ISSN NO: 2770-0003
Date of Publication: 06-05-2023

The Least Quadratic Nonresidue and Vinogradov's Hypothesis

Abdunabiyev Jamshid Olimjon o'g'li
TerDU matematika yo'nalishi
1-kurs magistranti

Abstract. These are rough notes covering the second block of lectures in the "Elementary Methods in Analytic Number Theory" course. In these lectures we will develop several forms of the large sieve inequality, which assert that no sequence can be well correlated with many exponentials or poorly distributed in many arithmetic progressions. By combining the large sieve with Vaughan's Identity and the Siegel– Walfisz theorem, we will deduce the Bombieri–Vinogradov theorem on the average distribution of primes in progressions.

(No originality is claimed for any of the contents of these notes. In particular, they borrow from the books of Davenport [1] and Iwaniec and Kowalski [3].)

Key words:

The idea of the large sieve

In the previous chapter we used information about the distribution of sequences A in arithmetic progressions (usually just the zero residue class), together with combinatorial constructions of sieve weights λ_d, to deduce things about the number of primes in those sequences. The same arguments give information about other kinds of numbers defined by multiplicative conditions. In this chapter we develop another approach to investigating and exploiting the distribution of A in progressions.

We start by observing that we can detect the distribution of sequences in arithmetic progressions using sums of complex exponentials. This is a form of Fourier analysis, but with only finite sums appearing and therefore no convergence issues.

Definition 5.1. We write $e(\theta) := e^{2\pi i\theta}$ for the *complex exponential*, where $\theta \in$ R.

Note that $e(\theta)$ is a 1-periodic function.

Lemma 5.2 (Discrete Parseval identity)**.** *Let* A = (a_n) *be any finite sequence, and let* $q \in$ N. *Write* $S(\theta) = S_{\text{A}}(\theta)$:= $^{\text{P}}{}_{n}\, a_n e(n\theta)$ *for each* $\theta \in$ R. *Then*

$$q\sum_{a=1}^{q}\left|\sum_{n\equiv a\,mod\,q} a_n\right|^2 = \sum_{a=1}^{q}|S(a/q)|^2$$

.

Date: 25th February 2015.

1

Proof of Lemma 5.2. If we expand the sum on the right hand side we obtain

$$\sum_{a=1}^{q} S(a/q)\overline{S(a/q)} = \sum_{a=1}^{q}\sum_{n} a_n e(na/q)\sum_{m}\overline{a_m}e(-ma/q) = \sum_{n} a_n\sum_{m}\overline{a_m}\sum_{a=1}^{q} e((n-m)a/q)$$

.

Now if $n-m$ is divisible by q then $(n-m)a/q$ is always an integer, and so $e((n-m)a/q)$ is always 1, and the sum over a is just q. On the other hand, if $n-m$ is not divisible by q then $(n-m)/q$ is a non-trivial fraction (say b/r, with $(b,r) = 1$ and $r \geq 2$ and $r \mid q$), so

q r

$$\underset{a=1}{\overset{X}{}} e((n-m)a/q) = (q/r) \underset{a=1}{\overset{X}{}} e(ab/r) = 0,$$

since on summing the geometric progression the end cancels the first term.
We have shown that

$$\sum_{a=1}^{q} |S(a/q)|^2 = q \sum_{n,m:n\equiv m \bmod q} a_n\overline{a_m} = q\sum_{a=1}^{q}\left|\sum_{n\equiv a \bmod q} a_n\right|^2,$$

as claimed.

In view of Lemma 5.2, we see that to understand the distribution of A in progressions to different moduli q (at least in a mean square sense) we must understand sums of $|S(\theta)|^2$ for various points θ on the unit circle. The idea of the *large sieve inequality* is that one can compare a sum$\sum_{r=1}^{R}|S(\theta_r)|^2$, where the θ_r are "well-spaced" points, with the integral around the whole circle which we can understand precisely:

$$\int_0^1 |S(\theta)|^2 d\theta = \sum_n a_n \sum_m \overline{a_m}\int_0^1 e((n-m)\theta)d\theta = \sum_n |a_n|^2,$$

as the integral vanishes unless $n - m = 0$. (This is a continuous version of Parseval's identity).
We will prove a general version of the inequality first, and then specialise to the case where the θ_r are various rationals a/q.

Definition 5.3. Let $\delta > 0$. We say points $\theta_1,...,\theta_R \in$ R are *δ-spaced* if

$$\|\theta_r - \theta_s\| \geq \delta \qquad \forall r\ 6= s,$$

where $\|\cdot\|$ denotes distance to the nearest integer.

Theorem 5.4 (Large Sieve inequality, Exponential Sums Version)**.** *Let $\delta > 0$, and suppose $\theta_1,...,\theta_R \in$ R are δ-spaced points. Also let $M \in$ Z, and let A = $(a_n)_{M<n\leq M+N}$ be any complex numbers. Then*

$$\sum_{r=1}^{R} |S_{\mathcal{A}}(\theta_r)|^2 \leq \left(\frac{1}{\delta} + 2\pi N\right)\sum_{M<n\leq M+N} |a_n|^2 .$$

Proof of Theorem 5.4. Notice first that for any real θ,

$$S_A(\theta) := \underset{M<n\leq M+N}{\overset{X}{}} a_n e(n\theta) = \underset{0<n\leq N}{\overset{X}{}} a_{n+M}e((n+M)\theta) = e(M\theta) \underset{0<n\leq N}{\overset{X}{}} \tilde{a}_n e(n\theta),$$

where $\tilde{a}_n := a_{n+M}$ for all n. So we have $|S_A(\theta)| = |S_{\tilde{A}}(\theta)|$ for all θ (where $\tilde{A} = (\tilde{a}_n)_{0<n\leq N}$), so we see it will suffice to prove the theorem in the case where $M = 0$.
We shall give a proof due to Gallagher in 1967, that relies on a simple real analysis lemma comparing the value at a point with an average value.

Lemma 5.5 (Sobolev–Gallagher inequality)**.** *Let $f : [0,1] \to$ C be a function whose first derivative is continuous. Then*

$$|f(1/2)| \leq \int_0^1 |f(t)|dt + \frac{1}{2}\int_0^1 |f'(t)|dt.$$

More generally, for any $\delta \leq 1$ we have

$$|f(1/2)| \leq \frac{1}{\delta}\int_{1/2-\delta/2}^{1/2+\delta/2} |f(t)|dt + \frac{1}{2}\int_{1/2-\delta/2}^{1/2+\delta/2} |f'(t)|dt.$$

Proof of Lemma 5.5. We have

Z 1 Z 1 Z 1 *f*(1/2) = *f*(1/2)*du* = *f*(*u*)*du* + (*f*(1/2) − *f*(*u*))*du.*
0 0 0

However, we also have $f(1/2) - f(u) = \int_u^{1/2} f'(t)dt$ for all *u* ∈ [0,1], so

Z 1 Z 1/2 Z 1/2 Z 1 Z *u*
f(1/2) = *f*(*u*)*du* + $f^0(t)dtdu$ + (− $f^0(t)dt)du$
0 0 *u* 1/2 1/2

Z 1 Z 1/2 Z 1
= *f*(*u*)*du* + $tf^0(t)dt$ + $(-(1-t)f^0(t))dt.$
0 0 1/2

The first statement in the lemma follows easily if we insert absolute values everywhere.

To prove the second statement, define *g*(*t*) := *f*(1/2 + *δ*(*t* − 1/2)) for all *t* ∈ [0,1], and note that $g^0(t) = \delta f^0(1/2 + \delta(t-1/2))$. Applying the first part of the lemma to *g*, we obtain

$$\begin{aligned}|f(1/2)| = |g(1/2)| &\leq \int_0^1 |g(t)|dt + \frac{1}{2}\int_0^1 |g'(t)|dt \\ &= \int_0^1 |f(1/2+\delta(t-1/2))|dt + (\delta/2)\int_0^1 |f'(1/2+\delta(t-1/2))|dt.\end{aligned}$$

The result follows if we make the substitution *u* = 1/2+*δ*(*t*−1/2) in the integrals.

