LARGE SIEVE AND BOMBIERI–VINOGRADOV THEOREM

NICK HARLAND

1. Introduction

Let $\pi(x;q;a)$ be the number of primes in an arithmetic progression up to x. More specifically, it is the cardinality of the set $\{p \mid p \text{ prime}, p \equiv a \pmod q, p \leq x\}$. Notationally we can write this as

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

where it's understood that for the purposes of this paper that p will always denote a prime. A similar function is

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

where

$$\Lambda(n) = \begin{cases} \log p & n = p^k \\ 0 & \text{otherwise.} \end{cases}$$

Facts about the number of primes in an arithmetic progression can often be more easily studied by looking at the function $\psi(x;q;a)$ rather than $\pi(x;q;a)$, and we can use Euler Summation to go from the latter to the former. The prime number theorem for arithmetic progressions asserts that for $(a,q) = 1$ and $q \leq \log^A x$ we have (see [6, Corollary 11.19])

$$\psi(x;q;a) = \frac{x}{\phi(q)} + O_A\bigg(x \log^{-A} x\bigg)$$

for any constant A. The error term, however is quite far from what we would expect. For example under the Generalized Riemann Hypothesis, we get that for $(a,q) = 1$, (see [6, Corollary 13.8])

$$\psi(x;q;a) = \frac{x}{\phi(q)} + O\bigg(\sqrt{x} \log^2 x\bigg). \tag{1}$$

The Bombieri–Vinogradov Theorem is an unconditional result which says that while for any specific q, the error term might be large, on average over a certain range, the error term resembles that of (1). That is over a sum of $q \leq Q$, we should get that the error is $Qx^{1/2}log^2x$. We will get something close to that by proving the following theorem

Theorem 1. *[Bombieri–Vinogradov] Let $A > 0$ be fixed, and let $E(x;q,a) = \psi(x;q;a) - \frac{x}{\phi(q)}$. Then*

$$\sum_{q \leq Q} \max_{(a,q)=1} |E(x;q,a)| \ll_A x \log^{-A} x, \tag{2}$$

where $Q = x^{1/2} \log^{-B} x$, for $B = 2A + 8$.

Most proofs of the theorem involve the Large Sieve which we'll start by discussing below.

2. The Additive Large Sieve

A sieve in number theory is one meant to take a set S of natural numbers, and "sift out" all the numbers lying in another set Ω_p modulo primes p. For example the Sieve of Eratosthenes eliminates all numbers congruent to 0 modulo p except p itself. What's left are the primes.

Recall the notation

$$e(z) = e^{2\pi i z}$$

which will be used in the following corollary as well as throughout the rest of the paper. The following is taken from [4, Chapter 7] We'll start by proving the following, called the additive large sieve.

Theorem 2. *Let a_n be complex numbers, M and N be natural numbers with $n \in \{M+1, ..., M+N\}$. Then we have the following*

$$\sum_{q \le Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| \sum_{n=M+1}^{M+N} a_n e\left(\frac{an}{q}\right) \right| \le (Q^2 + N - 1)\|\vec{a}\|^2, \tag{3}$$

where $\|\vec{a}\| = \sum_i |a_i|^2$.

The idea behind the proof is that we want to show if numbers are not too close together, then the sum in (3) can't get too large. Then we use that the rational numbers are spaced out well enough. Note that the reason we call it the "additive" large sieve is the use of the additive characters in the sum. In order to prove Theorem 2 we'll need some preliminary definitions and lemmas. Before we get into our first lemma, recall the Cauchy–Schwarz inequality which we will use repeatedly.

$$\left| \sum_{i=1}^{n} x_i \bar{y}_i \right|^2 \le \sum_{i=1}^{n} |x_i|^2 \sum_{i=1}^{n} |y_i|^2 \tag{4}$$

We want to start with some lemmas first. They will be useful as we will have some double sums that we want to separate into cases where the indexes are equal to each other and where they are not. The next few lemmas are going to help treat the case where the indexes are not equal. Our first lemma is due to Montgomery and Vaughan [5] and is a generalization of the Hilbert Inequality.

Lemma 3. *Suppose λ_r are a finite collection of distinct real numbers with $|\lambda_r - \lambda_s| \ge \delta$ if $r \ne s$ for $0 < \delta < \frac{1}{2}$. Then for any complex numbers z_r we have*

$$\left| \sum_{\substack{(r,s) \\ r \ne s}} \frac{z_r \overline{z_s}}{\lambda_r - \lambda_s} \right| \le \frac{\pi}{\delta} \sum_r |z_r|^2. \tag{5}$$

Proof. First note that without loss of generality, we can order the λ_r such that $\lambda_1 < \lambda_2 < ...$ and so we can replace the condition $|\lambda_r - \lambda_s| \ge \delta$ with the condition $|\lambda_r - \lambda_s| \ge \delta |r - s|$. We'll begin by proving

$$\sum_r \left| \sum_{r \ne s} \frac{\overline{z_s}}{\lambda_r - \lambda_s} \right|^2 \le \frac{\pi^2}{\delta^2} \sum_r |z_r|^2. \tag{6}$$

Note that

$$\left|\sum_s f(s)\right|^2 = \left(\sum_s f(s)\right)\overline{\left(\sum_t f(t)\right)} = \left(\sum_s f(s)\right)\left(\sum_t \overline{f(t)}\right) = \sum_s\sum_t f(s)\overline{f(t)},$$

so

$$\sum_r \left|\sum_{r\neq s} \frac{\overline{z_s}}{\lambda_r - \lambda_s}\right|^2 = \sum_{(s,t)} \overline{z_s} z_t \sum_{r\neq s,t} \frac{1}{(\lambda_r-\lambda_s)(\lambda_r-\lambda_t)},$$

noting that there are no conjugates on the λ as they are real, we now separate the sum into two pieces based on whether $s=t$ or $s\neq t$. The first sum is

$$\sum_s |z_s|^2 \sum_{r\neq s} \frac{1}{(\lambda_r-\lambda_s)^2},$$

while the second is

$$\begin{aligned}\sum_{\substack{(s,t)\\ s\neq t}} \overline{z_s} z_t \sum_{r\neq s,t} \frac{1}{(\lambda_r-\lambda_s)(\lambda_r-\lambda_t)} &= \sum_{\substack{(s,t)\\ s\neq t}} \frac{\overline{z_s} z_t}{\lambda_s-\lambda_t} \sum_{r\neq s,t}\left(\frac{1}{\lambda_r-\lambda_s} - \frac{1}{\lambda_r-\lambda_t}\right)\\ &= \sum_{\substack{(s,t)\\ s\neq t}} \frac{\overline{z_s} z_t}{\lambda_s-\lambda_t}\left[\sum_{r\neq s}\frac{1}{\lambda_r-\lambda_s} - \sum_{r\neq t}\frac{1}{\lambda_r-\lambda_t} + \frac{2}{\lambda_s-\lambda_t}\right].\end{aligned}$$

Now as we sum over all s,t we notice that the two inner sums will cancel each other. Hence we are left with

$$\sum_{\substack{(s,t)\\ s\neq t}} \frac{\overline{z_s} z_t}{\lambda_s-\lambda_t}\left[\frac{2}{\lambda_s-\lambda_t}\right] = \sum_{\substack{(s,t)\\ s\neq t}} \frac{2\overline{z_s} z_t}{(\lambda_s-\lambda_t)^2}.$$

Now by using $2|z_s z_t| \leq |z_s|^2 + |z_t|^2$, we get

$$\left|\sum_{\substack{(s,t)\\ s\neq t}} \frac{2\overline{z_s} z_t}{(\lambda_s-\lambda_t)^2}\right| \leq \sum_s\sum_{t\neq s} \frac{|z_s|^2+|z_t|^2}{(\lambda_s-\lambda_t)^2} = 2\sum_s |z_s|^2 \sum_{t\neq s}\frac{1}{(\lambda_s-\lambda_t)^2},$$

where the last equality is due to the pair (i,j) occurring once as $s=i, t=j$ and once as $s=j, t=i$. Therefore we get

$$\begin{aligned}\sum_r \left|\sum_{r\neq s} \frac{\overline{z_s}}{\lambda_r-\lambda_s}\right|^2 &= \sum_s |z_s|^2 \sum_{r\neq s}\frac{1}{(\lambda_r-\lambda_s)^2} + 2\sum_s |z_s|^2 \sum_{t\neq s}\frac{1}{(\lambda_s-\lambda_t)^2}\\ &= 3\sum_s |z_s|^2 \sum_{r\neq s}\frac{1}{(\lambda_r-\lambda_s)^2}.\end{aligned}$$

Since $|\lambda_r - \lambda_s| \geq \delta|r-s|$, we get

$$
\begin{aligned}
3\sum_s |z_s|^2 \sum_{r\neq s} \frac{1}{(\lambda_r-\lambda_s)^2} &\le \frac{3}{\delta^2}\sum_s |z_s|^2 \sum_{r\neq s}\frac{1}{(r-s)^2} \\
&\le \frac{3}{\delta^2}\sum_s |z_s|^2 \left(2\sum_{n=1}^{\infty}\frac{1}{n^2}\right) \\
&= \frac{6}{\delta^2}\frac{\pi^2}{6}\sum_s |z_s|^2 \\
&= \frac{\pi^2}{\delta^2}\sum_s |z_s|^2.
\end{aligned}
$$

By Cauchy–Schwarz (4) we get

$$
\left|\sum_{\substack{(r,s)\\ r\neq s}} \frac{z_r\overline{z_s}}{\lambda_r-\lambda_s}\right|^2 \le \sum_r |z_r|^2 \sum_r \left|\sum_{\substack{(r,s)\\ r\neq s}} \frac{\overline{z_s}}{\lambda_r-\lambda_s}\right|^2 \le \sum_r |z_r|^2 \left(\frac{\pi^2}{\delta^2}\sum_s |z_s|^2\right) = \left(\frac{\pi}{\delta}\sum_r |z_r|^2\right)^2
$$

which implies our theorem. □

We will also need the following identity in order to help with the next lemma. The proof of it will require Liousville's Theorem in complex analysis.

