2019

Primes in Arithmetical Progression

Edward C. Wessel
Colby College

Primes in Arithmetical Progression

Edward (Teddy) Wessel

April 2019

Contents

1 Message to the Reader

This thesis will tackle *Dirichlet's Theorem on Primes in Arithmetical Progressions*. The majority of information that follows below will stem from Tom M. Apostol's *Introduction to Analytical Number Theory* [Apo76]. This is the main source of all definitions, theorems, and method. However, I would like to assure the reader that prior knowledge of neither the text nor analytical number theory in general is needed to understand the result. A rough background in Abstract Algebra and a moderate grasp on Complex and Real Analysis are more than sufficient. In fact, my project's intent is to introduce Dirichlet's ideas to the mathematics student who may be discouraged by the complex presentations of the topic that are commonplace.

2 Introduction

2.1 Primes

Prime numbers have been a subject of human fascination since their discovery/definition by the school of Pythagoras in 600 BC. To this day, numbers such as 3, 7, and 13 are considered mystical omens of good or bad luck. Mathematically, there are a number of questions that arise regarding primes, especially in relation to cardinality. These questions have gotten increasingly difficult to understand and prove. For example, the Twin Prime Conjecture suggests that primes a distance of two apart from one another (i.e. 3 and 5, 11 and 13, or 71 and 73) continue to appear throughout the natural numbers. This conjecture has gone unproven for decades. However, there are a number of simpler theorems that are just as interesting.

2.2 Euclid

Since Euclid's 300 BC proof, the existence of infinitely many prime numbers has been well established. A version converted into modern notation can be seen below.

Theorem 2.2.1. *There are infinitely many prime numbers.*

Proof: Assume for a contradiction, that there are finitely many primes, $p_0, p_1, p_2, ..., p_n$. Let $P = (\prod_{i=0}^{n} p_i) + 1$. If P is prime, then we are done, since P is greater than all p_i for $i \in \{1, 2, 3, ..., n\}$. If P is not prime, then P is

divisible by some prime p. However, $p \neq p_k$ for any k such that $0 \leq k \leq n$ because otherwise $p|1$. To see this, observe that

$$\frac{P}{p_k} = \frac{(\prod_{i=0}^{n} p_i) + 1}{p_k} = \frac{1}{p_k} + \prod_{\substack{0 \leq i \leq n \\ i \neq k}} p_i$$

Thus, $\frac{1}{p_k}$ must be an integer. Thus, p_k must divide 1, which is not possible by definition of primes. Hence, by contradiction, P must be a prime number. Thus, there are infinitely many primes. Q.E.D.

2.3 Euler and Dirichlet

Euler furthered the discussion of infinite primes by showing that $\sum \frac{1}{p}$, extended over the primes, diverges. Dirichlet proved a similar statement with one key difference. Instead of summing over all of the primes, he decided to restrict the sum to primes within a specific arithmetic progression. This will be formally defined in a later section, but for now think of an arithmetic progression as a sequence of positive integers with a starting point and a common difference. The first term, a_0, is the starting point, and $a_{n+1} = a_n +$ (the common difference). We will denote the arithmetic sequence with starting point h and common difference k by $AP(h, k)$. We defined this progression below:

$$AP(h, k) = (kn + h | n \in \mathbb{N})$$

Dirichlet successfully proved that there are infinite primes in arithmetic progressions. He specifies that h and k are relatively prime (ie. $(h, k) = 1$). This is a necessary condition for the proof because if h and k have a common divisor, d, then every term in $AP(h, k)$ is divisible by d. Thus, no term (except perhaps the first) is prime, posing a major contradiction.

2.4 Shapiro

Harold N. Shapiro later improved Dirichlet's result with an easier proof by working with $\sum \frac{\log p}{p}$ instead of $\sum \frac{1}{p}$. He showed that this new sum also diverges, also implying the infinite cardinality of the prime numbers. The reader may question this implication, but it is easy to see that if the sum diverges, then there must be infinitely many terms. Thus, there must be infinitely many primes to contribute to each infinite $\frac{\log p}{p}$ term in the series.

Therefor, divergence of this sum implies infinite primes. More importantly for our work, the divergence of this sum restricted to an arithmetic progression implies infinite primes within that progression.

2.5 Formal Statement

The formal statement of Shapiro's version of Dirichlet's Theorem is actually quite a bit stronger than simply proving infinite primes. The implications of the statement will be examined in the "Discussion" section at the end of this thesis. The direct result is as follows.

Theorem 2.5.1. *If $k > 0$ and $(h, k) = 1$, then for all $x > 1$*

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

Now Theorem 2.4.1 implies the infinite cardinality of primes within an arithmetic progression, but it should be unclear to the reader at this point. What is missing is an understanding of φ, *Euler's Totient Function* and of $O(1)$, the *Big Oh* term. These will be defined in a later section, but for now, assume that $\varphi(k)$ is finite for finite k. As discussed earlier, k is the common difference between terms of $AP(h, k)$. Thus it is arbitrarily specified to be a natural number with the only requirement being that $(h, k) = 1$, meaning $\varphi(k)$ is a constant. Also assume that $O(1)$ is a bounded function, which is true, in fact by definition.

Those assumptions having been made, it is clear (since $\log x$ diverges) that

$$\lim_{x \to \infty} \frac{1}{\varphi(k)} \log x + O(1) = \infty$$

Thus,

$$\lim_{x \to \infty} \sum_{\substack{p \leq x \\ p \equiv h \pmod{k}}} \frac{\log p}{p} = \sum_{p \equiv h \pmod{k}} \frac{\log p}{p} = \infty$$

Thus, there are infinitely many primes within the arithmetic progression, $AP(h, k)$. Now, before we enter into the details of proving this, the reader should have a sense of how this theorem came to fruition. A lot of human effort has been put into the result of this theorem, and the mathematical community owes many great minds of the past (including Dirichlet and Shapiro) for its discovery.

3 Arithmetical Functions

We begin by defining some arithmetical functions, that will be relevant when discussing Theorem 2.5.1. One, Euler's Totient Function, is actually written into the statement of theorem we saw, but all of them are relevant in our proof to come. This process will be tedious to read through, but a good grasp of the definitions and nature of these functions will allow the reader to move through the later sections more fluidly. However, one can always return to this section for a refresher.

Definition 3.0.1. An **arithmetical function**, f, is a complex-valued function defined on the positive integers. They are also sometimes called *number-theoretic functions.*

To clarify, this definition (from Apostol's book) implies that $f : \mathbb{N} \to \mathbb{C}$. The domain is the set of natural numbers (or positive integers), and the range is the complex plane. There are a wide variety of functions that meet this requirement, but these are the ones that we need.

3.1 Euler's Totient Function

Definition 3.1.1. If $n \geq 1$, the **Euler Totient**, φ, is defined to be the number of positive integers not exceeding n, which are relatively prime to n; thus

$$\varphi(n) = \sum_{\substack{1 \leq k \leq n \\ (k,n)=1}} 1$$

Interestingly, while the $\lim_{n\to\infty} \varphi(n) = \infty$, it is periodically the case that $\varphi(i+1) \leq \varphi(i)$ For example, $\varphi(7) = 6$ and $\varphi(8) = 4$.

3.2 The Mobius Function

Definition 3.2.1. The **Mobius Function**, μ, is defined as follows:

$$\mu(1) = 1$$

Furthermore, if $n > 1$, then we factor n and write it as a product of powers of primes, as allowed by *The Fundamental Theorem of Arithmetic*, so that $n = p_1^{a_1} \times p_2^{a_2} \times ... \times p_k^{a_k}$. And we have that

$$\mu(n) = \begin{cases} (-1)^k & \text{if } a_1 = a_2 = ... = a_k = 1 \\ 0 & \text{else} \end{cases}$$

One interesting thing to point out about this function is that $\mu(n) = 0$ if and only if n has a square factor greater than one.

3.3 A Relationship Between φ and μ

Theorem 3.3.1. *If $n \geq 1$, then*

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

Before we begin our proof, we need a lemma.

Lemma 3.1. *If $n \geq 1$, then*

$$\sum_{d|n} \mu(d) = \left[\frac{1}{n}\right]$$

where the brackets are a function that outputs the greatest non-negative integer less than the quantity within. For example, $\left[\frac{7}{3}\right] = 2$ and $\left[\frac{1}{4}\right] = 0$.

Proof of Lemma:

If $n = 1$, then the equation is trivially true. Assume $n > 1$, and let $n = p_1^{a_1} \times p_2^{a_2} \times ... \times p_k^{a_k}$. We now notice that $\mu(d)$ is nonzero if and only if $d = 1$ or d is a product of distinct primes. Thus,

$$\sum_{d|n} \mu(d) = \mu(1) + \mu(p_1) + ... + \mu(p_k) + \mu(p_1p_2) + \mu(p_2p_3) + ... \\ + \mu(p_{k-1}p_k) + ... + \mu(p_1p_2...p_k)$$

We see that this is just equal to:

$$1 + \binom{k}{1}(-1) + \binom{k}{2}(-1)^2 + ... + \binom{k}{k}(-1)^k = (1-1)^k = 0$$

Proof of Theorem:

We can rewrite our equation that we used to define φ as the following:

$$\varphi(n) = \sum_{k=1}^{n} \left[\frac{1}{(n,k)} \right]$$

again where the brackets indicate the greatest non-negative integer that is less than $\left[\frac{1}{(n,k)}\right]$, and k goes through all natural numbers less than n. We see this is true since if $(n,k) \neq 1$ (i.e. n and k are not relatively prime), then $\frac{1}{(n,k)} \leq 1$. Hence, $\left[\frac{1}{(n,k)}\right] = 0$, and thus, $\sum_{k=1}^{n} \left[\frac{1}{(n,k)}\right]$ counts the number of positive integers that are relatively prime to n.

Next, we use our Lemma with $n = (n,k)$, and we substitute $big[\frac{1}{(n,k)}]$ with $\sum_{d|(n,k)} \mu(d)$ like so:

$$\varphi(n) = \sum_{k=1}^{n} \sum_{d|(n,k)} \mu(d) = \sum_{k=1}^{n} \sum_{\substack{d|n \\ d|k}} \mu(d)$$

The second equality is trivial. If something divides the greatest common divisor of two numbers, then it must divide each of those two numbers.

The final step is to notice that the only terms of our double sum are k's, for which some divisor, d of n also divides k. In other words, for a fixed divisor of n, namely d, we are summing over all k such that $1 \leq k \leq n$ such that k is a multiple of d. We write $k = qd$, and note that $1 \leq k \leq n$ if and only if $1 \leq q \leq \frac{n}{d}$. Therefor,

$$\varphi(n) = \sum_{k=1}^{n} \sum_{\substack{d|n \\ d|k}} \mu(d) = \sum_{d|n} \sum_{q=1}^{\frac{n}{d}} \mu(d) = \sum_{d|n} \sum_{q=1}^{\frac{n}{d}} 1 = \sum_{d|n} \mu(d) \frac{n}{d}$$

Q.E.D.

If the reader was confused by this proof, please do not be worried. The result simply shows the intertwined relationship between Euler's Totient Function and the Mobius Function. It will also serve as an example of Dirichlet multiplication, which will be addressed in the proceeding section. However, it is not necessary for our main result (i.e the proof of infinite primes in arithmetic progression).

3.4 The Mangoldt Function

Definition 3.4.1. The next function we will look at is called the **Mangoldt Function**, Λ. Λ is defined as follows:

$$\Lambda(n) = \begin{cases} \log p & \text{if if} n = p^m \text{for some} m \geq 1 \\ 0 & \text{else} \end{cases}$$

Theorem 3.4.2. *If $n \geq 1$,*

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

Proof:

We first mention that when $n = 1$, this equality is obviously true, since both the right and left sides are equal to zero. Now if $n \geq 1$, then we write $n = \prod_{k=1}^{r} p_k^{a_k}$. Now we take the log of n.

$$\log n = \sum_{k=1}^{r} a_k \log p_k$$

Now we observe that terms of $\sum_{d|n} \Lambda(d)$ are only nonzero when $d = p_k^m$ for $k \in \{1, 2, 3, ..., r\}$ and $m \in \{1, 2, 3, ..., a_k\}$. Thus,

$$\begin{aligned} \sum_{d|n} \Lambda(d) &= \sum_{k=1}^{r} \sum_{m=1}^{a_k} \Lambda(p_k^m) \\ &= \sum_{k=1}^{r} \sum_{m=1}^{a_k} \log p_k \\ &= \sum_{k=1}^{r} a_k \log p_k = \log n \text{ Q.E.D.} \end{aligned}$$

Another formula for Λ is the following: $\Lambda(n) = \sum_{d|n} \mu(d) \log\left(\frac{n}{d}\right)$. This follows directly from the previous theorem and the Mobius inversion formula which will be proved later. However, it is often stated outright as the divisor sum definition of the Mangoldt Function. Observing the sum, it is obvious that the two definitions form a structurally identical function.

4 Dirichlet Convolution

The Dirichlet Convolution or Dirichlet Product is an operation on arithmetical functions that is relevant in the context of our main result. In the previous section, we used Dirichlet multiplication unknowingly in

our attempt to consider the relationship between φ and μ. In that situation we (and Dirichlet originally) multiplied the Mobius Function and the arithmetic function $g(x) = x$, the identity function. We then inferred the Dirichlet product to be Euler's Totient Function.

4.1 Definition

We will now define the Dirichlet convolution more generally, as well as describe some of the algebraic properties of the operation. I should note that this operation is a key aspect of our main result, and an understanding of its properties is necessary for manipulating the sums in order to get to our conclusion. The first step is to rigorously define the operation.

Definition 4.1.1. If f and g are two arithmetical functions, we define their Dirichlet product/convolution as h such that:

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

Notation-wise, please note two things: 1) We sometimes denote h as $f * g$ and $h(n)$ as $(f * g)(n)$; and 2) the outer parenthesis will be omitted from this point forward.

4.2 Some Basic Properties of Dirichlet Multiplication

In the following theorems, assume that f, g, and k are arithmetic functions. We will show that Dirichlet multiplication on the space of arithmetic functions is commutative and associative.

Theorem 4.2.1. *Dirichlet multiplication is commutative. In mathematical notion this means that,*

$$f * g = g * f$$

Proof: The key insight here is to notice that $d \times \left(\frac{n}{d}\right) = n$. Thus, we see that:

$$\begin{aligned}(f*g)(n) &= \sum_{d|n} f(d)g\left(\frac{n}{d}\right) \\ &= \sum_{\substack{a,b \text{ s.t.} \\ ab=n}} f(a)g(b) \\ &= \sum_{\substack{a,b \text{ s.t.} \\ ab=n}} g(a)f(b) \\ &= \sum_{d|n} g(d)f\left(\frac{n}{d}\right) = (g*f)(n) \text{ Q.E.D.}\end{aligned}$$

Theorem 4.2.2. *Dirichlet multiplication is associative or,*

$$(f*g)*k = f*(g*k)$$

Proof:

$$\begin{aligned}(f*(g*k)(n) &= \sum_{ad=n} f(a)(g*k))(d) \\ &= \sum_{ad=n} f(a) \sum_{bc=d} g(b)k(c) \\ &= \sum_{abc=n} f(a)g(b)k(c) \\ &= \sum_{ec=n} k(c) \sum_{ab=e} f(a)b(c) \\ &= \sum_{ec=n} (f*g)(e)k(c) = ((f*g)*k)(n) \text{ Q.E.D.}\end{aligned}$$

Note that we have used the commutative property from the last theorem in the final line of this proof.

