13. THE LARGE SIEVE.

13.1. Introduction. We have seen that the Generalized Riemann Hypothesis implies that

$$\pi(x;q,a) \sim \frac{\pi(x)}{\phi(q)}$$

whenever $(a,q)=1$ for x a little bigger than q^2. In fact this can be proved to hold except for a few rare moduli. A precise statement of the *Bombieri-Vinogradov theorem* is

$$\sum_{q\le Q} \max_{(a,q)=1} \left| \pi(x;q,a) - \frac{\pi(x)}{\phi(q)} \right| \ll_A \frac{x}{(\log x)^A}. \tag{13.1}$$

Here, for any $A>0$ we can take $Q=\sqrt{x}/(\log x)^{B(A)}$ where $B(A)$ is a constant that depends only on A. If we simply have a bound like $\pi(x;q,a) \ll \pi(x)/\phi(q)$ then the left side here is $\ll x\log x$; thus the Bombieri-Vinogradov theorem improves on this "trivial" estimate by an arbitrary power of $\log x$. The formulation of the result seems complicated, but this is useful for many applications: Its range, with q nearly up to $\sqrt{x}$ means that it can substitute for the Generalized Riemann Hypothesis in many important arguments. Early proofs of the Bombieri-Vinogradov theorem relied on the fact that few L-functions have zeros close to 1, by getting bounds for the number of zeros $\sigma+it$ with $\sigma>\alpha$ and $|t|<T$, over all characters $\chi \pmod q$ with $q\le Q$. Later proofs found that results such as (13.1) hold for many sequences which also satisfy a "Siegel-Walfisz theorem". In all of these proofs Q is restricted to be a little less than $\sqrt{x}$ and this barrier is one of the most important (and difficult) in the subject. For a fixed a one can get non-trivial results a little beyond $\sqrt{x}$ but these are not yet satisfactory for most applications. The Elliott-Halberstam conjecture claims that (13.1) holds for $Q=x^{1-\epsilon}$, for any fixed $\epsilon, A>0$. (The Bombieri-Vinogradov theorem is also valid replacing $|\pi(x;q,a)-\pi(x)/\phi(q)|$ by $\max_{y\le x}|\pi(y;q,a)-\pi(y)/\phi(q)|$ as we will prove below).

The Barban-Davenport-Halberstam theorem accounts for primes $\pmod q$ which are just a little bigger than q in a more conventional average sense:

$$\sum_{q\le Q} \sum_{\substack{1\le a\le q\\(a,q)=1}} \left| \pi(x;q,a) - \frac{\pi(x)}{\phi(q)} \right|^2 \sim Qx$$

for $x/(\log x)^A < Q < x$. This result is somewhat less applicable but is easier to prove.

13.2. Discussion. For a typical arithmetic function $\beta_n \in \mathbb{C}$, we might expect that $\sum_{n\leq x,\ n\equiv a \pmod q} \beta_n$ is about $\frac{1}{\phi(q)}\sum_{n\leq x,\ (n,q)=1} \beta_n$ whenever $(a,q)=1$; and so we study the difference. For most applications, it suffices to obtain results of type

$$\left| \sum_{\substack{n\leq x \\ n\equiv a \pmod q}} \beta_n - \frac{1}{\phi(q)} \sum_{\substack{n\leq x \\ (n,q)=1}} \beta_n \right| \ll \frac{1}{\phi(q)}\ \|\beta\|\ \frac{x^{\frac{1}{2}}}{(\log x)^A}$$

for any $A>0$, with the implied constant in $\ll$ depending only on A, where

$$\|\beta\|^2 := \sum_{n\leq x} |\beta_n|^2.$$

We want this result to be valid uniformly in q in a large range. A good example to keep in mind is where $\beta_n = \log n$ if n is prime, and $=0$ otherwise. In this case we have proved the above estimate in the range $q\leq (\log x)^B$, and so such an estimate is called of "Siegel-Walfisz type". As we saw with the primes it is difficult to prove such results in a wider range for all q, but it may be possible for "almost all" q. Thus, summing the above estimate, we might ask for a result of the form

$$\sum_{q<Q} \max_{(a,q)=1} \left| \sum_{\substack{n\leq x \\ n\equiv a \pmod q}} \beta_n - \frac{1}{\phi(q)} \sum_{\substack{n\leq x \\ (n,q)=1}} \beta_n \right| \ll \|\beta\|\ \frac{x^{\frac{1}{2}}}{(\log x)^A},$$

for appropriately large Q which is said to be of "Bombieri-Vinogradov" type. Another idea is to ask only for almost all q and almost all a, that is a result of the kind

$$\sum_{q<Q} \sum_{\substack{1\leq a\leq q \\ (a,q)=1}} \left| \sum_{\substack{n\leq x \\ n\equiv a \pmod q}} \beta_n - \frac{1}{\phi(q)} \sum_{\substack{n\leq x \\ (n,q)=1}} \beta_n \right|^2 \ll \|\beta\|^2\ \frac{x}{(\log x)^A},$$

for appropriately large Q which is said to be of "Barban-Davenport-Halberstam" type. In certain special circumstances one can even obtain an asymptotic for this sum.

