The Bombieri-Vinogradov Theorem

Anurag Sahay
SRF Application No. MATS 857
KVPY Registration No. SX-11011010

10th May - 20th July, 2013

About the Project

This report is a record of the reading project in analytic number theory undertaken by me during my visit as a summer student to the Institute of Mathematical Sciences, Chennai. The idea of the project was to gain some general maturity in the field of number theory (through lectures/talks on aspects of analytic and algebraic number theory by Mr. Sumit Giri and Dr. Prem Prakash Pandey at IMSc including proofs of the Prime Number Theorem), as well as to read the proof of a specific result, the Bombieri-Vinogradov theorem. In this report, (due to lack of space and ease of the general material compared to the specific result) a proof of the Bombieri-Vinogradov theorem is discussed, along with some other results in number theory necessary to establish the theorem. In the proof, we focus on the general ideas rather than focusing on the specifics, and the report is best read in conjunction with [1]. This project was the result of the Joint Summer Research Fellowship of the Indian Science Academies, and shall also be submitted to KVPY for the continuation of my fellowship.

Acknowledgements

I would like to thank Dr. Sanoli Gun for inviting me to IMSc for the summer, Prof. Balasubramanian for several fruitful mathematical discussions and Sumit Giri and Prem Prakash Pandey for their lectures/talks on number theory.

Notation

We shall describe here the notation that we will need that a layperson may not be familiar with.

We will use many standard notations from number theory. In particular, our asymptotic notation will always be as some parameter x goes to infinity. We shall use

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

to mean that there exists some positive constant C such that $|f(x)| \leq Cg(x)$ for sufficiently large x. Such an estimate is called a "big-oh estimate". We also use $f \ll g$ and $g \gg f$ to mean the same thing.

We use

$$f(x) = o(g(x))$$

to mean that $f(x)/g(x) \to 0$ as $x \to \infty$. Such an estimate is called a "little-oh estimate", and being $o(g(x))$ is strictly stronger than being $O(g(x))$. However, little-oh estimates are qualitative statements, and not very good for calculation. Hence, in practice, one always use more precise big-oh estimates for calculation (ie, with a smaller $g(x)$) and only return to the little-oh estimate in the last step to give a neater but strictly weaker estimate in the end, if at all. (See for example, the Prime Number Theorem).

We will often write

$$f(x) = g(x) + O(h(x)) \text{ or } f(x) = g(x) + o(h(x))$$

to mean that there exists a function $p(x)$ which is respectively $= O(h(x))$ or $= o(h(x))$ such that $f(x) = g(x) + p(x)$.

Finally we use

$$f(x) \sim g(x)$$

interchangeably with

$$f(x) = g(x) + o(g(x))$$

to denote the asymptotic equality $f(x)/g(x) \to 1$ as $x \to \infty$.

We use (a, b) to denote the greatest common divisor of a and b and $\varphi(n)$ for Euler's totient function,

$$\varphi(n) = \#\{x \in \mathbb{Z} : 1 \leq x \leq n, (x, n) = 1\}$$

For any $A \subset \mathbb{Z}$, we use $1_A(n)$ for its indicator function,

$$1_A(n) = \begin{cases} 1 & \text{if } n \in A \\ 0 & \text{otherwise} \end{cases}$$

We use p_n for the nth prime number. Furthermore, for us, n will always be an integer, p will always be prime, and $\mathcal{P}$ shall denote the set of prime numbers.

For summations and products, we shall use the standard practice of specifying the variable over which the operation is taking place under the $\sum$ or $\prod$ as well as specifying the other conditions the variable needs to satisfy. Furthermore, sums over p are over primes and sums over n are over positive integers. This may lead to sums of the form

$$\sum_{n \leq x}, \sum_{p \leq x}, \sum_{p \mid m}, \sum_{n \mid m}, \sum_{\chi \bmod q}$$

and so on, which are respectively sums over positive integers up to x, primes up to x, all prime divisors of m, all divisors of m, and all Dirichlet character modulo q.

Unless otherwise specified, all Dirichlet characters are modulo q. Further, for all Dirichlet characters χ, we denote

$$\delta(\chi) = \begin{cases} 1 & \text{if } \chi = \chi_0 \\ 0 & \text{otherwise} \end{cases}$$

where χ_0 is any principal character. For sums over primitive Dirichlet characters we add an asterisk to the sum like

$$\sideset{}{^*}\sum_{\chi \bmod q}$$

Finally, for any theorem beyond the scope of this report shall be marked with a $(\star)$ to indicate its difficulty. This means that the theorem was assumed as a black box when preparing this report because its proof is more complex than we have space for, and detracts rather than adds to the exposition of the Bombieri-Vinogradov theorem.

Contents

Chapter 1

Introduction

The Bombieri-Vinogradov theorem is a statement about the error term in Dirichlet's theorem which is useful in Number Theory. In this chapter, we will motivate the study of prime numbers, we will give the proper context in which this theorem lies, and we will introduce the notation and nomenclature that will let us state the Bombieri-Vinogradov theorem and explain its importance. The material in this report has been adapted from [1], with an intent to clarify the difficult steps in its exposition and make it easier for a non-expert to understand.

1.1 The Prime Number Theorem for Arithmetic Progressions

A central question in analytic number theory is that of the distribution of prime numbers among the positive integers. The "macrostructure" of this distribution is normally studied by examining the *prime-counting function* $\pi(x)$ given by

$$\pi(x) = \sum_{p \leq x} 1$$

where the summation is over primes less than or equal to x, and trying to determine its asymptotic behaviour as $x \to \infty$. One of the early achievements of analytic methods in number theory was the *Prime Number Theorem*

(PNT) proved independently by Hadamard and de la Valle-Poussin, which gives an asymptotic formula for $\pi(x)$ which says

$$\pi(x) \sim \frac{x}{\log x}$$

Another question of much importance in number theory is the distribution of prime numbers within arithmetic progressions. Information about this distribution can be used to prove a plethora of interesting facts about the prime numbers. Another early result (perhaps the seminal result in analytic number theory) proven by Dirichlet states that if

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

is the number of primes less than x in a given congruence class modulo q and further suppose that $(a, q) = 1$ (that is, a is *coprime* to q), then $\pi(x; q, a) \to \infty$ as $x \to \infty$.

If $(a, q) \neq 1$, there are obviously only finitely many primes in the congruence class containing a, since $p \equiv a \pmod q$ implies that any prime which divides both a and q must divide p. Thus, if $(a, q) > 1$ then the only primes which can be in the congruence class are the ones divisible by (a, q). If (a, q) is composite, then there are zero such primes, and if (a, q) is prime there is one such prime, and hence the number of primes in this congruence class is finite. Trivially, thus, any arithmetic progression has infinitely many primes if and only if the first term and common difference are coprime. This is known as “Dirichlet’s theorem on primes in arithmetic progressions”.

However, we can do much more than simply show infinitude for $(a, q) = 1$, and we are interested in obtaining a numerical estimate similar to PNT for primes in a particular progresson. There is no natural reason to expect that the primes would be more concentrated in one particular congruence class than the others. Thus, we would expect that all such congruence classes should roughly have the “same” number of primes. Since there are $\varphi(q)$ many such congruence classes we would expect that for some fixed a and sufficiently large x, $\pi(x; q, a)$ should roughly be $\pi(x)/\varphi(q)$. This turns out to be true, in what is a quantitative version of Dirichlet’s theorem which states that

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

This quantitative version of Dirichlet's theorem is known as the "Prime Number Theorem for arithmetic progressions".

While this is a deep theorem, the asymptotically equality is not sufficient and we need more information about the error in this theorem. We can write the above estimate as

$$\pi(x;q,a) = \frac{1}{\varphi(q)}\frac{x}{\log x} + o\left(\frac{x}{\log x}\right)$$

Where $o()$ is the little-oh asymptotic notation. The goal of many results is to replace the error term with a more precise big-oh estimate. In particular, the Siegel-Walfisz theorem gives such an estimate, and shall be used in our proof of the Bombieri-Vinogradov theorem.

Before moving on to that, however, we shall note that there exists an alternative way to state these theorems that is much easier to use and prove.

1.2 Chebyshev's ϑ and ψ Functions

It turns out that the prime-counting function $\pi(x;q,a)$ is very difficult to use in proofs. Instead, it has been typical since Chebyshev to replace them by the theta and psi functions, $\vartheta(x;q,a)$ and $\psi(x;q,a)$.

An alternative way to write π is the following:

$$\pi(x;q;a) = \sum_{\substack{n \le x \\ n \equiv a \bmod q}} 1_{\mathcal{P}}(n)$$

Where $1_A(n)$ is the indicator function of a set of integers A. Thus, π can be interpreted as a weighted sum over all elements in a congruence class with the prime elements weighted with 1 and the composite elements weighted with 0.

However, it turns out that this method of weighting is not ideal for proving results. Instead, a better weight is the von Mangoldt function, which we shall define presently. We thus consider instead the sum

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

where $\Lambda(n)$ is the more appropriate weight, the von Mangoldt function.

In this and the subsequent section we will provide a recipe for turning results about one of the above weighted sums to the other, and try to establish why the second sum is better suited for manipulation.

One way to motivate this is the following. Clearly, by PNT

$$\frac{\pi(x)\log x}{x} = 1 + o(1)$$

Taking natural logarithms both sides

$$\log \pi(x) - \log x + \log\log x = \log(1 + o(1)) = o(1)$$

where the last equality is easily established.[1] Now, if $x = p_n$, the nth prime number, then clearly $\pi(x) = n$. Thus we have

$$\log n - \log p_n + \log\log p_n = o(1)$$

Noting that $\log\log x = o(\log x)$, we thus get

$$\log n = \log p_n + o(\log p_n)$$

Or, in other words,

$$\log n \sim \log p_n$$

This suggests that if instead of giving all primes the same weight 1, we weight them by their logarithm, the higher primes would contribute more,

[1] As the logarithm is continuous at 1, if $f(x) = o(1)$, then $\lim_{x\to\infty} f(x) = 0$. Thus $\lim_{x\to\infty} \log(1+f(x)) = \log\left(1 + \lim_{x\to\infty} f(x)\right) = \log(1) = 0$. Hence, clearly, $\log(1+o(1)) = o(1)$.

multiplying a rough factor of a logarithm. We can formalize this heursitic by a partial summation[2] argument.

