

ELEMENTARY PROOF OF DIRICHLET THEOREM

ZIJIAN WANG

ABSTRACT. In this expository paper, we present the Dirichlet Theorem on primes in arithmetic progressions along with an elementary proof. We first show that $L(1, \chi) \neq 0$ for all non-principal characters and then use the basic properties of the Dirichlet series and Dirichlet characters to complete our proof of the Dirichlet Theorem.

CONTENTS

1. Introduction	1
2. Dirichlet Characters	2
3. Dirichlet series	4
4. The non-vanishing property of $L(1, \chi)$	9
5. Elementary proof of the Dirichlet Theorem	12
Acknowledgments	13
References	13

1. INTRODUCTION

It is undeniable that complex analysis is an extremely efficient tool in the field of analytic number theory. However, mathematicians never stop pursuing so-called "elementary proofs" of theorems which can be proved using complex analysis, since these elementary proofs can provide new levels of insight and further reveal the beauty of number theory. The famed analyst G.H. Hardy made a strong claim about the prime number theorem: he claimed that an elementary proof could not exist. Hardy believed that the proof of the prime number theorem used complex analysis (in the form of a contour integral) in an indispensable way. However, in 1948, Atle Selberg and Paul Erdős both presented elementary proofs of the prime number theorem. Unfortunately, this set off a major accreditation controversy around whose name should be attached to the proof, further underscoring the significance of an elementary proof of the prime number theorem. The beauty of that proof lies in the key formula of Selberg $\sum_{p \leq x} (\log p)^2 + \sum_{pq \leq x} (\log p)(\log q) = 2x \log x + O(x)$. Similarly, the brilliance of Dirichlet's proof of his theorem on primes in arithmetic progression lies in his creation of Dirichlet characters, which are used to detect primes in arithmetic progressions. In this expository paper, we follow Dirichlet's original proof except for his argument for the nonvanishing of $L(1, \chi)$. His original proof of the nonvanishing uses his class number formula. Here we provide a more

Date: AUGUST 28, 2017.

analytic proof of this result, following Iwaniec and Kowalski. The only algebraic ingredient in our proof is the orthogonality relations derived from Schur's lemma.

2. DIRICHLET CHARACTERS

Dirichlet characters are functions that act as the coefficients of a Dirichlet series. They not only provide us with the language to describe the Dirichlet series, but also turn out to be a useful tool with the help of the Schur orthogonality relations. Instead of providing a formal proof of the orthogonality relations, we will introduce a few applications of the Schur orthogonality relations which are relevant to our discussion of Dirichlet characters.

Definition 2.1. Given a finite abelian group G , a **character** is a homomorphism $\chi : G \rightarrow \mathbb{C}$.

Example 2.2. Given any cyclic group G of size k with generator g , we always have the character χ defined by $\chi(g^t) = e^{\frac{2\pi it}{k}}$. Actually any character defined on this cyclic group must be in the form of $\chi(g^t) = e^{\frac{2\pi i t t}{k}}$ since the character is a homomorphism.

Example 2.3. The principal character χ_0 is the trivial homomorphism such that

$$\chi_0(t) = 1, \forall t \in G.$$

Remark 2.4. We can define at least one character for any group, namely the principal character. However, there is a type of "characters" that are extremely useful to our discussion. We now introduce the Dirichlet characters.

Definition 2.5. Dirichlet Characters

A **Dirichlet character** χ is a function from integers \mathbb{N} to \mathbb{C} that satisfies the following conditions:

1. There exists a positive integer k such that $\chi(n) = \chi(n+k)$ for all n .
2. If $(n, k)^1 > 1$ then $\chi(n) = 0$; if $(n, k) = 1$ then $\chi(n) \neq 0$.
3. $\chi(mn) = \chi(m)\chi(n)$ for all integers m and n .

Remark 2.6. Dirichlet characters are actually not characters. However, they are deeply related to the characters on certain groups. For every period k , we have a canonical "trivial character mod k " defined by:

$$\chi(a) = \begin{cases} 1 & (a, k) = 1. \\ 0 & \text{otherwise.} \end{cases}$$

We call this **the principal character** of period k^2 . Moreover, one may notice that every Dirichlet character χ is built upon a character χ' on the group $\mathbb{Z}/k\mathbb{Z}^*$ for some integer k by the following construction:

$$\chi(a) = \begin{cases} \chi'(b) & (a, b) = 1, a \equiv b \pmod{k}. \\ 0 & \text{otherwise.} \end{cases}$$

Since Dirichlet characters are periodic, we usually think of them as characters on $\mathbb{Z}/k\mathbb{Z}^*$ in order to use the algebraic tools.

