Chapter 11

The Large sieve

11.1 Introduction

The goal of this chapter is to introduce the large sieve and some of its applications. The term "large sieve" was initially coined by Linnik. In some sense, the large sieve is neither large nor is it a sieve. As we will see in this section, the large sieve - in its most common formulations nowadays - is really just an inequality. Although some would not consider the large sieve to be a sieve at all, Linnik initially had an application in mind that was very sieve-like, so he decided to call it a sieve. He considered the large sieve to be "large" in the sense that it excludes a lot more congruence classes (mod p) than other sieves. Whereas many combinatorial sieves exclude (say) two residue classes, the large sieve can exclude (say) half of the residue classes (mod p). Later, large sieve results became powerful enough that they turned out to also be useful when a small number of residue classes (mod p) are excluded. Hence, the name "large sieve" can feel like a bit of a misnomer when one encounters it in certain contexts.

At the end of this chapter, we will present one of the most-important applications of the large sieve: the Bombieri-Vinogradov Theorem. The Bombieri-Vinogradov Theorem says that primes are more-or-less uniformly distributed (mod q) for integers q up to about $x^{1/2}$. This has many useful applications, as we will see later in this course.

11.2 The Large Sieve Inequality

Let's start with some motivation: Let $\mathscr{A} \subseteq \mathbb{Z}^+$, and suppose that we are in a situation where

$$\#\mathscr{A}_d \approx \frac{\#\mathscr{A}}{d} \qquad \text{(the linear sieve)}.$$

Recall that, in order to use the Fundamental Lemma of Sieve Methods, we need to understand $\mathscr{A}$ in arithmetic progressions by finding an estimate for

$$\sum_{d \leqslant Q} \left| \#\mathscr{A}_d - \frac{\#\mathscr{A}}{d} \right|, \tag{11.2.1}$$

where $Q = x^\gamma$ for some $\gamma \in (0,1)$. Let's try the following approach: by the orthogonality relations (Theorem 3.1.7), if $\mathrm{e}(x) := \mathrm{e}^{2\pi i x}$, then

$$\frac{1}{d} \sum_{b=0}^{d-1} \mathrm{e}\left(\frac{bn}{d}\right) = \begin{cases} 1 & \text{if } d|n, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$\begin{aligned} \#\mathscr{A}_d &= \sum_{\substack{n \in \mathscr{A} \\ d|n}} 1 \\ &= \sum_{n \in \mathscr{A}} \frac{1}{d} \sum_{b=0}^{d-1} \mathrm{e}\left(\frac{bn}{d}\right) \\ &= \frac{1}{d} \sum_{b=0}^{d-1} \sum_{n \in \mathscr{A}} \mathrm{e}\left(\frac{bn}{d}\right) \\ &= \frac{\#\mathscr{A}}{d} + \frac{1}{d} \sum_{b=1}^{d-1} \sum_{n \in \mathscr{A}} \mathrm{e}\left(\frac{bn}{d}\right), \end{aligned}$$

where the last equality follows from separating out the term with $b = 0$ from the sum. This together with the triangle inequality, shows that

$$\sum_{d \leqslant Q} \left| \#\mathscr{A}_d - \frac{\#\mathscr{A}}{d} \right| \leqslant \sum_{d \leqslant Q} \frac{1}{d} \sum_{b=1}^{d-1} \left| \sum_{n \in \mathscr{A}} \mathrm{e}\left(\frac{bn}{d}\right) \right|, \tag{11.2.2}$$

so it makes sense to study these exponential sums associated to $\mathscr{A}$. One example of an exponential sum associated to a set is the *Fourier transform* of $\mathscr{A}$:

Definition 11.2.1. Given a set $\mathscr{A} \subseteq \mathbb{Z}^+ \cap [1, x]$, we define the *Fourier transform* of $\mathscr{A}$ as

$$F_{\mathscr{A}}(t) := \sum_{n \in \mathscr{A}} \mathrm{e}(nt) = \sum_{n \leqslant x} \mathbb{1}_{\mathscr{A}}(n)\, \mathrm{e}(nt),$$

where $\mathbb{1}_{\mathscr{A}}$ is the characteristic function of $\mathscr{A}$, i.e.,

$$\mathbb{1}_{\mathscr{A}}(n) = \begin{cases} 1 & \text{if } n \in \mathscr{A}, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, understanding $\mathscr{A}$ in arithmetic progressions amounts to understanding a double sum for the Fourier transform of $\mathscr{A}$. We can also study this double sum more generally by replacing $\mathbb{1}_{\mathscr{A}}(n)$ with a sequence of complex numbers a_n, in which case we are interested in bounding a double sum with

$$\sum_{n \leqslant x} a_n\, \mathrm{e}(nt)$$

on the inside. This leads us to the main objective of this chapter, the *large sieve inequality.*

The large sieve inequality: Let (a_n) be a sequence of complex numbers and let $x, Q \in \mathbb{Z}^+$. Then,

$$\sum_{d \leqslant Q} \sum_{\substack{1 \leqslant b \leqslant d \\ (b,d)=1}} \left| \sum_{n \leqslant x} a_n\, \mathrm{e}\left(\frac{nb}{d}\right) \right|^2 \leqslant (Q^2 + 4\pi x) \sum_{n \leqslant x} |a_n|^2.$$

