore we will leave this out here and we will look into it later in this paper. The goal of this example was to gain an understanding of the outline of the proof of Dirichlet's theorem. Although different mathematicians have made their own versions of the proof, a lot of the ideas used come back in almost all the proofs. Now that we have an idea of the form, we can dive further into the proof. Firstly, we look at the general proof. Then we will look into the required convergence in the general case, the most complicated step that we have left out for now, and we will see how different writers approach this same problem in the chapters to come.

3 A classical proof

Throughout this whole thesis Definitions, Lemmas, Propositions and Theorems are numbered separately. The numbering is not reset at any point and transfers over chapters. The beginnings of proofs for Lemmas, Propositions and Theorems are recognized by the italic "*Proof.*" and the conclusion of a proof is recognized by the symbol "☼". Many proofs given here are inspired by and closely related to the proofs by various authors. If this is the case, at the start of a proof the relation is given between brackets.

In this chapter, we will study a complete proof of Dirichlet's Theorem. We analyse the proof as published by J.-P. Serre in *A Course in Arithmetic* [5] in 1973. We refer to his work throughout the chapter, the work being from this source. The proof starts by a paragraph on Group Theory. This will appear again later on, after a part on Dirichlet series. After the section on L-functions, we are ready to prove the theorem. But before that, let us study characters of groups:

3.1 Characters of a finite abelian group

Let G be an abelian group with multiplication as its group law and order $|G| < \infty$. We will be working with this group for the remainder of the section, simply referring to it as G.

Definition 2. A **character** of G is a homomorphism from G to the multiplicative group $\mathbb{C}^*$.

The set of characters of G forms a group which we call the dual of G, denoted $\hat{G}$. The group law in $\hat{G}$ is defined by $\chi_1\chi_2(g) = \chi_1(g)\chi_2(g)$. Let χ_1 and χ_2 be characters of G. Then $\chi_1\chi_2(g)$ is another character of G because

$$\chi_1\chi_2(gh) = \chi_1(gh)\chi_2(gh) = \chi_1(g)\chi_1(h)\chi_2(g)\chi_2(h) = \chi_1(g)\chi_2(g) \cdot \chi_1(h)\chi_2(h).$$

The identity element of $\hat{G}$ is the trivial character, denoted as

$$\begin{aligned} 1 \colon G &\to \mathbb{C}^* \\ g &\mapsto 1 \qquad \forall g \in G. \end{aligned}$$

Finally, an inverse element for any character $\chi \in \hat{G}$ can be defined:

$$\begin{aligned} \chi^{-1} \colon G &\to \mathbb{C}^* \\ g &\mapsto \chi(g)^{-1} \qquad \forall g \in G, \end{aligned}$$

which is a character because

$$\chi^{-1}(gh) = \chi(gh)^{-1} = (\chi(g)\chi(h))^{-1} = \chi(g)^{-1}\chi(h)^{-1} = \chi^{-1}(g)\chi^{-1}(h).$$

We can conclude that $\hat{G}$ is indeed a group.

Proposition 1. Let H be a subgroup of G. Every character of H can be extended to a character of G. This extension can be done in $[G : H]$ possible ways, where $[G : H]$ denotes the index of H in G.

Proof. (A variation on Serre)
We prove this using induction with respect to the index $[G : H]$. If $[G : H] = 1$, the groups are equal and the one way to extend the character is to itself. So for this case the argument is done. Now suppose $[G : H] > 1$ and let χ be a character of H. We will extend this character to a subgroup $H' \subseteq G$, which has more elements than H. Therefore, $[G : H']$ will be smaller than $[G : H]$. And thus the induction hypothesis allows us to extend the character to G. We approach proving this as follows:

let x be an element of G which is not contained in H. Let n be the smallest positive integer such that $x^n \in H$. Such n exists because G is a finite group and so $ord(n) < \infty$. Let H' be the group generated by H and x. H' is a subgroup of G and we claim H' is the following group:

$$H' = \{hx^a | h \in H,\ 0 \leq a < n\}$$

The group H' is defined in such a way that all elements are of the form hx^a, where $h \in H$ and $0 \leq a < n$. It is thus clear that

$$y \in H' \implies y \in \{hx^a | h \in H,\ 0 \leq a < n\}.$$

To prove that the same holds the other way, we have to prove that $\{hx^a | h \in H,\ 0 \leq a < n\}$ is a group. To show this, let $h_1, h_2 \in H$ and $0 \leq a \leq b < n$. Besides the identity element being in H', also

$$h_1x^a \cdot h_2x^b = h_1h_2x^{a+b} \in H',$$

as H is commutative. Furthermore,

$$(h_1x^a)^{-1} = (x^{-1})^a h_1^{-1} = h_1^{-1}x^{(n-1)a} \in H',$$

and so H' is a group. It is also true that $|H'| = n|H|$. This follows from the property that ever element of H' can be written in a unique way as $h' = hx^a$ for $h \in H$ and $0 \leq a < n$. We will prove this now: for $h_1, h_2 \in H$ and $0 \leq a \leq b < n$, we see that

$$h_1 x^a = h_2 x^b \Rightarrow x^{a-b} = (h_1)^{-1} h_2 \in H.$$

By the restrictions on a and b, it must hold that $a = b$. Hence $h_1 = h_2$ and so each element of H' can be written in a unique way as $h' = hx^a$. Therefore, it holds that $|H'| = n|H|$.

Now set $t = \chi(x^n) \in \mathbb{C}^*$. There exist n distinct elements $w \in \mathbb{C}^*$ such that $w^n = t$. Given any such w, we define the extended character $\chi_w : H' \to \mathbb{C}^*$ as follows:

$$\chi_w(h') = \chi_w(hx^a) = \chi_w(h)\chi_w(x)^a = \chi(h)w^a$$

This is unique for every element h' as all elements of H' can be written distinctly as $h' = hx^a$. It is indeed a character as

$$\chi_w(h_1' h_2') = \chi(h_1 h_2)w^{a+b} = \chi(h_1)w^a \cdot \chi(h_2)w^b = \chi_w(h_1')\chi_w(h_2').$$

Note that $\chi_w(h) = \chi(h)$ for all $h \in H$, as $a = 0$ for these elements. There are no other ways to define a character extended to H'. To prove this, let χ of H be extended to χ' of H'. Then,

$$\chi'(h') = \chi(h)\chi'(x^a) = \chi(h)\chi'(x)^a.$$

But $\chi'(x)^n = \chi(x^n) = t$ and therefore $\chi'(x) = w$ for some $w^n = t$. Thus, any character of H extended to H' must be of the form $\chi_w(h') = \chi(h)w^a$.

So the character $\chi \in \hat{H}$ can be extended to some $\chi' \in \hat{H}'$. The extended character χ' can be defined in $n = [H' : H]$ possible ways, because $\chi'(h') = \chi(h)w^a$ and there exist n distinct elements w such that $w^n = t$. As $[G : H'] < [G : H]$, by the induction hypothesis χ can be extended to all of G. Continuing the method, this can be done in $[G : H'][H' : H] = [G : H]$ possible ways. This concludes the proof. ☼

Proposition 2. There exist $|G|$ distinct characters of G.

Proof. Let $\{1\}$ be the trivial subgroup of G. Only the trivial character exists on this subgroup. By Proposition 1, this character can be extended in $[G : \{1\}] = |G|$ ways. This defines all characters on G. ☼

Proposition 3. Let $\chi \in \hat{G}$. Let $|G| = n$. Then

$$\sum_x \chi(x) = \begin{cases} n & \text{when } \chi = 1 \\ 0 & \text{when } \chi \neq 1, \end{cases}$$

where the sum is taken over all elements x of G.

Proof. (Serre's proof with added intermediate steps)
The first formula is clear as the trivial character $\chi = 1$ takes all n elements to 1. For the second formula, let χ be a non-trivial character and let y be an element of G such that $\chi(y) \neq 1$. This must exist, as we have taken a non-trivial character. Then,

$$\chi(y) \sum_x \chi(x) = \sum_x \chi(x)\chi(y) = \sum_x \chi(xy) = \sum_x \chi(x).$$

Because of this,

$$(1 - \chi(y)) \sum_x \chi(x) = 0.$$

By the choice of y, it follows that $\sum_x \chi(x) = 0$ and the proposition is proven. ☼

Proposition 4. Let $x \in G$. Let $|G| = n$. Then

$$\sum_{\chi} \chi(x) = \begin{cases} n & \text{when } x = 1 \\ 0 & \text{when } x \neq 1, \end{cases}$$

where the sum is taken over all characters χ of G.

