UNDERSTANDING SHAPIRO'S PROOF OF DIRCHLET'S THEOREM

ABHISRUTA MAITY

Abstract. In this brief report, we are going to present the proof of Dirichlet's theorem of infinitude of primes in arithmetic progression by Harold N. Shapiro (1950). In the first part, we briefly talk about the preliminary techniques and results which are essential for the proof. Subsequently, we finish the proof with constructing a logical flow between a few lemmas.

1. Introduction

Consider the sequence of prime numbers. Take any arbitrary natural number k, and create a new sequence of residues, dividing each member of the prime sequence by k. Except for the primes dividing k, we observe that each of the co-prime residues of k is occurring infinitely often in the sequence of residues — this simple observation, which was conjectured by Adrien-Marie Legendre (1785), remained unproven until the extraordinary paper by Lejeune-Dirichlet (27th of July, 1837) came into existence. In this paper, he formally introduced the techniques of analysis to solve such number theoretic problems and that's the emergence of a new field — Analytic Number Theory. In his proof he used completely new arithmetic functions, which are now known as Dirichlet characters χ and series involving it, Dirichlet L-functions ($L(s,\chi)$). We restate the theorem:

Theorem 1.1 (Dirichlet, 1837)
If $k \in \mathbb{N}$ and $\gcd(h,k) = 1$, then there are infinitely many primes of the form $mk + h$ ($m \in \mathbb{N} \cup \{0\}$), i.e., there are infinitely many primes in the arithmetic progression which includes h and having common difference k; in other words, there are infinitely many primes $p \equiv h \bmod k$.

Much later, Atle Selberg (1949) presented an elementary proof of the above theorem, which doesn't involve Dirichlet characters or Dirichlet L-functions and is restricted in finite sums only. But the bounds used in the proof are very non-trivial and therefore difficult to inculcate. In contrast, Harold N. Shapiro (1950) independently came up with another proof of the same theorem, which is elementary in the sense that although it's based on the series capturing Dirichlet characters, unlike any other classical proof, it doesn't encapsulate analysis of Dirichlet L-functions nearby $s = 1$. This is the proof we are going to present.

Part 1. Setting the Stage Up

The preliminary results which are the keys to unlock Theorem 1.1 are stated in this part of the report. Some of the important proofs are discussed.

Date: March, 2023.

UNDER THE SUPERVISION OF PROF. SOUMYA BHATTACHARYA, DEPARTMENT OF MATHEMATICS AND STATISTICS, IISER KOLKATA, MOHANPUR, WB: 741246, INDIA.

2. Generalized Möbius Inversion

With the Möbius inversion, we can recover arithmetic functions from a partial sum:

Theorem 2.1

Let α be a completely multiplicative function (i.e., $\alpha(xy) = \alpha(x)\alpha(y)$ for all $x, y \in \mathbb{N}$) with $\alpha(1) = 1$ and F be any complex-valued function, then if we define,

$$G(x) = \sum_{n \leq x} \alpha(n) F\left(\frac{x}{n}\right)$$

then we can conclude:

$$F(x) = \sum_{n \leq x} \mu(n)\alpha(n) G\left(\frac{x}{n}\right)$$

Proof. Passing the definition of G into the following expression:

$$\begin{aligned}
\sum_{n \leq x} \mu(n)\alpha(n) G\left(\frac{x}{n}\right) &= \sum_{n \leq x} \mu(n)\alpha(n) \sum_{m \leq x/n} \alpha(m) F\left(\frac{x}{mn}\right) \\
&= \sum_{n \leq x} \mu(n) \sum_{m \leq x/n} \alpha(n)\alpha(m) F\left(\frac{x}{mn}\right) \\
&= \sum_{n \leq x} \mu(n) \sum_{mn \leq x} \alpha(mn) F\left(\frac{x}{mn}\right) \\
&= \sum_{mn \leq x} \alpha(mn) F\left(\frac{x}{mn}\right) \sum_{n \leq x} \mu(n) \\
&= \sum_{d \leq x} \alpha(d) F\left(\frac{x}{d}\right) \sum_{n \mid d} \mu(n) = \alpha(1)F(x/1) = F(x).
\end{aligned}$$

$\square$

The heuristic which we will apply later is that if given F is a well-known function and so that the form of G turns out to be a well-studied partial sum, then plugging G back into the expression of F, we can estimate a potentially desired partial sum (see the proof of the Lemma 7.4 and 7.6).

