SIEVE METHODS IN NUMBER THEORY : THE SIEVE OF ERATOSTHENES AND BRUN'S PURE SIEVE

NOHA NASRI

ABSTRACT. In this paper, we present an introduction to sieve methods in number theory, with particular focus on the classical sieve of Eratosthenes and Brun's pure sieve. We detail the study of these two sieve methods in a rigorous and simplified approach,with emphasis on improving estimates for the count of integers free of small prime divisors. Beginning with a refined analysis of the classical Eratosthenes sieve, we apply Rankin's trick to obtain sharper upper bounds. These estimates lead to an improved asymptotic formula for the generalized sieve under standard hypotheses. The second part of the paper focuses on Brun's pure sieve and its application to problems related to twin primes and integers with few prime factors. We also introduce a truncated Möbius function and derive bounds for sifted sets with power-saving error terms. Our approach combines combinatorial identities with analytic estimates, yielding explicit bounds that approach the strength of more advanced sieve methods.

1. Introduction

Sieve methods are advanced techniques in number theory, used to count the number of elements with a certain characteristic within larger sets of numbers. The first sieve was developed by Eratosthenes in third century BC, its idea is intuitive and it works like a prime counting algorithm, which was used by Legendre in his studies of the prime number counting function $\pi(x)$. Later, sieve theory was revisited by Viggo Brun in the 20th century. He developed what is known as Brun's sieve and he applied it to deduce many interesting results, such as, the convergence of the sum of reciprocals of twin primes, and proving that there are infinitely many primes with exactly seven prime factors. Brun's groundbreaking work has led to intense investigation and newfound interest in sieve methods, see [Bru16], [Bru20], [Bru22]. Over the years, mathematicians came up with other sieve methods : Selberg's sieve, Turàn's sieve,Rosser's Sieve, the large sieve...

These tools have led to numerous remarkable results. Notable examples include the Brun-Titchmarsh theorem and Bombieri's influential theorem. Another powerful illustration is Chen's theorem [Che02], which asserts the existence of infinitely many primes p such that $p + 2$ is semiprime (the product of at most two primes). Furthermore, sieve methods have been particularly useful recently, and were used in many breakthroughs regarding prime gaps such as proving that there are infinitely many primes that differ by a gap no more than 246, see [Pol14].

Sieve methods also play an important part in applied fields of number theory such as Algorithmic Number Theory, and Cryptography. In fact, they are used directly, for example for finding all the prime numbers below a certain bound, or constructing numbers free of large prime factors, and indirectly, for example to deduce valuable information about the distribution of smooth numbers in short intervals in order to bound the running time of several factoring algorithms. These applications show how valuable sieve methods are.

Date: July 12, 2025.

2. **Walkthrough**

In Section 1, we revisit the classical sieve of Eratosthenes, derive its main counting estimate using Möbius inversion, and refine the error term via Rankin's method. We also introduce the concept of smooth numbers and prove an upper bound for their distribution. The generalized sieve of Eratosthenes is then developed in an abstract framework using multiplicative functions. In Section 2, we study Brun's pure sieve, beginning with the combinatorial identities underlying the truncated Möbius function. We conclude by deriving an asymptotic formula for the size of sifted sets under general conditions, with optimized error bounds.

3. **Background**

Before we delve into the topic of this paper, it is important to introduce some concepts and notations.

3.1. **Big O Notation.** Let $D \subseteq \mathbb{C}$ and $f : D \to \mathbb{C}$. We write

$$f(x) = O(g(x))$$

if there exists a positive constant A such that

$$|f(x)| \leq Ag(x) \quad \text{for all } x \in D,$$

where $g : D \to \mathbb{R}^+$. Typically, D is $\mathbb{N}$ or $\mathbb{R}_0^+$.

Sometimes, the notation

$$f(x) \ll g(x) \quad \text{or} \quad g(x) \gg f(x)$$

is used to mean $f(x) = O(g(x))$.

3.2. **The Möbius function.** The Möbius function $\mu(\cdot)$ is multiplicative and defined by:

$$\mu(1) = 1, \quad \mu(p) = -1 \text{ for every prime } p, \quad \mu(p^a) = 0 \text{ for } a \geq 2.$$

Lemma 3.1. *The fundamental property of the Möbius function states that:*

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

Proof. If $n = 1$, the statement is verified; otherwise, let $n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ and set $N = p_1 p_2 \cdots p_r$, the radical of n. Since $\mu(d) = 0$ unless d is squarefree, we have $\sum_{d|n} \mu(d) = \sum_{d|N} \mu(d)$. The sum on the RHS contains 2^r terms corresponding to the subsets of $\{p_1, \ldots, p_r\}$. Note that the number of subsets with k elements is $\binom{r}{k}$, and for a divisor d determined by such a subset, $\mu(d) = (-1)^k$. Hence,

$$\sum_{d|n} \mu(d) = \sum_{k=0}^{r} \binom{r}{k} (-1)^k = (1 - 1)^r = 0.$$

■

Thus: - $\mu(n) = 0$ if n is not squarefree, - $\mu(n) = (-1)^k$ if n is the product of k distinct primes.