¹ (a, b) is the greatest common denominator of integers a and b .

²Notice that here the principal character is not a character. It is just an periodic arithmetic function

³ b is the remainder of a divided by k .

Definition 2.7. $\bar{\chi}$ is the conjugate of χ if for all $g \in G$, $\bar{\chi}(g) = \overline{\chi(g)}$.

Remark 2.8. The conjugate of a character is also a character since complex conjugation is a multiplicative function.

Definition 2.9. The dual group \hat{G} is the set of all the characters over G . It is a group under the following group laws:

- a. χ_0 is the identity.
- b. $\forall g \in G, (\chi_1\chi_2)(g) := \chi_1(g)\chi_2(g)$.
- c. $\forall g \in G, \chi^{-1}(g) := (\chi(g))^{-1}$.

Example 2.10. If G is cyclic, then $G \simeq \hat{G}$. This is a direct result of (2.2).

The orthogonality relation is one of the most useful algebraic gadgets when we are dealing with Dirichlet characters. We start with the original statement of the Schur Orthogonality relation which works for any group.

Theorem 2.11. Schur orthogonality relations

Given a group G , C_G^g the centralizer of g , $h \in G$ and $\chi_i, \chi_j, \chi \in \hat{G}$, we have the following:

$$\begin{aligned}
 \text{a. } \sum_{g \in G} \chi_i(g) \overline{\chi_j(g)} &= \begin{cases} |G| & i = j. \\ 0 & i \neq j. \end{cases} \\
 \text{b. } \sum_{\chi \in \hat{G}} \chi(g) \chi(h) &= \begin{cases} |C_G^g| & g \text{ and } h \text{ are conjugate.} \\ 0 & \text{otherwise.} \end{cases}
 \end{aligned}$$

In this paper, G is considered to be abelian, so we can simply replace $|C_G^g|$ with $|G|$. We refine the orthogonality relation by setting j to be 0^4 and h to be the identity.

Lemma 2.12. Given group G and its dual \hat{G} , we have the following:

$$\begin{aligned}
 \text{a. } \sum_{x \in G} \chi(x) &= \begin{cases} |G| & \chi = \chi_0. \\ 0 & \text{otherwise.} \end{cases} \\
 \text{b. } \sum_{\chi \in \hat{G}} \chi(x) &= \begin{cases} |\hat{G}| & x = 1. \\ 0 & \text{otherwise.} \end{cases}
 \end{aligned}$$

Note that $|\mathbb{Z}/k\mathbb{Z}^*| = |\mathbb{Z}/\hat{k}\mathbb{Z}^*| = \varphi(k)$. Now we consider $\mathbb{Z}/k\mathbb{Z}^*$ and the corresponding Dirichlet characters. We have the following lemma.

Lemma 2.13. For any character χ on $G = \mathbb{Z}/k\mathbb{Z}^*$, we have:

$$\begin{aligned}
 \text{a. } \sum_{l \in G} \chi(l) &= \begin{cases} \varphi(k) & \chi = \chi_0. \\ 0 & \text{otherwise.} \end{cases} \\
 \text{b. } \sum_{\chi \in \hat{G}} \chi(l) &= \begin{cases} \varphi(k) & l \equiv 1 \pmod{k}. \\ 0 & \text{otherwise.} \end{cases}
 \end{aligned}$$

⁴Here χ_0 denotes the principal character i.e. the trivial homomorphism from G to \mathbb{C} .

Recall that every Dirichlet character χ corresponds to a character χ' on $\mathbb{Z}/k\mathbb{Z}^*$. We can deduce that the sum of χ on any interval of finite length is bounded by a constant if $\chi \neq \chi_0$ since the sum of $\sum_{i=a}^{a+k-1} \chi(i) = \sum_{i=1}^k \chi(i) = \sum_{i=1}^k \chi'(i) = 0$ by the orthogonality relations of χ' . This result is stated in the following lemma.