3. Dirichlet Characters and Its Orthogonality Relations

For a positive integer k, consider the set of its reduced residue classes under $\bmod k$ operation. This set forms a finite abelian group of order $\varphi(k)$ under a suitably defined multiplication. A member of the group looks like:

$$[a] = \{x : x \equiv a \bmod k\}$$

where $(k, a) = 1$. And the multiplication is defined as:

$$[a][b] = [ab]$$

which naturally follows from modular arithmetic. Clearly the idenity element is [1] and the inverse of [a] is [b] such that $ab \equiv 1 \bmod k$. Now we are ready to move onto character theory of a finite abelian group.

Definition 3.1 (Character)
Let G be an arbitrary multiplicative group and $\mathbb{C}^\times$ be the multiplicative group of non-zero complex numbers. A homomorphism $f : G \to \mathbb{C}^\times$, i.e.,

$$f(ab) = f(a)f(b)$$

for all $a, b \in G$ and with the property that $f(c) \neq 0$ for some c, is called character of G.

Next we state some simple results without proof.

Theorem 3.2
If f is a character defined on a finite group G with an identity element e, then $f(e) = 1$ and each value of f is a root of unity.

Hint. Fermat's little theorem. □

Every group has at least one character, namely the *principal character*, where f is identically equal to 1. The following theorem shows the existence of other characters on a finite abelian group.

Theorem 3.3
A finite abelian group G of order n has exactly n distinct characters.

Suppose G is a finite abelian group of order n with elements $a_1, \ldots, a_n$. Let f_1 denote the principal character and $f_2, \ldots, f_n$ denote *non-principal* characters.

Theorem 3.4 (Orthogonality Relations for Characters)
We have,

$$\sum_{r=1}^{n} \bar{f}_r(a_i) f_r(a_j) = \begin{cases} n & \text{if } a_i = a_j \\ 0 & \text{if } a_i \neq a_j \end{cases}$$

Now let us work with the group G of reduced residue classes $\bmod k$. Let $f_1, \ldots, f_{\varphi(k)}$ denote the characters G has.

Definition 3.5 (Dirichlet Characters)
For each $i \in \{1, \ldots, \varphi(k)\}$, define $\chi_i : \mathbb{N} \to \mathbb{C}^\times$ such that:

$$\chi_i(n) = \begin{cases} f_i([n]) & \text{if } \gcd(n, k) = 1 \\ 0 & \text{if } \gcd(n, k) > 1 \end{cases}$$

The arithmetic functions χ_is are called Dirichlet characters $\bmod k$.

Immediately we recieve the following theorem:

Theorem 3.6

There are $\varphi(k)$ distinct Dirichlet characters mod$k$, each of them is completely multiplicative and periodic with period k.

Thus we can find the orthogonality relations of Dirichlet characters using Theorem 3.4:

Theorem 3.7 (Orthogonality Relations for Dirichlet Characters)

If $m, n \in \mathbb{N}$ we have,

$$\sum_{i=1}^{n} \bar{\chi}_i(m)\chi_i(n) = \begin{cases} \varphi(k) & \text{if } m \equiv n \bmod k \\ 0 & \text{if } m \not\equiv n \bmod k \end{cases}$$

The above theorem is extremely important for extracting terms from a partial sum. It will be our starting point for proving Theorem 1.1.

4. Partial Sums Involving Dirchlet Characters

The following theorem shows the guaranteed convergence of sums involving non-principal Dirichlet characters:

Theorem 4.1

Let χ be any non-principal character mod$k$ and f be a real-valued non-negative function with continuous negative derivative for all $x \geq x_0$. Then for all $y \geq x \geq x_0$ we have,

$$\sum_{x<n\leq y} \chi(n)f(n) = O(f(x))$$

Furthermore, if $f(x) \to 0$ as $x \to \infty$, then the infinite series:

$$\sum_{n=1}^{\infty} \chi(n)f(n)$$

converges and we estimate the partial sum as:

$$\sum_{n\leq x} \chi(n)f(n) = \sum_{n=1}^{\infty} \chi(n)f(n) + O(f(x))$$

for all $x \geq x_0$.

