Chapter 10

The Fundamental Lemma of Sieve Theory

10.1 Introduction

In this chapter, we take a brief interlude from introducing new sieve methods in order to unify the ideas from the previous two chapters and present them as part of a more-general framework called the Fundamental Lemma of Sieve Methods. In the context of sieves, a fundamental lemma is any type of result that formulates a general type of sieve such as Brun's sieve or Selberg's sieve. The reason this is important is because many papers end up repeating the same arguments when setting up the sieving process. In Halberstam and Richert's [2] own words:

"A curious feature of sieve literature is that while there is frequent use of Brun's method there are only a few attempts to formulate a general Brun theorem; as a result there are surprisingly many papers which repeat in considerable detail the steps of Brun's argument"

In order to avoid repeating the steps of Brun's argument, we will use the general framework presented in this chapter for all of the remaining chapters. At the end of this chapter, we will also expand upon one of the major limitations of sieve methods (i.e., the so-called parity problem) that we introduced in the exercises in Chapter 9.

10.2 The Fundamental Lemma of Sieve Theory

As usual, we assume that $\mathscr{A}$ is a set of positive integers, and that

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

for some multiplicative function g. Recall that by inclusion exclusion, we have

$$S(\mathscr{A}, z) := \sum_{\substack{n \in \mathscr{A} \\ p \mid n \Rightarrow p > z}} 1 = \sum_{\substack{d \\ p \mid d \Rightarrow p \leqslant z}} \mu(d) \#\mathscr{A}_d.$$

So far, the basic idea of (combinatorial) sieve methods is that, if we have a moderate understanding of $g(d)$ and $r(d)$, then we can make the inclusion exclusion argument more accurate. Results of this type are said to satisfy the Fundamental Lemma of Sieve Methods. We encountered a first example of this phenomenon when we studied Brun's sieve:

Theorem 8.5.2 (Brun's sieve)**.** Assume that

(i) $|r(d)| \leqslant d g(d)$ for any squarefree d composed of primes of $\mathscr{P}$.

(ii) There is some $A_1 \geqslant 1$ such that

$$0 \leqslant g(p) \leqslant 1 - \frac{1}{A_1}.$$

(iii) There exists some $\kappa > 0$ and $A_2 \geqslant 1$ such that

$$\sum_{p \leqslant w} g(p) \log p \leqslant \kappa \log w + A_2.$$

Then, there exist constants c_1, c_2 depending only on A_1, A_2, κ such that

$$S(\mathscr{A}, z) \leqslant \#\mathscr{A} \prod_{p \leqslant z} (1 - g(p))(1 + \mathrm{e}^{c_1/\log z}) + O(z^{c_2}).$$

It is possible to obtain results as strong as Brun's sieve with weaker assumptions. It is not always necessary to obtain a pointwise bound such as $|r(d)| \leqslant d g(d)$; sometimes it suffices to understand $|r(d)|$ on average. The most common bound that we expect to hold is called a *type I estimate:*

Definition 10.2.1 (Type I estimate)**.** Given a set of positive integers $\mathscr{A}$, let $\mathscr{A}(x)$ denote the number of elements of $\mathscr{A}$ less than or equal to x, i.e., $\mathscr{A}(x) := \mathscr{A} \cap [1, x]$. A **type I estimate** is a bound of the form

$$\sum_{d<x^{\gamma}} |\#\mathscr{A}_d(x) - \#\mathscr{A}(x)g(d)| \ll_B \frac{\#\mathscr{A}(x)}{(\log x)^B} \tag{10.2.1}$$

for any $B > 0$, and for some $\gamma \in (0, 1)$. γ is called the **level of distribution**.

Example. Let $\mathscr{A} := \mathbb{Z} \cap [1, x]$ and let $\varepsilon > 0$. Then, $\#\mathscr{A}_d = \lfloor x/d \rfloor$, so that

$$\sum_{d \leqslant x^{1-\varepsilon}} \left| \#\mathscr{A}_d - \frac{\#\mathscr{A}}{d} \right| = \sum_{d \leqslant x^{1-\varepsilon}} \left| \left\{ \frac{x}{d} \right\} \right| \leqslant \sum_{d \leqslant x^{1-\varepsilon}} 1 \leqslant x^{1-\varepsilon} \ll_{\varepsilon, B} \frac{x}{(\log x)^B}.$$

This shows that the integers have level of distribution $1 - \varepsilon$ for any $\varepsilon > 0$. In the case where our set $\mathscr{A}$ satisfies a bound of this form, we can be more precise with our arguments and obtain the following version of the Fundamental Lemma (see Maynard [5] and Friedlander–Iwaniec [1] for details)

Lemma 10.2.2 (Fundamental Lemma of Sieve Methods)**.** *Let $\mathscr{A}$ be a set of positive integers such that*

