2017

An Exposition of Selberg's Sieve

Jack Dalton
University of Vermont

An Exposition of Selberg's Sieve

A Thesis Presented

by

Jack Robert Dalton

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
Specializing in Mathematics

May, 2017

Defense Date: March 23rd, 2017
Thesis Examination Committee:

Jonathan Sands, Ph.D., Advisor
Christian Skalka, Ph.D., Chairperson
Richard Foote, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of Graduate College

Abstract

A number of exciting recent developments in the field of sieve theory have been done concerning bounded gaps between prime numbers. One of the main techniques used in these papers is a modified version of Selberg's Sieve from the 1940's. While there are a number of sources that explain the original sieve, most, if not all, are quite inaccessible to those without significant experience in analytic number theory. The goal of this exposition is to change that. The statement and proof of the general form of Selberg's sieve is, by itself, difficult to understand and appreciate. For this reason, the inital exposition herein will be about one particular application: to recover Chebysheff's upper bound on the order of magnitude of the number of primes less than a given number. As Selberg's sieve follows some of the same initial steps as the more elementary sieve of Eratosthenes, this latter sieve will be worked through as well.

To help the reader get a better sense of Selberg's sieve, a few particular applications are worked through, including an upper bound on the number of twin primes less than a number. This will then be used to show the convergence of the reciprocals of the twin primes.

Acknowledgements

I am extremely grateful for the guidance of my advisor, Professor Jonathan Sands, who helped me along every step of the way. I would also like to thank Professor Richard Foote for help with understanding the prime number theorem.

Additionally, a special thanks goes out to my fellow graduate students, Sophie Gonet, Ryan Grindle, Kewang Chen, and Ada Morse for their advice, for being open to bouncing ideas off of, and for their invaluable help with formatting.

Table of Contents

List of Notational Conventions

$\mathbb{N}$: The set of all natural numbers: $\{1, 2, \ldots\}$.
$\mathbb{R}$: The set of all real numbers.
$\mathcal{A}$: Some specified set of integers to be sifted.
A_d : $\{a_n \in \mathcal{A} : d \mid a_n\}$.
$\mathcal{P}$: Some specified set of primes which will be used for sifting.
$P(z)$: The product of all primes in $\mathcal{P}$ less than z.
p_i : The ith smallest prime number in some list of primes.
x : The real number used as the upper bound for the sieve.
z : The upper bound on the primes in $\mathcal{P}$ used for sifting.
$S(\mathcal{A}, P(z))$: The number of integers in $\mathcal{A}$ that are relatively prime to $P(z)$.
$\pi(x)$: The number of primes less than x.
$\pi_2(x)$: The number of twin primes less than x.
$\pi(x; H)$: The number of primes in the interval $(x, x+H]$.
$\lfloor x \rfloor$: The greatest integer less than x.
$\nu(n)$: The number of distinct prime divisors of n.
$\phi(n)$: Euler's totient function $= |\{k \in \mathbb{N} : k \leq n, (n,k) = 1\}|$.
$f(n) = \mathcal{O}(g(n))$: $\exists K$ constant such that $\forall n, |f(n)| \leq K g(n)$.
$f(n) \ll g(n)$: $f(n) = \mathcal{O}(g(n))$.
$\log(x)$: The natural logarithm of x.
$d \mid n$: d divides $n \Rightarrow \exists m \in \mathbb{N}$ such that $dm = n$.
(x, y) : The greatest common divisor of the integers x and y.
$[x, y]$: The least common multiple of the integers x and y.
$\{x\}$: The fractional part of $x = x - \lfloor x \rfloor$.
$f(x) \sim g(x)$: $\lim\limits_{x \to \infty} \dfrac{f(x)}{g(x)} = 1$.
$\chi[\text{statement}]$: 1 if the statement is true, 0 if false.
$\sum\limits_{d \mid n}$: Sum over all positive divisors d of some fixed natural number n.
$\tau(n)$: The number of divisors of n.
$\tau_{\text{odd}}(n)$: The number of odd divisors of n.
$\Omega_{\text{odd}}(n)$: The number of odd prime powers dividing n.