Proof. (Inspired by Serre's proof for the previous proposition)
The proof is very similar to the proof for the previous proposition. Using Proposition 2, the first formula is clear: all characters take the identity element $x = 1$ to 1 and there are n characters of G. For the second formula, we take an arbitrary $x \neq 1$ in G. Let $\langle x \rangle$ be the cyclic subgroup of G generated by x. This subgroup is of order $k > 1$. Define a character on $\langle x \rangle$ as follows:

$$\begin{aligned} \widetilde{\chi_x} : \langle x \rangle &\to \mathbb{C}^* \\ x^a &\mapsto e^{\frac{2\pi i}{k} a}. \end{aligned}$$

Note that this indeed defines a character on $\langle x \rangle$. By Proposition 1 this character can be extended to a character χ_x of G. It holds that $\chi_x(x) = \widetilde{\chi_x}(x) = e^{\frac{2\pi i}{k}} \neq 1$. We see that:

$$\chi_x(x) \sum_{\chi} \chi(x) = \sum_{\chi} \chi_x(x)\chi(x) = \sum_{\chi} \chi_x\chi(x) = \sum_{\chi} \chi(x).$$

This gives:

$$(1 - \chi_x(x)) \sum_{\chi} \chi(x) = 0.$$

By the choice of χ_x, it follows that $\sum_{\chi} \chi(x) = 0$ and the proposition is proven. ☼

We will work with multiplicative groups of integers modulo m later, as this is what is needed for the argument. Therefore we define characters on these group as follows: We call an element χ of the dual of $(\mathbb{Z}/m\mathbb{Z})^*$ a *character modulo m*. In this case, consider χ as a map $\mathbb{Z} \to \mathbb{C}$ by defining

$$\chi(a) = \begin{cases} \chi(\bar{a}) & \text{when } \bar{a} \in (\mathbb{Z}/m\mathbb{Z})^* \\ 0 & \text{otherwise} \end{cases}$$

where $\bar{a} \equiv a \pmod{m}$. The order of the group $(\mathbb{Z}/m\mathbb{Z})^*$ is the amount of elements $n < m$ such that $gcd(n, m) = 1$: this is denoted $\phi(m)$, also called the Euler-ϕ function.

It is possible for the characters of a group to take only real values: the values 1 and -1. It is interesting to note that this happens for characters χ modulo m if and only if $m \mid 24$. Otherwise, a character takes values besides these real ones as well. We will prove this statement later and address the influence that the properties of a character have on our proof.

3.2 Dirichlet series

Definition 3. A **Dirichlet series** is a series of the form

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

where $a_n \in \mathbb{C}$ and $s \in \mathbb{C}$.

As s is a complex variable, we should make more clear what the part n^{-s} looks like. This can be rewritten as

$$\frac{1}{n^s} = \frac{1}{e^{s \log(n)}},$$

which should help in understanding the way the series behaves.

Many of the series in this proof are Dirichlet series. This is why we discuss some properties of them. The convergence of Dirichlet series allows us to define series later that we can work with. If there was no general convergence for this type of series, all functions would have to be proven to converge separately. We will show how Dirichlet series converge for values of s and how this convergence can be extended to other parts of the complex plane.

Proposition 5. If the exponents a_n of a Dirichlet series are bounded, there is absolute convergence for $R(s) > 1$.

(We denote the real part of the complex variable s by $R(s)$ and we will keep to this notation for the remainder of the paper.)

Proof. The convergence of the series $\sum_{n=1}^{\infty} 1/n^{\alpha}$ for $\alpha > 1$ is well known and we assume its proof known. Let $|a_n| \leq K$ for all n and some $K < \infty$. This implies that the Dirichlet series converges absolutely:

$$\left|\frac{a_n}{n^s}\right| \leq \frac{|a_n|}{n^{R(s)}} \leq \frac{K}{n^{R(s)}},$$

and the series $\sum K/n^{R(s)}$ converges for $R(s) > 1$ as K is finite. So $\sum |a_n/n^s|$ converges, and the proposition is proven. ☼

Lemma 1. Let U be an open subset of $\mathbb{C}$. Let f_n be a sequence of holomorphic functions on U which converges uniformly to a function f on every compact subset of U. Then f is holomorphic in U and derivatives f_n' of the functions f_n converge uniformly to the derivative f' of f on all compact sets.

Proof. This proof is quickly attained by the use of Cauchy's integral formula. For this, let D be a closed disc in U and let Γ be its boundary which is positively oriented. Then,

$$f_n(z_0) = \frac{1}{2\pi i} \int_{\Gamma} \frac{f_n(z)}{z - z_0} dz$$

for all z_0 in the interior of D. Taking the limit on both sides gives us

$$f(z_0) = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{z - z_0} dz$$

So f is holomorphic for all z_0 in the interior of D. This proves the first part of the lemma. For the second part, we use the formula

$$f_n'(z_0) = \frac{1}{2i\pi} \int_{\Gamma} \frac{f_n(z)}{(z - z_0)^2} dz.$$

This is obtained from the Taylor series of f_n around z_0 together with the residue theorem. As a result of this, the derivative f_n' converge uniformly to the derivative f' of f on all compact sets. ☼

Lemma 2 (Abel's lemma)**.** Let (a_n) and (b_n) be two sequences. Define

$$A_{m,m'} = \sum_{n=m}^{n=m'} a_n \qquad \text{and} \qquad S_{m,m'} = \sum_{n=m}^{n=m'} a_n b_n.$$

Then

$$S_{m,m'} = \sum_{n=m}^{n=m'-1} A_{m,n}(b_n - b_{n+1}) + A_{m,m'} b_{m'}$$

Proof. To achieve this result, we replace a_n by $A_{m,n} - A_{m,n-1}$ and regroup the brackets. This immediately provides us with the result. The lemma can be applied to the Dirichlet series $\sum a_n/n^s$. For this, set $b_n = n^{-s}$. Then we obtain the partial sum

$$\sum_{m}^{m'} a_n/n^s = S_{m,m'} = \sum_{n=m}^{n=m'-1} A_{m,n}\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) + A_{m,m'}\frac{1}{(m')^s}.$$

☼

Lemma 3. Let α, β be real numbers with $0 < \alpha < \beta$. Let z be a complex number such that $R(z) = x > 0$. Then,

$$|e^{-\alpha z} - e^{-\beta z}| \leq \frac{|z|}{x}(e^{-\alpha x} - e^{-\beta x}).$$

Proof. (The proof as done by Serre)
To prove this, we write

$$e^{-\alpha z} - e^{-\beta z} = z\int_{\alpha}^{\beta} e^{-tz}dt.$$

Note that $\left|e^{-tz}\right| = \left|e^{-t(R(z)+Im(z))}\right| = e^{-tx}$. Therefore, taking absolute values on either side gives

$$|e^{-\alpha z} - e^{-\beta z}| \leq |z|\int_{\alpha}^{\beta} e^{-tx}dt = \frac{|z|}{x}(e^{-\alpha x} - e^{-\beta x}),$$

which shows the result. ☼

Proposition 6. Let the Dirichlet series $f(s) = \sum a_n/n^s$ converge for $s = s_0$. Then f converges uniformly in all domains of the form $R(s - s_0) \geq 0$ and $Arg(s - s_0) \leq \theta$, with $\theta < \pi/2$.

Proof. (Similar to the proof by Serre, using a different series)
We make a translation on s such that we can suppose $s_0 = 0$. Then the series $f(0) = \sum a_n$ is convergent by hypothesis. The restrictions for the domain of uniform convergence become $R(s) \geq 0$ and $|s|/R(s) \leq k$. Since $\sum a_n$ converges, there exist an N and an $\varepsilon > 0$ such that

$$\text{if } m, m' \geq N, \qquad \text{then } \left|\sum_{n=m}^{n=m'} a_n\right| = |A_{m,m'}| \leq \varepsilon.$$

The notations are those of Lemma 2, which we applied to the Dirichlet series before to get

$$S_{m,m'} = \sum_{n=m}^{n=m'-1} A_{m,n}\left(e^{-\log(n)s} - e^{-\log(n+1)s}\right) + A_{m,m'}e^{-s\log(m')}.$$

This is rewritten, so that we can apply Lemma 3 to the exponentials with $\alpha = \log(n)$ and $\beta = \log(n+1)$. This lemma results in

$$|S_{m,m'}| \leq \varepsilon\left(1 + \frac{|s|}{R(s)}\sum_{n=m}^{n=m'-1}\left(e^{-\log(n)x} - e^{-\log(n+1)x}\right)\right).$$

Now we work out the series and simplify:

$$|S_{m,m'}| \leq \varepsilon\left(1 + k\left(e^{-\log(m)x} - e^{-\log(m')x}\right)\right)$$

Thus

$$|S_{m,m'}| \leq \varepsilon(1+k),$$

which proves the desired uniform convergence in the domain. ☼

Corollary 1. If f converges for $s = s_0$, it also converges for $R(s) > R(s_0)$ and the function thus defined is holomorphic.

This follows from Proposition 6 and Lemma 1.

Proposition 7. Let f be the Dirichlet series $f(s) = \sum a_n/n^s$ with real coefficients $a_n \geq 0$. Suppose the function f converges for $R(s) > \rho$, where ρ is real. Also suppose that f can be extended analytically to a function that is holomorphic in a neighbourhood of the point $s = \rho$, which lies on the real line. Then there exists a $\varepsilon > 0$ such that f converges for $R(s) > \rho - \varepsilon$.

(This means that the domain of convergence of f can be extended to the left in the complex plane as long as we do not run into a *singularity* of f on the real axis.)

Proof. (Similar to the proof by Serre, using a different series)
In similar fashion as we started the proof for Proposition 6, we make a translation on f. We can thus assume that $\rho = 0$. The function f is holomorphic for $R(s) > 0$ and also in a neighbourhood of 0 by assumption. Then f is holomorphic in a disc $|s-1| \leq 1+\varepsilon$, where $\varepsilon > 0$. As the function is holomorphic there, its Taylor series converges in the disc. The Taylor series around 1, is written as

$$f(s) = \sum_{p=0}^{\infty} \frac{1}{p!}(s-1)^p f^{(p)}(1) \qquad \text{for } |s-1| \leq 1+\varepsilon.$$

The point we want to extend our convergence to is $s = -\varepsilon$, here the Taylor series is

$$f(-\varepsilon) = \sum_{p=0}^{\infty} \frac{1}{p!}(1+\varepsilon)^p(-1)^p f^{(p)}(1).$$

The point lies in our disc and this series is thus convergent.