The most difficult question is to obtain a good upper bound for almost all q, for a fixed a. Here we seek to estimate

$$\sum_{\substack{q\leq Q \\ (q,a)=1}} \left| \sum_{\substack{n\leq x \\ n\equiv a \pmod q}} \beta_n - \frac{1}{\phi(q)} \sum_{\substack{n\leq x \\ (n,q)=1}} \beta_n \right|,$$

which we will discuss later.

13.3. The large sieve. We begin with a result from linear algebra:

The Duality Principle. *Let $x_{m,n} \in \mathbb{C}$ for $1 \le m \le M,\ 1 \le n \le N$. For any constant c we have*

$$\sum_n \left| \sum_m a_m x_{m,n} \right|^2 \le c\|a\|^2$$

for all $a_m \in \mathbb{C},\ 1 \le m \le M$ if and only if

$$\sum_m \left| \sum_n b_n x_{m,n} \right|^2 \le c\|b\|^2$$

for all $b_n \in \mathbb{C},\ 1 \le n \le N$. (Here $\|a\|^2 := \sum_n |a_n|^2$.)

Proof. Assume that the first inequality is true. Given $b_n \in \mathbb{C},\ 1 \le n \le N$ define $a_m = \sum_n b_n x_{m,n}$, so that

$$\sum_m \left| \sum_n b_n x_{m,n} \right|^2 = \sum_m \overline{a}_m \sum_n b_n x_{m,n} = \sum_n b_n \sum_m \overline{a}_m x_{m,n},$$

so by the Cauchy-Schwarz inequality, the above squared is

$$\|a\|^4 \le \|b\|^2 \sum_n \left| \sum_m \overline{a}_m x_{m,n} \right|^2 \le \|b\|^2 \cdot c\|a\|^2,$$

and the result follows. The reverse implication is completely analogous.

Proposition 13.1. *Let $a_n, M+1 \le n \le M+N$ be a set of complex numbers, and $x_r, 1 \le r \le R$ be a set of real numbers. Let $\delta := \min_{r \ne s} \|x_r - x_s\| \in [0, 1/2]$, where $\|t\|$ denotes the distance from t to the nearest integer. Then*

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

where $e(t) = e^{2i\pi t}$.

Proof. For any $b_r \in \mathbb{C},\ 1 \le r \le R$, we have

$$\sum_n \left| \sum_r b_r e(nx_r) \right|^2 = \sum_{r,s} b_r \overline{b}_s \sum_{n=M+1}^{M+N} e(n(x_r - x_s)) = N\|b\|^2 + E,$$

since the inner sum is N if $r = s$, where, for $L := M + \frac{1}{2}(N+1)$,

$$E \le \sum_{r \ne s} b_r \overline{b}_s e(L(x_r - x_s)) \frac{\sin(\pi N(x_r - x_s))}{\sin(\pi(x_r - x_s))}.$$

Taking absolute values we obtain

$$|E| \le \sum_{r\neq s} \frac{|b_r \overline{b}_s|}{|\sin(\pi(x_r - x_s))|} \le \sum_{r\neq s} \frac{|b_r \overline{b}_s|}{2\|x_r - x_s\|} \le \sum_r |b_r|^2 \sum_{s\neq r} \frac{1}{2\|x_r - x_s\|}$$

by the Cauchy-Schwarz inequality. Now for each x_r the nearest two x_s are at distance at least δ away, the next two at distance at least 2δ away, etc. Therefore,

$$|E| \le \sum_r |b_r|^2 \sum_{j=1}^{[1/\delta]} \frac{2}{2j\delta} \le \|b\|^2 \frac{\log(e/\delta)}{\delta},$$

so that

$$\sum_n \left| \sum_r b_r e(nx_r) \right|^2 \le \left(N + \frac{\log(e/\delta)}{\delta} \right) \|b\|^2.$$

The result, with $1/\delta - 1$ replaced by $\log(e/\delta)/\delta$, follows by the duality principle.

We now show how to get a constant $\ll N + 1/\delta$: Let $c_r = b_r e(Mx_r)$ so that

$$\begin{aligned}
\sum_{n=M+1}^{M+N} \left| \sum_r b_r e(nx_r) \right|^2 &\le \sum_{n=1}^{N} \left| \sum_r c_r e(nx_r) \right|^2 e^{\pi(1-(n/N)^2)} \\
&\le e^{\pi} \sum_{r,s} c_r \overline{c}_s \sum_{n\in\mathbb{Z}} e^{-\pi(n/N)^2} e(n(x_r - x_s)) \\
&= e^{\pi} \sum_{r,s} c_r \overline{c}_s \cdot N \sum_{n\in\mathbb{Z}} e^{-\pi N^2 (n + x_r - x_s)^2} \\
&= e^{\pi} \sum_{r,s} c_r \overline{c}_s \cdot N \{ e^{-\pi N^2 \|x_r - x_s\|^2} + O(e^{-\pi N^2/4}) \}
\end{aligned}$$

