Introduction to Sieves

Johnny Chen

March 2024

1 Introduction

In the previous weeks, we've been focused on representations of integers as sums of k-th powers, which culminated in the Hardy-Littlewood Asymptotic Formula. We will now shift gears into Sieve Methods, and particularly focus on prime numbers and the various interesting properties they hold. This talk will serve as an introduction for sieves and will focus mainly on the sieve of Eratosthenes, a simple and effective method for finding all primes up to a certain bound x.

Broadly speaking, imagine you were looking at a list of numbers, $2, 3, \cdots, \lfloor x \rfloor$. We will note that 2 is a prime, and then cross out all multiples of 2, and then move on to 3, which isn't crossed out, and so it must be a prime as well. We then cross out all the multiples of 3, and continue with this process until we arrive at a number greater than $\sqrt{x}$ before we stop. We will see that all the numbers not crossed out are primes. Legendre realized that this can be captured in a "sifting" process, and used this to study $\pi(x) = |\{p \leq x \mid p \text{ is prime}\}|$.

The motivation of this talk then is to apply this simple technique to study some more complex problems. First, we will try to estimate $\pi(x)$, then use it to bound k-smooth numbers, estimate "squarefree" numbers and "twin squarefree numbers", and then finally prove the general sieving problem.

2 The Sieve of Eratosthenes-Legendre

In this section, we will explore the sieve of Eratosthenes and find a bound for the number of primes within a given set. As a preliminary, let $P_z = \prod_{p<z} p$. The sieve of Eratosthenes essentially deletes from a list of numbers all of the ones that are not relatively prime to P_z, except the primes that divide P_z. Thus, we can restate our task here as finding the bounds on the cardinality of the set $S = \{n \mid n \leq x, (n, P_z) = 1\}$. We can then define

$$s(n) = \begin{cases} 1 & n \in S \\ 0 & \text{otherwise.} \end{cases}$$

as the characteristic function of S. We can then use the properties of the Mobius function (see appendix; the idea is that the Mobius function $\mu(n)$ returns $(-1)^k$ if n can be expressed as the product of k distinct primes, and 0 otherwise) to write an explicit expression for $s(n)$, which we will call the *sifting function*:

$$s(n) = \sum_{d \mid \gcd(n, P_z)} \mu(d)$$

Through some simple algebra, we can find an expression for $|S|$:

$$\begin{aligned}
|S| &= \sum_{n} s(n) \\
&= \sum_{n\le x} \sum_{d|\ \gcd(n,P_z)} \mu(d) \\
&= \sum_{d|P_z} \mu(d) \left(\sum_{\substack{n\le x \\ d|n}} 1 \right) \\
&= \sum_{d|P_z} \mu(d) \lfloor \frac{x}{d} \rfloor \\
&= \sum_{d|P_z} \mu(d) \left(\frac{x}{d} + \lfloor \frac{x}{d} \rfloor - \frac{x}{d} \right) \\
&= \sum_{d|P_z} \mu(d) \frac{x}{d} + \sum_{d|P_z} \mu(d) \left(\lfloor \frac{x}{d} \rfloor - \frac{x}{d} \right).
\end{aligned}$$

It's clear that the terms of the second sum are at most 1, and we can rewrite the first sum using trick of algebra (included in appendix) which allows us to rewrite

$$\begin{aligned}
|S| &\le x \sum_{d|P_z} \frac{\mu(d)}{d} + O(2^{\pi(z)}) \\
&= x \prod_{p|P_z} \left(1 - \frac{1}{p} \right) + O(2^{\pi(z)}).
\end{aligned}$$

By another number-theoretic application (Merten's Theorem),

$$\prod_{p<z} \left(1 - \frac{1}{p} \right) \sim \frac{e^{\gamma}}{\ln z}$$

which yields

$$|S| \le x \frac{e^{\gamma}}{\ln z} + O(2^{\pi(z)})$$

as $z, x \to \infty$.