3.3. **Abel summation or Partial summation.**

Theorem 3.2. *Let* $c_1, c_2, \ldots$ *be a sequence of complex numbers and set*

$$S(x) := \sum_{n \le x} c_n.$$

Let n_0 *be a fixed positive integer. If* $c_j = 0$ *for* $j < n_0$ *and* $f : [n_0, \infty) \to \mathbb{C}$ *has continuous derivative in* $[n_0, \infty)$*, then for* x *an integer* $> n_0$ *we have*

$$\sum_{n \le x} c_n f(n) = S(x)f(x) - \int_{n_0}^{x} S(t)f'(t)\,dt.$$

4. **The Sieve of Eratosthenes**

The Eratosthenes sieve is the oldest sieve method (around 3rd century BC), named after Eratosthenes (276–194 BCE), a Greek mathematician. It is an ancient algorithm used to find prime numbers up to a certain limit, and it first appeared in the work of Nicomedes (280–210 BCE), entitled *Introduction to Arithmetic*.

The idea of the sieve is very intuitive. Take a list of numbers $2, 3, \ldots, x$, with x an integer , and start by calling 2 a prime and crossing out all its multiples, the same as 3, then picking the next uncrossed number and repeat. We stop the process at the next uncrossed integer m such that $m \geq \sqrt{x}$. At this stage, all of the uncrossed numbers are prime.

A.M. Legendre (1752–1833) included a modern form of the sieve in his book *La Théorie des Nombres* [Leg30], 1808. In this section, we will describe this form, and we will show how the sieve of Eratosthenes becomes as powerful as Brun's sieve, which we will talk about in the next section, when it is combined with 'Rankin's trick', through some applications.

4.1. **The sieve of Eratosthenes.** Let $P(z) = \prod_{p<z} p$. The sieve of Eratosthenes deletes from the list of numbers all those that are not relatively prime to $P(z)$, except the primes dividing $P(z)$ itself. This motivates us to study the function,

$$\Phi(x, z) = \#\{n \leq x : n \text{ is not divisible by any prime } p < z\}.$$

This will serve as a motivating example, before the formal generalization of the sieve.

Theorem 4.1.

$$\Phi(x, z) = x \prod_{p<z} \left(1 - \frac{1}{p}\right) + O(2^z)$$

Proof. Using lemma 3.1:

$$\phi(x, z) = \sum_{n \leq x} \left(\sum_{d | \gcd(n, P(z))} \mu(d) \right) = \sum_{n \leq x} \left(\sum_{\substack{d|n \\ d|P(z)}} \mu(d) \right)$$

We rearrange the sum, by putting the sum over d dividing $P(z)$ on the outside, and for each of these divisors, we sum over all n less than or equal to x that are divisible by d.

$$\Phi(x, z) = \sum_{d|P(z)} \mu(d) \left(\sum_{\substack{n \leq x \\ d|n}} 1 \right) = \sum_{d|P(z)} \mu(d) \left(\sum_{m \leq \frac{x}{d}} 1 \right)$$

The last step follows by substituting $n = m \cdot d$, so that the sum over all positive integers $n \leq x$ such that n is a multiple of d becomes the sum over all positive integers $m \leq \frac{x}{d}$.

$$\sum_{m \leq \frac{x}{d}} 1 = \left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} + O(1)$$

Then using the fact that $|\mu(d)| \leq 1$ for all d,

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

$$= x \sum_{d|P(z)} \frac{\mu(d)}{d} + O\left(\sum_{d|P(z)} 1 \right)$$

The sum in the error term is to count the number of subsets of the set of primes less than or equal to z, which bounded by 2^z. For the main term:

$$x \sum_{d|P(z)} \frac{\mu(d)}{d} = x \left(1 + \sum_{p|P(z)} \frac{\mu(p)}{p} + \sum_{p_1<p_2|P(z)} \frac{\mu(p_1p_2)}{p_1p_2} + \cdots \right)$$

$$= x \left(1 - \sum_{p|P(z)} \frac{1}{p} + \sum_{p_1<p_2|P(z)} \frac{1}{p_1p_2} - \cdots \right)$$

Observe that in this sum, we either choose a $\frac{1}{p}$ or a $\frac{-1}{p}$, so:

$$x \sum_{d|P(z)} \frac{\mu(d)}{d} = x \prod_{p \leq z} \left(1 - \frac{1}{p} \right)$$

Thus:

$$\phi(x,z) = x \prod_{p<z} \left(1 - \frac{1}{p} \right) + O(2^z)$$

■

4.2. **Rankin's Trick and the Function $\psi(x,z)$.** We want to refine the error term obtained in the previous subsection, and this can be done using a clever idea due to Rankin. Let's introduce the function $\psi(x,z)$:

$$\psi(x,z) = \#\{n \leq x : p \mid n \Rightarrow p \leq z\}$$

In that case, n is called a z-smooth number: only divisible by primes at most z. Recall that:

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

Note that the term $\left\lfloor \frac{x}{d} \right\rfloor$ is only nonzero when $d \leq x$. This implies that we are only summing over divisors of $P(z)$ that are less than or equal to x. Hence:

$$\phi(x,z) = \sum_{\substack{d \leq x \\ d|P(z)}} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor = \sum_{\substack{d \leq x \\ d|P(z)}} \mu(d) \left(\frac{x}{d} + O(1) \right)$$

$$\phi(x,z) = x \sum_{\substack{d \leq x \\ d|P(z)}} \frac{\mu(d)}{d} + O\left(\sum_{\substack{d \leq x \\ d|P(z)}} 1 \right)$$

Thus:

$$\phi(x,z) = x \sum_{\substack{d\le x \\ d|P(z)}} \frac{\mu(d)}{d} + O\left(\psi(x,z)\right) \tag{4.1}$$

In order to show that the error term above is actually better than the exponential error term, we will try to bound $\psi(x,z)$.

Theorem 4.2.

$$\psi(x,z) \ll x(\log z)\exp\left(-\frac{\log x}{\log z}\right)$$

Proof.

