Chapter 8

Brun's (Pure) Sieve

8.1 Introduction and Notation

In this chapter, we will look at our first example of a 20th century sieve: Brun's (pure) sieve. Developed by Viggo Brun in 1915, Brun's (pure) sieve is a truncated version of inclusion-exclusion. For many purposes, it will be easier to use than standard inclusion-exclusion.

Before proceeding further, we will define the basic notation that will be used throughout these lecture notes. Some of these definitions arise naturally from our discussion in Chapter 7.

Let $\mathscr{A}$ be a set of integers and let $\mathscr{P}$ be a set of primes. We will denote by $S(\mathscr{A}, \mathscr{P})$ the number of terms in $\mathscr{A}$ that are not divisible by any primes $p \in \mathscr{P}$. One of the goals in this chapter will be to estimate $S(\mathscr{A}, \mathscr{P})$ for various choices of $\mathscr{A}$ and $\mathscr{P}$. The general philosophy is that, if a set $\mathscr{A}$ has approximate size x and if the events "being divisible by a prime $p \in \mathscr{P}$" are close to being independent, with each occurring with probability roughly $g(p)$, then we would expect something like

$$S(\mathscr{A}, \mathscr{P}) \approx x \prod_{p \in \mathscr{P}} (1 - g(p))$$

to hold. The goal of Brun's pure sieve (and its variants) is to justify this approximation, and make the result more-quantitative (e.g., with reasonable error terms). And, of course, we would like to be able to use these kinds of arguments to handle a broad array of different counting problems.

Let $\mathscr{A}_d := \{a \in \mathscr{A} : d \mid a\}$ and take $P := \prod_{p\in\mathscr{P}} p$. We will sometimes only need to exclude integers with prime factors below a certain parameter, $z > 0$. Thus, we define $S(\mathscr{A}, \mathscr{P}, z) := S(\mathscr{A}, \mathscr{P} \cap [1, z])$ and $P(z) := \prod_{\substack{p\in\mathscr{P}\\ p\leqslant z}} p$. In the case that $\mathscr{P}$ is the set of all primes, we denote $S(\mathscr{A}, \mathscr{P}, z)$ by $S(\mathscr{A}, z)$.

To estimate $S(\mathscr{A}, \mathscr{P})$, we will usually assume that $\#\mathscr{A}_d \approx g(d)\#\mathscr{A}$ for some multiplicative function $g\colon \mathbb{N} \to [0, 1]$ in the sense that

$$\#\mathscr{A}_d = \#\mathscr{A}\, g(d) + r(d), \tag{8.1.1}$$

for squarefree d, where $r(d)$ is an error term that will usually be small. The general idea is to define $g(d)$ and $r(d)$ so that (8.1.1) holds. For future reference, we record the sieve notation in the following table:

Notation	Meaning
$\mathscr{A}$	a set of integers
$\mathscr{A}_d$	$\{a \in \mathscr{A} \ : \ d\|a\}$
p	a prime number
$\mathscr{P}$	a set of primes
P	$\prod_{p\in\mathcal{P}} p$
$P(z)$	$\prod_{\substack{p\in\mathscr{P}\\ p\leqslant z}} 1$
$S(\mathscr{A}, \mathscr{P})$	$\sum_{\substack{n\in\mathscr{A}\\ p\|n\Rightarrow p\notin\mathscr{P}}} 1$
$S(\mathscr{A}, \mathscr{P}, z)$	$S(\mathscr{A}, \mathscr{P} \cap [1, z])$
$S(\mathscr{A}, z)$	$\sum_{\substack{n\in\mathscr{A}\\ p\|n\Rightarrow p>z}} 1$

8.2 Motivating Examples

Example 1. (Sieve of Eratosthenes) For a positive integer x, let $\mathscr{A} = \mathbb{Z} \cap [1, x]$, and let $\mathscr{P}$ be the set of all primes. Then,

$$\#\mathscr{A}_d = \left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} - \left\{ \frac{x}{d} \right\} = \frac{\#\mathscr{A}}{d} - \left\{ \frac{x}{d} \right\},$$

so we can take $g(d) = 1/d$, and $r(d) = -\{x/d\}$.

Suppose now that we want to count how many integers remain after carrying out the Sieve of Eratosthenes. Recall that we only carry out the Sieve of Eratosthenes for primes up to $\sqrt{x}$, so if we want to know what remains after removing such primes, $S(\mathscr{A}, \sqrt{x})$ precisely counts this quantity. Then, we have

$$S(\mathscr{A}, \sqrt{x}) = \pi(x) - \pi(\sqrt{x}) + 1. \tag{8.2.1}$$

Later in this chapter, we will use inclusion-exclusion in order to try to obtain an upper bound for the number of integers that remain after performing the Sieve of Eratosthenes.

Example 2. (Twin Primes) Let $x \in \mathbb{Z}^+$, let $\mathscr{A} = \{n(n+2) : 1 \leqslant n \leqslant x\}$, and let $\mathscr{P}$ be the set of all primes. Let

$$\pi_2(x) = \#\{p \leqslant x : p + 2 \text{ is prime}\}.$$

Then,

$$S(\mathscr{A}, \mathscr{P}, \sqrt{x+2}) = \pi_2(x) - \pi_2(\sqrt{x+2}).$$

It is still an open problem whether there are infinitely many pairs of twin primes (i.e., as $x \to \infty$, does $\pi_2(x) \to \infty$?). Brun showed that most primes do not belong to a twin prime pair by showing that the sum of the reciprocals of twin primes converges. We will prove that later in this chapter.