In Exercises 11.3 and 11.4 we will see how to use the methods from this chapter in order to obtain a bound for (11.2.2). For the remainder of this section, we focus on proving the large sieve inequality. Our proof follows the approach of Cojocaru and Murty. We begin with the following lemma:

Lemma 11.2.2. *Let $f \in C^1([0,1], \mathbb{C})$ be a periodic function with period 1, and let $Q \in \mathbb{Z}^+$. Then,*

$$\sum_{d \leqslant Q} \sum_{\substack{1 \leqslant b \leqslant d \\ (b,d)=1}} \left| f\left(\frac{b}{d}\right) \right| \leqslant Q^2 \int_0^1 |f(t)|\, \mathrm{d}t + \int_0^1 |f'(t)|\, \mathrm{d}t.$$

Proof. By the Fundamental Theorem of Calculus,

$$-f\Big(\frac{b}{d}\Big) = -f(t) + \int_{\frac{b}{d}}^{t} f'(x)\,\mathrm{d}x,$$

so that, by the triangle inequality, we have

$$\Big|f\Big(\frac{b}{d}\Big)\Big| \leqslant |f(t)| + \int_{\frac{b}{d}}^{t} |f'(x)|\,\mathrm{d}x. \tag{11.2.3}$$

Now, let $\delta > 0$. Then, integrating (11.2.3) with respect to t over the interval

$$I = I(b,d,\delta) := \Big(\frac{b}{d} - \frac{\delta}{2}, \frac{b}{d} + \frac{\delta}{2}\Big),$$

shows that

$$\delta\Big|f\Big(\frac{b}{d}\Big)\Big| \leqslant \int_I |f(t)|\,\mathrm{d}t + \int_I \int_{\frac{b}{d}}^{t} |f'(x)|\,\mathrm{d}x\,\mathrm{d}t,$$

and since $x \in \Big(\frac{b}{d}, t\Big)$, then $x \in I$, so that

$$\begin{aligned}\delta\Big|f\Big(\frac{b}{d}\Big)\Big| &\leqslant \int_I |f(t)|\,\mathrm{d}t + \int_I \int_I |f'(x)|\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_I |f(t)|\,\mathrm{d}t + \delta \int_I |f'(x)|\,\mathrm{d}x.\end{aligned}$$

Dividing by δ allows us to obtain

$$\Big|f\Big(\frac{b}{d}\Big)\Big| \leqslant \frac{1}{\delta}\int_I |f(t)|\,\mathrm{d}t + \int_I |f'(t)|\,\mathrm{d}t. \tag{11.2.4}$$

We now choose $\delta := \dfrac{1}{Q^2}$. By Exercise 11.1, the intervals I don't overlap for $1 \leqslant b \leqslant d$, with $(b,d) = 1$ and $d \leqslant Q$. Moreover, these I are also contained in the interval $[0,1]$. Therefore, summing (11.2.4) over all such intervals shows that

$$\sum_{d \leqslant Q} \sum_{\substack{1 \leqslant b \leqslant d \\ (b,d)=1}} \Big|f\Big(\frac{b}{d}\Big)\Big| \leqslant Q^2 \int_0^1 |f(t)|\,\mathrm{d}t + \int_0^1 |f'(t)|\,\mathrm{d}t. \qquad \square$$

We now apply this lemma to

$$f(t) := \Big(\sum_{n \leqslant x} a_n\,\mathrm{e}(nt)\Big)^2,$$

where (a_n) is a sequence of complex numbers and $x \in \mathbb{Z}^+$. For simplicity, let

$$S(t) := \sum_{n \leqslant x} a_n \, \mathrm{e}(nt),$$

so that

$$f(t) = S(t)^2 \quad \text{and} \quad f'(t) = 2S(t)S'(t).$$

Then, Lemma 11.2.2 shows that

$$\sum_{d \leqslant Q} \sum_{\substack{1 \leqslant b \leqslant d \\ (b,d)=1}} |S(t)|^2 \leqslant Q^2 \int_0^1 |S(t)|^2 \, \mathrm{d}t + 2 \int_0^1 |S(t)S'(t)| \, \mathrm{d}t. \tag{11.2.5}$$

By Parseval's identity (Exercise 11.2),

$$\int_0^1 |S(t)|^2 \, \mathrm{d}t = \sum_{n \leqslant x} |a_n|^2, \tag{11.2.6}$$

and by the Cauchy–Schwartz inequality,

$$\int_0^1 |S(t)S'(t)| \, \mathrm{d}t \leqslant \left(\int_0^1 |S(t)|^2 \, \mathrm{d}t \right)^{\frac{1}{2}} \left(\int_0^1 |S'(t)|^2 \, \mathrm{d}t \right)^{\frac{1}{2}}.$$

Now, since

$$S'(t) = \sum_{n \leqslant x} 2\pi i n a_n \, \mathrm{e}(nt),$$

once again Parseval's identity shows that

$$\int_0^1 |S(t)S'(t)| \, \mathrm{d}t \leqslant \left(\sum_{n \leqslant x} |a_n|^2 \right)^{\frac{1}{2}} \left(\sum_{n \leqslant x} 4\pi^2 n^2 |a_n|^2 \right)^{\frac{1}{2}} \leqslant 2\pi x \sum_{n \leqslant x} |a_n|^2. \tag{11.2.7}$$