Chapter 1

Introduction

Recent developments in the field of sieve theory have led to some exciting results concerning positive lower bounds on the number of primes in bounded intervals. Arguably the most exciting development was the discovery by Yitang Zhang in 2013 that there exists a bound $h_k \leq 70,000,000$ which has 2 prime numbers in the interval $[n, n+h_k]$ for an infinite number of different values of n. Since then, this h_k bound has been reduced to 246 by the Polymath Project, a collaborative effort of mathematicians all over the world. These works built partly on the work of Goldston, Pintz and Yıldırım, who proved in 2005 that

$$\liminf_n \frac{p_{n+1} - p_n}{\log(p_n)} = 0.$$

One underlying tool used in a number of these works is a suitably altered version of Atle Selberg's Λ^2 Sieve from the 1940s. While there are a number of sources available to read about Selberg's Sieve, none are accessible without much study of analytic number theory. The goal of this exposition is to change that.

We will start, as is classically done, with a presentation of the sieve of Eratosthenes: an ancient method that basically uses inclusion/exclusion. Legendre formalized the underlying mathematics, and as such we will refer to the sieve of Eratosthenes/Legendre. Using this sieve, we will work out the somewhat crude upper bound on the number of primes less than a given real number:

$$\pi(x) \ll \frac{x}{\log(\log(x))}.$$

Selberg's Sieve initally follows the same methods as that of the sieve of Eratosthenes/Legendre, but replaces the use of the Möbius function with a sequence of real numbers, which lead to a similar conclusion with a much improved error term. This method will recover Chebycheff's upper bound on the number of primes less than a given number:

$$\pi(x) \ll \frac{x}{\log(x)}.$$

The proof of this specific statement from basic assumptions mirrors the proof of the general form of Selberg's Sieve in such a way that it would be repetitive to work through the latter proof as well. However, we will state the general theorem, as it will be useful in applications.

A pair of additional applications will then be worked through. An upper bound on the number of primes in a given interval turns out to be a direct application with little modifications. An upper bound on the number of twin primes less than a given number requires a couple extra lemmas, but enlightens the reader on the potential applications of Selberg's Sieve.

A prime number p is called a twin prime if either $p+2$ or $p-2$ is also prime.

For example, 5 and 7 are both twin primes. The Twin Prime Conjecture states that there are infinitely many twins primes and has eluded mathematicians for centuries. However an upper bound on the number of twin primes less than a given x is worked through with sieves in the applications chapter. This upper bound will then be used to recover Brun's theorem that the sum of the reciprocals of the twin primes converges. Brun proved this in 1915 with a slightly more crude upper bound which makes the proof here more immediate.

Chapter 2

Preliminaries

This section will be statements and proof of some preliminary lemmas used in the later proofs to make the latter more readable.

Lemma.

$$\sum_{d|n} \mu(d) = \chi[n = 1].$$

Proof. If $n = 1$, then since $\mu(1) = 1$, we have that $\sum_{d|n} \mu(d) = 1$. Now if $n > 1$, let $n = p_1^{a_1} p_2^{a_2} \dots p_r^{a_r}$ be the unique factorization of n into distinct primes powers. The terms in the sum $\sum_{d|n} \mu(d)$ with any power of a prime greater than 1 will be zero, so we have that the sum will be equal if taken over divisors of n or divisors of $N = p_1 p_2 \dots p_r$. Thus $\sum_{d|n} \mu(d) = \sum_{d|N} \mu(d)$. Each unique divisor d of N is the product of elements of a unique subset of $\{p_1, p_2, \dots, p_r\}$, so if we count how many elements are in each subset, i.e., how many primes divide d, we get that there are $\binom{r}{k}$ divisors whose prime factorization has k primes in it. The value of $\mu(d)$ for this d with k primes dividing it will be $(-1)^k$. Therefore the sum can be written as $\sum_{d|N} \mu(d) = \sum_{k=0}^{r} \binom{r}{k} (-1)^k = \sum_{k=0}^{r} \binom{r}{k} (-1)^k (1)^{r-k}$. Applying the binomial theorem to this last sum, we get $(1-1)^r = 0$. □

Lemma.

$$\sum_{d|P(z)} 1 = 2^{\pi(z)}.$$

Proof. The left hand side is counting the number of possible subsets of $\mathcal{P}$, which is just $2^{|\mathcal{P}|} = 2^{\pi(z)}$. □