Thus, we define a new function, called the Chebyshev ϑ-function in the literature as follows

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

which weights each prime by their logarithm instead of 1.

As mentioned above, using partial summation, we can establish the following two identities

$$\pi(x) = \frac{\vartheta(x)}{\log x} + \int_2^x \frac{\vartheta(t)}{t(\log t)^2} dt$$

$$\vartheta(x) = \pi(x) \log x - \int_2^x \frac{\pi(t)}{t} dt$$

Using these, we can convert any estimate on the first function into one of the second, and vice-versa.

In particular, it is easily shown that

$$\lim_{x \to \infty} \frac{1}{x} \int_2^x \frac{\pi(t)}{t} dt = \lim_{x \to \infty} \frac{\log x}{x} \int_2^x \frac{\vartheta(t)}{t(\log t)^2} dt = 0$$

which shows that PNT is equivalent to $\vartheta(x) \sim x$. In any case, the first identity can be used to change any estimate for ϑ to one for π.

Analogous to the prime-counting function for progressions, $\pi(x; q, a)$, we can define a $\vartheta(x; q, a)$ for progressions as follows

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

The above identities can then be proved in exactly the same way by replacing $\pi(x)$ with $\pi(x; q, a)$ and $\vartheta(x)$ with $\vartheta(x; q, a)$.

[2] See Appendix

In some sense, it is natural to work with logarithms of primes when working with weighted sums. Primes are essentially multiplicative objects, and the logarithm allows one to pass from the multiplicative to the additive, and thus form a natural candidate for dealing with sums over primes. However, it turns out even weighting all primes by their logarithms and all composites by 0 does not give the most convenient form. The most convenient form is given instead by Chebyshev's ψ-function,

$$\psi(x) = \sum_{p^k \leq x} \log p$$

where the sum is over all primes p and all positive integers k such that $p^k \leq x$. In other words, we weight all prime powers by the logarithm of the prime of which they are a power, and all other numbers by 0. The hope then, is that since the prime powers contribute a smaller amount than the primes, the contribution from them can be controlled.

Clearly,

$$\psi(x) = \sum_{k=1}^{\infty} \sum_{p^k \leq x} \log p = \sum_{k=1}^{\infty} \sum_{p \leq \sqrt[k]{x}} \log p = \sum_{k=1}^{\infty} \vartheta(x^{1/k})$$

Here note that since for a fixed positive x, $\lim_{k\to\infty} x^{1/k} = 1$ thus for sufficiently large k, $x^{1/k} < 2$, and thus $\vartheta(x^{1/k}) = 0$. Thus, all but finitely many terms vanish, and in particular, the terms are non-vanishing if and only if $x^{\frac{1}{k}} \geq 2$. Taking logarithm to the base 2 on both sides, we see this is the same as requiring $k \leq \log_2 x$.

Thus,

$$\psi(x) = \sum_{k \leq \log_2 x} \vartheta(x^{\frac{1}{k}})$$

Now, trivially, $\vartheta(x) = \sum_{p \leq x} \log p \leq \sum_{p \leq x} \log x \leq x \log x$. Also, we know that $\vartheta(x)$ is increasing and thus, $\vartheta(x^{1/2}) \geq \vartheta(x^{1/k})$ for $k \geq 2$. With this we can see that

$$
\begin{aligned}
\psi(x) - \vartheta(x) &= \sum_{2 \le k \le \log_2 x} \vartheta(x^{1/k}) \\
&\le \sum_{2 \le k \le \log_2 x} \vartheta(x^{1/2}) \\
&\le \vartheta(x^{1/2}) \log_2 x \\
&\le x^{1/2} (\log_2 x)(\log x^{\frac{1}{2}}) \\
&= O\left(x^{1/2} (\log x)^2\right)
\end{aligned}
$$

Thus, any estimate for ψ can be converted into an estimate for ϑ, provided the estimate has an error larger than $O(\sqrt{x})$ by at least two logarithmic factors. In particular, since logarithms always grow slower than powers, for any $\epsilon > 0$, an error of the form $O(x^{1/2+\epsilon})$ can be tolerated. This is much tighter than most bounds we have, and thus in any theorem we shall prove here, ψ may be interchanged with ϑ and vice-versa. This also means that the PNT is equivalent to $\psi(x) \sim x$. Using the bound $\vartheta(x) = O(x)$, which is substantially weaker than PNT and was proven by Chebyshev using elementary methods, we can sharpen the estimate to $\psi(x) - \vartheta(x) = O(\sqrt{x})$.[3]

Identically to π and ϑ, we define $\psi(x; q, a)$

$$
\psi(x; q, a) = \sum_{\substack{p^k \le x \\ p^k \equiv a \bmod q}} \log p
$$

Furthermore, as above

$$
\psi(x; q, a) = \sum_{k \le \log_2 x} \vartheta(x^{1/k}; q, a)
$$

and thus,

[3] Clearly

$$
\psi(x) - \vartheta(x) = \vartheta(x^{1/2}) + \sum_{k=3}^{\lfloor \log_2 x \rfloor} \vartheta(x^{1/k}) \le \vartheta(x^{1/2}) + \vartheta(x^{1/3}) \log_2 x = O(x^{1/2})
$$

$$\begin{aligned}
\psi(x;q,a) - \vartheta(x;q,a) &= \sum_{2\le k\le \log_2 x} \vartheta(x^{\frac{1}{k}};q,a) \\
&\le \sum_{2\le k\le \log_2 x} \vartheta(x^{1/k}) \\
&= \psi(x) - \vartheta(x)
\end{aligned}$$

Hence, all comments as above apply to the Chebyshev functions of a particular progression as well.

1.3 The von Mangoldt Function

We are now in a position to define the von Mangoldt function. This function is the weight by which the ψ-function had been defined, above. In other words,

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ for some } p \in \mathcal{P} \text{ and } k \in \mathbb{Z}^+ \\ 0 & \text{otherwise} \end{cases}$$

Thus we have

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

The reason $\Lambda(n)$ is used is because it arises naturally in the Dirichlet series of the logarithmic derivative of the Riemann Zeta function. The Riemann Zeta function is defined as

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

which is absolutely convergent for $\Re(s) > 1$. In this same region, it can be shown that

$$-\frac{\zeta'}{\zeta}(s) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$$

The following identity is equivalent to the above Dirichlet series equality, and can be interpreted as an analytic statement of the fundamental theorem of arithmetic.

Theorem 1.3.1. *For any $n \in \mathbb{N}$,*

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

Proof. By the fundamental theorem,

$$n = \prod_{p^a||n} p^a$$

Hence, taking logarithms both sides

$$\begin{aligned}\log n &= \sum_{p^a||n} a \log p \\ &= \sum_{p^a||n} \sum_{k \leq a} \log p \\ &= \sum_{p^k|n} \log p \\ &= \sum_{d|n} \Lambda(d)\end{aligned}$$

where the last equality follows from the definition.

□

This theorem gives another example of how $\Lambda(n)$ can arise naturally in situations involving divisibility.

1.4 The Error Term in PNT

Now, as stated earlier, we are interested in the error term in the PNT for arithemtic progressions. The strongest known result for an individual pair a and q is the following result by Siegel and Walfisz.

Theorem 1.4.1 ($\star$). *Fix a real number $A > 0$. If $(a, q) = 1$ and $q \leq (\log x)^A$, we have*

$$\psi(x; q, a) = \frac{x}{\varphi(q)} + O_A\left(\exp(-c_1\sqrt{\log x})\right)$$

where c_1 is some absolute positive constant, and $O_A()$ means that the implicit constant depends on A, but is uniform in a and q.

The above theorem, known in the literature as the Siegel-Walfisz theorem is quite deep and we will thus omit its proof.

The *Generalized Riemann Hypothesis* (GRH), which is one of the biggest unsolved problems in number theory, is a statement about the zeroes of analytic objects associated with every arithmetic progression which are called L-functions. If GRH holds for every L-function of a given congruence class (say, $a \pmod q$), then we can show that

$$\psi(x; q, a) = \frac{x}{\varphi(q)} + O(x^{1/2}(\log x)^2)$$

where the error term is uniform for all q. In many applications, the GRH is important because of these strong bounds it gives us for the PNT in arithmetic progressions, and it essentially says that the error is $O(x^{1/2})$, ignoring logarithmic factors.

Now fix some real number x, and look at

$$\Delta(x; q) = \max_{(a,q)=1} \sup_{y \leq x} \left| \psi(y; q, a) - \frac{y}{\varphi(q)} \right|$$

then $\Delta(x; q)$ represents the maximum possible error in the PNT for any congruence class modulo q for numbers $\leq x$. If GRH holds, we would expect this to be $O(x^{1/2})$ barring logarithmic factors. Thus, if we sum all these error terms until some real positive number Q, we would expect the error to be $O(x^{1/2}Q)$. That is,

$$\sum_{q\leq Q} \Delta(x;q) = O(x^{1/2}Q)$$

barring logarithms.

GRH, however, is a deep conjecture and notoriously difficult. Number theorists assume GRH all the time to prove extremely strong results. However, the hypothesis itself is still well out of the reach of current methods.

The theorem of Enrico Bombieri and A.I. Vinogradov that we wish to discuss however establishes something similar in spirit to the above calculations *unconditionally*, that is, without using GRH in the proof in any way. With Δ as defined above, we can now state their theorem.