4.3 Identity and Inverses within Dirichlet Multiplication

Definition 4.3.1. We let the arithmetical function I be defined as seen below, and we call this the *identity*.

$$I(n) = \left[\frac{1}{n}\right] = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$$

Theorem 4.3.2. *Consider an arbitrary arithmetical function, f. We claim that $I * f = f * I = f$.*

Proof:

$$(f * I)(n) = \sum_{d|n} f(d) I\left(\frac{d}{n}\right) = \sum_{d|n} f(d) \left[\frac{d}{n}\right] = f(n) \text{ Q.E.D.}$$

The final equation becomes immediately apparent upon noting that $\left[\frac{d}{n}\right]$ is nonzero only when $d = n$. Thus, the sum has only one nonzero term, namely, $f(n)\left[\frac{n}{n}\right] = f(n)$.

Theorem 4.3.3. *If f is an arithmetical function with $f(1) \neq 0$, then there exists a unique arithmetical function that is the Dirichlet inverse of f, denoted by f^{-1} such that,*

$$f * f^{-1} = f^{-1} * f = I$$

First, we will define the Dirichlet inverse of f, recursively. Then, we will show that it satisfies the property above.

Definition 4.3.4. Let f be an arbitrary arithmetical function. Define f^{-1} in the following way,

$$f^{-1}(1) = \frac{1}{f(1)}$$
$$f^{-1}(n) = \frac{-1}{f(1)} \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) f^{-1}(d) \text{ for all } n > 1.$$

Proof:

Our first step is proving that our base case satisfies the desired property. So, we are claiming that $(f * f^{-1})(1) = I(1)$. We can distribute the 1 amongst our two arithmetical functions on the left. Thus, $f(1) * f^{-1}(1) = I(1)$. From here, we simply divide both sides of the equation by $f(1)$ to obtain our desired result. Recall, we presupposed that $f(1)$ was nonzero in our Theorem, so there are no problems with dividing by $f(1)$.

Next, we will inductively prove the recursive portion of the definition. Assume that $f^{-1}(k)$ has been defined for all $k < n$. We also must assume that these values are unique, since this is a existence and uniqueness proof. From here we will show that $(f * f^{-1})(n) = I(n)$. Since this case is

considering $n > 1$, we note that $I(n) = 0$. Putting this together with the definitions of f^{-1} and Dirichlet convolution, we see that $(f * f^{-1})(n) = I(n)$ is equivalent to saying that,

$$\sum_{d|n} f\left(\frac{n}{d}\right) f(d) = 0$$

Next, we isolate the term where $d = n$. The remaining sum is restricted to all divisors of n such that $d < n$.

$$f(1)f^{-1}(n) + \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) f^{-1}(d) = 0$$

We note that f^{-1} is uniquely defined for all $k < n$, thus all $d < n$ where $d|n$ are also uniquely defined. Thus, the previous equation in our proof can be used to find a unique value for $f^{-1}(n)$, specifically,

$$f^{-1}(n) = \left(\frac{-1}{f(1)}\right) \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) f^{-1}(d) \text{ Q.E.D.}$$

We should pause a moment after this proof and consider it in the context of everything else we have done with Dirichlet convolutions of arithmetical functions. We have proved associativity, and the existence and uniqueness of an identity and inverse. Furthermore that arithmetical functions are closed under Dirichlet multiplication is trivial. So, this space forms a group. In tandem with our proof of commutativity, we can conclude that the space is an abelian group. If the reader is confused by this discussion, he/she should review group axioms and the definition of abelian.

4.4 Multiplicative Functions and their Relationship with Dirichlet Convolution

We will now consider a subset of the arithmetical functions we have defined previously. They are called *multiplicative functions*. They may not be particularly interesting or new to the reader on their own, but the relationship these functions have with Dirichlet convolutions certainly is. In fact, while we will not prove this, the multiplicative functions form a subgroup of the group of arithmetical functions under Dirichlet multiplication.

Definition 4.4.1. An arithmetical function f is *multiplicative* if f is not the zero function and if,

$$f(mn) = f(m)f(n) \text{ when } (m,n) = 1$$

We say f is *completely multiplicative* if m and n need not be relatively prime.

Theorem 4.4.2. *If f and g are both multiplicative, then so is their Dirichlet convolution, $h = f * g$*

Proof:

We begin by writing $h(mn)$ using our definition of Dirichlet convolution. We let m and n, be relatively prime of course, since that was part of our definition of *multiplicative*. Thus, $h(mn) = \sum_{c|mn} f(c)g\left(\frac{mn}{c}\right)$. Now we use a trick, and express every divisor, c, of mn as a product ab where $a|m$ and $b|n$. We note that $(m,n) = 1$ thus $(a,b) = 1$ and $(\frac{a}{m}, \frac{b}{n}) = 1$. Thus, there exists an injection from the products, ab, to the divisors, c, of mn. From these deductions, the remainder of the proof is just algebra;

$$\begin{aligned} h(mn) &= \sum_{\substack{a|m \\ b|n}} f(ab)g\left(\frac{mn}{ab}\right) = \sum_{\substack{a|m \\ b|n}} f(a)f(b)g\left(\frac{m}{a}\right)g\left(\frac{n}{b}\right) \\ &= \sum_{a|m} f(a)g\left(\frac{m}{a}\right) \sum_{b|n} g(b)g\left(\frac{n}{b}\right) = h(m)h(n) \text{ Q.E.D.} \end{aligned}$$

Theorem 4.4.3. *If f is completely multiplicative, then $f^{-1}(n) = \mu(n)f(n)$ for all $n \geq 1$.*

Proof:

Let $g(n) = \mu(n)f(n)$. If f is completely multiplicative, then

$$\begin{aligned} (g * f)(n) &= \sum_{d|n} g(d)f\left(\frac{n}{d}\right) = \sum_{d|n} (\mu(d)f(d))f\left(\frac{n}{d}\right) \\ &= \sum_{d|n} \mu(d)f(d)f\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d)f\left(\frac{dn}{d}\right) \\ &= \sum_{d|n} \mu(d)f(n) = f(n)\sum_{d|n} \mu(d) \end{aligned}$$

$$= f(n)I(n)$$

The last line of the equality above is given by Lemma 1 from Theorem 3.3.1. Now this may seem like a weird place to be, but it is exactly the equality we need. The final simplification step is to notice that $I(n) = 0$ for all $n > 1$. Additionally, $f(1) = 1$. Thus, $f(n)I(n)$ can be simplified to just $I(n)$. Thus, we have that $(g * f)(n) = f(n)I(n) = I(n)$ Q.E.D.

4.5 Generalized Convolutions

Up until now, we have applied the convolution operation to two arithmetical functions only. We now consider a convolution between two functions, only one of which will be arithmetical. The other function will be defined on the positive real numbers, and thus, its domain will encompass that of the arithmetical functions. The cardinally of this new domain is much larger however.

Definition 4.5.1. We let F be a complex-valued functions defined on $(0, +\infty)$ such that $F(x) = 0$ for $0 < x < 1$. Let α be an arbitrary arithmetical function. We will denote the operation of this generalized convolution with, $\star$. Just as with the Dirichlet convolutions from before, we define,

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

There are a few things to say about this new function $G(x)$. First, we note that G takes the argument x as opposed to n, like the Dirichlet product. This is because this new G is defined on $(0, +\infty)$, as F was. Another thing we can point out is that if we restrict F to the natural numbers or let $F(x) = 0$ for all non integral x, $(\alpha \star F)(x) = (\alpha * F)(x)$ for all natural number x. Now this operation is neither commutative nor associative, as we proved Dirichlet products are. Regardless there are a few things we can prove about this $\star$, including a quasi-associativity.

Before we get into some theorems, we need to mention a special arithmetical function under this operation. This discussion will be relatively informal compared to the proofs to come. The identity element for the group of arithmetical functions under Dirichlet products is important here. $I = \left[\frac{1}{n}\right]$ also happens to be the left identity for the operation, $\star$. We see this is true since,

$$(I \star F)(x) = \sum_{n \leq x} I(n) F\left(\frac{x}{n}\right) = \sum_{n \leq x} \left[\frac{1}{n}\right] F\left(\frac{x}{n}\right)$$

Well, $\left[\frac{1}{n}\right]$ is only nonzero when $n = 1$, in which case $F\left(\frac{x}{n}\right) = F(x)$. Thus, we see that, $(I \star F)(x) = F(x)$.

Theorem 4.5.2. *For arithmetical functions, α and β, and a function F, as defined above,*

$$\alpha \star (\beta \star F) = (\alpha * \beta) \star F$$

Proof:

We begin the proof by taking an arbitrary $x \in (0, +\infty)$ and applying our definition for $\star$. When stacking these definitions inside of one another, we must take into consideration the fractional argument inside the second function within the sum. This will become more clear in the algebra.

$$\begin{aligned}
(\alpha \star (\beta \star F))(x) &= \sum_{n \leq x} \alpha(n) \sum_{m \leq x/n} \beta(m) F\left(\frac{x}{mn}\right) \\
&= \sum_{mn \leq x} \alpha(n)\beta(m) F\left(\frac{x}{mn}\right) \\
&= \sum_{l \leq x} \left(\sum_{n|l} \alpha(n)\beta\left(\frac{l}{n}\right)\right) F\left(\frac{x}{l}\right) \\
&= \sum_{l \leq x} (\alpha * \beta)(l) F\left(\frac{x}{l}\right) = ((\alpha * \beta) \star F)(x)
\end{aligned}$$

Note that the second inequality is just a matter of manipulating the sums. We notice that $m \leq \frac{x}{n}$ is equivalent to saying that $mn \leq x$, and if $mn \leq x$ then $n \leq x$. From there, the manipulation is trivial. The third equation is simply letting $l = mn$ and reorganizing the sum accordingly. All other equalities are condensing and expanding definitions for $*$ and $\star$.

The next theorem we will discuss is sometimes referred to as the *generalized inversion formula*. It uses our quasi-associativity result to pull inverses through our new operator, $\star$. Our discussion of the left identity for $\star$ is also pertinent.

Theorem 4.5.3. *Let α be an arithmetical function with Dirichlet inverse, α^{-1}. Then we have that,*

$$G(x) = \sum_{n \leq x} \alpha(n) F\left(\frac{x}{n}\right) \iff F(x) = \sum_{n \leq x} \alpha^{-1}(n) G\left(\frac{x}{n}\right)$$
$$\textit{i.e. } G = (\alpha \star F) \iff F = (\alpha^{-1} \star G)$$

Proof:

$$\begin{aligned} \text{If } G = (\alpha \star F) \text{ then, } \alpha^{-1} \star G &= \alpha^{-1} \star (\alpha \star F) \\ &= (\alpha^{-1} \star \alpha) \star F = I \star F = F \\ \text{If } F = (\alpha^{-1} \star G) \text{ then, } \alpha \star F &= \alpha \star (\alpha^{-1} \star G) \\ &= (\alpha \star \alpha^{-1}) \star G = I \star G = G \end{aligned}$$

Q.E.D.

This is an critical result for some of our theorems to come. A particularly important case will be the *Generalized Mobius inversion formula.*

Theorem 4.5.4. *If α is a completely multiplicative function, then,*

$$G(x) = \sum_{n \leq x} \alpha(n) F\left(\frac{x}{n}\right) \iff F(x) = \sum_{n \leq x} \mu(n)\alpha(n) G\left(\frac{x}{n}\right)$$
$$\textit{i.e. } G = (\alpha \star F) \iff F = (\mu\alpha \star G)$$

Proof:

We recall from Theorem 4.4.3. that if an arithmetical function, f, is completely multiplicative, then its inverse is μf. Thus, $\alpha^{-1} = \mu\alpha$. The proof then follows from Theorem 4.5.3.

4.6 Partial Sums of Dirichlet Convolution

The following theorem will discuss the relationship between the partial sums of arbitrary arithmetical functions and the partial sums of their Dirichlet product. This theorem will be used in Shapiro's proof of infinite primes, which shows the divergence of the sum, $\sum_{p \leq x} \frac{\log p}{p}$.

Theorem 4.6.1. *If $h = f * g$, assign*

$$H(x) = \sum_{n \leq x} h(n) \qquad F(x) = \sum_{n \leq x} f(n) \qquad G(x) = \sum_{n \leq x} g(n)$$

Then the following equation is true:

$$H(x) = \sum_{n \leq x} h(n) = \sum_{n \leq x} f(n) G\left(\frac{x}{n}\right) = \sum_{n \leq x} g(n) F\left(\frac{x}{n}\right)$$

Proof:

We will construct an equation in order to prove this with Theorem 4.5.2., the quasi-associativity of generalized convolutions. We will call this equation U. Let,

$$U(x) = \begin{cases} 0 & \text{if } 0 < x < 1, \\ 1 & \text{if } x \geq 1. \end{cases}$$

Notice that U is defined on the positive real line and is zero-valued for $0 < x < 1$. Also, notice that, $(f \star U)(x) = \sum_{n \leq x} f(n) U\left(\frac{x}{n}\right)$. But since $\frac{x}{n} \geq 1$ for all $n \leq x$, we have that,

$$\begin{aligned} (f \star U)(x) &= \sum_{n \leq x} f(n) U(\frac{x}{n}) \\ &= \sum_{n \leq x} f(n) = F(x) \end{aligned}$$

By similar logic $G(x) = (g \star U)(x)$. Thus see that,

$$\begin{aligned} f \star G = f \star (g \star U) = (f * g) \star U = H \text{ and} \\ g \star F = g \star (f \star U) = (g * f) \star U = H \end{aligned}$$

Q.E.D.

This result has an important corollary that we will state here. It will be referenced and restated in a later section aimed at proving Shapiro's result, $\sum \frac{\log p}{p}$ diverges.

Corollary 1. *If $F(x) = \sum_{n \leq x} f(n)$ then,*

$$\sum_{n \leq x} \sum_{d|n} f(d) = \sum_{n \leq x} f(n) \left[\frac{x}{n}\right] = \sum_{n \leq x} F\left(\frac{x}{n}\right)$$

Proof:

Let $g(n) = 1$ for all n. Thus, $G(x) = \sum_{n \leq x} 1 = [x]$ and,

$$h(n) = (f * g) = \sum_{d|n} \left(f\left(d\right) g\left(\frac{n}{d}\right) \right) = \sum_{d|n} f(d)1 = \sum_{d|n} f(d)$$

Thus we have that,

$$\begin{aligned} H(x) &= \sum_{n \leq x} \sum_{d|n} f(d) \\ &= \sum_{n \leq x} f(n) G\left(\frac{x}{n}\right) = \sum_{n \leq x} f(n) \left[\frac{x}{n}\right] \\ &= \sum_{n \leq x} g(n) F\left(\frac{x}{n}\right) = \sum_{n \leq x} F\left(\frac{x}{n}\right). \end{aligned}$$

Removing the intermediary/explanatory steps of that equation, we see that our corollary is proved.

