The Bombieri-Vinogradov Theorem

Krishna Praneeth Sidde

July 11, 2022

Abstract

The purpose of this paper is to provide an exposition of the Bombieri Vinogradov Theorem with an intent to explain its importance and make it easier for a non-expert to understand. The paper also discusses a few techniques from Sieve Theory and gives a brief introduction to Analytic Number Theory as it provides the nomenclature and the setting for proving the Bombieri-Vinogradov Theorem. Furthermore, the paper would also briefly discuss about the Goldston-Yildrim-Pintz Theorem in relation with the Bombieri-Vinogradov Theorem.

1 Introduction

In this section, we will most likely focus on the history behind the study of primes, and also on the motivation or intuition behind the Bombieri-Vinogradov Theorem. We would also like to note that most of the content in this paper has likely been adapted from various textbooks and papers with an intent to clarify the difficult topics and make it easier for a non-expert (preferably a beginner in Number Theory) to understand.

The Mystery of Primes.

Right from when the Greeks classified integers into various ways, prime numbers have certainly been one of the most popular and puzzled studies of integers. The fact that they are quite literally the building blocks of integers yet their indiscernible pattern has overwhelmed mathematicians for centuries and the "Mystery of Primes" remains to be unsolved till date. However, a significant amount of progress has been made since the Greeks and a huge effort is being taken by modern-day mathematicians to uncover The Mystery of Primes.

It started with Euclid who proved that there are infinitely many primes in the IX book of *Euclid's Elements*. Then, mathematicians pondered whether there could be a formula that could represent only primes and significant ideas born out of this were the Mersenne's conjecture and Fermat's Conjecture. Mersenne conjectured that $2p - 1$ is a prime where p is also a prime while Fermat conjectured that $2^{2^n} + 1$ is a prime for all integers $n \geq 1$. However, Mersenne's conjecture holds true only for the first 24 primes and it is known to give a composite number for all other values $p \leq 257$. Fermat's conjecture has also later been disproved by Euler who found out that the number is composite for $n = 5$. Since then, no primes have been discovered through Fermat's formula for $n > 4$.

Various mathematicians then considered to answer the question of how many primes there are less than a given positive integer z, and *Carl Friedrich Gauss* (1777-1855) was the first who looked at the proportion of primes. As a teenager, Gauss was obsessed with primes and counted up to the first 3 million primes in hopes of finding patterns. We denote this prime-counting step function as $\pi(x)$ where all the primes up to x are counted. It was later observed by Gauss that the slope of the graph for the prime-counting function looked roughly similar to the graph of $\frac{1}{\log x}$. Hence, Gauss conjectured that the number of primes less than n was roughly equal to $\int \frac{dn}{\log n}$. In today's notation, we denote this prime-counting function as $\pi(x)$ and show that Gauss's conjecture is equal to

$$\pi(x) \sim \int_2^x \frac{dx}{\log x} \sim \frac{x}{\log x}. \tag{1.1}$$

Eventually, Gauss's conjecture was proved by *Jacques Hadamard*(1865-1963) and *Charles Jean de la Vellée Poussin*(1866-1962) in 1896 and it is now famously known as the *Prime Number Theorem.* Another early question that arose eventually concerning the distribution of prime numbers was the distribution of primes within arithmetic progressions. Dirichlet's work was mainly focused on primes within arithmetic progressions and his studies included a slightly modified version of $\pi(x)$ that is denoted as

$$\pi(x; q, a) = \sum_{\substack{p \leq x \\ p \equiv a \bmod q}} 1.$$

$\pi(x; q, a)$ counts the number of primes less than x in a given congruence class modulo q where $(a, q) = 1$. Dirichlet's theorem on primes in arithmetic progressions stated that $\pi(x; q, a) \to \infty$ as $x \to \infty$.

However, we are actually interested in obtaining a numerical estimate for primes in a particular arithmetic progression rather than just showing the infinitude of primes in a particular arithmetic progression. To do so, we would expect all congruence classes to have roughly the "same" number of primes and since there would be $\phi(q)$ congruence classes according to *Euler's Totient Function*(2.2)[1], we would expect that $\pi(x; q, a)$ is roughly equal to $\frac{\pi(x)}{\phi(q)}$ for some sufficiently large x. This version of Dirichlet's theorem for primes in arithmetic progressions is usually known as the Prime Number Theorem for primes in arithmetic progressions and it is denoted as

$$\pi(x; q, a) \sim \frac{\pi(x)}{\phi(x)} \sim \frac{1}{\phi(q)} \frac{x}{\log x} \tag{1.2}$$

While this seems to show everything about the Distribution of primes, the *asymptotic equality*(2.11) is not quite sufficient and we would need to more information about 1.2. Hence, we denote 1.2 using an error term through the *little-oh*(2.10) notation as

$$\pi(x; q, a) \sim \frac{\pi(x)}{\phi(x)} \sim \frac{1}{\phi(q)} \frac{x}{\log x} + o\left(\frac{x}{\log x}\right).$$

[1]We will be discussing about $\phi(n)$ later in the paper but for now, we can note that $\phi(n)$ denotes the number of positive integers ≥ 1 and $\leq n$ that are relatively prime to n.

The goal of many results that we would discuss in this paper to prove the Bombieri-Vinogradov Theorem is to replace this error term with a more precise *big-oh*(2.10) estimate. Before moving on to that, however, we will introduce and discuss several functions and theorems that could be used to explain (or comprehend rather) these concepts in an alternative way. Along the way, we will also introduce concepts from Sieve theory that play a vital role in Analytic Number Theory and particularly in the proof of the Bombieri-Vinogradov Theorem.

2 Elementary functions and theorems from Analytic Number Theory

This section focuses on introducing various concepts from Analytic Number Theory which involve the definitions, nomenclature, and theorems that pertain to the distribution of prime numbers and the Bombieri-Vinogradov Theorem. The material in this section would likely be referenced along various proofs as we advance through the paper and we would adapt the explanation of the concepts from [Apo98]. We would also omit the proofs for most of the theorems in this section since their knowledge is not directly related to the paper and we can continue without the knowledge of the proof. However, an interested reader could find a more robust explanation on these concepts from [Apo98], or preferably any textbook on Analytic Number Theory ([Dav13]; [Mur08]); as far as this section is concerned with. During the end of this section, we will be expanding more about the "error term" we mentioned about in the introduction.

2.1 Arithmetical Functions and their Properties.

Like many other branches of Mathematics, Number theory also deals with various sequences of complex or real numbers. These "number-theoretic" sequences are often called as Arithmetical functions and are denoted as real or complex valued functions defined over positive integers.

$$f := \mathbb{Z}^+ \to \mathbb{R} \vee \mathbb{C}$$

A widely studied question, property rather, of arithmetical functions is whether if it's Multiplicative. A nonzero arithmetical function is associated to be multiplicative if, $f(mn) = f(m)f(n)$ whenever $(m,n) = 1$[2] and is called a completely multiplicative function if we have $f(mn) = f(m)f(n)$ for all m, n. For the rest of this section, we will introduce and study such *multiplicative* arithmetical functions that play an important role in the distribution of primes.

We begin with two important examples, the Möbius function $\mu(n)$ and the Euler totient function $\phi(n)$.

[2] (a,d) represent the *gcd*(Greatest Common Divisor) of a and d

Definition 2.1. The Möbius function $\mu(n)$ is defined as,

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

$where, n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ for $n > 1$.

A very celebrated and fundamental property of the Möbius function in Number Theory for $n \geq 1$, is denoted by a simple formula.

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & n = 1, \\ 0 & n > 1. \end{cases} \tag{2.1}$$

This formula is very obvious for $n = 1$ but for $n > 1$, we consider the prime power factorisation[3] of n and prove the formula using the fact that the Möbius function is multiplicative. Let $F(n) = \sum_{d|n} \mu(d)$. Then,

$$F(n) = F(p_1^{a_1}) F(p_2^{a_2}) \cdots F(p_k^{a_k}).$$

Since,

$$F(p^k) = \sum_{d|p^k} \mu(d) = \mu(1) + \mu(p) + \cdots + \mu(p^k) = 1 + (-1) + 0 + \cdots + 0 = 1 - 1 = 0,$$

We therefore conclude that $F(n) = 0$ due to its multiplicative property. ■

Definition 2.2. The Euler's totient function, $\phi(n)$, is defined to be the number of positive integers not exceeding n, that are coprime to n. It can therefore defined as:

$$\phi(n) = |\mathcal{K}|.$$

$where, \mathcal{K} := \{k : (n,k) = 1\ \forall k \in \mathbb{Z}_{\geq 1} | k \leq n\}$.

The Möbius Function arises in various places where a divisor sum is involved and a fundamental relation of the Möbius function with Euler's totient function is given below. With the help of 2.1, we will state a formula that denotes $\phi(n)$ in a way that is much easier to use in proofs.

Theorem 2.3. *The Euler totient function is related to the Möbius function for $n \geq 1$ through the following formula:*

$$\phi(n) = \sum_{d|n} \mu(d) \frac{n}{d}. \tag{2.2}$$

[3]Factorisation of an integer n of the form $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ where p_i are primes

Due to the popularity of the sums of the kind mentioned above, we denote them in a whole different notation in Analytic Number Theory. For two arithmetical functions, f and g, we say that their *Dirichlet product* is equal to the function h, defined by the equation,

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

Notation. We write $(f * g)$ for h and $(f * g)(n)$ for $h(n)$ to denote the Dirichlet product of two arithmetical functions f and g. We would also take note of the symbol N that denotes the arithmetical function $N(n) = n$ for all n.

In this notation, we therefore denote the formula in Theorem 2.3 as

$$\phi = \mu * N \qquad \phi(n) = (\mu * N)(n)$$

We will now describe a few properties of the Dirichlet Product that are actually quite analogous to Matrix Multiplication, and introduce the Identity function in Dirichlet Products.

Theorem 2.4. *For any arithmetical functions f, g, k,*

$$f * g = g * f \qquad \text{(Commutativity)}$$

$$(f * g) * k = f * (g * k) \qquad \text{(Associativity)}$$

Furthermore, for all f, we have $f * I = I * f = f$, where the I is denoted as the identity function,

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

Theorem 2.5. *For an arithmetical function f, where $f(1) \neq 0$, there is a unique arithmetical function f^{-1}, called the Dirichlet inverse such that,*

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

f^{-1} is given by the recursion formula as,

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

We will now look at a fundamental property of the Möbius function involved with the Dirichlet Product.

With the help of the unit function N we discussed previously and the Identity function I, we define 2.1 as $\sum_{d|n} \mu(d) = I(n)$. In the notation of Dirichlet convolution, we denote this with the help of the unit function N as $\mu * N = I$ Hence, with the help of 2.5, we say that N and μ are Dirichlet inverses of each other.

$$N = \mu^{-1} \qquad \text{and} \qquad \mu = N^{-1}$$

This remarkably simple property of the Möbius function and along with the properties of Dirichlet product, we define the Möbius inversion formula in the next theorem which states that every sum of the form $f(n) = \sum_{d|n} g(d)$ can be denoted using the Möbius function μ.

Theorem 2.6 (The Möbius Inversion Formula)**.** *The equation,*

$$f(n) = \sum_{d|n} g(d)$$

implies

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

Proof. Using the unit function u, we can write $f(n)$ as $g * u$. Multiplying both sides by μ, we note that

$$\begin{aligned} f * \mu &= (g * u) * \mu = g * (u * \mu) \\ &= g * I = g. \end{aligned}$$

■

Next, we introduce the von Mangoldt's function Λ, Chebyshev's ψ, and Chebyshev's θ functions (also called as Chebyshev's first-order and second-order functions). These functions play a very important role in the distribution of prime numbers.

