Some Sieve Theory...

Sachin Kumar
Faculty of Mathematics, Univeristy of Waterloo

April 1, 2024

Abstract

In the essay, I will talk about a fundamental area in analytic number theory called Sieve theory and its relating asymptotics in arithmetic functions. A sieve is a tool for separating desired objects from other objects. The most well-known sieve being the Sieve of Eratosthenes (which we will discuss in detail later), which states that if $n \leq x$ is not prime, then $n = pa$ where p is a prime, $p \leq \sqrt{x}$, i.e., an algorithm for finding all prime numbers up to any given limit.

Contents

1 Asymptotics and Arithmetic Functions

DEFINITION 1.1. An ***arithmetic function*** is a function $f : \mathbb{N} \to \mathbb{C}$. These functions can be used to capture and study certain arithmetic behaviour.

EXAMPLE 1.2. Here are some examples:

(a) $\nu(n) = \#\{\text{distinct prime divisors } p \mid n\}$ or $\omega(n)$.
(b) $d(n) = \#\{\text{divisors } d \mid n\}$ or $\sigma_0(n)$
(c) $\varphi(n) = \#\{1 \leq d < n : \gcd(d, n) = 1\} = |(\mathbb{Z}/n\mathbb{Z})^\times|$ or $\phi(n)$
(d) If $\mathcal{A} \subseteq \mathbb{N}$,

$$1_{\mathcal{A}}(n) = \begin{cases} 1 & n \in \mathcal{A} \\ 0 & n \notin \mathcal{A} \end{cases}$$

Arithmetic functions often have very erratic behavior, which makes them more difficult to deal with using analytic techniques. Consider the divisor function $d(n) = \#\{\text{positive divisors of } n\}$. Let p be a prime,

- If $n = p$, then $d(n) = 2$.
- If $n = 2^k$, then $d(2^k) = k + 1$.

DEFINITION 1.3. Let $f(x)$ and $g(x)$ be two functions and let $x \to \infty$. We say $f(x)$ is ***asymptotic to*** $g(x)$ and write $f(x) \sim g(x)$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$$

So using asymptotic theory, we can smooth out the information contained in an arithmetic function by considering the function of a real variable x,

$$\sum_{n\leq x} d(n) \sim x \log x \iff \lim_{x\to\infty} \frac{\sum_{n\leq x} d(n)}{x \log x} = 1$$

Here is a example. Consider the floor function $\lfloor x \rfloor$ i.e., the greatest integer $\leq x$. Alternatively,

$$\lfloor x \rfloor = \sum_{n\leq x} 1$$

LEMMA 1.4. $\lfloor x \rfloor \sim x$

PROOF. By definition, we know $\lfloor x \rfloor \leq x$. Also, $\lfloor x \rfloor + 1 > x$ implies $\lfloor x \rfloor > x - 1$. Now, we will find $\lim_{x\to\infty} \frac{\lfloor x \rfloor}{x}$, so we have

$$1 - \frac{1}{x} = \frac{x-1}{x} < \frac{\lfloor x \rfloor}{x} \leq \frac{x}{x} = 1$$

Also, we have $\lim_{x\to\infty} 1 - \frac{1}{x} = 1$ and $\lim_{x\to\infty} 1 = 1$. So by Squeeze Theorem, we can conclude that

$$\lim_{x\to\infty} \frac{\lfloor x \rfloor}{x} = 1$$

Hence proved. □

DEFINITION 1.5. Let $f(x)$ and $g(x)$ be function of the real variable x. We define the following:

(a) $f(x)$ is ***little-oh*** of $g(x)$, written $f(x) = o(g(x))$, if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$$

In this case, $f(x)$ is ***asymptotically smaller*** than $g(x)$.

(b) $f(x)$ is ***big-oh*** of $g(x)$, written $f(x) = O(g(x))$ or $f(x) \ll g(x)$, if there exists a constant $C > 0$ such that $|f(x)| \leq C \cdot |g(x)|$ for all $x \geq x_0$. Equivalently,

$$\limsup_{x\to\infty} \frac{f(x)}{g(x)} < \infty$$

In this case, $f(x)$ is ***asymptotically the same order of magnitude or smaller*** than $g(x)$.

LEMMA 1.6. *If $f(x) \sim g(x)$, then $f(x) = O(g(x))$.*

We write

$$f(x) = \underbrace{g(x)}_{\text{main term}} + \underbrace{O(h(x))}_{\text{error term}} \iff f(x) - g(x) = O(h(x))$$

Similar notation applies to ***little-oh***.

LEMMA 1.7. *$O(h(x))$ and $o(h(x))$ are ideals in the ring of functions defined for x sufficiently large. The above notation is then equivalent to stating that $f(x)$ and $g(x)$ belong to the same coset when quotienting by the ideal $O(h(x))$.*

LEMMA 1.8. $\lfloor x \rfloor = x + O(1)$

PROOF. We want to bound $|\lfloor x \rfloor - x|$. We know that $\lfloor x \rfloor \leq x$ and $\lfloor x \rfloor > x - 1$, implying $-1 < \lfloor x \rfloor - x \leq 0$, concluding that $|\lfloor x \rfloor - x| \leq 1$. So, there exists a constant $C > 0$ (specifically $C = 1$) such that $|\lfloor x \rfloor - x| \leq C \cdot 1$ for all $x \geq 0$, hence implying $\lfloor x \rfloor - x = O(1)$. Therefore, $\lfloor x \rfloor = x + O(1)$. Hence proved. □

LEMMA 1.9. *Let $f(x)$, $g(x)$ and $h(x)$ be functiosn and let $x \to \infty$.*

(a) *Let $f(x) \cdot O(g(x)) = O(f(x)g(x))$ and $f(x) \cdot o(g(x)) = o(f(x) \cdot g(x))$. If $h(x) = O(g(x))$, then $f(x)h(x) = O(f(x)h(x))$.*

(b) *If $f(x) = O(g(x))$ and $h(x) = O(g(x))$, then $f(x) + h(x) = O(g(x))$.*

(c) *If $f(x) = O(g(x))$ and $g(x) = O(h(x))$, then $f(x) = O(h(x))$.*

(d) *If $f(x) = O(g(x))$, then*

$$\sum_{n\leq x} f(n) = O\left(\sum_{n\leq x} g(n)\right)$$

(e) *If $f(x) = O(g(x))$ and some $y \in \mathbb{R}$, then*

$$\int_y^x f(t)\, dt = O\left(\int_y^x g(t)\, dt\right)$$

2 Abel Summation

THEOREM 2.1. *Write* $A(x) = \sum_{n \le x} a_n$ *and suppose* $f(t)$ *is a differentiable function on the interval* (y, x) *for* $y < x < \infty$. *Then,*

$$\sum_{y<n\le x} a_n f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)\,dt$$

*$A(t)$ is like a "**discrete antiderivative**" of a_n, and we can recognize the familar integration by parts formula with $u = f(t)$ and $\varepsilon dv\varepsilon = a_n$ (so, $du = f'(t)\,dt$ and $\varepsilon v\varepsilon = \sum_{n\le t} a_n = A(t)$):*

$$\sum_{y<n\le x} a_n f(n) = A(t)f(t)\Big|_y^x - \int_y^x A(t)f'(t)\,dt$$

COROLLARY 2.2. $\sum_{n\le x} \frac{1}{n} = \log x + O(1)$

PROOF. By Abel Summation, we have $a_n = 1$, $f(t) = \frac{1}{t} \Rightarrow f'(t) = -\frac{1}{t^2}$, $A(t) = \lfloor t \rfloor$.