Lemma 4. *If z is a complex number with $z \notin \mathbb{Z}$, then*

$$
\frac{1}{z} + 2z\sum_{k=1}^{\infty}\frac{(-1)^k}{z^2-k^2} = \frac{\pi}{\sin \pi z}.
$$

Proof. First note from the power series expansion that $\sin \pi z$ has simple zeros at $k \in \mathbb{Z}$. Also note that

$$
\lim_{z\to k}\frac{\pi(z-k)}{\sin \pi z} = \lim_{z\to k}\frac{\pi}{\pi\cos \pi z} = (-1)^k
$$

so $\frac{\pi}{\sin \pi z}$ has simple poles at $k \in \mathbb{Z}$ with residue $(-1)^k$. Hence if we let

$$
S = \frac{1}{z} + 2z\sum_{k=1}^{\infty}\frac{(-1)^k}{z^2-k^2} = \lim_{K\to\infty}\sum_{k=-K}^{K}\frac{(-1)^k}{k-z},
$$

we get that $\frac{\pi}{\sin \pi z} - S$ is an entire function. It's clear that as $|z| \to \infty$ away from the integers, that S is bounded. Also

$$
\sin z = \frac{e^{iz}-e^{-iz}}{2i}
$$

which is bounded away from 0 as long as z is not near πk for integer k, since as $|Im(z)| \to \infty \Rightarrow |\sin z| \to \infty$, and as $|Re(z)| \to \infty$ for $Im(z) = b \neq 0$ we get

$$
|\sin z| = \lim_{|a|\to\infty}\frac{e^{-b}e^{ia}-e^{b}e^{-ia}}{2i}.
$$

If this went to 0, then $|e^{-b}e^{ia}| \to |e^{b}e^{-ia}|$ However that implies $e^b = e^{-b}$ which is impossible for $b \neq 0$. Hence $\sin z$ is bounded away from zero making $\frac{\pi}{\sin \pi z} - S$ a bounded entire function. By Liousville's Theorem in complex analysis, that means that $\frac{\pi}{\sin \pi z} - S$ is constant. Hence

$$\frac{\pi}{\sin \pi z} - \frac{1}{z} = C + 2z \sum_{k=1}^{\infty} \frac{(-1)^k}{z^2 - k^2},$$

so taking limits as $z \to 0$ yields

$$\begin{aligned} C &= \lim_{z\to 0} \frac{\pi}{\sin \pi z} - \frac{1}{z} \\ &= \lim_{z\to 0} \frac{\pi z - \sin \pi z}{z \sin \pi z} \\ &= \lim_{z\to 0} \frac{\pi - \pi \cos \pi z}{\sin \pi z + \pi z \cos \pi z} \\ &= \lim_{z\to 0} \frac{\pi^2 \sin \pi z}{2\pi cos\pi z - z \sin \pi z} = 0. \end{aligned}$$

Hence $C = 0$ yielding our identity.

□

Lemma 3 yields some corollaries. Note that our summations are involving sine functions which are part of the definition of $e(z)$.

Corollary 5. *If z_r are complex numbers and $\alpha_r \in \mathbb{R}/\mathbb{Z}$ with $|\alpha_r - \alpha_s| \geq \delta$, for $0 < \delta < \frac{1}{2}$, then*

$$\left| \sum_{\substack{(r,s) \\ r \neq s}} \frac{z_r \overline{z_s}}{\sin \pi(\alpha_r - \alpha_s)} \right| \leq \delta^{-1} \sum_r |z_r|^2.$$

Proof. Using Lemma 3 with the values $z_{m,r} = (-1)^m z_r$ and $\lambda_{m,r} = m + \alpha_r$ for $1 \leq m \leq K$, we get

$$\begin{aligned} \left| \sum_{\substack{(r,m),(s,n) \\ (r,m)\neq(s,n)}} (-1)^{m-n} \frac{z_r \overline{z_s}}{m - n + \alpha_r - \alpha_s} \right| &= \left| \sum_{\substack{(r,m),(s,n) \\ (r,m)\neq(s,n)}} \frac{z_{m,r} \overline{z_{n,s}}}{\lambda_{m,r} - \lambda_{n,s}} \right| \\ &\leq \frac{\pi}{\delta} \sum_r |z_{m,r}|^2 \\ &\leq \frac{\pi K}{\delta} \sum_r |z_r|^2, \end{aligned}$$

noting that the δ hasn't changed since adding an integer or multiplying by (-1) won't effect the spacing. Now if $r = s$ we note that the term in the sum for $m = i, n = j$ is

$$(-1)^{i-j} \frac{|z_r|^2}{i - j} = -(-1)^{j-i} \frac{|z_r|^2}{j - i}$$

which is the negative of the term for $m = j, n = i$. Therefore they cancel out in the sum and hence the condition $(r, m) \neq (s, n)$ can be simply replaced by $r \neq s$. Therefore we have

$$\frac{\pi}{\delta}\sum_r |z_r|^2 \geq \frac{1}{K}\left|\sum_{\substack{r,s,m,n\\ r\neq s}} (-1)^{m-n}\frac{z_r\overline{z_s}}{m-n+\alpha_r-\alpha_s}\right|$$
$$= \left|\sum_{\substack{r,s\\ r\neq s}} z_r\overline{z_s}\sum_{k=-K}^{K}\frac{N(k)}{K}\frac{(-1)^k}{k+\alpha_r-\alpha_s}\right|$$

where $N(k)$ is the number of pairs (m,n) with $1 \leq m,n \leq K$ such that $m-n=k$. Clearly $N(k) = K - |k|$ and so we get

$$\frac{\pi}{\delta}\sum_r |z_r|^2 \geq \left|\sum_{\substack{r,s\\ r\neq s}} z_r\overline{z_s}\sum_{k=-K}^{K}\left(1-\frac{|k|}{K}\right)\frac{(-1)^k}{k+\alpha_r-\alpha_s}\right|.$$

By letting $K \to \infty$ and using Lemma 4 we get

$$\frac{\pi}{\delta}\sum_r |z_r|^2 \geq \left|\sum_{\substack{r,s\\ r\neq s}} \frac{\pi z_r\overline{z_s}}{\sin\pi(\alpha_r-\alpha_s)}\right|$$

yielding our corollary. □

Now the previous corollary leads directly to

Corollary 6. *For any real number x, complex numbers z_r and $\alpha_r \in \mathbb{R}/\mathbb{Z}$ with $|\alpha_r - \alpha_s| \geq \delta$ for $0 < \delta < \frac{1}{2}$, we have*

$$\left|\sum_{\substack{(r,s)\\ r\neq s}} \frac{z_r\overline{z_s}\sin 2\pi x(\alpha_r-\alpha_s)}{\sin\pi(\alpha_r-\alpha_s)}\right| \leq \delta^{-1}\sum_r |z_r|^2.$$

Proof. Using Lemma 5 with $z_r' = z_r e(x\alpha_r)$ and $z_r'' = z_r e(-x\alpha_r)$ and noticing that $|z_r''| = |z_r'| = |z_r|$ we get

$$\left|\sum_{\substack{(r,s)\\ r\neq s}} \frac{z_r\overline{z_s}e(x\alpha_r - x\alpha_s)}{\sin\pi(\alpha_r-\alpha_s)}\right| \leq \delta^{-1}\sum_r |z_r|^2 \text{ and}$$
$$\left|\sum_{\substack{(r,s)\\ r\neq s}} \frac{z_r\overline{z_s}e(-x\alpha_r + x\alpha_s)}{\sin\pi(\alpha_r-\alpha_s)}\right| \leq \delta^{-1}\sum_r |z_r|^2.$$

Now using that $e(z) - e(-z) = 2i\sin(2\pi z)$ we get

$$\left|\sum_{\substack{(r,s)\\ r\neq s}} \frac{z_r\overline{z_s}\sin 2\pi x(\alpha_r-\alpha_s)}{\sin\pi(\alpha_r-\alpha_s)}\right| = \left|\sum_{\substack{(r,s)\\ r\neq s}} \frac{z_r\overline{z_s}\Big(e(x(\alpha_r-\alpha_s))-e(x(\alpha_r-\alpha_s))\Big)}{2\sin\pi(\alpha_r-\alpha_s)}\right|$$

$$\leq \left|\sum_{\substack{(r,s)\\ r\neq s}} \frac{z_r\overline{z_s}e(x\alpha_r-x\alpha_s)}{2\sin\pi(\alpha_r-\alpha_s)}\right| + \left|\sum_{\substack{(r,s)\\ r\neq s}} \frac{z_r\overline{z_s}e(-x\alpha_r+x\alpha_s)}{2\sin\pi(\alpha_r-\alpha_s)}\right|$$

$$\leq \delta^{-1}\sum_r |z_r|^2.$$

□

Now we show the Duality Lemma. It allows us to interchange the order of summation. This will allow us to sum over the residues first in our theorem as opposed to over the n which will be useful.

Lemma 7. *Given a function $\phi(m,n)$. Suppose for any complex numbers β_n we have*

$$\sum_m \left|\sum_n \beta_n\phi(m,n)\right|^2 \leq \Delta\|\vec{\beta}\|^2 \tag{7}$$

then for any complex numbers α_n we get

$$\sum_n \left|\sum_m \alpha_m\phi(m,n)\right|^2 \leq \Delta\|\vec{\alpha}\|^2.$$

Proof. If we let $\beta_n = \sum_m \overline{\alpha_m\phi(m,n)}$, then

$$\sum_n \left|\sum_m \alpha_m\phi(m,n)\right|^2 = \sum_n \left(\sum_m \alpha_m\phi(m,n)\right)\left(\sum_{m'} \overline{\alpha_{m'}\phi(m',n)}\right)$$

$$= \sum_n\sum_m \alpha_m\beta_n\phi(m,n)$$

$$= \sum_m \alpha_m \sum_n \beta_n\phi(m,n).$$

Now let this sum be $\Phi(m,n)$ then using Cauchy–Schwarz (4) we get,

$$|\Phi(m,n)|^2 \leq \sum_m|\alpha_m|^2 \sum_m \left|\sum_n \beta_n\phi(m,n)\right|^2 \leq \Delta\|\vec{\alpha}\|^2\|\vec{\beta}\|^2$$

by hypothesis (7). However note that

$$\Phi(m,n) = \sum_n \left|\sum_m \alpha_m\phi(m,n)\right|^2 = \sum_n |\beta_n|^2 = \|\vec{\beta}\|^2$$

and so we are left with

$$\Phi(m,n) \le \Delta \|\vec{\alpha}\|^2$$

as needed.

