THE SELBERG SIEVE: LECTURE 2

STEVE LESTER

1. Minimizing quadratic forms

In the last lecture we saw how to reduce the problem of bounding $\pi(x)$ to minimizing a quadratic form

$$Q(\lambda) = \sum_{\substack{d_1, d_2 \le z \\ d_1, d_2, \text{ square-free}}} \frac{\lambda_{d_1} \lambda_{d_2}}{[d_1, d_2]},$$

subject to the constraint that $\lambda_1 = 1$ (recall $[d_1, d_2]$ denotes the least common multiple of d_1 and d_2. We were able to solve the optimization problem using a clever diagonalization procedure and prove $\pi(x) \ll x/\log x$ (without using Chebyshev's argument). By repeating the same argument (with some minor adjustments) it is possible to minimize certain other forms, whose coefficients also are expressed in terms of multplicative functions.

Theorem 1.1. *Let f be a multiplicative function (i.e. $f(mn) = f(m)f(n)$ whenever $\gcd(m, n) = 1$). Also, let $\mathcal{P}$ denote a set of primes and*

$$P(z) = \prod_{\substack{p \le z \\ p \in \mathcal{P}}} p.$$

Suppose that $\{\lambda_d\}_{d \ge 1} \subset \mathbb{R}$ such that $\lambda_1 = 1$ and $\lambda_d = 0$ if $d > z$. Then the minimum of the quadratic form

$$Q(\lambda) = \sum_{\substack{d_1, d_2 \le z \\ d_1, d_2 | P(z)}} \frac{\lambda_{d_1} \lambda_{d_2}}{f([d_1, d_2])}$$

subject to the constraint above is

$$S(z) = \sum_{\substack{d \le z \\ d | P(z)}} \frac{\mu^2(d)}{(\mu * f)(d)}.$$

Moreover, the minimizing vector satisfies $|\lambda_d| \le 1$.

We will not give a proof of this result since it follows from the ideas from the argument given in the last lecture. The fact that f is multiplicative is essential to the argument. This allows us to diagonalize $Q(\lambda)$ with an extremely simple argument, so the optimization problem is not difficult to

Date: May 7, 2015.

solve. Crucially, one takes advantage of the extra arithmetic structure of the coefficients of the quadratic form $Q(\lambda)$.

2. A sieving problem.

We will now show how to apply Theorem (1.1) to the following sieving problem. Given the following:

- $\mathcal{A} \subset \mathbb{Z}$ with $\#\mathcal{A} = X$;
- $\mathcal{P}$ a set of primes and
$$P(z) = \prod_{\substack{p \le z \\ p \in \mathcal{P}}} p;$$
- for each square-free d such that $p|d \Rightarrow p \in \mathcal{P}$ define
$$\mathcal{A}_d = \{n \in \mathcal{A} : d|n\}.$$
and assume
$$\#\mathcal{A}_d = \frac{X}{f(d)} + R_d \tag{1}$$
where f is a multiplicative function;
-
$$\mathcal{S}(\mathcal{A}, \mathcal{P}, z) = \#\{n \in \mathcal{A} : \gcd(n, P(z)) = 1\}$$

the problem then is to estimate $\mathcal{S}(\mathcal{A}, \mathcal{P}, z)$. We will see in Zeev's lecture how to relate questions on twin primes to bounding $\mathcal{S}(\mathcal{A}, \mathcal{P}, z)$.

Remark. *This sieving method begins with a set $\mathcal{A}$ and a set of primes $\mathcal{P}$. Then one sieves out the elements of $\mathcal{A}$ which are divisible by some $p \in \mathcal{P}$ with $p \le z$. After this sieving procedure what remains is $\{n \in \mathcal{A} : \gcd(n, P(z)) = 1\}$ and this is sometimes called the sifted set, and $\mathcal{P}$ is referred to as the sifting set.*

One can interpret $1/f(d)$ as the probability that $n \in \mathcal{A}_d$. The terms remainder term R_d will be relatively small in applications and constitute an error term.

Theorem 2.1. *In the notation as above, we have*
$$\mathcal{S}(\mathcal{A}, \mathcal{P}, z) \le \frac{X}{S(z)} + R(z),$$
where
$$S(z) = \sum_{\substack{d \le z \\ d | P(z)}} \frac{\mu^2(d)}{(\mu * f)(d)}$$
and
$$R(z) = \sum_{\substack{d_1, d_2 \le z \\ d_1, d_2 | P(z)}} |R_{[d_1, d_2]}|.$$

Proof. Let $\{\lambda_d\}_{d\geq 1}\subset\mathbb{R}$ be such that $\lambda_1=1$ and $\lambda_d=0$ for $d>z$. Then

$$\delta(m)\leq\left(\sum_{d|m}\lambda_d\right)^2.$$

Using the above inequality one has that

$$\begin{aligned}\mathcal{S}(\mathcal{A},\mathcal{P},z)=&\sum_{\substack{n\in\mathcal{A}\\ \gcd(n,P(z))=1}}1\\ =&\sum_{n\in\mathcal{A}}\delta(\gcd(n,P(z)))\\ \leq&\sum_{n\in\mathcal{A}}\left(\sum_{d|\gcd(n,P(z))}\lambda_d\right)^2.\end{aligned}\tag{2}$$