$$\begin{aligned}
\frac{1}{1} + \sum_{1<n\le x} \frac{1}{n} &= \frac{1}{1} + \lfloor t \rfloor \cdot \frac{1}{t}\Big|_1^x - \int_1^x \lfloor t \rfloor \left(-\frac{1}{t^2}\right) dt = \frac{\lfloor x \rfloor}{x} + \int_1^x \frac{\lfloor t \rfloor}{t^2}\,dt \\
&= \frac{x + O(1)}{x} + \int_1^x \frac{t + O(1)}{t^2}\,dt \\
&= 1 + O\left(\frac{1}{x}\right) + \int_1^x \frac{1}{t}\,dt + \underbrace{O\left(\int_1^x \frac{1}{t^2}\,dt\right)}_{O\left(1-\frac{1}{x}\right)}
\end{aligned}$$

Hence proved. □

THEOREM 2.3. $\sum_{n\le x} d(n) = x \log x + O(x)$

PROOF. We know that $d(n) = \#\{\text{positive } d \mid n\} = \sum_{d|n} 1 = \sum_{da=n} 1$, for some $a \in \mathbb{Z}$. So, we

have

$$\begin{aligned}\sum_{n\le x} d(n) = \sum_{n\le x}\sum_{da=n} 1 &= \sum_{da\le x} 1\\ &= \sum_{d\le x}\left(\sum_{a\le \frac{x}{d}} 1\right)\\ &= \sum_{d\le x}\left\lfloor \frac{x}{d}\right\rfloor\\ &= \sum_{d\le x}\left(\frac{x}{d}+O(1)\right)\\ &= x\left(\sum_{d\le x}\frac{1}{d}\right)+O\left(\sum_{d\le x}1\right)\\ &= x(\log x + O(1)) + O(\lfloor x\rfloor)\\ &= x\log x + O(x) + O(\lfloor x\rfloor)\end{aligned}$$

since $\lfloor x\rfloor \le x$, so $\lfloor x\rfloor = O(x)$. Hence proved. □

3 Möbius Function

DEFINITION 3.1. The ***Möbius function*** is defined as follows:

$$\mu(n) = \begin{cases}(-1)^k & n = \prod_{i=1}^{k} p_i \text{ is a product of } k \text{ distinct primes}\\ 0 & \text{otherwise}\end{cases}$$

LEMMA 3.2. *Let $\mu(d)$ be the Möbius function,*

$$\sum_{d|n}\mu(d) = \begin{cases}1 & n = 1\\ 0 & n \ne 0\end{cases}$$

PROOF. Suppose $n = 1$, we have

$$\sum_{d|1}\mu(d) = \mu(1) = 1$$

Suppose $n = \prod_{i=1}^{k} p_i^{e_{p_i}} \ne 0$, using the binomial theorem we get

$$\begin{aligned}\sum_{d|n}\mu(d) &= \mu(1) + \sum_{i=1}^{k}\mu(p_i) + \sum_{\{i,j\}}\mu(p_ip_j) + \cdots + \mu(p_1\cdots p_k)\\ &= \binom{k}{0} + \binom{k}{1}(-1) + \binom{k}{2}(-1)^2 + \cdots + \binom{k}{k}(-1)^k\\ &= ((-1)+1)^k = 0\end{aligned}$$

Hence proved. □

THEOREM 3.3. *If* $f(n) = \sum_{d|n} g(d)$, *then* $g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right)$

PROOF. We have

$$\begin{aligned}
\sum_{d|n} \mu(d) f\left(\frac{n}{d}\right) &= \sum_{da=n} \mu(d) f(a) \\
&= \sum_{da=n} \mu(d) \left(\sum_{b|a} g(b) \right) \\
&= \sum_{da=n} \mu(d) \left(\sum_{bc=a} g(b) \right) \\
&= \sum_{dbc=n} \mu(d) g(b) \\
&= \sum_{b|n} g(b) \left(\sum_{dc=\frac{n}{b}} \mu(d) \right) \\
&= \sum_{b|n} g(b) \left(\sum_{d \left| \frac{n}{b} \right.} \mu(d) \right) \\
&= 0 + 0 + \cdots + g(n) \cdot 1
\end{aligned}$$

where,

$$\sum_{d \left| \frac{n}{b} \right.} \mu(d) = \begin{cases} 1 & \text{if } \frac{n}{b} = 1 \Rightarrow b = n \\ 0 & \text{if } \frac{n}{b} \neq 1 \Rightarrow b \neq n \end{cases}$$

Hence proved. □

EXAMPLE 3.4. $\sigma_0(n) = \sum_{d|n} 1 \Rightarrow 1 = \sum_{d|n} \mu(d) \sigma_0\left(\frac{n}{d}\right)$

a. *Squarefree Numbers*

DEFINITION 3.5. An integer n is called squarefree if for all prime numbers p we have $p^2 \nmid n$. Let's denote $\S = \{s_1, s_2, \dots\}$ be the set of squarefree integers.

Is there a sieve for squarefree numbers? Let's try something (method of inclusion-exclusion) …

$$\begin{aligned}
&\lfloor x \rfloor - \left\lfloor \frac{x}{2^2} \right\rfloor - \left\lfloor \frac{x}{3^2} \right\rfloor - \left\lfloor \frac{x}{5^2} \right\rfloor - \left\lfloor \frac{x}{7^2} \right\rfloor - \cdots \\
&\quad + \left\lfloor \frac{x}{2^2 \cdot 3^2} \right\rfloor + \left\lfloor \frac{x}{2^2 \cdot 5^2} \right\rfloor + \cdots + \left\lfloor \frac{x}{3^2 \cdot 5^2} \right\rfloor + \cdots \\
&\quad - \left\lfloor \frac{x}{2^2 \cdot 3^2 \cdot 5^2} \right\rfloor - \cdots + \cdots = \sum_d \mu(d) \left\lfloor \frac{x}{d^2} \right\rfloor
\end{aligned}$$

THEOREM 3.6. $\#\{\S \le x\} = \left(\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2}\right) x + O(\sqrt{x})$

PROOF. First, we can observe that $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = \frac{6}{\pi^2}$. Also, If $d^2 > x$ (i.e., $d > \sqrt{x}$), then $\left\lfloor \frac{x}{d^2} \right\rfloor = 0$. We have the following:

$$\sum_d \mu(d) \left\lfloor \frac{x}{d^2} \right\rfloor = \sum_{d \le \sqrt{x}} \mu(d) \left\lfloor \frac{x}{d^2} \right\rfloor = \sum_{d \le \sqrt{x}} \mu(d) \left(\frac{x}{d^2} + O(1) \right) = \left(\sum_{d \le \sqrt{x}} \frac{\mu(d)}{d^2} \right) x + \underbrace{O\left(\sum_{d \le \sqrt{x}} 1 \right)}_{O(\lfloor \sqrt{x} \rfloor)}$$

$$\left(\sum_{d \le \sqrt{x}} \frac{\mu(d)}{d^2} \right) x = \left(\sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} - \sum_{\sqrt{x} < d} \frac{\mu(d)}{d^2} \right) x = \left(\sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} \right) x - \left(\sum_{\sqrt{x} < d} \frac{\mu(d)}{d^2} \right) x$$