□

We require one more lemma. It will use the duality lemma and the additive large sieve will be a simple corollary of it using rational numbers.

Lemma 8. *For any $\alpha_r \in \mathbb{R}/\mathbb{Z}$ with $|\alpha_r - \alpha_s| \ge \delta$ for $0 < \delta < \frac{1}{2}$ and any complex numbers a_n with $M < n \le M+N$, we have*

(a)

$$\sum_r \Bigg| \sum_{M<n<M+N} a_n e(\alpha_r n) \Bigg|^2 \le (\delta^{-1} + N)\|\vec{a}\|^2.$$

(b) *Replace $(\delta^{-1}+N)$ with $(\delta^{-1}+N-1)$.*

Proof. (a) Using duality, we can instead show the inequality

$$\sum_{N=M+1}^{M+N} \Bigg| \sum_r z_r e(\alpha_r n) \Bigg|^2 \le (\delta^{-1} + N)\|\vec{z}\|^2.$$

The summation can be expanded to get

$$\sum_{N=M+1}^{M+N} \Bigg(\sum_r z_r e(\alpha_r n) \Bigg) \Bigg(\sum_s \overline{z_s} e(-\alpha_s n) \Bigg).$$

For the terms where $r = s$ we have

$$\sum_{N=M+1}^{M+N} \sum_r |z_r|^2 = N\|\vec{z}\|^2,$$

and for the terms with $r \ne s$ we have

$$\sum_{N=M+1}^{M+N} \Bigg(\sum_r z_r e(\alpha_r n) \Bigg) \Bigg(\sum_{s\ne r} \overline{z_s} e(-\alpha_s n) \Bigg) = \sum_{\substack{(r,s)\\ r\ne s}} z_r \overline{z_s} \sum_{N=M+1}^{M+N} e(n(\alpha_r - \alpha_s)).$$

Looking at the inner sum we note that

$$\begin{aligned}
\sum_{c<n\le d} e(n\alpha) &= \frac{e((c+1)\alpha) - e((d+1)\alpha)}{1 - e(\alpha)} \\
&= e((-d+3c+1)\alpha/2) \frac{e(-(d-c)\alpha/2) - e((d-c)\alpha/2)}{e(-\alpha/2) - e(\alpha/2)} \\
&= e((-d+3c+1)\alpha/2) \frac{\sin(\pi(d-c)\alpha)}{\sin \pi\alpha},
\end{aligned}$$

so we get that our sum is equal to

$$\sum_{\substack{(r,s)\\ r\ne s}} z_r \overline{z_s} e((M + \frac{1}{2}(N+1))(\alpha_r - \alpha_s)) \frac{\sin \pi N(\alpha_r - \alpha_s)}{\sin \pi(\alpha_r - \alpha_s)}.$$

By noting that our sum is less than the the sum in 6 with $x = \frac{N}{2}$ we get that

$$\sum_{N=M+1}^{M+N} \left(\sum_r z_r e(\alpha_r n) \right) \left(\sum_{s \neq r} \overline{z_s} e(-\alpha_s n) \right) \leq \delta^{-1} \|\vec{z}\|^2$$

finishing the proof of (a).

(b) Now given the a_n, α_r and the δ from part (a), let $\beta_m = \frac{\alpha_r + k}{K}$, where $1 \leq k \leq K$. The $\beta_{r,k}$ are separated by $\frac{\delta}{K}$. If we let

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

then we get another trigonometric polynomial $T(\alpha)$ where

$$\sum_r \left| S(\alpha_r) \right|^2 = \frac{1}{K} \sum_{k,r} \left| T(\beta_{r,k}) \right|^2.$$

Since $T(\alpha)$ ranges over the $m = nK$ where $MK + K \leq m \leq (M+N)K$, the number of terms in the summation is $N' = NK - K + 1$. Hence part (a) says that

$$\begin{aligned} \sum_r \left| \sum_{N=M+1}^{M+N} a_n e(\alpha_r n) \right|^2 &\leq \frac{1}{K} (K\delta^{-1} + NK - K + 1) \|\vec{a}\|^2 \\ &= (\delta^{-1} + N - 1 - 1/K) \|\vec{a}\|^2. \end{aligned}$$

As $K \to \infty$ we get

$$\sum_r \left| \sum_{N=M+1}^{M+N} a_n e(\alpha_r n) \right|^2 \leq (\delta^{-1} + N - 1) \|\vec{a}\|^2$$

finishing our lemma.

□

Note that while part (b) is required for the proof of Theorem 2 and Theorem 9 below, the -1 won't be necessary for the proof of Theorem 1. We now have enough to prove our main theorem of this section.

Proof of Theorem 2. Let α_r be the rational numbers $\frac{a}{q}$ (mod 1) with $1 \leq q \leq Q$ and $(a, q) = 1$. We need to figure out how spaced they are away from each other. If $\frac{a}{c} \neq \frac{b}{d}$, then

$$\left| \frac{a}{c} - \frac{b}{d} \right| = \left| \frac{ad - bc}{cd} \right|.$$

Since $ad - bc$ is an integer and is not zero since $\frac{a}{c} \neq \frac{b}{d}$ we know that $|ad - bc| \geq 1$ and so

$$\left| \frac{a}{c} - \frac{b}{d} \right| \geq \frac{1}{cd} \geq Q^{-2}$$

since $1 \leq c, d \leq Q$. Hence we can take δ to be Q^{-2} so by Lemma 8 we get our theorem.

□

3. The Multiplicative Large Sieve

Throughout the remainder of the paper, we will denote $\|\vec{a}\| = \left(\sum_i |a_i|^2\right)^{1/2}$. The main theorem we wish to prove in this section is:

Theorem 9. *For natural numbers M, N, let a_n be complex numbers with $n \in \{M+1, ..., M+N\}$. Then we have the following:*

$$\sum_{q \le Q} \frac{q}{\phi(q)} \sideset{}{'}\sum_{\chi \text{ (mod } q)} \left| \sum_{N=M+1}^{M+N} a_n \chi(n) \right| \le (Q^2 + N - 1)\|\vec{a}\|^2, \tag{8}$$

where $\sum'_{\chi \text{ (mod } q)}$ indicates the sum over primitive characters χ.

Note that we have multiplicative characters $\chi(n)$ explaining the terminology. To prove this theorem, we will need some basic facts about character sums. The first is orthogonality, which are just stated. The proofs are easy and can be found in any analytic number theory book, for example [6, Pages 108,117].

Lemma 10 (Orthogonality)**.** (a)

$$\sum_{\chi \text{ (mod } q)} \chi(n)\overline{\chi}(a) = \begin{cases} \phi(q) & n \equiv a \text{ (mod } q) \\ 0 & \textit{otherwise.} \end{cases}$$

(b)

$$\sum_{a=1}^{q} \chi(a) = \begin{cases} \phi(q) & \chi = \chi_0 \\ 0 & \textit{otherwise} \end{cases}$$

where χ_0 is the principal character.

(c)

$$\sum_{a=1}^{q} e(an/q)e(-am/q) = \begin{cases} q & m \equiv n \text{ (mod } q) \\ 0 & \textit{otherwise.} \end{cases}$$

Definition 11. *Given a character χ (mod q), the Gauss Sum is*

$$\tau(\chi) = \sum_{a \text{ (mod } q)} \chi(a) e\left(\frac{a}{q}\right).$$

Notice how τ relates multiplicative characters to additive ones, which is how we'll go from the additive large sieve to the multiplicative one. We'll need some facts about the function.

Lemma 12. (a) *If $(n, q) = 1$, then*

$$\tau(\chi)\overline{\chi}(n) = \sum_{a \text{ (mod } q)} \chi(a) e\left(\frac{an}{q}\right).$$

(b) *If χ is primitive, then the condition $(n, q) = 1$ is unnecessary.*
(c) *If χ is primitive, then*

$$|\tau(\chi| = \sqrt{q}.$$

Note that we won't actually use part (a) for our proof, but it's nice to know and not hard to prove.

Proof. (a)

$$\tau(\chi)\overline{\chi}(n) = \sum_{a \text{ (mod } q)} \chi(a)\overline{\chi}(n)e\left(\frac{a}{q}\right).$$

Now since $(n, q) = 1$, $\{an \text{ (mod } q) \mid (a, q) = 1\} = \{a \text{ (mod } q) | (a, q) = 1\}$. Therefore we can replace a by an in the sum. So

$$\begin{aligned} \tau(\chi)\overline{\chi}(n) &= \sum_{a \text{ (mod } q)} \chi(an)\overline{\chi}(n)e\left(\frac{an}{q}\right) \\ &= \sum_{a \text{ (mod } q)} \chi(a)\chi(n)\overline{\chi}(n)e\left(\frac{an}{q}\right) \\ &= \sum_{a \text{ (mod } q)} \chi(a)e\left(\frac{an}{q}\right), \end{aligned}$$

since $\chi(n)\overline{\chi}(n) = |\chi(n)|^2 = 1$ for $(n, q) = 1$.