Definition 2.7. The Mangoldt Function, $\Lambda(n)$, for every integer $n \geq 1$ is defined as:

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

For $n \geq 1$, we have $\log n = \sum_{d|n} \Lambda(d)$. Although the Mangoldt's function appears to be "out of the blue", it actually occurs naturally from the fundamental theorem of arithmetic. For $n > 1$, we write

$$n = \prod_{k=1}^{r} p_k^{a_k}.$$

Taking logarithms on both sides, we have,

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

Now consider the sum $\log n = \sum_{d|n} \Lambda(d)$. The only nonzero terms in this sum come from the divisors of d of the form p_k^m for $m = 1, 2, ..., a_k$ and $k = 1, 2, ..., r$. Hence,

$$\sum_{d|n} \Lambda(d) = \sum_{k=1}^{r} \sum_{m=1}^{a_k} \Lambda(p_k^m) = \sum_{k=1}^{r} \sum_{m=1}^{a_k} \log p_k = \sum_{k=1}^{r} a_k \log p_k = \log n,$$

showing that the von Mangoldt's function occurs naturally.

Definition 2.8. Chebyshev's ψ-function, $\psi(x)$, for $x > 0$ is defined as:

$$\psi(x) = \sum_{n \leq x} \Lambda(n).$$

Definition 2.9. Chebyshev's θ-function, $\theta(x)$, for $x > 0$ is defined as:

$$\theta(x) = \sum_{p \leq x} \log p,$$

where$p$ runs over all the primes $\leq x$.

Chebyshev's ψ and θ functions are primarily concerned with the Prime Number Theorem and it states the PNT in a slightly elementary manner. In the later subsections, we will introduce how Chebyshev's first-order and second-order functions can be replaced for $\pi(x)$, and expand more on the relation between Chebyshev's functions and the Prime Number Theorem. For now, we will observe that the Prime Number Theorem is also equivalent to the asymptotic formula,

$$\sum_{n \leq x} \Lambda(n) \sim x \quad as\ x \to \infty.$$

We will now look at the behaviour of these arithmetical functions for large values of n. For example, consider the function $\phi(n)$. This function takes the value $p - 1$ whenever p is a *prime* and also takes on arbitrarily lesser values for various composite numbers. This sort of fluctuation, leaving numerous "gaps" if plotted, definitely makes it very difficult to determine the behaviour of $\phi(n)$ for large values of n. Hence, it is sometimes more beneficial to study the arithmetic mean of an arithmetical function f,

$$\tilde{f}(n) = \frac{1}{n} \sum_{k=1}^{n} f(k).$$

However, before studying the averages of arithmetical functions, we will first need a knowledge of its partial sums $\sum_{k=1}^{n} f(k)$. This is what we would mostly discuss upon throughout this section as we define the average orders of $\Lambda(n)$.

Firstly, we will define an important **notation** that would be used throughout this paper and also while finding the average order of several arithmetical functions.

Definition 2.10 (The big oh and little oh notation)**.** We say that $f(x)$ is big oh of $g(x)$ and denote it as $f(x) = O(g(x))$ when $g(x) > 0$ for all $x \geq a$ to mean that $\frac{f(x)}{g(x)}$ is bounded for $x \geq a$ by a constant $M > 0$ such that,

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

However, we say that $f(x)$ is little oh of $g(x)$ and denote it as $f = o(g(x))$ when there exists some $k > 0$ and $c > 0$ such that,

$$f(x) < cg(x) \qquad \text{for all } x \geq k.$$

Note that if $f(x) = o(g(x))$, then $f(x) = O(g(x))$

Definition 2.11 (Asymptotic equality of functions)**.** If,

$$\lim_{x \to \infty} \frac{f(x)}{h(x)} = 1,$$

the function f(x) is said to be asymptotic to g(x) as $x \to \infty$. It is therefore denoted as,

$$f(x) \sim g(x) \ \text{ as } x \to \infty.$$

To study the asymptotic value of a partial sum, we use the summation formula of Euler to give the exact expression for the error term made in the approximation. In this formula $[t]$ denotes the greatest integer $\leq t$. We will omit the proof of the following theorem since it only serves as a tool for arithmetical functions as far as this paper is concerned. However, we can look at [Apo98, Chapter 3] to know more about the proof.

Theorem 2.12 (Euler's Summation formula)**.** *If f has a continuous derivative over the interval $[y, x]$, where $0 < y < x$, then,*

$$\sum_{y<n\leq x} f(n) = \int_y^x f(t)dt + \int_y^x (t-[t])f'(t)dt + f(x)([x]-x) - f(y)([y]-y). \tag{2.3}$$

Now, we will find the weighted average of $\Lambda(n)$.

Theorem 2.13 (Weighted averages of the $\Lambda(n)$)**.** *For $x \geq 1$, we have*

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

Proof.

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

■

Considering $f(t) = \log t$ and using the Euler's Summation formula, we obtain

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

Hence, we can finally write the weighted average of $\Lambda(n)$ as

$$\Lambda(n) = x\log x - x + O(\log x). \tag{2.4}$$

2.2 Dirichlet Characters

This subsection would focus on probably one of the most important topics in Multiplicative number theory since it is almost impossible to evade the concept of Dirichlet Characters while discussing about number theory. In this subsection, we will introduce Dirichlet characters and their properties, starting from the definition of Finite Abelian Groups. This subsection would also provide the associated mathematical nomenclature with Dirichlet Characters that would be referenced throughout the paper.

Although the study of Dirichlet Characters can be undertaken without any knowledge of groups, we would introduce a few elementary concepts from Group theory in order to simplify and gain a greater understanding of the latter discussions pertaining to the theory of Dirichlet Characters.

Firstly, we recall the definition of a group with a simple example: $\mathbb{Z}$ under group operation of addition. The sum of two integers, is always an integer and it is trivial that the addition of integers is Associative; satisfying the Closure and Associativity property of a Group. For

$\mathbb{Z}$ under addition, 0 acts as an identity element and for every integer n. Furthermore, there is always the negative of n, i.e, $-n$. When added, $n+(-n)=0=e$, acting as the inverse of n and satisfies the existence of inverse postulate. Therefore, we can note that any nonempty set of elements associated with a group operation that combines any two elements of the set to produce a third element of the set which satisfy the Closure, Associativity, Existence of Identity, and Existence of Inverse postulates can be termed as a group $\mathcal{G}$.

One particular type of group that the Dirichlet Characters are concerned with are *Finite Abelian groups.* A group is characterised as an Abelian group when every pair of its elements commute; that is, if $ab = ba$ for all a and b in $\mathcal{G}$ and such a group is termed as a Finite Abelian Group when $\mathcal{G}$ contains a finite number of elements.

Now, we will look into a particular set of functions defined over an arbitrary group that are defined as Characters of a group and later discuss upon the orthogonality relations for characters. Eventually, we will extend this discussion of Groups and Characters onto Dirichlet Characters and prove the orthogonality relations for Dirichlet Characters analogously.

Definition 2.14 (Characters of a group)**.** A complex-valued function f defined on an arbitrary group $\mathcal{G}$ is called a character of $\mathcal{G}$ if f has the multiplicative property $f(ab) = f(a)f(b)$ for all a, b in $\mathcal{G}$, and if $f(c) \neq 0$ for some c in $\mathcal{G}$.

We will now discuss about the Orthogonal relations for characters. However, we will omit the proof of this since the knowledge of the proof is not really essential in this paper.

Lemma 2.15 (Orthogonality relations for characters)**.** *For all the characters $f_1, f_2, \cdots, f_n$ of a Finite Abelian Group $\mathcal{G}$ with the elements $a_1, a_2, \cdots, a_n$, we have*

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

For defining the Dirichlet Characters, we will focus on a particular group of all the *reduced residue classes*[4] modulo k. We will then extend the discussion of the orthogonality relations for characters to Dirichlet Characters.

Definition 2.16 (Dirichlet Characters)**.** Let $\mathcal{G}$ be a group of reduced residue classes modulo k. Corresponding to each character f of $\mathcal{G}$, we define an arithmetical function $\chi = \chi_f$ such that

$$\chi(n) = f(\hat{n}) \quad if\ (n, k) = 1,$$

where $\hat{n} := \{x : x \equiv a \mod k\}$.

$$\chi(n) = 0 \qquad if\ (n, k) > 1.$$

The function χ is called a *Dirichlet Character* modulo k. The *principal character* χ_1 is that which has the properties,

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

[4]Recall that a reduced residue system modulo k is a set of $\phi(k)$ integers $\{a_1, a_2, ..., a_{\phi(k)}\}$ incongruent modulo k, each of which is relatively prime to k.

A very important property of a Dirichlet Character modulo k that we will state below is that they are *periodic modulo* k and that they are also *completely multiplicative*.

Using this result, we can also note that any completely multiplicative function defined on $\mathbb{N}$ that is periodic for a fundamental period q and does not vanish everywhere can actually be represented as a Dirichlet Character modulo q.

Theorem 2.17 (Orthogonality relations for Dirichlet Characters). *Let* $\chi_1, ..., \chi_{\phi(k)}$ *denote the* $\phi(k)$ *Dirichlet characters modulo* k*. Let* m *and* n *be two integers, with* $(n, k) = 1$*. Then we have*

$$\sum_{r=1}^{\phi(k)} \chi_r(m)\bar{\chi}_r(n) = \begin{cases} \phi(k) & if\ m \equiv n \pmod{k}, \\ 0 & if\ m \not\equiv n \pmod{k}. \end{cases}$$

Now, we further define various properties of Dirichlet Characters that generalize them and introduce the notation that would be used throughout this paper. These properties of Dirichlet Characters generalize Character Sums and introduce the conditions for the Separability of Gaussian Sums that would be discussed in the next subsection.

Definition 2.18 (Induced Modulus)**.** For some Dirichlet character, χ mod k, let d be any positive divisor of k. The number d is called an induced modulus for χ if we have,

$$\chi(a) = 1 \quad \text{whenever } (a, k) = 1 \text{ and } a \equiv 1 \pmod{d}.$$

Definition 2.19 (Primitive Characters)**.** A Dirichlet character χ mod k is said to be primitive mod k if it has no induced modulus $d < k$. In other words, χ is primitive mod k if, and only if, for every divisor d of k, $0 < d < k$, there exists an integer $a \equiv 1 \pmod{d}, (a, k) = 1$, such that $\chi(a) \not\equiv 1$.

Definition 2.20 (Conductor of a character)**.** The smallest induced modulus d for a Dirichlet character χ mod k is called the conductor of χ.

2.3 Gaussian Sums and the Pólya-Vinogradov Inequality

In this subsection, we will first take a look at how periodic arithmetical functions can be interpreted as a linear combination of Finite Fourier Series, and then extend the same discussion to Dirichlet Characters. We wish to represent the Finite Fourier Series of Dirichlet Characters using Gaussian Sums and discuss their properties. The end result of this section would be the Pólya-Vinogradov Inequality.

Before introducing the Finite Fourier Series for periodic arithmetical functions, we will go through some notation. We will say that an arithmetical function f is *periodic modulo* k if $f(n + k) = f(n)$. We note that any function similar to f can be expressed in the way $\sum_m c(m)e^{2\pi imn/k}$ where the summation is also periodic mod k for every choice of coefficients $c(m)$.

Theorem 2.21 (Finite Fourier Expansion of periodic arithmetical functions)**.** *For a periodic arithmetical function* f mod k*, there is a uniquely determined arithmetical function* g *periodic* mod k*, such that*

$$f(m) = \sum_{n=0}^{k-1} g(n)e^{2\pi imn/k}.$$

In fact, g is given by the formula

$$g(n) = \frac{1}{k} \sum_{m=0}^{k-1} f(m) e^{-2\pi imn/k}.$$

Note that the *Ramanujan Sum* is defined as the sum over the nth powers of the primitive kth roots of unity for a fixed positive integer n. It is therefore denoted as

$$c_k(n) = \sum_{\substack{m \bmod k \\ (m,k)=1}} e^{2\pi imn/k} \tag{2.5}$$

We will now shift our discussion for the rest of this subsection onto Gaussian Sums and their properties.