Theorem 1.4.2 (Bombieri-Vinogradov). *For any* $A > 0$

$$\sum_{q\leq Q} \Delta(x;q) \ll_A x(\log x)^{-A} + x^{1/2}Q(\log xQ)^4$$

Barring logarithmic factors, we can see that for $Q \leq x^{1/2}$, the error is roughly what we'd expect. In any case, we have the crude estimate $\Delta(x;q) \ll (x/q+1)\log x$. Thus,

$$\sum_{q\leq Q} \Delta(x;q) \ll \sum_{q\leq Q} \left(\frac{x\log x}{q} + \log x\right) \ll x(\log x)(\log Q) + Q\log x$$

Hence, we get

$$\sum_{q\leq Q} \Delta(x;q) \ll x(\log xQ)^2 + Q\log x$$

which is clearly stronger than our theorem for $Q > x^{1/2}$. Hence, for any proof we can assume the converse. The goal of the rest of this report is to prove this theorem while providing the background required for this proof.

Chapter 2

Dirichlet Characters

Virtually any discussion regarding multiplicative structure in arithmetic progressions must depend in some way on the concept of Dirichlet characters. In this chapter, we will introduce Dirichlet characters and associated mathematical furniture and prove some theorems about them we will be using in our exposition.

2.1 Definition

A Dirichlet character χ is an extension of a character of the multiplicative group $(\mathbb{Z}/q\mathbb{Z})^\times$ into one on the entirety of $\mathbb{Z}$.

Suppose $(G, \cdot)$ is a finite abelian group. Then a function $e : G \to \mathbb{T}$ is called a character if, for all $a, b \in G$

$$e(a \cdot b) = e(a)e(b)$$

or, in other words, e is a group homomorphism from G to $\mathbb{T}$. The character given by $e(a) = 1$ for all $a \in G$ is called the "trivial character".

Now, fix an integer q. For any character of $(\mathbb{Z}/q\mathbb{Z})^\times$, we can create a corresponding *Dirichlet character modulo q*, $\chi : \mathbb{Z} \to \mathbb{C}$ as follows:

$$\chi(n) = \begin{cases} e(n) & \textit{if } n \in (\mathbb{Z}/q\mathbb{Z})^\times \\ 0 & \textit{otherwise} \end{cases}$$

In other words, χ is supported on the integers coprime to q and is essentially the same as e at these points. The unique Dirichlet character associated with the trivial character is called the *principal character* and is denoted as χ_0. All other characters are known as *non-principal characters.*

The reader should verify that χ is completely multiplicative (ie, $\chi(mn) = \chi(m)\chi(n)$ for all integers m and n) and periodic with period q. It can be shown that, in fact, any completely multiplicative function on $\mathbb{N}$ which is periodic with minimal period q which does not vanish everywhere is actually a Dirichlet character modulo q.

We can then show the following orthogonality equation, that we will use throughout implicitly.

Theorem 2.1.1 (Orthogonality of Dirichlet Characters)**.** *For any fixed integer q, if χ and χ_1 are two Dirichlet characters modulo q, then*

$$\sum_{a \bmod q} \chi(a)\overline{\chi_1(a)} = \begin{cases} \varphi(q) & \text{if } \chi = \chi_1 \\ 0 & \text{if } \chi \neq \chi_1 \end{cases}$$

where the summation is over any complete residue class of integers modulo q. Furthermore, if χ is some Dirichlet characters modulo q and a and b are integers coprime to q, then

$$\sum_{\chi \bmod q} \chi(a)\overline{\chi(b)} = \begin{cases} \varphi(q) & \text{if } a \equiv b \pmod q \\ 0 & \text{otherwise} \end{cases}$$

We omit the proofs of the above theorem. The interested reader can find a proof in any book on analytic number theory, such as say [2] or [3].

2.2 The Twisted ψ Function

We are now in a position the twisted ψ-function, which is essentially Chebyshev's ψ-function, “twisted” by a factor of $\chi(n)$ for some Dirichlet character χ modulo q. That is, we define the summatory function $\psi(x;\chi)$ for a Dirichlet character χ as follows:

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

Now, clearly, like $\psi(x)$ and unlike $\psi(x;q,a)$, this function is a sum over all integers up to a given quantity and is not restricted at all in terms of which congruence class the integer lies in. This is thus much easier to handle in principle. This now shows the application of the orthogonality of Dirichlet characters - they can be used to "pick out" elements in a particular congruence class and convert a sum over them into one over all integers. In particular, a basic sum interchange combined with orthogonality can be used to easily establish the following identities

$$\psi(x;q,a) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \overline{\chi(a)}\psi(x;\chi)$$

$$\psi(x;\chi) = \sum_{a=1}^{q} \chi(a)\psi(x;q,a)$$

Thus information about ψ for all Dirichlet characters modulo q can be converted into information about congrunce classes modulo q, and vice-versa. In particular, the Siegel-Walfisz theorem admits the following variant.

Theorem 2.2.1. *Suppose that $A > 0$ is a fixed, real number. When $q \leq (\log x)^A$ and χ is a Dirichlet character modulo q, we have*

$$\psi(x;q) - \delta(\chi)x \ll_A x\exp(-c_1\sqrt{\log x})$$

where c_1 is an absolute positive constant.

Proof. We use directly the above mentioned identity, and apply the regular Siegel-Walfisz theorem. Fixing $A > 0$ and choosing any $q \leq (\log x)^A$, we have

$$\psi(x;q,a) = \frac{x}{\varphi(q)} + O_A\left(\exp(-c_1\sqrt{\log x})\right)$$

Thus,

$$
\begin{aligned}
\psi(x;\chi) &= \sum_{a=1}^{q} \chi(a)\psi(x;q,a) \\
&= \sum_{a=1}^{q} \left(\chi(a) \times \frac{x}{\varphi(q)} + \chi(a) \times O_A\left(\exp(-c_1\sqrt{\log x})\right)\right)
\end{aligned}
$$

Note that since each term in the sum $\sum_{a=1}^{q} \chi(a)$ is bounded by 1, thus, the sum is $\ll q \leq (\log x)^A$. To remove the dependency on A over here, note that since A is fixed, we have in fact that $(\log x)^A \ll x$. Hence, we obtain that

$$
\psi(x;\chi) = x\left(\frac{\sum_{a=1}^{q}\chi(a)}{\varphi(q)}\right) + O_A\left(x\exp(-c_1\sqrt{\log x})\right)
$$

Noting that the main term evaluates to $\delta(\chi)x$ by the orthogonality of characters, we obtain

$$
\psi(x;\chi) - \delta(\chi)x \ll_A x\exp(-c_1\sqrt{\log x})
$$

for $q \leq (\log x)^A$ as required.

□

This is the form of the Siegel-Walfisz theorem that we shall use in our proof.

2.3 Primitive and Imprimitive Dirichlet Characters

Fix a positive integer q, and consider some $q^*|q$, and a Dirichlet character χ^* modulo q^*. Consider the function χ defined as follows:

$$
\chi(n) = \begin{cases} \chi^*(n) & \text{if } (n,q) = 1 \\ 0 & \text{otherwise} \end{cases}
$$

In other words,

$$\chi(n) = \chi_0(n)\chi^*(n)$$

where χ_0 is the principal character mod q.

It is easily verified that χ is completely multiplicative, not everywhere vanishing and periodic with minimal period q, and hence a Dirichlet character.

Thus in some sense, we can "lift" from any Dirichlet character modulo q^*, a character modulo q. If q and q^* have the same prime divisors, then clearly $\chi(n) = \chi^*(n)$ for all n. In any case, χ and χ^* are remarkably similar, and differ only at points which have prime divisors common with q but not q^*. χ^* is said to *induce* χ.

We will now try to characterize the behaviour of primitive characters in contrast to impritive characters. If χ is a Dirichlet character modulo q, we say that d is a quasiperiod if $m \equiv n \pmod{d}$ and $(mn, q) = 1$ imply $\chi(m) = \chi(n)$. We can show that the least quasiperiod of any Dirichlet character divides it modulus.

Suppose $g = (d, q)$ where d is a quasiperiod of χ. We shall show that g must also be a quasiperiod of χ.

To see this, note that $g|(m-n)$ and that g is a linear combination of d and q by elementary properties of the greatest common divisor. Hence, there exist $x, y \in \mathbb{Z}$ such that

$$m - n = dx + qy$$

Thus $\chi(m) = \chi(m-qy) = \chi(n-dx) = \chi(n)$, and thus g is also a quasiperiod. Hence, if d is the least quasiperiod of χ, $(d, q) \leq d$ shall also be a quasiperiod. However this can only happen if $(d, q) = d$ or if $d|q$. Thus, the least quasiperiod of χ must divide q. This least quasiperiod is called the *conductor* of the Dirichlet character.

We now have the following theorem that connects conductors with the primitivity of a character.

Theorem 2.3.1. *Let χ be a Dirichlet character modulo q with conductor d. Then, there exists a unique primitive Dirichlet character $\chi^\star$ which is modulo d such that χ is induced by $\chi^\star$.*

Proof. We will go about this by constructing a Dirichlet character modulo d from χ which induces χ. The construction will demonstrate the uniqueness of this character.

The idea behind this proof is to suppose χ is induced by a character $\chi^\star$ modulo d, and try to determine what values $\chi^\star$ must take.

We first define, naturally, $\chi^\star(n) = 0$ if $(n, d) > 1$. Now suppose $(n, d) = 1$. Then $(n+kd, d) = 1$. Furthermore, by the Chinese Remainder Theorem, we may choose k so that, in fact, $(n+kd, q) = 1$. Now, if χ were induced by $\chi^\star$, $\chi^\star(n)$ must be the same as $\chi(n + kd)$. Hence we define $\chi^\star(n) = \chi(n + kd)$ for suitable k and note that due to the fact that d is the conductor of χ, this definition is independent of k.