This only makes sense if $\sum_{d=1}^{\infty} \frac{\mu(d)}{d^2}$ converges, $\left| \frac{\mu(d)}{d^2} \right| \le \left| \frac{1}{d^2} \right|$, so converges by comparison with $\sum_{d=1}^{\infty} \frac{1}{d^2}$. Taking, $a_n = 1 \Rightarrow A(t) = \lfloor t \rfloor$, $f(t) = \frac{1}{t^2} \Rightarrow f'(t) = \frac{-2}{t^3}\, dt$, we get

$$\begin{aligned} \left| \sum_{\sqrt{x} < d} \frac{\mu(d)}{d^2} \right| &\le \left| \sum_{\sqrt{x} < d \le \infty} \frac{1}{d^2} \right| \\ &\le \left| \frac{\lfloor t \rfloor}{t^2} \Big|_{\sqrt{x}}^{\infty} \right| + \left| 2 \int_{\sqrt{x}}^{\infty} \frac{\lfloor t \rfloor}{t^3}\, dt \right| \\ &\le \left| \frac{1}{t} \Big|_{\sqrt{x}}^{\infty} \right| + \left| 2 \int_{\sqrt{x}}^{\infty} \frac{1}{t^2}\, dt \right| \ll \frac{1}{\sqrt{x}} \end{aligned}$$

This holds, since $\lfloor t \rfloor = t + O(1) \Rightarrow \lfloor r \rfloor \le t$. Hence proved. □

4 Prime Numbers

Prime numbers has been a fundamental structure of study, since the period of Euclid (providing a proof of infinitude of primes). We know that Sieve of Eratosthenes, gives us a set of primes numbers upto a given limit. But how many are there? The prime number theorem gives an approximated answer to this question (which we will discuss in the next section).

a. *Foundational Results*

THEOREM 4.1 (CHEBYSHEFF'S THEOREM). $\pi(x) = O\left(\frac{x}{\log x}\right)$

PROOF. Next section! □

THEOREM 4.2. $\sum_{p\le x}\frac{\log p}{p} = \log x + O(1)$

PROOF. We know $n! = \prod_p p^{e_p} = \prod_{k=1}^{n} k$.

$$e_p = \underbrace{\left\lfloor \frac{n}{p} \right\rfloor}_{\#d\le n \text{ s.t. } p|d} + \underbrace{\left\lfloor \frac{n}{p^2} \right\rfloor}_{\#d\le n \text{ s.t. } p^2|d} + \cdots$$

$$\sum_{p\le n} e_p \log p = \sum_{p\le n} \log p \left(\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \cdots \right) = \sum_{k\le n} \log k = n\log n - n + O(\log n)$$

Taking the leading term,

$$\begin{aligned}\sum_{p\le n} \log p \left\lfloor \frac{n}{p} \right\rfloor &= \sum_{p\le n} \log p \left(\frac{n}{p} + O(1) \right) \\ &= n \sum_{p\le n} \frac{\log p}{p} + O\left(\sum_{p\le n} \log p \right) \\ &= n \sum_{p\le n} \frac{\log p}{p} + O(n)\end{aligned}$$

by Chebysheff's Theorem. For rest of the terms, we have

$$\begin{aligned}\sum_{p\le n} \log p \left(\left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots \right) &\le \sum_{p\le n} \log p \cdot \sum_{i=2}^{\infty} \frac{n}{p^i} \\ &= \sum_{p\le n} \log p \cdot \left(\frac{\frac{n}{p^2}}{1 - \frac{1}{p}} \right) \\ &= n \sum_{p\le n} \frac{\log p}{p(p-1)} \\ &= O(n)\end{aligned}$$

$\sum_{p\le n} \frac{\log p}{p(p-1)}$ converges as $n \to \infty \le \sum \frac{1}{n^{3/2}}$. Implying,

$$n \sum_{p\le n} \frac{\log p}{p} + O(n) + O(n) = n\log n - n + O(\log n)$$

$$\sum_{p\le n} \frac{\log p}{p} + O(n) + O(n) = \log n - 1 + O\left(\frac{\log n}{n} \right)$$

These are all $O(1)$. Hence proved. □

LEMMA 4.3. $\sum_{n\le x} \log n = x\log x - x + O(\log x)$

COROLLARY 4.4. $\sum_{p\le x} \frac{1}{p} = \log\log x + O(1)$

We could prove Corollary 4.4, using Abel Summation with $f(t) = (\log t)^{-1}$. We also get a nice intuitive argument for free!

COROLLARY 4.5. $\lim_{p\to\infty} \sum_{p} \frac{1}{p}$ *diverges.*

b. *Prime Number Theorem*

DEFINITION 4.6. Define $\pi(x)$ as the prime number counting function,

$$\pi(x) = \#\{p \le x \,:\, p \text{ prime}\} = \sum_{p\le x} 1$$

We will state and prove a Theorem by Chebysheff on the asymptotic bound for $\pi(x)$.

THEOREM 4.7 (CHEBYSHEFF'S THEOREM). $\pi(x) = O\left(\frac{x}{\log x}\right)$

LEMMA 4.8. $\pi(x) = O\left(\frac{x}{\log x}\right)$ *if and only if* $\theta(x) = \sum_{p\le x} \log p = O(x)$.

PROOF. We observe that,

$$\prod_{n<p\le 2n} p \,\Bigg|\, \binom{2n}{n} = \frac{(2n)!}{n!n!} \le 2^{2n}$$

where, $2^{2n} = (1+1)^{2n} = \sum_{i=0}^{2n} \binom{2n}{i}(1)^i$. So we have,

$$\sum_{n<p\le n} \log p \le 2n\log 2$$
$$\theta(2n) - \theta(n) \le 2n\log 2$$

We will prove $\theta(n) \le 4n\log 2$ via strong induction.
Base case: Suppose $n = 1$, then

$$\theta(1) = \sum_{p\le 1} \log p = 0 \le 4\cdot 1\cdot \log 2$$

Inductive step: Suppose for all $1 \le k < n$,

$$\theta(k) \le 4k\log 2$$

Suppose n is even, then

$$\begin{aligned}\theta(n) - \theta\left(\frac{n}{2}\right) &\leq n\log 2\\ \theta(n) &\leq n\log 2 + \theta\left(\frac{n}{2}\right)\\ &\leq n\log 2 + 4\left(\frac{n}{2}\right)\log 2 \quad \text{[Inductive Hypothesis]}\\ &\leq 3n\log 2\\ &\leq 4n\log 2\end{aligned}$$

Suppose $n \geq 3$ is odd. We observe that $n+1$ is even and $\geq 4 \implies n+1$ is not prime. So,

$$\theta(n) = \sum_{p\leq n} \log p = \sum_{p\leq n+1} \log p = \theta(n+1)$$

$n+1$ is even and $\frac{n+1}{2} < n$. Hence,

$$\begin{aligned}\theta(n) = \theta(n+1) &\leq (n+1)\log 2 + \theta\left(\frac{n+1}{2}\right)\\ &\leq (n+1)\log 2 + 4\left(\frac{n+1}{2}\right)\log 2 \quad \text{[Inductive Hypothesis]}\\ &\leq 3(n+1)\log 2\\ &\leq 4n\log 2 \qquad\qquad\qquad\qquad \text{[by } n \geq 3\text{]}\end{aligned}$$

implying $\theta(n) = O(n)$. Hence proved. □

We will state two equivalent statement of the Prime number theorem,

THEOREM 4.9 (PRIME NUMBER THEOREM VERSION 1). *Let $\pi(x)$ be the prime counting function. Then*

$$\pi(x) \sim \frac{x}{\log x}$$

DEFINITION 4.10. Let $\mathrm{Li}(x)$ be the logarithmic integral defined by $\displaystyle\int_2^x \frac{dt}{\log t}$.