For d squarefree, what is $\#\mathscr{A}_d$? This amounts to counting solutions to

$$n(n+2) \equiv 0 \pmod{d}.$$

Say that $N(d)$ is the number of such solutions. Then

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

In the notation defined at the beginning of this chapter, this means that we take $g(d) = \frac{N(d)}{d}$. We can actually ask the same question for any polynomial $f(x) \in \mathbb{Z}[x]$, i.e., for which values of n is $f(n) \equiv 0 \pmod{d}$? By the Chinese Remainder Theorem, the function $N(d)$ is multiplicative, hence $g(d)$ is multiplicative as well.

Example 3. (The $n^2 + 1$ problem) Let $\mathscr{A} = \{n^2 + 1 \ : \ 1 \leqslant n \leqslant x\}$ and let $\mathscr{P}$ be the set of all primes. Then $S(\mathscr{A}, \mathscr{P}, x) = S(\mathscr{A}, x)$ is the number of primes of the form $n^2 + 1$ which are greater than x. So integers n with $(x-1)^{1/2} < n \leqslant x$ survive. Notice that $\#\mathscr{A}_d = 0$ if d is divisible by a prime $p \equiv 3 \pmod{4}$.

When is $n^2 + 1 \equiv 0 \pmod{p}$? There is one solution if $p = 2$, no solutions if $p \equiv 3 \pmod{4}$, and two solutions if $p \equiv 1 \pmod{4}$. Thus, when d is squarefree, we have

$$N(d) = \prod_{p|d} N(p) = \begin{cases} 0 & \text{if } d \text{ divisible by some } p \equiv 3 \pmod{4} \\ 2^{W_0(d)} & \text{if else.} \end{cases}$$

As in the previous example, we take $g(d) = \frac{N(d)}{d}$.

Now that we understand how to set up these examples as sieving problems, we are ready to learn Brun's (pure) sieve, which will allow us to approximate $S(\mathscr{A}, \mathscr{P}, z)$.

8.3 Useful Lemmas

In this section, we present some lemmas that will be used for Brun's (pure) sieve.

Proposition 8.3.1. *With the notation defined as above, we have*

$$S(\mathscr{A}, \mathscr{P}, z) = \sum_{d|P(z)} \mu(d) \#\mathscr{A}_d.$$

Proof. Notice that

$$\begin{aligned} S(\mathscr{A}, \mathscr{P}, z) &= \sum_{\substack{a \in \mathscr{A} \\ \gcd(a, P(z))=1}} 1 \\ &= \sum_{a \in \mathscr{A}} \sum_{d \mid \gcd(a, P(z))} \mu(d) \\ &= \sum_{a \in \mathscr{A}} \sum_{\substack{d \mid a \\ d \mid P(z)}} \mu(d) \\ &= \sum_{d \mid P(z)} \mu(d) \#\mathscr{A}_d. \end{aligned}$$

□

We will also need the following combinatorial lemma:

Lemma 8.3.2. *Let $n \in \mathbb{N}$. Then*

$$\sum_{k=0}^{m} (-1)^k \binom{n}{k}$$

is positive or negative according to whether m is even or odd.

Proof. For $|x| < 1$, we have $(1-x)^{-1}(1-x)^n = (1-x)^{n-1}$. Hence, if we replace each term with its corresponding power series, we obtain

$$\sum_{k=0}^{\infty} x^k \sum_{k=0}^{n} (-1)^n \binom{n}{k} x^k = \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} x^k.$$

Let's look at the coefficients of x^m on both sides of the equation. On the lefthand side, we have $\sum_{k=0}^{m} (-1)^k \binom{n}{k}$. On the righthand side, we have $(-1)^m \binom{n-1}{m}$. Thus, if m is odd, then the sum $\sum_{k=0}^{m} (-1)^k \binom{n}{k}$ will be negative, and if m is even then the sum will be positive.

□

8.4 Brun's (pure) sieve

In the beginning of this section, we remarked that Brun's (pure) sieve is a truncated version of inclusion-exclusion. Now, we will finally see this in action. Observe that

$$S(\mathscr{A},\mathscr{P},z)=\sum_{d|P(z)}\mu(d)\#\mathscr{A}_d=\#\mathscr{A}_1-\sum_{p|P(z)}\#\mathscr{A}_p+\sum_{pq|P(z)}\#\mathscr{A}_{pq}\pm\cdots. \tag{8.4.1}$$

Useful Fact. If you stop at a "+" term in (8.4.1), you will have an overestimate, and if you stop at a "−" term, you will have an underestimate.

Let $\omega(n)$ denote the number of distinct prime divisors of n. Then,

$$S(\mathscr{A},\mathscr{P},z)=\sum_{k=0}^{\pi(z)}(-1)^k\sum_{\substack{d|P(z)\\ \omega(d)=k}}\#\mathscr{A}_d.$$

So, the Useful Fact says:

Proposition 8.4.1. *With the notation defined as above, we have*

$$S(\mathscr{A},\mathscr{P},z)\leqslant\sum_{k=0}^{m}(-1)^k\sum_{\substack{d|P(z)\\ \omega(d)=k}}\#\mathscr{A}_d$$

if m is even and

$$S(\mathscr{A},\mathscr{P},z)\geqslant\sum_{k=0}^{m}(-1)^k\sum_{\substack{d|P(z)\\ \omega(d)=k}}\#\mathscr{A}_d$$

if m is odd.