Now, by Lemma 1, we can compute the pth derivative of f. This derivative is

$$f^{(p)}(s) = \sum (-1)^p(\log(n))^p \frac{a_n}{n^s},$$

or for $s = 1$,

$$f^{(p)}(1) = \sum (-1)^p(\log(n))^p \frac{a_n}{n}.$$

Then $(-1)^p f^{(p)}(1) = \sum (\log(n))^p a_n/n$ and this is a convergent series with non-negative terms. This can be substituted into the Taylor series at $s = -\varepsilon$ to get the following convergent double series with non-negative coefficients:

$$\begin{aligned} f(-\varepsilon) &= \sum_{p,n} \frac{1}{p!}(1+\varepsilon)^p(\log(n))^p \frac{a_n}{n} \\ &= \sum_{n} \frac{a_n}{n} \sum_{p=0}^{\infty} \frac{1}{p!}(1+\varepsilon)^p(\log(n))^p. \end{aligned}$$

The positive coefficients make it so that we can rearrange the sums. This series over p can be recognized as the Taylor expansion of n^{1-s} at the point $s = -\varepsilon$. We use this to rewrite the function and get

$$f(-\varepsilon) = \sum_n \frac{a_n}{n} n^{1+\varepsilon} = \sum_n \frac{a_n}{n^{-\varepsilon}}.$$

This is the Dirichlet series at the point $s = -\varepsilon$. It came from a double series, which was convergent. So the Dirichlet series must also be convergent at $s = -\varepsilon$. By Corollary 1 of Proposition 6 this shows that the series also converges for $R(s) > -\varepsilon$ and the proof is complete. ☼

Proposition 8. Let f be the Dirichlet series $f(s) = \sum a_n/n^s$ with bounded partial sums $A_{m,m'} = \sum_m^{m'} a_n$. Then, the series converges for $R(s) > 0$.

Proof. (Serre's proof with added intermediate steps)
Let $|A_{m,m'}| \leq K$ for some $K < \infty$ as the partial sums are bounded. We apply Lemma 2 again and take absolute values to obtain

$$|S_{m,m'}| \leq K \left(\left| \frac{1}{(m')^s} \right| + \sum_{m}^{m'-1} \left| \frac{1}{n^s} - \frac{1}{(n+1)^s} \right| \right).$$

To continue, suppose s is real and non-negative. As a result of this, the absolute value brackets on the right-hand side can be removed as all terms in them will be positive. Then we can work out the series to get

$$|S_{m,m'}| \leq \frac{K}{m^s},$$

which makes the convergence clear for real s. By Proposition 6, the convergence holds for all complex s such that $R(s) > 0$. ☼

3.3 The zeta function

We repeat the definition of a multiplicative function from the previous chapter:

Definition 4. A function $g : \mathbb{N} \to \mathbb{C}$ is called **multiplicative** if $g(1) = 1$ and

$$g(mn) = g(m)g(n)$$

whenever integers n and m are relatively prime. If this equality holds for all pairs $n, m \in \mathbb{N}$, then g is said to be **strictly multiplicative**.

Let g be any bounded function that is strictly multiplicative.

Lemma 4. The Dirichlet series $f(s) = \sum g(n)/n^s$ converges absolutely for $R(s) > 1$. In this domain its sum is equal to

$$\prod_{p \in \mathcal{P}} \frac{1}{1 - g(p)/p^s},$$

which is a convergent infinite product. (Here, and in the rest of the paper, the set of prime numbers is denoted as $\mathcal{P}$.)

Proof. (Serre's proof with added intermediate steps)
By Proposition 5, the Dirichlet series converges absolutely for $R(s) > 1$ as the functions g are bounded. To prove equality to the infinite product, we define two sets: let S be a finite set of prime numbers and let $N(S)$ be the subset of positive integers all of whose prime factors belong to S. Then by the strictly multiplicative property of g, it holds for any $n \in N(S)$ such that $n = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k}$:

$$\frac{g(n)}{n^s} = \frac{g(p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k})}{(p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k})^s} = \frac{g(p_1^{\alpha_1})}{p_1^{\alpha_1 s}} \frac{g(p_2^{\alpha_2})}{p_2^{\alpha_2 s}} \ldots \frac{g(p_k^{\alpha_k})}{p_k^{\alpha_k s}}.$$

The prime numbers $p_1, p_2, \ldots, p_k$ are all elements of the set S and any natural number that can be written as a product of prime numbers from the set S is an element of the set $N(S)$. Therefore,

$$\sum_{n \in N(S)} \frac{g(n)}{n^s} = \prod_{p \in S} \left(\sum_{m=0}^{\infty} \frac{g(p^m)}{p^{ms}} \right).$$

When we increase both S and $N(S)$, the left-hand side tends to the Dirichlet series $\sum_{n=1}^{\infty} g(n)/n^s$. The right-hand side will then go to $\prod_{p\in\mathcal{P}} \left(\sum g(p^m)/p^{ms}\right)$. Because g is strictly multiplicative, we can rewrite this to get

$$\sum_{n=1}^{\infty} \frac{g(n)}{n^s} = \prod_{p\in\mathcal{P}} \left(\sum_{m=0}^{\infty} \frac{g(p)^m}{p^{ms}} \right).$$

By the geometric series, we see that this convergent infinite product is indeed equivalent to the required result. ☼

Definition 5. The **zeta function** is the function

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

where we ensure that $R(s) > 1$ for the formulas to converge and thus behave properly. This equality follows from Lemma 4 applied to $g = 1$.

Now that we have established the zeta function as a Dirichlet series, it is ensured that we can work with the function in a proper way. In the following part, we will learn more about the zeta function. This provides us with knowledge about this function and other functions which we will use later. The L-function, which we will be working with later, gains its properties directly from the zeta function. This is why it is so important to assert these facts now.

Proposition 9. The zeta function can be written as

$$\zeta(s) = \frac{1}{s-1} + \kappa(s),$$

where κ is holomorphic for $R(s) > 0$.

Proof. (Serre's proof with added intermediate steps)
We see that

$$\frac{1}{s-1} = \int_1^{\infty} t^{-s} dt = \sum_{n=1}^{\infty} \int_n^{n+1} t^{-s} dt.$$

Hence we can write

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

We set the latter part equal to κ in the following way:

$$\kappa_n(s) = \int_n^{n+1} (n^{-s} - t^{-s}) dt \qquad \text{and hence} \qquad \kappa(s) = \sum_{n=1}^{\infty} \kappa_n(s).$$

Now we need to show that κ is defined and holomorphic for $R(s) > 0$. This would be proven if we could show that all κ_n have these properties and $\sum \kappa_n$ converges uniformly. We can write

$$\kappa_n = \int_n^{n+1} (n^{-s} - t^{-s}) dt = \frac{1}{1-s} \left(n^{-s} - sn^{-s} - (n+1)^{1-s} + n^{1-s} \right),$$

which shows that all κ_n are defined and holomorphic for $R(s) > 0$. Then we will prove that the series $\sum \kappa_n$ converges uniformly. For this, see that

$$|\kappa_n(s)| \leq \sup_{t\in[n,n+1]} \left| n^{-s} - t^{-s} \right|.$$

The derivative of $n^{-s} - t^{-s}$ with respect to t is equal to s/t^{s+1}. The absolute value of this derivative is biggest at $t = n$:

$$\left|\frac{s}{t^{s+1}}\right| = \frac{|s|}{t^{R(s)+1}} \leq \frac{|s|}{n^{R(s)+1}} \qquad \text{for all } t \in [n, n+1]$$

Then, the absolute value of $\kappa_n(s)$ for any $t \in [n, n+1]$ can not be bigger than the sum of the absolute value at $t = n$ and 1 times the absolute value of the derivative here:

$$|\kappa_n(s)| \leq \left|n^{-s} - n^{-s}\right| + \left|\frac{s}{n^{s+1}}\right| = \frac{|s|}{n^{R(s)+1}}.$$

This shows the convergence of the series for $R(s) \geq \varepsilon$ for all $\varepsilon > 0$, which concludes the argument. ☼

For the next corollary, there is some more to be said about the logarithm used. We define $\log \frac{1}{1-\alpha}$ as $\sum_1^\infty \alpha^n/n$ for $|\alpha| < 1$. This is a holomorphic function with a continuation to a meromorphic function on the whole complex plane. We see that

$$e^{\sum \alpha^n/n} = 1 + \left(\alpha + \frac{\alpha^2}{2} + \ldots\right) + \left(\alpha + \frac{\alpha^2}{2} + \ldots\right)^2 + \ldots.$$

This is equal to $1 + \alpha + \alpha^2 + \ldots$, which is a geometric series and thus equals $1/1 - \alpha$ for $|\alpha < 1|$. Therefore, it holds that

$$e^{\sum \alpha^n/n} = \frac{1}{1-\alpha}.$$

and so we have a valid definition for our logarithm. Note that $\log uv = \log u + \log v$ as

$$e^{\log uv} = uv = e^{\log u} e^{\log v} = e^{\log u + \log v}.$$

Corollary 2. It holds that $\lim_{s\to 1} \left(\sum_p p^{-s}\right) / \log\left(\frac{1}{s-1}\right) = 1$ and $\sum_{p,k\geq 2} 1/p^{ks}$ stays bounded.