Corollary 4.2. For any non-principal χ mod k and $x \geq 1$ we have,

$$\sum_{n\leq x} \frac{\chi(n)}{n} = \sum_{n=1}^{\infty} \frac{\chi(n)}{n} + O\left(\frac{1}{x}\right)$$

$$\sum_{n\leq x} \frac{\chi(n)\log n}{n} = \sum_{n=1}^{\infty} \frac{\chi(n)\log n}{n} + O\left(\frac{\log x}{x}\right)$$

$$\sum_{n\leq x} \frac{\chi(n)}{\sqrt{n}} = \sum_{n=1}^{\infty} \frac{\chi(n)}{\sqrt{n}} + O\left(\frac{1}{\sqrt{x}}\right)$$

The proof of the first part of the above theorem is straightforward, uses the key fact:

$$A(x) = \sum_{n \leq x} \chi(n) = O(1)$$

and Abel's identity. The second part uses the Cauchy convergence criterion of an infinite series. These results will be used in the next section as well as the main proof.

5. Non-vanishing of $L(1, \chi)$ for Real Non-Principal χ

Consider $L(1, \chi)$ which is defined as:

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

which converges for all non-principal χ, by Theorem 4.1. In the proof of Theorem 1.1, we will require to show that $L(1, \chi) \neq 0$ for all non-principal χ. Here we will just show the non-vanishing of real non-principal χs by the virtue of the following theorem:

Theorem 5.1

For any real-valued non-principal character χ mod k and $x \geq 1$, define:

$$A(n) = \sum_{d|n} \chi(d); \quad B(x) = \sum_{n \leq x} \frac{A(n)}{n}$$

Then we have,

(1) $B(x) \to \infty$ as $x \to \infty$.
(2) $B(x) = 2\sqrt{x} L(1, \chi) + O(1)$

Therefore, $L(1, \chi) \neq 0$.

To prove the above theorem, we begin with the following lemma:

Lemma 5.2

Let χ be a real-valued character mod$k$ and define:

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

Then $A(n) \geq 0$ for all $n \in \mathbb{N}$. Moreover, $A(n) \geq 1$ if n is a perfect square.

Proof. Since $\chi(n)$ is multiplicative, $A(n)$ is so. Therefore consider the prime powers:

$$A(p^\alpha) = \sum_{i=0}^{\alpha} \chi(p)^i = 1 + \sum_{i=1}^{\alpha} \chi(p)^i$$

As χ takes only real values, $\chi(p) = 0, \pm 1$ are the only possibilities. If $\chi(p) = 0$, then $A(p^\alpha) = 1$; if $\chi(p) = 1$, then $A(p^\alpha) = 1 + \alpha$; and if $\chi(p) = -1$, then

$$A(p^\alpha) = \begin{cases} 0 & \text{if } \alpha \text{ is odd} \\ 1 & \text{if } \alpha \text{ is even.} \end{cases}$$

Therefore, $A(n) = \prod_{p^\alpha | n} A(p^\alpha) \geq 0$. If n is a perfect square, all the powers of the primes diving n, will be even; in that case $A(p^\alpha) \geq 1$. Hence, $A(n) \geq 1$. □

5.1. **Proof of Theorem 5.1.**

Proof. Using Lemma 5.2 for $x \geq 1$ we get

$$B(x) \geq \sum_{\substack{n \leq x \\ n = m^2}} \frac{1}{\sqrt{n}} = \sum_{m \leq \sqrt{x}} \frac{1}{m}$$

Since harmonic series diverges; as $x \to \infty$ we have $B(x) \to \infty$ proving part (1). Furthermore we notice that

$$B(x) = \sum_{n \leq x} \frac{1}{\sqrt{n}} \sum_{d|n} \chi(d) = \sum_{\substack{d,e \\ de \leq x}} \frac{\chi(d)}{\sqrt{de}}$$

Recall Dirichlet hyperbola method:

$$\sum_{\substack{d,e \\ de \leq x}} f(d)g(e) = \sum_{n \leq \sqrt{x}} f(n) G\left(\frac{x}{n}\right) + \sum_{n \leq \sqrt{x}} g(n) F\left(\frac{x}{n}\right) - F(\sqrt{x})G(\sqrt{x})$$

where $F(x) = \sum_{n \leq x} f(n)$ and $G(x) = \sum_{n \leq x} g(n)$. Setting

$$f(n) = \frac{\chi(n)}{\sqrt{n}}; \quad g(n) = \frac{1}{\sqrt{n}}$$

we arrive at:

$$B(x) = \sum_{\substack{d,e \\ de \leq x}} \frac{\chi(d)}{\sqrt{de}} = \sum_{n \leq x} \frac{\chi(n)}{\sqrt{n}} G\left(\frac{x}{n}\right) + \sum_{n \leq \sqrt{x}} \frac{1}{\sqrt{n}} F\left(\frac{x}{n}\right) - F(\sqrt{x})G(\sqrt{x})$$