Lemma. *Möbius inversion formula: Let $f, g : \mathbb{N} \to \mathbb{C}$. Then we have*

$$f(n) = \sum_{d|n} g(d) \iff g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right).$$

Proof. If $f(n) = \sum_{d|n} g(d)$, then

$$\begin{aligned}
\sum_{d|n} \mu(d) f\left(\frac{n}{d}\right) &= \sum_{d|n} \mu(d) \sum_{e|\frac{n}{d}} g(e) \\
&= \sum_{d} \chi[d \mid n] \mu(d) \sum_{e} \chi\left[e \mid \frac{n}{d}\right] g(e) \\
&= \sum_{d,e} \mu(d) g(e) \chi[d \mid n] \chi\left[e \mid \frac{n}{d}\right] \\
&= \sum_{d,e} \mu(d) g(e) \chi[d \mid n] \chi\left[de \mid n\right] \\
&= \sum_{e} g(e) \sum_{d} \mu(d) \chi\left[d \mid \frac{n}{e}\right] \chi\left[e \mid n\right] \\
&= \sum_{e} \chi\left[e \mid n\right] g(e) \sum_{d} \mu(d) \chi\left[d \mid \frac{n}{e}\right] \\
&= \sum_{e|n} g(e) \sum_{d|\frac{n}{e}} \mu(d) \\
&= \sum_{e|n} g(e) \chi[n = e] = g(n).
\end{aligned}$$

For the other direction, assume $g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right)$. Then

$$\begin{aligned}
\sum_{d|n} g(d) &= \sum_{d|n} \sum_{e|d} \mu(e) f\left(\frac{d}{e}\right) \\
&= \sum_{est=n} \mu(e) f(s) \\
&= \sum_{s|n} f(s) \sum_{e|\frac{n}{s}} \mu(e) \\
&= f(n).
\end{aligned}$$

□

Lemma.

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

Proof. Euler's Totent Function, $\phi(n)$, is defined as the number of integers less than n that are relatively prime to n. Clearly, on the primes, $\phi(p) = p - 1$. Since ϕ is multiplicative, this can be extended to show that, for any $n \in \mathbb{N}$, $\phi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$.

Now let S_d be the set of integers less than n that have greatest common divisor with n of d, i.e., $S_d = \{n \in \mathbb{N} : 1 \leq m \leq n, (m, n) = d\}$. Using the equivalence that $(m, n) = d \iff d \mid m$ and n, and $\left(\frac{m}{d}, \frac{n}{d}\right) = 1$, we have

$$|S_d| = |\{k \leq \frac{n}{d} : (k, \frac{n}{d}) = 1\}| = \phi(\frac{n}{d}).$$

Thus, $\forall m \leq n, \exists d$ such that $d \mid n$ and $m \in S_d$. Therefore, $\{1, \ldots, n\} = \bigcup_{d|n} S_d$. Now since all the S_d's are pairwise disjoint, taking the order of each side of this last set equation becomes $n = \sum_{d|n} |S_d| = \sum_{d|n} \phi(\frac{n}{d}) = \sum_{d|n} \phi(d)$.

Using Möbius inversion on this formula yields $\phi(n) = \sum\limits_{d|n} \mu(d)\frac{n}{d}$. Plugging in the above formula for $\phi(n)$ and using $n = P(z)$, the product of all primes less than z, we have $\sum\limits_{d|P(z)} \frac{\mu(d)}{d} = \prod\limits_{p<z}\left(1 - \frac{1}{p}\right)$, as desired. $\square$

Lemma. *The dual Möbius inversion formula states that if $\mathcal{D}$ is a divisor-closed subset of $\mathbb{N}$, and $f, g : \mathbb{N} \to \mathbb{C}$, then*

$$f(n) = \sum_{\substack{n|d \\ d\in\mathcal{D}}} g(d)$$

holds if and only if

$$g(n) = \sum_{\substack{n|d \\ d\in\mathcal{D}}} \mu\left(\frac{d}{n}\right) f(d),$$

assuming all the series are absolutely convergent.

