7. Applications of the Large Sieve to Number Theory

Linnik initially used the large sieve to study the distribution of quadratic residues. We will see that work on Problem set 4.

Perhaps the most important application of the large sieve in number theory concerns the distribution of primes mod q.

7.1. Distribution of primes mod q.

Let $\pi(N)$ denote the number of primes less than or equal to N. Let $\pi(N,q,a)$ be the number of primes p satisfying $p \leq N$ and $p = a \mod q$. We want to focus on $a \in \mathbb{Z}_q^*$, since if a and q are not relatively prime, $\pi(N,q,a)$ is at most one. So let $\phi(q) = |\mathbb{Z}_q^*|$. If the primes were evenly distributed mod q, then $\pi(N,q,a)$ would be close to $\frac{\pi(N)}{\phi(q)}$. To quantify how badly this fails, we introduce the function

$$\Delta_q(N) := \max_{a \in \mathbb{Z}_q^*} \left| \pi(N,q,a) - \frac{\pi(N)}{\phi(q)} \right|.$$

Here are some results on $\Delta_q(N)$:

Theorem 7.1 (Dirichlet). *For all q,*

$$\lim_{N \to \infty} \frac{\Delta_q(N)}{N/q} = 0.$$

Theorem 7.2 (Siegel-Walfisz). *For any A, there is some c_A such that*

$$\Delta_q(N) \leq c_A N (\log N)^{-A}.$$

This is the best result that applies to all q. If one assumes the generalized Riemann hypothesis, then it is true that

$$\Delta_q(N) \leq (C_\epsilon N^\epsilon) N^{1/2}$$

for any $\epsilon > 0$. Montgomery conjectured that for any $\epsilon > 0$, there is a constant C_ϵ such that $\Delta_q(N) \leq (C_\epsilon N^\epsilon) \left(\frac{N}{q}\right)^{1/2}$.

Instead of trying to understand what happens for all q, we will be concerned with the typical behavior of $\Delta_q(N)$. The theorem we will discuss is

Theorem 7.3 (Renyi, Bombiere–Vinogradov). *For all $\epsilon > 0$ and all A,*

$$\sum_{q \leq N^{1/2-\epsilon}} \Delta_q(N) \leq C(\epsilon, A) N (\log N)^{-A}.$$

This says that for most $q \leq N^{1/2-\epsilon}$, $\Delta_q(N) \leq \frac{N}{q}(\log N)^{-A} \ll \frac{N}{q}$. So the primes are close to equidistributed mod q for most q up to $N^{1/2-\epsilon}$.

We will not give the complete proof, which is somewhat messy, but we will discuss most of the main ideas. In particular, we will explain how the large sieve and projection theory enter the story.

7.2. **Multiplicative Convolution and Primes.** To prove this, we will use the **multiplicative convolution**, which interacts nicely with prime numbers and projections.

Definition 7.4. *If $f, g : \mathbb{N} \to \mathbb{C}$, then their **multiplicative convolution** is the function*

$$f *_M g(n) = \sum_{n_1, n_2, n_1 n_2 = n} f(n_1) g(n_2).$$

This is related to the prime numbers through the **sieve**. Sieving is the process of obtaining prime numbers by crossing off all the multiples of 2, then all the multiples of 3, and so on, until only the primes are left. If you try to write this down with a formula, the multiplicative convolution will appear. Let $1 = 1_{\mathbb{N}}$ and define

$$D_p(n) = \begin{cases} 1 & n = 1, \\ -1 & n = p, \\ 0 & n \neq 1, p. \end{cases}$$

Then we can calculate

$$1_{\mathbb{N}} *_M D_2 = 1_{\mathbb{N}} - 1_{2\mathbb{N}} = 1_{\text{odd}}.$$

Similarly, $1_{\mathbb{N}} *_M D_2 *_M D_3$ is the indicator function for n relatively prime to 2 and 3.

For a set of primes S, define

$$RP_S(n) = \begin{cases} 1 & (p, n) = 1 \ \forall p \in S, \\ 0 & \text{else.} \end{cases}$$

Note that if $S = P_{N^{1/2}}$ and $N^{1/2} < n \leq N$, then $RP_S(n) = P(n)$.