We apply the lemma to $f(\theta) := S_A(\theta)^2$. Note that, by the chain rule and by definition of $S_A(\theta)$, the derivative $f^0(\theta)$ is given by

$$f'(\theta) = 2S_{\mathcal{A}}(\theta)S'_{\mathcal{A}}(\theta),$$ where $S_A{}^0(\theta) = 2\pi i^X n a_n e(n\theta)$.
n

Thus the lemma (applied with a change of variables to shift θ_r to 1/2, using the fact that $S_A(\theta)$ is a 1-periodic function) implies that we always have

$$|S_{\mathcal{A}}(\theta_r)|^2 \leq \frac{1}{\delta}\int_{\theta_r-\delta/2}^{\theta_r+\delta/2} |S_{\mathcal{A}}(\theta)|^2 d\theta + \int_{\theta_r-\delta/2}^{\theta_r+\delta/2} |S_{\mathcal{A}}(\theta)S'_{\mathcal{A}}(\theta)|d\theta.$$

Now the crucial point is that the intervals $[\theta_r - \delta/2, \theta_r + \delta/2]$ for $1 \leq r \leq R$ *are non-overlapping modulo 1*, since we assume that the points θ_r are *δ*-spaced. Therefore if we sum over *r* we can upper bound the sum of all the integrals simply by the integral over [0,1], as follows:

$$\sum_{r=1}^{R} |S_{\mathcal{A}}(\theta_r)|^2 \leq \frac{1}{\delta}\int_0^1 |S_{\mathcal{A}}(\theta)|^2 d\theta + \int_0^1 |S_{\mathcal{A}}(\theta)S'_{\mathcal{A}}(\theta)|d\theta.$$

Finally, by the continuous Parseval identity (and our assumption that $M = 0$) we have $\int_0^1 |S_{\mathcal{A}}(\theta)|^2 d\theta = \sum_{0<n\leq N} |a_n|^2$, and

$$\int_0^1 |S_{\mathcal{A}}(\theta)S'_{\mathcal{A}}(\theta)|d\theta \leq \sqrt{\int_0^1 |S_{\mathcal{A}}(\theta)|^2 d\theta \cdot \int_0^1 |S'_{\mathcal{A}}(\theta)|^2 d\theta} = \sqrt{\sum_{0<n\leq N} |a_n|^2 \cdot (2\pi)^2 \sum_{0<n\leq N} n^2|a_n|^2},$$

from which the theorem follows.

Corollary 5.6. *Let* $Q \geq 1$, *and let* A = $(a_n)_{M<n\leq M+N}$ *be any complex numbers. Then*

$$\sum_{q\leq Q}\sum_{(a,q)=1} |S_{\mathcal{A}}(a/q)|^2 \leq \left(Q^2 + 2\pi N\right) \sum_{M<n\leq M+N} |a_n|^2 .$$

Proof of Corollary 5.6*.* In view of Theorem 5.4, we only need to check that all the points a/q are $1/Q^2$ spaced. However, if a/q 6= b/r then

$$\frac{a}{q} - \frac{b}{r} = \frac{ar - bq}{qr} ,$$

a fraction that is not zero mod 1 and therefore is at least $1/qr \geq 1/Q^2$.

In the next section we will apply the large sieve inequality to some arithmetic problems, but it is very interesting as a purely analytic statement. Note that at any fixed point θ we have

$$|S_{\mathcal{A}}(\theta)|^2 = \left|\sum_{M<n\leq M+N} a_n e(n\theta)\right|^2 \leq N \sum_{M<n\leq M+N} |a_n|^2 ,$$

by the Cauchy–Schwarz inequality. If $\delta \geq 1/N$, then Theorem 5.4 says that for any δ-spaced points $\theta_1,...,\theta_R$ we actually have

R

X 2 X 2

$|S_A(\theta_r)| \leq (1 + 2\pi)N$ $|a_n|$.

r=1 $M<n\leq M+N$

In other words, the entire sum over r cannot be much bigger than the Cauchy–Schwarz bound for a single term. This says that *no sequence* $(a_n)_{M<n\leq M+N}$ *can be strongly correlated with the exponentials* $(e(n\theta))_{M<n\leq M+N}$ *at more than a few well spaced points* θ*.* It will turn out that, on the arithmetic side, this implies that every sequence is well distributed among most arithmetic progressions.

5. The arithmetic large sieve, and applications

If we compare the Discrete Parseval identity (Lemma 5.2) with Corollary 5.6, we see that for any $Q \geq 1$ and any sequence A = $(a_n)_{M<n\leq M+N}$ we have

$$q\sum_{a=1}^{q}\left|\sum_{n\equiv a \bmod q} a_n\right|^2 = \sum_{a=1}^{q} |S_{\mathcal{A}}(a/q)|^2 \quad \forall q \leq Q,$$

$$\text{and} \quad \sum_{q\leq Q}\sum_{(a,q)=1} |S_{\mathcal{A}}(a/q)|^2 \leq \left(Q^2 + 2\pi N\right) \sum_{M<n\leq M+N} |a_n|^2 .$$

To combine these two statements we need to understand the impact of the condition $(a,q) = 1$, which appears in the sum on the second line but not the first. The simplest way to deal with this is to restrict attention to prime moduli q, which leads to the following theorem.

Theorem 6.1 (Large Sieve Inequality, Variance Version)**.** *Let* $Q \geq 1$, *let* $M \in$ Z, *and let* A = $(a_n)_{M<n\leq M+N}$ *be any sequence. Define* $X :=$ ${}^{\mathrm{P}}{}_{M<n\leq M+N}\, a_n$*. Then*

$$\sum_{\substack{p\le Q,\\ p\ prime}} p\sum_{a=1}^{p}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a \bmod p}} a_n - \frac{X}{p}\right|^2 \le (Q^2+2\pi N)\sum_{M<n\le M+N}|a_n|^2 .$$

Remark 6.2. If the sequence A were evenly distributed we might expect the sum in any congruence class mod *p* to be about 1/*p* times *X*, the total sum. Theorem 6.1 gives an upper bound for the discrepancy or “variance”, for *any* sequence A.

Proof of Theorem 6.1. Note that if *p* is prime then

$$\sum_{(a,p)=1}|S_A(a/p)|^2 = \sum_{a=1}^{p-1}|S_A(a/p)|^2 = \sum_{a=1}^{p}|S_A(a/p)|^2 - |S_A(1)|^2 = \sum_{a=1}^{p}|S_A(a/p)|^2 - |X|^2 .$$

Using the Discrete Parseval identity (Lemma 5.2) we can rewrite this as

$$p\sum_{a=1}^{p}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a \bmod p}} a_n\right|^2 - |X|^2 = p\left(\sum_{a=1}^{p}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a \bmod p}} a_n\right|^2 - \frac{|X|^2}{p}\right),$$

and one can check by expanding the square that the right hand side is equal to

$$p\sum_{a=1}^{p}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a \bmod p}} a_n - \frac{X}{p}\right|^2 .$$

The theorem follows by summing over primes $p \le Q$ and using the upper bound estimate $\sum_{p\le Q}\sum_{(a,p)=1}|S_A(a/p)|^2 \le (Q^2 + 2\pi N)\sum_{M<n\le M+N}|a_n|^2$, which follows from Corollary 5.6.

