Selberg's Sieve

University of Waterloo

Peiran Tao

1 Introduction

Recall in the Sieve of Eratosthenes, we have the setup:

Definition. Let A be a finite subset of $\mathbb{N}$. Let P be a set of primes and let $z > 0$ be a real number. Define:

$$S(A, P, z) = \sum_{\substack{a \in A \\ (a, P(z)) = 1}} 1$$

where:

$$P(z) = \prod_{\substack{p \in P \\ p < z}} p$$

With these setup, we can deduce that:

$$S(A, P, z) = \sum_{a \in A} \sum_{d \mid (a, P(z))} \mu(d) \tag{1.1}$$

using the property of the Möbius function that:

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

Selberg came up with this brilliant ideal to replace $\sum \mu(d)$ in (1.1) with a quadratic form, chosen optimally to make the result minimal. That is, let $(\lambda_d) \subseteq \mathbb{R}$ be a sequence such that $\lambda_1 = 1$, then:

$$\sum_{d \mid n} \mu(d) \leq \left(\sum_{d \mid n} \lambda_d \right)^2 \tag{1.2}$$

because the LHS is at most 1.

Recall the following setup we used to estimate $\pi(x)$. Let:

$$\pi(x, z) = \{n \leq x : p \mid n \Rightarrow p \geq z\}$$

be the number of $1 \leq n \leq x$ that are not divisible by any prime $p < z$. If we let $A = [1, x] \cap \mathbb{Z}$ and $P =$ all primes, then:

$$\pi(x, z) = S(A, P, z)$$

Then we have:

$$\pi(x, z) = \sum_{\substack{n \leq x \\ p \mid n \Rightarrow p \geq z}} 1 = 1 + \sum_{\substack{1 < n \leq z \\ p \mid n \Rightarrow p \geq z}} 1 + \sum_{\substack{z < n \leq x \\ p \mid n \Rightarrow p \geq z}} 1$$

The first sum is clearly 0. The second sum certainly counts all prime numbers p with $z < p \leq x$ and the number of such primes is $\pi(x) - \pi(z)$, hence:

$$\pi(x, z) \geq 1 + \pi(x) - \pi(z)$$

Rearrange them and use the fact that $\pi(z) \leq z$, we have:

$$\pi(x) \leq 1 + z + \pi(x, z) \tag{1.3}$$

Now it suffices to bound $\pi(x, z) = S(A, P, z)$. Let us see how to do this in full generality, then we come back to this problem.

2 Main Theorem

As always, let A, P, z be given as usual. For each $p \in P$, define:

$$A_p = \{a \in A : p \mid a\}$$

Moreover, for all squarefree integer d composed of primes in P, define $A_d = \bigcap_{p|d} A_p$. Suppose there is a multiplicative function f with $f(p) > 1$ for all $p \in P$, and for all d we have:

$$|A_d| = \frac{X}{f(d)} + R_d \tag{2.1}$$

to be the estimation of $|A_d|$, where X is an estimation of A and R_d is the error term.

Theorem 2.1 (**Selberg's Sieve**)**.** With the setting above. Let f_1 be the unique function such that:

$$f(n) = \sum_{d|n} f_1(d) \tag{2.2}$$

Also, we define:

$$V(z) = \sum_{\substack{d<z \\ d|P(z)}} \frac{\mu^2(d)}{f_1(d)} \tag{2.3}$$

Then we have:

$$S(A, P, z) \leq \frac{X}{V(z)} + \left(\sum_{\substack{d_1, d_2 \leq z \\ d_1, d_2 | P(z)}} |R_{[d_1,d_2]}| \right) \tag{2.4}$$

Lemma 2.2. Let f_1, f_2 be a multiplicative function and d_1, d_2 be positive squarefree integers, then:

$$f([d_1, d_2]) f((d_1, d_2)) = f(d_1) f(d_2) \tag{2.5}$$

Proof of Selberg's Sieve: Let (λ_d) be a sequence of real numbers with $\lambda_1 = 1$ and $\lambda_d = 0$ for all $d > z$. Then by (1.2) we have:

$$\begin{aligned} S(A, P, z) = \sum_{\substack{a \in A \\ (a, P(z))=1}} 1 &= \sum_{a \in A} \sum_{d|(a,P(z))} \mu(d) \leq \sum_{a \in A} \left(\sum_{d|(a,P(z))} \lambda_d \right)^2 = \sum_{a \in A} \left(\sum_{d_1, d_2 | (a, P(z))} \lambda_{d_1} \lambda_{d_2} \right) \\ &= \sum_{d_1, d_2 | P(z)} \lambda_{d_1} \lambda_{d_2} \sum_{\substack{a \in A \\ d_1, d_2 | a}} 1 = \sum_{d_1, d_2 | P(z)} \lambda_{d_1} \lambda_{d_2} \sum_{\substack{a \in A \\ [d_1, d_2] | a}} 1 = \sum_{d_1, d_2 | P(z)} \lambda_{d_1} \lambda_{d_2} |A_{[d_1, d_2]}| \end{aligned}$$