Lemma 7.5. *If $S = \{p_1, \ldots, p_r\}$, then*

$$RP_S(n) = 1 *_M D_{p_1} *_M \ldots *_M D_{p_r}.$$

7.3. **Multiplicative Convolution and Projections.** Now we will examine the relationship between multiplicative convolution and projection. Multiplicative convolution interacts nicely with the projection $\mathbb{Z} \to \mathbb{Z}_q$ because this projection is a ring homomorphism.

Lemma 7.6 (Lemma 1). *If $f, g : \mathbb{N} \to \mathbb{C}$, then $\pi_q(f *_M g) = \pi_q f *_M \pi_q g$.*

To be extra careful, we should say what we mean by multiplicative convolution in $\mathbb{Z}_q$:

$$F *_M G(a) = \sum_{a_1, a_2 \in \mathbb{Z}_q, a_1 a_2 = a} F(a_1) G(a_2)$$

for functions $F, G : \mathbb{Z}_q \to \mathbb{C}$.

Proof. Write

$$f = \sum_{n_1} \delta_{n_1} f(n_1), \; g = \sum_{n_2} \delta_{n_2} g(n_2).$$

Then

$$f *_M g = \sum_{n_1, n_2} \delta_{n_1 n_2} f(n_1) g(n_2).$$

Here δ_n is the **delta function** $\delta_n(m) = \begin{cases} 1 & n = m \\ 0 & \text{else} \end{cases}$. Then

$$\begin{aligned} \pi_q f(a) &= \sum_{n_1} \delta_{n_1 \mod q}(a) f(n_1), \\ \pi_q f *_M \pi_q g(a) &= \sum_{n_1, n_2} \delta_{n_1 n_2 \mod q} f(n_1) g(n_2) \\ &= \pi_q (f *_M g)(a). \end{aligned}$$

□

For our final result, we want L^∞ bounds, but our theory is geared toward L^2 bounds. Here's how we can get L^∞ bounds:

Lemma 7.7 (Lemma 2)**.** *If $f, g : \mathbb{N} \to \mathbb{C}$, then*

$$\|f *_M g\|_{L^\infty(\mathbb{Z}_q^*)} \leq \|f\|_{L^2} \|g\|_{L^2}.$$

Proof. For $a \in \mathbb{Z}_q^*$, $f *_M g(a) = \sum_{b \in \mathbb{Z}_q^*} f(b) g(ab^{-1}) \leq \|f\|_{L^2} \|g\|_{L^2}$ by Cauchy-Schwarz. □

There is also the minor technical annoyance of switching between $\mathbb{Z}_q$ and $\mathbb{Z}_q^*$. If $f : \mathbb{Z}_q \to \mathbb{C}$, let $f^* : \mathbb{Z}_q^* \to \mathbb{C}$ be the restriction. Then we can write $f = f_0 + f_h$ and $f^* = f_0^* + f_h^*$, where the starred functions are defined on $\mathbb{Z}_q^*$ and the unstarred functions are defined on $\mathbb{Z}_q$, the subscript zero indicates a constant function, and the subscript h indicates an average zero function.

Lemma 7.8 (Lemma 3)**.**

$$\|f_h^*\|_{L^2(\mathbb{Z}_q^*)} \leq \|f_h\|_{L^2(\mathbb{Z}_q)}.$$

Finally, taking the high frequency part commutes with multiplicative convolution:

Lemma 7.9 (Lemma 4)**.** *If $f^*, g^* : \mathbb{Z}_q^* \to \mathbb{C}$, then*

$$(f^* *_M g^*)_h = f_h^* *_M g_h^*.$$

If we combine all of these, we get the following proposition:

Proposition 7.10.

$$\|(\pi_q(f *_M g))_h^*\|_{L^\infty} \le \|(\pi_q f)_h\|_{L^2} \|(\pi_q g)_h\|_{L^2}.$$

Proof. By Lemma 1 then Lemma 4,

$$(\pi_q(f *_M g))_h^* = ((\pi_q f *_M \pi_q g))_h^* = (\pi_q f)_h^* *_M (\pi_q g)_h^*.$$

Then using Lemma 2 and Lemma 3, we get

$$\begin{aligned} \|(\pi_q(f *_M g))_h^*\|_{L^\infty} &\le \|(\pi_q f)_h^*\|_{L^2} \|(\pi_q g)_h^*\|_{L^2} \\ &\le \|(\pi_q f)_h\|_{L^2} \|(\pi_q g)_h\|_{L^2}. \end{aligned}$$

□

7.4. **Large Sieve and Multiplicative Convolution.** Our goal is to prove that $P(n)$ is evenly distributed mod q for most q of a given size. We will focus on the case that q is prime, which avoids technical issues but still shows the main proof ideas.