Definition 2.22 (Gaussian Sums)**.** For any Dirichlet Character χ mod k, the Gaussian Sum associated with χ is denoted as,

$$G(n, \chi) = \sum_{m=1}^{k} \chi(m) e^{2\pi imn/k}$$

Definition 2.23 (Separability of Gaussian Sums)**.** The Gaussian Sum $G(n, \chi)$ is said to be separable if,

$$G(n, \chi) = \bar{\chi}(n) G(1, \chi).$$

Lemma 2.24. *For any Dirichlet Character* χ mod k*,*

$$G(n, \chi) = \bar{\chi}(n) G(1, \chi) \qquad \text{whenever } (n, k) = 1.$$

Theorem 2.25 (Separability of Gaussian Sums of primitive Dirichlet characters)**.** *For a primitive Dirichlet character* χ *mod* k*, the following conditions hold true:*

(a) $G(n, \chi) = 0$ *for every* n *with* $(n, k) > 1$.

(b) $G(n, \chi)$ *is separable for every* n.

(c) $|G(1, \chi)|^2 = k$.

Using the knowledge of Gaussian Sums and the Finite Fourier series of periodic arithmetical functions, we can now work on the Finite Fourier Expansion of a primitive Dirichlet Character.

Theorem 2.26 (Finite Fourier expansion of a primitive Dirichlet Character)**.** *The finite Fourier expansion of a primitive Dirichlet Character* χ *mod* k *has the form,*

$$\chi(m) = \frac{\tau_k(\chi)}{\sqrt{k}} \sum_{n=1}^{k} \bar{\chi}(n) e^{-2\pi imn/k}$$

where,

$$\tau_k(\chi) = \frac{G(1, \chi)}{\sqrt{k}} = \frac{1}{\sqrt{k}} \sum_{m=1}^{k} \chi(m) e^{2\pi im/k}.$$

One application of the Gaussian Sums is the following fundamental inequality about character sums. The Pólya-Vinogradov Inequality was established by George Pólya and I.M. Vinogradov (not to be confused with A.I. Vinogradov who is concerned with the Bombieri-Vinogradov Theorem) in 1918 which improves the inequality

$$\left|\sum_{m\leq x}\chi(m)\right| \leq \phi(k)$$

for any Dirichlet Character χ mod k to a great margin. The Pólya-Vinogradov Theorem will also play a major rule as we progress through the next couple of sections.

Theorem 2.27 (Pólya-Vinogradov Inequality)**.** *For any primitive Dirichlet Character χ mod k,*

$$\left|\sum_{m\leq x}\chi(m)\right| < \sqrt{k}\log k$$

for all $x \geq 1$.

Proof. We will first express $\chi(m)$ by its finite Fourier expansion as

$$\chi(m) = \frac{\tau_k(\chi)}{\sqrt{k}}\sum_{n=1}^{k}\bar{\chi}(n)e^{-2\pi imn/k},$$

and sum over all $m \leq x$ to obtain

$$\sum_{m\leq x}\chi(m) = \frac{\tau_k(\chi)}{\sqrt{k}}\sum_{n=1}^{k-1}\bar{\chi}(n)\sum_{m\leq x}e^{-2\pi imn/k}.$$

Taking the absolute value on both sides of the above equation and multiplying it by $\sqrt{k}$, we obtain

$$\sqrt{k}\left|\sum_{m\leq x}\chi(m)\right| \leq \sum_{n=1}^{k-1}\left|\sum_{m\leq x}e^{-2\pi imn/k}\right| = \sum_{n=1}^{k-1}|f(n)| \tag{2.6}$$

where we denote $f(n) = \sum_{m\leq x}e^{-2\pi imn/k}$. Considering that $|f(k-n)| = |f(n)|$, we can write 2.6 as

$$\sqrt{k}\left|\sum_{m\leq x}\chi(m)\right| \leq \sum_{n\leq\frac{k}{2}}|f(n)|.$$

From the aforementioned definition of $f(n)$, we represent $f(n) = \sum_{m=1}^{r}y^m$ and consider $f(n)$ as a geometric sum of the form $f(n) = \sum_{m=1}^{r}y^m$ where $r = [x]$ and $y = e^{-2\pi in/k}$. We further represent $y = z^2$ where $z = e^{-\pi in/k}$ to obtain the following result using the sum of geometric series formula.

$$f(n) = y\frac{y^r - 1}{y - 1} = z^{r+1}\frac{z^r - z^{-r}}{z - z^{-1}}$$

Taking absolute values,

$$|f(n)| = \left|\frac{z^r - z^{-r}}{z - z^{-1}}\right| = \left|\frac{e^{-\pi r n/k} - e^{\pi r n/k}}{e^{-\pi i n/k} - e^{\pi i n/k}}\right| = \frac{\left|\sin \frac{\pi r n}{k}\right|}{\left|\sin \frac{\pi n}{k}\right|} \leq \frac{1}{\sin \frac{\pi n}{k}}.$$

Using the fact that $\sin t \geq 2t/\pi$ which is valid for $0 \leq t \leq \pi/2$,

$$|f(n)| \leq \frac{k}{2n}.$$

Using this result in 2.6, we get the following inequality,

$$\sqrt{k}\left|\sum_{m\leq x} \chi(m)\right| \leq k \sum_{n \leq \frac{k}{2}} \frac{1}{n} \leq k \log k.$$

Hence, we get

$$\left|\sum_{m\leq x} \chi(m)\right| < \sqrt{k} \log k,$$

and we complete the proof of the Pólya-Vinogradov Theorem. ■

2.4 Primes in Arithmetic Progressions

As the name suggests, this subsection would mainly focus on the distribution of prime numbers in Arithmetic Progressions and we will state the Siegel-Walfiz Theorem by the end of this section. The mathematical foundation for most of the part discussed during the Introduction on the History of Prime Numbers would be worked upon in this subsection and the properties of Chebyshev's first-order and second-order functions that we would discuss upon will play a major role in the proofs of The Large Sieve Inequality and the Barban-Davenport-Halberstam Theorem as we advance through the paper.

We first state a very important theorem that lets us replace $\psi(x)$ with $\theta(x)$ which in turn can be used to replace the prime-counting functions, $\pi(x)$ and $\pi(x; q, a)$.

Theorem 2.28. *For $x > 0$, we have*

$$\lim_{x\to\infty} \left(\frac{\psi(x)}{x} - \frac{\theta(x)}{x}\right) = 0.$$

Now, we will extend our definitions 2.8 and 2.9 of Chebyshev's functions to primes in arithmetic progressions and we note that the properties of Chebyshev's functions which hold for primes also hold for Chebyshev's functions associated with primes in arithmetic progressions.

Definition 2.29. Chebyshev's ψ and θ functions extended over primes in arithmetic progressions

$$\psi(x; q, a) = \sum_{\substack{n \leq x \\ n \equiv a \bmod q}} \Lambda(n),$$

$$\theta(x; q, a) = \sum_{\substack{p \leq x \\ p \equiv a \bmod q}} \log n.$$

The reason why we can replace $\pi(x)$ with $\theta(x)$ is because they both essentially mean the same thing since *pi* takes a step of 1 while θ takes a step of $\log p$. These properties also hold true for $\pi(x;q,a)$, $\theta(x;q,a)$, and $\psi(x;q,a)$.

We will now show an important relation between the ψ function, a Dirichlet Character χ modulo q, and the Λ function.

Definition 2.30. For a Dirichlet Character χ modulo q, we define,

$$\psi(x,\chi) = \sum_{n \le x} \Lambda(n)\chi(n)$$

We will now look at an expansion of $\psi(x;d,a)$

Theorem 2.31. *For any $d \le z$ and $1 \le a \le d$ with $(a,d) = 1$, we observe that*

$$\psi(x;d,a) = \frac{1}{\phi(d)} \sum_{\chi \ mod \ d} \bar{\chi}(a)\psi(x,\chi), \tag{2.7}$$

where $\sum_{\chi \ mod \ d}$ is over Dirichlet Characters χ modulo d.

An important estimate for the error term of $\pi(x;q,a)$ is given by the following theorem and we would be referencing it as we go through the proof of the Bombieri-Vinogradov Theorem.

Theorem 2.32 (Siegel-Walfiz Theorem)**.** *Let A be a real fixed constant and $(a,q) = 1$, where $q \le (\log x)^A$, then there exists a constant $C(A)$ depending only on A such that,*

$$\psi(x;q,a) = \frac{x}{\psi(q)} + O(xe^{-C(A)\sqrt{\log x}}). \tag{2.8}$$

We will omit the proof of the above theorem since it is quite deep but all we can generalize the the result from the above theorem to give a decent estimate for the error term in the PNT for primes in arithmetic progressions. The theorem mentioned below shows a few other variations of the Siegel-Walfiz Theorem that would be used in the latter sections of this paper.

Theorem 2.33. *For a Dirichlet Character χ mod q whose principal character mod q is denoted as χ_0 and is induced by a primitive character χ_1 mod r,where $1 < r < q$, the following two relations hold:*

$$\begin{aligned} &\psi(x,\chi_0) = x + xe^{-c\sqrt{\log x}}. \\ &|\psi(x,\chi) - \psi(x,\chi_1)| \le (\log qx)^2. \end{aligned} \tag{2.9}$$

for some constant, $c > 0$.

3 Results from Sieve Theory

Results from The Large Sieve are of key importance in proving the Bombieri-Vinogradov Theorem. In this section, we will first give a brief introduction to Sieve Theory and eventually work through The Large Sieve Inequality; from which we will deduce The Large Sieve. By the end of this section, we would establish the Barban-Davenport-Halberstam Theorem, which is a key result of the Large Sieve Inequality and an important portion of the Bombieri-Vinogradov Theorem. Various theorems and their proofs presented in this section have been adapted from [CM$^+$06] and [Ten15]. An interested reader in Sieve Theory can refer to [Ten15] or [CM$^+$06] for a more detailed introduction to Sieve Theory and their applications.

3.1 Not your typical household sieving

In simple words, the fundamental goal of Sieve Theory is to estimate the size of a particular set of sifted elements. A major role of sieves such as the sieve of Eratosthenes, Selberg's Sieve, and the Large Sieve in analytic number theory is to estimate the size of the distribution of prime numbers. In this subsection, we will first discuss the fundamental Sieve problem and give a brief idea about Sieves. On the latter part, we will introduce a few identities that would be used in proving The Large Sieve Inequality and the Barban-Davenport-Halberstam Theorem eventually.

Definition 3.1 (The Sieve Problem)**.** For a finite set of objects $\mathcal{A}$, let $\mathcal{P}$ be an index set of primes such that to each $p \in \mathcal{P}$ we have associated a subset $\mathcal{A}_p$ of $\mathcal{A}$. The *Sieve Problem* is to estimate, the size of the set

$$\mathcal{S}(\mathcal{A}, \mathcal{P}) := \mathcal{A} \setminus \cup_{p \in \mathcal{P}} \mathcal{A}_p. \tag{3.1}$$

For example, if we consider a finite set of positive integers $\leq x$, $\mathcal{A}$, and let $\mathcal{A}_p$ be a subset of $\mathcal{A}$ which contains the numbers in $\mathcal{A}$ divisible by $p \in \mathcal{P}$ where $\mathcal{P}$ is the set of all primes, $\mathcal{S}(\mathcal{A}, \mathcal{P})$ gives all the primes $\leq x$. The Sieve Problem would be to estimate this size of set and there are numerous types of techniques followed in Sieve Theory for this type of estimations.

By the inclusion-exclusion principle , we get a more explicit formula of 3.1 as

$$\#\mathcal{S}(\mathcal{A}, \mathcal{P}) = \sum_{I \subseteq A} (-1)^{\#I} \#\mathcal{A}_I.$$

The example of a Sieve Problem we looked at above is in fact famously known as the Sieve of Eratosthenes and is considered to be the foundation of Sieve Theory. We describe below a variation of the Sieve of Eratosthenes developed by A.M. Legendre.

$$\Phi(x, z) := \#\{n \leq x : n \text{ is not divisible by any prime } < z\}$$

A more formal definition of $\Phi(x, z)$ that was described with the help of the *Möbius Function* by A.M. Legendre is stated below.

$$\begin{aligned}\Phi(x, z) &= \sum_{n \leq x} \sum_{d | (n, P_z)} \mu(d) = \sum_{d | P_z} \mu(d) \left[\frac{x}{d} \right] \\ &= x \sum_{d | P_z} \frac{\mu(d)}{d} + O\left(2^z\right) = x \prod_{p < z} \left(1 - \frac{1}{p}\right) + O\left(2^z\right)\end{aligned}$$

where,

$$P_z := \prod_{p<z} p.$$

Now that we have a decent amount of understanding of Sieve Theory and Sieve Principles, we move onto discussing a few fundamental theorems that would be referenced throughout the next couple of subsections as we start discussing about the Large Sieve. We will also omit the proofs of the following theorems since the rest of the paper will only be discussing their applications which can be done without the knowledge of the proof. However, an interested reader can refer to [CM+06].