Theorem 4.3.

$$\sum_{p<z} \frac{\log p}{p} = \log z + O(1)$$

Proof. Consider the prime factorization of $n!$.

$$n! = \prod_{p\le n} p^{e_p},$$

where e_p is the multiplicity of p. Since only primes $p \le n$ can divide $n!$, we only need to track the product of primes up to n. There are $\left\lfloor \frac{n}{p} \right\rfloor$ multiples of p that are at most n, and $\left\lfloor \frac{n}{p^2} \right\rfloor$ multiples of p^2 that are at most n, and so on. Hence:

$$e_p = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \cdots,$$

Note that the sum is finite since for some power p^a of p we will have $n < p^a$ so that $\left\lfloor \frac{n}{p^a} \right\rfloor = 0$. Therefore:

$$\log n! = \sum_{p\le n} \left(\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \cdots \right) \log p.$$

We also have,

$$\log n! = \sum_{k\le n} \log k = n\log n - n + O(\log n),$$

which can be proved using Abel summation, and

$$\sum_{p\le n} \left(\left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots \right) \log p \le n \sum_{p} \frac{\log p}{p(p-1)}$$

We consider the series:

$$\sum_{p} \frac{\log p}{p(p-1)}$$

where the sum runs over all primes p.

Note that:

$$\frac{\log p}{p(p-1)} = \frac{\log p}{p^2 - p} \sim \frac{\log p}{p^2} \quad \text{as } p \to \infty.$$

But the following series is known to converge:

$$\sum_p \frac{\log p}{p^2} < \infty.$$

Hence, by comparison:

$$\sum_p \frac{\log p}{p(p-1)} < \infty.$$

Hence,

$$\sum_{p \leq n} \left(\left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots \right) \log p \ll n$$

. we get,

$$\sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p = n \log n + O(n).$$

■

Theorem 4.4.

$$\sum_{p<z} \frac{1}{p} = \log \log z + O(1)$$

Proof. We apply partial summation to deduce the result from Theorem 4.3. Setting $c_n := (\log n)/n$ if n is prime and zero otherwise,

$$\sum_{p \leq n} \frac{1}{p} = \sum_{k \leq n} c_k f(k) = C(n) f(n) - \int_2^n C(t) f'(t)\, dt,$$

where $C(t) := \sum_{p \leq t} \frac{\log p}{p} = \log t + O(1)$, by Theorem 4.3, and

$$f'(t) = -\frac{1}{t(\log t)^2}.$$

So,

$$\sum_{p \leq n} \frac{1}{p} = (\log n + O(1)) \cdot \frac{1}{\log n} + \int_2^n (\log t + O(1)) \cdot \frac{1}{t(\log t)^2}\, dt.$$

The first term is $1 + O\left(\frac{1}{\log n}\right)$, and the integral yields $\log \log n + O(1)$, so:

$$\sum_{p \leq n} \frac{1}{p} = \log \log n + O(1).$$

■

We are going to use Rankin's trick to estimate $\psi(x, z)$. For $\delta > 0$

$$\Psi(x, z) = \sum_{\substack{n \leq x \\ p | n \Rightarrow p < z}} 1 \leq \sum_{\substack{n \leq x \\ p | n \Rightarrow p < z}} \left(\frac{x}{n}\right)^\delta \leq x^\delta \prod_{p<z} \left(1 - \frac{1}{p^\delta}\right)^{-1}$$

Hence:

$$\begin{aligned}\psi(x,z) &\leq x^{\delta}\prod_{p<z}\left(1-\frac{1}{p^{\delta}}\right)^{-1}\\ &=\prod_{p<z}\left(1+\frac{1}{p^{\delta}}\right)\left(1-\frac{1}{p^{2\delta}}\right)^{-1}\\ &\ll\prod_{p<z}\left(1+\frac{1}{p^{\delta}}\right),\end{aligned}$$

because $\left(1-\frac{1}{p^{2\delta}}\right)^{-1}$ converges for $\delta > \frac{1}{2}$. Using the inequality $1+x \leq e^x$, we get

$$\begin{aligned}\Psi(x,z) &\ll x^{\delta}\prod_{p<z}\exp\left(\frac{1}{p^{\delta}}\right)\\ &= x^{\delta}\exp\left(\sum_{p<z}\frac{1}{p^{\delta}}\right)\end{aligned}$$

We set $\delta := 1-\eta$, with η "small" enough to ensure $\delta > \frac{1}{2}$ and $\eta \to 0$ as $z \to \infty$. Writing:

$$p^{-\delta} = p^{-1}p^{\eta} = p^{-1}e^{\eta\log p}$$

Note that: $e^x \leq 1 + xe^x$ Hence:

$$\Psi(x,z) \ll x^{1-\eta}\exp\left(\sum_{p<z}\frac{1}{p}(1+\eta\log p\cdot x^{\eta\log p})\right)$$

We choose η sufficiently small, $\eta := \frac{1}{\log z}$, which yields:

$$\Psi(x,z) \ll x^{1-\log z}\exp\left(\sum_{p<z}\frac{1}{p}\left(1+\frac{\log p}{\log z}\cdot e\right)\right)$$

Thus:

$$\Psi(x,z) \ll x\exp\left(-\frac{\log x}{\log z}\right)\left(\sum_{p<z}\frac{1}{p}+\sum_{p<z}\frac{\log p}{p}\cdot\frac{e}{\log z}\right)$$