Plugging (11.2.6) and (11.2.7) into (11.2.5) proves the large sieve inequality:

$$\sum_{d \leqslant Q} \sum_{\substack{1 \leqslant b \leqslant d \\ (b,d)=1}} \left| \sum_{n \leqslant x} a_n \, \mathrm{e}\left(\frac{nb}{d}\right) \right|^2 \leqslant (Q^2 + 4\pi x) \sum_{n \leqslant x} |a_n|^2.$$

Montgomery and Vaughan [12], and Selberg [17] independently showed that $Q^2 + 4\pi x$ can be replaced by $Q^2 + x$, so from now on we will use this improvement. We summarize what we have done in the following theorem:

Theorem 11.2.3 (The large sieve inequality)**.** *Let (a_n) be a sequence of complex numbers and let $x, Q \in \mathbb{Z}^+$. Then,*

$$\sum_{d \leqslant Q} \sum_{\substack{1 \leqslant b \leqslant d \\ (b,d)=1}} \left| \sum_{n \leqslant x} a_n \, \mathrm{e}\left(\frac{nb}{d}\right) \right|^2 \leqslant (Q^2 + x) \sum_{n \leqslant x} |a_n|^2.$$

11.3 The Arithmetic Large Sieve

In Exercise 11.5 you will be asked to prove the following version of the large sieve, which has a more sieve-like flavor:

Theorem 11.3.1. *Let $\mathscr{A}$ be a set of integers with elements $n \leqslant x$. Let $\mathscr{P}$ be a set of primes, and let*

$$P(z) := \prod_{p \in \mathscr{P} \cap [1,z]} p.$$

Now, suppose that for each prime $p \in \mathscr{P}$, we are given a set $\{w_{1,p}, \dots w_{k_p,p}\}$ of $k_p < p$ residue classes modulo p, and moreover assume that $k_p = 0$ for $p \notin \mathscr{P}$. Now, denote by $\tilde{S}(\mathscr{A}, \mathscr{P}, z)$ the number of elements of $\mathscr{A}$ avoiding all of these residue classes modulo p for $p|P(z)$. Explicitly,

$$\tilde{S}(\mathscr{A}, \mathscr{P}, z) = \#\{n \in \mathscr{A} \ : \ n \not\equiv w_{i,p} \pmod{p} \ \forall \ 1 \leqslant i \leqslant k_p, \ p|P(z)\}.$$

Then,

$$\tilde{S}(\mathscr{A}, \mathscr{P}, z) \leqslant \frac{x + z^2}{L(z)}, \tag{11.3.1}$$

where

$$L(z) = \sum_{1 \leqslant q \leqslant z} \mu(q)^2 \prod_{p|q} \frac{k_p}{p - k_p}.$$

The reason that this is a "large" sieve is because excluding more congruence classes makes the denominator $L(z)$ larger. Even though the large sieve gives better bounds when we exclude more residue classes, (11.3.1) still allows us to obtain very good bounds even when we exclude only 1 residue class:

Example 1 (Brun–Titchmarsh revisited)**.** Let

$$\begin{aligned}
\mathscr{A} &= \{a + mn \ : \ n \leqslant x/m\} \\
\mathscr{P} &= \{p \ : \ \gcd(p, m) = 1\} \\
k_p &= \begin{cases} 1 & \text{if } p \in \mathscr{P} \\ 0 & \text{otherwise} \end{cases} \\
w_{1,p} &= 0
\end{aligned}$$

Then

$$\tilde{S}(\mathscr{A},\mathscr{P},z)=\sum_{\substack{\alpha\in\mathscr{A}\\ \alpha\not\equiv 0 \pmod p\ \forall\ p|P(z)}}1=S(\mathscr{A},\mathscr{P},z).$$

Now, we know that for any $x>z\geqslant 2$, we have

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

Therefore, by the arithmetic large sieve, and the trivial bound $\pi(z;m,a)\leqslant z$, we have

$$\pi(x;m,a)\leqslant\frac{\frac{x}{m}+z^2}{L(z)}+z.$$

To deal with $L(z)$, note that if $(q,m)=1$ and $p|q$, then $(p,m)=1$, so that $k_p=1$. Similarly, if $(q,m)\neq 1$, then there is some $p|q$ such that $(p,m)\neq 1$, so that $k_p=0$ for said prime. This shows that

$$\begin{aligned}L(z)&=\sum_{\substack{1\leqslant q\leqslant z\\(q,m)=1}}\mu(q)^2\prod_{p|q}\frac{1}{p-1}\\&=\sum_{\substack{1\leqslant q\leqslant z\\(q,m)=1}}\frac{\mu(q)^2}{\varphi(q)}\\&=\sum_{\substack{1\leqslant q\leqslant z\\q|P(z)}}\frac{\mu(q)^2}{\varphi(q)},\end{aligned}$$

where the last equality follows from the fact that we are only summing over square-free q. Now, recall that in Selberg's sieve we studied the function

$$V(z)=\sum_{\substack{1\leqslant q\leqslant z\\q|P(z)}}\frac{\mu(q)^2}{f_1(q)},$$

where f_1 satisfies $f(n)=\sum_{d|n}f_1(d)$ for some multiplicative function f. Hence, $L(z)=V(z)$ with $f_1(q)=\varphi(q)$ and $f(n)=n$. In Exercise 9.3 you needed to bound $V(z)$, and it was possible to show that

$$V(z)\geqslant\frac{\varphi(m)}{m}\log z,$$

so that

$$\pi(x;m,a) \leqslant \frac{x+mz^2}{\varphi(m)\log z} + z.$$

Here, take $z := \Big(\frac{x}{m\log\frac{x}{m}}\Big)^{\frac{1}{2}}$. We see that

$$\pi(x;m,a) \leqslant \frac{x(1+\frac{1}{\log\frac{x}{m}})}{\varphi(m)\frac{1}{2}(\log\frac{x}{m} - \log\log\frac{x}{m})} + z.$$