Theorem 3.2 (Cauchy-Schwarz Inequality)**.** *For sequences of complex numbers a_i and b_i where $1 \leq i \leq n$, the following inequality holds true,*

$$\left| \sum_{1\leq i\leq n} a_i b_i \right|^2 \leq \left(\sum_{1\leq i\leq n} |a_i|^2 \right) \left(\sum_{1\leq i\leq n} |b_i|^2 \right).$$

Theorem 3.3 (Parseval's Identity)**.** *For a sequence of complex numbers $a_{n\geq 1}$ and a positive integer x, we have the following Fourier Series,*

$$\sum_{n\leq x} |a_n|^2 = \int_0^1 \left| \sum_{n\leq x} a_n e^{2\pi i n\alpha} \right|^2 d\alpha.$$

3.2 The Large Sieve

In this subsection, we will discuss about The Large Sieve, which was introduced by Yuri Linnik (1915-72) in 1941 and subsequently improved by many other mathematicians including Enrico Bombieri during 1965. We will deduce this sieve from the The Large Inequality whose significant application is the Bombieri-Vinogradov Theorem.

Before we start with the Large Sieve Inequality, we will begin with the following Lemma that would help us prove the Large Sieve Inequality.

Lemma 3.4. *Let $F : [0,1] \to \mathbb{C}$ be a differentiable function with continuous derivative, extended by periodicity to all $\mathbb{R}$ with period 1. Let z be a positive integer. Then*

$$\sum_{d\leq z} \sum_{\substack{1\leq a\leq d \\ (a,d)=1}} \left| F\left(\frac{a}{d}\right) \right| \leq z^2 \int_0^1 |F(\alpha)|\, d\alpha + \int_0^1 |F'(\alpha)|\, d\alpha.$$

Proof. By looking at the inequality, we can use the fundamental theorem of calculus to present $F(\frac{a}{d})$ using integrals in the form of an inequality. Since

$$\begin{aligned} -F\left(\frac{a}{d}\right) &= \left(\int_{\frac{a}{d}}^{\alpha} F'(t)dt \right) - F(\alpha), \\ \left| F\left(\frac{a}{d}\right) \right| &\leq |F(\alpha)| + \int_{\frac{a}{d}}^{\alpha} |F'(t)|dt. \end{aligned} \tag{3.2}$$

where, $d \le z, a \in [1,d] \cap \mathbb{N}$ with $(a,d) = 1$, and $\alpha \in [0,1]$.

Let

$$I\left(\frac{a}{d}\right) := \left(\frac{a}{d} - \delta, \frac{a}{d} + \delta\right)$$

for some $\delta > 0$ so that the intervals $I\left(\frac{a}{d}\right)$ are contained in $[0,1]$. We now integrate 3.2 over $I(a/d)$ with respect to α, and obtain

$$2\delta \left|F\left(\frac{a}{d}\right)\right| \le \int_{I\left(\frac{a}{d}\right)} |F(\alpha)| d\alpha + \int_{I\left(\frac{a}{d}\right)} \int_{\frac{a}{d}}^{\alpha} |F'(t)| \, \mathrm{d}t \, \mathrm{d}\alpha. \tag{3.3}$$

Since $\alpha \in I\left(\frac{a}{d}\right)$ and $t \in \left[\frac{a}{d}, \alpha\right]$, we can imply that $t \in I\left(\frac{a}{d}\right)$. We can now simplify the right hand side of the above inequality as,

$$\begin{aligned} &\le \int_{I\left(\frac{a}{d}\right)} |F(\alpha)| \mathrm{d}\alpha + \int_{I\left(\frac{a}{d}\right)} \int_{I\left(\frac{a}{d}\right)} |F'(t)| \, \mathrm{d}t \, \mathrm{d}\alpha \\ &= \int_{I\left(\frac{a}{d}\right)} |F(\alpha)| \mathrm{d}\alpha + 2\delta \int_{I\left(\frac{a}{d}\right)} |F'(\alpha)| \, \mathrm{d}\alpha \end{aligned}$$

Hence, we can simplify 3.3 as

$$2\delta \left|F\left(\frac{a}{d}\right)\right| \le \int_{I\left(\frac{a}{d}\right)} |F(\alpha)| d\alpha + 2\delta \int_{I\left(\frac{a}{d}\right)} |F'(\alpha)| \, \mathrm{d}\alpha. \tag{3.4}$$

Now, we consider $\delta = \frac{1}{z^2}$ and sum 3.4 over all intervals $I(a/d)$ to get,

$$\begin{aligned} \frac{1}{z^2} \sum_{\substack{d \le z \\ (\le a,d)=1}} \sum_{\substack{1 \le a \le d \\ (a,d)}} \left|F\left(\frac{a}{d}\right)\right| &\le \sum_{I\left(\frac{a}{d}\right)} \int_{I\left(\frac{a}{d}\right)} |F(\alpha)| \mathrm{d}\alpha + \frac{1}{z^2} \sum_{I\left(\frac{a}{d}\right)} \int_{I\left(\frac{a}{d}\right)} |F'(\alpha)| \, \mathrm{d}\alpha \\ \frac{1}{z^2} \sum_{\substack{d \le z \\ (\le a,d)=1}} \sum_{\substack{1 \le a \le d \\ (a,d)}} \left|F\left(\frac{a}{d}\right)\right| &\le \int_0^1 |F(\alpha)| \mathrm{d}\alpha + \frac{1}{z^2} \int_0^1 |F'(\alpha)| \, \mathrm{d}\alpha. \end{aligned}$$

This completes the proof of the lemma. ■

Theorem 3.5 (The Large Sieve Inequality)**.** *Let $(a_n)_{n \ge 1}$ be a sequence of complex numbers and let x, z be positive integers. Then*

$$\sum_{d \le z} \sum_{\substack{1 \le a \le d \\ (a,d)=1}} \left| \sum_{n \le x} a_n e^{2\pi i n a/d} \right|^2 \le \left(z^2 + 4\pi x\right) \sum_{n \le x} |a_n|^2.$$

Proof. Now, to prove the Large Sieve Inequality, we consider

$$F(\alpha) = S(\alpha)^2, \quad F'(\alpha) = 2S(\alpha)S'(\alpha)$$

where,

$$S(\alpha) := \sum_{n \le x} a_n e^{2\pi i n \alpha}.$$

By 3.4, we get the inequality,

$$\sum_{d\le z}\sum_{\substack{1\le a\le d\\(a,d)=1}}\left|S\left(\frac{a}{d}\right)\right|^2 \le z^2\int_0^1 |S(\alpha)|^2\,d\alpha + 2\int_0^1 |S(\alpha)S'(\alpha)|\,d\alpha.$$

Using Parseval's Identity (3.3), and Cauchy-Schwarz Inequality (3.2), we simplify the right-hand side of the above inequality and obtain

$$\begin{aligned}
\sum_{d\le z}\sum_{\substack{1\le a\le d\\(a,d)=1}}\left|S\left(\frac{a}{d}\right)\right|^2 &\le z^2\int_0^1 |S(\alpha)|^2\,d\alpha + 2\int_0^1 |S(\alpha)S'(\alpha)|\,d\alpha\\
&\le z^2\sum_{n\le x}|a_n|^2 + 2\left(\int_0^1 |S(\alpha)|^2\,d\alpha\right)^{1/2}\left(\int_0^1 |S'(\alpha)|^2\,d\alpha\right)^{1/2}\\
&\le z^2\sum_{n\le x}|a_n|^2 + 2\left(\sum_{n\le x}|a_n|^2\right)^{1/2}\left(\int_0^1\sum_{n\le x}4\pi^2n^2|a_n|^2\left(e^{2\pi i n\alpha}\right)^2 d\alpha\right)^{1/2}\\
&\le z^2\sum_{n\le x}|a_n|^2 + 2\left(\sum_{n\le x}|a_n|^2\right)^{1/2}\left(4\pi^2x^2\left(\int_0^1 |S(\alpha)|^2\,d\alpha\right)\right)^{1/2}\\
&\le z^2\sum_{n\le x}|a_n|^2 + 4\pi x\left(\sum_{n\le x}|a_n|^2\right)^{1/2}\left(\sum_{n\le x}|a_n|^2\right)^{1/2}
\end{aligned}$$

Simplifying the value on the left hand side of the inequality using the value of $S(\alpha)$ gives us

$$\sum_{d\le z}\sum_{\substack{1\le a\le d\\(a,d)=1}}\left|\sum_{n\le x}a_n e^{2\pi i n a/d}\right|^2 \le \left(z^2+4\pi x\right)\sum_{n\le x}|a_n|^2.$$

This completes the proof of the theorem and we establish The Large Sieve Inequality. ■

From this Inequality, we will now deduce the Large Sieve Method where we consider $\mathcal{A}$ to be a set of positive integers $n \le x$ and $\mathcal{P}$ to be a set of primes. For some positive integers $\le z$, The Sieve Problem of the Large Sieve Method is to estimate the set of positive integers $n \in \mathcal{A}$ such that n is incongruent to the set $\{w_{1,p}, w_{2,p}, \cdots, w_{\omega(p),p}\}$, $\omega(p)$ residue classes modulo p where $p \in \mathcal{P}$ and $p \le z$.

$$\mathcal{S}(\mathcal{A},\mathcal{P},z) := \{n\in\mathcal{A} : n \not\equiv w_{i,p}(\text{mod} p)\forall 1\le i\le\omega(p), \forall p<z\}$$

Hence, the sieve problem is to give an estimate of $|\mathcal{S}(\mathcal{A},\mathcal{P},z)|$. In simple words, we can note that the Large Sieve Problem is to provide an estimate for the positive integers n which do not occur in the form of $pn+(p-1)$ where $p\in\mathcal{P}$.

We will omit the proof of the following inequality which gives an estimate on The Large Sieve Problem since we will not really require it while proving the Bombieri-Vinogradov Theorem as it is mainly concerned with the Large Sieve Inequality and the Barban-Davenport-Halberstam Theorem. An interested reader could always refer to [CM+06, Chapter 8.2] for the proof of the following inequality.

$$|S(\mathcal{A},\mathcal{P},z)| \leq \frac{z^2+4\pi x}{\sum_{d\leq z}\mu^2(d)\prod_{p|d}\frac{\omega(p)}{p-\omega(p)}},$$

We will now introduce the modified versions of The Large Sieve Inequality that are given through Character Sums.

Theorem 3.6 (First Modified Large Sieve Inequality)**.** *For a sequence of complex numbers $a_{n\geq 1}$ and positive integers x, z, we have*

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi^*\, mod\, d}\left|\sum_{n\leq x}a_n\chi(n)\right|^2 \leq \left(z^2+4\pi x\right)\sum_{n\leq x}|a_n|^2,$$

where, $\sum_{\chi^*\, mod\, d}$ is the sum over primitive characters χ modulo k.

Proof. Using the representation of Finite Fourier Series for primitive Dirichlet Characters from 2.26, we multiply the expansion by a_n, sum it over $n\leq x$, and square it out to represent the left-hand side of the desired inequality.

$$\left|\sum_{n\leq x}a_n\chi(n)\right|^2 = \frac{1}{d}\left|\sum_{1\leq a\leq d}\bar{\chi}(a)\sum_{n\leq x}a_ne^{2\pi ina/d}\right|^2,$$

where, we consider χ to be a primitive character modulo d for some fixed $d\leq z$ and consider n be relatively prime to d.