To show the power of Theorem 6.1 we shall consider a famous example. Let *p* be an odd prime, and recall that a number *n* is said to be a *quadratic residue* modulo *p* if

$$n \equiv x^2 \bmod p \qquad \text{for some } x \in \mathbb{N},$$

and otherwise *n* is said to be a *quadratic non-residue* mod *p*. The non-zero quadratic residues form a subgroup of the invertible residue classes $(\mathbb{Z}/p\mathbb{Z})^*$, of order $(p-1)/2$.
We shall define

$n(p) := \min\{1 \le n \le p : n \text{ is a quadratic non-residue mod } p\}$,

the *least quadratic non-residue* mod *p*. It is an unsolved conjecture of Vinogradov that *n*(*p*) can never be too large relative to *p*, or more specifically that $n(p) \ll_\epsilon p^\epsilon$ for any > 0 and all *p*.

ISSN NO: 2770-0003
Date of Publication: 06-05-2023

Corollary 6.3 (Linnik, 1942). *For any N and any small* $\epsilon > 0$, *we have*

#$\{3 \le p \le N : n(p) > N^{\epsilon}\} \ll_{\epsilon} 1$.

The proof will be easier if we introduce a bit of standard notation.

Definition 6.4. For any *y* ≥ 2, a number *n* is said to be *y-smooth* if all of its prime factors are ≤ *y*.

Proof of Corollary 6.3. Let $\mathcal{P} \;:=\; \{3 \;\le\; p \;\le\; N \;:\; n(p) \;>\; N^{\epsilon}\}$, so our task is to show that #$\mathcal{P} \ll_{\epsilon}$ 1. The key observation is that if *p* ∈ P then every $n \;\le\; N^{\epsilon}$ is a quadratic residue modulo *p*, and therefore *every N-smooth number is a quadratic residue modulo p*. In particular, if we define a sequence A = (*a_n*)1≤*n*≤*N*2 by setting *an* = **1***n* is a quadratic residue modulo all *p*∈P, then we have

$$\sum_n a_n \ge \#\{1 \le n \le N^2 : n \text{ is } N^{\epsilon}$$ smooth} $$\sum_{N^{\epsilon/2-\epsilon^2/100} \le q_1,\dots,q_{\lfloor 4/\epsilon \rfloor} \le N^{\epsilon/2},} \left\lfloor \frac{N^2}{\prod q_i} \right\rfloor,$$

q_i prime

since $N^{2-\epsilon} \;\le\; \prod q_i \;\le\; N^2$ for any collection of *q_i* counted in the sum (and so any $m \le N^2/\prod q_i \le N^{\epsilon}$ will trivially be *N*-smooth). Consequently we have

$$\sum_n a_n \gg_{\epsilon} N^2 \Bigg(\quad \text{X}$$

$$\Bigg({}_{\substack{N^{\epsilon/2-\epsilon^2/100} \le q \le N^{\epsilon/2},\\ q \text{ prime}}} \Bigg) \quad \frac{1}{q} \Bigg)^{\lfloor 4/\epsilon \rfloor} \quad = \quad N^2 \left(\log\log(N^{\epsilon/2}) - \log\log(N^{\epsilon/2-\epsilon^2/100}) + o(1) \right)^{\lfloor 4/\epsilon \rfloor}$$

$\gg_{\epsilon} \quad N^2$,

in view of Fact 2 from Chapter 0. In other words, in the notation of Theorem 6.1 we have $X = \sum_{1 \le n \le N^2} a_n = \sum_{1 \le n \le N^2} |a_n|^2 \gg_{\epsilon} N^2$.

Now the sequence A has many non-zero terms, but is badly distributed in arithmetic progressions modulo all primes *p* ∈ P, so the large sieve will tell us that P must be small. Indeed, Theorem 6.1 implies that

$$\sum_{p\in\mathcal{P}} p \sum_{a=1}^{p} \left| \sum_{\substack{1 \le n \le N^2,\\ n \equiv a \bmod p}} a_n - \frac{X}{p} \right|^2 \le \left(N^2 + 2\pi N^2\right) X.$$

$n \equiv a \bmod p$

If *p* ∈ P then P1≤*n*≤*N*2, *a_n* = 0 whenever *a* is a quadratic non-residue modulo *p*, so

$$\sum_{p\in\mathcal{P}} p \sum_{a=1}^{p} \left| \sum_{\substack{1 \le n \le N^2,\\ n \equiv a \bmod p}} a_n - \frac{X}{p} \right|^2 \ge \sum_{p\in\mathcal{P}} p \quad {}_{a}$$

p
X
.

$${}_{p} \quad \frac{X^2}{p^2} \ge X^2 \sum_{p\in\mathcal{P}} \frac{p-1}{2p} \gg X^2 \#\mathcal{P}$$

a=1,
quadratic non-residue mod

It follows that #$\mathcal{P} \ll N^2/X \ll_{\epsilon} 1$.

Remark 6.5. The best upper bound we have for *n*(*p*) that is valid for all primes *p* is √–

roughly that $n(p) \le p^{1/(4\ e)}$. This is due to Burgess in 1957, and ultimately relies on a deep result from algebraic geometry (the Weil bound). The large sieve gives a much stronger estimate for all except a bounded number of primes.

Another result that we can obtain from Theorem 6.1 is the following.

Corollary 6.6. *Let N be large and let* $\mathcal{S} \subseteq \{1,2,...,N\}$. *Suppose that* $\#(\mathcal{S} \bmod p) \le 0.99p$ *for all primes* $p \le \sqrt{N}$. *Then*

$$\#\mathcal{S} \ll \sqrt{N}\log N.$$

Proof of Corollary 6.6. We can apply the large sieve (Theorem 6.1) with the a_n chosen to be the characteristic function $\mathbf{1}_{n\in\mathcal{S}}$, and with $Q = \sqrt{N}$. Thus we have $X = \sum_{1\le n\le N} a_n = \sum_{1\le n\le N} |a_n|^2 = \#\mathcal{S}$, so the large sieve implies that

$$\sum_{\substack{p\le\sqrt{N},\\ p \text{ prime}}} p\sum_{a=1}^{p}\left|\sum_{\substack{1\le n\le N,\\ n\equiv a \bmod p}} \mathbf{1}_{n\in\mathcal{S}} - \frac{\#\mathcal{S}}{p}\right|^2 \le (N+2\pi N)\,\#\mathcal{S}.$$

On the other hand, if $\#(\mathcal{S} \bmod p) \le 0.99p$ then there are at least $0.01p$ values $1 \le a \le p$ for which $\sum_{\substack{1\le n\le N,\\ n\equiv a \bmod p}} \mathbf{1}_{n\in\mathcal{S}} = 0$. So for each $p \le \sqrt{N}$ we have

$$p\sum_{a=1}^{p}\left|\sum_{\substack{1\le n\le N,\\ n\equiv a \bmod p}} \mathbf{1}_{n\in\mathcal{S}} - \frac{\#\mathcal{S}}{p}\right|^2 \ge p\cdot 0.01p\cdot\left(\frac{\#\mathcal{S}}{p}\right)^2 \gg (\#\mathcal{S})^2,$$

so we deduce that

$$\sum_{\substack{p\le\sqrt{N},\\ p \text{ prime}}} (\#\mathcal{S})^2 \ll N\#\mathcal{S}.$$

We know there $\asymp \sqrt{N}/\log N$ areprimes less than $\sqrt{N}$, so the bound $\#\mathcal{S} \ll \sqrt{N}\log N$ follows by rearranging.

Remark 6.7. In the proof of Corollary 6.3 we constructed a sequence A that we knew would have many terms, so the large sieve told us it couldn't be badly distributed modulo many primes. In Corollary 6.6 we assume that S *is* badly distributed modulo lots of primes, so the large sieve tells us that S cannot have too many elements. Actually Corollary 6.6 is close to best possible, since if we chose S := $\{n^2 \le N : n \text{ even}\}$ then we have $\#\mathcal{S} \gg \sqrt{N}$, whilst #(S mod p) is about $p/2$ (the quadratic residues) for all p.

Remark 6.8. Notice that in the above examples we assumed that our sequences missed a *large* number of residue classes (e.g. $0.01p$ classes) modulo different primes p, rather than just one or two classes as in Chapter 1. This is why these sorts of problems, and the associated methods, are called *large sieve* problems and methods.

