Chapter 9

Selberg's Sieve

Atle Selberg developed the sieve that is now referred to as "Selberg's sieve" in the 1940's, when he was trying to study the zeros of the Riemann zeta function. Like Brun's sieve, the Selberg sieve stems from the principle of inclusion-exclusion. However, Selberg introduces some clever innovations that allow for improving on the error term obtained via Brun's sieve (at least, in certain applications).

In this chapter, we will re-derive several of the results from the Brun's sieve chapters, but this time with better error terms. In particular, we will see how to obtain improved bounds for the count of twin prime pairs, and for the Brun-Titchmarsh inequality. We will also see how to use Selberg's sieve in order to obtain the upper bound in Chebyshev's inequality.

In the exercises at the end of this chapter, we will use ideas of Selberg (but not the Selberg sieve) to examine a serious limitation of sieve methods: the so-called parity problem. Even if we have very good estimates for the error term, sieves have a fundamental flaw: they cannot distinguish (on their own) between numbers with an odd number of prime factors and those with an even number of prime factors. In particular, sieves cannot really distinguish between (a) primes and (b) products of two primes of roughly equal size. As a result, sieves are only useful in handling certain types of problems. In order to get around these limitations, one has to cleverly combine sieves with other methods. We will see one such workaround in the final lecture of this course...

9.1 Chebyshev via Selberg

In this section, we will prove the upper bound from Chebyshev's inequality using Selberg's approach. Along the way, we will wind up proving a special case of Selberg's sieve, which will shed some light on the ideas that go into the proof of the more-general theorem.

We begin by recalling the basic setup from Chapters 8 and 8.6. Let $\mathscr{A} = \mathbb{N} \cap [1, x]$ so that

$$S(\mathscr{A}, z) = \#\{n \leqslant x \ : \ n \text{ is not divisible by any prime in } \mathscr{P} \leqslant z\}.$$

By inclusion-exclusion (cf. equation (8.4.1)), we have

$$S(\mathscr{A}, z) = \sum_{n \leqslant x} \sum_{d | (n, P(z))} \mu(d) = \sum_{d | P(z)} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor.$$

Then,

$$S(\mathscr{A}, z) = \sum_{d | P(z)} \mu(d) \frac{x}{d} + O(2^{\pi(z)}) = x \prod_{p \leqslant z} \left(1 - \frac{1}{p}\right) + O(2^{\pi(z)}).$$

Using a more refined analysis to bound

$$\sum_{d | P(z)} \mu(d) \left(\left\lfloor \frac{x}{d} \right\rfloor - \frac{x}{d} \right),$$

namely Exercise 7.1, it is possible to show that

$$S(\mathscr{A}, z) = x \prod_{p \leqslant z} \left(1 - \frac{1}{p}\right) + O\left(x \log z \exp\left(- \frac{\log x}{\log z} \right) \right).$$

Selberg's idea was to replace $\mu(d)$ with a quadratic form that is chosen in an optimal way. His method relies on the following key observation:

If $\lambda_1 = 1$ and (λ_n) is a sequence of arbitrary real numbers, then since

$$\sum_{d|m} \mu(d) = \begin{cases} 1 \text{ if } m = 1 \\ 0 \text{ otherwise,} \end{cases}$$

we have

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

Therefore,

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

where $[d_1, d_2] = \mathrm{lcm}[d_1, d_2]$.

Now, since $\#\{n \leqslant x : d \mid n\} = \left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} + O(1)$, then

$$S(\mathscr{A}, z) \leqslant x \sum_{d_1,d_2|P(z)} \frac{\lambda_{d_1}\lambda_{d_2}}{[d_1, d_2]} + O\left(\sum_{d_1,d_2|P(z)} |\lambda_{d_1}||\lambda_{d_2}| \right).$$

For convenience, assume that $\lambda_d = 0$ for all $d > z$. Then

$$S(\mathscr{A}, z) \leqslant x \sum_{d_1,d_2 \leqslant z} \frac{\lambda_{d_1}\lambda_{d_2}}{[d_1, d_2]} + O\left(\sum_{d_1,d_2 \leqslant z} |\lambda_{d_1}||\lambda_{d_2}| \right).$$