Now, we sum the above equation over all primitive characters χ^* modulo d.

$$\begin{aligned}
\sum_{\chi^*\text{ mod }d}\left|\sum_{n\leq x}a_n\chi(n)\right|^2 &= \frac{1}{d}\sum_{\chi^*\text{ mod }d}\left|\sum_{1\leq a\leq d}\bar{\chi}(a)\sum_{n\leq x}a_ne^{2\pi ina/d}\right|^2\\
&\leq \frac{1}{d}\sum_{\chi\text{ mod }d}\left|\sum_{1\leq a\leq d}\bar{\chi}(a)\sum_{n\leq x}a_ne^{2\pi ina/d}\right|^2\\
&\leq \frac{1}{d}\sum_{\chi\text{ mod }d}\sum_{1\leq a\leq d}\bar{\chi}(a)\sum_{n\leq x}a_ne^{2\pi ina/d}\sum_{1\leq b\leq d}\chi(b)\overline{\left(\sum_{n\leq x}a_ne^{2\pi inb/d}\right)}\\
&\leq \frac{1}{d}\sum_{1\leq a\leq d}\sum_{1\leq b\leq d}\left(\sum_{n\leq x}a_ne^{2\pi ina/d}\right)\overline{\left(\sum_{n\leq x}a_ne^{2\pi inb/d}\right)}\sum_{\chi\text{ mod }d}\bar{\chi}(a)\chi(b),
\end{aligned}$$

where $\bar{z}$ is the complex conjugate of z.

Now, using Theorem 2.17, we note that $\sum_{\chi \bmod d} \bar{\chi}(a)\chi(b) = \phi(d)$ since $a \equiv b \mod d$. We also reduce the right-hand side of the aforementioned inequality into a single Fourier expansion using the properties of Complex numbers and we get

$$\sum_{\chi^* \bmod d} \left| \sum_{n \le x} a_n \chi(n) \right|^2 \le \frac{\phi(d)}{d} \sum_{\substack{1 \le a \le d \\ (a,d)=1}} \left| \sum_{n \le x} a_n e^{2\pi i n a/d} \right|^2 .$$

Therefore, we get

$$\frac{\phi(d)}{d} \sum_{\chi^* \bmod d} \left| \sum_{n \le x} a_n \chi(n) \right|^2 \le \sum_{\substack{1 \le a \le d \\ (a,d)=1}} \left| \sum_{n \le x} a_n e^{2\pi i n a/d} \right|^2 .$$

By summing this result over $\sum_{d \le z}$ and applying the Large Sieve Inequality, Theorem 3.6, we can get the final result,

$$\sum_{d \le z} \frac{\phi(d)}{d} \sum_{\chi^* \bmod d} \left| \sum_{n \le x} a_n \chi(n) \right|^2 \le \left(z^2 + 4\pi x\right) \sum_{n \le x} |a_n|^2 .$$

■

An immediate consequence of applying the Cauchy-Schwarz Inequality for the First Modified Large Sieve Inequality in the theorem mentioned previously is the equation

$$\begin{aligned} &\sum_{d \le z} \frac{d}{\phi(d)} \sum_{\chi^* \bmod d} \left| \sum_{n \le x} \sum_{m \le y} a_n b_m \chi(nm) \right| \\ &\le \left(z^2 + 4\pi x\right)^{1/2} \left(z^2 + 4\pi y\right)^{1/2} \left(\sum_{n \le x} |a_n|^2 \right)^{1/2} \left(\sum_{m \le y} |b_m|^2 \right)^{1/2} , \end{aligned}$$

where $\sum_{\chi^* \bmod d}$ is the sum over all primitive characters χ modulo d, $a_{n \ge 1}$ and $b_{n \ge 1}$ are complex number sequences and x, y, z are positive integers. The second Modified Large Sieve Inequality that we would look at below is actually a variation of this inequality and the second modified Large Sieve Inequality would be very useful while working on the Bombieri-Vinogradov Theorem during the next section. However, we will omit the proof of the following theorem since it is slightly complicated and we can continue using it proving the Bombieri-Vinogradov Theorem without the knowledge of the proof.

Theorem 3.7 (Second Modified Large Sieve Inequality)**.** *For sequences of complex numbers $a_{n \ge 1}$ and $b_{n \ge 1}$ and positive integers x, y and z, we have*

$$\begin{aligned} &\sum_{ds \le z} \frac{d}{\phi(d)} \sum_{\chi^* mod\, d} \max_u \left| \sum_{n \le x} \sum_{\substack{m \le y \\ nm \le u}} a_n b_m \chi(nm) \right| \\ &\ll \left(z^2 + x\right)^{1/2} \left(z^2 + y\right)^{1/2} \left(\sum_{n \le x} |a_n|^2 \right)^{1/2} \left(\sum_{m \le y} |b_m|^2 \right)^{1/2} \log(2xy). \end{aligned}$$

where, $\sum_{\chi^* mod\, d}$ is the sum over primitive characters χ modulo k.

3.3 Barban-Davenport-Halberstam Theorem

In this subsection, we will look at a very important result deduced by Barban, Davenport, and Halberstam about the error term in the Prime Number Theorem for Primes in Arithmetic Progressions. We will use the fact from 2.4 that $\pi(x;d,a)$ can be replaced by $\psi(x;d,a)$ and hence, we will focus on the function $\psi(x;d,a)$ rather than $\pi(x;d,a)$.

Theorem 3.8 (The Barban-Davenport-Halberstam Theorem)**.** *For any $A > 0$ and for any z satisfying $x/\left(\log x\right)^A \leq z \leq x$, we have*

$$\sum_{d \leq z} \sum_{\substack{1 \leq a \leq d \\ (a,d)=1}} \left| \psi(x;d,a) - \frac{x}{\phi(d)} \right|^2 \leq xz \log x. \tag{3.5}$$

Proof. Using 2.7, we can write the error term $\psi(x;d,a) - \frac{x}{\phi(d)}$ in a more convenient form as

$$\psi(x;d,a) - \frac{x}{\phi(d)} = \frac{1}{\phi(d)} \sum_{\substack{\chi \bmod d \\ \chi \neq \chi_0}} \bar{\chi}(a)\psi(x,\chi) + \frac{\psi(x,\chi_0) - x}{\phi(d)}$$

By using orthogonality relations, we notice that

$$\sum_{\substack{\chi \bmod d \\ \chi \neq \chi_0}} \bar{\chi}(a)\psi(x,\chi) = \sum_{\substack{\chi \bmod d \\ \chi \neq \chi_0}} \psi(x,\chi).$$

We can now simplify the inequality 3.5 as

$$\sum_{d \leq z} \sum_{\substack{1 \leq a \leq d \\ (a,d)=1}} \left| \psi(x;d,a) - \frac{x}{\phi(d)} \right|^2 \ll \sum_{d \leq z} \frac{1}{\phi(d)} \sum_{\substack{\chi \bmod d \\ \chi \neq \chi_0}} \left|\psi(x,\chi)\right|^2 + \sum_{d \leq z} \frac{1}{\phi(d)} \left|\psi(x,\chi_0) - x\right|^2 \tag{3.6}$$

Now, we note that since the character $\chi \neq \chi_0$ modulo d is induced by some primitive character χ_1 modulo d_1,

$$\psi(x,\chi) = \psi(x,\chi_1) + O((\log x)(\log d)).$$

Hence, from 2.9, we have

$$\psi(x,\chi_1) - \psi(x,\chi) \ll (\log xd)$$

Using the above mentioned inequality, we can further simplify 3.6 as

$$\ll \sum_{d \leq z} \frac{1}{\phi(d)} \sum_{\substack{\chi \bmod d \\ \chi \neq \chi_0}} \left|\psi(x,\chi_1)\right|^2 + \sum_{d \leq z} \frac{1}{\phi(d)} \left|\psi(x) - x\right|^2 + z\left(\log zx\right)^2 \tag{3.7}$$

Recall from 2.8 that $\psi(x)$ can be estimated as $\psi(x) = x + O(xe^{-c\sqrt{\log x}})$. Thus,

$$\sum_{d \leq z} \frac{1}{\psi(d)} \left|\psi(x) - x\right|^2 \ll \frac{x^2 \log z}{\left(\log x\right)^A} \ll xz \log x, \tag{3.8}$$

where $A > 0$ is as in the statement of the theorem.

Since the term $z(\log zx)^2$ is negligible compared to the desired final estimate of $xz\log x$ and comparing the inequality of 3.8 with 3.7 and 3.5, we can see that in order to prove the theorem, it now suffices to show that

$$\sum_{d\leq z}\frac{1}{\phi(d)}\sum_{\substack{\chi \bmod d\\ \chi\neq\chi_0}}|\psi(x,\chi_1)|^2 \ll xz\log x. \tag{3.9}$$

For χ modulo d where $\chi \neq \chi_0$, let χ_1 modulo d_1 be its associated primitive character and let $d = d_1 k$ be the induced modulus for some positive integer k. We have

$$\begin{aligned}\sum_{d\leq z}\frac{1}{\phi(d)}\sum_{\substack{\chi \bmod d\\ \chi\neq\chi_0}}|\psi(x,\chi_1)|^2 &= \sum_{d_1\leq z}\sum_{\chi_1 \bmod d_1}|\psi(x,\chi_1)|^2\sum_{k\leq\frac{z}{d_1}}\frac{1}{\phi(d_1k)}\\ &= \sum_{d\leq z}\sum_{\chi^* \bmod d}|\psi(x,\chi)|^2\sum_{k\leq\frac{z}{d}}\frac{1}{\phi(dk)}\end{aligned}$$

where, $\sum_{\chi^* \bmod d}$ is the sum over primitive characters χ modulo d.

We will omit the proof of this but we note that $\sum_{k\leq\frac{z}{d}}\frac{1}{phi(dk)} \ll \frac{1}{phi(d)}\log\frac{2z}{d}$. Therefore, in order to prove 3.9,we note that it suffices to show that

$$\sum_{d\leq z}\frac{1}{\phi(d)}\log\frac{2z}{d}\sum_{\chi^* \bmod d}|\psi(x,\chi)|^2 \ll xz\log x. \tag{3.10}$$

Now, we turn onto the First Modified Large Sieve Inequality to prove 3.10. We choose $a_n := \Lambda(n)$ and we have

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d}|\psi(x,\chi)|^2 \ll \left(z^2+x\right)x\log x, \tag{3.11}$$

since we can easily observe that $\sum_{n\leq x}\Lambda(n)^2 \ll x\log x$ using the average order of $\Lambda(n)$ from 2.4

Now, we let $D := D(x)$ be a parameter to be chosen later such that $1 < D \leq z$, and let us divide the interval $(D, z]$ into *dyadic sub-intervals* $(U, 2U]$ with $U := z/2^k$, where k are integers running from 1 to $\log(z - 2D/D)$. D together with partial summation in 3.11 leads to

$$\sum_{U<d\leq 2U}\frac{1}{\phi(d)}\log\left(\frac{2z}{d}\right)\sum_{\chi^* \bmod d}|\psi(x,\chi)|^2 \ll \left(\frac{x^2}{U}+Ux\right)(\log x)\log\left(\frac{2z}{U}\right)$$

for each interval $(U, 2U]$. Summing over all intervals $U = z/2^k$, we obtain

$$\sum_{D<d\leq z}\frac{1}{\phi(d)}\log\left(\frac{2z}{d}\right)\sum_{\chi^* \bmod d}|\psi(x,\chi)|^2 \ll \frac{x^2}{D}(\log x)^2 + zx\log x.$$

Now, we will choose $D := (\log x)^{A+1}$ and recall that $\frac{x}{(\log x)^A} \leq z \leq x$ so that we can reduce the above inequality to $\ll zx \log x$ altogether. Now, we observe that it is enough to show

$$\sum_{2 \leq d \leq D} \frac{1}{\phi(d)} \log\left(\frac{2z}{d}\right) \sum_{\chi^* \bmod d} |\psi(x, \chi)|^2 \ll zx \log x. \tag{3.12}$$

in order to prove 3.9. To show that 3.11 holds true, we use an inequality followed from the Siegel-Walfiz Theorem (2.8) which states that $\psi(x, \chi) \ll xe^{-C(A)\sqrt{\log x}}$.