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

for some multiplicative function g. Suppose that

(i) *$\mathscr{A}$ satisfies a type I estimate with level of distribution γ.*

(ii) *$g(p) \approx \kappa/p$ for some $\kappa > 0$ in the sense that, for any $w \geqslant 2$,*

$$\sum_{p \leqslant w} g(p) \log p = \kappa \log w + O(1).$$

Then, for every $B, \eta > 0$,

$$S(\mathscr{A}, x^{\eta}) = \#\mathscr{A} \prod_{p \leqslant x^{\eta}} (1 - g(p))(1 + O_{\kappa}(e^{-\gamma/\eta})) + O_B\Big(\frac{x}{(\log x)^B}\Big).$$

The constant κ is called the **sieve dimension**. In the particular case where $\kappa = 1$, we have what is called the **linear sieve**. Linear sieves are much better-understood than sieves with other dimensions; we already know a great deal about their uses and limitations. That said, there are plenty of examples of sieves with other dimensions in the literature. Semi-linear sieves (with dimension $\kappa = 1/2$) are used, for example, for counting problems involving sums of two squares.

10.3 The Parity Problem Revisited

One application of the linear sieve is that it can be used to establish Chen's Theorem:

Theorem 10.3.1 (Chen's Theorem)**.** *There are infinitely many primes p such that $p + 2$ is the product of at most two primes.*

For a proof, see 254A, Supplement 5 on Terry Tao's blog. Notice that Chen's theorem does not distinguish between primes and numbers with two prime factors. This is due to the *parity problem* that was highlighted in Exercise 9.5. Selberg famously considered the set

$$\mathscr{A}^- := \{n \in \mathbb{Z} \cap [1, x] : n \text{ has an even number of distinct prime factors}\}.$$

It follows from Exercise 10.2 that the set $\mathscr{A}^-$ is extremely well-distributed in arithmetic progressions. In particular, for all $\varepsilon > 0$ and for all $B > 0$, we have

$$\sum_{d<x^{1-\varepsilon}} \left| \#\mathscr{A}_d^- - \frac{\#\mathscr{A}^-}{d} \right| \leqslant \frac{\#\mathscr{A}^-}{(\log x)^B}.$$

But clearly $\#\{p \in \mathscr{A}^-\} = 0$ since all primes have an odd number of prime factors. In other words, we have very good type I information for the set $\mathscr{A}^-$, but the linear sieve fails pretty spectacularly (as would any other sieves that we might consider using here instead). Even if we could obtain the best-possible estimates for $\#\mathscr{A}_d^-$, we do not have any hope of obtaining a nontrivial lower bound for the count of primes in $\mathscr{A}^-$ using only a type I estimate. The takeaway here is that one cannot prove asymptotic results about primes in a set using *only* type I information, as even some sets without primes come with lots of type I information.

Unfortunately, sieve methods do not distinguish between sets that do contain primes and sets that do not contain primes. In order to get around this problem, one would need to incorporate more arithmetic information. One possible workaround is to use a so-called type II estimate. One example of such an estimate is given by Vaughan's identity, which we will see as an ingredient in the proof of the Bombieri-Vinogradov Theorem. Unfortunately, type II estimates can become quite complicated and we will not be able to cover them in much depth in this course.

10.4 Applications of the Fundamental Lemma

A key point of the Fundamental Lemma is that we can still obtain an asymptotic with a good error term for $S(\mathscr{A}, z)$ even when z is as large as x^{η}, as long as we have a type I estimate for $\mathscr{A}$. This is what authors in the literature on sieve methods mean when they say that we need to understand $\mathscr{A}$ in arithmetic progressions. We have the following immediate consequence of the Fundamental Lemma:

Corollary 10.4.1. *Let $\mathscr{A}$ and g be as in the Fundamental Lemma. Then,*

$$\#\{p \leqslant x \ : \ p \textit{ is prime}\} \ll \frac{\#\mathscr{A}}{(\log x)^{\kappa}}.$$

Proof. See Exercise 10.1. □

There are a number of interesting conclusions that can be drawn from this corollary. For example, it shows that the number of twin primes up to x is $O(x/(\log x)^2)$, and that the number of primes of the form n^2+1 up to x is $O(x^{\frac{1}{2}}/\log x)$. Both of these estimates are conjectured to be sharp up to the multiplicative constants in the presumed asymptotics.

10.5 Exercises

Exercise 10.1. *Prove Corollary 10.4.1.*

Exercise 10.2. *Consider the sets*

$$\mathscr{A}^{-} := \{n \in \mathbb{Z} \cap [1, x] \ : \ n \textit{ has an even number of distinct prime factors}\},$$

and

$$\mathscr{A}^{+} := \{n \in \mathbb{Z} \cap [1, x] \ : \ n \textit{ has an odd number of distinct prime factors}\}.$$

Show that $\mathscr{A}^{+}$ and $\mathscr{A}^{-}$ satisfy a type I estimate with level of distribution $1-\varepsilon$ for any $\varepsilon \in (0,1)$.