This error term is actually better than what we obtain via standard inclusion-exclusion. Since $[d_1, d_2](d_1, d_2) = d_1 d_2$ and since $\sum_{e|d} \varphi(e) = d$, then

$$\begin{aligned}
\sum_{d_1, d_2 \leqslant z} \frac{\lambda_{d_1} \lambda_{d_2}}{[d_1, d_2]} &= \sum_{d_1, d_2 \leqslant z} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} (d_1, d_2) \\
&= \sum_{d_1, d_2 \leqslant z} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} \sum_{e|(d_1, d_2)} \varphi(e) \\
&= \sum_{e \leqslant z} \varphi(e) \sum_{\substack{d_1, d_2 \leqslant z \\ e|(d_1, d_2)}} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} \\
&= \sum_{e \leqslant z} \varphi(e) \left(\sum_{\substack{d \leqslant z \\ e|d}} \frac{\lambda_d}{d} \right)^2 = \sum_{e \leqslant z} \varphi(e) u_e^2,
\end{aligned}$$

where $u_e = \sum_{\substack{d \leqslant z \\ e|d}} \frac{\lambda_d}{d}$.

Our objective now is to minimize the quadratic form $\sum_{e \leqslant z} \varphi(e) u_e^2$. By Möbius inversion,

(9.1.1)
$$\frac{\lambda e}{e} = \sum_{e|d} \mu(d/e) u_d.$$

Consequently, we have the constraints $u_e = 0$ for any $e > z$ (by definition of u_e) and $\sum_{e \leqslant z} \mu(e) u_e = \lambda_1 = 1$ (using (9.1.1)).

Let $V(z) = \sum_{d \leqslant z} \frac{\mu^2(d)}{\varphi(d)}$. Then,

$$\begin{aligned}
\sum_{e \leqslant z} \varphi(e) \left(u_e - \frac{\mu(e)}{\varphi(e) V(z)} \right)^2 + \frac{1}{V(z)} &= \sum_{e \leqslant z} \left(\varphi(e) u_e^2 - \frac{2 u_e \mu(e)}{V(z)} + \frac{\mu^2(e)}{\varphi(e) V^2(z)} \right) + \frac{1}{V(z)} \\
&= \sum_{e \leqslant z} \varphi(e) u_e^2.
\end{aligned}$$

This shows that the minimum value of $\sum_{e \leqslant z} \varphi(d) u_e^2$ is $\frac{1}{V(z)}$ and occurs at

$$u_e = \frac{\mu(e)}{\varphi(e) V(z)}.$$

Therefore, the optimal choice of λ_e is

$$\lambda_e = e \sum_{\substack{e \leqslant z \\ e|d}} \frac{\mu(d/e) \mu(d)}{\varphi(d) V(z)}.$$

Hence,

$$S(\mathscr{A}, z) \leqslant \frac{x}{V(z)} + O\left(\sum_{d_1, d_2 \leqslant z} |\lambda_1||\lambda_2|\right).$$

Now,

$$\begin{aligned} V(z)\lambda_e &= e \sum_{\substack{d \leqslant z \\ e|d}} \frac{\mu(d/e)\mu(d)}{\varphi(d)} = e \sum_{t \leqslant \frac{z}{e}} \frac{\mu(t)\mu(et)}{\varphi(et)} \\ &= e \sum_{\substack{t \leqslant \frac{z}{e} \\ (t,e)=1}} \frac{\mu^2(t)\mu(e)}{\varphi(e)\varphi(t)} = \mu(e) \prod_{p|e} \left(1 + \frac{1}{p-1}\right) \sum_{\substack{t \leqslant \frac{z}{e} \\ (t,e)=1}} \frac{\mu^2(t)}{\varphi(t)}. \end{aligned}$$

Observe that

$$\prod_{p|e} \left(1 + \frac{1}{p-1}\right) \sum_{\substack{t \leqslant \frac{z}{e} \\ (t,e)=1}} \frac{\mu^2(t)}{\varphi(t)} \leqslant \sum_{d \leqslant z} \frac{\mu^2(d)}{\varphi(d)}.$$

(To see this, note that for any squarefree m dividing e, we have $\prod_{p|m} \frac{1}{p-1} = \prod_{p|m} \frac{1}{\varphi(p)}$. Then, using the change of variables $d = tm$, we see that if $t \leqslant \frac{z}{e}$ and $m \mid e$ then $tm \leqslant z$.) By taking absolute values, we obtain

$$|V(z)||\lambda_e| \leqslant |V(z)|,$$

and so $|\lambda_e| \leqslant 1$ for all e. Therefore,

$$\begin{aligned} S(\mathscr{A}, z) &\leqslant \frac{x}{V(z)} + O\left(\sum_{d_1, d_2 \leqslant z} |\lambda_{d_1}||\lambda_{d_2}|\right) \\ &= \frac{x}{V(z)} + O(z^2). \end{aligned}$$

We have just proven the following theorem:

Theorem 9.1.1. *Let* $V(z) = \sum_{d \leqslant z} \frac{\mu^2(d)}{\varphi(d)}$*, and let* $\mathscr{A} = \mathbb{N} \cap [1, z]$*. Then,*

$$S(\mathscr{A}, z) \leqslant \frac{x}{V(z)} + O(z^2)$$

as $x, z \to \infty$.