$$\sum_{2 \leq d \leq D} \frac{1}{\phi(d)} \log\left(\frac{2z}{d}\right) \sum_{\chi^* \bmod d} |\psi(x, \chi)|^2 \ll D(\log z)\, x^2 e^{-2C(A)\sqrt{\log x}}$$

Comparing the above inequality with 3.8, we note that $D(\log z)\, x^2 e^{-2C(A)\sqrt{\log x}}$ is similar to $O(\left(xe^{-2C(A)\sqrt{\log x}}\right)^2)$. Hence, we can simplify the above inequality as

$$\begin{aligned} \sum_{2 \leq d \leq D} \frac{1}{\phi(d)} \log\left(\frac{2z}{d}\right) \sum_{\chi^* \bmod d} |\psi(x, \chi)|^2 &\ll \frac{x^2}{(\log x)^A} \\ &\ll zx \log x. \end{aligned}$$

This completes the proof of the theorem and we establish the Barban-Davenport-Halberstam Theorem as we worked through 3.9, 3.10, and 3.12. ■

4 Bombieri-Vinogradov Theorem

We now move onto the crux of this paper, The Bombieri-Vinogradov Theorem. We will first dedicate a small subsection which demonstrates the plan for the proof of the Bombieri-Vinogradov Theorem and also introduces a few techniques that would be directly involved in the proof of the Bombieri-Vinogradov Theorem. Then, we will prove the Bombieri-Vinogradov Theorem using various tools that we've developed as we advanced throughout the paper. **Note** that the following proof is considered as *Vaughan's method* [Vau65] and it has been adapted from [CM+06, Chapter 9].

4.1 The plan for the proof of Bombieri's Theorem

We will be following a celebrated method derived by Vaughan to prove the Bombieri-Vinogradov Theorem. However, since the proof is quite long and complex, we will be dividing the proof of the theorem into a few lemmas which will contribute to the proof of the Bombieri-Vinogradov Theorem on the big picture. In this subsection we will introduce the infamous *Vaughan's Identity* and denote the properties of a class of functions that will help us in the next section which is dedicated to the proof of the Bombieri-Vinogradov Theorem.

We establish a class of functions,

$$\mathcal{D} := \{D : \mathbb{N} \to \mathbb{C} : \sum_{n \leq x} |D(n)|^2 = O(x(\log x)^\alpha) \text{ for some } \alpha > 0\},$$

and we will take note of the following basic properties for $D \in \mathcal{D}$.

1. For $D \in \mathcal{D}$ and $\theta \geq 0$,
$$\sum_{n \leq x} \frac{|D(n)|}{n^\theta} \ll x^{1-\theta} (\log x)^\alpha$$
for some $\alpha > 0$.

2. If $D_1, D_2 \in \mathcal{D}$, then
$$\sum_{ef \leq x} |D_1(e) D_2(f)| \, d(ef) \ll x (\log x)^\beta$$
for some $\beta > 0$, and
$$\sum_{ef \leq x} |D_1(e) D_2(f)|^2 \, d(ef) \ll x (\log x)^\gamma$$
for some $\gamma > 0$, where for a positive integer e, $d(e)$ denotes the number of the divisors of e.

Now, we will take note of a connection between normalized Dirichlet Series and the set of functions defined above with a set of hypotheses that we would assume to be true. Using these hypotheses, we will establish two fundamental inequalities (Lemma 4.1) in the next subsection that would help us in proving the Bombieri-Vinogradov Theorem.

Let x, z be positive integers and let
$$A(s) = \sum_{n \geq 1} \frac{a(n)}{n^s}, \quad B(s) = \sum_{n \geq 1} \frac{b(n)}{n^s}$$
be *normalized Dirichlet Series* for which we write
$$\frac{A(s)}{B(s)} = \sum_{n \geq 1} \frac{c(n)}{n^s}, \quad \frac{1}{B(s)} = \sum_{n \geq 1} \frac{\tilde{b}(n)}{n^s}$$
for some $c(n)$, $\tilde{b}(n) \in \mathbb{C}$. Followed from the half-plane of absolute convergence property of *Dirichlet Series*, we assume that these series are convergent for $Re(s) > \sigma_0$ for some σ_0. Furthermore, we assume that the they satisfy the following hypothesis, which partially gives a relation between the class of functions $\mathcal{D}$ we studies earlier and the normalized Dirichlet Series.

(H1) $a(n)_{n \geq 1}$ is an increasing sequence of positive real numbers.

(H2) The functions $b, \tilde{b}, c \in \mathcal{D}$.

(H3) There exists $0 \leq \theta < 1$ and $0 \leq \gamma < 1$ such that, for any non-trivial Dirichlet character *chi* modulo d,
$$\sum_{n \leq x} b(n) \chi(n) \ll x^\theta \sqrt{d} \log d + x^\gamma.$$

4.2 Proof of the Bombieri-Vinogradov Theorem

As discussed earlier, in this subsection, we will be working through a few lemmas first that would eventually contribute to the proof of Bombieri's Theorem. Assuming H(1), H(2), and H(3) from the first subsection to be true, we will continue the same notation establish Lemma 4.1. However, due to the complexity and the length of the proof of this lemma, we will only briefly discuss about its proof by stating only the important equations which progress towards proving the lemma.

Lemma 4.1. *If* $z \leq x^{\frac{1-\theta}{3-\theta}}$,

$$\sum_{d\leq z} \frac{d}{\phi(d)} \sum_{\chi^* \ mod\ d} \max_{y\leq x} \left| \sum_{n\leq y} c(n)\chi(n) \right| \\ \ll \left(z^2x^{1/2} + x + zx^{\frac{5-\theta}{2(3-\theta)}} + z^2x^{\frac{1-\theta+2\gamma}{3-\theta}} + z^{5/2}x^{\frac{1+\theta}{3-\theta}}a(x) \right) (\log x)^{\alpha'},$$

for some $\alpha' > 0$.

If $z > x^{\frac{1-\theta}{3-\theta}}$,

$$\sum_{d\leq z} \frac{d}{\phi(d)} \sum_{\chi^* \ mod\ d} \max_{y\leq x} \left| \sum_{n\leq y} c(n)\chi(n) \right| \\ \ll \left(z^2x^{1/2} + x + z^{\frac{9-4\theta}{2(3-2\theta)}}x^{\frac{2-\theta}{3-2\theta}}a(x)\,(\log z) + z^{\frac{3-4\theta+3\gamma}{3-2\theta}}x^{\frac{2-2\theta+\gamma}{3-2\theta}}\,(\log z) \right) (\log x)^{\alpha''}$$

for some $\alpha'' > 0$.

Proof. Continuing with the notation discussed in the first subsection, we set

$$F(s) := \sum_{n\leq U} \frac{c(n)}{n^s}, \quad G(s) := \sum_{n\leq V} \frac{\tilde{b}(n)}{n^s}$$

for some parameters $U = U(x,z)$ and $V = V(x,z)$, to be chosen later. We also note that $F(s)$ can be thought of as an approximation to the the normalized Dirichlet Series of $A(s)/B(s)$ and $G(s)$ can be thought of as an approximation of $1/B(s)$. Using *Vaughan's Identity*, we observe that,

$$\frac{A(s)}{B(s)} = F(s) - B(s)G(s)F(s) + A(s)G(s) + \left(\frac{A(s)}{B(s)} - F(s) \right) (1 - B(s)G(s)). \tag{4.1}$$

By comparing the coefficients of n^{-s} on both sides of the normalized Dirichlet Series in 4.1, we can deduce that

$$c(n) = a_1(n) + a_2(n) + a3(n) + a_4(n)$$

where

$$a_1(n) := \begin{cases} c(n) & if\ n \leq U, \\ 0 & if\ n > U, \end{cases} \tag{4.2}$$

$$a_2(n) := -\sum_{\substack{efg=n\\ f\le V\\ g\le U}} b(e)\tilde{b}(f)c(g), \tag{4.3}$$

$$a_3(n) := \sum_{\substack{ef=n\\ f\le V}} a(e)\tilde{b}(f), \tag{4.4}$$

$$a_4(n) := -\sum_{\substack{ef=n\\ e>U\\ f>V}} c(e) \sum_{\substack{gh=f\\ h\le V}} b(g)\tilde{b}(h). \tag{4.5}$$

Hence, we can use this expression of $c(n)$ to obtain

$$\sum_{n\le y} c(n)\chi(n) = \sum_{1\le i\le 4}\sum_{n\le y} a_i(n)\chi(n), \tag{4.6}$$

for any Dirichlet character χ modulo d. Furthermore, we denote the above mentioned equation as

$$\sum_{1\le i\le 4} S_i(y,x) := \sum_{1\le i\le 4}\sum_{n\le y} a_i(n)\chi(n).$$

The proof of the theorem hereafter is concerned with estimating each of the sums

$$\sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d}\max_{y\le x}\left|S_i(y,\chi)\right|,$$

for $1\le i\le 4$.

The estimate for $S_1(y,\chi)$. Using equation 4.2 and hypothesis 2, (H2) together with property 1 discussed in the aforementioned subsection, we obtain

$$|S_1(y,\chi)| = \left|\sum_{\substack{n\le y\\ n\le U}} c(n)\chi(n)\right| \ll \sum_{n\le U}|c(n)| \ll U\left(\log U\right)^{\alpha_0}$$

for some $\alpha_0>0$. Recalling that there are $\phi(d)$ Dirichlet Characters modulo d, we can obtain

$$\sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d}\max y\le x\left|S_1(y,\chi)\right| \ll z^2 U\left(\log U\right)^{\alpha_0}, \tag{4.7}$$

where the above bound is independent of $y, \chi \bmod d$.

The estimate for $S_2(y,\chi)$. Using equation 4.3, we denote

$$S_2(y,\chi) = -\sum_{\substack{efg\le y\\ f\le V\\ g\le U}} b(e)\tilde{b}(f)c(g)\chi(efg),$$

and split it into parts based on whether $fg\le U$ or $U<fg\le UV$. We denote the first part of sum obtained in this way as $S_2'(y,\chi)$ and the second sum as $S_2''(y,\chi)$.

For S_2', we write

$$|S_2'| \leq \sum_{g\leq U} |c(g)| \sum_{f\leq \min\left(V,\frac{U}{g}\right)} \left|\tilde{b}(f)\right| \left|\sum_{e\leq \frac{y}{fg}} b(e)\chi(e)\right|, \tag{4.8}$$

and for S_2'', we write

$$\begin{aligned}\sum_{d\leq z}\frac{d}{\phi(d)} &\sum_{\chi^* \bmod d} \max y \leq x \, |S_2''(y,\chi)| \\ &= \sum_{d\leq z}\frac{d}{\phi(d)} \sum_{\chi^* \bmod d} \max y \leq x \left| \sum_{\substack{eh\leq y\\ U<h\leq UV}} b(e) \left(\sum_{\substack{f\leq V, g\leq U\\ fg=h}} \tilde{b}(f)c(g)\right)\chi(eh)\right|.\end{aligned} \tag{4.9}$$

Now, we will work on the bounds independently for $S_2'(y,\chi)$ and $S_2''(y,\chi)$.