It can easily be seen that $\chi^\star$ so defined is periodic and completely multiplicative, with minimal period d, and satisfies $\chi(n) = \chi^\star(n)\chi_0(n)$. Furthermore, $\chi^\star$ is primitive as the conductor of $\chi^\star$ would be a quasiperiod of χ, and thus must be $\geq d$, giving that it must be d itself. The uniqueness of this construction can be easily seen as if there is another χ^* inducing χ, then for $(n, d) > 1$, they both vanish, while for $(n, d) = 1$ we get

$$\chi^\star(n) = \chi^\star(n + kd) = \chi(n + kd) = \chi^*(n + kd) = \chi^*(n)$$

where k has been chosen appropriately as above.

□

We now show a property of primitive characters, which shall be crucial when we are treating Gauss sums.

Theorem 2.3.2. *Suppose χ is a primitive Dirichlet character modulo q. Then, for $d|q$, $d \neq q$ and every integer a, we have*

$$\sum_{\substack{n=1 \\ n\equiv a \bmod d}}^{q} \chi(n) = 0$$

Proof. Since $d|q$, $d \neq q$, and χ is primitive, we see that d cannot be a quasiperiod. Hence, there must exist m and n such that $m \equiv n \pmod{d}$, $\chi(mn) \neq 0$ and $\chi(m) \neq \chi(n)$. Thus, since m must be invertible modulo q, we can choose some c such that $(c, q) = 1$ and $cm \equiv n \pmod{q}$. This c clearly satisfies $c \equiv 1 \pmod{d}$ and $\chi(c) \neq 1$.

Now note let us write the sum in the theorem as S. Note that as k runs through a complete set of residues $\pmod{q/d}$, the numbers $n = ac + kcd$ and $n = a + kd$ run through all residues $\pmod{q}$ for which $n \equiv a \pmod{d}$. Thus, we see that

$$S = \sum_{k=1}^{q/d} \chi(ac + kcd) = \chi(c) \sum_{k=1}^{q/d} \chi(a + kd) = \chi(c)S$$

As $\chi(c) \neq 1$, clearly $S = 0$.

□

2.4 Gauss Sums and Character Sums

The Dirichlet characters can be interpreted as an orthogonal basis of the function space on $(\mathbb{Z}/q\mathbb{Z})^\times)$, with respect to a particular inner product. However, another possible orthogonal basis can be created from $e(a/q) = e^{\frac{2\pi i a}{q}}$. These are in some sense "additive" characters, where Dirichlet characters are "multiplicative". It is strikingly clear that the additive characters are much easier to handle in certain settings than multiplicative characters, and thus we wish for some medium by which we can easily translate between the two. This is done primarily by the Gauss sum

$$\tau(\chi) = \sum_{a=1}^{q} \chi(a) e(a/q)$$

This can be thought of as the above inner product applied on additive and multiplicative characters, respectively.

We shall need the following theorem about separabilitiy of Gauss sums for our investigation of character sums

Theorem 2.4.1. *If χ is a Dirichlet character modulo q and n is a positive integer such that either χ is primitive or $(n, q) = 1$, then*

$$\sum_{a \bmod q} \chi(a) e(na/q) = \overline{\chi}(n) \tau(\chi)$$

Where the sum is over any complete residue system modulo q. In particular, when χ is primitive, $|\tau(\chi)| = \sqrt{q}$.

Proof. If $(n, q) = 1$, note that

$$\chi(n) \sum_{a \bmod q} \chi(a)e(na/q) = \sum_{a \bmod q} \chi(na)e(na/q) = \tau(\chi)$$

as na runs through a complete residue class when a does so. Thus, the theorem follows. For $(n, q) > 1$, and χ primitive, $\overline{\chi}(n) = 0$, hence we must show that the sum vanishes.

Now suppose m and d are the numerator and denominator of n/q in its reduced form. Then, clearly

$$\sum_{a=1}^{q} \chi(a)e(an/q) = \sum_{h=1}^{d} e(hm/d) \sum_{\substack{a=1 \\ a \equiv h \pmod d}}^{q} \chi(a)$$

By our previous theorem, the inner sum vanishes when χ is primitive.

To establish $|\tau(\chi)| = \sqrt{q}$, take the square modulus of the proposition and sum over $1 \leq n \leq q$.

Thus, we get

$$\sum_{n=1}^{q} \left| \sum_{a=1}^{q} \chi(a)e(an/q) \right|^2 = |\tau(\chi)|^2 \sum_{n=1}^{q} |\chi(n)|^2 = \varphi(q)|\tau(\chi)|^2$$

The sum on the left can be rewritten by using the fact that $|a|^2 = a\overline{a}$ as

$$\sum_{a=1}^{q} \sum_{b=1}^{q} \chi(a)\overline{\chi}(b) \sum_{n=1}^{q} e((a-b)n/q)$$

Clearly the innermost sum is q if $a \equiv b \pmod q$, and 0 otherwise. Thus, the sum is

$$q \sum_{a=1}^{q} |\chi(a)|^2 = q\varphi(q)$$

Thus, we get $\varphi(q)|\tau(\chi)|^2 = \varphi(q)q$, which establishes the theorem.

□

One application of the Gauss sums is the following fundamental inequality about character sums which we will use in our proof. This inequality is called the Pólya-Vinogradov inequality, where Vinogradov here is I.M. Vinogradov as opposed to A.I. Vinogradov who is eponymous in this report.

Theorem 2.4.2 (Pólya-Vinogradov Inequality). *Let χ be a non-principal Dirichlet character modulo q. Then for real $x \leq q$*

$$\sum_{x<n\leq y} \chi(n) \ll q^{1/2} \log q$$

where the implicit constant is uniform in x and y.

Proof. Clearly, it is sufficient to prove that

$$\sum_{n\leq x} \chi(n) \ll q^{1/2} \log q$$

where the sum is over all positive integers up to x.

We will use the orthogonality of additive characters to get the following

$$\sum_{n\leq x} \chi(n) = \sum_{n\leq x} \sum_{m=1}^{q} \frac{\chi(m)}{q} \sum_{h=1}^{q} e\left(\frac{h(m-n)}{q}\right)$$

We can now interchange sums to get

$$\begin{aligned} \sum_{n\leq x} \chi(n) &= \frac{1}{q} \sum_{h=1}^{q} \sum_{m=1}^{q} \chi(m) e(hm/q) \sum_{n\leq x} e(-hn/q) \\ &= \frac{1}{q} \sum_{m=1}^{q-1} \overline{\chi(h)} \tau(\chi) \sum_{n\leq x} e(-hn/q) \end{aligned}$$

where the last equality follows from our previous theorem about separable Gauss sums and the orthogonality of Dirichlet characters for $h = 1$.

Now we make note of the fact that for $h \neq 0$, we can explicitly calculate the inner sum using the geometric series formula to show that the sum is in fact $\ll \|h/q\|^{-1}$. This is a specific example of the general fact that $\sum_n e(n\alpha) = O(1/\alpha)$ for sums over all integers in an interval, and with $\alpha \neq 0$. We shall use this fact again later.

Combined with our previous theorem that $|\tau(\chi)| = \sqrt{q}$, we now get

$$\sum_{n \leq x} \chi(n) \ll q^{-1/2} \sum_{h=1}^{q-1} \|h/q\|^{-1}$$

Now we see that the quantity $\|h/q\|$ is simply $|h/q|$ for $h \leq q/2$, while the remaining part of the sum can atmost contribute as much as the first part does. This means that our sum is

$$\ll q^{1/2} \sum_{h \leq q/2} \frac{1}{h} \ll q^{1/2} \log q$$

which proves the theorem for primitive characters.

To prove the imprimitive case, suppose the conductor is r, we consider the unique primitive character $\chi^\star$ modulo r that induces χ. Then by the definition of $\chi(n)$, we may replace it by $\chi^\star(n)$ whenever $(n, q) = 1$

$$\sum_{n \leq x} \chi(n) = \sum_{\substack{n \leq x \\ (n,q)=1}} \chi^\star(n) = \sum_{\substack{n \leq x \\ (n,q/r)=1}} \chi^\star(n)$$

It is difficult to deal with the condition of coprimality in the sum as it stands right now. A common trick that is used to determine when an integer is unity comes from the theory of the Moebius function μ, which we will now use to convert the condition into a sum over a new index. We have the following identity

$$\sum_{m|n} \mu(m) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise} \end{cases}$$

which can be derived by noting that both sides are multiplicative (in the sense of $f(mn) = f(m)f(n)$ for $(m, n) = 1$), and comparing them on prime powers.

We can thus replace the condition $(n, q/r) = 1$ with a sum over μ of divisors of $(n, q/r)$, which we will hope are easier to handle. Hence,

$$\sum_{n \leq x} \chi(n) = \sum_{n \leq x} \chi^\star(n) \sum_{\substack{m|n \\ m|q/r}} \mu(m)$$

Interchanging sums, replacing the parameter n by lm and using the multiplicativity of Dirichlet characters, we get

$$\sum_{n \leq x} \chi(n) = \sum_{m|q/r} \mu(m)\chi^\star(m) \sum_{l \leq x/m} \chi^\star(l)$$

We now use the theorem for primitive characters, and the fact that the modulus of $\mu(m)\chi^\star(m)$ can never exceed 1 to get

$$\sum_{n \leq x} \chi(n) \ll r^{1/2} \log r \sum_{m|q/r} 1 = d(q/r) r^{1/2} \log r$$

Noting that $d(n) \leq 2\sqrt{n}$ (since for every divisor less than $\sqrt{n}$ there is one divisor greater than $\sqrt{n}$) and that $\log r \leq \log q$, we get that

$$\sum_{n \leq x} \chi(n) \ll q^{1/2} \log q$$

establishing the theorem.