(b) It's true by (a) if $(n, q) = 1$, so it's sufficient to show that both sides are 0 for $(n, q) \neq 1$. The left side is 0 since χ is supported on n where $(n, q) = 1$. As for the right side we first claim that if χ is primitive, then for $d \mid q, d < q$ and for every integer a,

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

This is true since χ primitive means there exists $m, n,\ m \equiv n \text{ (mod } d)$ with $\chi(m) \neq \chi(n), \chi(mn) \neq 0$ Hence there exists c with $(c, q) = 1, c \equiv 1 \text{ (mod } d)$ such that

$$\begin{aligned} cm = n \text{ (mod } q) &\Rightarrow \chi(c)\chi(m) = \chi(n) \\ &\Rightarrow \chi(c) \neq 1. \end{aligned}$$

Now as k runs through the residue system $\text{(mod } q/d)$ such that $(k, q/d) = 1$. Then the numbers $n = ac + kcd$ runs through all the residues $\text{(mod } q)$ with $n \equiv a \text{ (mod } d)$. So

$$\begin{aligned} \sum_{\substack{n=1 \\ n \equiv a \text{ (mod } d)}}^{q} \chi(n) &= \sum_{k=1}^{q/d} \chi(ac + kcd) \\ &= \chi(c) \sum_{k=1}^{q/d} \chi(a + kd) \\ &= \chi(c) \sum_{\substack{n=1 \\ n \equiv a \text{ (mod } d)}}^{q} \chi(n). \end{aligned}$$

Since $\chi(c) \neq 1$, our sum is equal to 0 proving the claim. Now if $(n, q) \neq 1$, then let $\frac{m}{d} = \frac{n}{q}$ with $(m, d) = 1$ implying $d \mid q, d < q$, and so

$$\sum_{a \text{ (mod } q)} \chi(a) e\left(\frac{an}{q}\right) = \sum_{h=1}^{d} e(hm/d) \sum_{\substack{a=1 \\ a \equiv h \text{ (mod } d)}}^{q} \chi(a) = 0$$

by the claim.

(c) Since there are $\phi(n)$ value of n (mod q) with $(n, q) = 1$ and

$$\chi(n)\overline{\chi}(n) = |\chi(n)| = \begin{cases} 1 & (n, q) = 1 \\ 0 & \text{otherwise.} \end{cases},$$

we have by part (b)

$$\begin{aligned}
\phi(q)|\tau(\chi)|^2 &= \sum_{n=1}^{q} \chi(n)\overline{\chi}(n)\tau(\chi)\tau(\overline{\chi}) \\
&= \sum_{n=1}^{q} \sum_{a \text{ (mod } q)} \chi(a) e\left(\frac{an}{q}\right) \sum_{b \text{ (mod } q)} \overline{\chi}(b) e\left(\frac{-bn}{q}\right) \\
&= \sum_{a \text{ (mod } q)} \sum_{b \text{ (mod } q)} \chi(a)\overline{\chi}(b) \sum_{n=1}^{q} e\left(\frac{(a-b)n}{q}\right).
\end{aligned}$$

By orthogonality, part (c) the last sum is 0 unless $a = b$ in which it is q, so we get

$$\phi(q)|\tau(\chi)|^2 = \sum_{a \text{ (mod } q)} |\chi(a)|^2 q = q\phi(q) \Rightarrow |\tau(\chi)| = \sqrt{q}.$$

□

Note that the claim in (b) is an if and only if statement, see [6, Page 284]. We can now prove the main theorem of this section. It's based off the proof in [4, Page 179] however they prove something stronger which we won't need.

Proof of Theorem 9. Let the a_n, M, N be as in the statement of the theorem. Let $S(\alpha)$ be the trigonometric sum $\sum_n a_n e(n\alpha)$ as in the proof of the additive large sieve. then by Lemma 12 (b) we get

$$\begin{aligned}
\sum_{N=M+1}^{M+N} a_n \chi(n) &= \frac{1}{\tau(\overline{\chi})} \sum_{N=M+1}^{M+N} \sum_{a \text{ (mod } q)} \overline{\chi}(a) e\left(\frac{an}{q}\right) \\
&= \frac{1}{\tau(\overline{\chi})} \sum_{a \text{ (mod } q)} \overline{\chi}(a) S\left(\frac{a}{q}\right).
\end{aligned}$$

Taking the modulus of the sum and using $|\tau(\chi))|^2 = q$ yields

$$\frac{q}{\phi(q)} \left| \sum_{N=M+1}^{M+N} a_n \chi(n) \right|^2 = \frac{1}{\phi(q)} \left| \sum_{a \text{ (mod } q)} \overline{\chi}(a) S\left(\frac{a}{q}\right) \right|^2.$$

By summing over $1 \leq q \leq Q$ and over all primitive characters (actually we'll achieve an upper bound by summing over all characters) we get the left side of our theorem is bounded above by

$$\frac{1}{\phi(q)}\sum_{q\le Q}\sum_{\chi \text{ (mod } q)}\left|\sum_{a \text{ (mod } q)}\overline{\chi}(a)S\left(\frac{a}{q}\right)\right|^2 \le \frac{1}{\phi(q)}\sum_{q\le Q}\sum_{\chi \text{ (mod } q)}1\left|\sum_{a \text{ (mod } q)}S\left(\frac{a}{q}\right)\right|^2.$$

By orthogonality, part (a) we have $\sum_{\chi \text{ (mod } q)} 1 = \phi(q)$ so Theorem 2 implies

$$\sum_{q\le Q}\left|\sum_{a \text{ (mod } q)}S\left(\frac{a}{q}\right)\right|^2 \le (Q^2+N-1)\|\vec{a}\|^2$$

completing our proof.

□

4. A Related Function

For the Bombieri–Vinogradov Theorem, we are trying to show that the difference between a function over a specific arithmetic progression and the average over all arithmetic progression is relatively small. We will give a specific function for this difference, and derive some properties.

Definition 13. *Given an arithmetic function f, we let*

$$D_f(x;q,a) = \sum_{\substack{n\le x\\ n\equiv a \text{ (mod } q)}} f(n) - \frac{1}{\phi(q)}\sum_{\substack{n\le x\\ (n,q)=1}} f(n).$$

Also recall the following definition:

Definition 14 (Dirichlet Convolution)**.** *Given arithmetic functions f, g, the Dirichlet Convolution is*

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d). \tag{9}$$

Now we will state our main theorem which will be helpful in showing the Bombieri–Vinogradov Theorem. It turns out the Dirichlet convolution gives an estimate which doesn't require the difference between two sums. This will be nicer to estimate, provided we can write our function as a convolution.

Theorem 15. *Suppose we have sequences $\alpha = (\alpha_n)$, $\beta = (\beta_n)$ which are supported on $\{1, ..., M\}$ and $\{1, ..., N\}$ respectively. Suppose*

$$|D_\beta(N;q,a)| \le \|\vec{\beta}\|N^{\frac{1}{2}}\Delta^9 \text{ for some } \Delta \in (0,1]. \tag{10}$$

Let $\gamma = \alpha * \beta$ be the Dirichlet convolution, then

$$\sum_{q\le Q}\max_{(a,q)=1}|D_\gamma(MN;q,a)| \ll \|\vec{\alpha}\|\|\vec{\beta}\|(\Delta M^{1/2}N^{1/2} + M^{1/2} + N^{1/2} + Q)\log^2 Q.$$

To prove this, we are going to need to make the summation condition $n \equiv a$ (mod q) easier to work with. We can use characters to do that, hence why we require the multiplicative large sieve. In the process of proving this, we will be splitting the interval $1, ..., Q$ into parts, so we'll need a lemma to help deal with the small values of q. First let's recall the formula for Möbius inversion of which the proof can be seen in any elementary number theory textbook, for example [6, Page 35].

Lemma 16 (Möbius Inversion). *If f, and F are two arithmetic functions such that*

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

then

$$f(n) = \sum_{d|n} \mu(d) F(n/d)$$

where μ is the Möbius μ-function.

Recall the elementary identity of whose proof can be found by using Möbius inversion with $F(d) = 1$ for all d.

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & n = 1 \\ 0 & n > 1 \end{cases} \tag{11}$$

We are going to need a technical lemma first.

Lemma 17. *Let $\beta = (\beta_n)$ be supported on $\{1, ..., N\}$, then for all natural numbers K, s we get*

$$\sum_{\substack{d|s \\ d \leq K}} \mu(d) \sum_{n \equiv 0 \ (\mathrm{mod}\ d)} \beta_n \chi(n) = \sum_{\substack{d|s \\ d \leq K}} \mu(d) \sum_{l|d} \mu(l) \sum_{(n,l)=1} \beta_n \chi(n).$$

Proof. By linearity of the finite sum, it suffices to show it's true for

$$\beta(n) = \begin{cases} 1 & n = n_0 \\ 0 & otherwise \end{cases}.$$

For this $\beta(n)$, the left hand side is

$$\sum_{\substack{d|s \\ d \leq K}} \mu(d) \begin{cases} 1 & (n_0, q) = 1, d \mid n_0 \\ 0 & otherwise \end{cases},$$

and the right hand side is

$$\sum_{\substack{d|s \\ d \leq K}} \mu(d) \sum_{l|d} \mu(l) \begin{cases} 1 & (n_0, q) = 1, (n_0, l) = 1 \\ 0 & otherwise \end{cases} = \sum_{\substack{d|s \\ d \leq K \\ (n_0,q)=1}} \mu(d) \sum_{\substack{l|d \\ (n_0,l)=1}} \mu(l).$$

Since $\mu(d)$ is only supported on square free numbers the inner sum is

$$\sum_{l | \prod_{\substack{p|d \\ (n,p)=1}} p} \mu(l),$$

hence by (11) we get that this is

$$\begin{cases} 1 & \prod_{\substack{p|d \\ (n,p)=1}} p = 1 \\ 0 & otherwise \end{cases},$$

in other words $d \mid n_0$. Hence the right hand side is

$$\sum_{\substack{d|s \\ d\le K \\ (n_0,q)=1}} \mu(d) \begin{cases} 1 & d \mid n_0 \\ 0 & otherwise \end{cases}$$

finishing our lemma. ☐

Now we will show a lemma that can help us deal with the resulting sum that will arise when changing the summation condition to sums involving characters.