Now we conclude this section with another theorem that is a more general version of the previous one. Recall that in Theorem 4.6.1., we wrote,

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

This is by definition of Dirichlet convolution. However, we can rearrange the sum on the left to be equal to,

$$\sum_{\substack{q,d \\ qd \leq x}} f(d)g(q)$$

Theorem 4.6.2. *Let all assumptions made in the previous theorem be the same. If a and b are positive real numbers such that $ab = x$, then*

$$\sum_{\substack{q,d \\ qd \leq x}} f(d)g(q) = \sum_{n \leq a} f(n) G\left(\frac{x}{n}\right) + \sum_{n \leq b} g(n) F\left(\frac{x}{n}\right) - F(a)G(b)$$

We leave this proof to the reader, since it takes a fair amount of time and is only used in one future result.

5 The Big Oh Notation

Thus far, we have discussed several arithmetical functions, their Dirichlet Convolutions, and their relationships with one another. However, the main result of this thesis concerns the behavior of these functions and sums involving these functions as $n \to \infty$. In other words, the sums will have infinitely many terms, and we need to consider the asymptotic behavior of the relevant functions. To do this we must introduce a new concept, the Big Oh notation. This notation shows up in our main result as well as the result, which we restrict to obtain our main result. This is crucial for our understanding of infinite primes in arithmetical progressions.

5.1 Definition

Definition 5.1.1. Let $g(x) > 0$ for all $x \geq a$. Then one writes,

$$f(x) = O(g(x))$$

to imply that $\frac{f(x)}{g(x)}$ is bounded for all $x \geq a$. In other words, there exists an $M > 0$, such that,

$$|f(x)| \leq Mg(x) \text{ for all } x \geq a.$$

Additionally, one writes,

$$f(x) = h(X) + O(g(x))$$

to imply that $|f(x) - h(X)| \leq M(g(x))$ for some $M > 0$.

The Big-Oh-Notation essentially tells us that the absolute value of the function equal to the big-oh-term is bounded by some constant times the argument function inside O.

5.2 Euler's Summation Formula

Our main result requires using Big Oh Notation to understand sums and partial sums. The following formula (attributed to Euler) is a tool for equating a partial sum and an integral. As the reader may know, we can approximate integrals with sums and vice versa. Euler's formula is a precise equation for the error in using a given approximation.

Theorem 5.2.1. *If f has a continuous derivative f' on the interval $[y, x]$, such that $0 < y < x$, then,*

$$\sum_{y<n\leq x} f(n) = \int_y^x f(t)dt + \int_y^x (t-[t])f'(t)dt + f(x)([x]-x) - f(y)([y]-y)$$

Proof: Let $m = [y]$ and $k = [x]$ for reference. For integers n and $n-1$ that are within the interval, $[x, y]$, we have that,

$$\begin{aligned}\int_{n-1}^{n} [t]f'(t)dt &= \int_{n-1}^{n} (n-1)f'(t)dt \\ &= (n-1)[f(n) - f(n-1)] = (nf(n) - (n-1)f(n-1)) - f(n)\end{aligned}$$

We note that the first equality is apparent upon realizing that there exists only one point $a \in [n-1, n]$ such that $[a] = n$, namely, n. And any integral from a point to itself is equal to zero.

Anyway, from the equation above we can say something about the integral over the interval [m, k]. Specifically we see that,

$$\begin{aligned}\int_m^k [t]f'(t)dt &= \sum_{n=m+1}^{k} (nf(n) - (n-1)f(n-1)) - \sum_{y<n\leq x} f(n) \\ &= kf(k) - mf(m) - \sum_{y<n\leq x} f(n)\end{aligned}$$

Thus, we have that,

$$\sum_{y<n\leq x} f(n) = -\int_m^k [t]f'(t)dt + kf(k) - mf(m)$$

Now I claim that $-\int_m^k [t]f'(t)dt + kf(k) - mf(m) = -\int_y^x [t]f'(t)dt + kf(x) - mf(y)$. To keep this proof rigorous, we will show this as a lemma.

Lemma 5.1.

$$-\int_m^k [t]f'(t)dt + kf(k) - mf(m) = -\int_y^x [t]f'(t)dt + kf(x) - mf(y)$$

Proof of Lemma:

Assign the following,

$$A = -\int_k^x [t]f'(t)dt = k\int_k^x f'(t)dt = k(f(x) - f(k))$$
$$B = -\int_m^y [t]f'(t)dt = m\int_m^y f'(t)dt = m(f(y) - f(m))$$

Notice that,

$$\int_m^k [t]f'(t)dt = \int_y^x [t]f'(t)dt - A + B$$

Thus we have that,

$$-\int_m^k [t]f'(t)dt + kf(k) - mf(m)$$
$$= -\int_y^x [t]f'(t)dt + (kf(k) + A) - (mf(m) - B)$$
$$= -\int_y^x [t]f'(t)dt + k(f(k) + f(x) - f(k)) - m(f(m) + f(y) - f(m))$$
$$= -\int_y^x [t]f'(t)dt + kf(x) - mf(y) \text{ Q.E.D.}$$

Now we use integration by parts to notice that, $-\int_y^x f(t)dt = xf(x) - yf(y) - -\int_y^x tf'(t)dt$. Thus, $0 = \int_y^x f(t)dt - xf(x) + yf(y) + \int_y^x tf'(t)dt$.

Adding that to our equation after substituting according to our lemma, we see that the sum, $\sum_{y<n\leq x} f(n)$

$$= \int_y^x f(t)dt + \int_y^x tf'(t)dt - \int_y^x [t]f'(t)dt + f(x)([x] - x) - f(y)([y] - y)$$
$$= \int_y^x f(t)dt + \int_y^x (t - [t])f'(t)dt + f(x)([x] - x) - f(y)([y] - y) \text{ Q.E.D.}$$

6 Shapiro's Tauberian Theorem

As discussed in our introduction, Euler proved the existence of infinite primes by showing that $\sum \frac{1}{p}$ diverges. Also aforementioned, we understand that to say, $\sum \frac{\log p}{p}$ diverges is an equivalent statement to Euler's (at least in the sense that both prove the existence of infinite primes). Of course, the

divergence of this series is important, since our main result, infinite primes in arithmetic progressions, is a restriction of this sum. In Section 7, we will discuss Harold N. Shapiro's Tauberian theorem, a particular case of which shows us that $\sum \frac{\log p}{p}$ diverges. Shapiro's theorem serves as the final tool we need to prove the divergence of $\sum \frac{\log p}{p}$. In the succeeding section, we will derive a result from our recently gathered tools (Sections 4-6) and use Shapiro's Tauberian theorem as a final step for this critical proof. Then we can modify the specific case of Shapiro's theorem to show our main desired result, $\sum_{p \equiv h \pmod{k}} \frac{\log p}{p}$ diverges.

6.1 Tauberian Theorems

Before we talk about Shapiro's theorem, one should know what a Tauberian theorem is.

Definition 6.1.1. A *Tauberian Theorem* is defined as a theorem that deduces the divergence of a series from the function it defines or is equivalent to.

6.2 Shapiro's Theorem

Shapiro's Tauberian Theorem actually gives us more information than we need. Only the first part of the three-part theorem will be necessary. Subsequently, we will only give a partial proof in this paper. The second part is proved in order to prove the first, and the third part is left without confirmation. However, the whole statement has been presented, and the curious reader may see Apostle's book for details.

The portion we are concerned with details the relationship between two series, which are defined on all n less than or equal to a given x. Note that we have seen many series of this form, and that is no coincidence. An arbitrary non-negative sequence, $\{a(n)\}$, is specified and appears in both sums. This is simply a function defined on the natural numbers with positive- or zero- valued outputs. The hypothesis requires that the first series, $\sum_{n \leq x} a(n) \left[\frac{x}{n}\right] = x \log x + O(x)$ for $x \geq 1$. Recall from Section 6, that this means that $\sum_{n \leq x} a(n) \left[\frac{x}{n}\right] - x \log x$ is bounded by Mx for some constant M . The conclusion states that we can drop the brackets in the first series, and the result is still valid. I believe the discussion within this paragraph will make more sense upon stating theorem.

6.3 Formal Statement and Partial Proof

Theorem 6.3.1. *Let $\{a(n)\}$ be a non negative sequence such that,*

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

for all $x \geq 1$.

Then the following 3 statements are true:

1. *For $x \geq 1$,*
$$\sum_{n\leq x} \frac{a(n)}{n} = \log x + O(1)$$

2. *There is a constant $B > 0$ such that for all $x \geq 1$,*
$$\sum_{n\leq x} a(n) \leq Bx$$

3. *There is a constant $A > 0$ and $x_0 > 0$ such that for all $x \geq x_0$,*
$$\sum_{n\leq x} a(n) \geq Ax$$

Proof:

We begin by proving part 2. of our theorem, which will serve as a lemma for part 1. As previously stated, we will not use the second or third parts of theorem, but they are stated here for the purpose of formality.

Let $S(x) = \sum_{n\leq x} a(n)$ and $T(x) = \sum_{n\leq x} a(n)\left[\frac{x}{n}\right]$. We consider the difference of $T(x) - 2T\left(\frac{x}{n}\right)$. Note that we are particularly interested in this difference because,

$$\begin{aligned} T(x) - 2T\left(\frac{x}{n}\right) &= x\log x + O(x) - 2\left(\frac{x}{2}\log\left(\frac{x}{2}\right) + O(x)\right) \\ &= x\left(\log x - \log\left(\frac{x}{2}\right)\right) + O(x) \\ &= x\left(\log\left(\frac{x}{\frac{x}{2}}\right)\right) + O(x) \\ &= x\log\left(\frac{1}{2}\right) + O(x) = O(x) \end{aligned}$$

So, if we can show that $S(x) - S\left(\frac{x}{2}\right) \leq T(x) - 2T\left(\frac{x}{n}\right)$, then we have shown that $S(x) - S\left(\frac{x}{2}\right) \leq O(x)$, which we will later use to deduce a conclusion about $S(x)$. We start by unpacking $T(x) - 2T\left(\frac{x}{n}\right)$ like so,

$$\begin{aligned} T(x) - 2T\left(\frac{x}{n}\right) &= \sum_{n \leq x} a(n)\left[\frac{x}{n}\right] - 2\sum_{n \leq \frac{x}{2}} a(n)\left[\frac{x}{2n}\right] \\ &= \sum_{n \leq \frac{x}{2}} a(n)\left(\left[\frac{x}{n}\right] - 2\left[\frac{x}{2n}\right]\right) + \sum_{\frac{x}{2} < n \leq x} a(n)\left[\frac{x}{n}\right] \end{aligned}$$

From here, we need a lemma to continue.

Lemma 6.1. $\sum_{n \leq \frac{x}{2}} a(n)\left(\left[\frac{x}{n}\right] - 2\left[\frac{x}{2n}\right]\right)$ *is non-negative.*

Proof of Lemma:

We consider the difference, $[2a] - 2[a]$, where a is arbitrary. Now let b be the greatest integer less than or equal to a. Let c be the difference $a - b$. Now we consider three cases,

1. $a = b$ thus, $c = 0$
2. $0 < c < \frac{1}{2}$
3. $c \geq \frac{1}{2}$

If a falls into case 1, then $[2a] - 2[a] = 2a - 2a = 0$. If a falls into case 2, then $[2a] - 2[a] = 2b - 2b = 0$. If a falls into case 3, then $[2a] - 2[a] = (2b + 1) - 2b = 1$. Therefor, $[2a] - 2[a]$ is never negative. The final step is to let $a = \frac{x}{2n}$, and we conclude that $\sum_{n \leq \frac{x}{2}} a(n)\left(\left[\frac{x}{n}\right] - 2\left[\frac{x}{2n}\right]\right)$ is non-negative, since $a(n)$ was specified to be non-negative in our hypothesis.

Now that our lemma is complete, we notice that,

$$\begin{aligned} T(x) - 2T\left(\frac{x}{n}\right) &= \sum_{n \leq \frac{x}{2}} a(n)\left(\left[\frac{x}{n}\right] - 2\left[\frac{x}{2n}\right]\right) + \sum_{\frac{x}{2} < n \leq x} a(n)\left[\frac{x}{n}\right] \\ &\geq \sum_{\frac{x}{2} < n \leq x} a(n)\left[\frac{x}{n}\right] \end{aligned}$$

This follows trivially from the definition of greater than or equal to, $\geq$.

Now we take a closer look at the final sum, $\sum_{\frac{x}{2}<n\leq x} a(n)\left[\frac{x}{n}\right]$. The trick here is to notice that $\left[\frac{x}{n}\right] > 1 \iff n \leq \frac{x}{2}$. Similarly, $\left[\frac{x}{n}\right] < 1 \iff n > x$. Since the series $\sum_{\frac{x}{2}<n\leq x} a(n)\left[\frac{x}{n}\right]$ runs through all terms such that, $\frac{x}{2} < n \leq x$, $\left[\frac{x}{n}\right] = 1$ in all terms. Thus we have that,

$$\sum_{\frac{x}{2}<n\leq x} a(n)\left[\frac{x}{n}\right] = \sum_{\frac{x}{2}<n\leq x} a(n) = S(x) - S\left(\frac{x}{2}\right)$$

Thus, we have shown that,

$$S(x) - S\left(\frac{x}{2}\right) \leq T(x) - 2T\left(\frac{x}{n}\right)$$

as desired. But now we turn to our hypothesis in order to conclude Part 2 of Theorem 6.3.1. Recall that $T(x) - 2T\left(\frac{x}{n}\right) = O(x)$ for all $x \geq 1$. Thus, for all $x \geq 1$ $S(x) - S\left(\frac{x}{2}\right) \leq Cx$ where C is a constant.

We now consider the sequence, x, $\frac{x}{2}$, $\frac{x}{4}$, $\frac{x}{8}$ and its relationship with our new inequality. We find that,

$$\begin{aligned} S(x) - S\left(\frac{x}{2}\right) &\leq Cx \\ S\left(\frac{x}{2}\right) - S\left(\frac{x}{4}\right) &\leq C\frac{x}{2} \\ S\left(\frac{x}{4}\right) - S\left(\frac{x}{8}\right) &\leq C\frac{x}{4} \\ &\vdots \end{aligned}$$

We simply add these inequalities to see that $S(x) \leq Cx\left(1 + \frac{1}{2} + \frac{1}{4} + ...\right)$. Note that $1 + \frac{1}{2} + \frac{1}{4} + ...$ is a geometric sequence with first term, 1, and common ratio, $\frac{1}{2}$. Thus it converges to 2. Thus $S(x) = \sum_{n\leq x} a(n) \leq 2C$. Thus, the Part 2. of Theorem 6.1.2. is proved with $B = 2C$.