Lemma 2.14. *Given $a, b \in \mathbb{N}$ and $a < b$, we have*

$$\sum_{n=a}^b \chi(n) = O(1).^5$$

3. DIRICHLET SERIES

This section provides a brief introduction to Dirichlet series and Dirichlet L-function. We will assume certain analytic properties such as the convergence of the Dirichlet L-function on the right half plane. Our main focus will be the different arithmetic functions that emerges from Dirichlet functions including the Möbius function μ and the Von Mangoldt function Λ . We will also delve a bit into Dirichlet convolution as well as useful techniques such as the Möbius inversion.

Definition 3.1. Dirichlet series

Given arithmetic function f , $D_f(s) = \sum_{n=1}^{\infty} (n)n^{-s}$ is a **Dirichlet series**. The **Riemann zeta function** is a special case of the Dirichlet series where all the weights are set to 1.

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

When $s \in (0, 1)$,

$$\zeta(s) = \lim_{x \rightarrow \infty} \sum_{n \leq x} (n^{-s} \frac{x^{1-s}}{1-s}).$$

Remark 3.2. Generally speaking, $s \in \mathbb{C}$ and D_f converge for any Dirichlet character f when $Re(s) > 1$.

Proposition 3.3. *$L(s, \chi)$ converge for $s > 0$ and non-principal character χ*

There are some useful asymptotic formulas related to the Riemann zeta function as defined above.

Theorem 3.4.

$$\sum_{n \leq x} n^{-s} = \frac{x^{1-s}}{1-s} + \zeta(s) + O(x^{-s}) \quad (s > 0, s \neq 1)$$

This theorem is a direct application of the following lemma.

Lemma 3.5. Euler summation formula

$f \in \mathbb{C}^1$ on $[a, b]$ with $0 < a < b$, we have

$$\sum_{a < n \leq b} f(n) = \int_a^b f(x) dx + \int_a^b (x - [x]) f'(x) dx + f(b)([b] - b) - f(a)([a] - a)$$

6

⁵For any two functions f and g , $f(x) = O(g(x))$ if and only if $|f(x)| \leq M|g(x)|$ for some constant M and all large x .

⁶ $[\]$ is the floor function. $[x]$ is the largest integer smaller or equal to x .

Proof. Note that $[x]$ is hard to deal with generally, but is easy to simplify if we confine x in some small and nice intervals.

$$\begin{aligned}\int_{[a]}^a [x]f'(x)dx &= [a] \int_{[a]}^a f'(x)dx = [a](f(a) - f[a]) \\ \int_{[b]}^b [x]f'(x)dx &= [b] \int_{[b]}^b f'(x)dx = [b](f(b) - f[b])\end{aligned}$$

More generally, for any integer n such that $a < n - 1 < n < b$, we have

$$\begin{aligned}\int_{n-1}^n [x]f'(x)dx &= \int_{n-1}^n (n-1)f'(x)dx \\ &= (n-1)(f(n) - f(n-1)) \\ &= nf(n) - (n-1)f(n-1) - f(n)\end{aligned}$$

We can now calculate $\int_{[a]}^{[b]} [x]f'(x)dx$.

$$\begin{aligned}\int_{[a]}^{[b]} [x]f'(x)dx &= \sum_{n=[a]+1}^{[b]} \int_{n-1}^n [x]f'(x)dx \\ &= \sum_{n=[a]+1}^{[b]} nf(n) - (n-1)f(n-1) - f(n) \\ &= [b]f[b] - [a]f[a] - \sum_{a < n \leq b} f(n)\end{aligned}$$

$$(6) \quad \sum_{a < n \leq b} f(n) = - \int_{[a]}^{[b]} [x]f'(x)dx + [b]f[b] - [a]f[a]$$

$$(7) \quad = - \int_a^b [x]f'(x)dx + [b]f(b) - [a]f(a)$$

We know the identity

$$(8) \quad \int_a^b f(x)dx = bf(b) - af(a) - \int_a^b xf(x)dx.$$

By adding (6) and (8), we finish the proof of the lemma. \square

Proof of 3.4. By applying the Euler summation formula to $f(n) = n^{-s}$ with $a = 1$ and $b = x$, we have