□

Chapter 3

Elementary Reductions

In this chapter, we reduce the Bombieri-Vinogradov theorem to what we call Vaughan's basic mean value theorem (BMVT).

3.1 The Mean Value Theorem

Theorem 3.1.1 (Vaughan's BMVT)**.** *Let*

$$T(x,Q) = \sum_{q \le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} \sup_{y \le x} |\psi(y;\chi)|$$

where the sum is over primitive characters modulo q. Then

$$T(x,Q) \ll \left(x + x^{5/6}Q + x^{1/2}Q^2\right)(\log xQ)^3$$

We will prove this theorem in the remaining chapters. In this chapter, we will establish that if the BMVT holds, then so does the Bombieri-Vinogradov theorem.

3.2 Reducing the Bombieri-Vinogradov Theorem

We start with the following identity for $(a,q) = 1$.

$$\psi(y;q,a) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \overline{\chi(a)}\psi(y;\chi)$$

By the definition of δ,

$$y \sum_{\chi \bmod q} \delta(\chi) = y$$

Note here that $\chi_0(a) = 1$, giving us

$$y = \sum_{\chi \bmod q} \overline{\chi(a)}\delta(\chi)y$$

Thus we have

$$\left|\psi(y;q,a) - \frac{y}{\varphi(q)}\right| \leq \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\psi(y;\chi) - \delta(\chi)y|$$

Note here that the right hand side does not depend on a but depends on y and the left hand side gives an error in a particular PNT for a particular AP. Thus to get bounds for $\sum_{q\leq Q} \Delta(x;q)$, it suffices to bound

$$\sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \sup_{y\leq x} |\psi(y;\chi) - \delta(\chi)y|$$

as any bound for that would yield the same bound for $\sum_{q\leq Q} \Delta(x;q)$. Here we have already thrown away the cancellation in the sum over characters.

We will now try to bound this sum in a manner using the BMVT. Clearly, we first need to reduce the problem to some sort of sum of over primitive characters. Pick an arbitrary character χ mod q. We know that for some $q^*|q$, there is some primitive character χ^* mod q^* that induces χ. We would want to believe that $\psi(y;\chi)$ and $\psi(y;\chi^*)$ must be close. In fact

$$\psi(y;\chi) - \psi(y;\chi^*) = \sum_{n\leq y} (\chi(n) - \chi^*(n))\Lambda(n) = \sum_{n\leq y} (\chi_0(n) - 1)\chi^*(n)\Lambda(n)$$

It is easy to see that the terms in the summation are non-zero only if n is a prime power (say p^k), and further, $p|q$ and $p \nmid q^*$. For such n, the value of the summand is $\log p$. Hence, this is

$$O\left(\sum_{p|q, p\nmid q^*} (\log p) \sum_{k\leq \frac{\log y}{\log p}} 1\right)$$

Thus

$$\psi(y;\chi) = \psi(y;\chi^*) + O\left((\log y)\sum_{p|q} 1\right)$$

Where the sum is over all primes p dividing q. Now by the fundamental theorem of arithmetic,

$$q = \prod_{p|q} p^{\alpha_p(q)}$$

for some $\alpha_p(q)$. For those p we are considering, obviously $p^{\alpha_p(q)}$, being a prime power is bigger than the least of the prime powers, viz. 2. Thus,

$$q \geq \prod_{p|q} 2 = 2^{\sum_{p|q} 1}$$

Taking logarithms, clearly $\sum_{p|q} 1 \ll \log q$, giving us that the error term when replacing χ by χ^* is $\ll (\log q)(\log y)$. Thus, we get

$$\sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{q^*|q} \sideset{}{^*}\sum_{\chi^* \bmod q^*} \left(\sup_{y\leq x} |\psi(y;\chi^*) - \delta(\chi^*)y| + O\left((\log q)(\log y)\right)\right)$$

Now, obviously,

$$\varphi(q) = \sum_{\chi \bmod q} 1 = \sum_{q^*|q} \sum_{\chi^* \bmod q^*} 1$$

Hence, the error term can be pulled out and seen to be $O(Q(\log xQ)^2)$, which on comparison to the bound in Bombieri-Vinogradov is more than tolerable. Hence, we discard this error term and examine only the main term[1].

We now interchange the order of summation, and replace q instead by q^*r and removing the * from the character variable, getting thus,

$$\sum_{q^* \leq Q} \sum_{r \leq Q/q^*} \frac{1}{\varphi(q^*r)} \sideset{}{^*}\sum_{\chi \bmod q^*} \sup_{y \leq x} |\psi(y;q) - \delta(\chi)y|$$

Now, it can easily be established that

$$\varphi(q^*r) \geq \varphi(q^*)\varphi(r)$$

by simply using the standard formula

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

and noting that the right hand side will have an extra factor of $\prod_{p|(r,q^*)}(1 - 1/p)$.

Furthermore, we also have

$$\sum_{n \leq Q} \frac{1}{\varphi(n)} \ll \log 2Q$$

which can be established by noting the identity

$$\frac{1}{\varphi(n)} = \frac{1}{n} \sum_{d|n} \frac{\mu(d)}{\varphi(d)}$$

where μ is the Moebius function, and then doing a simple sum interchange.

Thus, if we replace q^* by q for cleaner dummy variables, we see that our main term is

[1] We shall do this routinely, wherein once we show that the error in a particular estimation is drowned by the Bombieri-Vinogradov bound, we shall neglect it and focus on the remaining main term.

$$\ll \sum_{q \leq Q} \frac{\log 2Q}{\varphi(q)} \sum_{\chi \bmod q}^{*} \sup_{y \leq x} |\psi(y;\chi) - \delta(\chi) y|$$

At this point, we see that the summands bear some relation to expression in the variant of Siegel-Walfisz that we have proven. Thus, we could probably deal with the smaller terms (those $\leq (\log x)^A$ for some A) using Siegel-Walfisz. Indeed, let $R = (\log x)^{6+A}$. Then,

$$\sum_{q \leq R} \frac{\log 2Q}{\varphi(q)} \sum_{\chi \bmod q}^{*} \sup_{y \leq x} |\psi(y;\chi) - \delta(\chi) y| \ll_A (\log x) R x \exp(-c_2 \sqrt{\log x})$$

where we have implicitly used the fact that a sum over primitive characters of 1 may not exceed $\varphi(q)$, and that $Q \ll x$.

This part of the sum thus gives us a contribution that is

$$\ll x (\log x)^{-A}$$

which is acceptable. For the remaining part of the sum, clearly $q > 1$, and hence a primitive conductor χ modulo q is non-principal. In other words, $\delta(\chi) = 0$. It remains thus, to deal with

$$(\log 2Q) \sum_{R < q \leq Q} \frac{1}{\varphi(q)} \sum_{\chi \bmod q}^{*} \sup_{y \leq x} |\psi(y;\chi)|$$

It should now be obvious that an appeal to the BMVT, and a suitable partial summation argument would give us our desired conclusion. In fact, multiplying the numerator and denominator of the summand by q, and summing partially, we get

$$\begin{aligned}(\log 2Q) \sum_{R < q \leq Q} \frac{1}{\varphi(q)} \sum_{\chi \bmod q}^{*} \sup_{y \leq x} |\psi(y;\chi)| \quad &= \quad \frac{T(x,Q) - T(x,R)}{Q} \\ &\quad + \int_R^Q \frac{T(x,t) - T(x,R)}{t^2} dt\end{aligned}$$

here we have used the same notation $T(x,t)$ as in the BMVT.

Clearly, our sum is thus

$$\leq (\log 2Q)Q^{-1}T(x,Q) + \int_R^Q t^{-2}T(x,t)dt$$

Using $\log 2Q \ll \log x$ and the BMVT, we see this is

$$\ll Q^{-1}\left(x + x^{5/6}Q + x^{1/2}Q^2\right)(\log x)^4 + \int_R^Q t^{-2}\left(x + x^{5/6}t + x^{1/2}t^2\right)(\log x)^4 dt$$

$$\ll \left(xR^{-1} + x^{5/6}\log(2Q/R) + x^{1/2}Q\right)(\log x)^4$$

$$\ll x(\log x)^{-A} + x^{1/2}Q(\log x)^4$$

as required.

We have thus shown that it is sufficient to prove Vaughan's BMVT to establish the Bombieri-Vinogradov theorem. In the remaining part of this report we will examine the theory behind this inequality, and prove it.

Chapter 4

The Large Sieve

The Large Sieve is an analytic principle first invented by Linnik in the 1940s in his work on the least quadratic non-residue modulo a prime p. With time and exposition, the general principle and the usefulness of Linnik's large sieve is much better understood now, and it is a crucial ingredient in the proof of the BMVT. In this chapter, we shall exposit the basic idea behind the large sieve, we will establish a particular example of what is called a "Large Sieve Inequality" and enunciate a stronger inequality from which, by using Gauss sums, we can extract a general inequality for characters that will help establish the BMVT.

In general, let a_n be some sequence, and fix N. Further, let

$$S(\chi) = \sum_{n=M+1}^{M+N} a_n \chi(n)$$

We see that the summand of the BMVT is a term of this kind with a_n as the von Mangoldt function. Then a general bound of the form

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |S(\chi)|^2 \leq \lambda(N, Q) \sum_{n=M+1}^{M+N} |a_n|^2$$

would roughly mean that if we could control suitable sums of von Mangoldt functions, we can control the sum in the BMVT. Our goal in this chapter shall be to prove such a bound.