by Lemma 9.2. Now applying the Cauchy-Schwarz inequality as before, and the same analysis of the sequence of values of $\|x_r - x_s\|$ for each fixed r, this is

$$\begin{aligned}
&\le Ne^{\pi} \sum_r |c_r|^2 \sum_s \{ e^{-\pi N^2 \|x_r - x_s\|^2} + O(e^{-\pi N^2/4}) \} \\
&\le Ne^{\pi} \sum_r |b_r|^2 \left(\sum_{k\in\mathbb{Z}} e^{-\pi(\delta k N)^2} + O((1/\delta) e^{-\pi N^2/4}) \right) \\
&\le e^{\pi} \|b\|^2 \left(N + 1/\delta + O((N/\delta) e^{-\pi N^2/4}) \right).
\end{aligned}$$

The result, up to the constant, follows from the duality principle. (One can get the result claimed here by following the proof of Theorem 7.7 in [IK].)

Exercises

13.1a. Suppose that a_n are given. Given x_j define $y_j(t) = x_j + t$ (where $t \in \mathbb{R}$).

a) Show that if $\delta := \min_{r\neq s} \|x_r - x_s\|$ then $\min_{r\neq s} \|y_r(t) - y_s(t)\| = \delta$.

b) Prove that $\int_0^1 \left| \sum_{n=M+1}^{M+N} a_n e(ny_r(t)) \right|^2 dt = \|a\|^2$.

c) Deduce that for any $\delta > 0$ there exist x_r such that $\sum_r \left| \sum_{n=M+1}^{M+N} a_n e(nx_r) \right|^2 \ge (1/\delta - 1)\ \|a\|^2$.

13.1b. Suppose that x_j are given. For any given M, N select complex numbers $a_n, M < n \le M + N$, each of absolute value 1, such that $\sum_r \left| \sum_{n=M+1}^{M+N} a_n e(nx_r) \right|^2 \ge N^2 = N\|a\|^2$.

Proposition 13.2. *Let $\beta_n, M+1 \le n \le M+N$ be a set of complex numbers. Then*

$$\sum_{q\le Q} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \left| \sum_{n=M+1}^{M+N} \beta_n \chi(n) \right|^2 \le (N+Q^2)\|\beta\|^2. \tag{13.2}$$

Proof. By (3.5.1) we have

$$\sum_{n=M+1}^{M+N} \beta_n \chi(n) = \frac{1}{g(\overline{\chi})} \sum_{a \pmod q} \overline{\chi}(a) \sum_{n=M+1}^{M+N} \beta_n e\left(\frac{an}{q}\right).$$

By (3.5.2) we therefore deduce that

$$\begin{aligned} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \left| \sum_{n=M+1}^{M+N} \beta_n \chi(n) \right|^2 &\le \frac{1}{q} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \left| \sum_{a \pmod q} \overline{\chi}(a) \sum_{n=M+1}^{M+N} \beta_n e\left(\frac{an}{q}\right) \right|^2 \\ &= \frac{\phi(q)}{q} \sum_{\substack{a \pmod q \\ (a,q)=1}} \left| \sum_{n=M+1}^{M+N} \beta_n e\left(\frac{an}{q}\right) \right|^2 \end{aligned}$$

so that exercise 13.2a.a implies

$$\sum_{q\le Q} \frac{q}{\phi(q)} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \left| \sum_{n=M+1}^{M+N} \beta_n \chi(n) \right|^2 \le (N+Q^2)\|\beta\|^2.$$

Exercises

13.2a. Let $a_n, M+1 \le n \le M+N$ be a set of complex numbers. Deduce from Proposition 13.1 that

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

13.3a.a) Recall that $\phi(q) \gg q/\log\log Q$ for all $q \le Q$. By cutting the sum over q up into dyadic intervals, deduce from (13.2) that

$$\sum_{R<q\le Q} \frac{1}{\phi(q)} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \left| \sum_{n=M+1}^{M+N} \beta_n \chi(n) \right|^2 \ll \left(\frac{N}{R} + Q\right) \|\beta\|^2 \log\log Q.$$

b) Suppose that α_ℓ is supported on an interval of length L, where $LN = x$. Use the Cauchy-Schwarz inequality to deduce from (13.2) that

$$\sum_{q\le Q} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \left| \sum_\ell \alpha_\ell \chi(\ell) \right| \cdot \left| \sum_n \beta_n \chi(n) \right| \le (x^{1/2} + (L+N)^{1/2} Q + Q^2)\ \|\alpha\|\ \|\beta\|.$$

c) By combining these methods deduce that

$$\sum_{R<q\le Q} \frac{1}{\phi(q)} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \left| \sum_\ell \alpha_\ell \chi(\ell) \right| \cdot \left| \sum_n \beta_n \chi(n) \right| \ll \left(\frac{x^{1/2}}{R} + (L+N)^{1/2} \log Q + Q\right)\ \|\alpha\|\ \|\beta\| \log\log Q.$$