Now we will prove Part 1 of our theorem. We rewrite $\left[\frac{x}{m}\right] = \frac{x}{n} = O(1)$. If it is unclear why we may do this, please review the Big Oh section (Section 5). Now we consider our series, $T(x)$.

$$T(x) = \sum_{n\leq x} a(n)\left[\frac{x}{n}\right]$$

$$= \sum_{n \leq x} a(n) \left(\frac{x}{n} + O(1) \right)$$
$$= x \sum_{n \leq x} \frac{a(n)}{n} = O\Big(\sum_{n \leq x} a(n) \Big)$$

Now, Part 2 of our theorem says that there exists a constant, B, such that, $\sum_{n \leq x} a(n) \leq Bx$ for all $x \geq 1$. Thus, for all $x \geq 1$ $O\big(\sum_{n \leq x} a(n)\big) = O(x)$. Thus,

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

Finally, we divide by x and replace $T(x)$ with its equal from our hypothesis to finish our proof of Part 1. and partial proof of this theorem. Observe,

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

Q.E.D.

7 Using Shapiro's Theorem to Prove Infinite Primes

We need to show that $\sum \frac{\log p}{p}$ diverges. The theorem we will prove actually says something a little stronger than this.

Theorem 7.0.1. *For all $x \geq 1$,*

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

I think it important to preface further discussion with an understanding of why we are proving this equation. Otherwise, it can be difficult to understand why we are limiting the sum to primes less than a particular x. We are after all trying to prove a statement about the existence of infinite primes. However, we note that $\log x \to \infty$ as $x \to \infty$, albeit extremely slowly. And while $O(1)$ may be a function of x, it is a bounded function. Thus, the divergence of $\sum \frac{\log p}{p}$ is a corollary of Theorem 7.0.1. To spell

it out, this is because as $x \to \infty$, $(\log x + O(1)) \to \infty$. Thus, as $x \to \infty$, $\sum_{p \leq x} \frac{\log p}{p} \to \infty$.

Now, we have already proved some useful results help with Theorem 7.0.1, but we will need to discuss a few more that are particular to this proof. Once that is completed, I will guide the reader through fitting each piece together and completing the proof. Ultimately, we will prove a necessary hypothesis that will be used in Shapiro's theorem to conclude that $\sum_{p \leq x} \frac{\log p}{p} = x \log x + O(x)$.

7.1 Don't Forget that Corollary

Section 5 was ended with an important corollary, and we mentioned that it would come up again. The corollary is as follows,

Corollary 2. *If $F(x) = \sum_{n \leq x} f(n)$ then,*

$$\sum_{n \leq x} \sum_{d|n} f(d) = \sum_{n \leq x} f(n) \left[\frac{x}{n} \right] = \sum_{n \leq x} F\left(\frac{x}{n} \right)$$

Suppose we let $f(n) = \Lambda(n)$. Note, that means that $F(x) = \sum_{d|n} \Lambda(d) = \log n$ by Theorem 3.4.2. What can we say about the sum, $\sum_{n \leq x} \Lambda(n) \left[\frac{x}{n} \right]$?

Theorem 7.1.1.

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

Proof:

The proof of this theorem directly follows from our Corollary. We see that,

$$\begin{aligned} \sum_{n \leq x} \Lambda(n) \left[\frac{x}{n} \right] &= \sum_{n \leq x} \sum_{d|n} \Lambda(d) \\ &= \sum_{n \leq x} F\left(\frac{x}{n} \right) \\ &= \sum_{n \leq x} \log n = \log[x]! \text{ Q.E.D.} \end{aligned}$$

This is a nice result, and our next two will be very related. In fact, the three will later be combined to achieve our goal.

7.2 Euler's Summation Formula and $\log[x]!$

The following theorem gives an asymptotic formula for $\log[x]!$. Of course, this means that it will give an asymptotic formula for $\sum_{n\leq x}\Lambda(n)\left[\frac{x}{n}\right]$ as well. This is the connection between these two results.

Theorem 7.2.1.

$$\log[x]! = x\log x - x + O(\log x)$$

Proof:

Recall that Euler's Summation formula states that, if f has a continuous derivative f' on the interval $[y,x]$, such that $0<y<x$, then,

$$\sum_{y<n\leq x} f(n) = \int_y^x f(t)dt + \int_y^x (t-[t])f'(t)dt + f(x)([x]-x) - f(y)([y]-y)$$

Let $f(t) =$.

Thus we see that,

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

Now, notice two things. The first is that $\int_1^x \frac{t-[t]}{t}dt$ has a max value of $\int_1^x \frac{1}{t}dt$. Thus, $\int_1^x \frac{t-[t]}{t}dt = O(\int_1^x \frac{1}{t}dt) = O(\log x)$. Also $1 = O(\log x)$, since $1 \leq \log x$ for x big enough. Therefor,

$$\begin{aligned}
\log[x]! &= x\log x - x + 1 + \int_1^x \frac{t-[t]}{t}dt + O(\log x) \\
&= x\log x - x + O(\log x) \text{ Q.E.D.}
\end{aligned}$$

There are two important corollaries that come out of this theorem. One will be a manipulation of the big-oh-term. The other relates theorem from the previous subsection to 8.2.1. to our most recent result.

Corollary 3.

$$\log[x]! = x\log x + O(x)$$

Proof:

Notice that functions within the big-oh-term, $O(\log x)$, are functions of $\log x$. Thus, they are functions of x. Hence, $x + O(\log x)$ is a function of x. Therefor, we can rewrite $x + O(\log x)$ as $O(x)$; it just gives us less information. Thus,

$$\log[x]! = x\log x - x + O(\log x) = x\log x + O(x) \text{ Q.E.D.}$$

Corollary 4.

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

Proof:

$\sum_{n\leq x} \Lambda(n)\left[\frac{x}{n}\right] = \log[x]!$ by Theorem 7.1.1. And, $\log[x]! = x\log x - x + O(\log x)$ by Theorem 7.2.1. By the previous corollary, $x\log x - x + O(\log x) = x\log x + O(x)$. Thus,

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

7.3 First Theorem about Primes

This entire thesis so far has proved many results about convolutions and other summations of functions of natural numbers. It is now time to introduce a theorem about primes. This theorem will be our final step toward Shapiro's result. Then we will be able to conclude the existence of infinite primes. More importantly, we will be able to prove a stronger result, that is critical to our main result.

Theorem 7.3.1. *For $x \geq 2$,*

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

where the sums are extended over all primes, $p \leq x$.

Proof:

Well, $\Lambda(n) = 0$ unless n is a prime or power of a prime. Thus,

$$\sum_{n \leq x} \Lambda(n) \left[\frac{x}{n} \right] = \sum_{p \leq x} \sum_{\substack{m \text{ s.t.} \\ p^m \leq x}} \left[\frac{x}{p^m} \right] \log p$$

Now, we notice that $p^m \leq x$ implies that $p \leq x$. Also, $\left[\frac{x}{p^m} \right]$ is zero-valued when $p^m > x$. We mention this so we can redefine the sum $\sum_{p \leq x} \sum_{\substack{m \text{ s.t.} \\ p^m \leq x}} \left[\frac{x}{p^m} \right] \log p$ as $\sum_{p \leq x} \sum_{m=1}^{\infty} \left[\frac{x}{p^m} \right] \log p$.

Now, we will isolate the terms of this sum that where the power of the prime input is 1. That is to say, we will isolate the terms with p^m where $m = 1$ like so,

$$\sum_{p \leq x} \sum_{m=1}^{\infty} \left[\frac{x}{p^m} \right] \log p = \sum_{p \leq x} \left[\frac{x}{p} \right] \log p + \sum_{p \leq x} \sum_{m=2}^{\infty} \left[\frac{x}{p^m} \right] \log p$$

Now if we can only prove that $\sum_{p \leq x} \sum_{m=2}^{\infty} \left[\frac{x}{p^m} \right] \log p \leq O(x)$, then we are set. Well,

$$\begin{aligned}
\sum_{p \leq x} \sum_{m=2}^{\infty} \left[\frac{x}{p^m} \right] \log p &\leq \sum_{p \leq x} \sum_{m=2}^{\infty} \frac{x}{p^m} \log p && (1) \\
&= x \sum_{p \leq x} \log p \sum_{m=2}^{\infty} \left(\frac{1}{p} \right)^m && (2) \\
&\leq x \sum_{p \leq x} \log p \left(\frac{1}{1 - \frac{1}{p}} \right) && (3) \\
&= x \sum_{p \leq x} \left(\frac{\log p}{1 - \frac{1}{p}} \right) && (4) \\
&\leq x \sum_{n \leq x} \left(\frac{\log n}{1 - \frac{1}{n}} \right) && (5)
\end{aligned}$$

Note that the third step is made possible by the fact that a geometric series converges to $\frac{a}{1 - \frac{1}{r}}$ where a is the initial term and r is the ratio between terms. $\sum_{m=2}^{\infty} \left(\frac{1}{p} \right)^m$ is a geometric series with $a = 1$ and $r = \frac{1}{p}$. However, the series is missing the first two terms. Since each term is positive, the inequality follows.

Now all that is left is to notice that $\sum_{n\leq x}\left(\frac{\log n}{1-\frac{1}{n}}\right)$ is a series with finitely many terms, so it is equal to some constant. Hence,

$$x\sum_{n\leq x}\left(\frac{\log n}{1-\frac{1}{n}}\right)=O(x)$$

and consequently,

$$\sum_{p\leq x}\sum_{m=2}^{\infty}\left[\frac{x}{p^m}\right]\log p=O(x)$$

Putting this all together, we see that,

$$\begin{aligned}\sum_{n\leq x}\left[\frac{x}{n}\right]\Lambda(n)&=\sum_{p\leq x}\left[\frac{x}{p}\right]\log p+\sum_{p\leq x}\sum_{m=2}^{\infty}\left[\frac{x}{p^m}\right]\log p\\&=\sum_{p\leq x}\left[\frac{x}{p}\right]\log p+O(x)\text{ Q.E.D.}\end{aligned}$$

The corollary for this is perhaps the most important result we have proven thus far. We will deduce an asymptotic formula for $\sum_{p\leq x}\left[\frac{x}{p}\right]\log p$, that will serve as our hypothesis statement to use in the Tauberian theorem.

Corollary 5. *For $x\geq 2$,*

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

Proof:

We know that $\sum_{n\leq x}\Lambda(n)\left[\frac{x}{n}\right]=x\log x+O(x)$ from a corollary of Theorem 7.2.1. We also know that $\sum_{n\leq x}\left[\frac{x}{n}\right]\Lambda(n)=\sum_{p\leq x}\left[\frac{x}{p}\right]\log p+O(x)$ from the latest theorem. This implies that $\sum_{n\leq x}\left[\frac{x}{n}\right]\Lambda(n)+O(x)=\sum_{p\leq x}\left[\frac{x}{p}\right]\log p$ as well. Thus,

$$\begin{aligned}\sum_{p\leq x}\left[\frac{x}{p}\right]\log p&=(x\log x+O(x))+O(x)\\&=x\log x+O(x)\text{ Q.E.D.}\end{aligned}$$

8 Particular Case of Shapiro's Theorem

Recall the hypothesis of Shapiro's Tauberian theorem. It says "let $\{a(n)\}$ be a non negative sequence such that, $\sum_{n\leq x} a(n)\left[\frac{x}{n}\right] = x\log x + O(x)$ for all $x \geq 1$." We note that $\log p$ is non-negative and $\sum_{p\leq x}\left[\frac{x}{p}\right]\log p = x\log x + O(x)$. So, $\log p$ almost meets the requirement for Shapiro's Tauberian theorem. Unfortunately, our function, $a_0(p) = \log p$, is only defined on the primes and thus is not a sequence. However, we may extend it into a sequence and just let it equal 0 for all non-prime n. This would not mess with the sum, but would make our function applicable to Shapiro's theorem.

Definition 8.0.1. We define the function $\Lambda_1(n)$ as follows,

$$\Lambda_1(n) = \begin{cases} \log n & n \text{ is prime} \\ 0 & \text{else} \end{cases}$$

Notice that,

$$\sum_{n\leq x}\left[\frac{x}{n}\right]\Lambda_1(n) = \sum_{p\leq x}\left[\frac{x}{p}\right]\log p = x\log x + O(x)$$

Now, we will finally prove the divergence of $\sum\frac{\log p}{p}$ as a corollary to the stronger theorem below.

Theorem 8.0.2. *For all $x \geq 1$,*

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

Proof:

We have already shown that Λ_1 satisfies the condition for our hypothesis. Now, we turn our attention to the first part of the conclusion of Shapiro's theorem. It says, "$\sum_{n\leq x}\frac{a(n)}{n} = \log x + O(1)$." Thus,

$$\sum_{n\leq x}\frac{\Lambda_1(n)}{n} = \log x + O(1)$$

But, just as we expanded $\sum_{p\leq x}\left[\frac{x}{p}\right]\log p$ into $\sum_{n\leq x}\left[\frac{x}{n}\right]\Lambda_1(n)$, we now restrict the sum $\sum_{n\leq x}\frac{\Lambda_1(n)}{n}$ to the primes. This is possible, since $\Lambda_1(n) = 0$ for all composite n, as we defined it. Thus,

$$\sum_{p \leq x} \frac{\log p}{p} = \sum_{n \leq x} \frac{\Lambda_1(n)}{n} = \log x + O(1)$$

Q.E.D.

Corollary 6.

$$\sum \frac{\log p}{p} \textit{ diverges.}$$

Proof:

Note that $\log x \to \infty$ as $x \to \infty$. well,

$$\sum \frac{\log p}{p} = \lim_{x \to \infty} \sum_{p \leq x} \frac{\log p}{p} = \lim_{x \to \infty} (\log x + O(1)) = \lim_{x \to \infty} \log x = \infty$$

Thus, $\sum \frac{\log p}{p}$ diverges as desired. Q.E.D.

Note that divergence of $\sum \frac{\log p}{p}$ also implies the existence of infinite primes. This is because there must be infinitely many terms of the form $\frac{\log p}{p}$ for the sum to be infinite. Each term is a function of a unique prime, since that is how summation works. Thus, we conclude this section with evidence that there must be infinitely many prime numbers.

9 Arithmetic Progressions

We know that there are infinitely many primes, but this has been long established, since the days of Euclid. Our more interesting goal is to prove infinitely many primes in arithmetic progression. We will discuss the definition of an arithmetic progression once the appropriate preparatory material has been presented.

9.1 Modular Congruence

A basic understanding of congruence classes and modular arithmetic is assumed in this paper. However, we will go over some of the basics and the properties for review. theory of congruence was established by Gauss in order to simplify work with divisibility of integers. It has since become one of the central branches of number theory. We begin with a definition.