$$\begin{aligned}\sum_{n \leq x} n^{-s} &= 1 + \sum_{1 < n \leq x} n^{-s} \\ &= 1 + \int_1^x t^{-s}dt + \int_1^x (-s)(t - [t])t^{-s-1}dt + x^{-s}([x] - x) - 1^{-s}([1] - 1) \\ &= 1 + \frac{x^{1-s}}{1-s} - \frac{1}{1-s} - s \int_1^\infty (t - [t])t^{-s-1}dt + O(x^{-s}) \\ &= \frac{x^{1-s}}{1-s} + C_s + O(x^{-s})\end{aligned}$$

$$\text{where } C_s = 1 - \frac{1}{1-s} - s \int_1^\infty (t - [t])t^{-s-1} dt.$$

Now we compare C_s with the Riemann zeta function $\zeta(s)$ with $x \rightarrow \infty$.
If $s > 1$,

$$\zeta(s) = \lim_{x \rightarrow \infty} \sum_{n \leq x} n^{-s} = 0 + C_s + 0.$$

If $s \in (0, 1)$, by (3.23),

$$C_s = \lim_{x \rightarrow \infty} \left(\sum_{n \leq x} n^{-s} - \frac{x^{1-s}}{1-s} \right) = \zeta(s).$$

□

Proposition 3.9. Euler Product

If f is multiplicative, we have

$$D_f(s) = \prod_p \left(1 + \sum_{n=1}^{\infty} f(p^n) p^{-ns} \right),$$

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1}.$$

Proof. The Euler product is a direct consequence of the unique prime factorization of integers. □

Definition 3.10. Dirichlet Convolution

Given arithmetic functions f and g , the **Dirichlet convolution** of f and g $f * g$ is defined by

$$f * g = \sum_{d|n} f(d)g\left(\frac{n}{d}\right).$$

Remark 3.11. The set of arithmetic functions actually forms a ring under canonical addition and Dirichlet convolution. More specifically, if we multiply two Dirichlet series D_f and D_g , we get a new Dirichlet series D_h where the arithmetic function h is equivalent to the convolution of f and g . This result is stated as the following proposition.

Proposition 3.12.

$$D_{f*g} = D_f D_g$$

Definition 3.13. Möbius function

The **Möbius function** $\mu : \mathbb{N}^* \rightarrow \{0, -1, 1\}$ is defined by

$$\mu(m) = \begin{cases} (-1)^t & \text{if } m \text{ is the product of } t \text{ distinct primes.} \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 3.14. For $n \in \mathbb{Z}$,

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1. \\ 0 & \text{if } n > 1. \end{cases}$$

Proof.

If $n = 1$, $\sum_{d|1} \mu(d) = \mu(1) = 1$.

If $n > 1$, suppose $n = \prod_{i=1}^r p_i^{\alpha_i}$, we have

$$\begin{aligned} \sum_{d|n} \mu(d) &= \sum_{k=1}^r r(-1)^k \binom{r}{k} \\ &= (1-1)^r \\ &= 0. \end{aligned}$$

□

The Möbius function appears as the coefficient of the Dirichlet series obtained by taking the reciprocal of the Riemann zeta function. We now prove this result in the following proposition.

Proposition 3.15.

$$\zeta^{-1} = D_\mu^7$$

Proof. Using the Euler product of zeta function, we have

$$\begin{aligned} \zeta^{-1}(s) &= \left(\sum_{n=1}^{\infty} n^{-s} \right)^{-1} \\ &= \prod_p (1 - p^{-s}) \end{aligned}$$

1. If $p^2 \mid n$, then $p^{2s} \mid n^s$. Therefore, the coefficient of n^{-s} must be 0.

2. If n is the product of t distinct primes, coefficient of such n^{-s} is $(-1)^t$.

Therefore the weight function of the Dirichlet series generated by the reciprocal of Riemann zeta function is the same as that of D_μ . □

One of the most well-known results on arithmetic functions is the Möbius Inversion, which can be verified by direct calculation.