Proof. (Serre's proof with added intermediate steps)
We see that

$$\log \zeta(s) = \log\left(\prod_{p\in\mathcal{P}} \frac{1}{1-\frac{1}{p^s}}\right) = \sum_{p\in\mathcal{P}} \log\left(\frac{1}{1-\frac{1}{p^s}}\right).$$

Using our definition of the logarithm, we rewrite this as

$$\log \zeta(s) = \sum_{p,k\geq 1} \frac{1}{kp^{ks}} = \sum_p \frac{1}{p^s} + \psi(s).$$

For this, introduce the series $\psi(s) = \sum_{p,k\geq 2} 1/kp^{ks}$. For the absolute values of the terms $\left|1/kp^{ks}\right| < \left|1/p^{ks}\right|$ as $k \geq 2$. If we can prove that $\sum_{p,k\geq 2} 1/p^{ks}$ converges absolutely, so will $\psi(s)$. First, prove this convergence in the following way:

$$\sum_{p,k\geq 2} \left|\frac{1}{p^{ks}}\right| = \sum_p \left|\frac{1/p^s}{1-1/p^s} - \frac{1}{p^s}\right| = \sum_p \left|\frac{1}{p^s(p^s-1)}\right| \leq \sum_p \frac{1}{p(p-1)} \leq \sum_{n=2}^\infty \frac{1}{n(n-1)}.$$

This sum converges to 1 and so,

$$\sum_{p,k\geq 2} \frac{1}{p^{ks}} \leq \sum_{n=2}^\infty \frac{1}{n(n-1)} = 1.$$

Thus, this sum is bounded and the second result has been shown. This also implies that $\psi(s)$ is bounded. To prove the first result, we use that $(s-1)\zeta(s)$ is bounded for s in a neighbourhood of 1.

This holds as the zeta function has only a simple pole at $s = 1$. So for some neighbourhood around $s = 1$ and some K with $|K| < \infty$:

$$(s-1)\zeta(s) = K.$$

Taking the logarithm on either side and making sure these exist in the neighbourhood, it follows that

$$\log(s-1) + \log\zeta(s) = \log K.$$

Now divide each side by $\log\frac{1}{(s-1)}$ and rearrange the terms to get

$$\log\zeta(s)/\log\left(\frac{1}{s-1}\right) = 1 + \log K/\log\left(\frac{1}{s-1}\right).$$

If we take the limit as $s \to 1$, the value of $\log K$ stays bounded and so

$$\lim_{s\to 1}\log\zeta(s)/\log\left(\frac{1}{s-1}\right) = 1.$$

We divide the equality from before by $\log\frac{1}{(s-1)}$:

$$\log\zeta(s)/\log\left(\frac{1}{s-1}\right) = \sum_p \frac{1}{p^s}/\log\left(\frac{1}{s-1}\right) + \psi(s)/\log\left(\frac{1}{s-1}\right).$$

As $s \to 1$, the value of $\psi(s)$ stays bounded and so:

$$\lim_{s\to 1}\sum_p \frac{1}{p^s}/\log\left(\frac{1}{s-1}\right) = \lim_{s\to 1}\log\zeta(s)/\log\left(\frac{1}{s-1}\right) = 1.$$

This concludes the proof. ☼

3.4 L-functions

Definition 6. Let $m \geq 1$ be an integer and let χ be a character modulo m. The **L-function** corresponding to m and χ is defined by the series

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

This is a Dirichlet series. Such a character χ modulo m takes values from the set of natural numbers to the complex plane, but only the integers prime to m will give non-zero terms. The function χ is defined to be strictly multiplicative and is bounded.

The L-function is closely related to the zeta function and we will see that some of the zeta functions properties transfer. Afterwards, we will use the L-function to prove the hardest step in the proof of Dirichlet's Theorem. By the way this function is defined, we can use it later to provide the much needed convergence of a series. Firstly, it has to be shown that the L-function does not vanish at $s = 1$ for any character $\chi \neq 1$. This is the aim of this paragraph.

Proposition 10. For $\chi = 1$,

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

Thus, $L(s,1)$ extends analytically for $R(s) > 0$ and it has a simple pole at $s = 1$.

Proof. The equality comes from the definition of the character: using the product expansion of the zeta function:

$$F(s)\zeta(s) = \prod_{p|m}\left(1-\frac{1}{p^s}\right)\prod_{p\in\mathcal{P}}\frac{1}{1-\frac{1}{p^s}} = \prod_{p\nmid m}\frac{1}{1-\frac{1}{p^s}}.$$

This is equal to $L(s,1)$, the expansion can be found in the same way that the expansion for the zeta function was found. As the function $F(s)$ is holomorphic for $R(s) > 0$ and so the properties transfer directly from the zeta function. This shows why the simple pole is there at $s = 1$ and the fact that $L(s,1)$ extends analytically. ☼

Proposition 11. For $\chi \neq 1$, the series $L(s,\chi)$ converges for $R(s) > 0$. For $R(s) > 1$, the function converges absolutely and

$$L(s,\chi) = \prod_{p\in\mathcal{P}}\frac{1}{1-\frac{\chi(p)}{p^s}} \qquad \text{for } R(s) > 1$$

Proof. (Serre's proof with added intermediate steps)
The properties of the series for $R(s) > 1$ are clear by Lemma 4 as $L(s,\chi)$ is a Dirichlet function with bounded, strictly multiplicative exponents χ. Showing the convergence for $R(s) > 0$ requires some more work: using Proposition 8, convergence follows from boundedness of the partial sums

$$A_{u,v} = \sum_{n=u}^{v}\chi(n) \qquad \text{where } u \leq v.$$

As $\chi \neq 1$, from Proposition 3 it follows that

$$\sum_{n=u}^{u+m-1}\chi(n) = 0.$$

Because of this, we only need to show that $A_{u,v}$ is bounded for $v - u < m$. For these values, we can easily majorize the sums: we have

$$|A_{u,v}| \leq \phi(m),$$

and the proposition is proven. ☼

We will now go on to prove a notoriously hard step in the proof of Dirichlet's theorem. This is the non-vanishing of $L(s,\chi)$ at $s = 1$ for the characters $\chi \neq 1$. For this, set m to be a fixed positive integer. For any integer $p \nmid m$ and the group $G(m) = (\mathbb{Z}/m\mathbb{Z})^*$, we denote by $\bar{p}$ the image of p in $G(m)$. Denote by $f(p)$ the order of $\bar{p}$ in the group $G(m)$, and put

$$g(p) = \frac{\phi(m)}{f(p)}.$$

Thus $g(p)$ is the order of the group $G(m)/\langle\bar{p}\rangle$, where $\langle\bar{p}\rangle$ is the subgroup of $G(m)$ generated by $\bar{p}$.

Lemma 5. If $p \nmid m$, the following identity holds:

$$\prod_{\chi}(1-\chi(p)T) = \left(1-T^{f(p)}\right)^{g(p)},$$

where the product is taken over all characters χ of $G(m)$.

Proof. (Serre's proof with added intermediate steps)
We start the proof by looking at the $1/T$-th roots of 1:

$$\left(\left(\frac{1}{T}\right)^{f(p)} - 1\right) = \prod_{\{w|w^{f(p)}=1\}}\left(\frac{1}{T} - w\right)$$

If we multiply both sides by $T^{f(p)}$, we get the following equality:

$$\left(1 - T^{f(p)}\right) = \prod_{\{w | w^{f(p)}=1\}} (1 - wT)$$

This $f(p)$ is defined to be the order of $\langle p \rangle$. On this group, there exist $f(p)$ characters. There exists one character χ of $\langle p \rangle$ such that $\chi(\bar{p}) = w$ for any such root w. Any such character can be extended to $[G(m) : f(p)] = g(p)$ characters of $G(m)$ by Proposition 1. From this, it follows that

$$\prod_{\chi} (1 - \chi(p)T) = \left(\prod_{w \in W} (1 - wT) \right)^{g(p)} = \left(1 - T^{f(p)}\right)^{g(p)}.$$

The lemma is shown. Now define a new function $\zeta_m(s)$ as follows:

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

where the product is extended over all characters χ of $G(m)$. ☼

Proposition 12. We have

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

This is a Dirichlet series with positive coefficients, that converges for $R(s) > 1$.

Proof. Replacing the L-functions by their product expansion yields

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

Now, apply Lemma 5 with $T = p^{-s}$ to achieve the required result. To show this is a Dirichlet series, first expand one term of the product:

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

We can use this as $\left| p^{-f(p)s} \right| = (p^{-f(p)R(s)}) < 1$ for all p. Work out this sum to get

$$\left(\sum_{m=0}^{\infty} p^{-f(p)ms} \right)^{g(p)} = \left(1 + p^{-f(p)s} + p^{-2f(p)s} + \dots \right)^{g(p)} = 1 + b_{p_1}^{-f(p)s} + b_{p_2}^{-2f(p)s} + \dots,$$

where $b_{p_1}, b_{p_2}, \dots$ are non-negative and depend on p. We plug this in to the product that is ζ_m:

$$\begin{aligned}
\zeta_m(s) &= \prod_{p \in \mathcal{P}} 1 + b_{p_1}p^{-f(p)s} + b_{p_2}p^{-2f(p)s} + \dots \\
&= \left(1 + b_{2_1}2^{-f(2)s} + b_{2_2}2^{-2f(2)s} + \dots \right) \left(1 + b_{3_1}3^{-f(3)s} + b_{3_2}3^{-2f(3)s} + \dots \right) \dots \\
&= \sum_{n=1}^{\infty} \frac{a_n}{n^s},
\end{aligned}$$

where the coefficients a_n are products of different b_{p_i} and are thus also non-negative. The coefficients are also bounded. This shows that $\zeta_m(s)$ is a Dirichlet series with non-negative coefficients and it follows that ζ_m converges for $R(s) > 1$ by Proposition 5. ☼

Theorem 2. The function $L(1,\chi) \neq 0$ for all $\chi \neq 1$.