THEOREM 4.11 (PRIME NUMBER THEOREM VERSION 2). *Let $\pi(x)$ be the prime counting function and $Li(x)$ be the logarithmic integral. Then*

$$\pi(x) \sim Li(x)$$

There are variety of proofs of the ***Prime Number Theorem***, the most accessible of which require complex analytic techniques. There does exist an "elementary proof" (i.e. one that does not appeal to complex analysis). Before proving this theorem, we will state and proving some required background theorems and lemmas.

Recall the Riemann zeta function, which is defined as

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s}$$

on the strip $\Re(s) > 1$. As we will see later, the zeta function can then be meromorphically extended to the complex plane, with exactly a simple pole at $s = 1$ of residue 1, and with trivial zeros at the negative even integers.

Define the two auxiliary functions,

$$\Phi(s) = \sum_p \frac{\log p}{p^s} \qquad \theta(x) = \sum_{p \le x} \log p$$

Note that the series defining $\Phi(s)$ converges absolutely uniformly in any compact subset of $\Re(s) > 1$ and hence $\Phi(s)$ is holomorphic there. Through a series of steps, we will prove the asymptotic relation $\theta(x) \sim x$, which then immediately implies the prime number theorem.

Moreover, define the xi function as

$$\xi(s) = \pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)$$

in the half-plane $\Re(s) > 1$. We will state a property of the xi function without proof.

The function $\xi(s)$ has an analytic continuation into a meromorphic on the complex plane with only simple poles at 0,1, and

$$\xi(s) = \frac{1}{s-1} - \frac{1}{s} + F(s)$$

where $F(s)$ is entire.

COROLLARY 4.12. *The zeta function can then be meromorphically extended to the complex plane, with exactly a simple pole at $s = 1$ of residue 1, and with trivial zeros at the negative even integers.*

LEMMA 4.13. *The relation $\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$ holds for $\Re(s) > 1$.*

PROOF. Expand and use the Fundamental Theorem of Arithmetic to show that the LHS is the uniform limit of the partial products, $\prod_{p \le x} \left(1 - \frac{1}{p^s}\right)^{-1}$. □

LEMMA 4.14. *The function $\zeta(s) - \frac{1}{s-1}$ extends holomorphically to an entire function.*

PROOF. Using ***Fact 4.2.3.***, and since $\frac{\pi^{s/2}}{\Gamma(s/2)}$ is entire, looking at the residues implies that $\zeta(s) - \frac{1}{s-1}$ extends to an entire function. □

LEMMA 4.15. *We have the relation $\theta(x) \in O(x)$.*

PROOF. If $n \in \mathbb{N}$, then

$$4^n = (1+1)^{2n} = \sum_{i=0}^{2n} \binom{2n}{i} > \binom{2n}{n} = \frac{(2n)!}{(n!)^2} \ge \prod_{n < p \le 2n} p = e^{\theta(2n) - \theta(n)}$$

Therefore,

$$\theta(2n) - \theta(n) \leq (2\log 2)n \tag{1}$$

Observe that for $\delta \leq 1$, we have that $\theta(x+\delta) - \theta(x)$ is at most $\log(x+1)$. This allows us to bound $\theta(x)$ in terms of $\theta(\lceil x \rceil)$, and an application of the "Master Theorem" to (1), implies that $\theta(x) \in O(x)$. □

LEMMA 4.16. *In the closed half plane $\Re(s) \geq 1$, the zeta function is nonzero and $\Phi(s) - \dfrac{1}{s-1}$ is holomorphic.*

PROOF. Recall that $\zeta(s) = \prod_p \left(1 - \dfrac{1}{p^s}\right)^{-1}$ holds for $\Re(s) > 1$. By Prop 5.3.2 (in E. M. Stein and R. Shakarchi, Complex Analysis), we have

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_p \frac{\frac{d}{ds}(1-p^{-s})^{-1}}{(1-p^{-s})^{-1}} = \sum_p \frac{-\log p}{p^s - 1} = -\Phi(s) - \sum_p \frac{\log p}{p^s(p^s-1)} \tag{2}$$

Since, $\sum_p \dfrac{\log p}{p^s(p^s-1)}$ converges to a holomorphic function in $\Re(s) > \frac{1}{2}$, and using ***Corollary 4.2.4***, we have that Φ extends meromorphically unto $\Re(s) > \frac{1}{2}$.

In particular, this implies that $\zeta(s)$ has no roots in $\Re(s) > 1$, as otherwise the RHS would have a pole there, but $\Phi(s)$ is holomorphic in $\Re(s) > 1$ by a previous observation.

Note that

$$\mathrm{Res}_{s=1}\Phi(s) = \lim_{\epsilon \to 0} \epsilon\Phi(1+\epsilon) = 1$$

That is, $\Phi(s)$ has a simple pole of residue 1 at $s = 1$, and by the analytic continuation of the zeta function and (2), this is the only pole of Φ in $\Re(s) > \frac{1}{2}$. Therefore, $\Phi(s) - \dfrac{1}{s-1}$ has a holomorphic extension in $\Re(s) > \frac{1}{2}$.

Now assume that ζ has a zero of multiplicity μ at $s = 1 + i\alpha$. Since ζ is real on the real line, the Schwarz reflection principle implies that $1 - i\alpha$ is also a root of multiplicity μ. Let ν be the multiplicity of the root at $1 \pm 2i\alpha$.

Similarly, using (2), we see that

$$\lim_{\epsilon \to 0} \epsilon\Phi(1 + \epsilon \pm i\alpha) = \mu$$

and that

$$\lim_{\epsilon \to 0} \epsilon\Phi(1 + \epsilon \pm 2i\alpha) = -\nu$$

We can therefore compute,

$$\sum_{r=-2}^{2} \binom{4}{2+r} \Phi(1 + \epsilon + ri\alpha) = \sum_{r=-2}^{2} \binom{4}{2+r} \sum_p \frac{\log p}{p^{1+\epsilon+ri\alpha}} = \sum_p \frac{\log p}{p^{1+\epsilon}} \left(p^{i\alpha/2} + p^{-i\alpha/2}\right)^4 \geq 0$$

On the other hand, expanding the following expression yields,

$$\lim_{\epsilon\to 0}\epsilon\sum_{r=-2}^{2}\binom{4}{2+r}\Phi(1+\epsilon+ri\alpha)6-8\mu-2\nu$$

Combining the two results yields that $6-8\mu-2\nu\geq 0$ and hence $4\mu+\nu\leq 3$. Since, μ, ν are non-negative integers (multiplicities of roots), we must have $\mu=0$, and so $\zeta(s)$ has no roots on the line $\Re(s)=1$. □

THEOREM 4.17 (ANALYTIC THEOREM). *Let $f(t)$ be a bounded, locally integrable function over the non-negative reals. Suppose that the function*

$$g(z)=\int_0^\infty e^{-xt}f(t)\,dt$$

defined on $\Re(z)>0$ extends holomorphically to $\Re(z)\geq 0$. Then $\int_0^\infty f(t)\,dt$ converges and equals $g(0)$.