Now using (2.1) and (2.5) we have:

$$\begin{aligned}
S(A,P,z) &= X\sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} f([d_1,d_2]) + \sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} r_{[d_1,d_2]} \\
&= X\sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} \frac{f(d_1)f(d_2)}{f((d_1,d_2))} + \sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} r_{[d_1,d_2]} \\
&= XT + R
\end{aligned}$$

where we defined:

$$T = \sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} \frac{f(d_1)f(d_2)}{f((d_1,d_2))} = \sum_{\substack{d_1,d_2\le z \\ d_1,d_2|P(z)}} \lambda_{d_1}\lambda_{d_2} \frac{f(d_1)f(d_2)}{f((d_1,d_2))} \tag{2.6}$$

so that XT is our main term, and:

$$R = \sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} r_{[d_1,d_2]} = \sum_{\substack{d_1,d_2\le z \\ d_1,d_2|P(z)}} \lambda_{d_1}\lambda_{d_2} r_{[d_1,d_2]} \tag{2.7}$$

to be our error term. Let us analyze T first. Our main term is a quadratic form in (λ_d), and remember, we want to minimize it to get a good upper bound. To do this, we will first transform it into a diagonal form.

$$\begin{aligned}
T &= \sum_{\substack{d_1,d_2\le z \\ d_1,d_2|P(z)}} \frac{\lambda_{d_1}\lambda_{d_2}}{f(d_1)f(d_2)} f((d_1,d_2)) \\
&= \sum_{\substack{d_1,d_2\le z \\ d_1,d_2|P(z)}} \frac{\lambda_{d_1}\lambda_{d_2}}{f(d_1)f(d_2)} \sum_{\delta|(d_1,d_2)} f_1(\delta) && \text{(by (2.2))} \\
&= \sum_{\substack{\delta\le z \\ \delta|P(z)}} f_1(\delta) \sum_{\substack{d_1,d_2\le z \\ d_1,d_2|P(z) \\ \delta|(d_1,d_2)}} \frac{\lambda_{d_1}\lambda_{d_2}}{f(d_1)f(d_2)} \\
&= \sum_{\substack{\delta\le z \\ \delta|P(z)}} f_1(\delta) u_\delta^2
\end{aligned}$$

where u_δ is defined by:

$$u_\delta = \sum_{\substack{d\le z \\ d|P(z) \\ \delta|d}} \frac{\lambda_d}{f(d)} \tag{2.8}$$

Hence we have transformed our quadratic form to a diagonal form:

$$T = \sum_{\substack{\delta\le z \\ \delta|P(z)}} f_1(\delta) u_\delta^2$$

By dual Möbius Inversion Formula on (2.8) we have:

$$\frac{\lambda(\delta)}{f(\delta)} = \sum_{\substack{d|P(z) \\ \delta|d}} \mu\left(\frac{d}{\delta}\right) u_d \tag{2.9}$$

since $\lambda_d/f(d)$ and u_δ are well-defined on the divisor-closed set $\{\delta < z : \delta \mid P(z)\}$. Let $\delta = 1$, we have:

$$1 = \frac{1}{f(1)} = \sum_{\substack{d|P(z) \\ \delta|d}} \mu(d) u_d = \sum_{d|P(z)} \mu(d) u_d$$

Also, by (2.8), if $\delta \geq z$, then the sum is empty since $z \leq \delta < d < z$. Therefore $u_\delta = 0$ for $\delta \geq z$. Using this, we can write the above equality as:

$$\sum_{\substack{\delta \leq z \\ \delta | P(z)}} \mu(\delta) u_\delta = 1 \tag{2.10}$$