Squaring out, switching order of summation gives, and applying (1) gives

$$\begin{aligned}\sum_{n\in\mathcal{A}}\left(\sum_{d|\gcd(n,P(z))}\lambda_d\right)^2=&\sum_{\substack{d_1,d_2\leq z\\ d_1,d_2|P(z)}}\lambda_{d_1}\lambda_{d_2}\sum_{\substack{n\in\mathcal{A}\\ [d_1,d_2]|n}}1\\ =&\sum_{\substack{d_1,d_2\leq z\\ d_1,d_2|P(z)}}\lambda_{d_1}\lambda_{d_2}\left(\frac{X}{f([d_1,d_2])}+R_{[d_1,d_2]}\right)\\ =&XQ(\lambda)+\sum_{\substack{d_1,d_2\leq z\\ d_1,d_2|P(z)}}\lambda_{d_1}\lambda_{d_2}R_{[d_1,d_2]}\end{aligned}\tag{3}$$

where

$$Q(\lambda)=\sum_{\substack{d_1,d_2\leq z\\ d_1,d_2|P(z)}}\frac{\lambda_{d_1}\lambda_{d_2}}{f([d_1,d_2])}.$$

Therefore, by (2), (3) and Theorem 1.1 we have

$$\mathcal{S}(\mathcal{A},\mathcal{P},z)\leq\frac{X}{\displaystyle\sum_{\substack{d\leq z\\ d|P(z)}}\frac{\mu^2(d)}{(\mu*f)(d)}}+\sum_{\substack{d_1,d_2\leq z\\ d_1,d_2|P(z)}}|R_{[d_1,d_2]}|.$$

□

3. Application to primes in short intervals

Let

$$\pi(x;H)=\pi(x+H)-\pi(x)=\#\{x<p\leq x+H:p\text{ prime}\}.$$

A classical problem in analytic number theorem is to estimate $\pi(x;H)$, in particular for H small relative to x. Recall the prime number theorem states that

$$\pi(x) = \int_2^x \frac{dt}{\log t} + O\left(xe^{-\sqrt{\log x}}\right)$$

Thus, it follows for $x/(\log x)^A < H < x$ that

$$\pi(x;H) \sim \frac{H}{\log x}. \tag{4}$$

In fact, this estimate has been improved significantly and it has been proved by Hoheisel in 1930 [1] that (4) holds for $x^\theta < H < x$ for $\theta > 1 - 1/33,000$, today it is known that (4) holds for $\theta > 7/12$ (see Huxley [2]). The Riemann hypothesis states

$$\pi(x) = \int_2^x \frac{dt}{\log t} + O\left(x^{1/2+\varepsilon}\right),$$

so RH implies (4) holds for $x^{1/2+\varepsilon} < H < x$. However, we believe that more is true and it is conjectured that (4) holds for $x^\varepsilon < H < x$ for any $\varepsilon > 0$. This would essentially be the largest range possible, since a result of Maier [3] implies that (4) cannot hold for $H = (\log x)^A$ for any $A > 1$. In fact Maier showed that

$$G(A) = \limsup_{x\to\infty} \frac{\pi(x;H)}{H/\log x} > 1$$

His method also shows that for $1 < A < e^\gamma$, where γ is Euler's constant that

$$G(A) > \frac{e^\gamma}{A}.$$

Establishing (4) for small H is an extremely difficult problem. We will give an upper bound for $\pi(x;H)$.

Theorem 3.1. *For $H \geq 2$*

$$\pi(x;H) \leq 2\zeta(2) \cdot \frac{H}{\log H}(1 + o(1)).$$

Remark. *With a more careful argument one can replace the $\zeta(2)$ factor with 1. We essentially already proved this, but we will go over this example again to show how to apply the abstract sieve.*

Proof. Let $P(z) = \prod_{p\leq z} p$ and observe that if p is prime then $p|P(z)$ or $\gcd(n, P(z)) = 1$. Thus,

$$\pi(x;H) \leq \#\{x < n \leq x + H : \gcd(n, P(z)) = 1\} + z.$$

Now we will apply the Selberg sieve.

Let $\mathcal{A} = (x, x + H]$, $\mathcal{P} = \{p : p \text{ is prime}\}$. To use the Selberg sieve we also need to estimate $\mathcal{A}_d = \{n \in \mathcal{A} : d|n\}$ for squarefree d. It is easy to see that

$$\#\mathcal{A}_d = \frac{X}{d} + R_d \qquad \text{where} \qquad R_d = O(1).$$

Thus, Theorem 3.1 gives

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

Where

$$S(z) = \sum_{d \leq z} \frac{\mu^2(d)}{(\mu * \iota)(d)}.$$

Since $(\mu * \iota)(d) = \varphi(d) \leq d$ we have by partial summation

$$S(z) \geq \frac{1}{\zeta(2)} \log z(1 + o(1)).$$

So taking $z = H^{1/2+o(1)}$ gives

$$S(\mathcal{A}, \mathcal{P}, z) \leq 2\zeta(2) \cdot \frac{H}{\log H}(1 + o(1)).$$

□

References

1. G. Hoheisel, *Primzahlprobleme in der analysis.s*, Sitz. Preuss. Akad. Wiss. phys.-math. Klasse (1930), 580–588.
2. M. N. Huxley, *On the difference between consecutive primes*, Invent. Math. **15** (1972), 164–170. MR 0292774 (45 #1856)
3. Helmut Maier, *Primes in short intervals*, Michigan Math. J. **32** (1985), no. 2, 221–225. MR 783576 (86i:11049)