Now, if $\frac{x}{m} \to \infty$,

$$\frac{1}{\log\frac{x}{m} - \log\log\frac{x}{m}} = \frac{1}{\log\frac{x}{m}} + o(1).$$

Thus,

$$\pi(x;m,a) \leqslant \frac{2x}{\varphi(m)}\Big(\frac{1}{\log\frac{x}{m}} + o\Big(\frac{1}{\log\frac{x}{m}}\Big)\Big) + z.$$

Of course, $z = o\Big(\frac{x}{\varphi(m)\log\frac{x}{m}}\Big)$, and so

$$\pi(x;m,a) \leqslant \frac{(2+o(1))x}{\varphi(m)\log\frac{x}{m}}. \tag{11.3.2}$$

Using more careful bounds, via the large sieve, Montgomery and Vaughan [13] were able to remove the $o(1)$ from the inequality, i.e., they showed that

$$\pi(x;m,a) \leqslant \frac{2x}{\varphi(m)\log\frac{x}{m}}.$$

Note that this is the best bound that we can reasonably hope for, in the sense that any improvements on the constant 2 would imply the non-existence of Siegel zeros (which were first introduced in these notes in the discussion following Theorem 6.3.3); see [14] for details.

11.4 The Bombieri-Vinogradov Theorem

The goal of this section is to study an important application of the large sieve: the Bombieri-Vinogradov Theorem. The theorem is now named after Enrico Bombieri and A. I. Vinogradov, who independently proved this result in 1965 and 1966, respectively. (Due to the Iron Curtain, mathematicians in the Soviet Union were often unaware of recent results on the other side. Mathematicians in the West were also sometimes unaware, or even neglectful, of what was happening on the Soviet side. As a result, it makes sense to give credit to both Bombieri and Vinogradov for this remarkable theorem.)

The Bombieri-Vinogradov Theorem is used in many important proofs in modern analytic number theory, including the celebrated proofs on bounded gaps between primes due to Zhang and Maynard. It is a result on primes in arithmetic progressions. It says, roughly speaking, that primes are reasonably uniformly distributed (mod q) for "small" integers q (where "small" here means values of q at least up to $\sqrt{x}$).

Let

$$\psi(x;q,a) := \sum_{\substack{n\leqslant x\\ n\equiv a \pmod q}} \Lambda(n).$$

We begin by recalling the following refinement of the prime number theorem for arithmetic progressions (Theorem 6.3.2):

Theorem 11.4.1 (Siegel–Walfisz Theorem)**.** *For all $A > 0$ there is a positive constant C_A (depending only on A) such that*

$$\psi(x;q,a) = \frac{x}{\varphi(q)} + O\big(x\exp(-C_A\sqrt{\log x})\big)$$

for all $(a,q) = 1$ and $q \leqslant (\log x)^A$.

Of course we expect a better error term in a much wider range for q:

Remark 11.4.2. Assuming the Generalized Riemann Hypothesis, for every $(a,q) = 1$, we have

$$\psi(x;q,a) - \frac{x}{\varphi(q)} = O(x^{\frac{1}{2}+\delta})$$

for all $\delta > 0$.

Assume for a moment that GRH is true and let $\varepsilon > 0$. Then, letting $\delta = \varepsilon/2$ in the previous remark, we have

$$\sum_{q \leqslant x^{\frac{1}{2}-\varepsilon}} \sup_{(a,q)=1} \left|\psi(x;q,a) - \frac{x}{\varphi(q)}\right| \ll x^{1-\frac{\varepsilon}{2}} \ll_{A,\varepsilon} \frac{x}{(\log x)^A} \tag{11.4.1}$$

for any $A > 0$. It turns out that we don't need GRH to show that (11.4.1) is true, and this result is known as the *Bombieri-Vinogradov theorem*; in other words, (11.4.1) says that GRH is true on average. We now state the result in the following slightly stronger form:

Theorem 11.4.3 (The Bombieri–Vinogradov Theorem)**.** *For any $A, \varepsilon > 0$, we have*

$$\sum_{q \leqslant x^{\frac{1}{2}-\varepsilon}} \sup_{\substack{y \leqslant x \\ (a,q)=1}} \left|\psi(y;q,a) - \frac{y}{\varphi(q)}\right| \ll_{A,\varepsilon} \frac{x}{(\log x)^A}.$$

The main idea of the proof is to use a method developed by Vaughan to deal with sums of the form

$$\sum \Lambda(n) f(n).$$

Unfortunately, these results are a bit technical, and they go beyond the scope for this course. The interested reader can see the full proof in Appendix A. For the purposes of the exercises, it will only be necessary to understand the statement of the Bombieri-Vinogradov theorem.