To finish this section, we will formulate some versions of Theorem 6.1 that work when we don't restrict to prime moduli p. There are various ways to proceed, but we will start by doing the obvious thing and just summing over all moduli $q \le Q$, and seeing what happens.

Lemma 6.9. *Let $Q \ge 1$, let $M \in \mathbb{Z}$, and let* A = $(a_n)_{M<n\le M+N}$ *be any sequence.*
Define X := P*M<n≤M+N an. Then*

$$\sum_{q\le Q} q\sum_{a=1}^{q}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a\, mod\, q}} a_n - \frac{X}{q}\right|^2 \le Q(8Q+2\pi N)\sum_{M<n\le M+N}|a_n|^2 .$$

Proof of Lemma 6.9. The same proof as in Theorem 6.1 shows that for any $q \le Q$,

$$q\sum_{a=1}^{q}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a \bmod q}} a_n - \frac{X}{q}\right|^2 = q\left(\sum_{a=1}^{q}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a \bmod q}} a_n\right|^2 - \frac{|X|^2}{q}\right) = \sum_{a=1}^{q-1}|S_{\mathcal{A}}(a/q)|^2 .$$

For general q the sum on the right is not restricted to $(a,q) = 1$, but by writing each fraction a/q in lowest terms we can rewrite it as
q−1

X 2 X X 2
$|S_A(a/q)|$ = $|S_A(b/r)|$, a=1 $r|q$, r6=1 (b,r)=1

so if we sum over all $q \le Q$ we obtain

$$\sum_{q\le Q} q\sum_{a=1}^{q}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a \bmod q}} a_n - \frac{X}{q}\right|^2 = \sum_{q\le Q}\ \sum_{r|q,\ r\neq 1}\ \sum_{(b,r)=1}|S_{\mathcal{A}}(b/r)|^2 \le \sum_{2\le r\le Q}\frac{Q}{r}\sum_{(b,r)=1}|S_{\mathcal{A}}(b/r)|^2 .$$

Finally, notice that for any $R \ge 2$ we have

$$\sum_{R\le r\le 2R}\frac{Q}{r}\sum_{(b,r)=1}|S_{\mathcal{A}}(b/r)|^2 \le \frac{Q}{R}((2R)^2+2\pi N)\sum_{M<n\le M+N}|a_n|^2 = Q\left(4R+\frac{2\pi N}{R}\right)\sum_{M<n\le M+N}|a_n|^2 ,$$

in view of Corollary 5.6. If we sum over all R of the form $2^j \le Q$, we obtain

$$\begin{aligned}\sum_{2\le r\le Q}\frac{Q}{r}\sum_{(b,r)=1}|S_{\mathcal{A}}(b/r)|^2 &\le Q\left(4\sum_{2^j\le Q}2^j + 2\pi N\sum_{2^j\le Q}\frac{1}{2^j}\right)\sum_{M<n\le M+N}|a_n|^2\\ &\le Q(8Q+2\pi N)\sum_{M<n\le M+N}|a_n|^2,\end{aligned}$$

as claimed.

Lemma 6.9 is quite a lot weaker than Theorem 6.1, because instead of a term $2\pi N$ in the bound we have a term $2\pi QN$. This reflects the fact that if the sequence A is badly distributed to some small modulus r, it will also be badly distributed modulo all multiples q of r (an issue which doesn't arise when summing over prime moduli). If we know in advance that A is well distributed to small moduli, then we get a stronger bound.

Lemma 6.10. *Let the situation be as in Lemma 6.9. Also let $2 \le R \le Q$, and suppose that*

$$\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a\, mod\, r}} a_n - \frac{X}{r}\right| \le \frac{|X|}{R^2} \qquad \forall 1\le a\le r, \quad \forall r\le R.$$

Then we have

$$\sum_{q\le Q} q\sum_{a=1}^{q}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a\, mod\, q}} a_n - \frac{X}{q}\right|^2 \ll Q(Q+\frac{N}{R})\sum_{M<n\le M+N}|a_n|^2 .$$

Proof of Lemma 6.10. Following the proof of Lemma 6.9, but separating out those $r \le R$ and then summing over values 2^j starting from R rather than from 2, we have

$$\sum_{q\le Q} q\sum_{a=1}^{q}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a\, \mathrm{mod}\, q}} a_n - \frac{X}{q}\right|^2 \ll \sum_{2\le r\le R}\frac{Q}{r}\sum_{(b,r)=1}|S_{\mathcal{A}}(b/r)|^2 + Q(Q+\frac{N}{R})\sum_{M<n\le M+N}|a_n|^2 .$$

For any $r \le R$ and any $(b,r) = 1$ we have

$$S_{\mathcal{A}}(b/r) := \sum_{M<n\le M+N} a_n e(nb/r) = \sum_{a=1}^{r} e(ab/r)\sum_{\substack{M<n\le M+N,\\ n\equiv a\, \mathrm{mod}\, r}} a_n = \sum_{a=1}^{r} e(ab/r)\left(\frac{X}{r} + O\left(\frac{|X|}{R^2}\right)\right) ,$$

by hypothesis. However, we have $\sum_{a=1}^{r} e(ab/r)\frac{X}{r}$= 0, just by summing the geometric progression, and so we see $|S_{\mathcal{A}}(b/r)| \ll |X|/R$. It follows that

$$\sum_{2\le r\le R}\frac{Q}{r}\sum_{(b,r)=1}|S_{\mathcal{A}}(b/r)|^2 \ll \sum_{2\le r\le R}\frac{Q}{r}\sum_{(b,r)=1}\frac{|X|^2}{R^2} \le Q|X|^2/R.$$

Finally, we have $|X|^2 = \left|\sum_{M<n\le M+N} a_n\right|^2 \le N\sum_{M<n\le M+N}|a_n|^2$ by the Cauchy–Schwarz inequality, so the total contribution from those $r \le R$ is $\ll (QN/R)\sum_{M<n\le M+N}|a_n|^2$, which is acceptable.

Another way to handle general moduli q is to stop trying to relate $^{\mathrm{P}}{}_{(a,q)=1}\,|S_{\mathrm{A}}(a/q)|^2$ to $\sum_{a=1}^{q}\left|\sum_{n\equiv a\, \mathrm{mod}\, q} a_n\right|^2$, and instead to relate $^{\mathrm{P}}{}_{(a,q)=1}\,|S_{\mathrm{A}}(a/q)|^2$ to the distribution of A in other ways. We finish this section by stating such a result.

Theorem 6.11 (Large Sieve Inequality, Arithmetic Version, Montgomery, 1968). *Let*

M ∈ Z, let A = $(a_n)_{M<n\le M+N}$ be any sequence, and define X := $^{\mathrm{P}}{}_{M<n\le M+N}\, a_n$. Suppose that for each prime p, there is a set of $0 \le \omega(p) < p$ residue classes mod p such that $a_n = 0$ whenever n mod p lies in such a residue class. Then for any squarefree q we have

$$\sum_{(a,q)=1}|S_{\mathcal{A}}(a/q)|^2 \ge |X|^2\prod_{p|q}\frac{\omega(p)}{p-\omega(p)} .$$

Consequently, for any $Q \ge 1$ we have

$$|X|^2\sum_{q\le Q,}\prod_{p|q}\frac{\omega(p)}{p-\omega(p)} \le (Q^2+2\pi N)\sum_{M<n\le M+N}|a_n|^2 .$$

q squarefree

Proof of Theorem 6.11. The proof will be an exercise on the second problem sheet (using induction on the number of prime factors of q).