LEMMA 4.18. *The integral $\int_1^\infty \frac{\theta(x)-x}{x^2}\,dx$ converges.*

PROOF. Note that $\theta(x)=\sum_p \mathbb{1}_{x\geq p}\log p$, and so if $\Re(s)>1$, then

$$s\int_1^\infty\frac{\theta(x)}{x^{s+1}}\,dx=\sum_p s\int_p^\infty\frac{dx}{x^{s+1}}\log p=\sum_p\frac{\log p}{p^s}=\Phi(s)$$

Making a change of variable $x=e^t$, we get that this is equal to

$$s\int_0^\infty e^{-st}\theta(e^t)\,dt$$

Applying the ***analytic theorem*** to $f(t)=\theta(e^t)e^{-t}-1$, which is clearly locally integrable and is bounded by ***Lemma 4.15***, we get that

$$f(z)=\int_0^\infty e^{-zt}\Big(\theta(e^t)e^{-t}-1\Big)\,dt=\frac{\phi(z+1)}{z+1}-\frac{1}{z}$$

This is holomorphic by ***Lemma 16***, and the poles at zero cancel by comparing residues, so the analytic theorem implies that

$$\int_0^\infty\Big(\theta(e^t)e^{-t}-1\Big)\,dt<\infty$$

Making a change of variables once again, we get that

$$\int_1^\infty\frac{\theta(x)-x}{x^2}\,dx$$

converges, as required. Hence proved. □

LEMMA 4.19. *The asymptotic equivalence $\theta(x) \sim x$ holds.*

PROOF. Assume, for the sake of contradiction, that there exists some $\lambda > 1$ such that $\theta(x) \geq \lambda x$ for infinitely many x. Then, by ***Lemma 4.18***, we have

$$\int_x^{\lambda x} \frac{\theta(t) - t}{t^2}\, dt \geq \int_x^{\lambda x} \frac{\lambda(t) - t}{t^2}\, dt = \int_1^{\lambda} \frac{\lambda - t}{t^2}\, dt > 0$$

However, this contradicts the convergence of the integral from ***Lemma 4.18***.

Similarly, we get a contradiction if we assume that there exists some $\lambda < 1$ such that $\theta(x) \leq \lambda x$ for infinitely many x. □

PROOF AND CONSEQUENCES OF THE PRIME NUMBER THEOREM

Now, we have the required analytical and asymptotic background to prove the Prime number Theorem (***Theorem 4.9*** and ***Theorem 4.11***, which are two equivalent statements).

THEOREM 4.20 (PRIME NUMBER THEOREM VERSION 1). *Let $\pi(x)$ be the prime counting function. Then*

$$\pi(x) \sim \frac{x}{\log x}$$

PROOF. Note that

$$\theta(x) = \sum_{p \leq x} \log p \leq \sum_{p \leq x} \log x = \pi(x) \log x \tag{3}$$

However, we also have that for any $0 < \epsilon < 1$,

$$\theta(x) = \sum_{p \leq x} \log p \geq \sum_{x^{1-\epsilon} \leq p \leq x} \log p \geq (1 - \epsilon) \log x \Big(\pi(x) - \pi(x^{1-\epsilon}) \Big) \tag{4}$$

Combining (3) and (4), we get that for any $0 < \epsilon < 1$,

$$\frac{\theta(x)}{\log x} \leq \pi(x) \leq \frac{\theta(x)}{(1 - \epsilon) \log x} + O(x^{1-\epsilon})$$

Now, since $\theta(x) \sim x$, we can conclude that $\pi(x) \sim \dfrac{x}{\log x}$. Hence proved. □

THEOREM 4.21 (PRIME NUMBER THEOREM VERSION 2). *Let $\pi(x)$ be the prime counting function and $Li(x)$ be the logarithmic integral. Then*

$$\pi(x) \sim Li(x)$$

PROOF. It suffices to prove that $\mathrm{Li}(x) \sim \dfrac{x}{\log x}$. Using integration by parts, we have

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t} = \frac{x}{\log x} - \frac{2}{\log 2} + \int_2^x \frac{dt}{(\log t)^2}$$

Note that to prove $\mathrm{Li}(x) \sim \frac{x}{\log x}$, it suffices to show that

$$\lim_{x\to\infty} \frac{\log x}{x} \int_2^x \frac{dt}{(\log t)^2} = 0$$

Applying l' Hôpital's rule, this is equal to

$$\lim_{x\to\infty} \left(\frac{1}{x} \int_2^x \frac{dt}{(\log t)^2} + \frac{1}{\log x} \right) = \lim_{x\to\infty} \frac{\int_2^x \frac{dt}{(\log t)^2}}{x}$$

Once again, applying l' Hôpital's rule, we get that this is equal to

$$\lim_{x\to\infty} \frac{1}{(\log x)^2} = 0$$

which concludes the proof. □

In fact, $\mathrm{Li}(x)$ provides an even better approximation to $\pi(x)$ than $\frac{x}{\log x}$. It also turns out that $\mathrm{Li}(x)$ belongs to the class $O\left(\frac{x}{(\log x)^2}\right)$, although this was not required for our proof.

We will now discuss some consequences of the Prime number theorem,

THEOREM 4.22. *Let p_n be the n^{th} prime number. Then $p_n \sim n \log n$.*

PROOF. The prime number theorem implies that $\pi(x) \log x \sim x$, and hence, by properties of asymptotic equivalence, we have that $\log \pi(x) + \log\log x \sim \log x$. However, $\lim_{x\to\infty} \frac{\log\log x}{\log x} = 0$, and so $\log \pi(x) \sim \log x$. By the prime number theorem, we have that $n \sim \frac{p_n}{\log p_n}$, and so $p_n \sim n \log p_n \sim n \log n$. Hence proved. □

A longer proof that contains a recurring idea in analytic number theory is also presented.

PROOF. Fix $\epsilon > 0$. By the prime number theorem, we have that $n \sim \frac{p_n}{\log p_n}$. In particular, for large enough n, we have that $p_n^{\epsilon} \geq \log p_n$ and hence $p_n^{1-\epsilon} \leq n \leq p_n$. Taking logarithms, we see that

$$(1-\epsilon) \log p_n \leq \log n \leq \log p_n$$

and so $\log p_n \sim \log n$. Therefore, we get that $p_n \sim n \log n$, as required. □

THEOREM 4.23. *For all $\epsilon > 0$, there exist N such that for all $n \geq N$, the interval $[n, (1+\epsilon)n]$ contains a prime number.*

PROOF. The prime number theorem, and elementary properties of asymptotic equivalence, imply that

$$\frac{\pi\big((1+\epsilon)x\big)}{\pi(x)} \sim \frac{\frac{(1+\epsilon)x}{\log\left((1+\epsilon)x\right)}}{\frac{x}{\log x}} = \frac{(1+\epsilon)\log x}{\log\big((1+\epsilon)x\big)}$$

which by l'Hôpital's rule, tends to $1+\epsilon$. Hence, we have that $\lim_{x\to\infty} \frac{\pi\big((1+\epsilon)x\big)}{\pi(x)} = 1+\epsilon$, and so there exists N such that $\pi\big((1+\epsilon)x\big) > \pi(x)$, for all $x \geq N$. This directly implies the result. □

COROLLARY 4.24. *Given any string of decimal digits $a_1 \cdots a_k$, there exists infinitely many primes whose decimal representation begins with $a_1 \cdots a_k$.*

PROOF. Let $M = \overline{a_1 \cdots a_k}$. The decimal representation of an integer n begins with $a_1 \cdots a_k$ if and only if $10^k M \leq n < 10^k(M+1)$ for some $k \in \mathbb{N}$. Applying the previous theorem, we see that the interval $[10^k M, 10^k(M+1)]$ contains a prime for all large enough k. □

5 Sieve of Eratosthenes

Eratosthenes sieved out numbers divisible by small primes. We can this by considering the function (an intermediate step of the sieve)

$$\phi(x,z) = \#\{n \leq x \,:\, p \nmid n, \text{where } p < z\}$$

where $x, z \in \mathbb{R}^+$.