We can use Theorem 9.1.1 to prove the upper bound from Chebyshev's inequality. Recall that

$$\pi(x) \leqslant S(\mathscr{A}, \mathscr{P}, z) + z.$$

Note that

$$V(z) \geqslant \sum_{d \leqslant z} \frac{\mu^2(d)}{d} = \sum_{\substack{d \leqslant z \\ d \text{ squarefree}}} \frac{1}{d} \gg \log z,$$

where the last bound follows from partial summation and the fact that

$$\sum_{\substack{d \leqslant z \\ d \text{ squarefree}}} 1 \sim \frac{z}{\zeta(2)}$$

Hence,

$$\pi(x) \ll \frac{x}{\log z} + z^2.$$

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

$$\pi(x) \ll \frac{2x}{\log x - \log\log x} + \frac{x}{\log x} \ll \frac{x}{\log x}.$$

Therefore, we have

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

as desired.

Let us pause briefly to compare what we obtain using Selberg's Sieve versus with Brun's Sieve. Brun's sieve yields a main term of

$$x \prod_{p \leqslant z} \left(1 - \frac{1}{p}\right),$$

while Selberg's sieve yields a main term of

$$\frac{x}{\displaystyle\sum_{d \leqslant z} \frac{\mu^2(d)}{\varphi(d)}}.$$

Notice that these are both $O(x/\log z)$, so they have the same order of magnitude. The important distinction is that Selberg's method gives a smaller error term than what one obtains using Brun's sieve.

9.2 Selberg's Sieve

Now we generalize the ideas from the previous section. Recall that for a set of integers $\mathscr{A}$, we assume that

$$\#\mathscr{A}_d = \#\mathscr{A}\, g(d) + r(d),$$

where $g : \mathbb{N} \to [0,1]$ was a multiplicative function and $r(d)$ plays the role of an error term. For Selberg's sieve, it will be natural to introduce a new function, f, where $g(d) = 1/f(d)$. Then we have

$$\#\mathscr{A}_d = \frac{\#\mathscr{A}}{f(d)} + r(d),$$

and f is a multiplicative function with $f(p) > 1$ for all $p \in \mathscr{P}$.

Theorem 9.2.1 (Selberg's sieve). *Let $X > 0$ and let f be a multiplicative function satisfying $f(p) > 1$ for any prime $p \in \mathscr{P}$ such that for any squarefree integer d composed of primes in $\mathscr{P}$ we have*

$$\#\mathscr{A}_d = \frac{\#\mathscr{A}}{f(d)} + r(d)$$

for some real number $r(d)$. Let f_1 be the function satisfying

$$f(n) = \sum_{d \mid n} f_1(d)$$

that is uniquely determined by the Möbius inversion formula. Let

$$V(z) := \sum_{\substack{d \leqslant z \\ d \mid P(z)}} \frac{\mu^2(d)}{f_1(d)}.$$

Then

$$S(\mathscr{A}, \mathscr{P}, z) \leqslant \frac{\#\mathscr{A}}{V(z)} + O\Big(\sum_{\substack{d_1, d_2 \leqslant z \\ d_1, d_2 \mid P(z)}} |r([d_1, d_1])| \Big).$$

Proof. The proof follows in much the same manner as the proof of Theorem 9.1.1. For details, see pp. 120-123 in Cojocaru and Murty's book. □

In order to use Selberg's sieve, we will need to bound $\frac{1}{V(z)}$. To do this, we make use of the following lemma:

Lemma 9.2.2. *Let $\tilde{f}$ be a completely multiplicative function with $\tilde{f}(p) := f(p)$ for all primes p and let*

$$\overline{P}(z) := \prod_{\substack{p \notin \mathscr{P} \\ p \leqslant z}} p.$$

Then:

(i) $V(z) \geqslant \sum_{\substack{e \leqslant z \\ p|e \Rightarrow p|P(z)}} \frac{1}{\tilde{f}(e)}$

(ii) $f(\overline{P}(z))V(z) \geqslant f_1(\overline{P}(z)) \sum_{e \leqslant z} \frac{1}{\tilde{f}(e)}.$