Proposition 3.16. Möbius Inversion

Given arithmetic functions $f, g : \mathbb{N} \rightarrow \mathbb{C}$, we have

$$g(n) = \sum_{d|n} f(d) \Leftrightarrow f(n) = \sum_{d|n} \mu(d)g\left(\frac{n}{d}\right).$$

Definition 3.17. von Mangoldt function

The **von Mangoldt** function Λ is defined as

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^t, t \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

Proposition 3.18.

$$-(\zeta)' \zeta^{-1} = D_\Lambda^8$$

Proof. It is equivalent to prove $D_\Lambda \zeta^{-1} = -(\zeta)'$. Let $D_h = D_\Lambda \zeta^{-1}$. We need to show h is the same as \log . $\forall n$ with prime factorization $\prod_{i=1}^r p_i^{\alpha_i}$, for each i , we have

⁷Here ζ^{-1} is the reciprocal of the Riemann zeta function.

⁸ $(\zeta)'$ is the derivative of the Riemann zeta function.

$$n = p_1^{\alpha_1} \left(\frac{n}{p_1}\right) = p_1^2 \left(\frac{n}{p_1^2}\right) = p_1^3 \left(\frac{n}{p_1^3}\right) \dots = p_i^{\alpha_i} \left(\frac{n}{p_i^{\alpha_i}}\right).$$

$$\begin{aligned} h(n) &= \sum_{i=1}^r \sum_{j=1}^{\alpha_i} \Lambda(p_i^j) \\ &= \sum_{i=1}^r \alpha_i \log p_i \\ &= \sum_{i=1}^r p_i^{\alpha_i} \\ &= \log n. \end{aligned}$$

□

Proposition 3.19.

$$\log n = \sum_{d|n} \Lambda(d)$$

Proof. By the previous proposition we know that $\Lambda = \mu * \log$, in other words,

$$\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d}$$

By Möbius inversion, we have $\log n = \sum_{d|n} \Lambda(d)$ □

At the end of this section, we introduce the Dirichlet L function, which will later replace Dirichlet series.

Definition 3.20. Dirichlet L-function

$$L(s, \chi) = D_\chi(s)$$

where χ is the Dirichlet character defined in Section 2.

Remark 3.21. Complex analysis is not the focus of this paper, so here we only provide a simple and intuitive definition of the L-function instead of a full and rigorous construction.

If χ is multiplicative, we have the Euler product for L-function

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

When χ is the principal character χ_0 of mod k , we can factor out the Riemann zeta function and have

$$L(s, \chi_0) = \zeta(s) \prod_{p|k} (1 - p^{-s}).$$

We now prove some useful asymptotic properties on the L-function.

Proposition 3.22.

$$\sum_{k \leq \frac{x}{d}} \chi(k) k^{-1} = L(1, \chi) + O\left(\frac{d}{x}\right).$$

Proof.

$$\begin{aligned}
 |L(1, \chi) - \sum_{k \leq \frac{x}{d}} \chi(k)k^{-1}| &= |\sum_{k > \frac{x}{d}} \chi(k)k^{-1}| \\
 &= \frac{|\sum_{k=1}^{\infty} \chi(k)k^{-1}|}{\frac{x}{d}} \\
 &= \frac{d}{x} L(1, \chi) \\
 &= O\left(\frac{d}{x}\right) \text{ since } L(1, \chi) \text{ converges.}
 \end{aligned}$$

□

4. THE NON-VANISHING PROPERTY OF $L(1, \chi)$

In this section we prove that $L(1, \chi) \neq 0$ for all non-principle characters. This turns out to be the key of our proof of the Dirichlet Theorem. We do not follow the original strategy of Dirichlet, which uses the Class Number Formula. Instead, we use an elementary method⁹.

Theorem 4.1.

$$\forall \chi \neq \chi_0, L(1, \chi) \neq 0.$$

Proposition 4.2.

$$a. \text{ If } L(1, \chi) \neq 0, \text{ then } \sum_{l \leq x} \chi(l)\Lambda(l)l^{-1} = O(1).$$

$$b. \text{ If } L(1, \chi) = 0, \text{ then } \sum_{l \leq x} \chi(l)\Lambda(l)l^{-1} = -\log x + O(1).$$

Proof of a. We estimate the series $\sum_{n \leq x} \chi(n)n^{-1} \log n$ using the formulas we derived in previous sections.