Proof. First, we have

$$\begin{aligned}\sum_{k=0}^{m}(-1)^k\sum_{\substack{d|P(z)\\ \omega(d)=k}}\#\mathscr{A}_d&=\sum_{\substack{d|P(z)\\ \omega(d)\leqslant m}}\mu(d)\#\mathscr{A}_d\\ &=\sum_{\substack{d|P(z)\\ \omega(d)\leqslant m}}\mu(d)\sum_{\substack{a\in\mathscr{A}\\ d|a}}1\\ &=\sum_{a\in\mathscr{A}}\sum_{\substack{d|(a,P(z))\\ \omega(d)\leqslant m}}\mu(d),\end{aligned}$$

where the final step follows from swapping the order of summation.

Assume that n is squarefree. If $m \geqslant \omega(n)$, then

$$f(n) = \sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1. \end{cases}$$

If $m < \omega(n)$: say $\omega(n) = W$ and n has the distinct prime factors $p_1, p_2, ..., p_W$. Then

$$\begin{aligned} f(n) &= \sum_{k=0}^{m} (-1)^k \sum_{\substack{d|n \\ \omega(d)=k}} 1 \\ &= \sum_{k=0}^{m} (-1)^k \binom{W}{k}. \end{aligned}$$

By Lemma 8.3.2, we have:

$$\begin{aligned} &f(1) = 1, \\ &f(n) = 0 \text{ if } n > 1 \text{ and } \omega(n) \leqslant m, \\ &f(n) \geqslant 0 \text{ if } m \text{ is even and } \omega(n) > m, \\ &f(n) \leqslant 0 \text{ if } m \text{ is odd and } \omega(n) > m. \end{aligned}$$

Then,

$$\sum_{\substack{d|P(z) \\ \omega(d)\leqslant m}} \mu(d) \#\mathscr{A}_d = \sum_{a\in\mathscr{A}} f(\gcd(a, P(z))).$$

If m is even then this sum is at least

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

In other words, we have

$$S(\mathscr{A}, \mathscr{P}, z) \leqslant \sum_{\substack{d|P(z) \\ \omega(d)\leqslant m}} \mu(d) \#\mathscr{A}_d$$

in the case where m is even. On the other hand, if m is odd,

$$S(\mathscr{A}, \mathscr{P}, z) \geqslant \sum_{\substack{d|P(z) \\ \omega(d)\leqslant m}} \mu(d) \#\mathscr{A}_d.$$

□

Remarks:

(1) The error term in Proposition 8.4.1 won't be too large since we have at most m terms that we are summing.
(2) In general, it is difficult to show that the lower and upper bounds are approximately equal.

Theorem 8.4.2 (Brun's pure sieve)**.** *Let* $\#\mathscr{A}_d = \#\mathscr{A} g(d) + r(d)$. *For all* $m \in \mathbb{Z}_{\geqslant 0}$, *we have*

$$S(\mathscr{A}, \mathscr{P}, z) = \#\mathscr{A} \prod_{\substack{p \in \mathscr{P} \\ p \leqslant z}} (1 - g(p)) + O\Big(\sum_{\substack{d \mid P(z) \\ \omega(d) \leqslant m}} |r(d)| \Big) + O\Big(\#\mathscr{A} \sum_{\substack{d \mid P(z) \\ \omega(d) \geqslant m}} g(d) \Big).$$

Proof. From Proposition 8.4.1, we know that if m is even, then

$$S(\mathscr{A}, \mathscr{P}, z) \leqslant \sum_{\substack{d \mid P(z) \\ \omega(d) \leqslant m}} \mu(d) \#\mathscr{A}_d.$$

Since $m - 1$ is odd, we have

$$\sum_{\substack{d \mid P(z) \\ \omega(d) \leqslant m-1}} \mu(d) \#\mathscr{A}_d \leqslant S(\mathscr{A}, \mathscr{P}, z).$$

An analogous argument applies when m odd. Therefore,

$$\begin{aligned}
S(\mathscr{A}, \mathscr{P}) &= \sum_{\substack{d \mid P(z) \\ \omega(d) \leqslant m}} \mu(d) \#\mathscr{A}_d + O\Big(\sum_{\substack{d \mid P(z) \\ \omega(d) = m}} \mathscr{A}_d \Big) \\
&= \sum_{\substack{d \mid P(z) \\ \omega(d) \leqslant m}} \mu(d) (\#\mathscr{A} g(d) + r(d)) + O\Big(\sum_{\substack{d \mid P(z) \\ \omega(d) = m}} \#\mathscr{A} g(d) + r(d) \Big) \\
&= \#\mathscr{A} \sum_{\substack{d \mid P(z) \\ \omega(d) \leqslant m}} \mu(d) g(d) + O\Big(\sum_{\substack{d \mid P(z) \\ \omega(d) \leqslant m}} |r(d)| \Big) + O\Big(\sum_{\substack{d \mid P(z) \\ \omega(d) = m}} \#\mathscr{A} g(d) \Big) \\
&= \#\mathscr{A} \Big(\sum_{d \mid P(z)} \mu(d) g(d) - \sum_{\substack{d \mid P(z) \\ \omega(d) > m}} \mu(d) g(d) \Big) + O\Big(\sum_{\substack{d \mid P(z) \\ \omega(d) \leqslant m}} |r(d)| \Big) + O\Big(\sum_{\substack{d \mid P(z) \\ \omega(d) = m}} \#\mathscr{A} g(d) \Big) \\
&= \#\mathscr{A} \prod_{p \in \mathscr{P}} (1 - g(p)) + O\Big(\sum_{\substack{d \mid P(z) \\ \omega(d) \leqslant m}} |r(d)| \Big) + O\Big(\#\mathscr{A} \sum_{\substack{d \mid P(z) \\ \omega(d) \geqslant m}} g(d) \Big).
\end{aligned}$$