Using 4.4 and 4.3, we get the following:

$$\begin{aligned}\Psi(x,z) &\ll x\exp\left(-\frac{\log x}{\log z}\right)\left(\log\log z + O(1) + \frac{e}{\log z}\left(\log z + O(1)\right)\right)\\ &\ll x\exp\left(-\frac{\log x}{\log z}\right)\exp(\log\log z)\\ &\ll x(\log z)\exp\left(-\frac{\log x}{\log z}\right)\end{aligned}$$

■

Lemma 4.5. *Let $C(x) = \sum_{n \le x} c_n$, and let $f(t)$ be a differentiable function with continuous derivative. Suppose that*

$$\lim_{Y \to \infty} C(Y)f(Y) = 0 \quad \textit{and} \quad \int_1^\infty |C(t)f'(t)|\, dt < \infty.$$

Then, for all $x \ge 1$,

$$\sum_{n > x} c_n f(n) = -C(x)f(x) - \int_x^\infty C(t)f'(t)\, dt.$$

Proof. By Abel summation, we have

$$\sum_{n \le x} c_n f(n) = C(x)f(x) - \int_1^x C(t)f'(t)\, dt \tag{4.2}$$

Now observe that:

$$\begin{aligned}
\sum_{n > x} c_n f(n) &= \sum_{n=1}^{\infty} c_n f(n) - \sum_{n \le x} c_n f(n) \\
&= \sum_{n=1}^{\infty} c_n f(n) - \left[C(x)f(x) - \int_1^x C(t)f'(t)\, dt \right] \\
&= \sum_{n=1}^{\infty} c_n f(n) - C(x)f(x) + \int_1^x C(t)f'(t)\, dt
\end{aligned}$$

To proceed, it suffices to show that:

$$\sum_{n=1}^{\infty} c_n f(n) = -\int_1^\infty C(t)f'(t)\, dt$$

which follows by sending $x \to \infty$ in equation (4.2). Since:

$$\lim_{x \to \infty} \int_1^x C(t)f'(t)\, dt = \int_1^\infty C(t)f'(t)\, dt < \infty,$$

we conclude that:

$$\sum_{n > x} c_n f(n) = -C(x)f(x) - \int_x^\infty C(t)f'(t)\, dt.$$

∎

Theorem 4.6.

$$\sum_{\substack{d \mid P(z) \\ d \le Z}} \frac{\mu(d)}{d} = \prod_{p < z} \left(1 - \frac{1}{p}\right) + O\left(x(\log x)^2 \exp\left(-\frac{\log x}{\log z}\right)\right).$$

Proof. Note that

$$\sum_{\substack{d \mid P(z) \\ d \le Z}} \frac{\mu(d)}{d} = \sum_{d \mid P(z)} \frac{\mu(d)}{d} - \sum_{\substack{d \mid P(z) \\ d > Z}} \frac{\mu(d)}{d},$$

and

$$\left|\sum_{\substack{d|P(z)\\d>Z}}\frac{\mu(d)}{d}\right|\leq\sum_{\substack{d|P(z)\\d>Z}}\frac{1}{d}.$$

Using Lemma 4.5, this sum is bounded by

$$\sum_{\substack{d|P(z)\\d>Z}}\frac{1}{d}\leq-\frac{\psi(x,z)}{x}+\int_x^\infty\frac{\psi(t,z)}{t^2}\,dt.$$

The integral satisfies the following.

$$\log z\int_x^\infty\exp\left(-\frac{\log t}{\log z}\right)\frac{dt}{t}=\log z\int_x^\infty\frac{dt}{t^{1+\frac{1}{\log z}}}\leq(\log z)^2\exp\left(-\frac{\log x}{\log z}\right),$$

which completes the proof. ∎

Theorem 4.7.

$$\Phi(x,z)=x\prod_{p<z}\left(1-\frac{1}{p}\right)+O\left(x(\log z)^2\exp\left(-\frac{\log x}{\log z}\right)\right),\quad as\ \frac{\log x}{\log z}\to\infty.$$

Proof. The proof follows from recalling 4.1 and combining 4.2 and 4.6. ∎

4.3. **The generalized sieve of Eratosthenes.** We can now introduce the formal generalized form of the sieve of Eratosthenes. Let $\mathcal{A}$ be any set of natural numbers $\leq x$ and let $\mathcal{P}$ be a set of primes. To each prime $p\in\mathcal{P}$, associate $\omega(p)$ distinguished residue classes modulo p. Let $\mathcal{A}_p$ denote the subset of elements in $\mathcal{A}$ that belong to at least one of the residue classes modulo p. Set $\mathcal{A}_1:=\mathcal{A}$, and for any squarefree integer d composed only of primes from $\mathcal{P}$, define:

$$\mathcal{A}_d:=\bigcap_{p|d}\mathcal{A}_p$$

and

$$\omega(d):=\prod_{p|d}\omega(p).$$

Let z be a positive real number and define:

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

We denote by $S(\mathcal{A},\mathcal{P},z)$ the number of elements of the sifted set:

$$\mathcal{A}\setminus\bigcup_{p|P(z)}\mathcal{A}_p.$$

We assume that there exists a constant X such that for every squarefree d composed of primes from $\mathcal{P}$, the cardinality of $\mathcal{A}_d$ satisfies:

$$\#\mathcal{A}_d=\frac{\omega(d)}{d}X+R_d\tag{5.3}$$

for some remainder term R_d.

Theorem 4.8. *The sieve of Eratosthenes*
In the above setting, suppose the following conditions hold,

(1) $|R_d| = O(\omega(d))$;

(2) *for some* $\kappa \geq 0$,

$$\sum_{p|P(z)} \frac{\omega(p)\log p}{p} \leq \kappa \log z + O(1);$$

(3) *for some positive real number* y, $\#\mathcal{A}_d = 0$ *for every* $d > y$.