Definition 9.1.1. Let $a, b, k \in \mathbb{Z}$ such that $k > 0$. We say that a is *congruent to b modulo k* if $k|(a-b)$. We write this as,

$$a \equiv b \pmod{k}$$

We call k the **modulus** of this congruence.

We will now say some things about modular congruence. First, notice that $a \equiv 0 \pmod{k} \iff k|a$. Similarly, $a \equiv b \pmod{k} \iff (a-b) \equiv 0 \pmod{k}$. Congruence has many other interesting qualities, and a whole other paper about this relation could be written. However, we will state some of the more basic properties, since congruence is not where the complexity of our result lies. We begin with the following theorem.

Theorem 9.1.2. *Modular Congruence is an equivalence relation. That is to say,*

- *It is reflexive;* $a \equiv a \pmod{k}$ *for all* $a \in \mathbb{Z}$
- *It is symmetric;* $a \equiv b \pmod{k} \implies b \equiv a \pmod{k}$
- *And it is transitive;* $a \equiv b \pmod{k}$ *and* $b \equiv c \pmod{k} \implies a \equiv c \pmod{k}$

Proof:

That the relation is reflexive is trivial; 0 is divisible by k for all $k \neq 0$ and k was specified as non-zero in the definition. That it is symmetric is almost trivial; $k|(a-b) \iff k|(b-a)$. That it is transitive is the only property that is not crazy obvious; $k|(a-b)$ and $k|(b-c) \implies k|((a-b)+(b-c)) = a-b+b-c = a-c$. Q.E.D.

We obviously used the properties of the integers, which are assumed to be true and known by the reader.

The next theorem that we will prove involves adding, subtracting, and multiplying modular congruences of the same modulus, k. We will see that modular equivalence on $\mathbb{Z}$ has many of the same properties as the standard equality, "=," on the set of all quantities.

Theorem 9.1.3. *Assume that $a \equiv b \pmod{k}$ and $c \equiv d \pmod{k}$. Let x and y be arbitrary integers. Let n be an arbitrary natural number. Let f be an arbitrary polynomial with integer coefficients. Then the following are true:*

1. $(ax + cy) \equiv (bx + dy) \pmod{k}$;
2. $ac \equiv bd \pmod{k}$;
3. $a^n \equiv b^n \pmod{k}$
4. $f(a) \equiv f(b) \pmod{k}$

Proof:

1. $a \equiv b \pmod{k}$ and $c \equiv d \pmod{k}$ implies that $k|(a-b)$ and $k|(c-d)$. Thus, $k|x(a-b)$ and $k|y(c-d)$. Hence, $k|(x(a-b)+y(c-d)) = ax+cy-bx-dy = (ax+cy)-(bx+dy)$. Thus, $(ax+cy) \equiv (bx+dy) \pmod{k}$ as desired.

2. Note, $ac-bd = c(a-b)+b(c-d)$, and by part 1, $(c(a-b)+b(c-d)) = (ac+cb)-(bc+db) \equiv 0 \pmod{k}$. Thus, $ac-bd \equiv 0 \pmod{k}$. Thus, $ac \equiv bd \pmod{k}$ as desired.

3. Let $c = a$ and $d = b$ from part 2. Then we induct on n for the result.

4. We note that a polynomial with integer coefficients is just an expression with addition (and subtraction but negatives are ignored in modular congruence), scalar multiplication, and non-negative powers. Since we have already proved each of these operations individually in parts 1-3, we implement them one by one to prove part 4.

Q.E.D.

9.2 Residue Classes

Now that we have some properties of modular arithmetic, we are ready to discuss the concept of *residue classes*. I first would like to give the reader an idea of why we use the term, "residue," since I think that is useful. It suggests something left over after an event. In this case the "residue" is the remainder after dividing by the modulus. We will allude to this with a theorem.

Theorem 9.2.1. *$a \equiv b \pmod{k}$ if and only if a and b yield the same remainder when divided by k.*

Proof:

Let $a = kq + r$ and $b = kQ + R$, where q, r, Q, R are all integers and $0 \leq r, R < k$. Thus, $0 \leq |r-R| < k$. By Theorem 9.1.3., $(a-b) \equiv (r-R) \pmod{k} \implies r \equiv R \pmod{k} \implies k|(r-R)$. Thus, $k \leq |r-R|$ unless $r = R$. Since we already deduced that $k > |r-R|$, we conclude that $r = R$.

The converse of this is much easier. If $r = R$, then $r - R = 0$ and $k|(r-R)$. Thus, $r \equiv R \pmod{k}$. Of course, $k|kq$ and $k|kQ \implies kq \equiv$

$0 \pmod{k} \equiv kQ \pmod{k}$. Thus, $a \equiv kq + r \equiv kQ + R \equiv b$ all $\pmod{k}$ of course.
Q.E.D.

Now that we have an understanding of where the term *residue class* comes from, we need to define them. We will do so formally, but a residue class is simply an equivalence class on the integers defined by modular congruence. As suggested by the latest theorem, each class yields the same remainder when dividing by the modulus.

Definition 9.2.2. Let $k > 0$ be an arbitrary modulus. The *residue class* of an arbitrary $a \pmod{k}$ is the set, $\{x : x \equiv a \pmod{k}\}$. We denote this class with $\hat{a}$.

Residue classes have all the properties of equivalence classes, which we will state without proof here.

Theorem 9.2.3. *For a given modulus, k, the following are true:*

1. $\hat{a} = \hat{b}$ *if and only if* $a \equiv b \pmod{k}$
2. *Let* $x, y \in \mathbb{Z}$. *Then* x *and* y *are in the same residue class if and only if* $x \equiv y \pmod{k}$
3. *There are k residue classes* $\hat{0}, \hat{1}, \hat{2}, ..., k \hat{-} 1$, *they are disjoint, and their union is* $\mathbb{Z}$, *the set of all integers.*

Now, we will talk about congruence $\pmod{k}$ and residue classes, a when $(a, k) = 1$. We will use residue classes to talk about the infinitely many integers that are congruent $\pmod{k}$ as one entity. We make the convention (for now) that integers within the same congruence class are not distinct. Adjusting our thinking like this, we are allowed the following theorem.

Theorem 9.2.4. *Assume* a, b *are integers. They may be distinct (by our new definition) or not. Also, assume that* $(a, k) = 1$. *Then the congruence,*

$$ax \equiv b \pmod{k}$$

has exactly one distinct solution, x. *That is to say, all solutions to the congruence are in the same residue class.*

Proof:

We only need to test the numbers $\{1, 2, 3,...,k\}$, since this set contains representative integers from all the residue classes $(\text{mod } k)$. We multiply each element from this set by a. We have created the set, $\{1a, 2a, 3a,...,ka\}$. We claim that this set also contains representative integers from all the residue classes.

Assume for a contradiction that $ia \equiv ja \pmod{k}$ for some $i, j \in \{1, 2, 3,...,k\}$ such that $i \neq j$. Then, $i \equiv j \pmod{k}$ since $(a, k) = 1$. But this can't be true since i and j are distinct elements of a complete residue system, so $i \not\equiv j \pmod{k}$. By contradiction, we conclude that $\{1a, 2a, 3a,...,ka\}$ contains representative integers from all residue classes.

Hence, exactly one of the elements in $\{1a, 2a, 3a,...,ka\}$ is congruent to $b \pmod{k}$. Thus, there is a unique element from $\{1, 2, 3,...,k\}$ that can be multiplied by a to get b. Q.E.D.

For further details on residue classes, please see Apostle's book or other resources. While, the topic is quite interesting, we cannot get bogged down by it for this paper.

9.3 Reduced Residue System

We will now discuss the concept of a *reduced residue system*. A reduced residue system modulo k is a subset of the residue classes modulo k. It is very interconnected to Euler's totient function, φ, which was introduced in Section 4. This is because it's members are relatively prime to the modulus. Here is the formal definition,

Definition 9.3.1. A **reduced residue system** modulo k is any set of $\varphi(k)$ many integers that are incongruent to each other modulo k and are relatively prime to k.

Theorem 9.3.2. *If* $\{a_1, a_2, a_3, ..., a_{\varphi(k)}\}$ *is a reduced residue system and* $(c, k) = 1$, *then* $\{ca_1, ca_2, ca_3, ..., ca_{\varphi(k)}\}$ *is also a reduced residue system modulo k.*

Proof:

From Theorem 9.1.3., we know that $ca_i \not\equiv ca_j$ for all $i, j \in \{1, 2, 3, ..., \varphi(k)\}$ unless $i = j$. Also, $(a_i, k) = 1$ and $(c, k) = 1$, so $(ca_i, k) = 1$. Thus, $\{ca_1, ca_2, ca_3, ..., ca_{\varphi(k)}\}$ is a reduced residue system modulo k. Q.E.D.

Theorem 9.3.3. *Assume* $(c, k) = 1$, *then,*

$$c^{\varphi(k)} \equiv 1 \pmod{k}$$

Proof:

Let $\{a_1, a_2, a_3, ..., a_{\varphi(k)}\}$ be a reduced residue system modulo k. As we just proved, $\{ca_1, ca_2, ca_3, ..., ca_{\varphi(k)}\}$ is also a reduced residue system modulo k. So, for all $i \in \{1, 2, 3, ..., \varphi(k)\}$, $ca_i \equiv a_j \pmod{k}$ for some $j \in \{1, 2, 3, ..., \varphi(k)\}$. Thus,

$$\prod_{i=1}^{\varphi(k)} ca_i \equiv \prod_{j=1}^{\varphi(k)} a_j \pmod{k}$$

Thus,

$$c^{\varphi(k)} \times \prod_{i=1}^{\varphi(k)} a_i \equiv \prod_{j=1}^{\varphi(k)} a_j \pmod{k}$$

Since i and j are just dummie variables, and we can cancel out $\prod_{i=1}^{\varphi(k)} a_i$ and $\prod_{j=1}^{\varphi(k)} a_j$ to obtain our result. Q.E.D.

9.4 Definition

We are now ready to state the formal definition for an arithmetic progression and discuss a relationship with groups that will be helpful in future results.

Definition 9.4.1. An **arithmetic progression** with first term h and common difference k consists of all numbers of the form, $kn + h$, where $n \in \mathbb{N} \cup \{0\}$. We will denote this progression with $AP(k, h)$.

This might seem familiar, since it is simply all the positive members of the residue class, $\hat{k}$. We will of course attempt to prove infinite primes within an arbitrary $AP(k, h)$. We have actually already done so for a particular progression, namely $AP(2, 1)$. This is all of the odd positive integers. Since the even numbers contain no prime numbers other than 2, and we proved the existence of infinite primes, all infinitely many primes must be somwhere in the progression $AP(2, 1)$.

Now there is one important thing to mention when trying to prove infinite primes within an arithmetic progression. We require that $(k, h) = 1$. Let's assume that $(k_0, h_0) = d$ for some $d \neq 1$ and try to prove infinite primes. Well every term in $AP(k_0, h_0)$ must be divisible by d since each term is a linear combination of k_0, h_0, which are both divisible by d. Thus, there are no primes in $AP(k_0, h_0)$ other than perhaps h_0. So when we prove infinite primes in arithmetic progression in a later section we will keep this in mind and require that $(k, h) = 1$.

10 Dirichlet Characters

10.1 Reduced Residue Systems as Groups

It is expected that the reader have an understanding of groups before beginning this thesis. We will simply state the axioms here and not much else. We will then prove that a reduced residue system modulo k is a group with order $\varphi(k)$.

Definition 10.1.1. A **group**, G, is a nonempty set of elements in tandem with a binary operation, which we denote with $\circ$. The following axioms must be met in order for the set and operation to be considered a group.

1. Closure; for all $a, b \in G$, $a \circ b \in G$
2. Associativity; for all $a, b, c \in G$, $(a \circ b) \circ c = a \circ (b \circ c)$
3. Existence and Uniqueness of Identity; there exists a unique element, $e \in G$, such that for all $a \in G$, $a \circ e = e \circ a = a$
4. Existence and Uniqueness of Inverses; for all $a \in G$, there exists a unique element $a^{-1} \in G$, such that $a \circ a^{-1} = a^{-1} \circ a = e$

Now there is a special type of group that we should mention. A reduced residue system forms this type of group with a special but not unfamiliar type of multiplication, which we will define soon.

Definition 10.1.2. An **abelian group** is a group, G, such that all elements commute under the operation, $\circ$. This is to say that $a \circ b = b \circ a$ for all $a, b \in G$

Before we can prove that a reduced residue system modulo k can form a group, we have to define our operation to accompany our $\varphi(k)$ elements.

Definition 10.1.3. Let $\hat{a}$ and $\hat{b}$ be two arbitrary residue classes modulo k. Define,

$$\hat{a} \odot \hat{b} = \widehat{ab}$$

That is, the product, $\hat{a}\hat{b}$, is defined as the class $\widehat{ab}$.

Theorem 10.1.4. *This multiplication that we have defined for residue classes is well-defined.*

Proof:

Note that the equivalence relation of modular congruence between classes almost immediately implies that this type of multiplication is well-defined. Observe, that if $\hat{a} = \hat{\alpha}$ and $\hat{b} = \hat{\beta}$ then $a \equiv \alpha \pmod{k}$ and $b \equiv \beta \pmod{k}$. By part 2. of Theorem 9.1.3., we see that $ab \equiv \alpha\beta \pmod{k}$. Thus, $\widehat{ab} = \widehat{\alpha\beta}$. Thus, $\hat{a} = \hat{\alpha}$ and $\hat{b} = \hat{\beta} \implies \hat{a} \odot \hat{b} = \hat{\alpha} \odot \hat{\beta}$. Q.E.D.

Theorem 10.1.5. *The set of reduced residue classes modulo k forms a finite abelian group, G, with order, $\varphi(k)$, and operation, $\odot$.*

Proof:

First, we must show that the elements and operation meet the four group axioms.

1. Assume $\hat{a}$ and $\hat{b}$ are reduced residue classes modulo k. Thus, $(a,k) = 1$ and $(b,k) = 1$. Thus, $(ab,k) = 1$. Thus $\hat{a} \odot \hat{b} = \widehat{ab}$ is a reduced residue class modulo k.

2. Since multiplication of residue classes, $\odot$, is defined by regular multiplication of integers, the associative property is immediate.

3. Consider the residue class $\hat{1}$. Clearly, $\hat{1} \odot \hat{a} = \hat{a}$ for all reduced residue classes, $\hat{a}$. Assume for a contradiction that this is not unique. Thus, there exists some other element $\hat{e_0}$ such that $\hat{e_0} \odot \hat{a} = \hat{a}$ for all $\hat{a} \in G$. Well then, $\widehat{e_0 a} = \hat{e_0} \odot \hat{a} = \hat{a} = \hat{1} \odot \hat{a} = \widehat{1a}$. Thus, $e_0 a \equiv a \pmod{k}$ and $1a \equiv a \pmod{k}$. But, by Theorem 9.2.4., since $(a,k) = 1$, $1a$ and $e_0 a$ cannot be congruent to the same thing unless $1 \equiv e_0 \pmod{k}$. Thus, $\hat{1} = \hat{e_0}$, a contradiction. Thus, $\hat{1}$ is unique.