Proof. First assume $f(n) = \sum\limits_{\substack{n|d \\ d\in\mathcal{D}}} g(d)$, then

$$\begin{aligned}
\sum_{\substack{n|d \\ d\in\mathcal{D}}} \mu\left(\frac{d}{n}\right) f(d) &= \sum_{\substack{n|d \\ d\in\mathcal{D}}} \mu\left(\frac{d}{n}\right) \sum_{\substack{d|e \\ e\in\mathcal{D}}} g(e) \\
&= \sum_{d\in\mathcal{D}} \mu\left(\frac{d}{n}\right) \chi[n \mid d] \sum_{e\in\mathcal{D}} g(e)\chi[d \mid e] \\
&= \sum_{e\in\mathcal{D}} g(e) \sum_{d\in\mathcal{D}} \mu\left(\frac{d}{n}\right) \chi[n \mid d]\chi[d \mid e], \qquad\qquad \text{let } m = \frac{d}{n} \\
&= \sum_{e\in\mathcal{D}} g(e) \sum_{mn\in\mathcal{D}} \mu(m)\chi[n|mn]\chi[mn|e] \\
&= \sum_{e\in\mathcal{D}} g(e) \sum_{m|\frac{e}{n}} \mu(m) \\
&= \sum_{e\in\mathcal{D}} g(e)\chi[e = n] \\
&= g(n).
\end{aligned}$$

For the other direction, assume $g(n) = \sum_{\substack{n|d \\ d\in\mathcal{D}}} \mu\left(\frac{d}{n}\right) f(d)$, then

$$\begin{aligned}
\sum_{\substack{n|d \\ d\in\mathcal{D}}} g(d) &= \sum_{\substack{n|d \\ d\in\mathcal{D}}} \sum_{\substack{d|e \\ e\in\mathcal{D}}} \mu\left(\frac{e}{d}\right) f(e) \\
&= \sum_{d\in\mathcal{D}} \sum_{e\in\mathcal{D}} \mu\left(\frac{e}{d}\right) f(e)\chi[n \mid d]\chi[d \mid e] \\
&= \sum_{e\in\mathcal{D}} f(e) \sum_{d\in\mathcal{D}} \mu\left(\frac{e}{d}\right) \chi[n \mid d]\chi[d \mid e] && \text{let } m = \frac{e}{d} \\
&= \sum_{e\in\mathcal{D}} f(e) \sum_{\frac{e}{m}\in\mathcal{D}} \mu(m)\chi\left[n \mid \frac{e}{m}\right] \chi[d \mid md] \\
&= \sum_{e\in\mathcal{D}} f(e) \sum_{\substack{m|\frac{e}{n} \\ \frac{e}{n}\in\mathcal{D}}} \mu(m) \\
&= \sum_{e\in\mathcal{D}} f(e)\chi[e = n] \\
&= f(n).
\end{aligned}$$

$\square$

Lemma. *Partial Summation.*

If $\{a_n\}_1^\infty \subseteq \mathbb{C}$ and if $f : \mathbb{N} \to \mathbb{C}$ is continuously differentiable and we define:

$$S(x) = \sum_{1\leq n\leq x} a_n,$$

then, $\forall A, B \in \mathbb{N}$ with $A < B$, we have:

$$\sum_{A<n\leq B} a_n f(n) = f(B)S(B) - f(A)S(A) - \int_A^B S(x)f'(x)dx.$$

Proof. We have

$$\begin{aligned}\sum_{A<n\le B} a_n f(n) &= \sum_{A<n\le B} f(n)(S(n)-S(n-1)) \\ &= \sum_{A<n\le B} f(n)S(n) - \sum_{A-1<n\le B-1} f(n+1)S(n) \\ &= f(B)S(B) - f(A)S(A) - \sum_{A-1<n\le B-1} S(n)(f(n+1)-f(n)).\end{aligned}$$