Proof. (i) For $d \mid P(z)$, using the fact that $f(p) = f_1(p) + 1$, we have

$$\frac{f(d)}{f_1(d)} = \prod_{p|d} \frac{f(p)}{f_1(p)} = \prod_{p|d} \left(1 - \frac{1}{f(p)}\right)^{-1}$$
$$= \prod_{p|d} \sum_{n \geqslant 0} \frac{1}{f(p)^n} = \sum_k{}' \frac{1}{\tilde{f}(k)},$$

where the sum is composed of prime divisors of d. Then

$$V(z) = \sum_{\substack{d \leqslant z \\ d|P(z)}} \frac{\mu^2(d)}{f(d)} \sum_k{}' \frac{1}{\tilde{f}(k)} \geqslant \sum_{\substack{e \leqslant z \\ p|e \Rightarrow p|P(z)}} \frac{1}{\tilde{f}(e)}.$$

(ii) As in (i), we have

$$\frac{f(\overline{P}(z))}{f_1(\overline{P}(z))} V(z) = \prod_{\substack{p \notin \mathscr{P} \\ p < z}} \left(\sum_{n \geqslant 0} \frac{1}{f(p)^n}\right) \sum_{\substack{d \leqslant z \\ d|P(z)}} \frac{\mu^2(d)}{f(d)} \sum_k{}' \frac{1}{\tilde{f}(k)}$$
$$\geqslant \sum_{e \leqslant z} \frac{1}{\tilde{f}(e)}.$$

□

9.3 Twin primes, revisited

Recall the setting of Example 2 from Brun's sieve: let

$$\mathscr{A} := \{n(n+2) \ : \ 1 \leqslant n \leqslant x\}$$

Then,

$$\#\mathscr{A}_d = \frac{N(d)}{d}x + r(d),$$

where $r(d) \leqslant N(d) \leqslant 2^{\omega(d)}$ for $d|P(z)$. From this, $f(d) = \frac{d}{N(d)}$. Hence, by Selberg's sieve, we have

$$S(\mathscr{A}, z) \leqslant \frac{x}{V(z)} + O\Big(\sum_{\substack{d_1, d_2 \leqslant z \\ d_1, d_2 | P(z)}} 2^{\omega([d_1, d_2])} \Big).$$

We will use Lemma 9.2.2 to bound $V(z)$, but first let's study the error term. Since $2^{\omega(d)} = \tau(d)$ for d squarefree, then

$$\begin{aligned} \sum_{\substack{d_1, d_2 \leqslant z \\ d_1, d_2 | P(z)}} 2^{\omega([d_1, d_2])} &\leqslant \Bigg(\sum_{\substack{d \leqslant z \\ d \text{ squarefree}}} 2^{\omega(d)} \Bigg)^2 \\ &\leqslant \Bigg(\sum_{d \leqslant z} \tau(d) \Bigg)^2 \\ &\leqslant (z \log z)^2, \end{aligned}$$

where the last bound follows from Exercise 9.2. Now we bound $V(z)$: note that the condition $p|e \Rightarrow p|P(z)$ is trivially satisfied for $e \leqslant z$. Now, recall that $\tilde{f}$ is the completely multiplicative function defined by $\tilde{f}(p) = f(p)$, and since

$$f(p) = \begin{cases} p \text{ if } p = 2, \\ \frac{p}{2} \text{ if } p > 2, \end{cases}$$

then

$$\frac{1}{\tilde{f}(d)} \leqslant \frac{2^{\Omega(d)}}{d}.$$

Therefore, by (i) of Lemma 9.2.2,

$$\begin{aligned} V(z) &\geqslant \sum_{d \leqslant z} \frac{1}{\tilde{f}(d)} \\ &\geqslant \sum_{d \leqslant z} \frac{2^{\Omega(d)}}{d} \\ &\geqslant \sum_{d \leqslant z} \frac{\tau(d)}{d} \\ &\gg (\log z)^2, \end{aligned}$$

where the last bound follows from Exercise 9.2. This shows that

$$S(\mathscr{A}, z) \ll \frac{x}{\log^2 z} + O(z^2 \log^2 z).$$

Now, let $\pi_2(x)$ denote the number of twin primes less than or equal to x and note that if $n, n+2$ are both prime then $n \leqslant z$ or both $n, n+2$ have only prime factors exceeding z. Hence,

$$\pi_2(x) \leqslant S(\mathscr{A}, z) + z.$$

Letting $z = x^{\frac{1}{4}}$, we obtain the upper bound

$$\pi_2(x) \ll \frac{x}{\log^2 x}.$$

Compare this bound with the one obtained from Brun's sieve: In Corollary 8.6.2, we showed that

$$\pi_2(x) \ll \frac{x \log_2(x)}{\log^2 x}.$$

Therefore, via Selberg's sieve we were able to improve the bound for twin primes.