4. By the same logic and using the same theorem, we see that for each $\hat{a} \in G$, there is a unique element $\hat{a^{-1}} \in G$ such that $\hat{a} \odot \hat{a^{-1}} = \hat{1}$.

We only have two more steps to finish our proof. We note that $|G| = \varphi(k)$, so G is finite. We also see that G is abelian; like associativity, commutativity carries over from multiplication of integers. Thus, G is a finite abelian group. Q.E.D.

10.2 Character Functions

We will revisit the group of reduced residue class. However, we first need to talk about something more generally: character functions. We will end

this section by defining particular type of these functions called Dirichlet characters. They will be used to restrict our sum from Shapiro's theorem. However, there are many types of character functions, and they have some useful properties.

Definition 10.2.1. Let G be an arbitrary group. Let f be a complex-valued function defined on G. Then f is a **character** of G if f has the multiplicative property,

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

for all $a, b \in G$, and $f(c) \neq 0$ for any element, $c \in G$.

Theorem 10.2.2. *If f is a character function of G such that G is a finite group, then, $f(e) = 1$ where e is the identity element in G and $f(a)$ is a root of unity for all $a \in G$. Furthermore, if $a^n = e$, then $f(a)^n = 1$.*

Proof:

Recall, that $f(c) \neq 0$ for all $c \in G$. Note that $ce = c$ for all $c \in G$ as well. By the multiplicative property of character functions,

$$f(c)f(e) = f(ce) = f(c)$$

Thus, $f(e) = 1$ as desired. Also by the multiplicative property, if $a^n = e$, then

$$f(a)^n = f(a^n) = f(e) = 1 \text{ Q.E.D.}$$

The following theorem gives us information about the number of character functions of a finite abelian group. We will omit the proof of this in the interest of time, but it is a well established result. The curious reader may refer to Apostle's book or other resources to find the justification.

Theorem 10.2.3. *If G is a finite abelian group of order n, then there are exactly n distinct character functions of G.*

The proof of this theorem mostly follows from the multiplicative property of character functions, but as mentioned, we will not go into details. However, we will prove the next theorem.

Theorem 10.2.4. *The set of characters of a finite abelian group, G, form a finite abelian group in their own right with the operation defined below:*

$$(f_i f_j)(a) = f_i(a)f_j(a) \text{ for all } a \in G.$$

Proof:

1. Closure; We need to show that $f_i f_j$ is a character if f_i and f_j are. Well,

$$\begin{aligned}(f_i f_j)(ab) &= f_i(a) f_i(b) f_j(a) f_j(b) \\ &= f_i(a) f_j(a) f_i(b) f_j(b) \\ &= (f_i f_j)(a)(f_i f_i)(b)\end{aligned}$$

   Thus, $f_i f_j$ is a character function for arbitrary $a, b \in G$.

2. Associativity; This property follows directly from the associative property of $\mathbb{C}^*$

3. Identity; The principal character defined by $f_1(a) = 1$ for all $a \in G$ is clearly a character function of all groups. This function is our identity element in the character group, $\hat{G}$. Notice, that it is unique since any hypothetical other function that acted as an identity element on this set would have to be defined in the same way.

4. Inverses; We define the inverse of a given character, f, to be $\frac{1}{f}$. Since G is finite and abelian, for each character, f, $|f(a)| = 1$ for all $a \in G$. Thus, the reciprocal of $f(a)$ is equal to the complex conjugate of $f(a)$, $\overline{f(a)}$. Well this is also a character of G and $f(a)\overline{f(a)} = 1$ for all a. Thus, $f(a)\overline{f(a)} = f_1(a)$. So the function f^{-1}, defined by $f^{-1}(a) = \overline{f(a)} = \frac{1}{f(a)} = f(a^{-1})$ is the inverse of f in $\hat{G}$.

Finally, the last theorem states that $|G| = |\hat{G}|$, so $\hat{G}$ is finite. And the commutative property transfers over from the commutative property of $\mathbb{C}^*$. Thus, $\hat{G}$ is a finite abelian group. Q.E.D.

10.3 Orthogonality Relations

Before diving into the orthogonal relations for characters, we must explain the following notation. In this section, we let G be a finite abelian group of order n. The elements of G are denoted by the set, $\{a_1, a_2, a_3, ..., a_n\}$ and the character function group $\hat{G} = \{f_1, f_2, f_3, ..., f_n\}$. We now create an $n \times n$ matrix A such that the entry a_{ij} in the ith row and jth column is $f_i(a_j)$. Now we are ready to state the following theorems.

Theorem 10.3.1. *The sum of the elements in the ith row of A is equal to the following,*

$$\sum_{r=1}^{n} f_i(a_r) = \begin{cases} n & \text{if } i = 1 \text{ } (f_i \text{ is the principal character)} \\ 0 & \text{else} \end{cases}$$

Proof:

Case 1: Let $S = \sum_{r=1}^{n} f_i(a_r)$. If $i = 1$ then each term of the sum is equal to 1. Thus, $S = n$ and Case 1 is proven.

Case 2: Now, if $i \neq 1$, then we need a trick. We will use the property of zero that tells us that if α is not the identity and $\alpha \times \beta = \beta$, then $\beta = 0$. We note that if $i \neq 1$, then there exists an element $b \in G$ such that $f_i(b) \neq 1$. Since G is finite, the set of terms in the sequence (a_r) is equivalent to that of the sequence, (ba_r). Thus, $\sum_{r=1}^{n} f_i(a_r) = \sum_{r=1}^{n} f_i(ba_r)$. Hence, we have the following equation,

$$S = \sum_{r=1}^{n} f_i(ba_r) = f_i(b) \sum_{r=1}^{n} f_i(a_r) = f_i(b)S.$$

Thus by the aforementioned property of zero with $\alpha = f_i(b)$ and $\beta = S$, we have that $S = 0$. Thus, Case 2 is proven. Therefore,

$$\sum_{r=1}^{n} f_i(a_r) = \begin{cases} n & \text{if } i = 1 \text{ } (f_i \text{ is the principal character)} \\ 0 & \text{else} \end{cases} \quad \text{Q.E.D.}$$

Theorem 10.3.2. *Let A^* denote the conjugate transpose of A. Then the following equation holds,*

$$AA^* = nI$$

where I is the identity matrix of dimensions $n \times n$.

Proof:

Denote the matrix AA^* with $B = AA^*$. Let us look at an arbitrary entry b_{ij} of B. We multiply the row and column vectors to get the entry in the following way:

$$b_{ij} = \sum_{r=1}^{n} f_i(a_r)\overline{f_j(a_r)} = \sum_{r=1}^{n} f_i(a_r)f_j^{-1}(a_r)$$

$$= \sum_{r=1}^{n} (f_i f_j^{-1})(a_r) = \sum_{r=1}^{n} f_k(a_r)$$
$$= \begin{cases} n & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

Therefor, we see that the only non-zero entries of the $n \times n$ matrix, B are those in the diagonal from top-left to bottom-right. The non-zero entries are all equal to n. Thus, $B = nI$.

Corollary 7. *$n^{-1}A^*$ is the inverse of A.*

Proof:

This proof is trivial. We simply notice that $A = (A^*)^*$. Thus $A^*A = A^*(A^*)^* = nI$ as well. This allows us to get around the issue of left/right multiplication and conclude the statement of our corollary.

Theorem 10.3.3.

$$\sum_{r=1}^{n} \overline{f_r(a_i)} f_r(a_j) = \begin{cases} n & \textit{if } i = 1 \textit{ (f_i is the principal character)} \\ 0 & \textit{else} \end{cases}$$

Proof:

We already showed that if $AA^* = nI$ then $A^*A = nI$ also. Well, the element in the ith row and jth column of A^*A is exactly the sum on the left of our equation. The equation in Theorem 10.3.3. follows. Q.E.D.

Corollary 8. *The sum of the entries in the jth column of A is equal to the following,*

$$\sum_{r=1}^{n} f_r(a_j) = \begin{cases} n & \textit{if } a_j = e \\ 0 & \textit{else} \end{cases}$$

Proof:

Note that $\overline{f_r(a_j)} = (f_r(a_i))^{-1} = f_r(a_i^{-1})$ Thus, we have that the sum form Theorem 10.3.3. is equal to the following,

$$\sum_{r=1}^{n} \overline{f_r(a_i)} f_r(a_j) = \sum_{r=1}^{n} f_r(a_i^{-1}) f_r(a_j) = \sum_{r=1}^{n} f_r(a_i^{-1} a_j)$$

It follows that,

$$\sum_{r=1}^{n} f_r(a_i^{-1}a_j) = \begin{cases} n & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

Thus, we conclude that,

$$\sum_{r=1}^{n} f_r(a_j) = \begin{cases} n & \text{if } a_j = e \\ 0 & \text{else} \end{cases}$$

Q.E.D.

10.4 Definition

As we showed earlier in this section, the set of reduced residue classes with modulus k forms a finite abelian group of order $\varphi(k)$. Let us denote this group with R, and let RR be the character group of R. So the characters are functions from $R = U(\mathbb{Z}_k) \to \mathbb{C}^*$. Dirichlet characters extend the domain of character functions to all of $\mathbb{N}$. This is really important because when we are questioning infinite primes in arithmetic progressions, we obviously must concern ourselves with all integers in that arithmetic progression not just the representative element of the congruence class, which is that arithmetic progression. This comment will make more sense down the line.

Definition 10.4.1. Let f_i be a character of R; the Dirichlet character associated with f_i is the arithmetic function, χ_i, defined by:

$$\chi_i(n) = \begin{cases} f_i(\hat{n}) & \text{if } (n,k) = 1 \\ 0 & \text{else} \end{cases}$$

We point out that this definition obviously implies two things. The first is that there are exactly $\varphi(k)$ distinct Dirichlet characters modulo k. The second is that,

$$\chi_1(n) = \begin{cases} 1 & \text{if } (n,k) = 1 \\ 0 & \text{else} \end{cases}$$

There are a few elementary but interesting facts about Dirichlet characters that we should point out.

Theorem 10.4.2. *Dirichlet characters are completely multiplicative and periodic with period k. In mathematical notation, this means that for a given Dirichlet character, χ,*

$$\chi(mn) = \chi(m)\chi(n)$$

and

$$\chi(n) = \chi(n+k) \text{ for all } n \in \mathbb{N}.$$

Proof:

Multiplicative: If $(m,k) > 1$ then $(mn,k) > 1$. Thus, $\chi(m)\chi(n) = 0 = \chi(mn)$. The equation is identical if $(n,k) > 1$. If $(m,k) = 1$ and $(n,k) = 1$, then $\chi(mn) = f(\widehat{mn}) = f(\hat{m})f(\hat{n}) = \chi(m)\chi(n)$. All cases accounted for, we see that Dirichlet characters are multiplicative.

Periodic: Assume $\chi(n) = 0$ for a given n. Thus, $(n,k) = d > 1$, then $d|n$ and $d|k$. Thus, $d|(n+k)$. Thus, $((n+k),k) = d > 1$ and $\chi(n+k) = 0$. Now assume $\chi(n) \neq 0$. Then, $\chi(n) = f(\hat{n})$. Well, since $(n+k) \in \hat{n}$, we have that $\chi(n) = f(\hat{n}) = \chi(n+k)$.

Thus, Dirichlet characters are multiplicative and periodic with period k. Q.E.D.

10.5 Orthogonality and Dirichlet Characters

The following theorem details an important result that we will use in our main result. It relates the orthogonality relation for character functions to Dirichlet characters themselves.

Theorem 10.5.1. *Let $\chi_1, \chi_2, \chi_3, ..., \chi_{\varphi(k)}$ denote the Dirichlet characters modulo k, and let m and n be two integers such that $(n,k) = 1$. Then the following holds,*

$$\sum_{r=1}^{\varphi(k)} \chi_r(m)\overline{\chi_r(n)} = \begin{cases} \varphi(k) & \text{if } m \equiv n \pmod{k} \\ 0 & \text{if } m \not\equiv n \pmod{k} \end{cases}$$

Proof:

Case 1: If $(m,k) = 1$, then $\chi_r(m)\overline{\chi_r(n)} = f_r(\hat{m})\overline{f_r(\hat{n})}$. The case follows from Theorem 10.3.3., since $m \equiv n \pmod{k} \iff \hat{m} = \hat{n}$.

Case 2: If $(m,k) \neq 1$, then $m \not\equiv n \pmod{k}$, since our hypothesis specified that n and k were relatively prime. Also, by the definition of Dirichlet character, each term of our sum on the left is zero.

Putting these two cases together, we see that our equation holds. Q.E.D.

11 Sums involving Dirichlet Characters

In this section, we will discuss Dirichlet characters and their relationship with series. Throughout Shapiro's proof of Dirichlet's theorem, we will need these results to get from one lemma to the next. The majority of these theorems show convergence of sums involving Dirichlet characters or that other sums of the same nature are non-zero. Convergence allows one to absorb constants (which are the limits of the convergent series) into different big-oh-terms. On the other hand, the non-zero property is necessary when dividing both sides of an equation, since dividing by zero is obviously undefined.

11.1 Some Convergence Theorems

The next theorem is a general theorem that we will use to prove the convergence of some particular series that we need for our main result. It uses Abel's identity, which we will assume to be true in the interest of the reader's time.

Theorem 11.1.1. *Let χ be a non-principal Dirichlet character modulo k. Let f be a non-negative function with a continuous negative derivative, $f'(x)$, for all x large enough. Let $y \geq x$. Then the following equation holds,*

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

Proof:

Consider the sum, $S(k) = \sum_{n=1}^{k} \chi(n)$. Since χ is non-principal by supposition, we have that this sum is equal to zero by Theorem 10.3.1. Note that the n that are not relatively prime are equal to zero and the sum of those that are relatively prime meets the hypothesis for the 11.3.1. Since Dirichlet characters are periodic, it fallows that $S(mk) = \sum_{n=1}^{mk} \chi(n) = 0$ for all $m \in \mathbb{N}$. Now, since $\chi(n)$ is a root of unity or zero for all n by Theorem 10.2.2., we have that $S(x)$ is bounded, specifically by $\varphi(k)$. Thus, $S(x) = O(1)$.