Sometimes it feels more natural to think about the Bombieri-Vinogradov theorem in terms of the $\pi(x)$ function instead of Chebyshev's $\psi(x)$ function. However, it is important to note that the result is NOT true when we merely replace $\psi(x;q,a)$ with $\pi(x;q,a)$ and $\frac{x}{\varphi(q)}$ with $\frac{x}{\varphi(q)\log x}$. The issue is that the error term is too large when we try to approximate $\pi(x;q,a)$ with $\frac{x}{\varphi(q)\log x}$. Instead, we need the precision of the logarithmic integral, $\mathrm{Li}(x)$:

Theorem 11.4.4 (Bombieri-Vinogradov, 1965/6)**.** *For every constant $A > 0$, there exists a constant $B = B(A)$ such that*

$$\sum_{q \leqslant Q} \max_{\substack{a \bmod q \\ (a,q)=1}} \left|\pi(x;q,a) - \frac{\mathrm{Li}(x)}{\varphi(q)}\right| \ll_A \frac{x}{(\log x)^A},$$

where $Q = \frac{x^{1/2}}{(\log x)^B}$.

We will prove a special case Theorem 11.4.4 in Exercise 11.7.

Theorem 11.4.4 tell us that the primes are well-distributed in residue classes modulo q for "small" values of q (up to about $x^{1/2}$). In fact, we expect that much more is true. There is a famous conjecture of Elliot and Halberstam which says that one should be able to replace the $\frac{x^{1/2}}{(\log x)^B}$ bound for Q by $x^{1-\varepsilon}$ for any $\varepsilon > 0$ instead. If the Elliot-Halberstam Conjecture is true, it would mean that primes are well-distributed in residue classes modulo q for "large" values of q as well. So far, this has not been proven for a single value of $\varepsilon < 1/2$. However, Yitang Zhang showed in 2013 that, by adding two extra conditions to the Theorem 11.4.4 (namely, that the moduli q are squarefree and y-smooth; that is, they do not contain any prime factors larger than y) then the Elliott-Halberstam Conjecture holds with $\varepsilon = \frac{1}{584}$. He used this in order to prove, for the very first time, that there are bounded gaps between primes.

We will examine some of the groundbreaking work on bounded gaps between primes, and see the essential role that sieve methods played in these proofs, in the final lecture of this course.

11.5 Exercises

Exercise 11.1. *For $\delta > 0$ and $b, d \in \mathbb{Z}$, define the intervals*

$$I(b,d,\delta) := \big(\frac{b}{d} - \frac{\delta}{2}, \frac{b}{d} + \frac{\delta}{2}\big).$$

(i) *Show that for $\delta := \dfrac{1}{Q^2}$ and $1 \leqslant b \leqslant d, (b,d) = 1, d \leqslant Q$, the intervals $I(b,d,\delta)$ do not overlap.*

(ii) *Let f be as in Lemma 11.2.2 Show that*

$$\sum_{1 \leqslant b \leqslant d} \Big| f\Big(\frac{b}{d}\Big)\Big| \leqslant d \int_0^1 |f(t)|\,\mathrm{d}t + \int_0^1 |f'(t)|\,\mathrm{d}t.$$

Hint: Take $\delta = \dfrac{1}{d}$ in (11.2.4).

Exercise 11.2 (Parseval's Identity)**.** *Let (a_n) be a sequence of complex numbers, and let*

$$S(t) := \sum_{n \leqslant x} a_n\, \mathrm{e}(nt).$$

Show that

$$\int_0^1 |S(t)|^2\,\mathrm{d}t = \sum_{n \leqslant x} |a_n|^2.$$

Hint: $|S(t)|^2 = S(t)\overline{S(t)}$.

Exercise 11.3. *The purpose of this exercise is to give an upper bound for* (11.2.2) *using the methods developed in this chapter for dealing with exponential sums. Given a set of integers $\mathscr{A} \subseteq [1,x]$, let $F_{\mathscr{A}}(t) = \sum_{n \in \mathscr{A}} \mathrm{e}(nt)$ be its Fourier transform.*

(i) *Show that*

$$\sum_{d \leqslant Q} \frac{1}{d} \sum_{b=1}^{d-1} \Big| F_{\mathscr{A}}\Big(\frac{b}{d}\Big)\Big| \leqslant \log Q \sum_{d \leqslant Q} \frac{1}{d} \sum_{\substack{b=1 \\ (b,d)=1}}^{d-1} \Big| F_{\mathscr{A}}\Big(\frac{b}{d}\Big)\Big|.$$

Hint: write $\dfrac{b}{d} = \dfrac{b'}{d'}$ with $(b',d') = 1$ and write $d = d'd''$.

(ii) *Use the same method as in the proof of the large sieve inequality to show that*

$$\sum_{d\leqslant Q}\sum_{\substack{b=1\\(b,d)=1}}^{d-1}\Big|F_{\mathscr{A}}\Big(\frac{b}{d}\Big)\Big| \ll (Q^2+x)\#\mathscr{A}^{\frac{1}{2}}.$$

Hint: You can trivially bound the integrals by the Cauchy–Schwartz inequality and then use Parseval's identity.