□

8.5 Generalizations

The term "Brun's sieve" is often used to describe a number of different (but related) sieves. In each of these sieves, the goal is more-or-less the same: to obtain estimates of the form

$$S(\mathscr{A}, \mathscr{P}) \approx x \prod_{p \in \mathscr{P}} (1 - g(p)),$$

where g is as defined in the beginning of this chapter. Brun's pure sieve is the simplest version, and the one most often taught to students. But, it is worth emphasizing that sometimes it isn't powerful enough for a particular application. See, for example, Exercise 8.3. As hard as one might try, Brun's pure sieve alone just is not sufficient to prove this (very useful) result. Instead, we turn to a series of improvements on Brun's original idea. One such improvement is called the Brun-Hooley sieve. It comes from a 1994 paper of Christopher Hooley in which he started with Brun's pure sieve and derived a multidimensional version that allows for sharper bounds. It is important to note that Hooley proved both upper and lower bounds. However, the proof of the lower bound is quite complicated, so we omit describing it here. (Only the upper bound will be used in this course.)

Suppose that we partition a set of primes $\mathscr{P}$ into r disjoint sets,

$$\mathscr{P} = \bigcup_{j=1}^{r} \mathscr{P}_j.$$

Then n is divisible by no $p \in \mathscr{P}$ if and only if n is divisible by no $p \in \mathscr{P}_j$ for all $1 \leqslant j \leqslant r$. Define $P_j := \prod_{p \in \mathscr{P}_j} p$ and apply the machinery (with some adaptations) from the proof of Brun's pure sieve to $\mathscr{P}_j$ and P_j in place of $\mathscr{P}$ and P (for details, see Pollack pp. 182 - 185). This allows us to obtain:

Theorem 8.5.1 (Brun-Hooley Upper Bound Sieve)**.** *Let $\mathscr{P} = \bigcup_{j=1}^{r} \mathscr{P}_j$ be a partition of a set $\mathscr{P}$. Suppose that $g(p) < 1$ for all $p \in \mathscr{P}$. For any choice of nonnegative even integers $m_1, ..., m_r$, we have*

$$S(\mathscr{A}, \mathscr{P}) \leqslant x \prod_{p \in \mathscr{P}} (1 - g(p)) \exp\Big(\sum_{j=1}^{r} \Big(\sum^{(j)} / \prod^{(j)} \Big) \Big) + O\Big(\sum_{\substack{d_1, ..., d_r \\ d_j | P_j,\ \omega(d_j) \leqslant m_j}} |r(d_1 \cdots d_r)| \Big),$$

where $\prod^{(j)} := \prod_{p \in \mathscr{P}_j} (1 - g(p))$ *and* $\sum^{(j)} := \sum_{\substack{d_j | P_j \\ \omega(d_j) = m_j + 1}} g(d_j)$.

There is another variant that is often referred to simply as "Brun's sieve." In order to motivate it, we will present a more-general framework than what we encountered in Brun's pure sieve.

Recall that the main ingredient in the proof of Brun's pure sieve was to use Proposition 8.4.1 to show that, if m is even, then we have

$$\sum_{\substack{d|P(z)\\ \omega(d)\leqslant m-1}} \mu(d)\#\mathscr{A}_d \leqslant S(\mathscr{A},\mathscr{P},z) \leqslant \sum_{\substack{d|P(z)\\ \omega(d)\leqslant m}} \mu(d)\#\mathscr{A}_d.$$

In other words, for even m, we can define the functions

$$f_U(d) = \begin{cases} 1 & \text{if } \omega(d)\leqslant m \\ 0 & \text{otherwise,} \end{cases}$$

and

$$f_L(d) = \begin{cases} 1 & \text{if } \omega(d)\leqslant m-1 \\ 0 & \text{otherwise,} \end{cases}$$

which yields

$$\sum_{d|P(z)} \mu(d)f_L(d)\#\mathscr{A}_d \leqslant S(\mathscr{A},\mathscr{P},z) \leqslant \sum_{d|P(z)} \mu(d)f_U(d)\#\mathscr{A}_d. \tag{8.5.1}$$

Brun's idea for improving on his original "pure" sieve was to replace f_U and f_L by other functions that give sharper bounds. Namely, we can write $d = p_1p_2\cdots p_r$ with $p_1 > \cdots > p_r$ and set $y_1, y_2, y_3, \ldots$ in some way (independent of d). For an upper bound, we use

$$f_U(d) = \begin{cases} 1 & \text{if } p_m < y_m \text{ for all odd } m \\ 0 & \text{otherwise.} \end{cases}$$

For a lower bound, we take

$$f_L(d) = \begin{cases} 1 & \text{if } p_m < y_m \text{ for all even } m \\ 0 & \text{otherwise.} \end{cases}$$

Then, these choices of f_U and f_L satisfy (8.5.1). Brun had lots of different ideas on how to choose the y_m's; one of these choices yields what we now call

"Brun's sieve." However, the business of choosing the y_m's has led to many different variants on Brun's original ideas. For example, Iwaniec and Rosser noticed that one can allow the y_m's to depend on the previous primes. In particular, they took $y_m = (y/p_1 \cdots p_m)^{1/\beta}$ where $\beta > 1$ and y satisfies $p_1 \cdots p_{m-1} p_m^{\beta} < y$. This gives best-possible bounds in certain cases. Note that the Iwaniec-Rosser variant of Brun's sieve is often called "Brun's sieve" in the literature. However, this version of Brun's sieve was not known to Brun.