Therefore, we have:

$$\begin{aligned} \sum_{\substack{\delta \leq z \\ \delta | P(z)}} f_1(\delta) \left(u_\delta - \frac{\mu(\delta)}{f_1(\delta)V(z)} \right)^2 &= \sum_{\substack{\delta \leq z \\ \delta | P(z)}} f_1(\delta) u_\delta^2 - 2 \sum_{\substack{\delta \leq z \\ \delta | P(z)}} \frac{f_1(\delta)\mu(d)}{f_1(\delta)V(z)} u_\delta + \sum_{\substack{\delta \leq z \\ \delta | P(z)}} f_1(\delta) \frac{\mu(\delta)^2}{f_1(\delta)^2 V(z)^2} \\ &= T - \frac{2}{V(z)} \sum_{\substack{\delta \leq z \\ \delta | P(z)}} \mu(d) u_\delta + \frac{1}{V(z)^2} \sum_{\substack{\delta \leq z \\ \delta | P(z)}} \frac{\mu(\delta)^2}{f_1(\delta)} \end{aligned}$$

By (2.10) and (2.3), the above sum is equal to:

$$T - \frac{2}{V(z)} + \frac{1}{V(z)} = T - \frac{1}{V(z)}$$

Therefore we have:

$$T = \sum_{\substack{\delta \leq z \\ \delta | P(z)}} f_1(\delta) \left(u_\delta - \frac{\mu(\delta)}{f_1(\delta)V(z)} \right)^2 + \frac{1}{V(z)} \tag{2.11}$$

Note that since $\sum_{d|n} f_1(d) = f(n)$, by Möbius inverison we have:

$$f_1(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) f(d)$$

so when $n = p$ is prime:

$$f_1(p) = \mu(p)f(p) + \mu(1)f(1) = f(p) - 1 > 0$$

By multiplicativity, $f_1(d) > 0$ for all d. Therefore the first sum in (2.10) is always non-negative, so T is minimized when the sum is 0, which is when:

$$u_\delta = \frac{\mu(\delta)}{f_1(\delta)V(z)} \tag{2.12}$$

because $f_1(d)$ is always positive. The minimal value of T is $1/V(z)$.

Now let us look at the error term R. By (2.12) and (2.9) we have:

$$\begin{aligned} V(z)\lambda_\delta = f(\delta) \sum_{\substack{d \leq z \\ d | P(z) \\ \delta | d}} \frac{\mu(d/\delta)\mu(d)}{f_1(\delta)} &= f(\delta) \sum_{\substack{t \leq z/\delta \\ t | P(z) \\ (t,\delta)=1}} \frac{\mu^2(t)\mu(\delta)}{f_1(t) f_1(\delta)} \\ &= \mu(\delta) \left(\sum_{p|\delta} \frac{f(p)}{f_1(p)} \right) \sum_{\substack{t \leq z/\delta \\ t | P(z) \\ (t,\delta)=1}} \frac{\mu^2(t)}{f_1(t)} \\ &= \mu(\delta) \left(\sum_{p|\delta} \left(1 + \frac{1}{f_1(p)} \right) \right) \sum_{\substack{t \leq z/\delta \\ t | P(z) \\ (t,\delta)=1}} \frac{\mu^2(t)}{f_1(t)} \end{aligned}$$

Therefore we ge t $|V(z)||\lambda_\delta| \leq |V(z)|$ so $|\lambda_\delta| \leq 1$. Hence:

$$R = O\left(\sum_{\substack{d_1,d_2 \leq z \\ d_1,d_2 | P(z)}} |\lambda_{d_1}\lambda_{d_2}||R_{[d_1,d_2]}|\right) = \left(\sum_{\substack{d_1,d_2 \leq z \\ d_1,d_2 | P(z)}} |R_{[d_1,d_2]}|\right)$$

As desired. □

Remark. In fact, in the proof of the theorem, if we analyze R more carefully, we can get a better bound:

$$R = \sum_{\substack{d \leq z^2 \\ d | P(z)}} 3^{\omega(d)} |R_d| \tag{2.13}$$

where $\omega(d) = \sum_{p|d} 1 =$ the number of prime divisors of d.

To use Selberg's Sieve, we need to find a lower bound on $V(z)$. So we have the following lemma:

Lemma 2.3. Let $\tilde{f}$ be a completely multiplicative function such that $\tilde{f}(p) = f(p)$ for all primes p. Let:

$$\overline{P}(z) = \prod_{\substack{p \notin P \\ p < z}} p$$

Then we have:

$$V(z) \geq \sum_{\substack{\delta \leq z \\ p|\delta \Rightarrow p|P(z)}} \frac{1}{\tilde{f}(\delta)} \tag{2.14}$$

and that:

$$f(\overline{P}(z))V(z) \geq f_1(\overline{P}(z)) \sum_{\delta \leq z} \frac{1}{\tilde{f}(\delta)} \tag{2.15}$$