To estimate F and G we observe that:

$$F(x) = \sum_{n \leq x} \frac{\chi(n)}{\sqrt{n}} = B + O\left(\frac{1}{\sqrt{x}}\right)$$

where $B = \sum_{n=1}^{\infty} \frac{\chi(n)}{\sqrt{n}}$ and

$$G(x) = \sum_{n \leq x} \frac{1}{\sqrt{n}} = 2\sqrt{x} + A + O\left(\frac{1}{\sqrt{x}}\right)$$

for some constant A. First its visible that $F(\sqrt{x})G(\sqrt{x}) = 2Bx^{\frac{1}{4}} + O(1)$. Therefore,

$$B(x) = \sum_{n \leq x} \frac{\chi(n)}{\sqrt{n}} \left\{ 2\sqrt{\frac{x}{n}} + A + O\left(\sqrt{\frac{n}{x}}\right) \right\} + \sum_{n \leq \sqrt{x}} \frac{1}{\sqrt{n}} \left\{ B + O\left(\sqrt{\frac{n}{x}}\right) \right\} - 2Bx^{\frac{1}{4}} + O(1)$$

$$= 2\sqrt{x} \sum_{n \leq x} \frac{\chi(n)}{n} + AF(\sqrt{x}) + O\left(\frac{1}{\sqrt{x}} \sum_{n \leq x} |\chi(n)|\right)$$

$$+ BG(\sqrt{x}) + O\left(\frac{1}{\sqrt{x}} \sum_{n \leq \sqrt{x}} 1\right) - 2Bx^{\frac{1}{4}} + O(1)$$

which simplifies to

$$\begin{aligned} B(x) &= 2\sqrt{x}L(1,\chi) + BG(\sqrt{x}) - 2Bx^{\frac{1}{4}} + O(1) \\ &= 2\sqrt{x}L(1,\chi) + 2Bx^{\frac{1}{4}} - 2Bx^{\frac{1}{4}} + O(1) \\ &= 2\sqrt{x}L(1,\chi) + O(1) \end{aligned}$$

Therefore, $L(1,\chi)$ can't vanish as required. □

6. An Application of Shapiro's Tauberian Theorem

Oftentimes instead of dealing with the partial sums of arithmetic functions, we play with weighted partial sums. Surprisingly a set of theorems build a connection between apparently different weighted sums of arithmetic functions, which are called *Tauberian Theorems*. Shapiro (1950) gave his Tauberian theorem (stated without proof):

Theorem 6.1 (Shapiro's Tauberian Theorem; One of the Implication)
Let f be a real-valued arithmetic function such that for all $x \geq 1$:

$$\sum_{n\leq x} f(n)\left[\frac{x}{n}\right] = x\log x + O(x)$$

Then for all $x \geq 1$ we have,

$$\sum_{n\leq x} \frac{f(n)}{n} = \log x + O(1)$$

Now we recall a known result (stated without proof):

$$\log [x]! = \sum_{n\leq x} \Lambda(n)\left[\frac{x}{n}\right] = x\log x - x + O(\log x).$$

Observe that for $x > 1$ we have,

$$\sum_{n\leq x} \Lambda(n)\left[\frac{x}{n}\right] = \sum_{p\leq x}\sum_{\substack{\alpha\geq 1\\ p^\alpha\leq x}} \log p\left[\frac{x}{p^\alpha}\right] = \sum_{p\leq x}\log p\left[\frac{x}{p}\right] + \sum_{p\leq x}\sum_{\substack{\alpha\geq 2\\ p^\alpha\leq x}} \log p\left[\frac{x}{p^\alpha}\right]$$

where the second term is $O(x)$. This can be proved using a trick which is used in proving Lemma 7.3. Therefore we obtain the following (after plugging back and rearranging) for $x > 1$:

$$\sum_{n\leq x} \Lambda^\star(n)\left[\frac{x}{n}\right] = \sum_{p\leq x}\log p\left[\frac{x}{p}\right] = x\log x + O(x)$$

where we have defined an arithmetic function $\Lambda^\star : \mathbb{N} \to \mathbb{R}$ such that

$$\Lambda^\star(n) = \begin{cases} \log n & \text{if } n \text{ is prime} \\ 0 & \text{otherwise.} \end{cases}$$