While we have just outlined the ideas behind Brun's sieve, the observant reader will notice that we still have not given a precise statement of Brun's sieve in a form similar to Theorem 8.4.2. To do so takes a bit of work. The parts that remain for us to show are "just" analysis. Unfortunately, the analysis is rather technical. One possible statement and proof of Brun's sieve can be seen in Section 6.2 of Cojocaru and Murty. This version is based in Brun's original work (i.e., it uses one of Brun's choices of the y_m's). We give a somewhat simplified version of the result from Cojocaru and Murty below. It is worth emphasizing that our simplified version is much weaker than what is given by Cojocaru and Murty, but it will be sufficient for all of the exercises in this chapter.

Theorem 8.5.2 (Brun's sieve)**.** *As usual, suppose that* $\#\mathscr{A}_d = \#\mathscr{A} g(d) + r(d)$ *for some multiplicative function* g*. 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 if* $2 \leqslant w < z$*, then*

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

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

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

8.6 Applications

For the remainder of this chapter, we will carefully examine several applications of inclusion-exclusion and Brun's (pure) sieve.

8.6.1 Sieve of Eratosthenes

In this section, we will attempt to use inclusion-exclusion to bound the count of integers that remain after the Sieve of Eratosthenes is performed. Let

$$\pi(x,z) := \#\{n \leqslant x : p \mid n \Rightarrow p > z\}.$$

Some people refer to the Sieve of Eratosthenes as a method for computing $\pi(x,z)$ in this way. Let us see how Proposition 8.3.1 can be applied to Example 1. Namely, for $z = x^{1/2}$, we have

$$\begin{aligned}\pi(x) - \pi(x^{1/2}) + 1 &= S(\mathscr{A}, \mathscr{P}, z) \\ &= \sum_{d \mid P(z)} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor.\end{aligned}$$

We have just obtained an exact formula for $\pi(x,z)$. However, it is a bit unsatisfying, since the result does not give us a sense of how large $\pi(x,z)$ really is (in terms of x and z). Observe that $\lfloor \frac{x}{d} \rfloor = \frac{x}{d} + (\lfloor \frac{x}{d} \rfloor - \frac{x}{d})$. We will use this to rewrite $\pi(x,z)$ in a way that allows us to avoid having floor functions in our main term. In particular, we have

$$\pi(x,z) = x \sum_{d \mid P(z)} \frac{\mu(d)}{d} + \sum_{d \mid P(z)} \mu(d) \left(\left\lfloor \frac{x}{d} \right\rfloor - \frac{x}{d} \right).$$

In the main term, we can replace the Dirichlet series with its corresponding Euler product, $\prod_{p \leqslant z}(1 - 1/p)$. We can crudely bound the error term by taking its absolute value and observing that this is at most $2^{\pi(z)}$. Therefore, we obtain

$$\pi(x,z) = x \prod_{p \leqslant z} \left(1 - \frac{1}{p}\right) + O(2^{\pi(z)}). \tag{8.6.1}$$

If z tends to infinity with x sufficiently slowly, then we will obtain an asymptotic here. Namely, by Mertens' Second Theorem, we have

$$x\prod_{p\leqslant z}\left(1-\frac{1}{p}\right)\sim e^{-\gamma}\frac{x}{\log z},$$

as $x\to\infty$. However, in this particular case (asking how many integers remain after performing the Sieve of Eratosthenes), we take $z=x^{1/2}$, which is rather large relative to x. In that case, the so-called "error term" is of size $O(2^{\sqrt{x}})$, which is much larger than the main term! As a result, the equation (8.6.1) does not give us the asymptotic formula that we were hoping for. In fact, it is not even true that $\pi(x,x^{1/2})\sim x\prod_{p\leqslant x^{1/2}}(1-1/p)$. By the prime number theorem, we see that

$$\pi(x,x^{1/2})=\pi(x)-\pi(x^{1/2})+1\sim x/\log x.$$

8.6.2 Primes of the form n^2+1

Let $\pi_{n^2+1}(x)$ denote the number of $n\leqslant x$ for which n^2+1 is prime. Here, we take $\mathscr{A}=\{n^2+1:n\leqslant x\}$ and $\mathscr{P}$ to be the set of all primes. Then, for $z\leqslant x$, we have

$$\pi_{n^2+1}(x)\leqslant S(\mathscr{A},\mathscr{P},z)+z^{1/2}.$$

As in Example 3, let $N(p)$ be the number of solutions to the congruence $n^2+1\equiv 0 \pmod p$. We saw that $N(2)=1$ and, for odd primes p, $N(p)=0$ or 2, depending on whether $p\equiv 3 \pmod 4$ or $1 \pmod 4$. In general, with $N(d)$ defined analogously for any positive integer d, note that

$$\left\lfloor\frac{x}{d}\right\rfloor N(d)\leqslant \#\mathscr{A}_d\leqslant\left\lceil\frac{x}{d}\right\rceil N(d).$$

Hence,

$$|r(d)|=\left|\#\mathscr{A}_d-x\frac{N(d)}{d}\right|\leqslant N(d).$$

Therefore, by Proposition 8.3.1 (inclusion-exclusion),

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

This shows that for $x \geqslant 0$ and $z \geqslant 2$, we have

$$\pi_{n^2+1}(x) \leqslant S(\mathscr{A}, \mathscr{P}, z) + z^{1/2} \tag{8.6.2}$$

$$= \frac{1}{2}x \prod_{\substack{p \leqslant z \\ p \equiv 1 \pmod 4}} \Big(1 - \frac{2}{p}\Big) + O\Big(\sum_{d \mid P(z)} N(d)\Big) + z^{1/2}. \tag{8.6.3}$$