THEOREM 5.1. $\phi(x,z) = x\prod_{p<z}\left(1-\frac{1}{p}\right) + O(2^z)$

EXAMPLE 5.2. Suppose $z = \log\log x$. We have, $O(2^{\log\log x}) = O(\log x)$.

QUESTION 5.3. How does $x \prod_{p<\log\log x}\left(1-\frac{1}{p}\right)$ grow?

We can observe that $\prod_{p<z}\left(1-\frac{1}{p}\right) \leq 1$ as $z \to \infty$, and prod $\to 0$. Also, $O(2^z)$ grows very fast in z if z is too big, $O(2^z) \geq x\prod_{p<z}\left(1-\frac{1}{p}\right)$, i.e., bigger than the main term.

PROOF. We will prove $\phi(x,z) = x\prod_{p<z}\left(1-\frac{1}{p}\right) + O(2^z)$. Let $P(z) = \prod_{p<z} p$ and $(a,b) = \gcd(a,b)$. We also observe that,

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

Also, if $d \mid n \iff n = md$ for some $m \in \mathbb{Z}$, implying $n \leq x \iff m \leq \frac{x}{d}$. Therefore, We have

$$\begin{aligned}
\phi(x,z) &= \sum_{\substack{n\leq x \\ (n,P(z))=1}} 1 \\
&= \sum_{n\leq x}\left(\sum_{d\mid(n,P(z))} \mu(d)\right) \\
&= \sum_{n\leq x}\sum_{\substack{d\mid n \\ d\mid P(z)}} \mu(d) \\
&= \sum_{d\mid P(z)} \mu(d)\left(\sum_{\substack{n\leq x \\ d\mid n}} 1\right) = \sum_{m\leq \frac{x}{d}} 1 = \left\lfloor \frac{x}{d} \right\rfloor \neq 0, \quad \text{when } d \leq x
\end{aligned}$$

Since, if $d \mid n \iff n = md$, for some $n \in \mathbb{Z}$, then we observe that $n \leq x \iff m \leq \frac{x}{d}$.

$$\begin{aligned}
\phi(x,z) &= \sum_{d\mid P(z)} \mu(d)\left\lfloor \frac{x}{d} \right\rfloor \\
&= \sum_{d\mid P(z)} \left(\frac{x}{d} + O(1)\right) \\
&= x\sum_{\substack{d\mid P(z) \\ d\leq x}} \frac{\mu(d)}{d} + O\left(\sum_{\substack{d\mid P(z) \\ d\leq x}} 1\right) \\
&= x\left(1 - \sum_{p\mid P(z)} \frac{1}{p} + \sum_{p_1p_2\mid P(z)} \frac{1}{p_1p_2} - \cdots\right) \\
&= x\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\left(1-\frac{1}{p_3}\right)\cdots
\end{aligned}$$

We know that $P(z) = \prod_{p<z} p$, so we observe that $d \mid P(z)$ implies $d = \prod_{p\in S} p$, where S is the subset of primes, $p < z$. So, the number of subsets of set of size z is 2^z, so we get the error as $O(2^z)$. Hence,

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

Hence proved. □

To improve on the error $O(2^z)$, consider the function

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

where n is a z-smooth number. So, this tells us that

$$\begin{aligned}\phi(x,z) &= \sum_{d|P(z)} \mu(d)\lfloor\frac{x}{d}\rfloor \\ &= \sum_{\substack{d\le x\\ d|P(z)}} \mu(d)\left(\frac{x}{d}+O(1)\right) \\ &= x\sum_{\substack{d\le x\\ d|P(z)}} \mu(d) + O(\Psi(x,z))\end{aligned}$$

We have the following two goals to make sure to improve the bound we had in ***Theorem 5.1***:

(a) We want to bound $O(\Psi(x,z))$.

(b) To make sure $d \le x$, doesn't break anything.

THEOREM 5.4. $\Psi(x,z) \ll x(\log z)\exp\left(-\frac{\log x}{\log z}\right)$ *compare to* 2^z

We observe that $\Psi(x,z) \le \#(n \le x) = O(x)$. We will use the Rantkin's trick.

PROOF. We have

$$\begin{aligned}\Psi(x,z) &= \sum_{\substack{n\le x\\ p|n \Rightarrow p<z^1}} \\ &\le \sum_{\substack{n\le x\\ p|n \Rightarrow p<z}} \left(\frac{x}{n}\right)^{\delta} \quad \text{for some } \delta>0 \\ &\le x^{\delta} \sum_{\substack{n\\ p|n \Rightarrow p<z}} \frac{1}{n^{\delta}} \\ &= x^{\delta}\prod_{p<z}\left(1+\frac{1}{p^{\delta}}+\frac{1}{p^{2\delta}}+\frac{1}{p^{3\delta}}+\dots\right) \\ &= x^{\delta}\prod_{p<z}\left(1-\frac{1}{p^{\delta}}\right)^{-1} \\ &\le x^{\delta}\prod_{p<z}\left(1-\frac{1}{p^{\delta}}\right)^{-1} \\ &= x^{\delta}\prod_{p<z}\left(1+\frac{1}{p^{\delta}}\right)\prod_{p<z}\left(1-\frac{1}{p^{2\delta}}\right)^{-1}\end{aligned}$$

converges as $z \to \infty$, if $\delta > \frac{1}{2}$ and using $1 + x \leq e^x$, so

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

Set $\delta = 1-\eta$, for η small. So,

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

using the fact that $e^x \leq 1 + xe^x$, we have

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

Taking $\eta = \dfrac{1}{\log z}$, we know that

$$\begin{aligned}e^{\eta\log p} &= e^{\frac{\log p}{\log z}}\\ &\leq p^{\frac{1}{\log z}}\\ &\leq z^{\frac{1}{\log z}}\\ &= e\end{aligned}$$

So, we have

$$\begin{aligned}\Psi(x,z) &\ll x^{1-\frac{1}{\log z}}\exp\left(\sum_{p<z}\frac{1}{p}\left(1+\frac{\log p}{\log z}\cdot e\right)\right)\\ &\ll x^1\exp\left(-\frac{\log x}{\log z}\right)\exp\left(\sum_{p<z}\frac{1}{p}+\frac{e}{\log z}\sum_{p<z}\frac{\log p}{p}\right) \rightsquigarrow (\log\log z + O(1)) + \underbrace{\frac{e}{\log z}(\log z + O(1))}_{O(1)}\\ &\ll \underbrace{x^1\exp\left(-\frac{\log x}{\log z}\right)}_{\ll x^1}\underbrace{\exp(\log\log z)}_{\log z}\end{aligned}$$

Hence proved! □

THEOREM 5.5. $\phi(x,z) = x\prod_{p<z}\left(1-\frac{1}{p}\right) + O\left(x(\log z)^2\exp\left(-\frac{\log x}{\log z}\right)\right)$

PROOF.