6. Primes in arithmetic progressions

Recall that we write $\pi(x;q,a) := \#\{p \le x : p \equiv a \bmod q\}$, where p denotes primes. In Chapter 1 we showed that

$$\pi(x;q,a) \ll \frac{x}{\phi(q)\log(x/q)} \qquad \forall 2 \le q \le x/1000, \;\; \forall (a,q)=1$$
,

where $\varphi(q)$ is the Euler totient function. Using zeta function type methods one can prove the *Siegel–Walfisz Theorem*, which says that for any fixed $A > 0$ we have

$$\pi(x;q,a) = \frac{1}{\phi(q)}\int_2^x \frac{dt}{\log t} + O_A(xe^{-c\sqrt{\log x}}) \qquad \forall 2 \le q \le \log^A x, \;\; \forall (a,q)=1$$
,

where $c > 0$ is a constant. For many applications (such as bounded gaps between primes) we need to know that $\pi(x;q,a)$ is close to $\frac{1}{\phi(q)}\int_2^x \frac{dt}{\log t}$ for certain fixed a and for *most values of q up to a power of x*. More specifically, in sieve arguments we need to bound remainder sums like

$$\sum_{\substack{d\le D,\\ d\,\text{squarefree},\\ d \quad \in\mathcal{P}}} 3^{\omega(d)}|r(d)| = \sum_{\text{squarefree}} 3^{\omega(d)}\left|\sum_{n:d|n} a_n - g(d)X\right|$$

,
$d \le D$,

d ,

$p| \Rightarrow p$ $p|d \Rightarrow p \in$P

and if the sequence a_n is somehow related to the primes (e.g. the primes shifted by a fixed integer a) this leads to remainder sums like

$$\sum_{\substack{q\le D,\\ (q,a)=1}} \left|\pi(x;q,a) - \frac{1}{\phi(q)}\int_2^x \frac{dt}{\log t}\right|.$$

To bound such remainder sums we shall prove the following famous theorem.

Theorem 7.1 (Bombieri–Vinogradov Theorem, 1965)**.** *For any fixed $A > 0$, there exists a constant $B = B(A) > 0$ such that, for all large x,*

$$\sum_{q\le\sqrt{x}/\log^B x} \max_{(a,q)=1} \left|\pi(x;q,a) - \frac{1}{\phi(q)}\int_2^x \frac{dt}{\log t}\right| \ll \frac{x}{\log^A x}.$$

This was proved independently by Bombieri and by A. I. Vinogradov. The original proofs used the large sieve and lots of zeta function ideas, but we shall give a proof where the only zeta function input is in the form of Siegel–Walfisz type results for small moduli.

To prove Theorem 7.1, we first need to address the fact that the primes can only possibly be equidistributed in the *coprime* residue classes mod q, whereas our large sieve results were formulated for sequences that we expect to be equidistributed over *all* residue classes mod q (so the term we subtracted in the "variance" was X/q rather than $X/\varphi(q)$). The following form of the large sieve is what we shall need.

Theorem 7.2 (Large Sieve Inequality, Coprime Version)**.** *Let $Q \ge 1$, let $M \in$ Z, and let* A $= (a_n)_{M<n\le M+N}$ *be any sequence. Then*

$$\sum_{q\le Q} q \sum_{(a,q)=1} \left|\sum_{\substack{M<n\le M+N,\\ n\equiv a \bmod q}} a_n - \frac{1}{\phi(q)}\sum_{\substack{M<n\le M+N,\\ (n,q)=1}} a_n\right|^2 \ll Q(Q+N)\sum_{M<n\le M+N} |a_n|^2.$$

Moreover, if $2 \le R \le Q$ and if

$$\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a\,mod\,r}} a_n - \frac{1}{\phi(r)}\sum_{\substack{M<n\le M+N,\\ (n,r)=1}} a_n\right| \le \frac{\sqrt{N\sum_{M<n\le M+N}|a_n|^2}}{R^4} \qquad \forall (a,r)=1, \quad \forall r\le Q,$$

then

$$\sum_{q\le Q} q \sum_{(a,q)=1}\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a\ mod\ q}} a_n - \frac{1}{\phi(q)}\sum_{\substack{M<n\le M+N,\\ (n,q)=1}} a_n\right|^2 \ll Q(Q+\frac{N\log^3 Q}{R}) \sum_{M<n\le M+N} |a_n|^2 \quad .$$

Proof of Theorem 7.2. The proof is omitted [[and therefore non-examinable, of course]] to save time— it works exactly like the various proofs we have already seen, but translating from exponentials $e(na/q)$ to so-called *Dirichlet characters* which detect the coprimality conditions.

Remark 7.3. Note that Theorem 7.2 is highly analogous to Lemma 6.10. In the second part we ask for a bound that is valid for all $r \le Q$, rather than just $r \le R$, but in fact the bound is trivial for r large enough since (if $N \ge r$) we have

$$\left|\sum_{\substack{M<n\le M+N,\\ n\equiv a\,\mathrm{mod}\,r}} a_n\right| \le \sqrt{\left(\sum_{M<n\le M+N} |a_n|^2\right)\left(\sum_{\substack{M<n\le M+N,\\ n\equiv a\,\mathrm{mod}\,r}} 1\right)} \ll \sqrt{\frac{N\sum_{M<n\le M+N}|a_n|^2}{r}} \quad ,$$

by the Cauchy–Schwarz inequality.

The next and biggest issue in proving Theorem 7.1 is that *we don't want a bound where we sum over q and over* $(a,q) = 1$, *but where we sum over q only and look at the worst residue class a for each q*. There is a nice general idea that will eventually let us handle this, which is that *if the sequence a_n is "well factorable", so that sums of a_n can be rewritten as double sums of two sequences α_u,β_v, then one of the variables u,v can take the place of the sum over residue classes.*

Proposition 7.4. *Let* $2 \le R \le Q$, *and let* $(\alpha_u)_{1\le u\le U}$ *and* $(\beta_v)_{1\le v\le V}$ *be any sequences.*

Suppose that

$$\left|\sum_{\substack{v\le V,\\ v\equiv a\,mod\,r}} \beta_v - \frac{1}{\phi(r)}\sum_{\substack{v\le V,\\ (v,r)=1}} \beta_v\right| \le \frac{\sqrt{V\sum_{v\le V}|\beta_v|^2}}{R^4} \qquad \forall (a,r)=1, \quad \forall r\le Q.$$

Then

$$\sum_{q\le Q}\max_{(a,q)=1}\left|\sum_{\substack{1\le u\le U, 1\le v\le V,\\ uv\equiv a\ mod\ q}} \alpha_u\beta_v - \frac{1}{\phi(q)}\sum_{\substack{1\le u\le U, 1\le v\le V,\\ (uv,q)=1}} \alpha_u\beta_v\right|$$
$$\ll \quad (Q+\sqrt{U+V}\log^2(QUV)+\frac{\sqrt{UV}\log^3 Q}{\sqrt{R}})\sqrt{\sum_{u\le U}|\alpha_u|^2\sum_{v\le V}|\beta_v|^2} \quad .$$

Remark 7.5. Notice that if $UV \approx x$, and if the α_u and β_v are on average of size about 1, then the bound in Proposition 7.4 will be roughly $\ll \ (Q+\sqrt{U+V}\log^2(Qx)+\frac{\sqrt{x}\log^3 Q}{\sqrt{R}})\sqrt{x}$ √_ . So if Q is a bit less than x, and U and V are both a bit less than x, and R isn't too small, then we can obtain a bound that is a bit less than x (as we want in the Bombieri–Vinogradov theorem).

ISSN NO: 2770-0003
Date of Publication: 06-05-2023

Sketch Proof of Proposition 7.4. (We shall prove the weaker bound $\ll (Q+\sqrt{Q}\sqrt{U+V}\log^2(QUV)+\frac{\sqrt{UV}\log^3 Q}{\sqrt{R}})\sqrt{\sum_{u\le U}|\alpha_u|^2\sum_{v\le V}|\beta_v|^2}$, which would suffice to prove the Bombieri–Vinogradov

√ — *B*

theorem with the sum over $q \le x/\log x$ replaced by a sum over q up to a small power of x. The proof of the stronger bound would require us to look inside the proof of Theorem 7.2, which we omitted.)