Proposition 13.3. *Let $\beta_n, M+1 \le n \le M+N$ be a set of complex numbers such that $\beta_n = 0$ if n has a prime factor $< Q$. Then*

$$\sum_{q\le Q} \log\frac{Q}{q} \sum_{\substack{\chi \pmod q\\ \chi \text{ primitive}}} \left| \sum_{n=M+1}^{M+N} \beta_n \chi(n) \right|^2 \le (N+Q^2)\|\beta\|^2.$$

Let $\beta_n = 1$ if n is prime and $> Q$, for $n \in [M+1, M+N]$, and let $\beta_n = 0$ otherwise. Taking $Q = (N/\log N)^{1/2}$ in Proposition 13.3, and bounding the left side by the $q = 1$ term, we obtain:

The Brun-Titchmarsh Theorem. *For any $M, N \ge 1$ we have*

$$\pi(M+N) - \pi(M) \le \frac{2N}{\log N} + O\left(\frac{N \log\log N}{(\log N)^2}\right).$$

Remark. Note that the upper bound given here depends only on the number of terms being considered, and is uniform in M. It is of great interest to determine the smallest constant that can replace the 2 in the Brun-Titchmarsh theorem. From the prime number theorem that the 2 cannot be replaced by any number smaller than 1. In fact the proof we gave counted the number of integers in this interval with no prime factor $< Q$. An old conjecture of Hardy and Littlewood stated that

$$\max_M \#\{n \in [M+1, M+N] :\ p|n \implies p > N\} \le \pi(N).$$

This was proved to be wrong by Hensley and Richards (though not necessarily by a lot).

Proof of Proposition 13.3. By exercise 13.3b the left side above is, writing $\ell = qr$,

$$\le \sum_{\ell\le Q} \sum_{\substack{q|\ell\\ (q,\ell/q)=1}} \frac{q}{\phi(q)} \frac{\mu(\ell/q)^2}{\phi(\ell/q)} \sum_{\substack{\chi \pmod q\\ \chi \text{ primitive}}} \left| \sum_{n=M+1}^{M+N} \beta_n \chi(n) \right|^2.$$

Now let $\psi \pmod \ell$ be the character induced by $\chi \pmod q$. From the discussion in section 3.5 we have $g(\psi) = \mu(\ell/q)\chi(\ell/q)g(\chi)$, so that if $(q, \ell/q) > 1$ then $g(\psi) = 0$, and otherwise $|g(\psi)|^2 = q\mu(\ell/q)^2$ and $\phi(q)\phi(\ell/q) = \phi(\ell)$. Therefore the last line equals

$$\sum_{\ell\le Q} \frac{1}{\phi(\ell)} \sum_{\psi \pmod \ell} |g(\psi)|^2 \left| \sum_{n=M+1}^{M+N} \beta_n \psi(n) \right|^2,$$

using the fact that $\beta_n\psi(n) = \beta_n\chi(n)$ by the hypothesis on β_n. Then, by (3.5.1) we see that this equals

$$\sum_{\ell\le Q} \sum_{\substack{a \pmod \ell\\ (a,\ell)=1}} \left| \sum_{n=M+1}^{M+N} \beta_n e\left(\frac{an}{q}\right) \right|^2,$$

which gives the result by exercise 13.2a.

Exercises

13.3b. Prove that for any $m, N \geq 1$ we have

$$\frac{m}{\phi(m)} \sum_{\substack{r \leq N \\ (r,m)=1}} \frac{\mu(r)^2}{\phi(r)} \geq \log N.$$

(Hint: Expand each term as a sum of reciprocals of integers.)

13.4. Barban-Davenport-Halberstam, I.

Definition. *The sequence $\beta_n, n \leq N$ is said to satisfy a Siegel-Walfisz condition if for any $d \geq 1, q \geq 1$ and a with $(k,a) = 1$ we have*

$$\left| \sum_{\substack{n \equiv a \pmod q \\ (n,d)=1}} \beta_n - \frac{1}{\phi(q)} \sum_{n:\ (n,dq)=1} \beta_n \right| \ll \tau(d)^{B_1} \|\beta\| \frac{N^{\frac{1}{2}}}{(\log N)^C}.$$

Here $\tau(d)$ is the number of divisors of d.

Exercises

13.4a.a) Suppose that χ is a character $\pmod q$. Prove that for any integer $d \neq 0$ we have

$$\sum_{(n,d)=1} \beta_n \chi(n) = \sum_{(a,q)=1} \chi(a) \left(\sum_{\substack{n \equiv a \pmod q \\ (n,d)=1}} \beta_n - \frac{1}{\phi(q)} \sum_{n:\ (n,dq)=1} \beta_n \right).$$

b) Deduce that if β_n satisfies the Siegel-Walfisz condition then

$$\left| \sum_{(n,d)=1} \beta_n \chi(n) \right| \ll \phi(q) \tau(d)^{B_1} \|\beta\| \frac{N^{\frac{1}{2}}}{(\log N)^C}.$$

Theorem 13.1. *Suppose that the sequence of complex numbers $\beta_n, n \leq x$ satisfies a Siegel-Walfisz condition. For any $A > 0$ there exists $B = B(A) > 0$ such that*

$$\sum_{q \leq Q} \sum_{a:\ (a,q)=1} \left| \sum_{n \equiv a \pmod q} \beta_n - \frac{1}{\phi(q)} \sum_{(n,q)=1} \beta_n \right|^2 \ll \|\beta\|^2 \frac{x}{(\log x)^A} \tag{13.3}$$

where $Q = x/(\log x)^B$.