Lemma 18. *If $\beta = (\beta_n)$ is a sequence supported on $\{1, ..., N\}$ satisfying the conditions in equation* (10) *then for all χ, a non-principal character* $\pmod r$ *and positive integers s we have*

$$\left| \sum_{(n,s)=1} \beta_n \chi(n) \right| \le \|\vec{\beta}| N^{1/2} \Delta^3 r \tau(s),$$

where $\tau(s)$ is the number of positive divisors of s.

Proof. Using (11) we get

$$\begin{aligned} \sum_{(n,s)=1} \beta_n \chi(n) &= \sum_n \beta_n \chi(n) \sum_{d|(n,s)} \mu(d) \\ &= \sum_{d|s} \mu(d) \sum_{n\equiv 0 \pmod d} \beta_n \chi(n). \end{aligned}$$

We are going to split the sum over d at the value K. For $d \le K$ we have

$$\sum_{\substack{d|s \\ d\le K}} \mu(d) \sum_{n\equiv 0 \pmod d} \beta_n \chi(n) = \sum_{\substack{d|s \\ d\le K}} \mu(d) \sum_{l|d} \mu(l) \sum_{(n,l)=1} \beta_n \chi(n)$$

by Lemma 17. Looking at the inner sum over residue classes $\pmod{lr}$ and looking at the condition given in (10) we know that

$$\begin{aligned} \sum_{(n,l)=1} \beta_n \chi(n) &= \sum_{\substack{a=1 \\ (a,lr)=1}}^{lr} \sum_{n\equiv a \pmod{lr}} \beta_n \chi(n) \\ &= \sum_{\substack{a=1 \\ (a,lr)=1}}^{lr} \chi(a) \sum_{n\equiv a \pmod{lr}} \beta_n \\ &= \sum_{\substack{a=1 \\ (a,lr)=1}}^{lr} \chi(a) \left(\sum_{(n,lr)=1} \beta_n + O(\|\vec{\beta}\| N^{\frac{1}{2}} \Delta^9) \right). \end{aligned}$$

The leading term goes to 0 since

$$\sum_{\substack{a=1 \\ (a,lr)=1}}^{lr} \chi(a) = 0,$$

giving us

$$\sum_{(n,l)=1} \beta_n \chi(n) \ll \phi(lr) \|\vec{\beta}\| N^{\frac{1}{2}} \Delta^9.$$

Therefore our sum over $d \leq K$ is

$$\ll \|\vec{\beta}\| N^{\frac{1}{2}} \Delta^9 \sum_{\substack{d|s \\ d \leq K}} |\mu(d)| \sum_{l|d} |\mu(l)| \phi(lr) \leq \|\vec{\beta}\| N^{\frac{1}{2}} \Delta^9 K \phi(r) \tau(s),$$

using that

$$\sum_{l|d} \phi(l) = d \leq K \text{ and } \sum_{\substack{d|s \\ d \leq K}} 1 \leq \tau(s).$$

As for the part where $d > K$, we have

$$\sum_{\substack{d|s \\ d>K}} \mu(d) \sum_{n \equiv 0 \ (\mathrm{mod}\ d)} \beta_n \chi(n).$$

Using Cauchy–Schwarz (4) on the inside summation, we get

$$\begin{aligned} \Big| \sum_{n \equiv 0 \ (\mathrm{mod}\ d)} \beta_n \chi(n) \Big|^2 &\leq \sum_{n \equiv 0 \ (\mathrm{mod}\ d)} |\beta_n|^2 \sum_{n \equiv 0 \ (\mathrm{mod}\ d)} |\chi_n|^2 \\ &\leq \|\vec{\beta}\|^2 \frac{N}{d}. \end{aligned}$$

Hence our sum for $d > K$ yields

$$\|\vec{\beta}\| N^{1/2} \sum_{\substack{d|s \\ d>K}} d^{-1/2} \leq \|\vec{\beta}\| N^{1/2} K^{-1/2} \tau(s),$$

since there are at most $\tau(s)$ such d, all of which have $d^{-1/2} < K^{-1/2}$ By letting $K = \Delta^{-6}$ we get

$$\begin{aligned} \Big| \sum_{(n,s)=1} \beta_n \chi(n) \Big| &\leq \|\vec{\beta}\| N^{1/2} \Delta^3 \phi(r) \tau(s) + \|\vec{\beta}\| N^{1/2} \Delta^3 \tau(s) \\ &\leq \|\vec{\beta}\| N^{1/2} \Delta^3 r \tau(s). \end{aligned}$$

□

In the proof of our theorem, we wish to understand the function $D_\gamma(MN; q, a)$, so in our next lemma will express it as a product of sums instead of a difference which will be much easier to work with.

Lemma 19.

$$D_\gamma(MN; q, a) = \frac{1}{\phi(q)} \sum_{\substack{\chi \ (\mathrm{mod}\ q) \\ \chi \neq \chi_0}} \overline{\chi}(a) \Big(\sum_{m \leq M} \alpha_m \chi(m) \Big) \Big(\sum_{n \leq N} \beta_n \chi(n) \Big)$$

where χ_0 is the principal character (mod q).

Proof. We will prove this by expanding out the definition of $D_\gamma(MN;q,a)$ and noticing some cancellation in the resulting sums.

$$D_\gamma(MN;q,a) = \sum_{\substack{k\le MN \\ k\equiv a \ (\mathrm{mod}\ q)}} \gamma(k) - \frac{1}{\phi(q)} \sum_{\substack{k\le MN \\ (k,q)=1}} \gamma(k).$$

By orthogonality part (a) we can replace the condition $k \equiv a \pmod q$ with the sum

$$\frac{1}{\phi(q)} \sum_{\chi \ (\mathrm{mod}\ q)} \chi(k)\overline{\chi}(a),$$

to get

$$\frac{1}{\phi(q)} \sum_{\substack{k\le MN \\ (k,q)=1}} \sum_{\chi \ (\mathrm{mod}\ q)} \chi(k)\overline{\chi}(a)\gamma(k) - \frac{1}{\phi(q)} \sum_{\substack{k\le MN \\ (k,q)=1}} \gamma(k).$$

Note that for the principal character,

$$\sum_{\substack{k\le MN \\ (k,q)=1}} \chi_0(k)\gamma(k) = \sum_{\substack{k\le MN \\ (k,q)=1}} \gamma(k),$$

since $\chi_0(k) = 1$ for all $(k,q) = 1$. We are left with

$$\frac{1}{\phi(q)} \sum_{\substack{k\le MN \\ (k,q)=1}} \sum_{\chi\ne\chi_0 \ (\mathrm{mod}\ q)} \overline{\chi}(a)\chi(k)\gamma(k) = \frac{1}{\phi(q)} \sum_{\chi\ne\chi_0 \ (\mathrm{mod}\ q)} \overline{\chi}(a) \sum_{\substack{k\le MN \\ (k,q)=1}} \chi(k)\gamma(k).$$

We're left to show

$$\sum_{\substack{k\le MN \\ (k,q)=1}} \chi(k)\gamma(k) = \left(\sum_{m\le M} \alpha_m\chi(m)\right)\left(\sum_{n\le N} \beta_n\chi(n)\right).$$

To this end we get

$$\left(\sum_{m\le M} \alpha_m\chi(m)\right)\left(\sum_{n\le N} \beta_n\chi(n)\right) = \sum_{m\le M}\sum_{n\le N} \alpha_m\beta_n\chi(mn).$$

Letting $k = mn$ the coefficient of $\chi(k)$ is the sum of all $\alpha_m\beta_n$ with $mn = k$.However since $\gamma = \alpha * \beta$, we get that the coefficient is

$$\sum_{m|k} \alpha(m)\beta(k/m) = \gamma(k)$$

finishing the lemma.

□

Due to two sums which will come up in the proof we will establish the following estimates, whose proofs rely on properties of multiplicative functions.

Lemma 20. (a)

$$\sum_{s\le Q}\frac{\tau(s)}{\phi(s)}\le \log^2 Q. \tag{12}$$

(b)

$$\sum_{s\le Q}\frac{1}{\phi(s)}\le \log Q. \tag{13}$$

Proof. For any non-negative multiplicative function $f(n)$ we first claim that

$$\sum_{n\le x} f(n)\le \prod_{p\le x}\sum_{k=0}^{\infty} f(p^k),$$

since by letting

$$n=\prod p^{a_p}$$

be the prime factorization of n, then since $p\le n$, all p have $p\le x$ and all products of those p^{a_p} occur on the right hand side of the inequality. Using

$$f(n)=\prod f(p^{a_p})$$

finishes the claim.