$$\phi(x,z) = x\sum_{\substack{d\le x\\ d\mid P(z)}}\frac{\mu(d)}{d} + \underbrace{O(\Psi(x,z))}_{\ll x\log z\exp\left(-\frac{\log x}{\log z}\right)}$$

$$\sum_{\substack{d\le x\\ d\mid P(z)}}\frac{\mu(d)}{d} = \underbrace{\sum_{d\mid P(z)}\frac{\mu(d)}{d}}_{=\prod_{p<z}\left(1-\frac{1}{p}\right)} - \sum_{\substack{d>x\\ d\mid P(z)}}\frac{\mu(d)}{d}$$

We have that

$$a_d = \begin{cases}1 & d \mid P(z)\\ 0 & \text{otherwise}\end{cases}$$

$$A(t) = \sum_{d\mid P(z)//d\le z} 1 \ll \Psi(t,z)$$

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

Hence, we get

$$\left|\sum_{\substack{d>x\\ d\mid P(z)}}\frac{\mu(d)}{d}\right| \le \sum_{\substack{d>x\\ d\mid P(z)}}\frac{1}{d} = \frac{\Psi(t,z)}{t}\Bigg|_x^\infty + \int_x^\infty \frac{\Psi(t,x)}{t^2}\,dt$$

Hence proved! □

THEOREM 5.6 (MERTEN'S THEOREM). *Let γ be the the Euler-Mascheroni constant, where*

$$\gamma = \lim_{n\to\infty}\left(-\ln n + \sum_{k=1}^{n}\frac{1}{k}\right) = \int_1^\infty\left(\frac{1}{\lfloor x\rfloor} - \frac{1}{x}\right)dx \approx 0.577216.$$

Then,

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

PROOF. Using the fact that $1+x \le e^x$, we expand

$$\begin{aligned}\prod_{p<z}\left(1-\frac{1}{p}\right) \le \prod_{p<z}\exp\left(-\frac{1}{p}\right) &= \exp\left(-\sum_{p<z}\frac{1}{p}\right)\\ &= \exp(-\log\log z + O(1))\\ &= O\left(\frac{1}{\log z}\right)\end{aligned}$$

Taking $\log z = \frac{\log x}{A \log\log x}$, where A is a arbitrary large-enough constant. So, we have

$$x \prod_{p<z} \left(1 - \frac{1}{p}\right) \ll \frac{Ax \log\log x}{\log x}$$

Therefore, we get

$$O\left(x(\log z)^2 \exp\left(-\frac{\log x}{\log z}\right)\right) \ll x \left(\frac{\log x}{A \log\log x}\right)^2 \exp\left(\underbrace{-A \log\log x}_{\log((\log x)^{-A})}\right)$$
$$\ll \frac{x}{A(\log x)^{A-2}(\log\log x)^2}$$

So done! □

Let $\mathcal{A}$ be a set of integers $\leq x$, $\mathcal{P}$ a set of primes, and $P(z) = \prod_{\substack{p \in \mathcal{P} \\ p<z}} p$. For each prime $p \in \mathcal{P}$, let $A_p \subset \mathcal{A}$ be a subset of integers belonging to $\omega(p)$ distinct residue classes modulo p. Define $S(\mathcal{A}, \mathcal{P}, z) = \#\left(A \setminus \bigcup_{p|P(z)} \mathcal{A}_p\right)$. For example, suppose $\mathcal{A} = \{n \leq x\}$, then $\mathcal{A}_p = \{n \leq x : p \mid n\}$. Now, suppose d is a squarefree number divisible by primes of $\mathcal{P}$, define $\omega(d) = \prod_{p|d} \omega(p)$ and $\mathcal{A}_d = \bigcap_{p|d} \mathcal{A}_p$.

Here is an example (idea): Suppose, we have

$$S(\mathcal{A}, \mathcal{P}, z) = \sum_{a \in \mathcal{A}} \left(\sum_{d|(n, P(z))} \mu(d) \right)$$
$$= \sum_{\substack{d \leq x \\ d|P(z)}} \mu(d) \underbrace{\left(\sum_{\substack{a \in \mathcal{A} \\ d|n}} 1 \right)}_{\#\mathcal{A}_d}$$

In general,

$$S(\mathcal{P}, \mathcal{A}, z) = \sum_{\substack{d \leq x \\ d|P(z)}} \mu(d) \#\mathcal{A}_d \rightsquigarrow \text{like } \phi(x, z) = \sum_{\substack{d \leq x \\ d|P(z)}} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor$$

THEOREM 5.7 (THE SIEVE OF ERATOSTHENES). *Suppose the following conditions hold:*

(a) *There exists an X such that* $\#\mathcal{A}_d = \frac{\omega(d)}{d} X + O(\omega(d)) \rightsquigarrow$ *like* $\left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} + O(1)$.

(b) *For some $\kappa \geq 0$,* $\sum_{p|P(z)} \frac{\omega(p) \log p}{p} \leq \kappa \log z + O(1)$.

(c) *For some $y > 0$, $\#\mathcal{A}_d = 0$ 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 $W(z) = \prod_{\substack{p \in \mathcal{P} \\ p<z}} \left(1 - \frac{\omega(p)}{p}\right)$.

Instead of $\sum_{\substack{d \le x \\ d \mid P(z)}} 1 \ll \Psi(x, z)$, use $\sum_{\substack{d \le x \\ d \mid P(z)}} \omega(d) := F_\omega(x, z)$.

6 Brun's Sieves

Brun's sieve is set up in essentially the same way as Eratosthenes. Given some set $\mathcal{A}$ of integers $\le x$, we have some collection of $\mathcal{A}_p$ of elements we want to remove, and measure the size of

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

The idea/punchline behind Brun's results is that under similar, but slightly relaxed, hypotheses to the sieve of Eratosthenes, Burn proves that

$$S(\mathcal{A}, \mathcal{P}, z) = XW(z) + O(\text{better error})$$

where $X = \#\mathcal{A}$ and $W(z) = \prod_{p \mid P(z)} \left(1 - \frac{\omega(p)}{p}\right)$.

THEOREM 6.1. *Suppose the following conditions hold:*

(a) *There exists an X such that $\#\mathcal{A}_d = \frac{\omega(d)}{d} X + O(\omega(d))$ $\rightsquigarrow$ like $\left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} + O(1)$.*

(b) *There exists a constant C such that $\omega(p) < C$.*

(c) *There exist constants C_1 and C_2 such that $\sum_{p \mid P(z)} \frac{\omega(p)}{p} \le C_1 \log\log z + C_2$.*

Then,

$$S(\mathcal{A}, \mathcal{P}, z) = \underbrace{XW(z)}_{\text{main term}} + \underbrace{XW(z)O((\log z)^{-\eta \log \eta}) + O(z^{\eta \log\log z})}_{\text{error terms}}$$

where η is any positive number (possibly depending on x and z).

Previously, we saw that $\log z = \dfrac{\log x}{A \log\log x}$. So, we have

$$\begin{aligned} z^{\eta \log\log z} &= \exp(\eta \log z \log\log z) \\ &= \exp\left(\frac{\eta \log x}{A \log\log x}(\log\log x - \log A - \log\log\log x)\right) \\ &= X^{\frac{\eta}{A} - \frac{\eta \log A}{A \log\log x} - \frac{\eta \log\log\log x}{A \log\log x}} \ll X^{\frac{\eta}{A}} \end{aligned}$$

small if, A is large.