Proof. We begin with the identity

$$\sum_{a:\ (a,q)=1} \left| \sum_{n \equiv a \pmod q} \beta_n - \frac{1}{\phi(q)} \sum_{(n,q)=1} \beta_n \right|^2 = \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0} \left| \sum_n \beta_n \chi(n) \right|^2.$$

Now if $\chi \pmod q$ is induced by $\psi \pmod m$ then $\sum_n \beta_n\chi(n) = \sum_{n:\ (n,q/m)=1} \beta_n\psi(n)$, and $\phi(q) \geq \phi(m)\phi(q/m)$ so that the left side of (13.3) is

$$= \sum_{q\leq Q} \frac{1}{\phi(q)} \sum_{\substack{m|q\\ m>1}} \sum_{\substack{\psi \pmod m\\ \psi \text{ primitive}}} \left| \sum_{n:\ (n,q/m)=1} \beta_n\psi(n)\right|^2$$

$$\leq \sum_{r\leq Q} \frac{1}{\phi(r)} \sum_{1<m\leq Q/r} \frac{1}{\phi(m)} \sum_{\substack{\psi \pmod m\\ \psi \text{ primitive}}} \left| \sum_{n:\ (n,r)=1} \beta_n\psi(n)\right|^2$$

From exercise 13.3a.a we deduce that this sum restricted to $m > M := (\log x)^{B+1}$ is

$$\ll \sum_{r\leq Q} \frac{1}{\phi(r)} \left(\frac{x}{M} + \frac{Q}{r}\right) \log\log Q\ \|\beta\|^2 \ll Q\log\log Q\ \|\beta\|^2$$

For the sum restricted to $m \leq M$ we use the above identity to get the upper bound

$$\leq \sum_{r\leq Q} \frac{1}{\phi(r)} \sum_{1<m\leq M} \frac{1}{\phi(m)} \sum_{\substack{\psi \pmod m\\ \psi\neq\psi_0}} \left| \sum_{n:\ (n,r)=1} \beta_n\psi(n)\right|^2$$

$$= \sum_{r\leq Q} \frac{1}{\phi(r)} \sum_{1<m\leq M} \sum_{a:\ (a,m)=1} \left| \sum_{\substack{n\equiv a \pmod m\\ (n,r)=1}} \beta_n - \frac{1}{\phi(m)} \sum_{(n,mr)=1} \beta_n\right|^2$$

and this is

$$\ll \sum_{r\leq Q} \frac{\tau(r)^{2B_1}}{\phi(r)} M^2\|\beta\|^2 \frac{x}{(\log x)^{2C}} \ll \|\beta\|^2 \frac{x}{(\log x)^A}$$

by the Siegel-Walfisz condition, provided $2C \geq A + 2B + 2 + 2^{2B_1}$. The result follows by taking $B > A$

Theorem 13.2. *Suppose that we have two sequences of complex numbers $\alpha_\ell,\ L < \ell \leq 2L$, and $\beta_n, N < n \leq 2N$ which satisfies the Siegel-Walfisz condition. For any $A > 0$ there exists $B = B(A) > 0$ such that if $f(r) = \sum_{\ell n=r} \alpha_\ell\beta_n$ and $x = LN$ then*

$$\sum_{q\leq Q} \max_{a:\ (a,q)=1} \left| \sum_{n\equiv a \pmod q} f(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} f(n)\right| \ll \|\alpha\|\|\beta\| \frac{x^{1/2}}{(\log x)^A} \tag{13.4}$$

where $Q = x^{1/2}/(\log x)^B$, provided $N \geq \exp((\log x)^\epsilon)$ and $L \geq (\log x)^{2B+4}$.

Proof. We begin by observing that

$$\sum_{r\equiv a \pmod q} f(r) - \frac{1}{\phi(q)} \sum_{(r,q)=1} f(r) = \frac{1}{\phi(q)} \sum_{\chi\neq\chi_0} \overline{\chi}(a) \left(\sum_m \alpha_m \chi(m)\right) \left(\sum_n \beta_n \chi(n)\right).$$

In absolute value this is, proceeding as in the proof of Theorem 13.1,

$$\begin{aligned}
&\leq \frac{1}{\phi(q)} \sum_{\chi\neq\chi_0} \left|\sum_m \alpha_m \chi(m)\right| \cdot \left|\sum_n \beta_n \chi(n)\right| \\
&\leq \sum_{rm=q} \frac{1}{\phi(r)} \frac{1}{\phi(m)} \sum_{\substack{\psi \pmod m \\ \psi \text{ primitive}}} \left|\sum_{\ell:\ (\ell,r)=1} \alpha_\ell \psi(\ell)\right| \cdot \left|\sum_{n:\ (n,r)=1} \beta_n \psi(n)\right|
\end{aligned}$$