Now we apply Shapiro's Tauberian Theorem, which yields:

Theorem 6.2

For $x > 1$ we have,

$$\sum_{n \leq x} \frac{\Lambda^\star(n)}{n} = \sum_{p \leq x} \frac{\log p}{p} = \log x + O(1)$$

Part 2. The Finale

7. Overview of the Strategy of the Proof

To prove Theorem 1.1, our aim is to end up with a similar asymptotic formula, as we saw in Theorem 6.2, restricting the index of the sum to the primes which satisfy $p \equiv h \bmod k$:

Theorem 7.1

For $x > 1$ we have,

$$\sum_{\substack{p \leq x \\ p \equiv h \bmod k}} \frac{\log p}{p} = \frac{1}{\varphi(k)} \log x + O(1).$$

The right hand side diverges as $x \to \infty$, therefore, the index of the sum in the left cannot be finite. Therefore, there will be infinitely many primes with $p \equiv h \bmod k$, implying Theorem 1.1. So, to prove Theorem 7.1, we start with the following lemma:

Lemma 7.2

For $x > 1$ we have,

$$\sum_{\substack{p \leq x \\ p \equiv h \bmod k}} \frac{\log p}{p} = \frac{1}{\varphi(k)} \log x + \frac{1}{\varphi(k)} \sum_{i=2}^{\varphi(k)} \bar{\chi}_i(h) \sum_{p \leq x} \frac{\chi_i(p) \log p}{p} + O(1).$$

We can easily see that proving Theorem 7.1 reduces down to proving

$$\sum_{p \leq x} \frac{\chi(p) \log p}{p} = O(1)$$

for each $\chi \neq \chi_1$. The sum on the left is a bit difficult to handle, because it's over the primes only. Therefore, first we convert it to a sum the naturals:

Lemma 7.3

For $\chi \neq \chi_1$ and $x > 1$ we have,

$$\sum_{p \leq x} \frac{\chi(p) \log p}{p} = -L'(1, \chi) \sum_{n \leq x} \frac{\mu(n)\chi(n)}{n} + O(1)$$

Therefore, the primary problem reduces to show that

$$\sum_{n \leq x} \frac{\mu(n)\chi(n)}{n} = O(1)$$

for each $\chi \neq \chi_1$. But the close we can get is the following:

Lemma 7.4
For $\chi \neq \chi_1$ and $x > 1$ we have,

$$L(1, \chi) \sum_{n \leq x} \frac{\mu(n)\chi(n)}{n} = O(1)$$

If $L(1, \chi)$ does not vanish for all non-principal χs, then we are done. We already showed that for real non-principal χ. An important observation is that if χ is non-principal and non-real, then $L(1, \chi) = 0$ if and only if $L(1, \bar{\chi}) = 0$, i.e., the vanishing χ arise in pairs. Therefore defining the number $N(k)$ which counts the number of χs for which $L(1, \chi)$ vanishes, picks even non-negative integers only. Then we proceed further to show that

Lemma 7.5
For $x > 1$ we have,

$$\sum_{\substack{p \leq x \\ p \equiv 1 \bmod k}} \frac{\log p}{p} = \frac{1 - N(k)}{\varphi(k)} \log x + O(1)$$

which tells us that if $N(k) \neq 0$, i.e., $N(k) \geq 2$, the principal term on the right diverges to $-\infty$, as $x \to \infty$, whereas the left hand side is always positive, which is a contradiction. Therefore, $N(k) = 0$, which completes the proof of Theorem 7.1. But to prove the asymptotic formula involving $N(k)$ in the Lemma 7.5, we need an intermediate step:

Lemma 7.6
If $\chi \neq \chi_1$, $L(1, \chi) = 0$ and $x > 1$ we have,

$$L'(1, \chi) \sum_{n \leq x} \frac{\mu(n)\chi(n)}{n} = \log x + O(1)$$

This was the overall flowchart to achieve the Theorem 1.1.