First, we focus on the main term. By Exercises 8.1 and 8.2,

$$\prod_{\substack{p \leqslant z \\ p \equiv 1 \pmod 4}} \Big(1 - \frac{2}{p}\Big) \sim \frac{C}{\log z}$$

for some $C > 0$. To handle the O-term, we observe that

$$\sum_{d \mid P(z)} N(d) = \prod_{p \leqslant z} (1 + N(p)) < 3^z.$$

Inserting these into (8.6.2), and taking $z = \frac{1}{2}\log x$, allows us to conclude that $\pi_{n^2+1}(x) \ll x/\log\log x$. In particular, this tells us that the set of integers n such that $n^2 + 1$ is prime has asymptotic density zero.

8.6.3 Twin primes

Let

$$\pi_2(x) := \#\{n \leqslant x : n \text{ and } n+2 \text{ prime}\}.$$

Let's first convince ourselves heuristically that there should be infinitely many pairs of twin primes. By the prime number theorem, the probability that a random number $n \leqslant x$ is prime should be about $\frac{1}{\log x}$. By the same logic, the probability that a random number $n+2$ is prime should also be $\frac{1}{\log x}$. If these two events were independent, then the probability that both n and $n+2$ are simultaneously prime should be $\frac{1}{(\log x)^2}$, which would imply that there are about $\frac{x}{(\log x)^2}$ primes in the interval $[1, x]$. Since this tends to infinity with x, it would seem to imply that there are infinitely many pairs of twin primes... However, this is wrong, since the same argument would show that there are infinitely many pairs of primes $n, n+1$, and there is exactly one value of n with this property (namely, $n = 2$). Clearly, we were incorrect to assume that the events "n is prime" and "$n+2$ is prime" are independent!

We can revise our heuristic argument to try to correct for the problem of non-independence. Let p and p' be independently chosen random integers. Look at:

$$\frac{P(p, p+2 \text{ not both divisible by } q)}{P(p, p' \text{ not both divisible by } q)}, \tag{8.6.4}$$

for each small prime q. Since

$$P(q \mid p) = \frac{1}{q},$$

then

$$P(q \nmid p) = 1 - \frac{1}{q}.$$

Thus, the denominator in equation (8.6.4) should be

$$P(q \nmid p \text{ and } q \nmid p') = \left(1 - \frac{1}{q}\right)^2.$$

On the other hand, let's consider the numerator:

$$P(q \nmid p \text{ and } q \nmid (p+2)) = P(p \not\equiv 0 \text{ or } -2 \pmod{q}).$$

Observe that

$$P(p \not\equiv 0 \text{ or } -2 \pmod q) = \begin{cases} 1-2/q & \text{if } q>2 \\ 1-1/2 & \text{if } q=2. \end{cases}$$

Hence, if $q>2$ then the correction factor for divisibility by q is

$$\frac{(1-\frac{2}{q})}{(1-\frac{1}{q})^2}.$$

If $q=2$ then the correction factor is

$$\frac{1-\frac{1}{2}}{(1-\frac{1}{2})^2} = 2.$$

Thus, we define

$$c_2 := 2 \prod_{\substack{q \text{ prime} \\ q \geqslant 3}} \frac{(1-2/q)}{(1-1/q)^2} \approx 1.3203236....$$

This suggests that

$$\#\{p \leqslant x : p \text{ and } p+2 \text{ prime}\} \approx c_2 \frac{x}{(\log x)^2}.$$

We call c_2 the Twin Prime Constant.

Unfortunately, this is still not a rigorous proof, since it relies on the assumption that the primes p are uniformly distributed among the residue classes mod q for all $q \leqslant x$. We won't be able to give a precise asymptotic in this course, as doing so would amount to proving the Twin Prime Conjecture! However, we can use Brun's pure sieve in order to obtain an upper bound for $\pi_2(x)$.

Let us recall our notation from Example 2. We let $\mathscr{A} = \{n(n+2) : n \leqslant x\}$ and $\mathscr{P} = \{p \leqslant z : p \text{ prime}\}$. Moreover, let $N(p)$ be the number of solutions to the congruence $n(n+2) \equiv 0 \pmod p$. Recall that $N(2)=1$ and that $N(d) = 2^{\omega(d)}$ for $d|P$. Assume that $z \leqslant x^{1/20\log_2 x}$ and that x and z both tend to infinity. By Brun's (pure) sieve, we obtain

$$S(\mathscr{A},\mathscr{P}) = x\prod_{p\in\mathscr{P}}(1-g(p)) + O\Big(\sum_{\substack{d|P \\ \omega(d)\leqslant m}} |r(d)|\Big) + O\Big(x\sum_{\substack{d|P \\ \omega(d)\geqslant m}} g(d)\Big)$$

for all $m \geqslant 0$. We will first focus on bounding the main term, and then we will handle the error terms. For ease of reference, let

$$\beta_1 = \sum_{\substack{d|P \\ \omega(d)\leqslant m}} |r(d)| \quad \text{and} \quad \beta_2 = x \sum_{\substack{d|P \\ \omega(d)\geqslant m}} g(d).$$

Recall that

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

where $|r(d)| \leqslant N(d) = 2^{\omega(d)}$. Then by the corrected version of the heuristic argument above, we have

$$\begin{aligned}
\prod_{p\in\mathscr{P}}(1-g(p)) &= \frac{1}{2}\prod_{2<p\leqslant z}\left(1-\frac{2}{p}\right) \\
&= 2\prod_{2<p\leqslant z}\frac{1-\frac{2}{p}}{(1-\frac{1}{p})^2}\prod_{p\leqslant z}\left(1-\frac{1}{p}\right)^2 \\
&\sim \frac{2c_2e^{-2\gamma}}{(\log z)^2},
\end{aligned}$$

where the final asymptotic follows from Mertens' Second Theorem (and c_2 is the Twin Prime Constant).