Since χ is multiplicative, for $l|n$ we have

$$\begin{aligned}
 \chi(n)n^{-1}\Lambda(l) &= \chi\left(\frac{n}{l}\right)\chi(l)\left(\frac{n}{l}\right)^{-1}l^{-1}\Lambda(l) \\
 &= \chi(l)\Lambda(l)l^{-1}\chi\left(\frac{n}{l}\right)\left(\frac{n}{l}\right)^{-1}.
 \end{aligned}$$

Using this, we can deduce that

$$\begin{aligned}
 \sum_{n \leq x} \chi(n)n^{-1} \log n &= \sum_{n \leq x} (\chi(n)n^{-1} \sum_{l|n} \Lambda(l)) \\
 &= \sum_{l \leq x} (\chi(l)\Lambda(l)l^{-1} \sum_{m \leq \frac{x}{l}} \chi(m)m^{-1}).
 \end{aligned}$$

By (3.22) we can replace $\sum_{m \leq \frac{x}{l}} \chi(m)m^{-1}$. Therefore,

$$\begin{aligned}
 \sum_{l \leq x} \chi(l)\Lambda(l)l^{-1}(L(1, \chi) + O\left(\frac{l}{x}\right)) &= \sum_{l \leq x} \chi(l)\Lambda(l)l^{-1}L(1, \chi) + \sum_{l \leq x} \chi(l)\Lambda(l)l^{-1}O\left(\frac{l}{x}\right) \\
 &= L(1, \chi) \sum_{l \leq x} \chi(l)\Lambda(l)l^{-1} + O(1).
 \end{aligned}$$

⁹We follow the idea in the book Analytic Number theory by Iwaniec and Kowalski.

In all, we have

$$\sum_{n \leq x} \chi(n)n^{-1} \log n = L(1, \chi) \sum_{l \leq x} \chi(l)\Lambda(l)l^{-1} + O(1).$$

Since $\sum_{n \leq x} \chi(n)n^{-1} \log n$ and $L(1, \chi)$ are both bounded, we have $\sum_{l \leq x} \chi(l)\Lambda(l)l^{-1} = O(1)$. \square

Proof of b.

$$\begin{aligned} & \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) \log \frac{x}{d} - \sum_{n \leq x} \chi(n)\Lambda(n)n^{-1} \\ &= \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) \log \frac{x}{d} - \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) \log \frac{n}{d} \\ &= \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) \log \frac{x}{n} \\ &= \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) \log x - \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) \log n. \end{aligned}$$

Using (3.14), we have

$$\begin{aligned} & \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) \log x - \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) \log n \\ &= \log x \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) - \sum_{n \leq x} \chi(n)n^{-1} \log(n) \sum_{d|n} \mu(d) \\ &= \log x \chi(1)1^{-1} - \chi(1)1^{-1} \log(1) \\ &= \log x. \end{aligned}$$

Therefore,

$$\begin{aligned} \log x + \sum_{n \leq x} \chi(n)\Lambda(n)n^{-1} &= \sum_{n \leq x} \chi(n)n^{-1} \sum_{d|n} \mu(d) \log \frac{x}{d} \\ &= \sum_{d \leq x} \mu(d)\chi(d)d^{-1} \log \frac{x}{d} \sum_{l \leq \frac{x}{d}} \chi(l)l^{-1} \\ &= L(1, \chi) \sum_{d \leq x} \mu(d)\chi(d)d^{-1} \log \frac{x}{d} + O(1). \end{aligned}$$

By our assumption $L(1, \chi) = 0$, we know that

$$\log x + \sum_{n \leq x} \chi(n)\Lambda(n)n^{-1} = O(1).$$

In other words,

$$\sum_{l \leq x} \chi(l)\Lambda(l)l^{-1} = -\log x + O(1).$$

\square

Remark 4.3.

We can conclude (4.2) by the following.

$$\sum_{l \leq x} \chi(l) \Lambda(l) l^{-1} = -\lambda_\chi \log x + O(1),$$

where the coefficient λ_χ depends on χ

$$\lambda_\chi = \begin{cases} -1 & L(1, \chi) = 0 \text{ and } \chi \neq \chi_0. \\ 0 & L(1, \chi) \neq 0 \text{ and } \chi \neq \chi_0. \\ 1 & \chi = \chi_0. \end{cases}$$

Proof of 4.1.

Now we sum (4.3) over all characters.