For each $q \le Q$, let a_q denote the coprime residue class mod q at which $\max_{(a,q)=1}$ of the inner modulus is attained, so the sum we are trying to bound can be rewritten as

$$\sum_{q\le Q}\left|\sum_{\substack{1\le u\le U,1\le v\le V,\\ uv\equiv a_q \bmod q}}\alpha_u\beta_v-\frac{1}{\phi(q)}\sum_{\substack{1\le u\le U,1\le v\le V,\\ (uv,q)=1}}\alpha_u\beta_v\right|.$$

Notice also that $uv \equiv a_q \bmod q$ if and only if there exists $(h,q) = 1$ such that $u \equiv h \bmod q$ and $v \equiv a_q h^{-1} \bmod q$ (where h^{-1} denotes the inverse of $h \bmod q$). So we can rewrite the sum again as

$$\sum_{Q/2\le q\le Q}\sum_{(h,q)=1}\left|\sum_{\substack{1\le v\le V,\\ v\equiv a_q h^{-1}\bmod q}}\beta_v-\frac{1}{\phi(q)}\sum_{\substack{1\le v\le V,\\ (v,q)=1}}\beta_v\right|^2 \le \frac{2}{Q}\sum_{Q/2\le q\le Q}q\sum_{(h,q)=1}\left|\sum_{\substack{1\le v\le V,\\ v\equiv h\bmod q}}\beta_v-\frac{1}{\phi(q)}\sum_{\substack{1\le v\le V,\\ (v,q)=1}}\beta_v\right|$$

$$\ll (Q+\frac{V\log^3 Q}{R})\sum_{v\le V}|\beta_v|^2.$$

$$\sum_{q\le Q}\left|\sum_{(h,q)=1}\sum_{\substack{1\le u\le U,1\le v\le V,\\ u\equiv h,\ v\equiv a_q h^{-1}\bmod q}}\alpha_u\beta_v-\frac{1}{\phi(q)}\sum_{(h,q)=1}\sum_{\substack{1\le u\le U,1\le v\le V,\\ u\equiv h\bmod q,\ (v,q)=1}}\alpha_u\beta_v\right|,$$

and using the triangle inequality this is

$$\le\sum_{q\le Q}\sum_{(h,q)=1}\sum_{\substack{1\le u\le U,\\ u\equiv h\bmod q}}|\alpha_u|\left|\sum_{\substack{1\le v\le V,\\ v\equiv a_q h^{-1}\bmod q}}\beta_v-\frac{1}{\phi(q)}\sum_{\substack{1\le v\le V,\\ (v,q)=1}}\beta_v\right|,$$

and using the Cauchy–Schwarz inequality it is

$$\le\sqrt{\sum_{q\le Q}\sum_{(h,q)=1}\left(\sum_{\substack{1\le u\le U,\\ u\equiv h\bmod q}}|\alpha_u|\right)^2}\sqrt{\sum_{q\le Q}\sum_{(h,q)=1}\left|\sum_{\substack{1\le v\le V,\\ v\equiv a_q h^{-1}\bmod q}}\beta_v-\frac{1}{\phi(q)}\sum_{\substack{1\le v\le V,\\ (v,q)=1}}\beta_v\right|^2}.$$

Next, using the second part of the Coprime Large Sieve (Theorem 7.2) we have

The same argument shows that when we sum over $Q/2^{j+1} \le q \le Q/2^j$ we have a bound $\ll (\frac{Q}{2^j}+\frac{V\log^3 Q}{R})\sum_{v\le V}|\beta_v|^2$, so summing our bounds over all $0 \le j \le (\log Q)/\log 2$ we obtain

$$\sum_{q\le Q}\sum_{(h,q)=1}\left|\sum_{\substack{1\le v\le V,\\ v\equiv a_q h^{-1}\bmod q}}\beta_v-\frac{1}{\phi(q)}\sum_{\substack{1\le v\le V,\\ (v,q)=1}}\beta_v\right|^2\ll(Q+\frac{V\log^4 Q}{R})\sum_{v\le V}|\beta_v|^2$$

2 .

To handle the other sum $\sum_{q\leq Q}\sum_{(h,q)=1}\left(\sum_{\substack{1\leq u\leq U,\\ u\equiv h \bmod q}}|\alpha_u|\right)$, we note it is

$$\ll \sum_{q\leq Q}\sum_{(h,q)=1}\left(\sum_{\substack{1\leq u\leq U,\\ u\equiv h \bmod q}}|\alpha_u| - \frac{1}{\phi(q)}\sum_{\substack{1\leq u\leq U,\\ (u,q)=1}}|\alpha_u|\right)^2 + \sum_{q\leq Q}\sum_{(h,q)=1}\left(\frac{1}{\phi(q)}\sum_{\substack{1\leq u\leq U,\\ (u,q)=1}}|\alpha_u|\right)^2 .$$

The second term here is

$$= \sum_{q\leq Q}\frac{1}{\phi(q)}\left(\sum_{\substack{1\leq u\leq U,\\ (u,q)=1}}|\alpha_u|\right)^2 \leq \sum_{q\leq Q}\frac{1}{\phi(q)}U\sum_{u\leq U}|\alpha_u|^2 \ll U\log Q\sum_{u\leq U}|\alpha_u|^2 ,$$

by the Cauchy–Schwarz inequality and the fact that

$$\sum_{q\leq Q}\frac{1}{\phi(q)} = \sum_{q\leq Q}\frac{1}{q}\prod_{p|q}\left(1-\frac{1}{p}\right)^{-1} \ll \sum_{q\leq Q}\frac{1}{q}\prod_{p|q}\left(1+\frac{1}{p}\right) \leq \sum_{q\leq Q}\frac{1}{q}\sum_{d|q}\frac{1}{d} = \sum_{d\leq Q}\frac{1}{d^2}\sum_{m\leq Q/d}\frac{1}{m} \ll \log Q.$$

The first term can be bounded using the first part of the Coprime Large Sieve, similarly as above, showing it is$\ll (Q+U\log Q)\sum_{u\leq U}|\alpha_u|^2$.

Putting everything together, we finally deduce that the sum we wanted to bound is

$$\ll \sqrt{(Q+U\log Q)\sum_{u\leq U}|\alpha_u|^2(Q+\frac{V\log^4 Q}{R})\sum_{v\leq V}|\beta_v|^2}$$

$$\ll \quad (Q+\sqrt{QU\log Q}+\sqrt{\frac{QV\log^4 Q}{R}}+\sqrt{\frac{UV\log^5 Q}{R}})\sqrt{\sum_{u\leq U}|\alpha_u|^2\sum_{v\leq V}|\beta_v|^2} ,$$

which is good enough for the weaker bound that we said we would prove.

To prove Theorem 7.1, we must now show that we can express the sum over primes less than x in terms of double sums $^{\mathrm{P}}{}_{u\leq U,v\leq V}\alpha_u\beta_v$, where each of U and V is a bit less than x and where the sequence β_v is well distributed to small moduli (so we can take R moderately large). To do this it is helpful to work with prime powers as well as primes.

Definition 7.6. We define the *von Mangoldt function* $\Lambda(n)$ to be log$p$ if $n = p^k$ for some prime p and some $k \geq 1$, and to be zero if n is not a prime power.