To handle the error terms β_1 and β_2, we will take $m = 10\lfloor \log\log z \rfloor$. Then

$$\begin{aligned}
\beta_1 &= \sum_{\substack{d|P \\ \omega(d)\leqslant m}} |r(d)| \\
&\leqslant \sum_{\substack{d|P \\ \omega(d)\leqslant m}} 2^{\omega(d)} \\
&= \sum_{k=0}^{m} 2^k \binom{\pi(z)}{k} \\
&\leqslant \sum_{k=0}^{m} 2^k \pi(z)^k \\
&\leqslant \sum_{k=-\infty}^{m} (2\pi(z))^k = (2\pi(z))^m \cdot \frac{1}{1-\frac{1}{2\pi(z)}} \\
&\leqslant 2(2\pi(z))^m \\
&\leqslant 2z^m.
\end{aligned}$$

Hence, it follows that

$$\beta_1 \leqslant 2z^m \leqslant 2z^{10\log\log z} \leqslant 2z^{10\log\log x} \leqslant 2x^{1/2},$$

where the final inequality uses that $z \leqslant x^{1/20\log_2 x}$ (here we define $\log_2(x) = \log\log x$).

We will use this last inequality to show that the error term β_1 is small relative to the main term. Observe that

$$\frac{2x^{1/2}}{\frac{x}{(\log z)^2}} \leqslant \frac{2x^{-1/2}(\log x)^2}{(20\log_2 x)^2} \to 0$$

as $x \to \infty$. Therefore, $\beta_1 = o(\frac{x}{(\log z)^2})$.

Next, we will show that β_2 is also small relative to the main term. Recall that

$$\beta_2 = x \sum_{\substack{d|P \\ \omega(d)\geqslant m}} g(d) = x \sum_{k\geqslant m} \sum_{\substack{d|P \\ \omega(d)=k}} g(d).$$

By unique factorization and multiplicativity of g, we have

$$\sum_{\substack{d|P \\ \omega(d)=k}} g(d) = \sum_{p_1<p_2<\cdots<p_k\leqslant z} g(p_1)g(p_2)\cdots g(p_k).$$

Then, by the multinomial theorem, we have

$$\begin{aligned}\Big(\sum_{p\leqslant z} g(p)\Big)^k &= \sum_{\substack{k_1+\cdots+k_m=k \\ k_1,\ldots,k_m\geqslant 0}} \frac{k!}{k_1!\cdots k_m!}\prod_{j=1}^{m} g(p_j)^{k_j} \\ &\geqslant k! \sum_{p_1<p_2<\cdots<p_k\leqslant z} g(p_1)g(p_2)\cdots g(p_k).\end{aligned}$$

Therefore, we have

$$\sum_{\substack{d|P \\ \omega(d)=k}} g(d) \leqslant \frac{1}{k!}(\sum_{p\leqslant z} g(p))^k.$$

By Mertens' First Theorem,

$$\sum_{p\leqslant z} \frac{1}{p} \leqslant \log_2 z + c$$

for all $z \geqslant 3$, were c is a constant. Recalling that $g(p) \leqslant \frac{2}{p}$, we obtain

$$\beta_2 \leqslant x \sum_{k \geqslant m} \frac{1}{k!} (\sum_{p \leqslant z} g(p))^k \leqslant x \sum_{k \geqslant m} \frac{(2 \log_2 z + 2c)^k}{k!}.$$

Let $a_k = \frac{(2 \log_2 z + 2c)^k}{k!}$. Then

$$\frac{a_{k+1}}{a_k} = \frac{2 \log_2 z + 2c}{k+1} \leqslant \frac{2 \log_2 z + 2c}{10 \lfloor \log_2 z \rfloor + 1} \leqslant \frac{1}{2}$$

for z sufficiently large. Hence, $a_{k+1} \leqslant a_k/2$. As a result, we have

$$\begin{aligned} \beta_2 &\leqslant x \sum_{k \geqslant m} a_k \leqslant x(a_m + \frac{a_m}{2} + \frac{a_m}{4} + \cdots) \\ &= 2x a_m = 2x \frac{(2 \log_2 m + 2c)^m}{m!}. \end{aligned}$$

Now, since

$$e^m = 1 + m + \frac{m^2}{2!} + \cdots \geqslant \frac{m^m}{m!},$$

we have

$$m! \geqslant \left(\frac{m}{e} \right)^m.$$

Thus,

$$\beta_2 \leqslant 2x \left(\frac{2e \log_2 z + 2ec}{m} \right)^m \leqslant 2x \left(\frac{3}{5} \right)^m.$$

Now, since $m = 10 \lfloor \log_2 z \rfloor$, we have

$$\begin{aligned} 2x \left(\frac{3}{5} \right)^m &\ll 2x \left(\frac{3}{5} \right)^{10 \log_2 z} \\ &= 2x e^{10 \log \frac{3}{5} \log_2 z} \\ &= 2x e^{-5 \log_2 z} = \frac{2x}{(\log z)^5} = o\left(\frac{x}{(\log z)^2} \right), \end{aligned}$$

where the third inequality follows from the fact that $10 \log \frac{3}{5} < -5$. Therefore, $\beta_2 = o(\frac{x}{(\log z)^2})$, so both error terms are negligible relative to the main term.