Then

$$S(\mathcal{A},\mathcal{P},z) = XW(z) + O\left(\left(X + \frac{y}{\log z}\right)(\log z)^{\kappa+1}\exp\left(-\frac{\log y}{\log z}\right)\right),$$

where

$$W(z) := \prod_{p\in\mathcal{P},\ p<z}\left(1 - \frac{\omega(p)}{p}\right).$$

Proof.

Lemma 4.9. *Under the assumptions of 4.8, define*

$$F(t,z) := \sum_{\substack{d\leq t\\ d|P(z)}} \omega(d).$$

Then we have

$$F(t,z) = O\left(t(\log z)^{\kappa}\exp\left(-\frac{\log t}{\log z}\right)\right).$$

Proof. We apply Rankin's trick. For any $\delta > 0$,

$$F(t,z) \leq \sum_{d|P(z)} \omega(d)\left(\frac{t}{d}\right)^{\delta}.$$

Since ω is multiplicative, the sum can be factorized, and we obtain

$$F(t,z) \leq t^{\delta}\prod_{p|P(z)}\left(1 + \frac{\omega(p)}{p^{\delta}} + \frac{\omega(p^2)}{p^{2\delta}} + \cdots\right).$$

Using the inequality $1 + x \leq e^x$ and the fact that the Dirichlet series converges, we deduce

$$F(t,z) \leq \exp\left(\delta\log t + \sum_{p|P(z)}\frac{\omega(p)}{p^{\delta}}\right).$$

Now set $\delta = 1-\theta$ for some small $\theta > 0$. Using the inequality $e^x \leq 1 + xe^x$, we find

$$F(t,z) \leq t\exp\left(-\theta\log t + \sum_{p\leq z}\frac{\omega(p)}{p} + \theta z^{\theta}\sum_{p\leq z}\frac{\omega(p)\log p}{p}\right).$$

From the second assumption in Theorem 5.4.1 and partial summation, we have

$$\sum_{p\leq z}\frac{\omega(p)}{p} \leq \kappa\log\log z + O(1).$$

So we get

$$F(t,z) \ll t\exp\left(-\theta\log t + \kappa\log\log z + \kappa\theta(\log z)z^{\theta}\right).$$

Choosing $\theta = 1/\log z$ yields the desired estimate:

$$F(t,z) = O\left(t(\log z)^{\kappa}\exp\left(-\frac{\log t}{\log z}\right)\right).$$

■

Lemma 4.10. *Under the same assumptions,*

$$\sum_{\substack{d>y\\ d|P(z)}} \frac{\omega(d)}{d} = O\left((\log z)^{\kappa+1}\exp\left(-\frac{\log y}{\log z}\right)\right).$$

Proof. By partial summation,

$$\sum_{\substack{d>y\\ d|P(z)}} \frac{\omega(d)}{d} \ll \int_y^{\infty} \frac{F(t,z)}{t^2}\,dt.$$

Now apply 4.9 to estimate $F(t,z)$ and conclude the result. ■

Using the inclusion–exclusion principle and the first and third hypotheses of the theorem, we write:

$$S(\mathcal{A},\mathcal{P},z) = \sum_{\substack{d\le y\\ d|P(z)}} \mu(d)\frac{|\mathcal{A}_d|}{d} = \sum_{\substack{d\le y\\ d|P(z)}} \mu(d)\frac{X\omega(d)}{d} + O\left(F(y,z)\right).$$

Then by 4.9 and 4.10, we deduce

$$S(\mathcal{A},\mathcal{P},z) = XW(z) + O\left(\left(X + \frac{y}{\log z}\right)(\log z)^{\kappa+1}\exp\left(-\frac{\log y}{\log z}\right)\right).$$

■

4.4. **Application of the Sieve of Eratosthenes.** A twin prime is a prime number p such that $p+2$ is also prime. Viggo Brun proved that $\sum_{p,p+2\text{prime}} \frac{1}{p} < \infty$ in his 1919 paper see [HR13], using the sophisticated Brun's sieve which we will discuss in the next section. However, the interesting thing is that the same result can be derived using the elementary sieve of Eratosthenes. The following upper bound is not the best that we can find, but it is enough to prove that the series mentioned above converges.

Theorem 4.11. *The number of primes $p \le x$ such that $p+2$ is prime is*

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

Proof. Let $\mathcal{A}$ be the set of natural numbers n less than or equal to x and let $\mathcal{P}$ be the set of all primes. Let z be a positive real number, to be chosen soon. For each prime $p < z$, we distinguish the residue classes 0 and -2 modulo p. Since A_p (the set of n less than or equal to x belonging to at least one of these residue classes) is empty for $p > x+2$, we apply the sieve of Eratosthenes $\kappa = 2$ to deduce that

$$S(\mathcal{A},\mathcal{P},z) = xW(z) + O\left(x(\log z)^3\exp\left(-\frac{\log x}{\log z}\right)\right),$$

where

$$W(z) := \prod_{p<z}\left(1 - \frac{2}{p}\right).$$

Now

$$W(z) = \prod_{p<z}\left(1 - \frac{2}{p}\right) \le \exp\left(-\sum_{p<z}\frac{2}{p}\right) \ll (\log z)^{-2}.$$

We choose z such that

$$\log z = \log x / A \log\log x$$

for some large positive constant A and deduce

$$S(\mathcal{A}, \mathcal{P}, z) \ll \frac{x(\log\log x)^2}{\log^2 x}.$$

Thus,the number of primes does not exceed:

$$\pi(z) + S(\mathcal{A}, \mathcal{P}, z) \le z + S(\mathcal{A}, \mathcal{P}, z),$$

with z as above. The result follows directly. ■

Theorem 4.12. *(Brun's theorem) The sum,*

$$\sum_{\substack{p \\ p+2 \ prime}} \frac{1}{p}$$

converges.