Example. Let us look back at the problem in Section 1. We want to estimate $S(A, P, z)$ with $A = [1, x] \cap \mathbb{Z}$ and $P =$ all primes and $z > 0$. We have:

$$A_d = \{n \leq x : d \mid n\} \implies |A_d| = \left[\frac{x}{d}\right] = \frac{x}{d} + \left\{\frac{x}{d}\right\}$$

Therefore let $X = x$ and $f(d) = d$ and $R_d = \left\{\frac{x}{d}\right\}$. Therefore since $\sum_{d|n} f_1(d) = n$, we have $f_1(d) = \phi(d)$.

$$V(z) = \sum_{\substack{d \leq z \\ d | P(z)}} \frac{\mu^2(d)}{\phi(d)} \geq \sum_{\substack{d \leq z \\ d | P(z)}} \frac{\mu^2(d)}{d} = \sum_{d \leq z} \frac{1}{d} - {\sum_{d \leq z}}' \frac{1}{d}$$

where the sum $\sum'$ is over all non-squarefree integers d. Since:

$$\sum_{d \leq z} \frac{1}{d} = \log z + O(1)$$

and also notice that:

$${\sum_{d \leq z}}' \frac{1}{d} \leq \frac{1}{4} \sum_{d \leq z/4} \frac{1}{d}$$

It follows that:

$$V(z) = \sum_{\substack{d \leq z \\ d | P(z)}} \frac{\mu^2(d)}{\phi(d)} \gg \log z$$

Hence by Selber's Sieve we have:

$$\pi(x,z) = S(A,P,z) \ll \frac{x}{\log z} + z^2$$

here the error term is $\ll z^2$ since $R_d \ll 1$. Pick:

$$z = \left(\frac{x}{\log x}\right)^{1/2}$$

Note that $\log z \gg \log x$, and $z^2 = x/\log x$, so we have:

$$\pi(x,z) \ll \frac{x}{\log x}$$

Hence, combined with (1.3) it follows that:

$$\pi(x) \ll 1 + \left(\frac{x}{\log x}\right)^{1/2} + \frac{x}{\log x} \ll \frac{x}{\log x}$$

As desired!

Remark. Recall that using the Sieve of Eratosthenes, we can only get:

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

This suggests that Selberg's Sieve can give us better upper bound! (Even though it is way harder to derive).

3 The Brun-Titchmarsh Theorem

In this section we will use Selberg's Sieve to estimate the number of primes $p \leq x$ in an arithmetic progession. For $a, k \in \mathbb{Z}$ with $(a,k) = 1$, we define:

$$\pi(x;k,a) = |\{p \leq x : p \equiv a \pmod{k}\}|$$

It was conjectured that $\pi(x;k,a)$ is unbounded as $x \to \infty$, that is, there are infinitely many primes p such that $p \equiv a \pmod{k}$. It was proved by Dirichlet in 1930 that:

Theorem 3.1 (Dirichlet). Let $a, k \in \mathbb{Z}$ be coprime, then:

$$\pi(x;k,a) \sim \frac{1}{\phi(k)} \operatorname{li}(x)$$

as $x \to \infty$. Here $\operatorname{li}(x)$ is the logarithmic integral defined by:

$$\operatorname{li}(x) = \int_2^x \frac{1}{\log t}\, dt$$

In particular, $\lim\limits_{x\to\infty} \operatorname{li}(x) = \infty$, so $\pi(x;k,a) \to \infty$.

But "how many" primes of this form are there? That is, what is its density in $\mathbb{N}$?

Definition. Let A be a subset of $\mathbb{N}$ and $A(n) = A \cap [1,n]$ for $n \in \mathbb{N}$. The **natural density** of A is:

$$\lim_{n\to\infty} \frac{|A(n)|}{n}$$

provided that the limit exists.

Definition. Let A be a set of prime numbers, the **analytic density** of A is:

$$\lim_{s \to 1^+} \frac{\sum\limits_{p \in A} 1/p^s}{\log 1/(s-1)}$$

provided that the limit exists.

It can be shown that, if a set of primes has natural density δ, then it also has analytic density δ. Let P be the set of primes that are $\equiv a \pmod{k}$. Dirichlet proved that P has analytic density $1/\phi(k)$.

There is an effective asymptotic formulae for $\pi(x; k, a)$.