Proof. (Serre's proof with added intermediate steps) For contradiction, assume that $L(1,\chi) = 0$ for some $\chi_1 \neq 1$. Then writing χ_1 as its Taylor expansion around $s = 1$ allows for a factor $s - 1$ to be factorized out. As a result of this, the simple pole of $L(s,1)$ at $s = 1$ that was shown in Proposition 10 is cancelled out when computing ζ_m. This, in combination with Proposition 11, implies that ζ_m is holomorphic at $s = 1$ and can thus be extended analytically to all of $R(s) > 0$. The function ζ_m also converges for $R(s) > 1$. By Proposition 7, this implies that ζ_m converges for all of $R(s) > 0$ as we run into no singularities. This can not hold however, we will show that this contradicts with the definition of our function: the pth factor of ζ_m is

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

This dominates the series

$$\sum_{k=0}^{\infty} p^{-k\phi(m)s}.$$

As a result of this, our function ζ_m has all its coefficients greater than those of the series

$$\sum_{(n,m)=1} n^{-\phi(m)s} = \prod_{p \nmid m} \sum_{k=0}^{\infty} p^{-k\phi(m)s}.$$

But this series diverges for $s = 1/\phi(m)$. Thus ζ_m must also diverge for this value of s, for which it is true that $R(s) > 0$. We have a contradiction and our assumption must be negated. The proposition follows. ☼

Corollary 3. The function ζ_m has a simple pole at $s = 1$.

As $L(s,\chi)$ has a simple pole for $\chi = 1$ and is nonzero for all other characters χ, this result is shown.

3.5 Proving Dirichlet's Theorem

As we saw in Corollary 2 to Proposition 9,

$$\lim_{s \to 1} \left(\sum_p p^{-s}\right) / \log\left(\frac{1}{s-1}\right) = 1.$$

Here $\mathcal{P}$ is the set of prime numbers. We now define s to be real and bigger than 1 to not cause any problems. The limit is taken as $s \to 1$ from above.

Definition 7. Let A be a subset of $\mathcal{P}$. The set A has a **density** of k, when

$$\left(\sum_{p \in A} \frac{1}{p^s}\right) \Big/ \left(\log\left(\frac{1}{s-1}\right)\right)$$

tends to k as $s \to 1$. As $A \subset \mathcal{P}$, this means that $0 \leq k \leq 1$. This is also called *Dirichlet density*.

This concept is defined to state a relation between the set A and the set $\mathcal{P}$. Using this definition, we can rewrite Dirichlet's Theorem as follows:

Theorem 3 (Dirichlet's theorem)**.** Let $m \geq 1$ and a be positive integers such that $(a,m) = 1$. Let $\mathcal{P}_a$ be the set of prime numbers p such that $p \equiv a \pmod{m}$. Then, the set $\mathcal{P}_a$ has density $1/\phi(m)$.

This theorem implies a corollary, which shows that there is an infinite amount of primes in the arithmetic progression $a \pmod{m}$.

Corollary 4. The set $\mathcal{P}_a$ is infinite.

Any finite set has density zero.

Lemmas

To prove the theorem, some lemmas are needed. Let χ be a character of $G(m)$. Define

$$f_\chi(s) = \sum_{p \nmid m} \frac{\chi(p)}{p^s},$$

which converges for $R(s) > 1$. In the following lemma 7, it becomes clear why the L-function is defined to be how it is. Taking the logarithm results in a combination of series that we want to show convergent at $s = 1$. As it is shown earlier that the L-function does not vanish at this point, its logarithm must be bounded in the way we define it.

Lemma 6. If $\chi = 1$, then $f_\chi \sim \log\left(\frac{1}{s-1}\right)$ for $s \to 1$ from the right side of the complex plane.

For the trivial character $\chi = 1$, our function becomes $f_1 = \sum_{p \nmid m} \frac{1}{p^s}$. This differs from the series $\sum_{p \in \mathcal{P}} \frac{1}{p^s}$ only by a finite number of terms and thus the equivalence is implied.

Lemma 7. If $\chi \neq 1$, the function f_χ remains bounded when $s \to 1$ from the right side of the complex plane.

Proof. (Serre's proof with added intermediate steps)
To prove the boundedness of f_χ, we use the logarithm of the function $L(s, \chi)$. Again, we use our definition of the logarithm:

$$\log\left(\frac{1}{1-\alpha}\right) = \sum_{n=1}^{\infty} \frac{\alpha^n}{n} \qquad \text{for } |\alpha| < 1.$$

We can apply this to the L-function, by using the product expansion of $L(s, \chi)$ and taking $\alpha = \chi(p)p^{-s}$:

$$\begin{aligned}\log L(s,\chi) &= \sum_{p \in P} \log \frac{1}{1-\chi(p)p^{-s}} \qquad \text{when } R(s) > 1 \\ &= \sum_{p,n} \frac{\chi(p)^n}{np^{ns}}.\end{aligned}$$

Note that this works as $|\chi(p)p^{-s}| < 1$. To gain information about f_χ, we split this sum up into two parts to get

$$\log L(s,\chi) = f_\chi(s) + F_\chi(s).$$

For this, we set

$$F_\chi(s) = \sum_{p, n \geq 2} \frac{\chi(p)^n}{np^{ns}}.$$

Theorem 2 states that $\log L(s, \chi)$ remains bounded as $s \to 1$ as $\chi \neq 1$ and the L-function converges here. By Corollary 2 to Proposition 9 the series $\sum_{p,n \geq 2} 1/p^{ns}$ remains bounded as $s \to 1$. This implies that $F_\chi(s)$ stays bounded then as

$$\left|F_\chi(s)\right| = \left|\sum_{p,n\geq 2} \frac{\chi(p)^n}{np^{ns}}\right| < \left|\sum_{p,n \geq 2} \frac{1}{p^{ns}}\right|.$$

. Now that $\log L(s, \chi)$ and $F_\chi(s)$ both stay bounded as $s \to 1$, so must $f_\chi(s)$. This holds for all $\chi \neq 1$ and the proposition is proven.

We will now prove Theorem 3. For this, introduce the following function:

$$g_a(s) = \sum_{p \in \mathcal{P}_a} \frac{1}{p^s}.$$

To study the behaviour of this function as $s \to 1$, we use the following Lemma 8. We see that this relies on the function that we defined and showed properties of in the previous lemmas. We can then use the convergence and divergence shown there to prove Dirichlet's Theorem. ☼

Lemma 8. We can rewrite the function $g_a(s)$ as follows:

$$g_a(s) = \frac{1}{\phi(m)} \sum_\chi \chi(a)^{-1} f_\chi(s),$$

where the sum is taken over all characters χ of $G(m)$.

Proof. (Serre's proof with added intermediate steps) This is most clear when proven in a backwards manner and thus we start by rewriting $\sum_\chi \chi(a)^{-1} f_\chi(s)$ by replacing f_χ by its definition, we get:

$$\sum_\chi \chi(a)^{-1} \sum_{p \nmid m} \frac{\chi(p)}{p^s} = \sum_{p \nmid m} \frac{\sum_\chi \chi(a)^{-1}\chi(p)}{p^s}$$

But $\chi(a)^{-1}\chi(p) = \chi(a^{-1})\chi(p)$. Here a^{-1} denotes the unique inverse of the element a in the group $(\mathbb{Z}/m\mathbb{Z})^*$. This is the element in $(\mathbb{Z}/m\mathbb{Z})^*$ such that $a^{-1}a \equiv 1 \pmod{m}$. We see that $\chi(a^{-1})\chi(p) = \chi(a^{-1}p)$ and by Proposition 4:

$$\sum_\chi \chi(a^{-1}p) = \begin{cases} \phi(m) & \text{when } a^{-1}p \equiv 1 \pmod{m} \\ 0 & \text{otherwise.} \end{cases}$$

The sum is thus only non-zero when $a^{-1}p \equiv 1 \pmod{m}$, which is when $p \equiv a \pmod{m}$. Hence, it holds that

$$\sum_\chi \chi(a)^{-1} f_\chi(s) = \sum_{p \nmid m} \frac{\sum_\chi \chi(a^{-1}p)}{p^s} = \phi(m) g_a(s),$$

and the lemma is proven. ☼

We are now ready to prove Theorem 3:

Proof. As $s \to 1$, for $\chi = 1$, the value $\chi(a)^{-1} = 1$ and the constant $1/\phi(m)$ stays the same. By Lemma 6 it is known that $f_\chi(s) \sim \log\left(\frac{1}{s-1}\right)$. Lemma 7 shows that all other f_χ stay bounded. Thus by Lemma 8, the following equivalence relation holds:

$$g_a(s) \sim \frac{1}{\phi(m)} \log \frac{1}{s-1} \qquad \text{as } s \to 1.$$

This proves Theorem 3 and the proof of Dirichlet's Theorem is concluded by Corollary 4. ☼

4 Other versions of the non-vanishing of L-functions

Besides the proof by Serre, there are many more proofs of Dirichlet's Theorem. In this chapter we will focus on a few different approaches. These are other ways of proving the result of Theorem 2 in the work by Serre. This theorem states that $L(1, \chi)$ is not equal to 0 for all non-trivial characters χ. The result is notoriously the hardest step in the proof. There are no ways around the proof however, as this step is crucial in proving the divergence we are after later on. This is why many mathematical writers have come up with their own way of proving the theorem. We focus on work by Monsky [3] and Veklych [6]. Both authors had their respective arguments published in the American Mathematical Monthly.