By the orthogonality relations (2.13),

$$\begin{aligned} LHS &= \sum_{\chi \bmod k} \sum_{l \leq x} \chi(l) \Lambda(l) l^{-1} \\ &= \varphi(k) \sum_{l \leq x, l \equiv 1 \pmod k} \Lambda(l) l^{-1}. \end{aligned}$$

$$RHS = \sum_{\chi \bmod k} \lambda_\chi \log x + O(1).$$

We have

$$\varphi(k) \sum_{l \leq x, l \equiv 1 \pmod k} \Lambda(l) l^{-1} = \sum_{\chi \bmod k} \lambda_\chi \log x + O(1).$$

As a result, $\sum_{\chi \bmod k} \lambda_\chi$ must be nonnegative. Since $\lambda_\chi > 0$ only for the principal character, there is at most one non-principal character for which $L(1, \chi) = 0$. If there exist such non-principal character, then it must be real because $L(s, \bar{\chi}) = \overline{L(s, \chi)}$ for real s . Therefore, all we need to prove now is that real character χ s.t. $L(1, \chi) = 0$ does not exist.

Assume $L(1, \chi)$ is real. We consider the function

$$F(x) = \sum_{n \leq x} f_\chi(n) n^{-\frac{1}{2}}$$

where

$$f_\chi(n) = \sum_{d|n} \chi(d).$$

Since χ is multiplicative, given $n = \prod_{i=1}^r p_i^{\alpha_i}$, we know that

$$f_\chi(n) = \prod_{i=1}^r \sum_{j=1}^{\alpha_i} \chi(p_i^j).$$

Therefore, $f_\chi(n)$ is always non-negative. Notice that $f_\chi(m^2) \geq 1$. We have the following estimation of the lower bound for F ,

$$(4) \quad F(x) \geq \sum_{m^2 \leq \sqrt{x}} f_\chi(m^2) m \geq \sum_{m \leq \sqrt{x}} m^{-1} > \int_1^{\sqrt{x}} t^{-1} dt = \log \sqrt{x} = \frac{\log x}{2}.$$

On the other hand, we can split $F(x)$ into two parts

$$\begin{aligned} F(x) &= \sum_{ab \leq x} \chi(a)(ab)^{-\frac{1}{2}} \\ &= \sum_{a \leq \sqrt{x}} \chi(a)a^{-\frac{1}{2}} \sum_{b \leq \frac{x}{a}} b^{-\frac{1}{2}} + \sum_{b \leq \sqrt{x}} b^{-\frac{1}{2}} \sum_{\sqrt{x} < a < \frac{x}{b}} \chi(a)a^{-\frac{1}{2}}. \end{aligned}$$

Now using (3.4) and set $s = \frac{1}{2}$, we have

$$\begin{aligned} &\sum_{a \leq \sqrt{x}} \chi(a)a^{-\frac{1}{2}} \sum_{b \leq \frac{x}{a}} b^{-\frac{1}{2}} + \sum_{b \leq \sqrt{x}} b^{-\frac{1}{2}} \sum_{\sqrt{x} < a < \frac{x}{b}} \chi(a)a^{-\frac{1}{2}} \\ &= \sum_{a \leq \sqrt{x}} \chi(a)a^{-\frac{1}{2}} \left(2\left(\frac{x}{a}\right)^{-\frac{1}{2}} + c + O\left(\sqrt{\frac{m}{x}}\right)\right) + O\left(\sum_{n < \sqrt{x}} n^{-\frac{1}{2}}x^{-\frac{1}{4}}\right) \\ &= 2L(1, \chi)x^{-\frac{1}{2}} + O(1). \end{aligned}$$

Compare this with (4), we have

$$4L(1, \chi)x^{-\frac{1}{2}} + O(1) > \log x.$$

Since $\lim_{x \rightarrow \infty} \log x = \infty$, we can safely conclude from the above inequality that $L(1, \chi) > 0$ \square

Now we know that $\sum_{\chi} \lambda_{\chi} = 1^{10}$, so we have the following lemma

Lemma 4.5. *For all $x \in \mathbb{R}$, if $(l, k) = 1$, we have*

$$\sum_{n \leq x, n \equiv l \pmod{k}} \Lambda(n)n^{-1} = \varphi(k)^{-1} \log x + O(1)$$

5. ELEMENTARY PROOF OF THE DIRICHLET THEOREM

Finally we are able prove the Dirichlet Theorem in the way that Dirichlet proved it, using the Dirichlet characters and the Dirichlet L-function. The proof uses the non-vanishing result from the last section and does not involve any complex analysis.