Proof. By partial summation and Theorem 4.11, the sum is bounded by

$$\sum_{\substack{p \\ p+2 \text{ prime}}} \frac{1}{p} \ll \int_2^\infty \frac{\pi_2(t)}{t^2} dt \ll \int_2^\infty \frac{(\log\log t)^2}{t \log^2 t}\, dt,$$

which is finite. ■

5. **Brun's Pure Sieve**

In 1915, Viggo Brun (1885–1978) introduced the sieve method that now bears his name, as documented in his paper [Bru15]. Prior to Brun's contribution, Jean Merlin [Mer11] had made what is considered the first serious attempt to improve upon the sieve of Eratosthenes. Merlin was killed during World war I (refer to [HR13]), leaving only two of his manuscripts, which was prepared for publication posthumously by Jacques Hadamard (1865–1963). Brun reportedly studied Merlin's work closely and found inspiration in it. This influence likely led to his seminal 1915 paper, which later evolved into a more refined sieving method [Bru20].

In his groundbreaking research, Brun proved that there exist infinitely many integers n such that both $n+2$ have no more than nine prime factors. He also showed that any sufficiently large even integer can be expressed as the sum of two integers, each having at most nine prime factors. These results constitute relevant progress towards solving the twin prime conjecture and Goldbach's conjecture.As mentioned in the introduction, one of his notable conclusions was that the sum of the reciprocals of the twin primes is finite.

Despite his groundbreaking work, Brun's original papers received little attention at first. There is a story that Edmund Landau (1877–1938) left the papers unread on his desk for eight years, and a share of this neglect could partly be explained by Brun's use of complicated notation.

As previously mentioned, some of Brun's early results can be replicated using the sieve of Eratosthenes combined with Rankin's trick, though this was not realized until the publication of [MS87]. Nevertheless, Brun's later developments in sieve theory go well beyond basic techniques and are essential to the modern study of sieve methods.

5.1. **Starting idea.**

In this section, we will detail Brun's starting idea for the development of Brun's pure sieve. Observe that by comparing the coefficients of x^r on both sides of the identity

$$(1-x)^{-1}(1-x)^{\nu} = (1-x)^{\nu-1}.$$

We obtain that for any positive integers ν and r such that $0 \leq r \leq \nu - 1$,

$$\sum_{k \leq r} (-1)^k \binom{\nu/k}{k} = (-1)^r \binom{\nu-1}{r}.$$

Let n be a positive integer and let N be the *radical* of n. Using the above formula with $\nu = \nu(n)$ to deduce that, for any $0 \leq r \leq \nu(n) - 1$,

$$\sum_{\substack{d|n \\ \nu(d) \leq r}} \mu(d) = \sum_{\substack{d|N \\ \nu(d) \leq r}} \mu(d) = \sum_{k \leq r} (-1)^k \binom{\nu(n)}{r} = (-1)^r \binom{\nu(n)-1}{r}.$$

Let's introduce truncated Möbius function of d by

$$\mu_r(d) := \begin{cases} \mu(d) & \text{if } \nu(d) \leq r, \\ 0 & \text{if } \nu(d) > r, \end{cases}$$

and let us set

$$\psi_r(n) := \sum_{d|n} \mu_r(d).$$

Rewriting 5.1:

$$\psi_r(n) = (-1)^r \binom{\nu(n)-1}{r},$$

The above formula could be seen as a generalization of the fundamental property of the Möbius function.

From 5.1 we see that $\psi_r(n) \geq 0$ if r is even, and $\psi_r(n) \leq 0$ if r is odd. Hence for any positive integers n and r,

$$\psi_{2r+1}(n) \leq \sum_{d|n} \mu(d) \leq \psi_{2r}(n).$$

Notice that,

$$\psi_{2r+1}(n) = \sum_{\substack{d|n \\ r(d)\leq r}} \mu(d) + \sum_{\substack{d|n \\ r(d)=2r+1}} \mu(d) = \psi_r(n) + O\left(\sum_{\substack{d|n \\ r(d)=2r+1}} |\mu(d)| \right).$$

Therefore, for any positive integers n and r,

$$\sum_{d|n} \mu(d) = \psi_r(n) + O\left(\sum_{\substack{d|n \\ r(d)=2r+1}} |\mu(d)| \right).$$

Brun had the clever idea to use $\psi_r(n)$, through 5.1 in the sieve of Eratosthenes in order to improve the error terms. We will try to apply Brun's idea to obtain an upper bound for $\Phi(x,z)$. Set

$$P(z) := \prod_{p<z} p.$$

Then by 5.1, for r even,

$$\Phi(x,z) \leq \sum_{n\leq x} \sum_{\substack{d|n \\ d|P_z}} \mu_r(d) = \sum_{d|P_z} \mu_r(d) \left\lfloor \frac{x}{d} \right\rfloor = x \sum_{d|P_z} \frac{\mu_r(d)}{d} + O(z^r),$$

since $\mu_r(d) = 0$ unless $r(d) \leq r$.
Observe that by Möbius inversion,

$$\mu_r(d) = \sum_{\delta|d} \mu\left(\frac{d}{\delta}\right) \psi_r(\delta).$$

Hence,

$$\sum_{d|P_z} \frac{\mu_r(d)}{d} = \sum_{\delta|P_z} \sum_{\substack{d|P_z \\ \delta|d}} \mu\left(\frac{d}{\delta}\right) \frac{\psi_r(\delta)}{d} = \sum_{\delta|P_z} \psi_r(\delta) \sum_{\substack{d|P_z \\ \delta|d}} \frac{\mu(d/\delta)}{d} = \sum_{\delta|P_z} \frac{\psi_r(\delta)}{\delta} \sum_{f|P_z} \frac{\mu(f)}{f} = W(z) \sum_{\delta|P_z} \frac{\psi_r(\delta)}{\delta},$$

with

$$W(z) := \prod_{p<z} \left(1 - \frac{1}{p}\right)$$

and ϕ denotes the Euler function.