Now Abel's identity tells us that for any arithmetical function $a(n)$ and the corresponding sum, $A(x) = \sum_{n\leq x} a(n)$ and if $g(x)$ has a continuous derivative on the interval $[y, x]$ such that $0 < y < x$, then,

$$\sum_{x<n\leq y} a(n)g(n) = A(x)g(x) - A(y)g(y) - \int_{y}^{x} A(t)g'(t)dt$$

It follows that

$$\begin{aligned}\sum_{x<n\leq y}\chi(n)f(n) &= S(x)f(x) - S(y)f(y) - \int_y^x S(t)f'(t)dt \\ &= O(1)f(x) - O(1)f(y) - \int_y^x O(1)f'(t)dt \\ &= O(f(y)) + O(f(x)) + O\left(\int_y^x f'(t)dt\right) \\ &= O(f(y)) + O(f(x)) + O((O(1)f(x) - O(1)f(y)) \\ &= O(f(y)) + O(f(x)) + O(f(x)) = O(f(x)) \text{ Q.E.D.}\end{aligned}$$

Corollary 9. *If in addition to all assumption made above, $f(x) \to 0$ as $x \to 0$, then $\sum_{n=1}^{\infty}\chi(n)f(n)$ converges, and for x big enough we have that,*

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

Proof:

We note that,

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

We just proved that the second term on the right is equal to $O(f(x))$. Thus, our the second statement of our corollary is proven. The first statement is true by Cauchy's convergence test. Thus, we are done. These follow naturally from the previous theorem and corollary, so there is no need for a proof.

We are now ready to apply this theorem to some specific series that concern our main result.

Theorem 11.1.2. *Assuming χ is non-principal as we did so before,*

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

11.2 Non-Vanishing L-Functions

An L-function is a function defined to be a series involving Dirichlet characters on the complex plane. It takes, as its arguments, a Dirichlet character and complex number, s, such that the real portion $Re(s) > 0$. For our purposes, we will concentrate on the L-function $L(1, \chi)$. We will define this function as well as the function $L'(1, \chi)$. We have already shown the convergence of these two series as the reader will soon see. This will prove useful in our main proof. However, the majority of this section will be dedicated to proving another important fact: that $L(1, \chi)$ is non-zero for non-principal χ. We will do this here for real non-principal χ and then use a trick during our main proof to show the statement is true for complex non-principal χ.

Definition 11.2.1.

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

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

As one can see, we have been working with these series in the previous subsection. Then, we were showing convergence and asymptotic values of partial sums. Now, turning to the non-vanishing property, we need the following theorems:

Theorem 11.2.2. *Let χ be a non-principal character with modulus k. Let $D(n) = \sum_{d|n} \chi(d)$ be the divisor sum of $\chi(n)$. Then $D(n) \geq 0$ for all n, and $D(n) \geq 1$ if n is a square of some integer.*

Proof:

Consider all n such that n is a power of some prime p. Hence, $n = p^a$ for natural number a. Then since primes are indivisible by anything other than 1 and themselves we have that,

$$D(n) = D(p^a) = \sum_{t=0}^{a} \chi(p^t) = 1 + \sum_{t=t}^{a} \chi(p)^t$$

χ is real valued, and it must be a root of unity. Thus, $\chi(p) \in \{1, -1, 0\}$. Thus, $D(n) \in \{1, 1 + a, 0\}$. Well, for all $n \in \mathbb{N}$, $n = 1 \times p_1^{a_1} \times p_2^{a_2} \times p_3^{a_3} \times ... \times p_m^{a_m}$ for some non-negative integer m. Now since D is a divisor sum it

is multiplicative. Thus, $D(n) = D(p_1^{a_1}) \times D(p_2^{a_2}) \times D(p_3^{a_3}) \times ... \times D(p_m^{a_m})$. Thus, $D(n)$ is a product of non-negative integers. Thus, $D(n) > 0$.

Now, we notice $D(p^a)$ is only zero valued when $\chi(p) = -1$ and a is odd. If n is a square of say m. Then, we break down m into its l prime factors, and we see that n is the product of prime squares. Thus, $D(n) = D(p_1^{a_1}) \times D(p_2^{a_2}) \times D(p_3^{a_3}) \times ... \times D(p_m^{a_m})$ where every a_i is even. Thus, $D(n)$ is the product of positive integers, and $D(n) \geq 1$.

Thus, $D(n) \geq 0$ for all n, and $D(n) \geq 1$ if n is a square of some integer. Q.E.D.

Theorem 11.2.3. *Assume everything that was assumed in Theorem 10.2.2., but let χ be non-principal. Let $B(x) = \sum_{n \leq x} \frac{D(n)}{\sqrt{n}}$. Then, $B(x) \to \infty$ as $x \to \infty$ and $B(x) = 2\sqrt{x}L(1, \chi) + O(1)$ for all $x \geq 1$.*

Proof:

From the previous theorem we know that,

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

The series all the way on the right diverges, since it is simply the harmonic series. Thus, we see that $B(x)$ must diverge.

Now we use Theorem 4.6.2. Let $a = \sqrt{x} = b$, $f(n) = \frac{\chi(n)}{\sqrt{n}}$, $g(n) = \frac{1}{\sqrt{n}}$. Then define all other things accordingly. Thus, we have that,

$$\begin{aligned} B(x) &= \sum_{qd \leq x} \frac{\chi(d)}{\sqrt{qd}} \\ &= \sum_{n \leq \sqrt{x}} \frac{\chi(n)}{\sqrt{n}} G\left(\frac{x}{n}\right) + \sum_{n \leq \sqrt{x}} \frac{1}{\sqrt{n}} F\left(\frac{x}{n}\right) - F(\sqrt{x})G(\sqrt{x}) \end{aligned}$$

This next step is a bit hand wavy, but in the interest of avoiding unnecessarily complicated calculation, we will assume to be true that,

$$G(x) = \sum_{n \leq \sqrt{x}} \frac{1}{\sqrt{n}} = 2\sqrt{x} + C_0 + O\left(\frac{1}{\sqrt{x}}\right)$$

where C_0 is a constant. This can be proven using Euler's summation formula that we proved in Theorem 5.2.1. Now by Theorem 11.1.2. we also know that,

$$F(x) = \sum n \leq \sqrt{x}\frac{\chi(n)}{\sqrt{n}} = C_1 + O\left(\frac{1}{\sqrt{x}}\right)$$

where C_1 is a constant, but importantly, it is the constant that $\sum_{n=1}^{\infty} \frac{\chi(n)}{\sqrt{n}}$ converges to.

Thus, we have that $F(x)G(x) = 2C_1 x^{\frac{1}{4}} + O(1)$. We now plug these quantities back into the equation that we derived from Theorem 4.6.2. The result is the following:

$$\begin{aligned}
B(x) &= \sum_{qd \leq x} \frac{\chi(d)}{\sqrt{qd}} \\
&= \sum_{n \leq \sqrt{x}} \frac{\chi(n)}{\sqrt{n}} G\left(\frac{x}{n}\right) + \sum_{n \leq \sqrt{x}} \frac{1}{\sqrt{n}} F\left(\frac{x}{n}\right) - F(\sqrt{x})G(\sqrt{x}) \\
&= \sum_{n \leq \sqrt{x}} \frac{\chi(n)}{\sqrt{n}} 2\sqrt{\frac{x}{n}} + C_0 + O\left(\frac{\sqrt{n}}{\sqrt{x}}\right) \\
&+ \sum_{n \leq \sqrt{x}} \frac{1}{\sqrt{n}} C_1 + O\left(\frac{1}{\sqrt{x}}\right) \\
&- F(\sqrt{x})G(\sqrt{x})
\end{aligned}$$

Corollary 10. $L(1, \chi) \neq 0$

Proof:

Since $B(x) \to \infty$ as $x \to \infty$, we know that $\lim_{x \to \infty} 2\sqrt{x}L(1, \chi)$ must be finitely distanced away from infinity and thus be infinite itself. Thus, $2\sqrt{x}L(1, \chi)$ must diverge. Thus, $L(1, \chi) \neq 0$ as desired.

12 Primes in Arithmetic Progressions

In the next five subsections of this thesis, several lemmas will be presented. Instead of explaining each lemma along the way, it will be easier to prove these results and then connect them with an explanation. The explanation will follow in the Conclusion section. If the reader feels he/she can understand the proof as it is presented, that is great. However, I believe that to be incredibly difficult. Each piece can be connected logically, but the connections may be more subtle than one might anticipate.

12.1 Lemma 1

Lemma 12.1. *For $x > 1$, the following equation holds,*

$$\sum_{\substack{p \leq x \\ p \equiv h \pmod{k}}} \frac{\log p}{p} = \frac{\log x}{\varphi(k)} + \frac{1}{\varphi(k)} \sum_{r=2}^{\varphi(k)} \overline{\chi_r(h)} \sum_{p \leq x} \frac{\chi_r(p) \log p}{p} + O(1)$$

Proof:

The orthogonal relation for Dirichlet characters (Theorem 10.6.1) tell us that,

$$\sum_{r=1}^{n} \chi_r(m) \overline{\chi_r(n)} = \begin{cases} \varphi(k) & \text{if } m \equiv n \pmod{k} \\ 0 & \text{if } m \not\equiv n \pmod{k} \end{cases}$$

Now consider what happens when we let $m = p$ and $n = h$ and sum over all $p \leq x$. We get that,

$$\sum_{p \leq x} \sum_{r=1}^{\varphi(k)} \chi_r(p) \overline{\chi_r(h)} = \varphi(k) \sum_{\substack{p \leq x \\ p \equiv h \pmod{k}}} 1$$

We can of course multiply the terms of both series by $\frac{\log p}{p}$ to obtain,

$$\sum_{p \leq x} \sum_{r=1}^{\varphi(k)} \chi_r(p) \overline{\chi_r(h)} \frac{\log p}{p} = \varphi(k) \sum_{\substack{p \leq x \\ p \equiv h \pmod{k}}} \frac{\log p}{p}$$

This is exactly what we want on the right. Looking at the left-hand series, we isolate the principal character. This is because we will use the properties we know about the principal character to further simplify this equation.

$$\begin{aligned} \varphi(k) \sum_{\substack{p \leq x \\ p \equiv h \pmod{k}}} \frac{\log p}{p} &= \sum_{p \leq x} \sum_{r=1}^{\varphi(k)} \chi_r(p) \overline{\chi_r(h)} \frac{\log p}{p} \\ &= \sum_{p \leq x} \chi_1(p) \overline{\chi_r(h)} \frac{\log p}{p} + \sum_{p \leq x} \sum_{r=2}^{\varphi(k)} \chi_r(p) \overline{\chi_r(h)} \frac{\log p}{p} \end{aligned}$$

Now, we know that $\chi_1(h) = 1$, since $(h,k) = 1$, and the definition specified this value for the principal character. On another note,

$$\chi_1(p) = \begin{cases} 1 & \text{if } (p,k) = 1 \\ 0 & \text{else} \end{cases}$$

Thus, we see that,

$$\begin{aligned} \sum_{p \leq x} \chi_1(p)\overline{\chi_r(h)}\frac{\log p}{p} &= \sum_{\substack{p \leq x \\ (p,k)=1}} \frac{\log p}{p} \\ &= \sum_{p \leq x} \frac{\log p}{p} - \sum_{\substack{p \leq x \\ (p,k)\neq 1}} \frac{\log p}{p} \\ &= \sum_{p \leq x} \frac{\log p}{p} - \sum_{\substack{p \leq x \\ p|k}} \frac{\log p}{p} \end{aligned}$$

Note, the last step is true since p is prime so no p is divisible by k for any integer $k \neq 1$. But this also shows us that the second term of the last line of our equation is a finite series. Thus, it is equal to some constant. Thus, including our knowledge of the asymptotic formula for the sum $\sum_{p \leq x} \frac{\log p}{p}$, which was the focus of the first half of this paper, we see that,

$$\begin{aligned} \sum_{p \leq x} \frac{\log p}{p} - \sum_{\substack{p \leq x \\ p|k}} \frac{\log p}{p} &= \sum_{p \leq x} \frac{\log p}{p} + O(1) \\ &= \log x + O(1) + O(1) = \log x + O(1) \end{aligned}$$

Thus, we plug everything back into where it came from and we are left with

$$\varphi(k) \sum_{\substack{p \leq x \\ p \equiv h \pmod{k}}} \frac{\log p}{p} = \log x + O(1) + \sum_{p \leq x} \sum_{r=2}^{\varphi(k)} \chi_r(p)\overline{\chi_r(h)}\frac{\log p}{p}$$

Rearranging this equation and dividing both sides by $\varphi(k)$, we see that the equation of our lemma holds. Q.E.D.

12.2 Lemma 2

Lemma 12.2. *For $x > 1$ and non-principal χ,*

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

Proof:

Consider the sum, $\sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n}$. Recall that Λ is the Mangoldt function as defined (twice) in the section on arithmetical functions. From our first definition of Λ, we can expand this sum as follows,

$$\sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = \sum_{p \leq x} \sum_{\substack{a \text{ s.t.} \\ p^a \leq x}} \frac{\chi(p^a) \log p}{p^a}$$

Now, we will use a trick similar to our last lemma. We isolate the terms where a is 1. I should note that a must be a positive integer for the formula above to work. Basically, we isolate all terms where the power of the prime that goes into our functions is 1. This goes as follows.

$$\sum_{p \leq x} \sum_{\substack{a \text{ s.t.} \\ p^a \leq x}} \frac{\chi(p^a) \log p}{p^a} = \sum_{p \leq x} \frac{\chi(p) \log p}{p} + \sum_{p \leq x} \sum_{\substack{a \neq 1 \text{ s.t.} \\ p^a \leq x}} \frac{\chi(p^a) \log p}{p^a}$$

Now we will majorize the second sum of the right-hand side of our equation. This means that we will define a sum such that all partial sums of the newly defined series are greater than or equal to their respective partial sum in the pre-existing series.