4.1 The Analytical Principle of the Large Sieve

For any positive integers N, M and Q, and any sequence of real numbers $(a_n)_{n=1}^{\infty}$, suppose

$$S(\alpha) = \sum_{n=M+1}^{M+N} a_n e(n\alpha)$$

Then we seek some $\lambda(N, Q)$ (which can easily be seen to be uniform in M) such that a bound of the following form can be established

$$\sum_{q \le Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| S\left(\frac{a}{q}\right) \right|^2 \le \lambda(N, Q) \sum_{n=M+1}^{M+N} |a_n|^2$$

Any inequality of this form is called a "Large Sieve Inequality", and such a $\lambda(N, Q)$ clearly exists by applying Cauchy-Schwarz inequality on the left hand side. A little observation shows that we can rewrite the inequality in the following manner:

$$\sum_{\alpha \in \mathcal{F}_Q} |S(\alpha)|^2 \le \lambda(N, Q) \sum_{n=M+1}^{M+N} |a_n|^2$$

where $\mathcal{F}_Q$ is the sequence of Farey fractions[1] with denominator at most Q. This way of writing this equation suggests that such an inequality could be established for sets with more arbitrary structure than $\mathcal{F}_Q$. In particular, note that $S(\alpha)$ is clearly periodic with period 1. Furthermore, the intuition is that if too many α cluster around the same point, then those $|S(\alpha)|^2$ could possibly pick an a_n to be close enough to that the contribution of that term becomes very big, and thus cannot be bounded independently of the various α. Thus, any arbitrary set that we replace $\mathcal{F}_Q$ with must be equidistributed in some sense with respect to the metric on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, that is

$$\|\alpha\| = \sup_{n \in \mathbb{Z}} |\alpha - n|$$

[1] See Appendix

which is the distance of α from its nearest integer.

Thus, we replace the Farey fractions with an arbitrary set of real number $\mathcal{F}_\delta$ such that for distinct $x, y \in \mathcal{F}_\delta$ we have $\|x - y\| \geq \delta$. That is, any set which is equidistributed in $\|\cdot\|$ with level δ. We then look instead for a $\lambda_0(N, \delta)$ such that

$$\sum_{\alpha \in \mathcal{F}_\delta} |S(\alpha)|^2 \leq \lambda_0(N, \delta) \sum_{n=M+1}^{M+N} |a_n|^2$$

It is easily established that the Farey fractions $\mathcal{F}_Q$ are equidistributed with level $\delta = 1/Q^2$. Thus if any $\lambda_0(N, \delta)$ works, we can take $\lambda(N, Q) = \lambda_0(N, Q^{-2})$. Nowadays, even inequalities of the $\lambda_0(N, \delta)$ variety are called Large Sieve Inequalities. Thus the principle of the large sieve can be extracted to an analytical principle with only an equidistribution requirement on the set being considered. It was shown in 1991 that $\lambda_0(N, \delta) = N - 1 + \delta^{-1}$ works and that this is the best possible bound that can be obtained, even when only considering the Farey fractions. Thus, $\lambda(N, Q) = N - 1 + Q^2$ works. For the purpose of this proof, we will only use the much easier fact that $\lambda(N, Q) \ll N + Q^2$. In the next section, we shall provide an elementary proof that $\lambda(N, Q) = N + Q^2 \log 3Q^2$ suffices.

4.2 A Large Sieve Inequality

We now establish that

$$\lambda_0(N, \delta) = N + \frac{1}{\delta} \log \frac{3}{\delta}$$

works, which incidentally also establishes that

$$\lambda(N, Q) = N + Q^2 \log 3Q^2$$

also works.

Before we move on, we establish first the following useful theorem.

Theorem 4.2.1 (Duality Lemma)**.** *Suppose that c_{nr} are complex numbers for $n = 1, \cdots, N$ and $r = 1, \cdots, R$ and λ is a real number such that for all complex numbers z_r, we have*

$$\sum_{n=1}^{N} \left| \sum_{r=1}^{R} c_{nr} z_r \right|^2 \leq \lambda \sum_{r=1}^{R} |z_r|^2$$

then

$$\sum_{r=1}^{R} \left| \sum_{n=1}^{N} c_{nr} w_n \right|^2 \leq \lambda \sum_{n=1}^{N} |w_n|^2$$

holds for all complex numbers w_n.

Proof. Let LHS be the left hand side of the inequality we wish to establish. Now, using the fact that $|a|^2 = a\overline{a}$, and changing a few dummy variables, we have that

$$LHS = \sum_{m=1}^{N} w_m \sum_{r=1}^{R} c_{mr} \sum_{n=1}^{N} \overline{c_{nr} w_n}$$

Hence, by Cauchy-Schwarz, if we take $z_r = \sum_{n=1}^{N} c_{nr} w_n$

$$LHS^2 \leq \left(\sum_{m=1}^{N} |w_m|^2 \right) \sum_{m=1}^{N} \left| \sum_{r=1}^{R} c_{mr} \overline{z}_r \right|^2$$

which by the hypothesis does not exceed

$$\sum_{m=1}^{N} |w_m|^2 \lambda \sum_{r=1}^{R} |z_r|^2 = (LHS) \lambda \sum_{m=1}^{N} |w_m|^2$$

which establishes the theorem. □

The Duality Lemma may be interpreted as a statement from Linear Algebra about Hermitian matrices, but we shall not use this interpretation in any significant way.

We now establish our inequality

Theorem 4.2.2 (A Large Sieve). *Suppose that $0 < \delta \leq 1/2$, and that $\mathcal{F}$ is a set of real numbers for which $x, y \in \mathcal{F}$ and $x \neq y$ imply $\|x - y\| \geq \delta$. Further suppose that $(a_n)_{n=1}^{\infty}$ and $S(\alpha)$ are as defined above. Then*

$$\sum_{\alpha \in \mathcal{F}} |S(\alpha)|^2 \leq \lambda_0(N, \delta) \sum_{n=M+1}^{M+N} |a_n|^2$$

holds with

$$\lambda_0(N, \delta) = N + \frac{1}{\delta} \log \frac{3}{\delta}$$

Proof. Let $\mathcal{F} = \{x_1, x_2, \cdots, x_R\}$. Then, by the Duality Lemma it suffices to bound

$$\sum_{n=M+1}^{M+N} \left| \sum_{r=1}^{R} b_r e(nx_r) \right|^2 = \sum_{r=1}^{R} \sum_{s=1}^{R} b_r \overline{b_s} \sum_{n=M+1}^{M+N} e(n(x_r - x_s)$$

The diagonal terms ($s = r$) and non-diagonal terms ($s \neq r$) will clearly be different types of terms due to the cancellation or lack thereof from the trignometric sums. In particular, the $r = s$ terms contribute $N \sum_{r=1}^{R} |b_r|$.

For $r \neq s$, as previously noted, we can show that

$$\sum_{n=M+1}^{M+N} e(n(x_r - x_s)) \leq \frac{1}{2\|x_r - x_s\|}$$

Hence, from the diagonal terms we get at most

$$\sum_{r=1}^{R} \sum_{\substack{s=1 \\ s \neq r}}^{R} \frac{1}{2}(|b_r|^2 + |b_s|^2) \frac{1}{2\|x_r - x_s\|} = \sum_{r=1}^{R} |b_r|^2 \sum_{\substack{s=1 \\ s \neq r}}^{R} \frac{1}{2\|x_r - x_s\|}$$

Now we note that

$$\sum_{\substack{s=1 \\ s \neq r}}^{R} \frac{1}{2\|x_r - x_s\|} \leq 2 \sum_{k \leq 1/\delta} \frac{1}{2k\delta} \leq \frac{1}{\delta} \log \frac{3}{\delta}$$

establishing the theorem.

$\square$

Clearly this theorem also implies the large sieve inequality for Farey fractions with $\lambda(N, Q) = N + Q^2 \log 3Q^2$.

In the rest of this report, we could use this form of the Large Sieve to prove the Bombieri-Vinogradov theorem with a slightly inflated logarithmic factor, which is normally adequate for most applications of the theorem. However, we will use without proof the slightly stronger estimate that, in fact,

$$\lambda(N, Q) \ll N + Q^2$$

works for some appropriate implicit constant.

4.3 A Large Sieve Inequality for Characters

Assuming the stronger inequality of the previous section, we are now in a position to establish the type of bound we loosely described at the beginning of this chapter. We wish to go now from an inequality for additive characters to one for multiplicative characters. This shall be done by going through the useful tool of the Gauss sums. We have, thus, the following theorem:

Theorem 4.3.1 (Large Sieve for Characters)**.** *Suppose that for a Dirichlet character χ,*

$$S(\chi) = \sum_{n=M+1}^{M+N} a_n \chi(n)$$

Then,

$$\sum_{q \le Q} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^{*} |S(\chi)|^2 \le \lambda(N, Q) \sum_{n=M+1}^{M+N} |a_n|^2$$

holds with $\lambda(N, Q) \ll N + Q^2$.

Proof. Suppose χ is a primitive character. Then, by the separability of Gauss sums,

$$\chi(n) = \frac{1}{\tau(\overline{\chi})} \sum_{a=1}^{q} \overline{\chi}(a) e(na/q)$$

Thus, substituting this in $S(\chi)$,

$$S(\chi) = \sum_{n=M+1}^{M+N} \frac{a_n}{\tau(\overline{\chi})} \sum_{a=1}^{q} \overline{\chi}(a) e(na/q) = \frac{1}{\tau(\overline{\chi})} \sum_{a=1}^{q} \overline{\chi}(a) S(a/q)$$

by interchanging the sums and appealing to the definition of $S(\alpha)$. Hence,

$$\sum_{\chi \bmod q}{}^{*} |S(\chi)|^2 = \frac{1}{|\tau(\overline{\chi})|^2} \sum_{\chi \bmod q}{}^{*} \left| \sum_{a=1}^{q} \overline{\chi}(a) S(a/q) \right|^2$$

We have $|\tau(\overline{\chi})|^2 = q$, and the sum can be extended over all characters to get

$$\sum_{\chi \bmod q}{}^{*} |S(\chi)|^2 \leq \frac{1}{q} \sum_{\chi \bmod q} \left| \sum_{a=1}^{q} \overline{\chi}(a) S(a/q) \right|^2$$

We shall now show that the left hand side of this ineqality is

$$\frac{\varphi(q)}{q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| S\left(\frac{a}{q}\right) \right|^2$$

which combined for our large sieve inequality for exponential sums immediately established the theorem.