We will first deduce an estimate for the innermost sum of $S_2'(y,\chi)$, using hypothesis (H1). Then, by using hypothesis (H2) with property 1 of the class of functions defined in the aforementioned subsection, we expand more on this estimate to finally obtain

$$S_2'(y,\chi) \ll y^{\theta}\sqrt{d}(\log d)U^{1-\theta}(\log U)^{\alpha_3} + y^{\gamma}U^{1-\gamma}(\log U)^{\alpha_4}, \tag{4.10}$$

for some $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4 > 0$. From the above equation (4.10), we deduce the estimate for $S_2'(y,\chi)$ as

$$\begin{aligned}\sum_{d\leq z}\frac{d}{\phi(d)} &\sum_{\chi^* \bmod d} \max y \leq x \, |S_2'(y,\chi)| \\ &\ll x^{\theta}z^{5/2}(\log z)U^{1-\theta}(\log U)^{\alpha_3} + x^{\gamma}z^{2}(\log z)U^{1-\gamma}(\log U)^{\alpha_4}.\end{aligned}$$

For finding the estimate of S_2'', we start with using the second Modified Large Sieve Inequality from the last subsection to obtain

$$\begin{aligned}\sum_{d\leq z}\frac{d}{\phi(d)} &\sum_{\chi^* \bmod d} \max y \leq x \, |S_2''(y,\chi)| \\ &= \sum_{d\leq z}\frac{d}{\phi(d)} \sum_{\chi^* \bmod d} \max y \leq x \left| \sum_{\substack{eh\leq y\\ U<h\leq UV}} b(e) \left(\sum_{\substack{f\leq V, g\leq U\\ fg=h}} \tilde{b}(f)c(g)\right)\chi(eh)\right|.\end{aligned}$$

However, as we can observe, we definitely overestimated the expression as we directly considered the sequences

$$(b(e))_{\frac{x}{UV}} < e \leq \frac{x}{U} \quad \text{and} \quad \left(\tilde{b}(f)c(g)\right)_{fg=h, 2^k<h\leq 2^{k+1}}.$$

Hence, we will again divide the interval $(U, UV]$ into dyadic intervals $(2^k, 2^{k+1}]$ with $[\log_2 U < k < \log_2 UV]$ analogous to the proof of the Barban-Davenport-Halberstam Theorem in the last subsection. Applying the modified large sieve inequality to each of the pair of sequences

$$(b(e))_{\frac{x}{2^{k+1}}<e\leq\frac{x}{2^k}} and (\tilde{b}(f)c(g))_{fg=h,U<h\leq UV},$$

for each $[\log_2 U] < k < [\log_2 UV]$, we obtain that

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d}\max y \leq x\left|\sum_{\substack{eh\leq y\\ 2^k<h\leq 2^{k+1}}} b(e)\left(\sum_{\substack{f\leq V, g\leq U\\ fg=h}}\tilde{b}(f)c(g)\right)\chi(eh)\right|$$
$$\ll \left(z^2+\frac{x}{2^k}\right)^{1/2}\left(z^2+2^k\right)^{1/2}\left(\sum_{e\leq\frac{x}{2^k}}|b(e)|^2\right)^{1/2}\left(\sum_{2^k<h\leq 2^{k+1}}\left|\sum_{\substack{f\leq V, g\leq U\\ fg=h}}\tilde{b}(f)c(g)\right|^2\right)^{1/2}(\log x). \tag{4.11}$$

To further provide an accurate estimate of this bound, we use hypotheses (H2) to get

$$\sum_{e\leq\frac{x}{2^k}}|b(e)|^2 \ll \frac{x}{2^k}\left(\log\frac{x}{2^k}\right)^{\alpha_5},$$

for some $\alpha_5 > 0$. Furthermore, using the Cauchy-Schwarz inequality with hypothesis (H2) and the second property (2) of the class of functions defined in the first subsection, we obtain

$$\sum_{2^k<h\leq 2^{k+1}}\left|\sum_{\substack{f\leq V, g\leq U\\ fg=h}}\tilde{b}(f)c(g)\right|^2 \ll 2^k\left(\log 2^k\right)^{\alpha_6}$$

for some $\alpha_6 > 0$. Using these above inequalities and plugging them into 4.11 and summing the inequality over all k. We finally get the estimate for $S_2''(y,\chi)$; that is,

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d}\max y\leq x\,|S_2''(y,\chi)|$$
$$\ll\left(z^2+\frac{zx^{1/2}}{U^{1/2}}+z(UV)^{1/2}+x^{1/2}\right)x^{1/2}(\log x)^{\frac{\alpha_5}{2}+1}(\log UV)^{\frac{\alpha_6}{2}+1} \tag{4.12}$$

The estimate for $S_3(y,\chi)$ We begin by defining a step function $\mathcal{A}:\mathbb{R}\to\mathbb{R}$ by $\mathcal{A}(t) = a(1)$ if $t\leq 1$, and , generally, $\mathcal{A}(t) = a(n)-a(n-1)$ if $n-1<t\leq n$. We then observe that $a(n) = \int_0^n \mathcal{A}(t)dt$ and $\mathcal{A}$ is positive, since the sequence $(a(n))_{n\geq 1}$ is increasing. Using 4.4, we write

$$
\begin{aligned}
|S_3(y,\chi)| &= \left|\sum_{f\le V}\tilde{b}(f)\chi(f)\sum_{e\le \frac{y}{f}} a(e)\chi(e)\right| \\
&= \left|\sum_{f\le V}\tilde{b}(f)\chi(f)\sum_{e\le \frac{y}{f}} \chi(e)\int_0^e \mathcal{A}(t)dt\right| \\
&= \left|\int_0^y \mathcal{A}(t)\sum_{f\le V}\tilde{b}(f)\chi(f)\sum_{t\le e\le \frac{y}{f}} \chi(e)dt\right| \\
&\le \int_0^y |\mathcal{A}(t)|\sum_{f\le V}\left|\tilde{b}(f)\right|\left|\sum_{t\le e\le \frac{y}{f}} \chi(e)\right| dt.
\end{aligned}
$$

Using the Pólya-Vinogradov Theorem, we estimate one of the inner sum in the above equation and obtain

$$
\begin{aligned}
S_3(y,\chi) &\ll \sqrt{d}(\log d)\int_0^y |\mathcal{A}(t)|dt\sum_{f\le V}|\tilde{b}(f)| \\
&\ll \sqrt{d}(\log d)V(\log V)^{\alpha_7}a(y)
\end{aligned}
$$

for some $\alpha_7 > 0$.
By observing that $|\mathcal{A}(t)| = \mathcal{A}(t)$ and by using hypothesis (H1), we further obtain

$$
\begin{aligned}
\sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d}\max_{y\le x}|S_3(y,\chi)| &\ll z^{5/2}\,(\log z)\,V\,(\log V)^{\alpha_7}\max y\le x\,|a(y)| \\
&= z^{5/2}(\log z)V(\log V)^{\alpha_7}a(x).
\end{aligned}
$$

The estimate for $S_4(y,\chi)$. To begin with, we use equation 4.5 to write

$$
\begin{aligned}
&\sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d}\max_{y\le x}|S_4(y,\chi)| \\
&\quad = \sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d}\max_{y\le x}\left|\sum_{\substack{ef\le y\\ e>>U\\ f>V}} c(e)\left(\sum_{\substack{gh=f\\ h\le V}} b(g)\tilde{b}(h)\right)\chi(ef)\right|,
\end{aligned}
\tag{4.13}
$$

We can observe that the second modified large sieve inequality has been applied in the above equation for the pair of sequences of complex numbers

$$
(c(e))_{U<e<\frac{x}{V}} \text{ and } \left(b(g)\tilde{b}(h)\right)_{gh=f,h\le V,V<f<\frac{x}{U}}.
$$

However, to be more precise with the estimate, we proceed in the same way as we did for finding the estimate of $S_2''(y,\chi)$. We will divide the interval $(U, y/V]$ into dyadic intervals $(2^k, 2^{k+1}]$ with $[\log_2 U < k < \log_2 y/V]$. Applying the modified large sieve inequality to each of the pair of sequences

$$(c(e))_{2^k<e\leq 2^{k+1}} and (b(g)\tilde{b}(h))_{gh=f, h\leq V, \max\left(V, \frac{y}{2^{k+1}}\right)<f<\frac{y}{2^k}}.$$

For each $[\log_2 U] < k < [\log_2 y/V]$, we obtain that

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d} \max y \leq x \left| \sum_{\substack{ef\leq y \\ e>U \\ f>V \\ 2^k<e\leq 2^{k+1}}} c(e)\left(\sum_{\substack{gh=f \\ h\leq V}} b(g)\tilde{b}(h)\right)\chi(ef)\right|$$
$$\ll \left(z^2+2^k\right)^{1/2}\left(z^2+\frac{x}{2^k}\right)^{1/2}\left(\sum_{2^k<e\leq 2^{k+1}}|c(e)|^2\right)^{1/2}\left(\sum_{V<f\leq \frac{x}{2^k}}\left|\sum_{\substack{gh=f \\ h\leq V}} b(g)\tilde{b}(h)\right|^2\right)^{1/2}(\log x). \tag{4.14}$$

Again, we use hypotheses (H2) and Cauchy-Schwarz Inequality together with hypothesis (H2) and property 2 mentioned in the first subsection to eventually get the inequalities

$$\sum_{2^k<e\leq 2^{k+1}}|c(e)|^2 \ll 2^k\left(\log 2^k\right)^{\alpha_8}$$
$$\sum_{V<f\leq \frac{x}{2^k}}\left|\sum_{\substack{gh=f \\ h\leq V}} b(g)\tilde{b}(h)\right|^2 \ll \frac{x}{2^k}\left(\log\frac{x}{2^k}\right)^{\alpha_9}$$

for $\alpha_8, \alpha_9 > 0$. Finally we plug these inequalities into 4.14 and sum it over all $k \leq x$ to obtain the following estimate for $S_4(y,\chi)$,

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi^* \bmod d} \max y \leq x\, |S_4(y,\chi)|$$
$$\ll ()\, x^{1/2}\left(z^2+\frac{zx^{1/2}}{U^{1/2}}+\frac{zx^{1/2}}{V^{1/2}}+x^{1/2}\right)x^{1/2}(\log x)^{\frac{\alpha_8+\alpha_9}{2}+2}.$$

Now, we have completed estimating all the four sums $S_i(y,\chi)$, for all $1 \leq i < 4$. By putting these estimates together, we obtain

$$\begin{aligned}
&\sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi}^{*}\max_{y\le x}\left|\sum_{n\le y}c(n)\chi(n)\right| \\
&\ll z^2U(\log U)^{\alpha_0} \\
&\quad + x^{\theta}z^{5/2}(\log z)U^{1-\theta}(\log U)^{\alpha_3} + x^{\gamma}z^2(\log z)U^{1-\gamma}(\log U)^{\alpha_4} \\
&\quad + \left(z^2 + \frac{zx^{1/2}}{U^{1/2}} + z(UV)^{1/2} + x^{1/2}\right)x^{\frac{1}{2}}(\log x)^{\frac{\alpha_5}{2}+1}(\log UV)^{\frac{\alpha_6}{2}+1} \\
&\quad + z^{5/2}(\log z)V(\log V)^{\alpha_7}a(x) \\
&\quad + \left(z^2 + \frac{zx^{1/2}}{U^{1/2}} + \frac{zx^{1/2}}{V^{1/2}} + x^{1/2}\right)x^{\frac{1}{2}}(\log x)^{\frac{\alpha_8+\alpha_9}{2}+2}.
\end{aligned} \tag{4.15}$$

It only remains for us to choose the parameters U and V appropriately to prove the lemma. We look for U and V such that, $V = x^{\theta}U^{1-\theta}$. To do so, we begin with analyzing the expression

$$E(x,z,U) := \frac{zx}{U^{1/2}} + zU\left(z^{3/2}x^{\theta}U^{\theta} + x^{\frac{1+\theta}{2}}U^{-\frac{\theta}{2}}\right).$$

If

$$z \le x^{\frac{1-\theta}{3}}U^{\frac{\theta}{3}}, \tag{4.16}$$

$$E(x,z,U) \ll \frac{zx}{U^{1/2}} + zx^{\frac{1+\theta}{2}}U^{\frac{2-\theta}{2}}.$$

Hence, we choose U such that

$$\frac{zx}{U^{1/2}} = zx^{\frac{1+\theta}{2}}U^{\frac{2-\theta}{2}}.$$

Therefore, we define $U := x^{\frac{1-\theta}{3-\theta}}$ and substituting this choice of U in 4.16 implies that $z \le x^{\frac{1-\theta}{3-\theta}}$. With this choice of U, we attain the first desired inequality.