Merely these arguments and their proofs are not enough. The two authors prove the non-vanishing for real characters modulo m only. This would be enough if our unit modulo group $(\mathbb{Z}/m\mathbb{Z})^*$ had just real characters. This only holds for a handful of integers m however, as can be seen in the following lemma:

Lemma 9. All characters χ of the group $(\mathbb{Z}/m\mathbb{Z})^*$ are real if and only if $m \mid 24$.

Proof. For any positive integer m,

$$(\mathbb{Z}/m\mathbb{Z})^* = \prod_{p|m} (\mathbb{Z}/p^a\mathbb{Z})^*$$

by the Chinese Remainder Theorem for rings. Here a is the power corresponding to the prime p in the prime factorization of m. For $(\mathbb{Z}/m\mathbb{Z})^*$ to have only real characters, all its elements must have order 1 or 2. For if there exists an element x with $ord(x) > 2$, we can define a character $\chi_x : \langle x \rangle \to \mathbb{C}^*$ such that $\chi_x(x) \notin \mathbb{R}$. This character can be extended to a character on all of $(\mathbb{Z}/m\mathbb{Z})^*$, defining a non-real character. If all elements $y \in (\mathbb{Z}/m\mathbb{Z})^*$ have order 1 or 2, a character can only take real values as $\chi(y)^2 = \chi(y^2) = \chi(\overline{1}) = 1$.

Thus all elements of $(\mathbb{Z}/m\mathbb{Z})^*$ must have order 1 or 2. This holds if and only if all elements of $(\mathbb{Z}/p^a\mathbb{Z})^*$ have order 1 or 2 for all such groups in the Cartesian product. We will show that for $a \geq 1$, the group $(\mathbb{Z}/p^a\mathbb{Z})^*$ has only elements of order 1 or 2 if and only if p^a is 2, 3, 4 or 8. This proves the lemma, as then $m \mid 24$ for all $(\mathbb{Z}/m\mathbb{Z})^*$ for which this holds.

For any prime p and $a \geq 1$, we can define a surjective homomorphism

$$\begin{aligned} f : (\mathbb{Z}/p^a\mathbb{Z})^* &\to (\mathbb{Z}/p\mathbb{Z})^*, \\ x \pmod{p^a} &\mapsto x \pmod{p}. \end{aligned}$$

The group $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic since any finite group of units in any integral domain is. Therefore there exists a generating element $x \in (\mathbb{Z}/p\mathbb{Z})^*$ such that $ord(x) = p - 1$. Any element y in the inverse image of x has as order then a multiple of $p - 1$. This is an element $y \in (\mathbb{Z}/p^a\mathbb{Z})^*$. So for any $p \geq 5$, the group $(\mathbb{Z}/p^a\mathbb{Z})^*$ contains at least one element y with $ord(y) > 2$ for $a \geq 1$. What also follows from the result, is that the group $(\mathbb{Z}/3\mathbb{Z})^*$ has only elements of order 1 or 2, as it is cyclic.

For the first 3 powers of 2, check that the group $(\mathbb{Z}/p^a\mathbb{Z})^*$ only contains elements of order 1 and 2. For any greater power, an element of bigger order can be found. To show this, define two surjective homomorphisms that are similar to the function f above:

$$\begin{aligned} g_1 : (\mathbb{Z}/3^a\mathbb{Z})^* &\to (\mathbb{Z}/9\mathbb{Z})^* \mid x \pmod{3^m} \mapsto x \pmod{9} && \text{for } m \geq 2, \\ g_2 : (\mathbb{Z}/2^b\mathbb{Z})^* &\to (\mathbb{Z}/16\mathbb{Z})^* \mid x \pmod{2^n} \mapsto x \pmod{16} && \text{for } n \geq 4. \end{aligned}$$

In the group $(\mathbb{Z}/9\mathbb{Z})^*$, there is the element $\overline{2}$ with order 6. All elements in the inverse image of $\overline{2}$ then have as order a multiple of 6. Similarly, there is the element $\overline{3} \in (\mathbb{Z}/16\mathbb{Z})^*$ with order 4.

Therefore there exists at least one element of an order bigger than 2 in $(\mathbb{Z}/3^m\mathbb{Z})^*$ for $m \geq 2$ and in $(\mathbb{Z}/2^n\mathbb{Z})^*$ for $n \geq 4$.

Now, the group $(\mathbb{Z}/p^a\mathbb{Z})^*$ has only elements of order 1 and 2 for $a \geq 1$ if and only if $p^a \in \{2, 3, 4, 8\}$. This proves the lemma. ☼

As we are looking for a proof for all positive integers m, this is not enough. Both Monsky and Veklych note that this can be resolved easily. They note that the proof is significantly easier for non-real characters. For the completeness of this paper, we will also address that proof. To keep it approachable, we have combined work by Ram Murty [4] and Cohn [2] to construct an argument that fits in with the rest of our proof. We will start with this argument and treat the proofs for real characters after. In the end, a complete proof for Dirichlet's Theorem can be obtained by combining work by Serre, Ram Murty and Cohn and either Monsky or Veklych. In the next chapter, we will compare the arguments and comment more on the combination of these proofs.

4.1 M. R. Murty and H. Cohn on non-real characters

Lemma 10. For $R(s) > 1$, the Dirichlet series $\sum_1^\infty a_n n^{-s}$ that represents the product $\prod_\chi L(s, \chi)$ has the property that $a_1 = 1$ and $a_n \geq 0$ for all $n \geq 2$. The product of L-functions is taken over all characters χ modulo m for some positive integer m.

Proof. (Closely related to the proof by Ram Murty)
We start by writing $L(s, \chi)$ as its product expansion

$$L(s, \chi) = \prod_{p \in \mathcal{P}} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

Then, using the earlier defined logarithm

$$\log\left(\frac{1}{1-\alpha}\right) = \sum_{n=1}^{\infty} \frac{\alpha^n}{n} \qquad \text{for } |\alpha| < 1,$$

we can write the following:

$$\log \sum_{p \in \mathcal{P}} L(s, \chi) = \sum_\chi \sum_p \sum_{n=1}^{\infty} \frac{\chi(p^n)}{np^{ns}} = \sum_{n,p} \frac{1}{np^{ns}} \left(\sum_\chi \chi(p^n)\right).$$

The change in summation order is permitted as the series converges absolutely for $R(s) > 1$. This has been proven for the L-function in Proposition 11 and it transfers to this series. Now, the sum $\sum \chi(p^n)$ over all characters χ mod m is equal to $\phi(m)$ if $p^n \equiv 1 \pmod{m}$ and 0 otherwise. This follows from Proposition 4. We can thus rewrite as

$$\log \sum_{p \in \mathcal{P}} L(s, \chi) = \phi(m) \sum_n \sum_{p^n \equiv 1 \pmod{m}} \frac{1}{np^{ns}}.$$

If we exponentiate both sides and use the equality $e^x = 1 + x + x^2/2 + \dots$, we see that

$$\prod_\chi L(s, \chi) = 1 + \left(\phi(m) \sum_n \sum_{p^n \equiv 1 \pmod{m}} \frac{1}{np^{ns}}\right) + \dots.$$

This proves the lemma as the exponents a_n will all be non-negative and $a_1 = 1$. ☼

Theorem 4. The function value $L(1, \chi_1) \neq 0$ for all characters $\chi_1 \neq 1$ such that $\chi_1 \neq \overline{\chi_1}$. (That is, the character χ_1 is not real-valued.)

Proof. (Inspired by the proofs by Ram Murty and Cohn)
For contradiction, assume that for a non-real $\chi_1 \neq 1$, it holds that $L(1, \chi_1) = 0$. Then,

$$L(1, \overline{\chi_1}) = \sum_n \frac{\overline{\chi(n)}}{n} = \overline{\sum_n \frac{\chi(n)}{n}} = \overline{L(1, \chi_1)} = 0.$$

By the Taylor series of $L(s, \chi_1)$, we can write

$$L(s, \chi_1) = (s-1)g(s, \chi_1).$$

This $g(s, \chi)$ is continuous for all $R(s) > 0$, except possibly $s = 1$. By Lemma 1, we can compute the derivative of $L(s, \chi_1)$ at $s = 1$ and show that this is continuous for $R(s) > 0$. So if we set $g(1, \chi_1) = L'(1, \chi_1)$, the function $g(s, \chi_1)$ is continuous for all $R(s) > 0$. This procedure can also be applied to $\overline{\chi_1}$ to see that

$$L(s, \overline{\chi_1}) = (s-1)g(s, \overline{\chi_1}),$$

where $g(1, \chi_1) = L(1, \chi_1)$ at $s = 1$. This is thus also continuous for all $R(s) > 0$. Now,

$$\prod_{\chi} L(s, \chi) = L(s, 1)(s-1)^2 g(s, \chi_1) g(s, \overline{\chi_1}) \prod_{\chi \neq 1, \chi_1, \overline{\chi_1}} L(s, \chi).$$

Here the character 1 denotes the trivial character. The function $L(s, 1)(s-1)$ is continuous for all $R(s) > 0$ as the simple pole at $s = 1$ that $L(s, 1)$ has, is cancelled. Taking the limit on both sides gives

$$\lim_{s \to 1} \prod_{\chi} L(s, \chi) = \lim_{s \to 1} (s-1) L(s, 1)(s-1) g(s, \chi_1) g(s, \overline{\chi_1})$$

As $s \to 1$, all functions are continuous and can thus not yield infinity. But $(s-1) \to 0$, and so the whole product yields 0:

$$\lim_{s \to 1} \prod_{\chi} L(s, \chi) = 0.$$

But by Lemma 10,

$$\prod_{\chi} L(s, \chi) = 1 + \sum_{n=2}^{\infty} \frac{a_n}{n^s} \qquad \text{for } R(s) > 1.$$

The product function is thus greater than 1 real values of $s > 1$. If we would only have one non-trivial character χ such that $L(s, \chi) = 0$, the function would be continuous for all $R(s) > 0$, but this would not be any problem. Now that we have two such non-trivial characters however, there is a problem. We have now that the function is continuous, but also that the limit is 0 as $s \to 1$. But for real $s > 1$, the function is greater than 1. This would mean that the function is not continuous and this is a contradiction. We must thus conclude that $L(1, \chi_1) \neq 0$. ☼

4.2 P. Monsky on real characters

Lemma 11. Let χ be a real character modulo m, where m is a positive integer. Define $c_n := \sum \chi(d)$, where the sum is taken over all positive divisors d of n. Then the series $\sum_{n=1}^{\infty} c_n$ diverges.