As you might have already seen on the first problem sheet, for any natural number $n = p_1^{a_1}...p_k^{a_k}$, with the p_i distinct primes and $a_i \in$ N ∪ {0}, we have

X
$\Lambda(d) = a_1\log p_1 + a_2\log p_2 + ... + a_k\log p_k = \log n.$
$d|n$

Proposition 7.7 (Vaughan's Identity, 1977)**.** *Let $y,z \geq 1$ be any parameters. Then for any natural number $n > z$, we have*

□ □ □ □

$\Lambda(n) = {}^{\mathrm{X}}\mu(b)^{□}{}_{□}\log(n/b) - {}^{\mathrm{X}}\Lambda(c)^{□}{}_{□} + {}^{\mathrm{X}}\mu(b)^{□}{}_{□}\log(n/b) - {}^{\mathrm{X}}\Lambda(c)^{□}{}_{□},$
□ □ □ □ $b|n$, $c|(n/b)$, $b|n$, $c|(n/b)$, $b\leq y$ $c\leq z$ $b>y$ $c\leq z$

where $\mu(c)$ denotes the M¨obius function.

The parameter y could obviously be ignored in Vaughan's Identity, but we include it because in applications it is usual to split the sum into small and large b. Note that if $b \geq n/z$ then $\log(n/b) - {}^{\mathrm{P}}{}_{c|(n/b),}\ \Lambda(c) = \log(n/b) - {}^{\mathrm{P}}{}_{c|(n/b)}\ \Lambda(c) = 0$, so the sum
$c \leq z$
over b can anyway be restricted to those $b < n/z$. So Vaughan's Identity does give a decomposition of the von Mangoldt function into sums where the ranges of summation aren't too long.

Proof of Proposition 7.7. One can check (it is an exercise on the first problem sheet) that ${}^{\mathrm{P}}d|n\ \mu(d) = \mathbf{1}_{n=1}$, where $\mathbf{1}$ denotes the indicator function. Combining this observation with the fact that ${}^{\mathrm{P}}{}_{d|n}\ \Lambda(d) = \log n$, we see that for all $n \in$ N we have

$\Lambda(n) = {}^{\mathrm{X}}\Lambda(c)\ {}^{\mathrm{X}}\ \mu(b) = {}^{\mathrm{X}}\mu(b)\ {}^{\mathrm{X}}\ \Lambda(c) = {}^{\mathrm{X}}\mu(b)\log(n/b).$
$c|n \quad b|(n/c) \quad b|n \quad c|(n/b) \quad b|n$

(These manipulations are an example of *M¨obius inversion*.) Rewriting the above a little bit, we obtain that

□ □

$\Lambda(n) = {}^{\mathrm{X}}\mu(b)$□□$\log(n/b) - {}^{\mathrm{X}}\ \Lambda(c)$□□ $+ {}^{\mathrm{X}}\mu(b)\ {}^{\mathrm{X}}\ \Lambda(c).$

□ □

$b|n \quad c|(n/b), b|n \quad c|(n/b), c \leq z \quad c \leq z$

Then if we swap the order of the sums in the second term again, we find it is

X X X X X
$\mu(b) \quad \Lambda(c) = \quad \Lambda(c) \quad \mu(b) = \quad \Lambda(c)\mathbf{1}_{c=n}.$
$b|n \quad c|(n/b), c|n, \quad b|(n/c)\ c|n, c \leq z \quad c \leq z \quad c \leq z$

Since we assume that $n > z$, this sum is always zero.

Finally, Vaughan's Identity follows by simply splitting the sum over b into $b \leq y$ and $b > y$.

Using Vaughan's Identity, and a bit of tidying up to get rid of the prime powers and remove any extra conditions in our sums, we can express $\pi(x;q,a) - \frac{\pi(x)}{\phi(q)}$ in terms of
double sums of the form P$1 \leq u \leq U, 1 \leq v \leq V$, $\alpha_u\beta_v$.
$uv \equiv a$ mod q

Proposition 7.8. *Let x be large, and let $\delta > x^{-1/5}$ be a small parameter. Then for any*
$\sqrt{\_}$
$1 \leq q \leq x$ *and any* $(a,q) = 1$, *we have thatis* $\pi(x;q,a) - \frac{\pi(x)}{\phi(q)}$ *equal to*

$$\sum_{N,M} \frac{1}{\log(NM)} \left(\sum_{N<b\leq(1+\delta)N} \mu(b) \sum_{\substack{M<m\leq(1+\delta)M,\\ mb\equiv a\ mod\ q}} w(m) - \frac{1}{\phi(q)} \sum_{N<b\leq(1+\delta)N} \mu(b) \sum_{\substack{M<m\leq(1+\delta)M,\\ (mb,q)=1}} w(m) \right)$$

$$+O(\sqrt{x} + \frac{\delta x \log x}{q}),$$

where the outer sum is over values of N and M of the form $N = x^{1/5}(1+\delta)^j \leq x^{4/5}, M = x^{1/5}(1+\delta)^k \leq x/N$, and where

ISSN NO: 2770-0003
Date of Publication: 06-05-2023

$w(m) := \log m - {}^{X} \Lambda(c) \qquad \forall m \in \mathrm{N}.$

$c|m,\ c \le x1/5$

Proof of Proposition 7.8. Note first that $\pi(x;q,a)$ is

$$= \sum_{\substack{x^{1/5}<p\le x,\\ p\equiv a \bmod q}} 1 + O(1+\frac{x^{1/5}}{q}) = \sum_{\substack{x^{1/5}<n\le x,\\ n\equiv a \bmod q}} \frac{\Lambda(n)}{\log n} + O(\sum_{k\ge 2} \sum_{\substack{p^k\le x,\\ p^k\equiv a \bmod q}} 1+1+\frac{x^{1/5}}{q})$$

$$= \sum_{\substack{x^{1/5}<n\le x,\\ n\equiv a \bmod q}} \frac{\Lambda(n)}{\log n} + O(\sqrt{x}).$$

√

This is because we trivially have ${}^{\mathrm{P}}{}_{p}2_{\le x,}$ $1 \le$ x, and $\sum_{k\ge 3} \sum_{\substack{p^k\le x,\\ p^k\equiv a \bmod q}} 1 \ll x^{1/3}\log x \ll$

√x since the sum over p^k is empty if${}^{p}2^{\equiv a \bmod}k >^{q}$ (log$x$)/log2. (This treatment of the “big

Oh” term could be improved a lot, but we won’t need to do so.) The same argument shows that

$$\frac{\pi(x)}{\phi(q)} = \frac{1}{\phi(q)} \sum_{\substack{p\le x,\\ (p,q)=1}} 1 + O(\frac{\log q}{\phi(q)}) = \frac{1}{\phi(q)} \sum_{\substack{x^{1/5}<n\le x,\\ (n,q)=1}} \frac{\Lambda(n)}{\log n} + O(\sqrt{x})$$

.

Next we shall apply Vaughan’s Identity (Proposition 7.7) with the choices $y = z = x^{1/5}$. Together with our earlier observation that $w(m) = 0$ if $m \le x^{1/5}$, the proposition implies that

$$\sum_{\substack{x^{1/5}<n\le x,\\ n\equiv a \bmod q}} \frac{\Lambda(n)}{\log n} = \sum_{\substack{x^{1/5}<n\le x,\\ n\equiv a \bmod q}} \frac{1}{\log n} \left(\sum_{\substack{b|n,\\ b\le x^{1/5}}} \mu(b)w(n/b) + \sum_{\substack{b|n,\\ x^{1/5}<b<x^{4/5}}} \mu(b)w(n/b) \right)$$

$$= \sum_{b\le x^{1/5}} \mu(b) \sum_{\substack{x^{1/5}<m\le x/b,\\ mb\equiv a \bmod q}} \frac{w(m)}{\log(mb)} + \sum_{x^{1/5}<b<x^{4/5}} \mu(b) \sum_{\substack{x^{1/5}<m\le x/b,\\ mb\equiv a \bmod q}} \frac{w(m)}{\log(mb)}$$

.