(a) Now $\tau(s)$ and $\phi(s)$ are both non-negative multiplicative functions and $\phi(s)$ is never zero, hence their quotient is clearly non-negative and multiplicative and so the claim applies to give us

$$\sum_{s\le Q}\frac{\tau(s)}{\phi(s)}\le \prod_{p\le Q}\sum_{k=0}^{\infty}\frac{\tau(p^k)}{\phi(p^k)}.$$

Using that $\tau(p^k)=k+1$ and $\phi(p^k)=p^{k-1}(p-1)$ we get

$$\sum_{s\le Q}\frac{\tau(s)}{\phi(s)}\le \prod_{p\le Q}\left(1+\sum_{k=1}^{\infty}\frac{k+1}{p^{k-1}(p-1)}\right).$$

The sum is absolutely convergent for any p so we are permitted to rearrange the sum to get

$$1+\sum_{k=1}^{\infty}\frac{k+1}{p^{k-1}(p-1)}=\left(1+\sum_{k=1}^{\infty}\frac{1}{p^{k-1}(p-1)}\right)+\sum_{l=1}^{\infty}\sum_{k=l}^{\infty}\frac{1}{p^{k-1}(p-1)}$$

since each k term will occur once in the first sum and k times in the double sum, once for each $1\le l\le k$. Now using geometric series this is equal to

$$1+\frac{1}{p-1}\frac{1}{1-\frac{1}{p}}+\sum_{l=1}^{\infty}\frac{1}{p^{l-1}(p-1)}\frac{1}{1-\frac{1}{p}}$$

$$=1+\frac{p}{(p-1)^2}+\frac{1}{(p-1)^2}\sum_{l=1}^{\infty}p^{2-l}.$$

which is another geometric series yielding

$$1+\frac{p}{(p-1)^2}+\frac{p^2}{(p-1)^3}=1+\frac{2p^2-p}{(p-1)^3}.$$

The product of these factors can be written as the exponential of a sum of their logarithms to get

$$\exp\left(\sum_{p\leq Q}\log\left(1+\frac{2p^2-p}{(p-1)^3}\right)\right).$$

Using that $\log(1+x)\leq x$ we get that

$$\log\left(1+\frac{2p^2-p}{(p-1)^3}\right)\leq\frac{2p^2-p}{(p-1)^3}=\frac{2}{p-1}+O\left(\frac{1}{(p-1)^2}\right)$$
$$=\frac{2}{p}+O\left(\frac{1}{p^2}\right).$$

Since the sum $\sum_{n\geq 1}n^{-2}$ converges, so does the sum over primes, and using Merten's [6, Theorem 2.7] we get

$$\sum_{s\leq Q}\frac{\tau(s)}{\phi(s)}\leq e^{2\log\log Q+C}\ll\log^2 Q$$

concluding the lemma.

(b) Similar to part (a) we can get that

$$\sum_{s\leq Q}\frac{1}{\phi(s)}\leq\prod_{p\leq Q}\sum_{k=0}^{\infty}\frac{1}{\phi(p^k)}$$
$$=\prod_{p\leq Q}\left(1+\sum_{k=1}^{\infty}\frac{1}{p^{k-1}(p-1)}\right)$$
$$=\prod_{p\leq Q}\left(1+\frac{p}{(p-1)^2}\right).$$

Once again converting to sums of logarithms we get

$$\sum_{s\leq Q}\frac{1}{\phi(s)}\leq\exp\left(\sum_{p\leq Q}\log\left(1+\frac{p}{(p-1)^2}\right)\right)$$
$$\leq\exp\left(\sum_{p\leq Q}\frac{p}{(p-1)^2}\right)$$
$$=\exp\left(\sum_{p\leq Q}\frac{1}{p}+O\left(\sum_{p\leq Q}\frac{1}{(p-1)^2}\right)\right)$$
$$\ll\log Q$$

□

We now have enough to prove the main theorem of this section. The idea is to take our product form of the sum established in Lemma 19 and then some them over primitive characters so we can use some of our previous lemmas. We will also be splitting the interval from 1 to Q into small values, as well as dyadic intervals to help deal with larger values where we will use the multiplicative large sieve.

Proof of Theorem 15. By Lemma 19 we have

$$D_\gamma(MN;q,a) = \frac{1}{\phi(q)} \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \overline{\chi}(a)\Big(\sum_m \alpha_m \chi(m)\Big)\Big(\sum_n \beta_n\chi(n)\Big).$$

Now we wish to reduce the characters $\pmod q$ to primitive characters. Suppose the character χ is induced by a primitive character $\pmod r$. Then $r \mid q$ so let s be such that $q = rs$. Then since $\phi(q) \geq \phi(r)\phi(s)$ we get that

$$\sum_{q\leq Q} \max_{(a,q)=1} |D_\gamma(MN;q,a)| \leq \sum_{q\leq Q} \max_{(a,q)=1} \left|\frac{1}{\phi(q)} \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \overline{\chi}(a)\Big(\sum_m \alpha_m\chi(m)\Big)\Big(\sum_n \beta_n\chi(n)\Big)\right|$$
$$\leq \sum_{s\leq Q}\frac{1}{\phi(s)} \sum_{1<r\leq Q}\frac{1}{\phi(r)} \sideset{}{'}\sum_{\chi \pmod r}\Big|\sum_{(m,s)=1}\alpha_m\chi(m)\Big|\Big|\sum_{(n,s)=1}\beta_n\chi(n)\Big|,$$

noting that $r \neq 1$ since $\chi \neq \chi_0$. We will now separate the sum according to values of r. For example if $r \leq R$ (with R to be chosen later) we can use Lemma 18 to get

$$\sum_{s\leq Q}\frac{1}{\phi(s)} \sum_{1<r\leq R}\frac{1}{\phi(r)} \sideset{}{'}\sum_{\chi \pmod r}\Big|\sum_{(m,s)=1}\alpha_m\chi(m)\Big|\Big|\sum_{(n,s)=1}\beta_n\chi(n)\Big|$$
$$\leq \|\vec{\beta}\|N^{1/2}\Delta^3 \sum_{s\leq Q}\frac{\tau(s)}{\phi(s)} \sum_{1<r\leq R}\frac{r}{\phi(r)} \sideset{}{'}\sum_{\chi \pmod r}\Big|\sum_{(m,s)=1}\alpha_m\chi(m)\Big|.$$

However

$$\Big|\sum_{(m,s)=1}\alpha_m\chi(m)\Big| \leq \Big(\sum_{(m,s)=1}|\alpha_m\chi(m)|^2\Big)^{1/2} \leq \|\vec{\alpha}\|,$$

so by using (12) our partial sum is

$$\ll \|\vec{\alpha}\|\|\vec{\beta}\|N^{1/2}\Delta^3 \sum_{s\leq Q}\frac{\tau(s)}{\phi(s)}\sum_{1<r\leq R} r \ll \|\vec{\alpha}\|\|\vec{\beta}\|N^{1/2}\Delta^3 R^2\log^2 Q$$

Now for values of $R < r \leq Q$ we are going to separate them into intervals $\{P+1,...,2P\}$ where $P = R, 2R, 4R, ...$ giving us at most $O(\log Q)$ such intervals. We will finally be using the multiplicative large sieve that we discussed in Chapter 3. Over the dyadic intervals we get,

$$\sum_{s\leq Q}\frac{1}{\phi(s)} \sum_{P<r\leq 2P}\frac{1}{\phi(r)} \sideset{}{'}\sum_{\chi \pmod r}\Big|\sum_{(m,s)=1}\alpha_m\chi(m)\Big|\Big|\sum_{(n,s)=1}\beta_n\chi(n)\Big|.$$

By Cauchy–Schwarz (4) we get that

$$\sideset{}{'}\sum_{\chi \pmod r}\Big|\sum_{(m,s)=1}\alpha_m\chi(m)\Big|\Big|\sum_{(n,s)=1}\beta_n\chi(n)\Big|$$
$$\leq \Big(\sideset{}{'}\sum_{\chi \pmod r}\Big|\sum_{(m,s)=1}\alpha_m\chi(m)\Big|^2\Big)^{1/2}\Big(\sideset{}{'}\sum_{\chi \pmod r}\Big|\sum_{(n,s)=1}\beta_n\chi(n)\Big|^2\Big)^{1/2},$$

so by theorem 9 we get

$$\begin{aligned}
&\sum_{P<r\le 2P}\frac{1}{\phi(r)}\sideset{}{'}\sum_{\chi \,(\mathrm{mod}\, r)}\Big|\sum_{(m,s)=1}\alpha_m\chi(m)\Big|\Big|\sum_{(n,s)=1}\beta_n\chi(n)\Big| \\
&\qquad\le \frac{1}{P}\sum_{P<r\le 2P}\frac{r^2}{(\phi(r))^2}\sideset{}{'}\sum_{\chi \,(\mathrm{mod}\, r)}\Big|\sum_{(m,s)=1}\alpha_m\chi(m)\Big|\Big|\sum_{(n,s)=1}\beta_n\chi(n)\Big| \\
&\qquad\le \frac{1}{P}(4P^2+M)^{1/2}(4P^2+N)^{1/2}\|\vec{\alpha}\|\|\vec{\beta}\| \\
&\qquad\le \frac{1}{P}(2P+M^{1/2})(2P+N^{1/2})\|\vec{\alpha}\|\|\vec{\beta}\| \\
&\qquad\ll \|\vec{\alpha}\|\|\vec{\beta}\|(P+M^{1/2}+N^{1/2}+P^{-1}M^{1/2}N^{1/2}),
\end{aligned}$$

using $\sqrt{a+b}\le\sqrt{a}+\sqrt{b}$. Now since $R<P\le Q$, equation (13) and that there are at most $O(\log q)$ such sums to add together gives us that the sum for $r>R$ is

$$\ll \|\vec{\alpha}\|\|\vec{\beta}\|(Q+M^{1/2}+N^{1/2}+R^{-1}M^{1/2}N^{1/2})(\log^2 Q).$$

Choosing $R=\Delta^{-1}$ yields the theorem. □

5. The Proof of Bombieri–Vinogradov

We are now in a position to prove our main theorem. The idea may appear to be to use the above theorem with $f(n)=\Lambda(n)$, however the theorem doesn't directly apply to $\Lambda(n)$ since the condition $n\le x$ doesn't allow us to write it as a convolution. Hence we'll need to split up $\Lambda(n)$ into pieces. Some of which we can approximate directly, and several which we will use Theorem 15.

Proof of Theorem 1. We'll start by stating without proof an identity for $\Lambda(n)$ of Vaughan [7] of which you can also see a proof in [4, Page 344].

$$\Lambda(n)=\sum_{\substack{m\le y\\ m|n}}\mu(m)\log\left(\frac{n}{m}\right)-\sum_{\substack{m\le y\\ k\le z\\ km|n}}\mu(m)\Lambda(k)+\sum_{\substack{m>y\\ k>z\\ km|n}}\mu(m)\Lambda(k) \tag{14}$$

Also define

$$\lambda(l)=\log l-\sum_{\substack{k\le x^{1/5}\\ k|l}}\Lambda(k)$$

called an incomplete logarithm. Note that $0\le\lambda(l)\le\log l$, since the sum just subtracts $\log p$ (with multiplicity) for some of the primes dividing l. Also note that if $l\le x^{1/5}$, then $\lambda(l)=0$. We establish the following lemma to split up $\Lambda(n)$ into more tangible pieces.