We see that the series,

$$\sum_{p \leq x} \sum_{\substack{a \neq 1 \text{ s.t.} \\ p^a \leq x}} \frac{\log p}{p^a} \geq \sum_{p \leq x} \sum_{\substack{a \neq 1 \text{ s.t.} \\ p^a \leq x}} \frac{\chi(p^a) \log p}{p^a}$$

since each $\chi(p)$ is a root of unity and the absolute value must be less than or equal to 1. Thus, we have majorized the desired series. Now we point out that since none of the terms of our majorizing series are negative,

$$\sum_{p} \sum_{a \neq 1} \frac{\log p}{p^a} \geq \sum_{p \leq x} \sum_{\substack{a \neq 1 \text{ s.t.} \\ p^a \leq x}} \frac{\log p}{p^a}$$

Well, this infinite sum on the left is a convergent geometric series, but missing the initial two terms. Thus, it is equal to $\frac{p}{p-1} - 1 - \frac{1}{p} = \frac{1}{p(p-1)}$ after some simplification. Thus,

$$\sum_p \sum_{a \neq 1} \frac{\log p}{p^a} = \sum_p \frac{\log p}{p(p-1)} \leq \sum_n \frac{\log n}{n(n-1)}$$

and this is a well-known convergent series. Thus, $\sum_n \frac{\log n}{n(n-1)} = O(1)$. Therefor,

$$\begin{aligned}
\sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} &= \sum_{p \leq x} \sum_{\substack{a \text{ s.t.} \\ p^a \leq x}} \frac{\chi(p^a)\log p}{p^a} \\
&= \sum_{p \leq x} \frac{\chi(p)\log p}{p} + \sum_{p \leq x} \sum_{\substack{a \neq 1 \text{ s.t.} \\ p^a \leq x}} \frac{\chi(p^a)\log p}{p^a} \\
&= \sum_{p \leq x} \frac{\chi(p)\log p}{p} + O(1)
\end{aligned}$$

At this point, we remember our second definition for the Mangoldt function, $\Lambda(n) = \sum_{d|n} \mu(d) \log\left(\frac{n}{d}\right)$. Plugging this into our newfound equation, we have that,

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

Now we re-index this sum, letting $cd = n$. In other words, c and d run through all factors of n. Recall, χ is multiplicative, so $\chi(n) = \chi(c)\chi(d)$. We obtain the following equation,

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

Now, consider the inner sum on the right. Since $\frac{x}{d} \geq 1$, we may utilize equation 3 from Theorem 11.1.2. We see that.

$$\sum_{c \leq \frac{x}{d}} \frac{\chi(c)}{c} = -L'(1,\chi) + O\left(\frac{\log \frac{x}{d}}{\frac{x}{d}}\right)$$

Thus,

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

The big-oh-term can be heavily simplified.

$$O\Big(\sum_{d\le x}\frac{1}{d}\frac{\log\frac{x}{d}}{\frac{x}{d}}\Big) = O\Big(\sum_{d\le x}\frac{\log\frac{x}{d}}{x}\Big)$$
$$= O\Big(\frac{1}{x}\sum_{d\le x}(\log x -)\Big)$$
$$= O\Big(\frac{[x]\log x - \sum_{d\le x}}{x}\Big)$$
$$= O\Big(\frac{[x]\log x - \log[x]!}{x}\Big)$$
$$= O\Big(\frac{[x]\log x - (x\log x + O(x))}{x}\Big)$$
$$= O\Big(\frac{O(x)}{x}\Big) - O(1)$$

Recall that we proved the penultimate line of the above equation in a corollary of Theorem 7.2.1.

We may plug our new big-oh-term back into our equation. The result is that,

$$\sum_{p\le x}\frac{\chi(p)\log p}{p} = -L'(1,\chi)\sum_{d\le x}\frac{\chi(d)\mu(d)}{d} + O(1) \text{ Q.E.D.}$$

12.3 Lemma 3

Lemma 12.3. *For $x > 1$ and non-principal χ,*

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

Proof:

Recall, Theorem 4.4.4., which we have already used once in the previous lemma, says the following: If α is a completely multiplicative arithmetical function, and F is a complex-valued function defined on $(0, +\infty)$ such that $F(x) = 0$ for $0 < x < 1$, then,

$$G(x) = \sum_{n \leq x} \alpha(n) F\left(\frac{x}{n}\right) \iff F(x) = \sum_{n \leq x} \mu(n)\alpha(n) G\left(\frac{x}{n}\right)$$
$$\text{i.e. } G = (\alpha \star F) \iff F = (\mu\alpha \star G)$$

Let $\alpha = \chi$ and let $F(x) = \begin{cases} x & \text{if } x \notin (0,1) \\ 0 & \text{else} \end{cases}$ be a function defined on the positive real line. Applying the quoted theorem, we have that,

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

such that,

$$G(x) = x \sum_{n \leq x} \frac{\chi(n)}{n}$$

It follows from the first equation of Theorem 11.1.2. that $G(x) = xL(1,\chi)\sum_{n \leq x} \frac{\mu(n)\chi(n)}{n} + O(x)$. Substituting this back into our equation for x, we see that,

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

We divide both sides of the equation by x and bring the big-oh-term to the other side to obtain,

$$L(1,\chi)\sum_{n \leq x} \frac{\mu(n)\chi(n)}{n} = O(1) \text{ Q.E.D.}$$

12.4 Lemma 4

Lemma 12.4. *For non-principal χ, if $L(1,\chi) = 0$ then,*

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

Proof:

Again, we turn to Theorem 4.4.4.! Let $\alpha = \chi$ as before, but now let $F(x) = \begin{cases} x & \text{if } x\log x \notin (0,1) \\ 0 & \text{else} \end{cases}$. Thus,

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

such that,

$$\begin{aligned} G(x) &= \sum_{n\leq x}\chi(n)\frac{x}{n}\log\frac{x}{n} \\ &= \sum_{n\leq x}\frac{\chi(n)x}{n}(\log x - \log n) \\ &= x\log x\sum_{n\leq x}\frac{\chi(n)}{n} - x\sum_{n\leq x}\frac{\chi(n)\log n}{n} \end{aligned}$$

Applying the first and second parts of Theorem 11.1.2., we get the following:

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

The last line of the equation is by the fact that $L(1,\chi) = 0$, which was given. We plug in our asymptotic formula for $G(x)$ into our equation for F like so:

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

The final step is to divide the equation by x and bring the big-oh-term (which is now $O(1)$) over to the other side. Then,

$$\log x + O(1) = L'(1,\chi)\sum_{p\leq x}\frac{\mu(n)\chi(n)}{n} \text{ Q.E.D.}$$

12.5 Lemma 5

Lemma 12.5. *Define $N(k)$ to be the counting function that outputs the number of non-principal characters, χ with modulus k such that $L(1,\chi) = 0$. Then, $N(k) = 0$*

Proof:

Consider the result of Lemma 1, and let $h = 1$. We have that:

$$\sum_{\substack{p\leq x\\ p\equiv 1 \pmod k}}\frac{\log p}{p} = \frac{\log x}{\varphi(k)} + \frac{1}{\varphi(k)}\sum_{r=2}^{\varphi(k)}\overline{\chi_r(1)}\sum_{p\leq x}\frac{\chi_r(p)\log p}{p} + O(1)$$
$$= \frac{\log x}{\varphi(k)} + \frac{1}{\varphi(k)}\sum_{r=2}^{\varphi(k)}\sum_{p\leq x}\frac{\chi_r(p)\log p}{p} + O(1)$$

In Lemma 2, we showed that for non-principal χ we have,

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

Thus,

$$\sum_{\substack{p\leq x\\ p\equiv 1 \pmod k}}\frac{\log p}{p} = \frac{\log x}{\varphi(k)} + \frac{1}{\varphi(k)}\sum_{r=2}^{\varphi(k)} -L'(1,\chi_r)\sum_{n\leq x}\frac{\mu(n)\chi_r(n)}{n} + O(1)$$

Now, for a given χ_r, if $L(1, \chi_r) \neq 0$, then Lemma 3 tells us,

$$\begin{aligned} & L(1, \chi_r) \sum_{n \leq x} \frac{\mu(n)\chi_r(n)}{n} = O(1) \\ \implies & \sum_{n \leq x} \frac{\mu(n)\chi_r(n)}{n} = O(1) \\ \implies & -L'(1, \chi_r) \sum_{n \leq x} \frac{\mu(n)\chi_r(n)}{n} = O(1) \end{aligned}$$

This is due to the fact that $L(1, \chi)$ and $L'(1, \chi)$ both converge.

However, if $L(1, \chi) \neq 0$, then Lemma 4 tells us that,

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

So, we see that the sum, $\frac{1}{\varphi(k)} \sum_{r=2}^{\varphi(k)} -L'(1, \chi_r) \sum_{n \leq x} \frac{\mu(n)\chi_r(n)}{n}$ is essentially counting the number of Dirichlet characters χ such that $L(1, \chi) = 0$ and then multiplying them by $-\log x$. Notice if $L(1, \chi) \neq 0$, then $-L'(1, \chi_r) \sum_{n \leq x} \frac{\mu(n)\chi_r(n)}{n}$ is equal to $O(1)$ and gets absorbed by the already existing $O(1)$. Thus, our equation can be re-written in the following way:

$$\begin{aligned} \sum_{\substack{p \leq x \\ p \equiv 1 \pmod{k}}} \frac{\log p}{p} &= \frac{\log x}{\varphi(k)} + \frac{1}{\varphi(k)} \sum_{r=2}^{\varphi(k)} -L'(1, \chi_r) \sum_{n \leq x} \frac{\mu(n)\chi_r(n)}{n} + O(1) \\ &= \frac{\log x}{\varphi(k)} + \frac{1}{\varphi(k)} \sum_{\substack{2 \leq r \leq \varphi(k) \\ L(1, \chi_r) = 0}} -\log x + O(1) \\ &= \frac{\log x}{\varphi(k)} - \frac{N(k) \log x}{\varphi(k)} + O(1) \\ &= \frac{1 - N(k)}{\varphi(k)} \log x + O(1) \end{aligned}$$

Now assume $N(k) > 0$ for a contradiction. Well if $L(1, \chi) = 0$ for a given χ, then $L(1, \overline{\chi}) = 0$ as well. Now we know that no real-valued Dirichlet character has the property that $L(1, \chi) = 0$ because we proved it in Theorem 11.2.3. Thus, $\chi \neq \overline{\chi}$. Thus, the value of $N(k)$ must be even. Thus, $N(k) \geq 2$. However that would mean that the right-hand-side of our

final equation is negative. But every term in the sum on the left-hand-side is positive. Thus, we have a contradiction. Therefor, we conclude that $N(k) = 0$ Q.E.D.

12.6 Conclusion

The final lemma of the set gives a pretty solid idea of the motivation behind the itself and the other five. However, there is still some explaining to do. As mentioned prior to stating any lemma, in this subsection, we will put everything together.

The first lemma sets up the entire proof. Not only do we get incredibly close to our result from just the single result by itself, it beautifully illustrates the usage of Dirichlet characters as devices for isolating a particular congruence class with modulus k. The orthogonal relations of character functions and thus Dirichlet characters are really the major theorem here. This is how we begin to restrict Theorem 8.0.2. which tells us an asymptotic formula for $\sum_{p\leq x} \frac{\log p}{p}$ to the specific primes we desire (within a particular arithmetic progression.) At the end of this lemma, we are left with the equation,

$$\sum_{\substack{p\leq x \\ p\equiv h \pmod k}} \frac{\log p}{p} = \frac{\log x}{\varphi(k)} + \frac{1}{\varphi(k)} \sum_{r=2}^{\varphi(k)} \overline{\chi_r(h)} \sum_{p\leq x} \frac{\chi_r(p)\log p}{p} + O(1)$$

It seems that if we can prove that, $\sum_{p\leq x} \frac{\chi_r(p)\log p}{p} = O(1)$ then our result follows. Clearly, this would imply that $\frac{1}{\varphi(k)} \sum_{r=2}^{\varphi(k)} \overline{\chi_r(h)} \sum_{p\leq x} \frac{\chi_r(p)\log p}{p}$ is just a finite series of bounded constants. Thus, it would be some constant itself, and our theorem follows. While the end seems near, it become more difficult to prove this than one would anticipate.

The second lemma is an intermediary step that attempts to bring us closer to proving that $\sum_{p\leq x} \frac{\chi_r(p)\log p}{p} = O(1)$. It tells us that,

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

Since we have proven that $L'(1,\chi)$ converges, we can conclude from Lemma 2 that,

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

And so, this is our new goal.

Then we get to Lemma 3, which tells us that,

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

Again, from our previous work, we know that $L(1,\chi)$ converges. So what we would like to do is divide both sides by $L(1,\chi)$ and obtain our final result. However, we have to be careful that $L(1,\chi)\neq 0$. We had already proven this for real-valued Dirichlet characters, but most Dirichlet characters are not real. If $L(1,\chi)=0$ for a non-principal χ, then we cannot divide by the zero-valued quantity. The remaining lemmas are an attempt to prove that $L(1,\chi)\neq 0$ for non-principal χ as desired.

Lemma 4 gives us a asymptotic formula for $L'(1,\chi)\sum_{p\leq x}\frac{\mu(n)\chi(n)}{n}$ when $L(1,\chi)=0$. It says that if χ has a vanishing L-function then,

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

This does not seem particularly useful at first. However in conjunction with the next theorem, we see its purpose.

Lemma 5 says that $N(k)$, the number of χ with modulus k, such that $L(1,\chi)=0$ is zero valued. The way that it proves this is by comparing the value of $L'(1,\chi)\sum_{p\leq x}\frac{\mu(n)\chi(n)}{n}$ between Dirichlet character with vanishing L-functions and those without. Ultimately, it yields the contradiction we proved, and we conclude that $N(k)=0$ by necessity.

Finally, we are able to bring everything full circle. Since $N(k)=0$, we conclude that

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

Thus,

$$\sum_{p\leq x}\frac{\chi_r(p)\log p}{p}+O(1)$$

Thus,

$$\sum_{\substack{p\leq x \\ p\equiv h \pmod{k}}} \frac{\log p}{p} = \frac{\log x}{\varphi(k)} + \frac{1}{\varphi(k)} \sum_{r=2}^{\varphi(k)} \overline{\chi_r(h)} \sum_{p\leq x} \frac{\chi_r(p)\log p}{p} + O(1)$$

$$= \frac{\log x}{\varphi(k)} + O(1)$$

Thus, our main result holds. Q.E.D.

As mentioned before this asymptotic formula implies the existence of infinitely many primes. Since the function $\log x$ diverges, the sum, $\sum_{\substack{p\leq x \\ p\equiv h \pmod{k}}} \frac{\log p}{p}$ must diverge, and thus, there must be infinitely many primes in arithmetic progression, $AP(h,k)$, assuming that $(h,k) = 1$ of course.

13 Discussion

As one can see, Shapiro's decision to use $\frac{\log p}{p}$ proved very useful in the end. We were able to use both $L(1,\chi)$ and $L'(1,\chi)$ as tools to get the results that we want. This saved important time and energy, so that we could focus on the motivation behind each lemma. I think that Dirichlet's proof would be much more time consuming. Whatever the case, this proof is also quite elegant in its structure.

One other thing to point out is that the main result has another implication beyond infinite primes in arithmetic progression. It is slightly stronger than a direct proof of that statement. The equation,

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

tells us that each reduced residue class within the series on the left is weighted identically to all the rest. Notice, $\frac{\log x}{\varphi(k)}$ does not depend on h. For this reason, any h will produce the same asymptotic equation on the right. I think this is a really interest place to begin further study.

If you made it to the end of this paper, I want to sincerely thank you. You are probably my thesis advisor or reader, in which case I owe you thanks for your help. If you are not one of those two people, I want to thank you for sticking with the content and making it through to the end.

I know this is a lengthy and tiring result, but it is worthwhile academic endeavor.

13.1 Acknowledgements

I want to thank Gabi Bontea and Fernando Gouvea for the mentor-ship and guidance throughout. Of course, I need to thank Tom M. Apostol for writing my sole resource for this project. I underwent this thesis to try to make his book a little more undergraduate-friendly. Most of the proofs come from his book, but have intermittent steps that I added in between. Maybe somebody will use this as secondary resource for his book, *Introduction to Analytic Number Theory* some day. Who knows?

14 References

[Apo76] Tom M. Apostol. *Introduction to Analytical Number Theory*. Springer-Verlag New York Inc., 1976.