We have just shown:

Theorem 8.6.1. *If $z \to \infty$ as $x \to \infty$, with $z \leqslant x^{1/20 \log_2 x}$, then*

$$S(\mathscr{A}, \mathscr{P}) \sim \frac{2c_2 e^{-2\gamma} x}{(\log z)^2},$$

where γ is the Euler-Mascheroni constant and c_2 is the Twin Prime constant.

Corollary 8.6.2. *We have*

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

Proof. If $n, n+2$ are both prime then $n \leqslant z$ or both $n, n+2$ have only prime factors exceeding z. Therefore,

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

Taking $z = x^{1/20 \log_2 x}$, we have

$$\pi_2(x) \ll x \left(\frac{\log_2 x}{\log x} \right)^2 + x^{1/20 \log_2 x} \ll x \left(\frac{\log_2 x}{\log x} \right)^2.$$

□

Corollary 8.6.3. *Let $\mathscr{P}$ be the set of all primes p such that $p+2$ is also prime. Then $\sum_{p \in \mathscr{P}} \frac{1}{p}$ converges.*

Proof. Let p_n be the n^{th} prime such that $p_n + 2$ is also prime. Then

$$n = \pi_2(p_n) \ll p_n \left(\frac{\log_2 n}{\log n} \right)^2.$$

Hence,

$$\frac{1}{p_n} \ll \frac{1}{n} \left(\frac{\log_2 n}{\log n} \right)^2.$$

By the integral test, we see that $\sum \frac{1}{p_n}$ converges.

□

Remark 8.6.4. The fact that the sum of reciprocals of twin primes converges DOES NOT imply that there are only finitely many pairs of twin primes. It simply tells us that the set of twin primes is much sparser than the set of primes.

8.7 Exercises

Exercise 8.1. *Suppose that for each prime p, we have an integer k_p with $0 \leqslant k_p < p$, $k_p = O(1)$, and that for some real numbers $c, d > 0$,*

$$\sum_{p \leqslant x} \frac{k_p \log p}{p} = c \log x + d + o(1).$$

Prove that there is some number $C > 0$ such that

$$\prod_{p \leqslant x} (1 - k_p/p) \sim C/(\log x)^c$$

as $x \to \infty$.
Hint: Use Exercise 6.3.

Exercise 8.2. *Suppose $k_p = 2$ if $p \equiv 1 \pmod 4$, $k_p = 0$ if $p \equiv 3 \pmod 4$ and $k_2 = 1$. Show that this choice of numbers k_p satisfies the previous problem with $c = 1$. What is the relevance of this problem to primes of the form $n^2 + 1$?*

Exercise 8.3 (due to P. Pollack)**.** *The following is often referred to as "Brun's method." Fix a natural number k.*

(a) Let $A > 0$. Suppose that to each prime $p \leqslant x^A$, we associate $k_p \leqslant k < p$ residue classes modulo p. Show that the number of natural numbers $n \leqslant x$ avoiding all of these residue classes is

$$\ll_{k,A} x \prod_{p \leqslant x^A} \left(1 - \frac{k_p}{p}\right) \text{ (for } x > 0),$$

where the implied constant is independent of the particular choice of residue classes.

(b) Show that there is a constant $B > 0$, depending only on k, with the following property: If we choose $k_p \leqslant k$ residue classes modulo p for each prime $p \leqslant x^B$, then the number of natural numbers $n \leqslant x$ avoiding all these classes is

$$\gg_k x \prod_{p \leqslant x^B} \left(1 - \frac{k_p}{p}\right) \text{ (for } x \to \infty),$$

again uniformly in the particular choice of residue classes.

Hint: Use the Chinese Remainder Theorem to construct a polynomial F for which $p \mid F(n)$ precisely when n falls into one of the k_p chosen residue classes mod p. For part (a), use Brun's sieve to show that the result holds for any $A \leqslant A_0$, for some $A_0 > 0$, and then extend the product to $A > A_0$ using Exercise 8.1.

Remark: From (a) and (b) we may re-derive the result given in this chapter regarding the twin prime problem, with a slight loss of precision. In that case, the forbidden residue classes are 0 and -2 mod p.

Exercise 8.4 (Brun-Titchmarsh inequality)**.** *Let $x \geqslant 2$. Suppose that a and m are coprime integers with $1 \leqslant m < x$. Prove that*

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

where the implied constant is absolute. (Recall that $\pi(x; m, a)$ denotes the number of primes $p \leqslant x$ with $p \equiv a \pmod{m}$.) Is this still true without the assumption that a and m are relatively prime?

Hint: You may find Exercises 6.3(a), 8.1, and 8.3 to be useful here.

Exercise 8.5 (due to P. Pollack)**.** *Suppose that $F(T) \in \mathbb{Z}[T]$ is irreducible over $\mathbb{Q}$ and that the leading coefficient of $F(T)$ is positive. For each natural number d, let $\nu(d)$ denote the number of roots of F modulo d. A theorem of Landau asserts that for $x \geqslant 3$,*

$$\sum_{p \leqslant x} \frac{\nu(p)}{p} = \log\log x + C_F + O_F\left(\frac{1}{\log x}\right),$$

where C_F is a constant depending on F. Use this along with Exercise 8.3 to show that the number of $n \leqslant x$ for which $F(T)$ is prime is $\ll_F x/\log x$ for $x \geqslant 3$.