Lemma 21. *If $x^{1/5}<n\le x$,*

$$\Lambda(n)=\sum_{\substack{m\le x^{1/5}\\ lm=n}}\lambda(l)\mu(m)+\sum_{\substack{x^{1/5}<m\le x^{4/5}\\ lm=n}}\lambda(l)\mu(m).$$

Proof. We will show this by inserting the definition of $\lambda(l)$ and rearranging.

$$\sum_{\substack{m\le x^{1/5}\\ lm=n}} \lambda(l)\mu(m) + \sum_{\substack{x^{1/5}<m\le x^{4/5}\\ lm=n}} \lambda(l)\mu(m)$$

$$= \sum_{\substack{m\le x^{1/5}\\ lm=n}} \Bigg(\log l - \sum_{\substack{k\le x^{1/5}\\ k|l}} \Lambda(k)\Bigg)\mu(m) + \sum_{\substack{x^{1/5}<m\le x^{4/5}\\ lm=n}} \Bigg(\log l - \sum_{\substack{k\le x^{1/5}\\ k|l}} \Lambda(k)\Bigg)\mu(m)$$

$$= \sum_{\substack{m\le x^{1/5}\\ m|n}} \mu(m)\log\left(\frac{m}{n}\right) - \sum_{\substack{m\le x^{1/5}\\ lm=n}} \sum_{\substack{k\le x^{1/5}\\ k|l}} \Lambda(k)\mu(m) + \sum_{\substack{x^{1/5}<m\le x^{4/5}\\ lm=n}} \Bigg(\log l - \sum_{\substack{k\le x^{1/5}\\ k|l}} \Lambda(k)\Bigg)\mu(m).$$

In the second and fourth sums we notice that the conditions

$$lm = n, k \mid l \Leftrightarrow km \mid n,$$

so we get

$$\Lambda(n) = \sum_{\substack{m\le x^{1/5}\\ m|n}} \mu(m)\log\left(\frac{m}{n}\right) - \sum_{\substack{m\le x^{1/5}\\ k\le x^{1/5}\\ km|n}} \mu(m)\Lambda(k)$$
$$+ \sum_{\substack{x^{1/5}<m\le x^{4/5}\\ lm=n}} \mu(m)\log l - \sum_{\substack{x^{1/5}<m\le x^{4/5}\\ k\le x^{1/5}\\ km|n}} \Lambda(k)\mu(m)$$

Now note that if the prime factorization of l is $p_1^{e_1}...p_r^{e_r}$, so

$$\log l = e_1\log p_1 + ... + e_r\log p_r = \sum_{p_i^j|l} \log p_i = \sum_{k|l} \Lambda(k).$$

Hence

$$\sum_{\substack{x^{1/5}<m\le x^{4/5}\\ lm=n}} \mu(m)\log l = \sum_{\substack{x^{1/5}<m\le x^{4/5}\\ lm=n}} \mu(m)\sum_{k|l} \Lambda(k).$$

Therefore we get

$$\Lambda(n) = \sum_{\substack{m\le x^{1/5}\\ m|n}} \mu(m)\log\left(\frac{m}{n}\right) - \sum_{\substack{m\le x^{1/5}\\ k\le x^{1/5}\\ km|n}} \mu(m)\Lambda(k) + \sum_{\substack{x^{1/5}<m\le x^{4/5}\\ k>x^{1/5}\\ km|n}} \Lambda(k)\mu(m).$$

Note that the condition $x^{1/5} < m \le x^{4/5}$ can be replaced with $x^{1/5} < m$ as $m > x^{4/5}, k > x^{1/5} \Rightarrow km > x \ge n$ and so nothing will be added to the sum. Hence by using $y = z = x^{1/5}$ in (14) we get our result. □

If we let $\Lambda'(n), \Lambda''(n)$ be the two sums above for $x^{1/5} < n \le x$, then note that by trivial estimates

$$\sum_{\substack{n \le x^{1/5} \\ n \equiv a \ (\mathrm{mod}\ q)}} \Lambda(n) - \frac{1}{\phi(q)} \sum_{\substack{n \le x^{1/5} \\ (n,q)=1}} \Lambda(n) \ll \sum_{n \le x^{1/5}} \log(n) \le x^{1/5} \log x$$

and so

$$D_\Lambda(x; q, a) = D_{\Lambda'}(x; q, a) + D_{\Lambda''}(x; q, a) + O(x^{1/5} \log x).$$

For $D_{\Lambda'}(x; q, a)$ we have the following lemma.

Lemma 22.

$$\sum_{q \le Q} \max_{(a,q)=1} |D_{\Lambda'}(x; q, a)| \ll Q x^{2/5}.$$

Proof. We start by showing

$$D_\lambda(x; q, a) = \sum_{\substack{n \le x \\ n \equiv a \ (\mathrm{mod}\ q)}} \lambda(l) - \frac{1}{\phi(q)} \sum_{\substack{n \le x \\ (n,q)=1}} \lambda(l) \ll x^{1/5}. \tag{15}$$

The left hand side is

$$\Bigg(\sum_{\substack{n \le x \\ n \equiv a \ (\mathrm{mod}\ q)}} \log n - \frac{1}{\phi(q)} \sum_{\substack{n \le x \\ (n,q)=1}} \log n \Bigg) - \Bigg(\sum_{\substack{n \le x \\ n \equiv a \ (\mathrm{mod}\ q)}} \sum_{\substack{k \le x^{1/5} \\ k|n}} \Lambda(k) - \frac{1}{\phi(q)} \sum_{\substack{n \le x \\ (n,q)=1}} \sum_{\substack{k \le x^{1/5} \\ k|n}} \Lambda(k) \Bigg).$$

Since $\log n$ increases and is positive, the first bracket is bounded above by

$$\log x - \log Kq + \sum_{k=1}^{K} \Bigg(\log kq - \frac{1}{\phi(q)} (\phi(q)) \log(k-1)q \Bigg)$$

which telescopes to give us $\log x$. As for the second bracket, by switching the order of summation we can get

$$\sum_{\substack{n \le x \\ n \equiv a \ (\mathrm{mod}\ q)}} \sum_{\substack{k \le x^{1/5} \\ k|n}} \Lambda(k) - \frac{1}{\phi(q)} \sum_{\substack{n \le x \\ (n,q)=1}} \sum_{\substack{k \le x^{1/5} \\ k|n}} \Lambda(k) = \sum_{k \le x^{1/5}} \Lambda(k) \Bigg(\sum_{\substack{n \le x \\ n \equiv a \ (\mathrm{mod}\ q) \\ n \equiv 0 \ (\mathrm{mod}\ k)}} 1 - \frac{1}{\phi(q)} \sum_{\substack{n \le x \\ (n,q)=1 \\ n \equiv 0 \ (\mathrm{mod}\ k)}} 1 \Bigg)$$

Inside the bracket, over any residue system $rkq < n \le (r+1)kq$ the sums will cancel each other leaving us with a summation over a range of at most kq values, of which the first sum counts at most one of them as does the last. At worst we get the inner bracket being $\ll 1$ Hence we get

$$\sum_{k \le x^{1/5}} \Lambda(k) = \psi(x^{1/5}) \ll x^{1/5}$$

by the prime number theorem. Returning to our main summation, by rearranging the variables by using the condition that $lm = n$, we can get

$$\begin{aligned}
\sum_{q\le Q}\max_{(a,q)=1}|D_{\Lambda'}(x;q,a)| &= \sum_{q\le Q}\max_{a}\left|\sum_{\substack{n\le x\\ n\equiv a \ (\mathrm{mod}\ q)}}\ \sum_{\substack{m\le x^{1/5}\\ m|n\\ lm=n}}\lambda(l)\mu(m)-\frac{1}{\phi(q)}\sum_{\substack{n\le x\\ (n,q)=1}}\ \sum_{\substack{m\le x^{1/5}\\ m|n\\ lm=n}}\lambda(l)\mu(m)\right|\\
&= \sum_{q\le Q}\max_{a}\left|\sum_{\substack{m\le x^{1/5}\\ (m,q)=1}}\mu(m)\sum_{\substack{l\le x/m\\ l\equiv am^{-1} \ (\mathrm{mod}\ q)}}\lambda(l)-\frac{1}{\phi(q)}\sum_{\substack{m\le x^{1/5}\\ (m,q)=1}}\mu(m)\sum_{\substack{l\le x/m\\ (l,q)=1}}\lambda(l)\right|\\
&\le \sum_{q\le Q}\sum_{\substack{m\le x^{1/5}\\ (m,q)=1}}|\mu(m)|\max_{a}\left|\sum_{\substack{l\le x/m\\ l\equiv am^{-1} \ (\mathrm{mod}\ q)}}\lambda(l)-\frac{1}{\phi(q)}\sum_{\substack{l\le x/m\\ (l,q)=1}}\lambda(l)\right|\\
&\le \sum_{q\le Q}\sum_{\substack{m\le x^{1/5}\\ (m,q)=1}}|\mu(m)|x^{1/5}\ll Qx^{2/5}
\end{aligned}$$

using (15).