8. Proof of the Lemmas

8.1. Proof of Lemma 7.2.

Proof. We start with the asymptotic formula we obtained in Theorem 6.2, and extract the primes which are present only in the reduced residue class $h \bmod k$ (therefore with $(h, k) = 1$). This can be done invoking one of the orthogonality relations for Dirichlet characters. Observing the form of the statement which we desire, we notice that:

$$\sum_{i=1}^{\varphi(k)} \bar{\chi}_i(h)\chi_i(p) = \begin{cases} \varphi(k) & \text{if } p \equiv h \bmod k \\ 0 & \text{if } p \not\equiv h \bmod k \end{cases}$$

Our instinct is that if we inject this orthogonality relation as a weight to the sum of $\frac{\log p}{p}$, then the index of sum gets restricted only to the primes in h mod k. Therefore:

$$\varphi(k) \sum_{\substack{p \le x \\ p \equiv h \bmod k}} \frac{\log p}{p} = \sum_{p \le x} \frac{\log p}{p} \sum_{i=1}^{\varphi(k)} \bar{\chi}_i(h)\chi_i(p) = \sum_{i=1}^{\varphi(k)} \bar{\chi}_i(h) \sum_{p \le x} \frac{\chi_i(p)\log p}{p}$$

A careful observation tells us that if $\chi_i = \chi_1$, i.e., principal character, then:

$$\sum_{p \le x} \frac{\chi_1(p)\log p}{p} = \sum_{p \le x} \frac{\log p}{p} - \sum_{\substack{p \le x \\ p \mid k}} \frac{\log p}{p}$$

Since $\chi_1(p)$ vanishes when $\gcd(p,k) > 1$. As p is a prime, $\gcd(p,k) > 1$ implies $\gcd(p,k) = p$, thus $p|k$. Also, $\chi_1(h) = 1$ since $gcd(h,k) = 1$ But those are only finitely many, hence the sum is bounded. So we obtain:

$$\varphi(k) \sum_{\substack{p \le x \\ p \equiv h \bmod k}} \frac{\log p}{p} = \sum_{p \le x} \frac{\log p}{p} + \sum_{i=2}^{\varphi(k)} \bar{\chi}_i(h) \sum_{p \le x} \frac{\chi_i(p)\log p}{p} + O(1)$$

Using the asymptotic formula we started with and then dividing both sides by $\varphi(k)$ we obtain:

$$\sum_{\substack{p \le x \\ p \equiv h \bmod k}} \frac{\log p}{p} = \frac{1}{\varphi(k)} \log x + \frac{1}{\varphi(k)} \sum_{i=2}^{\varphi(k)} \bar{\chi}_i(h) \sum_{p \le x} \frac{\chi_i(p)\log p}{p} + O(1)$$

as required.

□

8.2. **Proof of Lemma 7.3.**

Proof. Looking at $\log p$ on the sum in the left, we recall von Mangoldt's function $\Lambda(n)$. Observe that the sum $\sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n}$ can be written as:

$$\sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n} = \sum_{p \le x} \sum_{j : p^j \le x} \frac{\chi(p^j)\log(p)}{p^j}$$

Since $\Lambda(n)$ vanishes at composite numbers and takes the value $\log p$ for $n = p^j$, $j \in \mathbb{N}$. The above expression immediately converts the sum on naturals to a sum over primes; Moreover, the term corresponding to $j = 1$ is exactly the sum we are looking for:

$$\sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n} = \sum_{p \le x} \frac{\chi(p)\log(p)}{p} + \sum_{p \le x} \sum_{\substack{j \ge 2 \\ p^j \le x}} \frac{\chi(p^j)\log(p)}{p^j}$$

The second term on the right is dominated from above:

$$\sum_{p\le x}\sum_{\substack{j\ge 2\\ p^j\le x}}\left|\frac{\chi(p^j)\log(p)}{p^j}\right| < \sum_{p\le x}\sum_{j=2}^{\infty}\left|\frac{\chi(p^j)\log(p)}{p^j}\right| \le \sum_{p\le x}\log(p)\sum_{j=2}^{\infty}\frac{1}{p^j} = \sum_{p\le x}\frac{\log p}{p(p-1)} < \sum_{n=2}^{\infty}\frac{\log n}{n(n-1)}$$

which is $O(1)$. Therefore,

$$\sum_{p\le x}\frac{\chi(p)\log(p)}{p} = \sum_{n\le x}\frac{\chi(n)\Lambda(n)}{n} + O(1)$$