It may seem that the error term $2^{\pi(z)}$ is huge, and if we take $z = O(\ln x)$, then we can get

$$\pi(x) - \pi(\ln x) = O\left(\frac{x}{\ln \ln x} \right).$$

With $\pi(x) \le x$, we can deduce that

$$\pi(x) = O\left(\frac{x}{\ln \ln x} \right).$$

To back up a bit, what we're doing here is finding the the "density" of primes. The conclusion above is clearly not as good as the well-known "Prime-Number Theorem," which tells us that $\pi(x) \sim \frac{x}{\ln x}$, but this still tells us something meaningful: that the density of primes is 0 for large x, i.e.

$$\frac{\pi(x)}{x} \to 0 \text{ as } x \to \infty$$

3 k-Smooth Numbers

First, let us define k-smooth numbers: a number n is k-smooth if

$$\forall p : (p|n) \Rightarrow (p < k).$$

What this means intuitively is that a number is k-smooth if all of its prime factors are less than k (some texts use $\leq k$, but we will stick with strictly less than). We can then define a function

$$\Psi(x, k) = |\{n \leq x | n \text{ is } k\text{-smooth}\},$$

or, intuitively, the number of k-smooth numbers up to and including x. We will be using our sieve argument to find a bound on Ψ.

Theorem 1 (k-smooth Numbers):

$$\Psi(x, k) = O\left(x\frac{\ln k}{\ln x} + 2^{\pi(x)-\pi(k)}\right)$$

Proof: We know a number is k-smooth if all its prime divisors are below $k \Rightarrow$ we can find the k-smooth numbers below a bound x by using our sifting set $P = \{p \mid k < p \leq x\}$. We can then define $P_{k,x} = \prod_{p \in P} p$. Essentially, our set P contains all the primes between k and x, which can be done using a similar sieve as before, filtering out all numbers that aren't k-smooth by removing numbers that are divisible by a prime larger than k. By taking the product of these numbers, we have a number that has as its factors all primes that we don't want to appear in the k-smooth numbers.
We now define $S = \{n \mid nisk\text{-smooth}\}$, and a new characteristic function

$$s(n) = \begin{cases} 1 & \text{if } n \in S \text{ or } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Following the work we did in the previous section, we can rewrite $s(n)$ with the Möbius function:

$$s(n) = \sum_{d | \gcd(n, P_{k,x})} \mu(d)$$

Through some algebra, we can deduce

$$\begin{aligned} S(n) = |S| &= \sum_{n \leq x} s(n) \\ &= \sum_{n \leq x} \sum_{d | \gcd(n, P_{k,x}} \mu(d) \\ &= \sum_{d | P_{k,x}} \mu(d) \left(\sum_{\substack{n \leq x \\ d | n}} 1 \right) \\ &= \sum_{d | P_{k,x}} \mu(d) \lfloor \frac{x}{d} \rfloor \\ &= x \prod_{k < p \leq x} \left(1 - \frac{1}{p}\right) + O(2^{\pi(x)-\pi(k)}) \\ &= O\left(x\frac{\ln k}{\ln x} + 2^{\pi(x)-\pi(k)}\right), \end{aligned}$$

which shows that the k-smooth numbers are bounded. □

4 Squarefree Numbers

In this section, we will be looking at the density of squarefree numbers. Squarefree numbers are ones that can be expressed as a product of unique primes (i.e. no squares of primes are present in its prime decomposition). Let $S = \{n \mid n \leq x, n \text{ is squarefree}\}$, and let $\kappa(x) = |S|$, the number of squarefree numbers smaller than or equal to x. To find S, we can just take the prime $p \leq \sqrt{x}$ and cross off multiples of p^2, using a variant of $s(n)$

$$s(n) = |\mu(n)|$$

Theorem 2 (Squarefree Numbers:

$$\kappa(x) = \frac{6}{\pi^2}x + O(\sqrt{x})$$

Proof: We can take

$$\begin{aligned}\kappa(x) &= \sum_{n \leq x} s(n) \\ &= \sum_{n \leq x} |\mu(n)|\end{aligned}$$

At this point, we shall prove a lemma that will allow us to continue:

Lemma: $s(n) = \sum_{d^2|n} \mu(d)$

Proof of lemma:

- If n is square-free, then the only divisor d of n for which d^2 also divides n is $d = 1$. That's because a square-free number has no prime squares in its factorization. Since $\mu(1) = 1$, the sum $\sum_{d^2|n} \mu(d)$ will be 1.
- If n is not square-free, it means there is some prime p such that p^2 divides n. For any divisor d of n that is not divisible by p^2, the Möbius function $\mu(d)$ will be either -1 or 1 (depending on the number of distinct prime factors of d). However, for any divisor d that is divisible by p^2, the Möbius function $\mu(d)$ will be 0, because μ is 0 for any number with a squared prime factor.

So, if n has any squared prime factors, there will be at least one term in the sum for which $\mu(d)$ is 0, and the properties of the Möbius function ensure that the sum of $\mu(d)$ over all divisors d of n for which d^2 divides n will cancel out to 0. This is because the Möbius function is designed to "detect" the presence of square factors in a number.
Now, we can set $m = \sqrt{x}$:

$$\kappa(x) = \sum_{n \leq x} s(n)$$

$$= \sum_{n \leq x} \sum_{d^2|n} \mu(d)$$

$$= \sum_{d \leq m} \mu(d) \sum_{d^2|n} 1$$

$$= \sum_{d \leq m} \mu(d) \left\lfloor \frac{x}{d^2} \right\rfloor$$

$$= x \sum_{d \leq m} \frac{\mu(d)}{d^2} + \sum_{d \leq m} \mu(d) \left(\left\lfloor \frac{x}{d^2} \right\rfloor - \frac{x}{d^2} \right)$$

$$= x \sum_{d \leq m} \frac{\mu(d)}{d^2} + O(m).$$

The algebra here is notably similar to the previous proofs. We will also use the fact that

$$\sum_{n\geq 1}\frac{\mu(n)}{n^2}=\prod_p\left(1-\frac{1}{p^2}\right)$$

and

$$\prod_p\left(1-\frac{1}{p^2}\right)=\frac{1}{\zeta(2)}$$

to show that, continuing from where we left off:

$$\kappa(x)=x\prod_p\left(1-\frac{1}{p^2}\right)-\sum_{d>m}\frac{\mu(d)}{d^2}+O(m)$$

But the second term is negligible, as the d^2 quickly overwhelms the Möbius function for large d, so we have

$$\kappa(x)=x\frac{1}{\zeta(2)}+O(\sqrt{x})\quad\Box$$

There is indeed a more generalized form of this theorem, but the proof is quite similar to the one above; we will only state the theorem here. Let $\kappa(x;a,l)=|\{n\leq x\mid n\text{ is squarefree, }n\equiv a\mod l\}|$

Theorem for the density of squarefree numbers in an arithmetic progression:
Let $q>2$ be prime, and let a be a positive integer coprime to q. Then there exists a constant $c>0$ depending only on q such that

$$\kappa(x;a,q)\geq cx+O(\sqrt{x})$$

More specifically,

$$\kappa(x;a,q)\geq x\left(\frac{1}{q\zeta(2)}-\frac{\zeta(2)}{q^2}\right)+O(\sqrt{x}).$$

Before moving on to the next section, we can go over a quick note regarding the error term. The full proof is technical and unnecessary for our objective, so we will just state the conclusion and leave it to the reader to look at the full notes for proof:

Lemma (Error Term): Assuming the Riemann Hypothesis,

$$\kappa(x)=\frac{x}{\zeta(2)}+O(x^{\frac{1}{3}+\delta})$$

5 Twin Squarefree Numbers

There is a famous conjecture of whether there are infinite twin primes, i.e. primes p such that $p+2$ is also prime. We will be going over this in next week's class, but for now, we will show that the result for squarefree numbers follows relatively easily using our past methods.
Let $\kappa_2(x)=|\{n(n+2)|\mu(n)^2=\mu(n+2)^2=1,n\leq x\}|$, the count of numbers n where both n and $n+2$ are both squarefree.