Theorem 3.2 (**Siegel-Walfisz).** For any $N > 0$, there exists $c(N) > 0$ such that if $k \leq (\log x)^N$, then:

$$\pi(x; k, a) = \frac{1}{\phi(k)} \operatorname{li}(x) + O(x \exp(-c(N)(\log x)^{1/2}))$$

uniformly in k.

Thus, the error terms of $|\pi(x; k, a) - \frac{1}{\phi(k)} \operatorname{li}(x)|$ are known (only) in a range of $k < (\log x)^N$. In this section we are going to obtain an upper bound of $\pi(x; k, a)$, using the Selberg's Sieve.

Theorem 3.3 (**Brun-Titchmarsh).** Let a, k be positive integers with $(a, k) = 1$ and let $x > 0$ such that $k \leq x^\theta$ for some $\theta < 1$. Then for any $\epsilon > 0$, there exists $x_0(\epsilon) > 0$ such that:

$$\pi(x; k, a) \leq \frac{(2+\epsilon)x}{\phi(k) \log(2x/k)}$$

for all $x > x_0(\epsilon)$.

Proof: For $z < x$, we note that:

$$\pi(x; k, a) = \pi(z; k, a) + (\pi(x; k, a) - \pi(z; k, a)) \leq z + (\pi(x; k, a) - \pi(z; k, a))$$

Now, call $B = \{z < p \leq x : p \equiv a \pmod{k}\}$, then:

$$\pi(x; k, a) - \pi(z; k, a) = |B| \leq S(A, P, z) \tag{3.1}$$

where A, P are defined by:

$$A = \{n \leq x : n \equiv a \pmod{k}\}$$

and:

$$P = \{p : (p, k) = 1\} = \{p : p \nmid k\}$$

Then $S(A, P, z)$ counts every integer $n \leq x$ with $n \equiv a \pmod{k}$ such that n is not divisible by any $p < z$ with $(p, k) = 1$, and every prime in B has this property, thus (3.1) is true. It now suffices to analyze $S(A, P, z)$.

$$P(z) = \prod_{\substack{p \in P \\ p < z}} p = \prod_{\substack{p < z \\ (p,k)=1}} p$$

For $p \in P$ we have:

$$A_p = \{n \in A : p \mid n\} = \{n \leq x : n \equiv a \pmod{k}, n \equiv 0 \pmod{p}\}$$

So to estimate the size of A_p, it suffices to find all solutions $\leq x$ to the simultaneous congruence:

$$n \equiv a \pmod{k}$$
$$n \equiv 0 \pmod{p}$$

Since $(p, k) = 1$, by Chinese Remainder Theorem, this has a unique solution in $\mathbb{Z}/kp\mathbb{Z}$, and since we need all solutions $\leq x$, there are $[x/kp]$ such solutions. Hence:

$$|A_p| = \left[\frac{x}{kp}\right] = \frac{x}{kp} + O(1)$$

Using the notation in Selberg's Sieve, we know $f(p) = p$ and $X = \frac{x}{k}$ is the esimation of $|A|$. Hence, we have:

$$|A_d| = \frac{x}{kd} + O(1) \tag{3.2}$$

Since $\sum_{d|n} f_1(d) = f(n) = n$, we must have $f_1(d) = \phi(d)$. Also, $R_d = O(1)$ by (3.2). Therefore by Selberg's Sieve we get:

$$S(A, P, z) \leq \frac{x}{kV(z)} + O(z^2)$$

where:

$$V(z) = \sum_{\substack{d \leq z \\ d | P(z)}} \frac{\mu^2(d)}{\phi(d)} = \sum_{\substack{d \leq z \\ (d,k)=1}} \frac{\mu^2(d)}{\phi(d)}$$

here $d \leq z$ and $d \mid P(z) \iff (d, k) = 1$ because $P(z)$ is composed of primes that are coprime to k, so $d \mid P(z)$ if and only if $(d, k) = 1$. Now we would like to use the lemma to estimate $V(z)$. By (2.15), we have:

$$\overline{P}(z)V(z) \geq \phi(\overline{P}(z)) \sum_{\delta \leq z} \frac{1}{\delta}$$

where:

$$\overline{P}(z) = \prod_{\substack{p \leq z \\ p \notin P}} p = \prod_{\substack{p \leq z \\ p | k}} p$$

Now we claim that:

$$\frac{\phi(\overline{P}(z))}{\overline{P}(z)} = \frac{\phi(k)}{k}$$

References

[1] Cojocaru, A.C. and Murty, M.R., An Introduction to Sieve Methods and their Applications. London Mathematical Society 66. Cambridge University Press, 2006.