The sum of this over $q \leq Q$, restricted to $m > M := (\log x)^{B+1}$ is, by exercise 13.3a.c,

$$\begin{aligned}
&\ll \sum_{r\leq Q} \frac{1}{\phi(r)} \left(\frac{x^{1/2}}{M} + (L+N)^{1/2} \log Q + \frac{Q}{r}\right) \|\alpha\|\ \|\beta\| \log\log Q \\
&\ll \left(\frac{x^{1/2}}{M} \log Q + (L+N)^{1/2} (\log Q)^2 + Q\right) \|\alpha\|\ \|\beta\| \log\log Q \\
&\ll Q \|\alpha\|\ \|\beta\| \log\log Q.
\end{aligned}$$

For the rest, using exercise 13.4a.b, and then the Cauchy-Schwarz inequality with (13.2), we obtain

$$\begin{aligned}
&\ll \sum_{r\leq Q} \frac{\tau(r)^{B_1}}{\phi(r)} \sum_{m\leq M} \sum_{\substack{\psi \pmod m \\ \psi \text{ primitive}}} \left|\sum_{\ell:\ (\ell,r)=1} \alpha_\ell \psi(\ell)\right| \cdot \|\beta\| \frac{N^{\frac{1}{2}}}{(\log N)^C} \\
&\ll M(L^{1/2} + M)\|\alpha\| \cdot \|\beta\| N^{\frac{1}{2}} \frac{(\log Q)^{2^{B_1}}}{(\log N)^C} \ll Q\|\alpha\|\ \|\beta\|
\end{aligned}$$

as $M \ll L^{1/2}$ and $\log N \geq (\log x)^\epsilon$ for $\epsilon C = 2B + 1 + 2^{B_1}$.

13.5. The Bombieri-Vinogradov theorem. We will prove (13.1) in the following form:

Theorem 13.3. *For any $A > 0$ there exists $B = B(A) > 0$ such that*

$$\sum_{q\leq Q} \max_{(a,q)=1} \left|\psi(x;q,a) - \frac{\psi(x)}{\phi(q)}\right| \ll_A \frac{x}{(\log x)^A}. \tag{13.5}$$

where $Q = x^{\frac{1}{2}}/(\log x)^B$.

The idea in the proof is to repeatedly use Theorem 13.2 after we have written $\Lambda(n)$ as a sum of such convolutions: Let $M(s) = \sum_{m\leq\sqrt{x}} \mu(m)/m^s$. As $\zeta(s)^{-1} = \sum_{m\geq 1} \mu(n)/n^s$ we see that the coefficient of $1/n^s$ in $\zeta(s)M(s)-1$ is 0 for $n \le \sqrt{x}$; and similarly the coefficients of $-\zeta'(s)/\zeta(s) - R(s)$ where $R(s) = \sum_{r\le\sqrt{x}} \Lambda(r)$. Multiplying the two together gives a Dirichlet series in which the coefficient of $1/n^s$ is 0 for $n \le x$. In particular we deduce that if $\sqrt{x} < n \le x$ then $\Lambda(n)$, the coefficient of $1/n^s$ in $-\zeta'(s)/\zeta(s)-R(s)$, equals the coefficient of $1/n^s$ in $(-\zeta'(s)/\zeta(s) - R(s))\zeta(s)M(s) = -\zeta'(s)M(s) - \zeta(s)M(s)R(s)$. Therefore

$$-\Lambda(n) = f_1(n) + f_2(n),$$

where

$$f_1(n) = \sum_{\substack{m\le\sqrt{x}\\ m|n}} \mu(m)\log(n/m) \quad\text{and}\quad f_2(n) = \sum_{\substack{m,r\le\sqrt{x}\\ mr|n}} \mu(m)\Lambda(r).$$

Now

$$\sum_{\substack{\sqrt{x}<n\le x\\ n\equiv a \pmod q}} f_1(n) = \sum_{\substack{m\le\sqrt{x}\\ (m,q)=1}} \mu(m) \sum_{\substack{\sqrt{x}<n\le x\\ n\equiv a \pmod q\\ m|n}} \log(n/m) = \sum_{\substack{m\le\sqrt{x}\\ (m,q)=1}} \mu(m) \sum_{\substack{\sqrt{x}/m<k\le x/m\\ k\equiv a/m \pmod q}} \log k$$

$$= \sum_{\substack{m\le\sqrt{x}\\ (m,q)=1}} \mu(m) \left(\frac{1}{q} \sum_{\sqrt{x}/m<k\le x/m} \log k + O(\log x) \right).$$