Theorem 5.1 (Dirichlet Theorem). *There exist infinitely many primes $p \equiv l \pmod{k}$ where l is coprime to k .*

Proof. For simplicity, we define P_k^l to be the set of primes $p \equiv l \pmod{k}$. Our goal is to show that P_k^l is infinite if l and k are coprime.

We first recall the Euler product for Dirichlet L-function

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

Using the Taylor expansion for $\log \frac{1}{1-t}$, we have,

$$\log \frac{1}{1 - \chi(p)p^{-s}} = \chi(p)p^{-s} + O((\chi(p)p^{-s})^2).$$

If we take the logarithm of both sides, we would have

$$\log L(s, \chi) = \sum_p \chi(p)p^{-s} + \sum_p (\chi(p)p^{-s})^2.$$

¹⁰ λ_{χ} is defined in 4.3

Notice that $\sum_p (\chi(p)p^{-s})^2 = \sum_p \chi(p)^2 p^{-2s}$ converges. Therefore,

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

Observe that the range of χ is within S^1 where $\chi(n)^{-1} = \overline{\chi(n)}$.¹¹ By the orthogonality relations, we know that $\sum_{\chi \bmod k} \overline{\chi}(l)\chi(p)$ does not vanish if and only if $p \in P_k^l$.

$$\begin{aligned} \varphi(k)^{-1} \sum_{\chi \bmod k} \overline{\chi}(l) \log L(s, \chi) &= \varphi(k)^{-1} \sum_{\chi \bmod k} \overline{\chi}(l) \left(\sum_p \chi(p)p^{-s} + O(1) \right) \\ &= \varphi(k)^{-1} \sum_{\chi \bmod k} \overline{\chi}(l) \sum_p \chi(p)p^{-s} + \varphi(k)^{-1} \sum_{\chi \bmod k} \overline{\chi}(l) O(1) \\ &= \sum_{p \in P_k^l} p^{-s} + O(1). \end{aligned}$$

If χ is the principal character, recall that we can factor out the Riemann zeta function from the L-function.

$$L(s, \chi_0) = \zeta(s) \prod_{p|k} (1 - p^{-s})$$

where

$$\zeta(s) = \sum_1^\infty n^{-s} = \int_1^\infty t^{-s} dt + O(1) = (s - 1)^{-1} + O(1).$$

As a result,

$$\begin{aligned} \varphi(k)^{-1} \overline{\chi_0}(l) \log L(s, \chi_0) &= \varphi(k)^{-1} \log L(s, \chi) \\ &= \varphi(k)^{-1} \log(s - 1)^{-1} + O(1). \end{aligned}$$

Now we need to show the rest of the sum is bounded/finite in order to prove that our target series $\sum_{p \in P_k^l} p^{-s}$ diverge as $s \rightarrow 1^+$.

Recall that $\lim_{s \rightarrow 1^+} L(s, \chi)$ exists and does not vanish. Therefore $\lim_{s \rightarrow 1^+} \log L(s, \chi)$ is finite for non-principal characters, in other words, when $s \rightarrow 1^+$,

$$\sum_{p \in P_k^l} p^{-s} = \varphi(k)^{-1} \log(s - 1)^{-1} + O(1).$$

The series $\sum_{p \in P_k^l} p^{-s}$ diverges as $s \rightarrow 1^+$. In other words, that there are infinitely many primes of the form $p \equiv l \pmod k$. \square

Acknowledgments. It is a pleasure to thank my mentor Owen Barrett for his time and effort in guiding me on the topic and helping me understanding the material. I also want to thank Professor May for helping me with my paper writing skills and giving me advice on academic writing in mathematics.

REFERENCES

- [1] Henryk Iwaniec, Emmanuel Kowalski. Analytic Number Theory. American Mathematical Society. 2004.
- [2] Tom M. Apostol. Introduction to Analytic Number Theory. Springer. 1976.
- [3] Wikipedia topic on Schur Orthogonality Relations. https://en.wikipedia.org/wiki/Schur_orthogonality_relations.

¹¹If $z \in S^1$, then $z\bar{z} = 1$.