Theorem (Pairs of Squarefree Numbers):

$$\kappa_2(x)=\prod_p\left(1-\frac{2}{p^2}\right)x+O(x^{\frac{2}{3}}\ln^{\frac{4}{3}}x).$$

The full proof is highly technical, and will be left to the reader. The general sketch of it is as follows:

1. Define $s(n)=\sum_{d^2|n}\mu(d)$ to represent $\kappa_2(x)$ as a Möbius sum of $\mu(a)$ and $\mu(b)$ with $a^2|n$ and $b^2|(n+2)$.

2. Express $\kappa_2(x)$ as a double sum over the Möbius functions $\mu(a)$ and $\mu(b)$, representing n and $n+2$ as k_1a^2 and k_2b^2 respectively, which works because $(a,b)=1$

3. We then bound the terms in the sum, noting that for terms where $ab > y$ for a certain threshold y, the sum over $\mu(a)\mu(b)$ tends to zero, as the Möbius function oscillates and tends to cancel out larger products. For the other terms, we use estimates on the number of solutions to a Diophantine equation

4. We can then use number-theoretic results, such as the Euler product formula, to bound the solutions to the Diophantine equation, and derive an asymptotic estimate for $\kappa_2(x)$.

Put simply, we use the previously proven lemma $s(n) = \sum_{d^2|n}\mu(d)$ applied to $\mu(a)$ and $\mu(b)$, where a, b are defined as $n = ka^2$ and $n+2 = kb^2$ and coprime. We then bound this sum because the Möbius function leads to significant cancellation for large products, and for the remaining terms, we can use a mix of number-theoretic results, such as Euler's product formula to derive this asymptotic formula.

6 The General Sieve Problem

We will now use the techniques we have seen to formulate the sieve problem in a more generic setting. Before getting into the theorem, we will introduce some more notation:

1. $\mathcal{A}, \mathcal{B}, \cdots$ will stand for integer sequences

2. $\mathcal{A}_d = \langle a \in \mathcal{A} : a \equiv 0 \mod d\rangle$, a congruence subsequence of integers in $\mathcal{A}$ that are divisible by d

3. $\mathcal{A}^z = \langle a \in \mathcal{A} : a \leq z\rangle$, an integer subsequence of $\mathcal{A}$ of all integers in $\mathcal{A}$ less than or equal to z

4. If $\mathcal{A}$ is a finite sequence, then $|\mathcal{A}|$ denotes its cardinality

5. $\mathcal{P} = \langle p_i : p_i \text{ is the } i\text{-th prime } \rangle$, a sequence of prime numbers

6. $P_z = \prod_{p\in\mathcal{P}^z} p$, the product of primes less than or equal to z

7. $S(\mathcal{A}; \mathcal{P}^z, x)$ is the number of elements in $\mathcal{A}^x$ that "survive" the sieving process by the sequence $\mathcal{P}^z$, i.e. the number of elements in $\mathcal{A}^x$ that are not divisible by any of the primes up to z.

8. $\omega(p)$ is defined such that $\frac{\omega(p)}{p}x$ is a good approximation of $|\mathcal{A}_p^x|$. More generally, is d is a squarefree integer, we can generalize this notation by defining $\omega(d) = \prod_{p|d}\omega(d)$. An intuitive way to think about it is the number of elements divisible by d.

9. $R_d(x) = |\mathcal{A}_d^x| - \frac{\omega(p)}{p}x$, i.e. the remainder in our estimate of $|\mathcal{A}_d^x|$

10. Define the weight function
$$W(z) = \prod_{p|P_z}\left(1 - \frac{\omega(p)}{p}\right)$$
which estimates the number of elements that survive the sieve

Theorem on General Sieves:
For all sufficiently large x and $z < x$, there is a θ with $|\theta| \leq 1$ (θ depending on z), such that
$$S(\mathcal{A}; \mathcal{P}^z, x) = xW(z) + \theta\sum_{d|P_z} R_d(x).$$
If we have $|R_d(x)| \leq \omega(d)$ and $\omega(p) \leq C_0$ then
$$S(\mathcal{A}; \mathcal{P}^z, x) = xW(z) + O((1+C_0)^{\pi(z)}).$$