To see this, note that

$$\sum_{\chi \bmod q} \left| \sum_{a=1}^{q} \overline{\chi}(a) S(a/q) \right|^2 = \sum_{\chi \bmod q} \sum_{a=1}^{q} \sum_{b=1}^{q} \overline{\chi}(a) \chi(b) S(a/q) \overline{S(b/q)}$$

Interchanging the sums, we have

$$\sum_{a=1}^{q}\sum_{b=1}^{q} S(a/q)\overline{S(b/q)} \sum_{\chi \bmod q} \overline{\chi}(a)\chi(b)$$

Now, using the orthogonality of Dirichlet characters, we see that the inner sum is 0 if $(a, q) > 1$ or $a \neq b$ and $\varphi(q)$ otherwise. Thus, we get the sum

$$\varphi(q) \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| S\left(\frac{a}{q}\right)\right|^2$$

which is as we wanted. This is the variant of Parseval's identity in this setting, and establishes our theorem. □

For our final application, we will actually use the following tweaked inequality which we can derive from our large sieve inequality for characters.

Theorem 4.3.2 (Tweaked Inequality). *Let $x \geq 2$, and $a_1, \cdots, a_M$ and $b_1, \cdots, b_N$ be complex numbers. Then,*

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}{}^{*} \sum_{y \leq x} \left| \sum_{\substack{m=1 \\ mn \leq y}}^{M} \sum_{n=1}^{N} a_m b_n \chi(mn) \right|^2$$
$$\ll (\log xMN) \sqrt{(M+Q^2)(N+Q^2) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2}$$

Proof. By the Cauchy-Schwarz inequality and our large sieve for characters, we have

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi}{}^{*} \left| \sum_{m=1}^{M} \sum_{n=1}^{N} a_m b_n \chi(mn) \right| \ll \sqrt{(M+Q^2)(N+Q^2) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2}$$

We now insert a maximality condition in this. Let

$$C = \int_{-\infty}^{\infty} \frac{\sin \alpha}{\alpha} d\alpha$$

Let $\gamma > 0$ so that we define

$$\epsilon(\beta) = \begin{cases} 1 & \text{if } 0 \leq \beta \leq \gamma \\ 0 & \text{if } \beta > \gamma \end{cases}$$

Now it can be seen that

$$\int_{-\infty}^{\infty} e^{i\beta\alpha} \frac{\sin \gamma\alpha}{C\alpha} d\alpha = \epsilon(\beta)$$

Since $\int_A^\infty \frac{\sin \lambda\alpha}{\alpha} d\alpha \ll 1/\lambda A$,

$$\epsilon(\beta) = \int_A^{-A} e^{i\beta\alpha} \frac{\sin \gamma\alpha}{C\alpha} d\alpha + O\left(\frac{1}{A|\gamma - \beta|}\right)$$

Putting $\gamma = \log(\lfloor y \rfloor + \frac{1}{2})$ and $\beta = \log mn$, we get

$$\epsilon(\log mn) = \begin{cases} 1 & \text{if } mn \leq y \\ 0 & \text{if } mn > y \end{cases}$$

Thus,

$$\epsilon(\log mn) = \int_{-A}^{A} (mn)^{i\alpha} \frac{\sin \gamma\alpha}{C\alpha} d\alpha + O\left(\frac{y}{A}\right)$$

Now, we can put this condition into the sum to get

$$\sum_{\substack{m=1 \\ mn \leq y}}^{M} \sum_{n=1}^{N} a_m b_n \chi(mn) = \sum_{m=1}^{M} \sum_{n=1}^{N} a_n b_m \chi(mn) \epsilon(\log mn)$$

$$= \int_{-A}^{A} \sum_{m=1}^{M} \sum_{n=1}^{N} a_m m^{i\alpha} b_n n^{i\alpha} \chi(mn) \frac{\sin(\gamma\alpha)}{C\alpha} d\alpha + O\left(\frac{y}{A} \sum_{m=1}^{M} \sum_{n=1}^{N} |a_m b_n|\right)$$

The error term here is manageable for $A = xMN$. Putting that in, and assuming $y \leq x$, the integral is

$$\ll \int_{-A}^{A} \left| \sum_{m=1}^{M} \sum_{n=1}^{N} a_m m^{i\alpha} b_n n^{i\alpha} \chi(mn) \right| \min\left(\log x, \frac{1}{|\alpha|} \right) d\alpha$$

The part of the integrand independent of α is clearly

$$\ll \sqrt{(M+Q^2)(N+Q^2) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2}$$

while

$$\int_{-A}^{A} \min\left(\log x, \frac{1}{|\alpha|} \right) d\alpha \ll \log xMN$$

proving the theorem.

□

Chapter 5

Estimates on the von Mangoldt Function

We are now in a position to establish the BMVT. The important idea here is what we call Vaughan's Identity which provides us with a way to separate sums involving the von Mangoldt function into 4 bilinear forms each of which we will later be able to deal with. The identity itself is a trivial equality of Dirichlet series, and hence equality of the coefficients of the corresponding series.

5.1 Vaughan's Identity

The identity is the following

Theorem 5.1.1 (Vaughan's Identity). *Suppose that u, v are positive reals and $y \geq 2$. Further suppose f is a complex-valued number theoretic function with finite support (ie, the number of points such that $f(n) \neq 0$ are finite). Then,*

$$\sum_n \Lambda(n) f(n) = S_1 - S_2 + S_3 - S_4$$

where

$$S_1 = \sum_{m \le u} \mu(m) \sum_{n \le y/m} (\log n) f(mn)$$

$$S_2 = \sum_{m \le uv} c_m \sum_{n \le y/m} f(mn) \text{ where } c_m = \sum_{\substack{k \le u, l \le v \\ kl = m}} \Lambda(k)\mu(l)$$

$$S_3 = \sum_{\substack{m > u \\ mn \le y}} \sum_{n > v} \left(\sum_{\substack{k|m \\ k > u}} \Lambda(k) \right) \mu(n) f(mn)$$

$$S_4 = \sum_{n \le v} \Lambda(n) f(n)$$

Proof. For any formal Dirichlet series $F(s)$ and $G(s)$, we have the identity

$$-\frac{\zeta'}{\zeta}(s) = G(s)(-\zeta'(s)) - F(s)G(s)\zeta(s) - (-\zeta'(s) - F(s)\zeta(s)) \left(G(s) - \frac{1}{\zeta(s)} \right) + F(s)$$

Now, we specialize with,

$$F(s) = \sum_{n \le u} \frac{\Lambda(n)}{n^s}$$

$$G(s) = \sum_{n \le v} \frac{\mu(n)}{n^s}$$

Let $\Lambda_j(n)$ be the coefficient of n^{-s} in the Dirichlet series of the jth term in the identity. Then,

$$\Lambda(n) = \Lambda_1(n) - \Lambda_2(n) + \Lambda_3(n) - \Lambda_4(n)$$

Using our knowledge of Dirichlet series, it's now very easy to see that

$$S_j = \sum_n \Lambda_j(n) f(n)$$

which proves our theorem.

□

Thus, using this theorem, we can convert sums of the type $\sum_n \Lambda(n) f(n)$ into bilinear forms. The expectation is that the contribution from S_4 shall be small, the smoothness of the coefficients of S_1 and S_2 will give control over them, while S_3 can be hopefully be bounded by some other methods.

5.2 Proof of the BMVT

We now return to our proof of the BMVT using Vaughan's identity and our tweaked inequality. In this portion of the exposition, we will follow [1] almost exactly.

We wish to show that if

$$T(x,Q) = \sum_{q \le Q} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^{*} \sup_{y \le x} |\psi(y;\chi)$$

then

$$T(x,Q) \ll \left(x + x^{5/6}Q + x^{1/2}Q^2\right) (\log xQ)^3$$

Now suppose $Q^2 > x$. By the tweaked inequality with $M = 1$, $a_1 = 1$, $N = \lfloor x \rfloor$ and $b_n = \Lambda(n)$, we get the bound

$$Q^2 (\log Q)^2 \sqrt{\sum_{n \le x} \Lambda(n)^2} \ll x^{1/2} Q^2 (\log xQ)^3$$

which is acceptable. We can thus suppose that $Q^2 \le x$.

Let $u = v = \min\left(Q^2, x^{1/3}, xQ^{-2}\right)$. Then using the same idea in the tweaked inequality, if we restrict the supremum to $y \le u^2$,

$$\sum_{q \le Q} \frac{q}{\varphi(q)} \sum_{\chi}^{*} \sup_{y \le u^2} |\psi(y;\chi)| \ll (u^2Q + uQ^2)(\log x)^3 \ll (x^{2/3}Q + x^{1/3}Q^2)(\log x)^3$$

which is again acceptable. Thus, it suffices to consider

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\sup_{u^2<y\le x}|\psi(y;\chi)|$$

We can now use Vaughan's identity with our chosen values for u, v and with

$$f(n)=\begin{cases}\chi(n) & \text{if } n\le y\\ 0 & \text{otherwise}\end{cases}$$

to separate $\psi(y;\chi)$ sum into four bilinear sums $S_j(\chi)$, $j=1,2,3,4$, whereby it suffices to bound instead

$$T_j=\sum_{q\le Q}\frac{q}{\varphi(q)}\sum_{\chi}{}^{*}\sup_{u^2<y\le x}|S_j(\chi)|$$

The case $j=4$ is easy and can be dealt with by appealing to the tweaked inequality.