Summing this up over all a with $(a,q)=1$ and dividing by $\phi(q)$ we deduce that

$$\sum_{q\le Q} \max_{(a,q)=1} \left| \sum_{\substack{\sqrt{x}<n\le x\\ n\equiv a \pmod q}} f_1(n) - \frac{1}{\phi(q)} \sum_{\substack{\sqrt{x}<n\le x\\ (n,q)=1}} f_1(n) \right| \ll Q\sqrt{x}\log x. \tag{13.6}$$

Now

$$\sum_{\substack{\sqrt{x}<n\le x\\ n\equiv a \pmod q}} f_2(n) = \sum_{\substack{m,r\le\sqrt{x},\ell\ge 1\\ \sqrt{x}<mr\ell\le x\\ mr\ell\equiv a \pmod q}} \mu(m)\Lambda(r).$$

In this latter sum we will cut the ranges for m, r, ℓ up into dyadic ranges, say $M < m \le 2M$, $R < r \le 2R$ and $L < \ell \le 2L$. To start with we have, for $\sqrt{x} < MRL \le x$

$$\sum_{\substack{M<m\le 2M\\ R<r\le 2R\\ (mr,q)=1}} \mu(m)\Lambda(r) \sum_{\substack{L<\ell\le 2L\\ \ell\equiv a/(mr) \pmod q}} 1 = \sum_{\substack{M<m\le 2M\\ R<r\le 2R\\ (mr,q)=1}} \mu(m)\Lambda(r) \left\{ \frac{L}{q} + O(1) \right\}.$$

Summing over all a with $(a,q)=1$ we get an error term

$$\ll \sum_{\substack{M<m\le 2M\\ R<r\le 2R}} \Lambda(r) \ll MR.$$

This is acceptably small provided $MR \le x/(\log x)^{A+2}$ (since there are $\ll (\log x)^2$ such pairs M, R). Therefore we may assume that $MR \ge x/(\log x)^{A+2}$: since $M, R \le \sqrt{x}$ this implies that $M, R \ge \sqrt{x}/(\log x)^{A+2}$. In this range we may employ Theorem 13.2, taking $\beta_r = \Lambda(r)$ which satisfies the Siegel-Walfisz criterion and

$$\alpha_n = \sum_{\substack{M<m\le 2M,\ L<\ell\le 2L \\ m\ell=n}} \mu(m),$$

so that $|\beta|^2 = \sum_{R<r\le 2R} \Lambda(r)^2 \ll \sum R\log x$ and $\|\alpha\|^2 \le \sum_n \tau(n)^2 \ll LM(\log x)^3$.

This is not quite a complete proof because if the dyadic intervals are given by $(L, 2L], (M, 2M], (R, 2R]$ with $LMR < x \le 8LMR$, we have counted sum terms corresponding to n that are larger than x. To correct for this we need cut the ranges up into finer intervals, say of the form $(L, (1+\delta)L], (M, (1+\delta)M], (R, (1+\delta)R]$, where $\delta = 1/(\log x)^C$, so that the total possible contribution of these intervals, whose contribution includes terms n that are greater than x, is sufficiently small.

13.6. Barban-Davenport-Halberstam, II. The Montgomery-Hooley refinement:

Theorem 13.4. *There exists a constant c such that if $1 \le Q \le x$ then*

$$\sum_{q\le Q}\ \sum_{a:\ (a,q)=1} \left|\theta(x;q,a) - \frac{x}{\phi(q)}\right|^2 = xQ(\log Q + c) + O\left(Q^2(\log x)^\epsilon + \frac{x^2}{(\log x)^A}\right)$$

for any fixed $A > 0$.

Proof. The result follows from Theorem 13.1 for $Q \le x/(\log x)^B$, so we can assume that $x/(\log x)^B < Q \le x$. It is convenient to $\beta_n = \log n$ if n is prime and 0 otherwise, so that the sequence $\beta_n, n \le N$ satisfies the Siegel-Walfisz condition. We start by noting that

$$\sum_{a:\ (a,q)=1} \left|\theta(x;q,a) - \frac{x}{\phi(q)}\right|^2 = \sum_{\substack{m,n\le x \\ m\equiv n \pmod q}} \beta_m\beta_n - \frac{x^2}{\phi(q)}\left\{1 + O\left(\frac{x}{(\log x)^{A+1}}\right)\right\}$$

by the Siegel-Walfisz theorem. Summing this up over all $Q < q \le x$ we obtain

$$\sum_{Q<q\le x}\ \sum_{a:\ (a,q)=1} \left|\theta(x;q,a) - \frac{x}{\phi(q)}\right|^2 = \sum_{Q<q\le x}\ \sum_{\substack{m,n\le x \\ m\equiv n \pmod q}} \beta_m\beta_n \\ - x^2 \sum_{Q<q\le x} \frac{1}{\phi(q)} + O\left(\frac{x^2}{(\log x)^A}\right).$$