(iii) *Use partial summation to show that*

$$\sum_{d\leqslant Q}\frac{1}{d}\sum_{\substack{b=1\\(b,d)=1}}^{d-1}\Big|F_{\mathscr{A}}\Big(\frac{b}{d}\Big)\Big| \ll (Q+x)\#\mathscr{A}^{\frac{1}{2}}.$$

Conclude that

$$\sum_{d\leqslant Q}\Big|\#\mathscr{A}_d-\frac{\#\mathscr{A}}{d}\Big| \ll \log Q(Q+x)\#\mathscr{A}^{\frac{1}{2}}. \tag{11.5.1}$$

Remark 11.5.1. Note that the trivial bound $|F_{\mathscr{A}}(t)|\leqslant \#\mathscr{A}$ gives us the estimate

$$\sum_{d\leqslant Q}\Big|\#\mathscr{A}_d-\frac{\#\mathscr{A}}{d}\Big| \leqslant \#\mathscr{A}Q,$$

and so (11.5.1) only beats the trivial bound when $\dfrac{Q}{\log Q}\gg\dfrac{x}{\#\mathscr{A}^{\frac{1}{2}}}$. Also, note that (11.5.1) is not $\ll_B \dfrac{\#\mathscr{A}}{(\log x)^B}$ for any $B>0$, so this bound is not good enough for a type I estimate. Nevertheless, in the next exercise we will see how these methods can be adapted to sometimes obtain a type I estimate.

Exercise 11.4. *The purpose of this exercise is to give the general idea on how we can use large sieve type machinery in order to obtain type I estimates. Let* $\mathscr{A}:=\mathbb{Z}^+\cap[1,x]$, *and as usual, for* $t\in[0,1]$, *let* $F_x(t):=\sum_{n\in\mathscr{A}}\mathrm{e}(nt)$ *denote the Fourier transform* $\mathscr{A}$.

(i) *Show that*

$$|F_x(t)|=\Big|\frac{\sin(\pi xt)}{\sin(\pi t)}\Big|.$$

Hint: Note that $\mathrm{e}(\alpha)-1=\mathrm{e}(-\frac{\alpha}{2})(\mathrm{e}(\frac{\alpha}{2})-\mathrm{e}(-\frac{\alpha}{2}))$.

(ii) *Prove that for all* $t \in [0,1]$,

$$|F_x(t)| \leqslant \min\Big\{x, \frac{1}{2\|t\|}\Big\},$$

where $\|\cdot\|$ *denotes the distance to the nearest integer function.*

Hint: Begin by comparing the graph of the functions $\sin(\pi t)$ and $2t$ in the interval $[0, 1/2]$.

(iii) *Show that*

$$\int_0^1 |F_x(t)|\,\mathrm{d}t \ll \log x.$$

Hint: Begin by writing

$$\int_0^{\frac{1}{2}} |F_x(t)|\,\mathrm{d}t = \int_0^{1/x} |F_x(t)|\,\mathrm{d}t + \int_{1/x}^{\frac{1}{2}} |F_x(t)|\,\mathrm{d}t,$$

and use the bound from part (ii). To deal with the integral in the interval $(1/2, 1)$, think about using symmetry.

(iv) *Using the definition of* $F_x'(t)$*, via partial summation, prove that*

$$\int_0^1 |F_x'(t)|\,\mathrm{d}t \ll x \log x.$$

(v) *Show that*

$$\sum_{d \leqslant Q} \sum_{\substack{b=1 \\ (b,d)=1}}^{d-1} \Big|F_x\Big(\frac{b}{d}\Big)\Big| \ll (Q^2 + x)\log x.$$

Deduce from this that for any $Q_1 \geqslant 1$,

$$\sum_{Q_1 \leqslant d \leqslant Q} \frac{1}{d} \sum_{\substack{b=1 \\ (b,d)=1}}^{d-1} \Big|F_x\Big(\frac{b}{d}\Big)\Big| \ll \Big(Q + \frac{x}{Q_1}\Big)\log x.$$

(vi) *Use (ii) to show that*

$$\sum_{d \leqslant Q_1} \frac{1}{d} \sum_{\substack{b=1 \\ (b,d)=1}}^{d-1} \Big|F_x\Big(\frac{b}{d}\Big)\Big| \ll Q_1 \log Q_1.$$

(vii) *Let $\varepsilon, B > 0$, and take $Q_1 = (\log x)^{B+1}$, and $Q = x^{1-\varepsilon}$. Combine (v) and (vi) to conclude that*

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

Remark 11.5.2. You may have noticed that the conclusion from this exercise is trivial in the sense that we can avoid all of these computations by noting that $\#\mathscr{A}_d = \lfloor x/d \rfloor$, and so

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

However, the idea was to show how the methods from this chapter can be applied to certain sets. A non-trivial example is the set of integers with missing digits in base 10 (for example, the set of integers with no 7's in their decimal expansions). In this case, it is not easy to describe $\#\mathscr{A}_d$, but nevertheless Maynard [10] was able to obtain a type I estimate for this set using the methods described in this exercise, which roughly amounts to giving L^∞ and L^1 bounds for the Fourier transform (parts (ii) and (iii) respectively). Moreover he was able to show that $\mathscr{A}$ contains infinitely many primes by combining the type I estimate with an estimate of a bilinear sum associated to $\mathscr{A}$.

The philosophy here is that, if we are able to understand the Fourier transform associated to a set well enough, then we can understand the set in arithmetic progressions.