For the case $j=1$,

$$\begin{aligned} S_1(\chi) &= \textstyle\sum_{m\le u}\mu(m)\chi(m)\sum_{n\le y/m}\chi(n)\int_1^n \frac{dt}{t}\\ &= \textstyle\int_1^y \sum_{m\le \min(u,y/t)}\mu(m)\chi(m)\sum_{t<n\le y/m}\chi(n)\frac{dt}{t}\\ &\ll \textstyle\int_1^y uq^{1/2}\log q\frac{dt}{t}\\ &\ll uq^{1/2}(\log q)(\log x)\end{aligned}$$

where the inequality holds for $q>1$. For $q=1$, we get an bound of $x(\log x)^2$, whereby

$$T_1\ll (x+uQ^{5/2})(\log xQ)^2\ll (x+x^{1/2}Q^2)(\log xQ)^2$$

on examining all the possible values of u.

For $j=3$, we consider

$$\mathcal{M}=\{2^k\lfloor u\rfloor : k=0,1,\cdots;2^k\lfloor u\rfloor\le x/u\}$$

Note that $|\mathcal{M}| \ll \log x$. Now, if

$$S_3(\chi; M) = \sum_{M<m\leq 2M} \sum_{\substack{u<n\leq x/M \\ mn\leq y}} \left(\sum_{\substack{k|m \\ k>u}} \Lambda(k) \right) \mu(n)\chi(mn)$$

then,

$$S_3(\chi) \ll \sum_{M\in\mathcal{M}} |S_3(\chi; M)|$$

We thus have

$$T_3 \ll \sum_{M\in\mathcal{M}} T_3(M)$$

where

$$T_3(M) = \sum_{q\leq Q} \frac{q}{\varphi(q)} \sum_{\chi}{}^{*} \sup_{u^2<y\leq x} |S_3(\chi; M)|$$

We now use the tweaked inequality to obtain

$$T_3(M) \ll (\log x)\sqrt{(M+Q^2)(x/M+Q^2) \sum_{m\leq 2M} (\log m)^2 \sum_{n\leq x/M} \mu(n)^2}$$

Using the easy facts that $\sum_{m\leq z}(\log m)^2 \ll z(\log 2z)^2$ and $\sum_{m\leq z}\mu(m)^2 \ll z$,

$$T_3(M) \ll (\log x)^2 \left(x + xM^{-1/2}Q + x^{1/2}M^{1/2}Q + x^{1/2}Q^2 \right)$$

Summing over elements of $\mathcal{M}$ will give

$$T_3 \ll (\log x)^3 \left(x + xu^{-1/2}Q + x^{1/2}Q^2 \right)$$

Thus,

$$T_3 \ll (\log x)^3 \left(x + x^{5/6}Q + x^{1/2}Q^2\right)$$

The final case, $j = 2$ can be dealt with by separating $S_2(\chi)$ as follows

$$S_2(\chi) = S_2'(\chi) + S_2''(\chi)$$

with S_2' with terms having $m \leq u$ and the S_2'' with terms having $u < m \leq u^2$. The corresponding sums T_2' and T_2'' can then be dealt with in a manner similar to T_1 and T_3 respectively.

This establishes the BMVT, and hence the Bombieri-Vinogradov theorem.

Appendix A

Appendix

A.1 Partial Summation

We now give an account for partial summation. In general, summation by parts is an identity similar to integration by parts, which relates the sum of the product of two functions with the sum of one function, and the difference of the other. However, for our purposes we shall need the following, much weaker version called partial summation.

Theorem A.1.1 (Partial Summation)**.** *Suppose $a_1, a_2, a_3 \cdots$ is a sequence of complex numbers, $A(x) = \sum_{n \leq x} a_n$ and $f(x)$ is some differentiable function on $(1, \infty)$. Then*

$$\sum_{n \leq x} a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)dt$$

Proof. Suppose x is a natural number. Therefore,

$$\begin{aligned} \sum_{n \leq x} a_n f(n) &= \sum_{n \leq x} \{A(n) - A(n-1)\} f(n) \\ &= A(x)f(x) - \sum_{n \leq x-1} A(n)\{f(n+1) - f(n)\} \end{aligned}$$

Now, using the fact that f is differentiable,

$$\sum_{n\leq x} a_n f(n) = A(x)f(x) - \sum_{n\leq x-1} A(n)\int_n^{n+1} f'(t)dt$$

Now, $A(x)$ is a step function changing values at positive integers. Hence $A(n)$ can be taken inside and replaced by $A(t)$.

$$\sum_{n\leq x} a_n f(n) = A(x)f(x) - \sum_{n\leq x-1} \int_n^{n+1} A(t)f'(t)dt$$

and thus

$$\sum_{n\leq x} a_n f(n) = A(x)f(x) - \int_1^{\infty} A(t)f'(t)dt$$

proving our theorem for integers. For non-integers, note that the theorem holds for $\lfloor x \rfloor$, the greatest integer less than x, and that

$$A(x)\{f(x) - f(\lfloor x \rfloor)\} - \int_{\lfloor x \rfloor}^{x} A(t)f'(t)dt = 0$$

which establishes the theorem.

□

This identity is a powerful tool for obtaining elementary estimates for many sums that arise in number theory, and is used in this report often without any specific appeal. We now use it to prove a fact we have assumed as common knowledge in the report.

Theorem A.1.2. *For $x > 0$,*

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

Proof. Putting $a_n = 1$ and $f(t) = 1/t$ in the partial summation identity, we see that $A(x) = \sum_{n\leq x} 1 = \lfloor x \rfloor$ and thus,

$$\sum_{n\le x}\frac{1}{x}=\frac{\lfloor x\rfloor}{x}+\int_1^x\frac{\lfloor t\rfloor}{t^2}dt$$

Now, using $\lfloor x\rfloor = x-\{x\}=x+O(1)$,

$$\sum_{n\le x}\frac{1}{x}=\frac{x+O(1)}{x}+\int_1^x\frac{t+O(1)}{t^2}dt$$

Thus, the first term is clearly $O(1)$. Furthermore, the error term in the integral evaluates to

$$\int_1^x\frac{dt}{t^2}=1-\frac{1}{x}$$

which contributes $O(1)$. Hence, the main term is

$$\int_1^x\frac{dt}{t}=\log x+O(1)$$

Hence our claim follows.

□

We now turn to the identities relating $\pi(x)$ and $\vartheta(x)$.

Theorem A.1.3. *For any real number $x>0$, we have*

$$\pi(x)=\frac{\vartheta(x)}{\log x}+\int_2^x\frac{\vartheta(t)}{t(\log t)^2}dt$$

$$\vartheta(x)=\pi(x)\log x-\int_2^x\frac{\pi(t)}{t}dt$$

Furthermore, these relations hold with $\pi(x)$ replaced by $\pi(x;q,a)$ and $\vartheta(x)$ replaced by $\vartheta(x;q,a)$.

Proof. Putting $a_n = 1_{\mathcal{P}}(n)\log n$ and $f(t)=1/\log t$, we get $A(x)=\vartheta(x)$ and $f'(t)=-1/t(\log t)^2$ in the partial summation identity. Furthermore, the left hand side become $\pi(x)$, giving us the first identity.

Similarly, putting $a_n = 1_{\mathcal{P}}(n)$ and $f(t) = \log t$, we get $A(x) = \pi(x)$ and $f'(t) = 1/t$. Furthermore the left hand side becomes $\vartheta(x)$, giving us the second identity.

In either identity if we replace $1_{\mathcal{P}}(n)$ with $1_{\mathcal{P}(a,q)}(n)$ where

$$\mathcal{P}(a,q) = \{p \in \mathcal{P} : p \equiv a \pmod q\}$$

then we obtain the relations between $\pi(x; q, a)$ and $\vartheta(x; q, a)$, analogously.

□

A.2 Farey Fractions

The Farey fractions are sequences of rational numbers which are recurring in number theory. For a fixed positive integer Q, the sequence of Farey fractions, $\mathcal{F}_Q$ is the set of all rational numbers in $[0,1]$ whose denominator in reduced form is $\leq Q$, ordered with respect to the natural ordering on the reals.

Thus, for example, $\mathcal{F}_1$ is

$$\frac{0}{1}, \frac{1}{1}$$

$\mathcal{F}_2$ is

$$\frac{0}{1}, \frac{1}{2}, \frac{1}{1}$$

$\mathcal{F}_3$ is

$$\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}$$

and so on.

There are many interesting facts about the Farey fractions, however we only need one fact for the report, which is the equidistribution of the fractions with respect to the norm $\|\cdot\|$ on $\mathbb{T}$ with level $1/Q^2$.

To see this, suppose $\alpha, \beta \in \mathcal{F}_Q$ are distinct. Without loss of generality, suppose that $\alpha > \beta$. Then,

$$\alpha = \frac{a}{b}$$

$$\beta = \frac{c}{d}$$

where $ad - bc \neq 0$. Hence,

$$|\alpha - \beta| = \left|\frac{a}{b} - \frac{c}{d}\right| = \frac{ad - bc}{bd}$$

Now note that $b, d \leq Q$, and that $ad - bc > 0$, and thus is ≥ 1. Hence,

$$|\alpha - \beta| \geq \frac{1}{Q^2}$$

Thus, since

$$\|\alpha - \beta\| = \min_{n \in \mathbb{Z}} |\alpha - \beta - n|$$

clearly

$$\|\alpha - \beta\| = |\alpha - \beta| \geq \frac{1}{Q^2}$$

proving that $\mathcal{F}_Q$ is equidistributed.

Bibliography

[1] R. C. Vaughan, **The Bombieri-Vinogradov Theorem**, `http://www.personal.psu.edu/rcv4/Bombieri.pdf`

[2] Tom M. Apostol, *Introduction to Analytic Number Theory* (Undergaduate Texts in Mathematics), Springer-Verlag, New York, NY, (1976).

[3] M. Ram Murty, *Problems in Analytic Number Theory* (Graduate Texts in Mathematics), Springer-Verlag, New York, NY, (2001).