Recalling that,

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

we plug it to the equation above:

$$\sum_{p\le x}\frac{\chi(p)\log(p)}{p} = \sum_{n\le x}\frac{\chi(n)}{n}\sum_{d|n}\mu(d)\log\left(\frac{n}{d}\right) + O(1)$$

which can be expressed as:

$$\begin{aligned}\sum_{p\le x}\frac{\chi(p)\log(p)}{p} &= \sum_{d\le x}\frac{\chi(d)\mu(d)}{d}\sum_{c\le x/d}\frac{\chi(c)\log(c)}{c} + O(1)\\ &= -L'(1,\chi)\sum_{d\le x}\frac{\chi(d)\mu(d)}{d} + O\left(\sum_{d\le x}\frac{\log(x/d)}{x}\right)\end{aligned}$$

We try to simplify the bound:

$$\begin{aligned}\sum_{d\le x}\frac{\log(x/d)}{x} &= \frac{1}{x}\sum_{d\le x}(\log x - \log d)\\ &= \frac{\lfloor x\rfloor}{x} - \frac{1}{x}\sum_{d\le x}\log d\\ &= \frac{\lfloor x\rfloor}{x}\log x - \frac{1}{x}\log x!\\ &= \frac{\lfloor x\rfloor}{x}\log x - \log x\\ &= -\frac{\{x\}\log x}{x} = O(1)\end{aligned}$$

Therefore we obtain:

$$\sum_{p\le x}\frac{\chi(p)\log(p)}{p} = -L'(1,\chi)\sum_{d\le x}\frac{\chi(d)\mu(d)}{d} + O(1) = -L'(1,\chi)\sum_{n\le x}\frac{\chi(n)\Lambda(n)}{n} + O(1)$$

as required. $\square$

8.3. **Proof of Lemma 7.4.**

Proof. The sum we have to estimate involves Möbius function and a completely multiplicative function χ, we proceed with the generalized Möbius inversion formula:

$$x = \sum_{n\leq x} \mu(n)\chi(n)G\left(\frac{x}{n}\right) \iff G(x) = \sum_{n\leq x} \chi(n)\frac{x}{n} = x\sum_{n\leq x} \frac{\chi(n)}{n}$$

Note that, since $\chi \neq \chi_1$ we have a guaranteed convergence:

$$G(x) = x\sum_{n\leq x} \frac{\chi(n)}{n} = xL(1,\chi) + O(1)$$

Plugging G into it, further we see:

$$x = \sum_{n\leq x} \mu(n)\chi(n)\left\{\frac{x}{n}L(1,\chi) + O(1)\right\} = xL(1,\chi)\sum_{n\leq x} \frac{\mu(n)\chi(n)}{n} + O\left(\sum_{n\leq x} \mu(n)\chi(n)\right)$$

Clearly, $\sum_{n\leq x} \mu(n)\chi(n) = O(1)$. Therefore, after dividing by x both sides and re-expressing it, we get:

$$L(1,\chi)\sum_{n\leq x} \frac{\mu(n)\chi(n)}{n} = O(1)$$

as required. □

8.4. **Proof of Lemma 7.6.**

Proof. Again we use generalized Möbius inversion formula:

$$x\log x = \sum_{n\leq x} \mu(n)\chi(n)G\left(\frac{x}{n}\right) \iff G(x) = \sum_{n\leq x} \chi(n)\frac{x}{n}\log\frac{x}{n} = x\log x\sum_{n\leq x} \frac{\chi(n)}{n} - x\sum_{n\leq x} \frac{\chi(n)\log n}{n}$$