We have seen that for a large range of n, $P(n)$ is equal to $RP_S(n)$, where $S = P_{N^{1/2}}$. The key property of $RP_S(n)$ is that it is a multiplicative convolution. Our next theorem shows that most projections of a multiplicative convolution are nearly constant – it is the main analytic ingredient in the proof of Bombieri-Vinogradov.

Theorem 7.11. *If $f : [N_1] \to \mathbb{C}$ and $g : [N_2] \to \mathbb{C}$, then $f *_M g : [N] \to \mathbb{C}$, where $N = N_1 N_2$, and*

$$\sum_{p \in P_M} \|(\pi_q(f *_M g))_h^*\|_{L^\infty}^2 \lesssim \left(\left(\frac{N_1}{M} + M \right) \left(\frac{N_2}{M} + M \right) \right)^{1/2} \|f\|_{L^2} \|g\|_{L^2}.$$

Proof. We apply the proposition, Cauchy-Schwarz, and then the large sieve:

$$\begin{aligned} \sum_{p \in P_M} \|(\pi_p(f *_M g))_h^*\|_{L^\infty}^2 &\le \sum_{p \in P_M} \|(\pi_p f)_h\|_{L^2} \|(\pi_p g)_h\|_{L^2} \\ &\le \left(\sum_{p \in P_M} \|(\pi_p f)_h\|_{L^2}^2 \right)^{1/2} \left(\sum_{p \in P_M} \|(\pi_p g)_h\|_{L^2}^2 \right)^{1/2} \\ &\lesssim \left(\left(\frac{N_1}{M} + M \right) \left(\frac{N_2}{M} + M \right) \right)^{1/2} \|f\|_{L^2} \|g\|_{L^2}. \end{aligned}$$

□

For the main theorem, we have $|f(n)|, |g(n)| \lessapprox 1$, so $\|f\|_{L^2}^2 \lessapprox N_1$ and $\|g\|_{L^2}^2 \lessapprox N_2$, so

$$\sum_{p \in P_M} \|(\pi_q(f *_M g))_h^*\|_{L^\infty}^2 \lessapprox \frac{N}{M} + \sqrt{N_1 N} + \sqrt{N_2 N} + M\sqrt{N}.$$

This will be good if $M \le N^{1/2-\epsilon}$ and $N_1, N_2 \ll N$. We cannot have N_1 or N_2 close to N, because in that case the other factor will be close to 1 and the multiplicative convolution will not result in a more evenly spread function. And the first condition must be true for the projection theory methods to be able to say anything.

Finally, we give a rough outline the proof of the Bombieri-Vinogradov theorem for q prime.

I am actually not sure whether the full BV theorem can be proven following this outline. The proof in books is based on a different way of finding multiplicative convolution structure in the primes, which is called Vaughn's identity. Vaughn's identity is more efficient and leads to fewer terms, but I found it a little harder to motivate.

Let $S = P_{<N^{1/2}}$. If $N^{1/2} < n < N$, $RP_S(n) = P(n)$. Also

$$\begin{aligned} RP_S(N) &= [1 *_M D_{p_1}] *_M [\ldots *_M D_{p_R}] \\ &= f *_M g \\ &= \left(\sum_{I_1} f 1_{I_1}\right) *_M \left(\sum_{I_2} g 1_{I_2}\right) \\ &= \sum_{I_1, I_2} f 1_{I_1} *_M g 1_{I_2}. \end{aligned}$$

Here I_1 and I_2 are intervals that are narrower than dyadic intervals. Let $N_1 = \min I_1$ and $N_2 = \min I_2$. For $n \le N$,

$$RP_S(n) = \sum_{I_1, I_2, N_1 \cdot N_2 \le N} f 1_{I_1} *_M g 1_{I_2}.$$

We can then apply the theorem above for each pair of intervals. This works when $1 \ll N_1, N_2 \ll N$. Otherwise, we must group the convolutions for $RP_S(N)$ differently. It is a possible course project to think this through carefully and see what bounds it gives.

18.156 Projection Theory
Spring 2025