$$\sum_{d|P_z} \frac{\mu_r(d)}{d} = W(z) + W(z) \sum_{\substack{\delta|P_z \\ \delta>1}} \frac{\psi_r(\delta)}{\delta\phi(\delta)}.$$

Notice that from 5.1,

$$\psi_r(\delta) \leq \binom{\nu(\delta)-1}{r},$$

Therefore,

$$\begin{aligned}
\sum_{\substack{\delta|P_z \\ \delta>1}} \frac{\psi_r(\delta)}{\delta\phi(\delta)} &\leq \sum_{\substack{\delta|P_z \\ \delta>1}} \frac{1}{\phi(\delta)} \binom{\nu(\delta)-1}{r} \\
&\leq \sum_{r+1\leq m\leq \pi(z)} \binom{m-1}{r} \sum_{\substack{\delta|P_z \\ \nu(\delta)=m}} \frac{1}{\phi(\delta)} \\
&\leq \sum_{r+1\leq m\leq \pi(z)} \binom{m-1}{r} \frac{1}{m!} \left(\sum_{p<z} \frac{1}{p-1} \right)^m \\
&\leq \frac{1}{r!} \sum_{n\geq r+1} \frac{(n-1)^r}{n!} (\log\log z + c_1)^n \\
&= \frac{(\log\log z + c_1)^r}{r!} \sum_{n\geq r+1} \frac{1}{n} (\log\log z + c_1)^{n-r} \\
&\leq \frac{(\log\log z + c_1)^r}{r!} \exp(\log\log z + c_1).
\end{aligned}$$

Where we have used the elementary estimate,

$$\sum_{p<z} \frac{1}{p} \leq \log\log z + c_1$$

for some positive constant c_1. C Thus,

$$\sum_{\substack{\delta|P_z \\ \delta>1}} \frac{\psi_r(\delta)}{\delta\phi(\delta)} \leq \frac{(\log\log z + c_1)^r}{r!} \exp(\log\log z + c_1).$$

We use the well-known bound:

$$\frac{1}{r!} \leq \left(\frac{e}{r}\right)^r$$

to further simplify 5.1. Hence,

$$\sum_{\substack{\delta|P_z \\ \delta>1}} \frac{\psi_r(\delta)}{\delta\phi(\delta)} \leq c_2 \exp\left(r - r\log r + r\log\Lambda\right) \log z,$$

where

$$\Lambda := \log\log z + c_1$$

and c_2 is a positive constant.

Combining estimates 5.1, ‘5.1, and 5.1, we derive an upper bound for $\Phi(x,z)$:

$$\Phi(x,z) \le xW(z) + xW(z)\cdot O\left(\exp(r - r\log r + r\log\Lambda)\log z\right) + O(z).$$

We want to minimize the exponential error term. To do this, we let:

$$r = \eta\frac{\log z}{\log\log z}$$

for some function $\eta = \eta(x,z)$, to be determined. Plugging this choice of r into 5.1,

$$x\exp\left(-\eta(\log\eta)(\log\log z) + 2\eta\log\log z\right),$$

To balance the exponential terms and minimize the error, we set:

$$\eta = \frac{\alpha\log x}{(\log z)(\log\log z)} \quad \text{for some } \alpha < 1.$$

for such a choice of r, it follows that:

$$\log z = O((\log x)^{1-\epsilon})$$

for some $\epsilon > 0$. For any $0 < \varepsilon < 1$,

$$\Phi(x,z) \le xW(z) + O\left(x\exp\left(-(\log x)^{\varepsilon}\right)\right).$$

After this introduction , we are now prepared to formally present the pure sieve. We begin with the following key identity:

Lemma 5.1. *Let n, r be positive integers with $r \le \nu(n)$. Then there exists $|\theta| \le 1$ such that*

$$\sum_{d|n}\mu(d) = \sum_{\substack{d|n\\ \nu(d)\le r}}\mu(d) + \theta\sum_{\substack{d|n\\ \nu(d)=r+1}}\mu(d).$$

Proof. This can be easily derived from the combinatorial identities stated at the beginning of this section. ■

5.2. **Brun's pure sieve.** Let $\mathcal{A}$ be any set of natural numbers $\le x$, and let $\mathcal{P}$ be a set of primes. For each prime $p \in \mathcal{P}$, let $\mathcal{A}_p$ be the set of elements of $\mathcal{A}$ that are divisible by p.