Proof. (Monsky's proof with added intermediate steps)
For any power of a prime p,

$$c_{p^a} = 1 + \chi(p) + \chi(p)^2 + \cdots + \chi(p)^a \geq 0.$$

The equality follows from the multiplicative property of the character χ. As χ is a real character, for any prime p the value of $\chi(p)$ is either 1, -1 or 0. Therefore,

$$c_{p^a} = \begin{cases} 1, 1 + a \text{ or } 0 & \text{when } a \text{ is odd} \\ 1 \text{ or } 1 + a & \text{when } a \text{ is even.} \end{cases}$$

For any distinct primes p and q and finite, non-negative $a, b \in \mathbb{R}$, it holds that $c_{p^a q^b} = c_{p^a} c_{q^b}$:

$$c_{p^a q^b} = \sum_{0 \le x \le a,\ 0 \le y \le b} \chi(p^x q^y) = \sum_{0 \le x \le a} \chi(p^x) \sum_{0 \le y \le b} \chi(q^y) = c_{p^a} c_{q^b}$$

Then for the prime factorization of any natural number $n = p_1^{a_1} p_2^{a_2} \dots p_k^{a_k}$,

$$c_n = c_{p_1^{a_1}} c_{p_2^{a_2}} \dots c_{p_k^{a_k}} \ge 0.$$

as all the values in this product are not negative. Furthermore, if p is a prime dividing m then $\chi(p) = 0$. Then for a power of such p,

$$c_{p^a} = 1 + \chi(p) + \chi(p)^2 + \dots + \chi(p)^a = 1.$$

For $m \ge 2$, there exists at least one prime p that divides m and therefore for such a prime p,

$$\lim_{N \to \infty} \sum_{n=1}^{N} c_n \ge \lim_{K \to \infty} \sum_{a=1}^{K} c_{p^a} = \lim_{K \to \infty} \sum_{a=1}^{K} 1.$$

It follows that the series diverges. For $m = 1$, the only character is the trivial character and divergence follows evidently for this case. As the series diverges for all positive integers m, the lemma is proven. ☼

Theorem 5. Let m be a positive integer. The function value $L(1, \chi) \neq 0$ for all real characters χ modulo m such that $\chi \neq 1$.

Proof. (The proof as done by Monsky)
Set f to be the following function:

$$f(t) = \sum_{n=1}^{\infty} \frac{\chi(n) t^n}{1 - t^n}.$$

This function converges absolutely for $t \in [0, 1)$: for $t = 0$, all terms are equal to 0 and convergence is evident. For $0 < t < 1$, we set

$$a_n := \left| \frac{\chi(n) t^n}{1 - t^n} \right| = \frac{t^n}{1 - t^n}.$$

If there exist exponents b_n such that a_n / b_n is bounded and $\sum b_n$ is convergent, then $\sum a_n$ must converge as well, this follows from the comparison test. For this, let $b_n = t^n$. Then,

$$\lim_{n \to \infty} \frac{a_n}{b_n} = \lim_{n \to \infty} \frac{1}{1 - t^n} = 1,$$

for all $0 < t < 1$. This shows that a_n / b_n is bounded and the series $\sum_1^{\infty} t^n$ is known to converge for $0 < t < 1$, as it is a geometric series. Therefore, the series

$$\sum_{n=1}^{\infty} \left| \frac{\chi(n) t^n}{1 - t^n} \right| = \sum_{n=1}^{\infty} \frac{t^n}{1 - t^n}$$

must also converge and the absolute convergence of $f(t)$ is proven. Then for $t \in [0, 1)$, we can use geometric series to write

$$\sum_{n=1}^{\infty} c_n t^n = \sum_{n} \sum_{d|n,\ d>0} \chi(d) t^n = \sum_{d} \sum_{k} \chi(d) t^{dk} = \sum_{d=1}^{\infty} \chi(d) \frac{t^d}{1 - t^d} = f(t).$$

For this, we need the series $\sum_1^\infty c_n t^n$ to be convergent. We can show it is by the root test for power series:

$$\lim_{n\to\infty} \sup \sqrt[n]{|c_n t^n|} = \lim_{n\to\infty} \sup t \sqrt[n]{c_n}$$

This root exists as $c_n \geq 0$ as was proven in Lemma 11. But $c_n = \sum \chi(d)$ over the positive divisors d of n. This sum is over no more than n divisors and all terms are at most 1, so the sum is not greater than n. Thus,

$$\lim_{n\to\infty} \sup \sqrt[n]{|c_n t^n|} \leq \lim_{n\to\infty} \sup t \sqrt[n]{n} \leq \sup t \leq 1,$$

and the convergence is shown. As a result of Lemma 11 and the equality to $\sum_1^\infty c_n t^n$, the function $f(t) \to \infty$ as $t \to 1^-$.

For contradiction, now suppose that $L(1,\chi) = \sum_1^\infty \chi(n)/n = 0$. This series is convergent for $R(s) > 0$ as was shown in Proposition 11. Then it holds that

$$-f(t) = \sum_{n=1}^{\infty} \chi(n) \left[\frac{1}{n(1-t)} - \frac{t^n}{1-t^n} \right],$$

which is then also convergent. We write this as $\sum_1^\infty \chi(n) b_n$ where b_n are functions of t. Then,

$$\begin{aligned}(1-t)(b_n - b_{n+1}) &= \frac{1}{n} - \frac{1}{n+1} - \frac{t^n}{1+t+\cdots+t^{n-1}} + \frac{t^{n+1}}{1+t+\cdots+t^n} \\ &= \frac{1}{n(n+1)} - \frac{t^n}{(1+t+\cdots+t^{n-1})(1+t+\cdots+t^n)}.\end{aligned}$$

By the inequality of the arithmetic and geometric means,

$$(1+t+\cdots+t^{n-1}) \geq n \left(1 \cdot t \cdot \cdots \cdot t^{n-1}\right)^{1/n} = n t^{(n-1)/2} \geq n t^{n/2}.$$

In the same way $(1+t+\cdots+t^n) \geq (n+1)t^{n/2}$. This, together with the equality from before, gives $b_n - b_{n+1} \geq 0$ and thus $b_1 \geq b_2 \geq b_3 \geq \ldots$.

It follows from Proposition 3 that $\sum_a^{a+m-1} \chi(n) = 0$ for $\chi \neq 1$ and any natural number a. Thus, the sum $\sum_1^N \chi(n)$ is bounded in absolute value by $\phi(m)$, independent of N. Together with the properties of b_n and the fact that $b_1 = 1$, this shows that

$$\left|\sum_{n=1}^{N} \chi(n) b_n\right| \leq |b_1| \left|\sum_{n=1}^{N} \chi(n)\right| \leq \phi(m), \qquad \text{for all } N \in \mathbb{N}.$$

Taking the limit as $N \to \infty$ shows that the function $f(t) = -\sum_1^\infty \chi(n) b_n$ is bounded, but before we have shown that $f(t) \to \infty$ when $t \to 1^-$. This is a contradiction and we must thus conclude that our assumption $L(1,\chi) = 0$ is false. The theorem is proven. ☼

4.3 B. Veklych on real characters

Lemma 12. Let χ be a real character modulo m, where m is a positive integer. Define $c_n = \sum \chi(d)$, where the sum is taken over all positive divisors d of n. Then, the series $\sum_{n=1}^\infty c_n n^{-1/2}$ diverges.

Proof. (A combination of the proofs by Monsky and Veklych)
In the proof of Lemma 11 that for a prime factorization of any natural number $n = p_1^{a_1} p_2^{a_2} \ldots p_k^{a_k}$ it holds that

$$c_n = c_{p_1^{a_1}} c_{p_2^{a_2}} \ldots c_{p_k^{a_k}} \geq 0.$$

Specifically, for a square n all powers $a_1, a_2, \ldots, a_k$ of the prime factorization are even and so $c_{p_1^{a_1}} \geq 1, c_{p_2^{a_2}} \geq 1, \ldots$. As a result, also $c_n \geq 1$. Then, for $N = M^2$,

$$\sum_{n=1}^{N} \frac{c_n}{n^{1/2}} \geq \sum_{m=1}^{M} \frac{c_{m^2}}{m} \geq \sum_{m=1}^{M} \frac{1}{m}.$$

Taking the limit as $M \to \infty$, the last series becomes the harmonic series, of which divergence is a well-known property. The series $\sum_{n=1}^{\infty} c_n n^{-1/2}$ thus also diverges and the lemma is proven. ☼

Theorem 6. Let m be a positive integer. The function value $L(1, \chi) \neq 0$ for all real characters χ modulo m such that $\chi \neq 1$.