Now if $m < n$ then we write $n - m = qr$ so that $r = (n-m)/q < x/Q$. Therefore, using the Siegel-Walfisz theorem,

$$\begin{aligned}\sum_{Q<q\leq x}\sum_{\substack{m<n\leq x\\ m\equiv n \pmod q}}\beta_m\beta_n &= \sum_{r\leq x/Q}\sum_{\substack{m+rQ<n\leq x\\ m\equiv n \pmod r}}\beta_m\beta_n\\ &= \sum_{r\leq x/Q}\sum_{m\leq x-rQ}\beta_m(\theta(x;r,m)-\theta(m+rQ;r,m))\\ &= \sum_{r\leq x/Q}\sum_{m\leq x-rQ}\beta_m\left(\frac{x-rQ-m}{\phi(r)}+O\left(\frac{x}{(\log x)^A}\right)\right)\\ &= \sum_{r\leq x/Q}\frac{(x-rQ)^2}{2\phi(r)}+O\left(\frac{x^2}{(\log x)^A}\right)\end{aligned}$$

For the last quantity we use a variant on Perron's formula: If $c>1$ then

$$\frac{1}{2i\pi}\int_{(c)}\frac{2y^{s+1}}{(s-1)s(s+1)}\,ds=\begin{cases}(y-1)^2 & \text{if } y>1\\ \frac{1}{2}(y-1)^2 & \text{if } y=1\\ 0 & \text{if } 0<y<1\end{cases}$$

Therefore, if R is not an integer then

$$\begin{aligned}\sum_{r\leq R}\frac{(R-r)^2}{2\phi(r)} &= \frac{1}{2i\pi}\int_{(c)}\sum_{r\geq 1}\frac{r^2}{\phi(r)}\left(\frac{R}{r}\right)^{s+1}\frac{ds}{(s-1)s(s+1)}\\ &= \frac{1}{2i\pi}\int_{(c)}\zeta(s)A(s)R^{s+1}\frac{ds}{(s-1)s(s+1)},\end{aligned}$$

where $A(s):=\prod_p(1+\frac{1}{p^s(p-1)})$. Pulling the contour back to the left we uncover poles at $s=1,0,-1$. At $s=1$ the integrand has a double pole, and so the residue is

$$\frac{1}{2}\ A(1)R^2\left(\log R+\frac{A'(1)}{A(1)}+\gamma-\frac{1}{2}\right).$$

Writing $A(s)=\zeta(s+1)B(s)$, we determine that the integrand also has a double pole at $s=0$ with residue

$$-\zeta(0)B(0)R\left(\log R+\frac{B'(0)}{B(0)}+\frac{\zeta'(0)}{\zeta(0)}+\gamma\right).$$

Now $B(0)=1, \zeta(0)=-1/2$ and $\zeta'(0)/\zeta(0)=\log(2\pi)$. One can show that the error term when incorporating these two residues in $O(R^\epsilon)$. Substituting this in above gives

$$\sum_{r\leq x/Q}\frac{(x-rQ)^2}{\phi(r)}=A(1)x^2\left(\log(x/Q)+c_1\right)+xQ\left(\log(x/Q)+c_2\right)+O(Q^2(x/Q)^\epsilon)$$

where $c_1 := A'(1)/A(1)+\gamma-1/2,\ c_2 := B'(0)/B(0)+\log(2\pi)+\gamma$. With a similar argument for $n<m$, and using the prime number theorem when $m=n$ we deduce that

$$\sum_{Q<q\le x}\ \sum_{\substack{m,n\le x\\ m\equiv n \pmod q}} \beta_m\beta_n =(x-Q)x(\log x-1)+A(1)x^2\left(\log(x/Q)+c_1\right)$$
$$+xQ\left(\log(x/Q)+c_2\right)+O(Q^2(x/Q)^{\epsilon})+O\left(\frac{x^2}{(\log x)^A}\right)$$

We also note that there exists a constant c_3 such that

$$\sum_{q\le x}\frac{1}{\phi(q)} = A(1)\log x+c_3+O(\log x/x). \tag{13.7}$$

Adding all of the above together and noting the symmetry in m and n, we obtain

$$\sum_{Q<q\le x}\ \sum_{a:\ (a,q)=1}\left|\theta(x;q,a)-\frac{x}{\phi(q)}\right|^2 = x^2(\log x+c_4)-xQ\left(\log Q-c_2-1\right)\big)$$
$$+O(Q^2(\log x)^{\epsilon}),$$

where $c_4 = A(1)c_1-1$. Adding in (13.3) with A sufficiently large implies that

$$\sum_{q\le x}\ \sum_{a:\ (a,q)=1}\left|\theta(x;q,a)-\frac{x}{\phi(q)}\right|^2 = x^2(\log x+c_4)+O\left(\frac{x^2}{(\log x)^A}\right).$$

Subtracting the last two equations achieves our objective (and seems to imply that $A(1)c_1 = c_4+1=-c_2$ which is dubious, so there may be an error.).

Exercises

13.6a. Prove the variant of Perron's formula given here. (Hint: You may wish to simply use the first version of Perron's formula directly rather than any calculus.)

13.6b.a) Use elementary methods to prove that $\sum_{q\le x}\ q/\phi(q) = A(1)x+O(\log x)$.

b) Use partial summation to deduce (13.7).