Let $\mathcal{A}_1 := \mathcal{A}$, and for any squarefree positive integer d composed of primes of $\mathcal{P}$, let

$$\mathcal{A}_d := \bigcap_{p|d}\mathcal{A}_p.$$

Let z be a positive real number, and define

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

We want to estimate

$$S(\mathcal{A},\mathcal{P},z) := \#\left(\mathcal{A}\setminus\bigcup_{p|P(z)}\mathcal{A}_p\right).$$

We suppose that there exists a multiplicative function $\omega(\cdot)$ such that, for any d as above,

$$\#\mathcal{A}_d = \frac{\omega(d)}{d}X + R_d$$

for some remainder term R_d, where

$$X := \#\mathcal{A}.$$

Theorem 5.2 (Brun's pure sieve)**.** *We keep the above setting and we make the additional assumptions that:*

(1) *$|R_d| \leq \omega(d)$ for any squarefree d composed of primes of $\mathcal{P}$;*
(2) *there exists a positive constant C such that $\omega(p) < C$ for any $p \in \mathcal{P}$;*
(3) *there exist positive constants C_1, C_2 such that*

$$\sum_{\substack{p<z \\ p\in\mathcal{P}}} \frac{\omega(p)}{p} < C_1 \log\log z + C_2.$$

Then

$$S(\mathcal{A},\mathcal{P},z) = XW(z)\left(1 + O\left((\log z)^{-A}\right)\right) + O\left(z^{\eta \log\log z}\right)$$

with $A = \eta \log \eta$. In particular, if $\log z \leq c \log x / \log\log x$ for a suitable positive constant c sufficiently small, we obtain

$$S(\mathcal{A},\mathcal{P},z) = XW(z)(1 + o(1)).$$

Proof. By 5.1, for any positive integer r, we have

$$S(\mathcal{A},\mathcal{P},z) = \sum_{a\in\mathcal{A}} \sum_{d|(a,P(z))} \mu(d).$$

Using the identity $\mu(d) = \mu_r(d) + \theta \sum_{\nu(d)=r+1} \mu(d)$, this becomes

$$= \sum_{a\in\mathcal{A}} \left(\sum_{d|(a,P(z))} \mu_r(d) + \theta \sum_{\substack{d|(a,P(z)) \\ \nu(d)=r+1}} \mu(d) \right).$$

Grouping terms by d, we obtain

$$= \sum_{d|P(z)} \mu_r(d)\#\mathcal{A}_d + O\left(X \frac{\pi(z)^{r+1}}{(r+1)!} \right),$$

where $\#\mathcal{A}_d$ is the number of elements $a \in \mathcal{A}$ divisible by d.

Now using 5.2, along with the first hypothesis and the multiplicativity of $\omega(\cdot)$, we find:

$$S(\mathcal{A},\mathcal{P},z) = X \sum_{d|P(z)} \frac{\mu_r(d)\omega(d)}{d} + O\left(\sum_{\substack{d|P(z) \\ \nu(d)\leq r}} |R_d| \right) + O\left(X \frac{z^{r+1}}{(r+1)!} \right).$$

By using the bound $|R_d| \leq \omega(d)$ from the first assumption, this simplifies further to:

$$= X \sum_{d|P(z)} \frac{\mu_r(d)\omega(d)}{d} + O\left(\left(1 + \sum_{p<z} \omega(p) \right) \frac{1}{r!} \right) + O\left(X \frac{z^{r+1}}{(r+1)!} \right).$$

Then, by once again applying the Möbius inversion formula, we deduce:

$$S(\mathcal{A},\mathcal{P},z) = X\sum_{\delta|P(z)}\frac{\psi_r(\delta)\omega(\delta)}{\delta}\sum_{d|\frac{P(z)}{\delta}}\frac{\mu(d)\omega(d)}{d}+O\left(\left(1+\sum_{p<z}\omega(p)\right)^r\frac{1}{r!}\right)+O\left(X\frac{z^{r+1}}{(r+1)!}\right).$$

Now, we introduce the function

$$\Omega(d) := \prod_{p|d}(p-\omega(p)).$$

With this definition, the first sum in the previous expression simplifies to

$$XW(z)\sum_{\delta|P(z)}\frac{\psi_r(\delta)\omega(\delta)}{\Omega(\delta)}.$$

Finally, using the first and second assumptions of the theorem, we arrive at the desired asymptotic formula. ■

6. Acknowledgments

I would like to thank my TA, Benjamin Vakil for helpful critiques and Simon Rubinstein-Salzedo for this class and for help in writing and organizing this paper.

References

[Bru15] Viggo Brun. Uber das goldbachsche gesetz und die anzahl der primzahlpaare. *Arch. Mat. Natur. B*, 1915.

[Bru16] Viggo Brun. Om fordelingen av primtallene i forskjellige talklasser. en övre begraensning. *Nyt tidsskrift for matematik*, 27:45–58, 1916.

[Bru20] Viggo Brun. Le crible d'eratosthène et le théorème de goldbach. *Videnskaps. Skr., Mat. Natur. Kl. Kristiana*, 1920.

[Bru22] Viggo Brun. Das siev des eratosthenes, 5. *Skand. Mat. Knogr., Helsingfors*, pages 197–203, 1922.

[Che02] Jing-run Chen. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. In *The Goldbach Conjecture*, pages 275–294. World Scientific, 2002.

[HR13] Heine Halberstam and Hans Egon Richert. *Sieve methods*. Courier Corporation, 2013.

[Leg30] Adrien Marie Legendre. *Théorie des nombres*, volume 2. Firmin Didot Frères, 1830.

[Mer11] J Merlin. Sur quelques théorèmes d'arithmétique et un énoncé qui les contient. *CR Acad. Sci. Paris*, 153:516–518, 1911.

[MS87] M Ram Murty and S Saradha. On the sieve of eratosthenes. *Canadian Journal of Mathematics*, 39(5):1107–1122, 1987.

[Pol14] DHJ Polymath. Variants of the selberg sieve, and bounded intervals containing many primes. *Research in the Mathematical sciences*, 1:1–83, 2014.