Exercise 11.5. *The purpose of this exercise is to prove the arithmetic large sieve inequality. Let*

$$S(t) := \sum_{n \leqslant x} a_n \, \mathrm{e}(nt),$$

where (a_n) is a non zero sequence of complex numbers.

(i) *Define*

$$\nu(p) := \#\{h \ : \ 0 \leqslant h < p, \ n \equiv h \pmod{p} \Rightarrow a_n = 0\}$$

and let

$$h(q) := \mu(q)^2 \prod_{p \mid q} \frac{\nu(p)}{p - \nu(p)}.$$

Show that the inequality

$$|S(0)|^2 h(q) \leqslant \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| S\Big(\frac{a}{q}\Big) \right|^2 \tag{11.5.2}$$

is equivalent to

$$|S(\beta)|^2 h(q) \leqslant \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| S\Big(\frac{a}{q} + \beta\Big) \right|^2 \quad (\beta \in \mathbb{R}) \tag{11.5.3}$$

(ii) *Suppose that* (11.5.3) *is satisfied for q and q' with $(q, q') = 1$. Show that* (11.5.2) *is true when you plug in qq' instead of q. Deduce from this that it suffices to establish* (11.5.2) *when q is prime.*

Hint: Begin by noting that by the Chinese Remainder Theorem,

$$\sum_{\substack{1 \leqslant c \leqslant qq' \\ (c, qq')=1}} \left| S\Big(\frac{c}{qq'}\Big) \right|^2 = \sum_{\substack{1 \leqslant a \leqslant q \\ (a,q)=1}} \sum_{\substack{1 \leqslant b \leqslant q' \\ (b,q')=1}} \left| S\Big(\frac{a}{q} + \frac{b}{q'}\Big) \right|^2$$

(iii) *For any prime number p, let*

$$S(p, h) := \sum_{\substack{n \leqslant x \\ n \equiv h \pmod{p}}} a_n.$$

Show that

$$|S(0)|^2 + \sum_{a=1}^{p-1} \left| S\Big(\frac{a}{p}\Big) \right|^2 = p \sum_{h=0}^{p-1} |S(p,h)|^2.$$

Hint: Use the orthogonality relations.

(iv) *Prove that*

$$|S(0)|^2 \leqslant (p - \nu(p)) \sum_{h=0}^{p-1} |S(p,h)|^2.$$

Hint: You do not need to use part (iii). Instead begin by observing that $S(0) = \sum_{h=0}^{p-1} S(p,h)$.

(v) *Show that* (11.5.2) *holds for all* $q \geqslant 1$. *Use this together with the large sieve inequality to conclude that*

$$\Big| \sum_{n \leqslant x} a_n \Big|^2 \leqslant \frac{x + Q^2}{L(Q)} \sum_{n \leqslant x} |a_n|^2, \tag{11.5.4}$$

where

$$L(Q) := \sum_{q=1}^{Q} h(q).$$

(vi) *Prove Theorem 11.3.1.*

Exercise 11.6. *For each odd prime* p, *let* $n_2(p)$ *denote the least quadratic non-residue, i.e., the smallest positive integer* $n_2(p)$ *such that* $(\frac{n_2(p)}{p}) = -1$. *Note that* $n_2(p)$ *is prime.*

A famous conjecture of Vinogradov states that $n_2(p) \ll_\varepsilon p^\varepsilon$ *for all* $\varepsilon > 0$. *Even though this conjecture remains open, in 1941 Linnik developed the large sieve and used it to show that large values of* $n_2(p)$ *are very rare. More specifically, he showed that*

$$\#\{p \leqslant x \ : \ n_2(p) > x^\varepsilon\} \ll_\varepsilon 1$$

for all $x \geqslant 1$. *The purpose of this exercise is to prove Linnik's result.*

(i) *Fix* $\varepsilon > 0$. *Let* N *be a positive integer and let*

$$\mathscr{A} := \Big\{ 1 \leqslant n \leqslant N \ : \ \Big(\frac{n}{p}\Big) = 1 \, \forall \, p \in \mathcal{P} \Big\},$$

where

$$\mathcal{P} := \Big\{ p \leqslant \sqrt{N} \ : \ \Big(\frac{n}{p}\Big) = 1 \, \forall \, n \leqslant N^{\varepsilon} \Big\}.$$

Show that $\#\mathscr{A} \gg_{\varepsilon} N$.

Hint: Start by showing that $\mathscr{A}$ contains all numbers $1 \leqslant n \leqslant N$ free of prime divisors $> N^{\varepsilon}$. Use this to deduce that $\mathscr{A}$ contains all the numbers of the form $n = mp_1 \cdots p_k$, where m is any integer such that $1 \leqslant m \leqslant \dfrac{N}{p_1 \cdots p_k}$, and $N^{\varepsilon - \varepsilon^2} < p_j < N^{\varepsilon}$ for $1 \leqslant j \leqslant k = \varepsilon^{-1}$.

(ii) *Let* a_n *be the characteristic function of* $\mathscr{A}$. *Use the arithmetic large sieve in the form of Exercise 11.5 (v) to show that*

$$L(\sqrt{N}) \ll \frac{N}{\#\mathscr{A}}.$$

(iii) *Prove that*

$$\sum_{p \in \mathcal{P}} \Big(1 - \frac{1}{p}\Big) \ll L(\sqrt{N}).$$

(iv) *Deduce Linnik's result:*

$$\#\{p \leqslant N \ : \ n_2(p) > N^{\varepsilon}\} \ll_{\varepsilon} 1.$$

Hint: Show that $\#\mathcal{P} \ll_{\varepsilon} 1$.