□

It remains to look at $D_{\Lambda''}(x;q,a)$. We chop up our interval from 1 to x into intervals $y\le n\le (1+\delta)y$ where $x^{-1/5}<\delta\le 1$. There are $O(\delta^{-1}\log x)$ such intervals since if there are J intervals, then

$$\begin{aligned}
&(1+\delta)^{\lfloor J\rfloor}=x^{3/5}\\
&\Rightarrow J\ll\frac{\log x}{\log(1+\delta)}\ll\delta^{-1}\log x.
\end{aligned}$$

We will now change the summation condition

$$lm=n, x^{1/5}<m\le x^{4/5}$$

to the condition

$$lm=n, L<l<(1+\delta)L, M<m<(1+\delta)M$$

where L, M are of the form $(1+\delta)^j$, $x^{1/5}<L, M<x^{4/5}$ and $LM<x$. Note that this won't cover all possible values of n. For example it won't cover $n<x^{1/5}$ however that wasn't part of the original condition anyways, and it won't properly cover (i.e. it won't cover with multiplicity one)

$$(1+\delta)^{-1}x<n<(1+\delta)x$$

Separately taking this piece leads us to the following lemma.

Lemma 23.

$$\sum_{\substack{(1+\delta)^{-1}x<n\le(1+\delta)x\\ n\equiv a \ (\mathrm{mod}\ q)}}\ \sum_{\substack{x^{1/5}<m\le x^{4/5}\\ lm=n}}\lambda(l)\mu(m)-\frac{1}{\phi(q)}\sum_{\substack{(1+\delta)^{-1}x<n\le x\\ (n,q)=1}}\ \sum_{\substack{x^{1/5}<m\le x^{4/5}\\ lm=n}}\lambda(l)\mu(m)\ll\delta q^{-1}x\log^2 x.$$

Proof. By rearranging the conditions on the summation in a similar way to how we did it when showing (15) we get

$$\sum_{\substack{(1+\delta)^{-1}x<n\leq(1+\delta)x \\ n\equiv a \pmod q}} \sum_{\substack{x^{1/5}<m\leq x^{4/5} \\ lm=n}} \lambda(l)\mu(m) - \frac{1}{\phi(q)} \sum_{\substack{(1+\delta)^{-1}x<n\leq x \\ (n,q)=1}} \sum_{\substack{x^{1/5}<m\leq x^{4/5} \\ lm=n}} \lambda(l)\mu(m)$$

$$= \sum_{\substack{x^{1/5}<m\leq x^{4/5} \\ (m,q)=1}} \mu(m) \sum_{\substack{(1+\delta)^{-1}x/m<l\leq(1+\delta)x/m \\ l\equiv am^{-1} \pmod q}} \lambda(l) - \frac{1}{\phi(q)} \sum_{\substack{x^{1/5}<m\leq x^{4/5} \\ (m,q)=1}} \mu(m) \sum_{\substack{(1+\delta)^{-1}x/m<l\leq(1+\delta)x/m \\ (l,q)=1}} \lambda(l).$$

Trivially estimating the inner sums gives us

$$\begin{aligned} &\ll \sum_{\substack{x^{1/5}<m\leq x^{4/5} \\ (m,q)=1}} \frac{|\mu(m)|\delta x \log x}{mq} \\ &\ll \delta q^{-1} x \log^2 x \end{aligned}$$

using $\sum_{m\leq x} m^{-1} \ll \log x$. Note that we may be able to do better with this lemma. In [4, Page 424] they claim that you can get $O(\delta q^{-1} x \log x)$ however I failed to see how you could get around that last $\log x$. However, it minimally impacts the result of the theorem and so we'll leave it as stated. □

Letting $D(LM; q, a)$ denote the sum

$$\sum_{\substack{l,m \\ L<l\leq(1+\delta)L \\ M<m\leq(1+\delta)M \\ lm\equiv a \pmod q}} \lambda(l)\mu(m) - \frac{1}{\phi(q)} \sum_{\substack{l,m \\ L<l\leq(1+\delta)L \\ M<m\leq(1+\delta)M \\ (lm,q)=1}} \lambda(l)\mu(m),$$

we have

$$D_{\Lambda''}(x; q, a) = \sum_{L,M} D(LM; q, a) + O(\delta q^{-1} x \log^2 x). \tag{16}$$

However, theorem 15 applies with $\beta(l) = \lambda(l)$, $\alpha(m) = \mu(m)$ since the prime number theorem for arithmetic progressions (specifically Siegel's Theorem) says that equation (10) is satisfied for $\lambda(l)$, if we let $\Delta = \log^{-B} x$. Also note that if we let $Q = \Delta x^{1/2}$, then

$$\|\vec{\alpha}\|^2 \leq \delta M, \|\vec{\beta}\|^2 \leq \delta L \log^2 x \text{ and } \log Q \ll \log x,$$

so we get

$$\begin{aligned} \sum_{q\leq Q} \max_{(a,q)=1} |D(LM; q, a)| &\ll (\delta M)^{1/2}(\delta L)^{1/2} \log x (\Delta(LM)^{1/2} + L^{1/2} + M^{1/2} + \Delta x^{1/2}) \log^2 x \\ &\ll \delta \Delta x \log^3 x \end{aligned}$$

using that $L, M, LM < x$. Summing over the $O(\delta^{-2} \log^2 x)$ such L, M such that $LM < x$, and including the error estimate from equation (16) gives us

$$\sum_{q\le Q}\max_{(a,q)=1}|D_{\Lambda''}(x;q,a)| \ll \delta^{-1}\Delta x\log^5 x + \delta x\log^3 x$$

using that $\sum_{q\le Q} q^{-1} \ll \log Q \le \log x$. Set $\delta = \Delta^{1/2}\log^2 x$, which at most 1 for $B \ge 4$. The bound then becomes $\Delta^{1/2}x\log^4 x$.

Therefore we conclude that

$$\sum_{q\le Q}\max_{(a,q)=1}\left|\psi(x;q,a) - \frac{1}{\phi(q)}\sum_{\substack{n\le x\\(n,q)=1}}\Lambda(n)\right| = \sum_{q\le Q}\max_{(a,q)=1}|D_\Lambda(x;q,a)|$$
$$\ll \Delta^{1/2}x\log^4 x.$$

However with the exception of the finite number of primes dividing q with are not counted in the sum, we have

$$\sum_{\substack{n\le x\\(n,q)=1}}\Lambda(n) = \psi(x) = x + O(x\log^{-C} x)$$

for any constant C by the prime number theorem. Hence we have

$$\sum_{q\le Q}\max_{(a,q)=1}\left|\psi(x;q,a) - \frac{x}{\phi(q)}\right| \ll_A x(\log x)^{-\frac{B}{2}+4}$$

which gives us Bombieri–Vinogradov for $B(A) = 2A + 8$.

□

We have been following the proof in [4, Chapter 17] however our result is slightly different. They got $B = 2A + 6$ however they appear to have an error. They say that there are $O(\delta^{-1})$ intervals dividing $x^{1/5} < n \le x^{4/5}$ which is clearly false when δ is constant and appears to be false in general. Hence our $O(\delta^{-1}\log x)$ gives us an extra logarithm in our theorem.

6. Notes, Additional Results and References

In Bombieri's original paper [1], he showed the constant $B = 3A + 23$. Note that it's possible to do better for the value of B than we did. The following has been shown in [2, Chapter 28].

$$\sum_{q\le Q}\max_{(a,q)=1}\max_{y\le x}|E(y;q,a)| \ll x^{1/2}Q\log^5 x \tag{17}$$

where $x^{1/2}\log^{-A} x \le Q \le x^{1/2}$. which is even stronger with $B = A + 5$.

Even under GRH we don't get what's considered to be the correct error term for $\psi(x;q,a)$. It's conjectured [6, Conjecture 13.9] that

$$E(x;q,a) \ll_\epsilon x^{1/2+\epsilon}/q^{1/2}$$

which since

$$\sum_{q\le x^{1-\epsilon}} q^{-1/2} \ll x^{1/2-\epsilon/2}$$

would imply the conjecture by Elliott–Halberstam that we can increase the range of values of Q all the way to $x^{1-\epsilon}$.

It might be thought that we can even do better than the Elliott–Halberstam conjecture, perhaps even having q up to x itself. However a result of Friedlander and Granville [3] showed the following.

Theorem 24. *For any $N > 0$, there exists arbitrarily large value of a and x for which*

$$\sum_{\substack{q<x/log^N x\\(q,a)=1}} \left|\psi(x;q,a) - \frac{x}{\phi(q)}\right| \gg_N x.$$

By using Euler Summation we can use Theorem 1 to say something similar about the prime counting function for arithmetic progressions.

Theorem 25. *Let $A > 0$ be fixed, and let $E(x;q,a) = \pi(x;q;a) - \frac{li(x)}{\phi(q)}$, then*

$$\sum_{q\leq Q} \max_{(a,q)=1} |E(x;q,a)| \ll_A x\log^{-A} x, \tag{18}$$

where $Q = x^{1/2}\log^{-B} x$, for B depending on A.

By mimicking the proof of Bombieri–Vinogradov we can prove a similar type result called the Barban–Davenport–Halberstam Theorem with the proof shown in [4, Page 424].

Theorem 26.

$$\sum_{q\leq Q} \sum_{\substack{a \pmod q\\(a,q)=1}} \left(\psi(x;q,a) - \frac{x}{\phi(q)}\right)^2 \ll_A x\log^{-A} x,$$

for and $A > 0$ where $Q = x^{1/2}\log^{-B} x$ with B depending on A.

References

[1] E. Bombieri, ”On the large sieve”. *Mathematica* **12** (1965), 201-225.

[2] H. Davenport, *Multiplicative Number Theory Third Edition*, Graduate Texts in Math., New York: Springer-Verlag, Vol. 74 (2000).

[3] J. Friedlander and A. Granville, ”Limitations to the equi-distribution of primes I”. *The Annals of Math.*, **129** (1989), 362-382.

[4] H.Iwaniec and E Kowalski, *Analytic Number Theory*, Colloquim Publications (American Mathematical Society) Vol. 53 (2004).

[5] H. L. Montgomery and R. C. Vaughan, ”Hilbert’s inequality”. *J. London Math. Soc.* (2) **8** (1974), 73–82.

[6] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge University Press (2007).

[7] R. C. Vaughan, ”Sommes trigonometriques sur les numbres premiers”. *C. R. Acad. Sci Paris Ser.* A-B **285** (1977), A981–A983.