THEOREM 6.2. *$\#\{p \le x : p \text{ and } p+2 \text{ are prime}\} \ll \frac{x(\log\log x)^2}{(\log x)^2}$*

We can observe that the above bound on the twin-primes is comparitively between than the bound on the number of primes, i.e., compared with $\pi(x) \sim \frac{x}{\log x}$

PROOF. Let $\mathcal{A}_p = \{n \le x : \underbrace{n \equiv 0 \pmod p}_{p|n} \text{ or } \underbrace{n \equiv -2 \pmod p}_{p|n+2}\}$. Using, our extimation for $\log z$, i.e., $\dfrac{\log x}{A \log\log x}$, implying $z = \exp\left(\dfrac{\log x}{A \log\log x}\right)$ and using the fact that $1 + x \le e^x$, we get

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, z) &\ll x \prod_{p<z}\left(1 - \frac{2}{p}\right) \\ &\ll x \exp\left(-2\sum_{p<z}\frac{1}{p}\right) \\ &\ll \frac{x}{(\log z)^2} = \frac{xA^2(\log\log x)^2}{(\log x)^2} \end{aligned}$$

$$\#\{p \le x : p, p+2 \text{ prime}\} \le S(\mathcal{A}, \mathcal{P}, z) + \pi(z) \ll \underbrace{\frac{x(\log\log x)^2}{(\log x)^2}}_{\text{main term}} + \underbrace{x^{\frac{1}{A\log x}} \ll x^{\frac{1}{2}}}_{\text{error term}}$$

□

COROLLARY 6.3. $\displaystyle\sum_{\substack{p \\ p+2 prime}} \frac{1}{p} < \infty$

PROOF. We know that $\displaystyle\sum_p \frac{1}{p}$ diverges. By Abel summation, we have

$$a_n = \begin{cases} 1 & n, n+2 \text{ prime} \\ 0 & \text{else} \end{cases}$$

$$A(t) = \#\{p \le x : p, p+2 \text{ prime}\}$$

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

Hence, we get

$$\begin{aligned}\sum_{\substack{p>1.5\\ p+2\text{prime}}} \frac{1}{p} &= \frac{A(t)}{t}\Big|_{1.5}^{\infty} + \int_{1.5}^{\infty} \frac{A(t)}{t^2}\,dt \\ &\ll \lim_{b\to\infty} \frac{(\log\log b)^2}{(\log b)^2} + O(1) + \int_{1.5}^{\infty} \frac{(\log\log t)^2}{t(\log t)^2}\,dt \\ &\ll 0 + O(1) + \int_{\log 1.5}^{\infty} \frac{(\log u)^2}{u^2}\,du\end{aligned}$$

so by taking $u = \log t$ and $du = \frac{dt}{t}$, we can observe that it converges by comparison with $\int \frac{1}{u^{\frac{3}{2}}}\,du$. □

We will now discuss about some ideas behind Brun's sieves. We are not going to prove the entire thing, but we will discuss why these techniques give a better bound.

The big idea being discussed now is the ***truncated Möbius Inversion***.

LEMMA 6.4. *Let $n, r \in \mathbb{Z}^+$ with $r \le \nu(n) = \#\{\text{distinct prime divisors of } n\}$. There exists $|\theta| \le 1$ such that,*

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

before proving the above claim, we will discuss our observations,

- $\sum_{d|P(z)}$ has 2^z terms.
- $\sum_{\substack{d\le x\\ d|P(z)}}$ has $\ll x\log z \exp\left(-\frac{\log x}{\log z}\right)$ terms.
- $\sum_{\substack{d|P(z)\\ \nu(d)\le r}}$ has $\ll 2^r$ terms.

PROOF. Let $n, r \in \mathbb{Z}^+$ with $r \leq \nu(n) = \#\{\text{distinct prime divisors of } n\}$.

$$\begin{aligned}
\sum_{\substack{d|n \\ \nu(d)\leq r}} \mu(d) &= 1 + \sum_{p|n}(-1) + \sum_{p_1p_2|n}(-1)^2 + \cdots + \sum_{p_1\cdots p_r|n}(-1)^r \\
&= 1 + \binom{\nu(n)}{1}(-1) + \binom{\nu(n)}{2}(-1)^2 + \cdots + \binom{\nu(n)}{r}(-1)^r \\
&= \sum_{k=0}^{r}\binom{\nu(n)}{k}(-1)^k \\
&= \binom{\nu(n)-1}{r}(-1)^r \\
\sum_{\substack{d|n \\ \nu(d)\leq 2r+1}} \mu(d) \leq \sum_{d|n}\mu(d) &\leq \sum_{\substack{d|n \\ \nu(d)\leq 2r}} \mu(d) \\
\sum_{\substack{d|n \\ \nu(d)=2r+1}} \mu(d) \leq \sum_{d|n}\mu(d) - \sum_{\substack{d|n \\ \nu(d)\leq 2r}} \mu(d) &\leq 0
\end{aligned}$$

Choosing $r = \lfloor \eta \log\log x \rfloor$, we get

$$\begin{aligned}
S(\mathcal{A}, \mathcal{P}, z) &= \sum_{a\in\mathcal{A}}\left(\sum_{d|(a,P(z))}\mu(d)\right) \\
&= \sum_{a\in\mathcal{A}}\left(\sum_{\substack{d|(a,P(z)) \\ \nu(d)\leq r}}\mu(d) + \theta\sum_{\substack{d|(a,P(z)) \\ \nu(d)=r+1}}\mu(d)\right) \\
&= \sum_{\substack{d|P(z) \\ \nu(d)\leq r}}\mu(d)\#\mathcal{A}_d + O\left(X\frac{\pi(z)^{r+1}}{(r+1)!}\right)
\end{aligned}$$

where $X = \#A$ and we chose $\#d \mid P(z)$ and $\nu(d) = r+1$, i.e., choosing $r+1$ primes from $\pi(z)$ primes. Now, we have

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

The next big idea we are discuss in order to improve our bound is ***replacing Möbius sums with an approximation***.

$$\sum_{d|n}\mu(d) \longleftrightarrow \sum_{d|n}\mu(d)g(d)$$

Strategic choices of "lower" and "upper" weight functions give bounds

$$\sum_{d|P(z)}\mu(d)g_L(d)\#\mathcal{A}_d \leq S(\mathcal{A}, \mathcal{P}, z) \leq \sum_{d|P(z)}\mu(d)g_U(d)\#\mathcal{A}_d$$

which are easier to count. So, we get the following theorem, based on this big idea.

THEOREM 6.5 (BRUN'S MAIN THEOREM). *There exist constants c_1 and c_2 such that*

$$S(\mathcal{A}, \mathcal{P}, z) \leq c_1 XW(z) + O(z^{\theta})$$

and

$$S(\mathcal{A}, \mathcal{P}, z) \geq c_2 XW(z) + O(z^{\theta-1})$$

where θ is given explicitely.

We can can choose $z = x^{\frac{1}{\theta}-\epsilon}$. Earlier, we saw that $\log z = \dfrac{\log x}{A \log\log x}$, implies $z = x^{\frac{1}{A \log\log x}}$

7 References

[1] Introduction to Analytic Number Theory Theory by Tom M. Apostol
[2] An Introduction to Sieve Methods and Their Applications by Murty and Cojocaru
[3] Sieves in Number Theory by Greaves
[4] Opera De Cribro by Friedlander and Iwaniec
[5] Sieve Methods by Halberstam and Richert