Exercise 11.7. *Let*

$$\vartheta(x; q, a) = \sum_{\substack{p \leqslant x \\ p \equiv a \pmod{q}}} \log p.$$

(i) *Show that*

$$\pi(x; q, a) = \frac{\vartheta(x; q, a)}{\log x} + \int_2^x \frac{\vartheta(t; q, a)}{t \log^2 t} \, \mathrm{d}t.$$

Hint: Use partial summation.

(ii) *Prove that*

$$\pi(x; q, a) = \frac{\psi(x; q, a)}{\log x} + \int_2^x \frac{\psi(t; q, a)}{t \log^2 t} \, \mathrm{d}t + O\Big(\frac{\sqrt{x}}{\log x}\Big).$$

Hint: First show that $\psi(x; q, a) - \vartheta(x; q, a) = O(\sqrt{x})$.

(iii) *Use the Bombieri Vinogradov theorem to show that*

$$\sum_{q\leqslant x^{\frac{1}{2}-\varepsilon}} \sup_{(a,q)=1} \Big|\pi(x;q,a) - \frac{\mathrm{Li}(x)}{\varphi(q)}\Big| \ll_{A,\varepsilon} \frac{x}{(\log x)^A}.$$

for all $A, \varepsilon > 0$.

Exercise 11.8. *Let* $\tau(n) = \sum_{d|n} 1$ *be the divisor counting function.*

(i) *Show that* $\tau(n) = 2\sum_{\substack{d|n\\ d\leqslant\sqrt{n}}} 1 - \delta(n)$, *where*

$$\delta(n) = \begin{cases} 1 \text{ if } n \text{ is a square,} \\ 0 \text{ otherwise.} \end{cases}$$

(ii) *Prove that for any* $a \in \mathbb{Z}$, *we have*

$$\sum_{p\leqslant x} \delta(p+a) \ll_a \sqrt{x}.$$

Deduce from this that

$$\sum_{p\leqslant x} \tau(p+a) = 2\sum_{d\leqslant\sqrt{x}} \pi(x;d,-a) + O_a(\sqrt{x}).$$

(We replace $d \leqslant \sqrt{x+a}$ *by* $d \leqslant \sqrt{x}$ *with an error term of* $O_a(\sqrt{x})$.*)*

(iii) *For this exercise, you may use the following version of the Bombieri Vinogradov theorem:*

$$\sum_{q\leqslant \frac{x^{\frac{1}{2}}}{(\log x)^8}} \sup_{(a,q)=1} \Big|\pi(x;q,a) - \frac{\mathrm{Li}(x)}{\varphi(q)}\Big| \ll \frac{x}{(\log x)^2}.$$

Show that there is some positive constant c *such that*

$$\sum_{p\leqslant x} \tau(p+a) = cx + O_a\Big(\frac{x\log\log x}{\log x}\Big).$$

Hint: Split the sum in the following way:

$$\sum_{d\leqslant\sqrt{x}}\pi(x;d,-a)=\sum_{d\leqslant\frac{x^{1/2}}{(\log x)^8}}\pi(x;d,-a)+\sum_{\frac{x^{1/2}}{(\log x)^8}\leqslant d\leqslant\sqrt{x}}\pi(x;d,-a).$$

For the first sum, use the given form of the Bombieri Vinogradov theorem and for the second sum, use (11.3.2). You may also use that there is a constant $c_0>0$ with

$$\sum_{n\leqslant x}\frac{1}{\varphi(n)}=c_0\log x+O(1)$$

for all $x\geqslant 1$, and that

$$\mathrm{Li}(x)=\frac{x}{\log x}+O\Big(\frac{x}{(\log x)^2}\Big).$$

Exercise 11.9. *Show that there is some positive constant c such that*

$$\sum_{n\leqslant x}\frac{1}{\varphi(n)}=c\log x+O(1)$$

for all $x\geqslant 1$.

Hint: First show that $\frac{n}{\varphi(n)}=\sum_{d|n}\frac{\mu(d)^2}{\varphi(d)}$.

Exercise 11.10. *Show that*

$$\sum_{m>x}\frac{1}{\varphi(m^2)}=O\Big(\frac{1}{\sqrt{x}}\Big).$$

Hint: Begin by showing that

$$\frac{\varphi(m)}{m}\gg\frac{1}{\log m}.$$

Exercise 11.11. *Show that every sufficiently large integer n can be written as the sum of a prime and a squarefree number.*

Hint: Let $Q(n)$ denote the number of representations of n as the sum of a prime number and a squarefree number. Begin by proving that

$$Q(n)=\sum_{\substack{m,r,p\\ n=m^2r+p}}\mu(m)=\sum_{m\leqslant\sqrt{n}}\mu(m)\pi(n;m^2,n).$$

Note that $\pi(n;m^2,n)=0$ if $m\geqslant\sqrt{n}$ and $\pi(n;m^2,n)\leqslant 1$ if $(m,n)>1$. Use these facts together with the ideas from Exercise 11.8 to show that $Q(n)\to\infty$ as $n\to\infty$. When applying the Brun Titchmarsh inequality, Exercise 11.10 can be useful.