Applying Vaughan’s Identity in the same way to $\frac{1}{\phi(q)} \sum_{\substack{x^{1/5}<n\le x,\\ (n,q)=1}} \frac{\Lambda(n)}{\log n}$, and taking the difference, we deduce that$\pi(x;q,a) - \frac{\pi(x)}{\phi(q)}$ is

$$= \sum_{b\le x^{1/5}} \mu(b) \left(\sum_{\substack{x^{1/5}<m\le x/b,\\ mb\equiv a \bmod q}} \frac{w(m)}{\log(mb)} - \frac{1}{\phi(q)} \sum_{\substack{x^{1/5}<m\le x/b,\\ (mb,q)=1}} \frac{w(m)}{\log(mb)} \right)$$

$$+ \sum_{x^{1/5}<b<x^{4/5}} \mu(b) \sum_{\substack{x^{1/5}<m\le x/b,\\ mb\equiv a \bmod q}} \frac{w(m)}{\log(mb)} - \frac{1}{\phi(q)} \sum_{x^{1/5}<b<x^{4/5}} \mu(b) \sum_{\substack{x^{1/5}<m\le x/b,\\ (mb,q)=1}} \frac{w(m)}{\log(mb)} + O(\sqrt{x})$$

.

It will be an exercise on the second problem sheet to show the first line above is $O(x^{2/5})$. The second line is almost what we want, except that the sums over m and b are linked by the factor $1/\log(mb)$ and the condition $m \le x/b$, which we shall fix by breaking the sums into short intervals. Indeed, we see

since $|w(m)| \le \log m$ and since the only error arises from terms with $m > x/b$. Since
√
one of the divisors b,m must always be ≤ $2x$, this “big Oh” term is

$$\sum_{x^{1/5}<b<x^{4/5}} \mu(b) \sum_{\substack{x^{1/5}<m\le x/b,\\ mb\equiv a \bmod q}} \frac{w(m)}{\log(mb)} = \sum_{N} \sum_{N<b\le(1+\delta)N} \mu(b) \sum_{\substack{x^{1/5}<m\le x/b,\\ mb\equiv a \bmod q}} \frac{w(m)}{\log(mb)}$$

$$= \sum_{N,M} \sum_{N<b\le(1+\delta)N} \mu(b) \sum_{\substack{M<m\le(1+\delta)M,\\ mb\equiv a \bmod q}} \frac{w(m)}{\log(mb)} + O\left(\sum_{\substack{x<bm\le(1+\delta)x,\\ mb\equiv a \bmod q}} 1 \right)$$

$$\ll \sum_{b\le\sqrt{2x}} \sum_{\substack{x<bm\le(1+\delta)x,\\ mb\equiv a \bmod q}} 1 \ll \sum_{b\le\sqrt{2x}} (\frac{\delta x}{bq}+1) \ll \frac{\delta x \log x}{q} + \sqrt{x},$$

which is acceptably small. The error term when replacing $\log(mb) = \log(NM) + O(\delta)$ by $\log(NM)$ can be handled in the same way, as can the other collection of sums

$$\frac{1}{\phi(q)} \sum_{x^{1/5}<b<x^{4/5}} \mu(b) \sum_{\substack{x^{1/5}<m\le x/b,\\ (mb,q)=1}} \frac{w(m)}{\log(mb)} .$$

Finally, by combining Proposition 7.8 (the consequence of Vaughan’s Identity) with Proposition 7.4 (the consequence of the large sieve) and a little bit of zeta function input, we can prove the Bombieri–Vinogradov theorem.

Proof of Theorem 7.1. Let $A > 0$. We want to show there exists $B = B(A)$ such that

$$\sum_{q\le\sqrt{x}/\log^B x} \max_{(a,q)=1} \left| \pi(x;q,a) - \frac{1}{\phi(q)} \int_2^x \frac{dt}{\log t} \right| \ll \frac{x}{\log^A x}$$

for all large x. By the triangle inequality, the left hand side is

$$\le \sum_{q\le\sqrt{x}/\log^B x} \max_{(a,q)=1} \left| \pi(x;q,a) - \frac{\pi(x)}{\phi(q)} \right| + \sum_{q\le\sqrt{x}/\log^B x} \frac{1}{\phi(q)} \left| \pi(x) - \int_2^x \frac{dt}{\log t} \right| ,$$

and here the second term is $\ll (\log x) \left| \pi(x) - \int_2^x \frac{dt}{\log t} \right|$. By a strong form of the Prime Number Theorem (which we shall assume from zeta function theory) this is $\ll_A x/\log^A x$ for any $A > 0$. By Proposition 7.8 and the triangle inequality, the first term is

$$\le \sum_{N,M} \frac{1}{\log(NM)} \sum_{q\le\sqrt{x}/\log^B x} \max_{(a,q)=1} \left| \sum_{\substack{N<b\le(1+\delta)N,\\ M<m\le(1+\delta)M,\\ mb\equiv a \bmod q}} \mu(b)w(m) - \frac{1}{\phi(q)} \sum_{\substack{N<b\le(1+\delta)N,\\ M<m\le(1+\delta)M,\\ (mb,q)=1}} \mu(b)w(m) \right|$$

$$+O\left(\sum_{q\le\sqrt{x}/\log^B x} (\sqrt{x} + \frac{\delta x \log x}{q}) \right),$$

ISSN NO: 2770-0003
Date of Publication: 06-05-2023

where $\delta > x^{-1/5}$ is a small parameter that we may choose, and the outer sum is over values $N = x^{1/5}(1 + \delta)^j \le x^{4/5}$,$M = x^{1/5}(1 + \delta)^k \le x/N$. In particular, the “big Oh” term is
$\ll x/\log^B x + \delta x \log^2 x,$

so it will be small enough if we choose $\delta = 1/\log^{A+2} x$ and $B \ge A$.

Finally we come to the heart of the proof, where we apply Proposition 7.4 to all the inner sums in the first term (with $(\beta_v) := (\mu(b))_{N<b\le(1+\delta)N}$ and $(\alpha_u) := (w(m))_{M<m\le(1+\delta)M}$), deducing it is

$$\ll \sum_{N,M} \frac{1}{\log(NM)} \left(\frac{\sqrt{x}}{\log^B x} + \sqrt{M+N}\log^2 x + \frac{\sqrt{MN}\log^3 x}{\sqrt{R}} \right) \sqrt{\sum_{M<m\le(1+\delta)M} w(m)^2 \sum_{N<b\le(1+\delta)N} 1} \, .$$

Here the quantity R measures the equidistribution of $(\mu(b))_{N<b\le(1+\delta)N}$ to small moduli, and it is a zeta function type fact (equivalent to the Siegel–Walfisz theorem) that we may take $R = \log^C x$ for any fixed $C > 0$, provided x is large enough. We have $w(m) \le \log m \le \log x$ and $M + N \le x^{4/5}$ and $x^{2/5} \le MN \le x$, so the above is

$$\begin{aligned} &\ll \sum_{N,M} \frac{1}{\log x} \left(\frac{\sqrt{x}}{\log^B x} + x^{2/5}\log^2 x + \frac{\sqrt{x}\log^3 x}{\log^{C/2} x} \right) \sqrt{x \log^2 x} \\ &\ll (\delta^{-1}\log x)^2 \left(\frac{x}{\log^B x} + x^{9/10}\log^2 x + \frac{x\log^3 x}{\log^{C/2} x} \right) \\ &\ll \log^{2(A+3)} x \left(\frac{x}{\log^B x} + x^{9/10}\log^2 x + \frac{x\log^3 x}{\log^{C/2} x} \right), \end{aligned}$$

since there are $\ll \delta^{-1}\log x$ values of N and of M. Provided $B \ge 3A + 6$ the first and second terms here will be acceptably small, and provided $C \ge 6A + 18$ the third term will be acceptably small.

References

1. H. Davenport. *Multiplicative Number Theory*. Third edition, revised by Hugh L. Montgomery, published by Springer Graduate Texts in Mathematics. 2000
2. J. Friedlander and H. Iwaniec. *Opera de Cribro*. AMS Colloquium Publications, vol. 57. 2010
3. H. Iwaniec and E. Kowalski. *Analytic Number Theory*. AMS Colloquium Publications, vol. 53. 2004