Since $\chi \neq \chi_1$, the convergences are guaranteed. Since $L(1,\chi) = 0$, we obtain:

$$G(x) = x\log xL(1,\chi) + xL'(1,\chi) + O(\log x) = xL'(1,\chi) + O(\log x)$$

Plugging this back, we get:

$$\begin{aligned} x\log x &= \sum_{n\leq x} \mu(n)\chi(n)\left\{\frac{x}{n}L'(1,\chi) + O\left(\log\frac{x}{n}\right)\right\} \\ &= xL'(1,\chi)\sum_{n\leq x} \frac{\mu(n)\chi(n)}{n} + O\left(\sum_{n\leq x}(\log x - \log n)\right) \\ &= xL'(1,\chi)\sum_{n\leq x} \frac{\mu(n)\chi(n)}{n} + O(x) \end{aligned}$$

where we had already obtained the bound in the proof of Lemma 7.3. Therefore, dividing by x, we finish with:

$$L'(1,\chi)\sum_{n\leq x} \frac{\mu(n)\chi(n)}{n} = \log x + O(1)$$

as required. □

8.5. **Proof of Lemma 7.5.**

Proof. Restricting our view to the reduced residue class 1 mod k, we consider Lemma 7.2 with $h = 1$:

$$\sum_{\substack{p \le x \\ p \equiv 1 \bmod k}} \frac{\log p}{p} = \frac{1}{\varphi(k)} \log x + \frac{1}{\varphi(k)} \sum_{i=2}^{\varphi(k)} \sum_{p \le x} \frac{\chi_i(p) \log p}{p} + O(1).$$

using the fact $\bar{\chi}_i(1) = 1$ for each i. To estimate the sum over the primes on the right hand side of the above equation, we use Lemma 7.3:

$$\sum_{p \le x} \frac{\chi_i(p) \log p}{p} = -L'(1, \chi_i) \sum_{n \le x} \frac{\mu(n)\chi_i(n)}{n} + O(1)$$

which is true for each $\chi_i \ne \chi_1$. Now, we observe if $L(1, \chi_i) \ne 0$ for some $i \ne 1$, then by the virtue of Lemma 7.4, we conclude

$$\sum_{p \le x} \frac{\chi_i(p) \log p}{p} = O(1)$$

which in turn reduces the sum we began with to:

$$\sum_{\substack{p \le x \\ p \equiv 1 \bmod k}} \frac{\log p}{p} = \frac{1}{\varphi(k)} \log x + \frac{1}{\varphi(k)} \sum_{\substack{2 \le i \le \varphi(k) \\ L(1,\chi_i) = 0}} \sum_{p \le x} \frac{\chi_i(p) \log p}{p} + O(1).$$

If $L(1, \chi_i) = 0$ for some $i \ne 1$, then by the virtue of Lemma 7.6, we note that:

$$\sum_{p \le x} \frac{\chi_i(p) \log p}{p} = -L'(1, \chi_i) \sum_{n \le x} \frac{\mu(n)\chi_i(n)}{n} + O(1) = -\log x + O(1)$$

which reduces our targetted sum to our desired asymptotic formula:

$$\sum_{\substack{p \le x \\ p \equiv 1 \bmod k}} \frac{\log p}{p} = \frac{1 - N(k)}{\varphi(k)} \log x + O(1).$$

where $N(k)$ is the number of χ_is such that $L(1, \chi_i)$ vanishes. □

9. Future Ideas

A neat observation in Theorem 7.1 is that the principal term on the right is independent of h, i.e., for each reduced residue class mod k, we have exactly same asymptotic formula. This somehow tells us to expect a same asymptotic density of primes in each of the classes, which turned out to be true, and known as *Prime Number Theorem in Arithmetic Progressions*. Studying this carefully with the proof of the Prime Number Theorem is our next exercise.

10. Acknowledgement

For completing this reading along with the coursework of our institute, I am really grateful to Prof. Soumya Bhattacharya for guiding me in the curriculum and keeping me enthusiastic every time with his down-to-earth smile. I thank my friends who always have been supporting me from the first day and last of all, I am indebted to my parents forever for their mental support in every situation in my life.

11. References

1 Tom M. Apostol. *Introduction to Analytic Number Theory.* Springer-Verlag, New York (1998).

2 Harold N. Shapiro. *On Primes In Arithmetic Progression (II)* . Annals of Mathematics (Vol. 52, No. 1, July, 1950)

3 Peter Gustav Lejeune Dirichlet. *There are infinitely many prime numbers in all arithmetic progressions with first term and difference coprime* (English Translation, 2014). Originally published in 1837. https://doi.org/10.48550/arXiv.0808.1408

Current address: A005 Nivedita Hall, Indian Institute of Science Education and Research, Kolkata, Mohanpur, West Bengal - 741246, India.