9.4 Exercises

Exercise 9.1. *Use the method of Lagrange multipliers to find the minimal value of the quadratic form*

$$\sum_{\substack{e \leqslant z \\ e | P(z)}} f_1(e) u_e^2.$$

This is the diagonal quadratic form that appears in the proof of the general version of Selberg's sieve.

Exercise 9.2. *Show that*

$$\sum_{n \leqslant x} \tau(n) \sim x \log x$$

as $x \to \infty$.

Exercise 9.3 (Brun-Titchmarsh)**.** *Let* a *and* k *be positive coprime integers. Let*

$$\mathscr{A} := \{n \leqslant x \ : \ n \equiv a \pmod{k}\},$$

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

and

$$P := \prod_{\substack{p \in \mathscr{P} \\ p < z}} p.$$

(a) *Let* d *be a squarefree number composed of primes in* $\mathscr{P}$. *Show:*

$$\#\mathscr{A}_d = \frac{x}{k} \cdot \frac{1}{d} + O(1).$$

(b) *Show that, for all* $0 < z < x$, *we have*

$$\pi(x; k, a) \leqslant z + S(\mathscr{A}, \mathscr{P}, z).$$

(c) *Use Selberg's sieve to show that, for all* $k < z$,

$$S(\mathscr{A}, \mathscr{P}, z) \leqslant \frac{x}{\varphi(k) \log z} + O(z^2).$$

Hint: Begin by proving that,

$$\frac{\varphi(\overline{P})}{\overline{P}} = \frac{\varphi(k)}{k},$$

where

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

In order to bound $V(z)$, use part (ii) of Lemma 9.2.2.

(d) Prove the Brun Titchmarsh inequality:

$$\pi(x; k, a) \ll \frac{x}{\varphi(k) \log(x/k)}$$

for all $k < x$. Hint: Take $z = \left(\frac{x}{k \log(x/k)}\right)^{\frac{1}{2}}$.

Exercise 9.4. *The prime number theorem with error term can be stated in the following way: There exists a positive constant* $c > 0$ *such that for all* $x \geqslant 2$*,*

$$\psi(x) = x + O(x \exp(-c\sqrt{\log x})).$$

Show that this statement implies that

$$\sum_{n \leqslant x} \lambda(n) \ll x \exp(-c\sqrt{\log x}),$$

where $\lambda(n) = (-1)^{\Omega(n)}$ *is the Liouville function.*

Hint: The starting point is the identity

$$-\mu(n) \log n = \sum_{d|n} \mu(d) \Lambda\left(\frac{n}{d}\right)$$

which can be proved by observing that

$$\left(\frac{1}{\zeta(s)}\right)' = \frac{1}{\zeta(s)} \frac{-\zeta'(s)}{\zeta(s)}.$$

To relate $\mu(n)$ *with* $\lambda(n)$*, note that*

$$\sum_{n|d} \lambda(n) = \begin{cases} 1, & \text{if } n \text{ is a perfect square,} \\ 0, & \text{otherwise.} \end{cases}$$

Exercise 9.5. *Let $\Phi_{odd}(x,z)$ (resp. $\Phi_{even}(x,z)$) denote the number of $n \leqslant x$ with an odd (resp. even) number of prime factors, counted with multiplicity, free of prime factors $\leqslant z$. Suppose that ρ_d is a bounded sequence of real numbers satisfying*

$$\sum_{d|n} \mu(d) \leqslant \sum_{d|n} \rho_d,$$

with $\rho_d = 0$ for $d > z$.

(i) *Use Exercise 9.4 to show that for any $\theta < 1$ and any $z < x^\theta$,*

$$\Phi_{odd}(x,z) \leqslant \frac{x}{2} \sum_d \frac{\rho_d}{d} + O(x(\log z)\exp(-c_1\sqrt{\log x})).$$

Establish a similar result for $\Phi_{even}(x,z)$.

Hint: The starting point is inclusion-exclusion.

(ii) *Deduce that*

$$\sum_{d\leqslant\sqrt{x}} \frac{\rho_d}{d} \geqslant \frac{2+o(1)}{\log x}.$$

Remark: *This exercise demonstrates the* ***parity problem*** *of sieve methods: we obtain the same bound for $\Phi_{odd}(x,\sqrt{x})$ and $\Phi_{even}(x,\sqrt{x})$, but $\Phi_{even}(x,\sqrt{x}) = 0$!*