To prove this theorem, we will first define the sifting function

$$\sigma(n) = \begin{cases} 1 & \text{if } (n, P_z) = 1 \\ 0 & \text{otherwise} \end{cases}$$

which can be rewritten as

$$\sigma(n) = \sum_{d \mid \gcd(n, P_Z)} \mu(d).$$

Then we can simply do:

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}^z, x) &= \sum_{\substack{n \in \mathcal{A} \\ n \leq x}} \sigma(n) \\ &= \sum_{n \in \mathcal{A}^x} \sum_{d | gcd(n, P_z)} \mu(d) \\ &= \sum_{d | P_z} \mu(d) \left(\sum_{\substack{n \in \mathcal{A}^x \\ d | n}} 1 \right) \\ &= \sum_{d | P_z} \mu(d) |\mathcal{A}_d^x| \\ &= \sum_{d | P_z} \mu(d) \left(\frac{\omega(d)}{d} x + R_d(x) \right) \\ &= x \sum_{d | P_z} \frac{\mu(d)\omega(d)}{d} + \sum_{d | P_z} \mu(d) R_d(x) \\ &= x \prod_{p | P_z} \left(1 - \frac{\omega(p)}{p} \right) + \sum_{d | P_z} \mu(d) R_d(x) \\ &= xW(z) + \sum_{d | P_z} \mu(d) R_d(x) \\ &= xW(z) + \theta \sum_{d | P_z} R_d(x) \text{ where } |\theta| \leq 1. \end{aligned}$$

We can assume $|R_d(x)| \leq \omega(d)$ and suppose there exists a constant C_0 such that $\omega(p) \leq C_0 \Rightarrow \omega(d) \leq C_0^{v(d)}$

$$\begin{aligned} \sum_{d | P_z} R_d(x) &\leq \sum_{d | P_z} C_0^{v(d)} \\ &= \prod_{p | P_z} (1 + C_0) \\ &= (1 + C_0)^{\pi(z)} \end{aligned}$$

which concludes the proof of the theorem $\square$

7 Appendix

7.1 Notation

For our purposes, p will always denote a prime number, and $\sum_{n \leq p \leq m} f(p)$ will denote a sum over the prime numbers within the given range. We will also use Vinogradov notation $\ll$ to mean that the inequality holds with a constant, i.e. $f(n) \ll g(n) \Rightarrow \exists c > 0 \mid f(n) \leq cg(n)$.

We will be using the floor function, which will be defined as the largest integer not exceeding x, denoted $\lfloor x \rfloor$. The Mobius function is denoted by $\mu(n)$, where

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1 \cdots p_k, \text{ for } 1 \leq i < j \leq k : p_i \neq p_j, \\ 0 & \text{otherwise.} \end{cases}$$

The prime counting function $\pi(x)$ is the cardinality of the set $P = \{p \leq x \mid p \text{ prime}\}$, and the prime counting function modulo q $\pi(x; q, a)$ will be defined as the cardinality of the set $P_q = \{p \leq x \mid p \equiv a \mod q\}$. The von-Mangoldt Function $\Lambda(n)$, where

$$\Lambda(n) = \begin{cases} \log(p) & \text{if } n = p^k \text{ for a prime} p \\ 0 & \text{otherwise} \end{cases}$$

and its cumulative function $\Psi(x) = \sum_{n \leq x} \Lambda(n)$.

7.2 Auxiliary Theorems

Theorem 1:

$$\sum_{d|N} \frac{\mu(d)}{d} = \prod_{p|n} \left(1 - \frac{1}{p}\right)$$
$$= \prod_{p|n} \left(1 + \frac{\mu(p)}{p}\right)$$

Proof: We know that $\sum_{d|n} \phi(d) = n$. Using Möbius inversion on this we get:

$$n \prod_{p|n} \left(1 - \frac{1}{p}\right) = \phi(n) = \sum_{d|n} \mu(d) \left(\frac{n}{d}\right). = n \sum_{d|n} \frac{\mu(d)}{d}.$$