Similarly, if

$$z^{3/2}x^{\theta}U^{-\theta} > x^{\frac{1+\theta}{2}}U^{-\frac{\theta}{2}},$$

we proceed to choose U such that

$$\frac{zx}{U^{1/2}} = z^{5/2}x^{\theta}U^{1-\theta}.$$

Therefore, we define

$$U := \frac{x^{\frac{2(1-\theta)}{3-2\theta}}}{z^{\frac{3}{3-2\theta}}}$$

and substituting this choice of U in 4.16 gives us the second desired inequality of this lemma. With this choice of U, we also get that $z > x^{\frac{1-\theta}{3-\theta}}$. ■

We will now look at one of the main ingredient need in Vaughan's proof of the Bombieri-Vinogradov Theorem that is actually a particular case of the inequalities we discussed in this section.

Lemma 4.2. *Let x and z be arbitrary positive integers. Then*

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi^*\ mod\, d}\max_{y\leq x}\left|\sum_{n\leq y}\Lambda(n)\chi(n)\right| \ll \left(z^2x^{1/2}+x+zx^{5/6}\right)(\log z)(\log x)^{\alpha} \tag{4.17}$$

for some $\alpha > 0$, where the summation $\sum_{\chi^*\ mod\, d}$ is over primitive characters χ modulo d and $\Lambda(n)$ denotes the von Mangoldt function.

Proof. Using the notation introduced in Lemma 4.1, we set

$$A(s) := -\zeta'(s) = \sum_{n\geq 1}\frac{\log n}{n^s},\ B(s) := \zeta(s) = \sum_{n\geq 1}\frac{1}{n^s}.$$

Hence, for any $n \geq 1$, we use Vaughan's Identity and denote

$$a(n) = \log n,\ b(n) = 1,\ c(n) = \Lambda(n),\ \tilde{b}(n) = \mu(n).$$

We also note that hypothesis (H1) and hypothesis (H2) are satisfied. Using the definition of $b(n) = 1$ from above, and the Pólya-Vinogradov Theorem, we can say that

$$\sum_{n\leq x} b(n)\chi(n) = \sum_{n\leq x}\chi(n) \ll \sqrt{d}\log d,$$

which satisfies hypothesis (H3) as well. Thus, with $\theta = 0 = \gamma$, we obtain from lemma 4.1 that if $x \leq x^{1/3}$

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi^*\ \mathrm{mod}\ d}\max y\leq x\left|\sum_{n\leq y}\Lambda(n)\chi(n)\right| \\ \ll \left(z^2x^{1/2}+x+zx^{5/6}+z^{5/2}x^{1/3}\right)(\log z)(\log x)^{\alpha'}$$

for some $\alpha' > 0$. If $z > x^{1/3}$, we obtain the other inequality

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi^*\ \mathrm{mod}\ d}\max y\leq x\left|\sum_{n\leq y}\Lambda(n)\chi(n)\right| \\ \ll \left(z^2x^{1/2}+x+z^{3/2}x^{2/3}\right)(\log z)(\log x)^{\alpha''}$$

for some $\alpha'' > 0$. Combining the two equations mentioned above, we get the desired inequality and hence, we complete the proof of this lemma. ■

An immediate consequence of the previous lemma is the inequality

$$\sum_{D<d\leq z}\frac{d}{\phi(d)}\sum_{\chi^*\ \mathrm{mod}\ d}\max_{y\leq x}\left|\sum_{n\leq y}\Lambda(n)\chi(n)\right| \\ \ll \left(zx^{1/2}+\frac{x}{z}+\frac{x}{D}+x^{5/6}\log z\right)(\log z)(\log x)^{\alpha}, \tag{4.18}$$

where x, z, D are positive integers such that $z > D$ and $\alpha > 0$.

Using the above mentioned lemma and the other main ingredient of Vaughan's proof, the Barban-Davenport-Halberstam Theorem (3.5), we will now complete the proof of the Bombieri-Vinogradov Theorem.

Theorem 4.3 (The Bombieri-Vinogradov Theorem)**.** *For $A > 0$, there exists $B = B(A) > 0$ such that*

$$\sum_{d \leq \frac{x^{1/2}}{(\log x)^B}} \max_{y \leq x} \max_{(a,d)=1} \left| \pi(y; d, a) - \frac{\operatorname{li} y}{\phi(d)} \right| \ll \frac{x}{(\log x)^A}, \tag{4.19}$$

where $\operatorname{li}(y)$ *is the* logarithmic integral *of* y.

Proof. Firstly, we note that proving 4.18 is the same as proving

$$\sum_{d \leq \frac{x^{1/2}}{(\log x)^B}} \max_{y \leq x} \max_{(a,d)=1} \left| \psi(y; d, a) - \frac{y}{\phi(d)} \right| \ll \frac{x}{(\log x)^A} \tag{4.20}$$

From 4.20, we can observe that the Bombieri-Vinogradov Theorem likely gives a stronger estimate than the Barban-Davenport-Halberstam theorem. Hence, we will start the proof from the Barban-Davenport-Halberstam Theorem and expand it using the lemmas we discussed in this section to obtain a stronger estimate. We notice that

$$\max_{(a,d)=1} \left| \psi(y; d, a) - \frac{y}{\phi(d)} \right| \leq \frac{1}{\phi(d)} \sum_{\substack{\chi \bmod d \\ \chi \neq \chi_0}} |\psi(y, \chi)| + \frac{\psi(y, \chi_0) - y}{\phi(d)} \tag{4.21}$$

We consider that the character $\chi \neq \chi_0$ modulo d is induced by some primitive character χ_1 modulo d_1. using 2.5 which states that $\psi(y, \chi_1) - \psi(y, \chi) \ll (\log y)(\log d)$ and summing over $d \leq z$ and taking the $\max y \leq x$, we get

$$\begin{aligned} \sum_{d \leq z} \max_{y \leq x} \max_{(a,d)=1} \left| \psi(y; d, a) - \frac{y}{\phi(d)} \right| &\ll \sum_{d \leq z} \frac{1}{\phi(d)} \sum_{\substack{\chi \bmod d \\ \chi \neq \chi_0}} \max_{y \leq x} |\psi(y, \chi_1)| \\ &+ \sum_{d \leq z} \frac{1}{\phi(d)} \max_{y \leq x} |\psi(y) - y| + z(\log y)(\log d), \end{aligned} \tag{4.22}$$

where $z = z(x)$ is a positive real number, depending on x, that we will specify soon. To estimate the second term in the above equation, we use an estimate from the prime number theorem which states that

$$\sum_{d \leq z} \frac{1}{\phi(d)} \max_{y \leq x} |\psi(y) - y| \ll \frac{x \log z}{(\log x)^{A+1}}. \tag{4.23}$$

Since the rightmost part of the inequality in 4.22, $z(\log y)(\log d)$ is negligible, it only now suffices to estimate the first term to prove the theorem.

We first write each modulus d as $d = d_1 k$ for some positive integer k and note that

$$\begin{aligned} \sum_{d \leq z} \frac{1}{\phi(d)} \sum_{\substack{\chi \bmod d \\ \chi \neq \chi_0}} \max_{y \leq x} |\psi(y, \chi_1)| &= \sum_{d_1 \leq z} \sum_{k \leq \frac{z}{d_1}} \frac{1}{\phi(d_1 k)} \sum_{\chi_1 \bmod d_1} \max_{y \leq x} |\psi(y, \chi_1)| \\ &\ll \sum_{d \leq z} \frac{1}{\phi(d)} \log \frac{2z}{d} \sum_{\chi^* \bmod d} |\psi(y, \chi)|, \end{aligned} \tag{4.24}$$

where we used the estimate

$$\sum_{k\le \frac{z}{d}} \frac{1}{\phi(dk)} \ll \frac{1}{\phi(d)} \log \frac{2z}{d}.$$

Now, to complete the proof of the theorem, we need to choose z of the form $z = x^{1/2}/(\log x)^B$ for some constant $B = B(A)$ to show that

$$\sum_{d\le z} \frac{1}{\phi(d)} \sum_{\chi^* \bmod d} \max_{y\le x} |\psi(y,\chi)| \ll \frac{x}{(\log x)^A}.$$

Since the Siegel-Walfiz Theorem states that there exists $B = B(A) > 0$ such that, if $d \le (\log x)^B$ and $\chi \neq \chi_0$ is a character modulo d, then

$$\psi(y,\chi) \ll \frac{x}{(\log x)^{A+1}}.$$

Therefore, we deduce that

$$\sum_{d\le(\log x)^B} \frac{1}{\phi(d)} \sum_{\chi^* \bmod d} \max_{y\le x} |\psi(y,\chi)| \ll \frac{x}{(\log x)^A}. \tag{4.25}$$

Now, we choose

$$z := \frac{x^{1/2}}{(\log x)^B}$$

and use equation 4.18 together with partial summation to obtain

$$\sum_{(\log x)^B<d\le z} \frac{1}{\phi(d)} \sum_{\chi^* \bmod d} \max_{y\le x} |\psi(y,\chi)| \ll \frac{x}{(\log x)^A}. \tag{4.26}$$

From 4.23, 4.25, and 4.26, we can conclude that

$$\sum_{d\le \frac{x^{1/2}}{(\log x)^B}} \max_{y\le x} \max_{(a,d)=1} \left| \pi(y;d,a) - \frac{li\, y}{\phi(d)} \right| \ll \frac{x}{(\log x)^A}.$$

Hence, we complete the proof the Bombieri-Vinogradov Theorem. ∎

4.3 What does all of this mean?

In this subsection, we will look at certain applications of the Bombieri-Vinogradov theorem and we will establish its prominence in research concerning the distribution of primes.

We can now observe that the Bombieri-Vinogradov Theorem actually provides a more precise estimate than the Siegel-Walfiz Theorem on the Error term for the Prime number theorem on primes in Arithmetic Progressions. To show this, we first represent

$$E(x;q) = \max_{(a,q)=1} \sup_{y\le x} \left| \psi(y;q,a) - \frac{y}{\phi(q)} \right|$$

as the maximum possible error term in the Prime Number Theorem for primes in arithmetic progressions for any congruence class modulo q.

If we observe closely, The Bombieri-Vinogradov Theorem actually establishes a very precise estimate for the error term. We will not be discussing about the Generalized Riemann Hypothesis since it is quite deep and complex to understand. However, the Generalised Riemann is considered to be this huge sorts of unsolved problem in Analytic Number Theory and is often assumed to be true in various proofs. Assuming the *GRH*, the Siegel-Walfiz Theorem we discussed previously is proved and it also gives a *big-oh* estimate of $O(x^{1/2})$ for the error term. A very prominent result of the Bombieri-Vinogradov Theorem is that it gives a much more precise estimate of the error term without assuming the *GRH* to be true.

Other major applications of the Bombieri-Vinogradov Theorem include the Titchmarsh Divisor Problem [CM$^+$06, Chapter 9.3], and the Elliot-Halberstam Conjecture. We will not be discussing about these topics since they are quite deep but an interested reader can refer [Sou06] for more information regarding the Elliot-Halberstam Conjecture and studies concerning the *Gaps between primes*.

References

[Apo98] Tom M Apostol. *Introduction to analytic number theory.* Springer Science & Business Media, 1998.

[CM$^+$06] Alina Carmen Cojocaru, M Ram Murty, et al. *An introduction to sieve methods and their applications*, volume 66. Cambridge University Press, 2006.

[Dav13] Harold Davenport. *Multiplicative number theory*, volume 74. Springer Science & Business Media, 2013.

[Mur08] M Ram Murty. *Problems in analytic number theory*, volume 206. Springer Science & Business Media, 2008.

[Sah] Anurag Sahay. The bombieri-vinogradov theorem.

[Sou06] K. Soundararajan. Small gaps between prime numbers: The work of goldston-pintz-yildirim, 2006.

[Ten15] Gérald Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 163. American Mathematical Soc., 2015.

[Vau65] RC Vaughan. The bombieri–vinogradov theorem, 1965.