Proof. (The proof as done by Veklych) Consider the function $F(s) = \zeta(s)L(s, \chi)$. The zeta function is holomorphic for $R(s) > 0$ besides a simple pole at $s = 1$, we showed this in Corollary 3 to Theorem 2. Assume that $L(1, \chi) = 0$. The expansion of $L(s, \chi) = 0$ at $s = 1$ cancels the pole of $\zeta(s)$ and so, the function $F(s)$ is holomorphic for all $R(s) > 0$. Specifically, this $F(s)$ is holomorphic in a disc centered at $s = 2$ that contains $s = 1/2$. Equivalently, the function $F(2 - s)$ is holomorphic in a disc centered at $s = 0$ that contains $s = 3/2$.

For $R(s) > 1$, we have the absolutely convergent Dirichlet series of $\zeta(s) = \sum_k k^{-s}$ and $L(s, \chi) = \sum_l \chi(l) l^{-s}$. Multiplying these, we obtain a Dirichlet series that represents $F(s)$ for at least $R(s) > 1$. This is the series $\sum_n c_n n^{-s}$, where $c_n = \sum_d \chi(d)$ as in Lemma 12. This equality can be shown in the following way:

$$\sum_{n=1}^{\infty} c^{-n} n^{-s} = \sum_{n} \sum_{d|n,\ d>0} \chi(d) n^{-s} = \sum_{d} \sum_{k} \chi(d)(dk)^{-s} = \sum_{d} \chi(d) d^{-s} \sum_{k} k^{-s}.$$

Now let $s \in [0, 1)$. Then $F(2 - s)$ is represented by

$$\sum_{n=1}^{\infty} \frac{c_n}{n^{2-s}} = \sum_{n} \frac{c_n}{n^2} \left(e^{s \log n} - 1\right) + \sum_{n} \frac{c_n}{n^2}.$$

Due to the expansion of $e^{s \log n}$, this is equal to

$$\sum_{n} \frac{c_n}{n^2} \sum_{k=1}^{\infty} \frac{s^k (\log n)^k}{k!} + F(2) = \sum_{k} s^k \sum_{n} \frac{c_n}{n^2} \frac{(\log n)^k}{k!} + F(2).$$

The change in summation order is permitted since all terms of the double series are non-negative as $s \in [0, 1)$. We have now found a power series in s that converges to $F(2 - s)$ for $s \in [0, 1)$. The radius of convergence of this power series must thus be at least 1 and so it defines a function holomorphic in $|s| < 1$. In this open disc, the power series coincides with $F(2 - s)$ as analytic continuations are unique. The series is thus the power series expansion of $F(2 - s)$ around 0. As a result, it must converge to $F(2 - s)$ in the whole disc around 0 where $F(2 - s)$ is holomorphic. We have shown that this disc contains $s = 3/2$ before. Therefore, we set s equal to $3/2$ in the power series and read the formula backwards. We then see that $\sum_1^{\infty} c_n n^{-1/2}$ converges to $F(1/2)$.

But by Lemma 12, the series $\sum_1^{\infty} c_n n^{-1/2}$ diverges already. This is a contradiction and we must therefore conclude that our assumption was false. Hence, it holds that $L(1, \chi) \neq 0$ for all real characters $\chi \neq 1$. ☼

5 A comparison between proofs

Ram Murty / Cohn on non-real characters

The argument on non-real characters is not a hard proof. It could possibly be made even easier, but the combination of work by Ram Murty and Cohn makes for a smooth addition to our earlier proof. What makes this proof so easy relatively, is the existence of a distinct character $\overline{\chi}$ for any character χ that is not real-valued. Both these series then vanish at $s = 1$ by assumption and we can subtract a factor $(s-1)$ from both their Taylor Series. As a result of this, there is no need for complicated arguments about converging and diverging series. The product of L-functions can just be shown to yield 0 as $s \to 1$ if we assume that $L(1,\chi) = 0$ for some non-real χ. This does not agree with the value of the corresponding Dirichlet series and a contradiction is quickly found.

This proof is simple and it would work well with the rest of the proof by Serre. We do need to combine it with either the proof by Monsky or by Veklych to be able to substitute out Theorem 2 of Serre's proof. I will include my opinion on the combination of the proofs with this one in the respective texts on the proofs by Monsky and Veklych.

Monsky on real characters

Recommended audience: readers with little knowledge of Complex Analysis, students of an introductory course in Number Theory.

The proof by Monsky is different from Serre and Veklych in that it does not rely on a product of L-functions. Monsky introduces a different function based on a character and its L-function. The argument thus does not have to do with the trivial character and the simple pole of its L-function. Instead, he introduces a series f. This f can be shown to diverge by a clever use of sums of characters. The divergence can only be shown in this way for real characters, which is a disadvantage of the argument. Afterwards, Monsky shows that f is bounded. Here the vanishing L-function comes in. He shows this boundedness in a quick and rather easy way, but some of the used steps are not very intuitive. The combination of divergence and boundedness then results in a contradiction, which proves the theorem.

This argument is understandable even for those who have less knowledge of L-functions and Complex Analysis. This makes it suitable for a classroom, but it has its downsides. As a result of the simplifications, it does not fit in too well with the rest of a proof for Dirichlet's Theorem. Although Monsky does not use so many properties of the L-function and results from Complex Analysis, they are used in the rest of such a proof. That means that the proof by Monsky is suitable for showing only this result in an easy and understandable way, but it might not be the best choice in a proof of the whole theorem. A last concern with this proof is that, although not specifically named, results from various fields are still used implicitly. This could give problems for less experienced readers, as they are not used to working with these results yet. An example of this is working with a series as the one defined as $f(t)$, which has to be shown convergent first for the given domain. The importance of a step like this could not be too clear for a reader. Summarizing, the proof by Monsky is a great option for readers who are merely interested in this step of the proof, they should just make sure that they do not simplify it too much for themselves.

Veklych on real characters

Recommended audience: readers with knowledge of and interest in Complex Analysis, students of a course in Complex Analysis.

The beginning of the proof by Veklych is similar to the beginning of Monsky's proof. Both use the same sums over characters to show a form of divergence. After this, the proofs go different ways. Veklych wants to contradict this form of divergence. To do this, he introduces a power series expansion that converges. Using results from Complex Analysis, like the uniqueness of analytic continuations, he expands the radius of convergence of this expansion. For this, he does need the fact that $L(1,\chi) = 0$ for some $\chi \neq 1$. Now that the convergence is extended, it can be used to contradict the earlier named divergence and the proof is concluded.

This proof relies a lot on Complex Analysis. This does fit in well with the rest of the proof, but can be difficult to understand for readers new to the field. The results that are used, are implemented in a clear way. This way, the proof would fit perfectly in a course on Complex Analysis. It could be a good addition to any course in the matter to peak an interest of Number Theory. The proof is clever, but the used power series expansion is not too easy to grasp. It is quite a long expansion which can cause confusion. The proof does tie in well with the use of a product of L-functions by Serre. This is also used in the proof on characters that are not real-valued and so it would fit in better than the proof by Monsky in my opinion. If the reader is not too familiar with Complex Analysis, the proof can be difficult. Otherwise, this is a proof that fits in very well and provides a good alternative, especially in a context of more complex theory.

Serre on all characters

Recommended audience: devoted readers.

Serre proves the non-vanishing of the L-function for all characters $\chi \neq 1$ at once. To show that the product over all L-functions diverges, he shows the divergence of another series that is dominated by this one. Afterwards, he shows the convergence of the product at the same point. He does this by cancelling the simple pole of the L-function of the trivial character at $s = 1$. By removing this singularity, the convergence of the product of L-functions can be extended to the point where it also diverges. This gives the required contradiction.

The proof by Serre is more broad than all other arguments that we include. It contains a proof for real and non-real characters, but it is even broad in another way: the propositions that Serre introduces to prove this theorem are also applicable to other cases. Proposition 7 for example, states a result about all Dirichlet series with positive coefficients instead of just stating it for the specific series. This is a strength of Serre's proof, but it does result in more complicated mathematics. The argument is long in comparison to the new arguments we have introduced, and more knowledge on complex functions is needed to understand all the steps. In showing the divergence, Serre uses an approach that deals with all characters at once. This does result in a more complicated argument in which more Group Theory is used. The proof by Serre is elegant and it carries the spirit of Dirichlet's own approach with it. Understanding this, does require a more devoted study than the other provided proofs do. For a reader that has the time and knowledge to do so, this proof is the one I would suggest. For an audience looking for a shorter or more modern proof, this proof is best left for later.

References

1. Jeremy Avigad and Rebecca Morris. The concept of “character” in dirichlet’s theorem on primes in an arithmetic progression. *Archive for history of exact sciences*, 68(3):265–326, 2014.

2. H. Cohn. *Advanced Number Theory*. Dover Books on Mathematics Series. Dover Publications, 1980.

3. Paul Monsky, Hiroshi Maehara, and Bruce Anderson. Notes. *The American Mathematical Monthly*, 100(9):861–866, 1993.

4. M.R. Murty, S.J. Axler, F.W. Gehring, and K.A. Ribet. *Problems in Analytic Number Theory*. Graduate texts in mathematics. Springer, 2001.

5. J.P. Serre. *A course in arithmetic*. Graduate texts in mathematics. Springer, 1973.

6. Bogdan Veklych. A one-formula proof of the nonvanishing of l-functions of real characters at 1. *The American Mathematical Monthly*, 122(5):484–485, 2015.