We evaluate the sum in this last line:

$$\begin{aligned}\sum_{A-1<n\le B-1} S(n)(f(n+1)-f(n)) &= \sum_{n=A}^{B-1} S(n) \int_n^{n+1} f'(x)dx \\ &= \sum_{n=A}^{B-1} \int_n^{n+1} S(x)f'(x)dx \\ &= \int_A^B S(x)f'(x)dx\end{aligned}$$

which completes the proof. $\square$

Chapter 3

The Sieve of Eratosthenes/Legendre

The goal of this section is to estimate $\pi(x)$, the number of primes less than a number x, using the simplest sieve, that of Eratosthenes. In general, sieves are used over some specified set $\mathcal{A}$. So for this example, $\mathcal{A} = \{n \in \mathbb{Z} : 1 \leq n \leq x\}$, the set of all positive integers up to x. In general, once again, we will need some way of sifting out the primes, so to start we will consider how many integers in $\mathcal{A}$ are relatively prime to some set of primes $\mathcal{P}$. In this case $\mathcal{P}$ will just be all the prime integers. One more piece we will need in general is a way of measuring the number of integers in $\mathcal{A}$ that are relatively prime to the primes in $\mathcal{P}$ that are less than some number z. So we form the product $P(z) = \prod_{p<z} p$ and consider the count of all integers in $\mathcal{A}$ that are relatively prime to $P(z)$, which we will denote $S(\mathcal{A}, P(z))$. So we have $S(\mathcal{A}, P(z)) = |\{n \in \mathbb{N} : 1 \leq n \leq x, (n, P(z)) = 1\}|$. So $S(\mathcal{A}, P(z))$ is the number of integers in $\mathcal{A}$ that are relatively prime to $P(z)$. In other words, it is the number of integers between 1 and x that have no prime factors smaller than z. To think of it another way, $S(A, P(z)) = \sum_{n \leq x} \chi[(n, P(z)) = 1]$. If everything works perfectly, the number we get from the sieve will be exactly $\pi(x) - \pi(z) + 1$, the number of

primes less than x but greater than z, where the 1 comes from the fact that for all n, $(n, 1) = 1$. This will then be used to get an upper bound on $\pi(x)$.

If we make the simplifying assumption that $x \in \mathbb{R}^+ \setminus \mathbb{N}$, then the statement $n < x$ is equivalent to $n \leq x$, and the same can be done for z. This will make it easier to keep track of later calculations.

So now to start counting the elements of $\mathcal{A}$ divisible by small primes in $\mathcal{P}$ starting with the smallest prime and working up towards the greatest. Clearly, there are $\lfloor x \rfloor$ integers less than x: $1, 2, \ldots, \lfloor x \rfloor$; so the count before any multiples of primes are removed will be $\lfloor x \rfloor$. Next, for each prime p_1 less than z, we "sift out" multiples of p_1. So we remove each $mp_1 \leq x$, where $m \in \mathbb{N}$. Clearly, we are deleting $\left\lfloor \frac{x}{p_1} \right\rfloor$ numbers from the list. This leaves us with $\lfloor x \rfloor - \sum_{p<z} \left\lfloor \frac{x}{p} \right\rfloor$ integers remaining in the list. Upon brief inspection, we notice that too many integers were deleted. For example, integers that are divisible by two primes less than z were counted twice. So we can add those back in, for which we have $\left\lfloor \frac{x}{p_1 p_2} \right\rfloor$ for each pair p_1, p_2. The new count is $\lfloor x \rfloor - \sum_{p_1<z} \left\lfloor \frac{x}{p_1} \right\rfloor + \sum_{p_2<p_1<z} \lfloor \frac{x}{p_1 p_2} \rfloor$. If this process is iterated, we get the following count for $S(\mathcal{A}, P) = \lfloor x \rfloor - \sum_{p_1<z} \left\lfloor \frac{x}{p_1} \right\rfloor + \sum_{p_2<p_1<z} \left\lfloor \frac{x}{p_1 p_2} \right\rfloor - \sum_{p_3<p_2<p_1<z} \left\lfloor \frac{x}{p_1 p_2 p_3} \right\rfloor + \ldots$

For example, looking at the integers less than $x = 25.9$, we have that $\lfloor 25.9 \rfloor = 25$. Below is a list of these integers with the primes underlined:

$$1, \underline{2}, \underline{3}, 4, \underline{5}, 6, \underline{7}, 8, 9, 10, \underline{11}, 12, \underline{13}, 14, 15, 16, \underline{17}, 18, \underline{19}, 20, 21, 22, \underline{23}, 24, 25.$$

So $\pi(x) = 9$. There are $\left\lfloor \frac{25.9}{2} \right\rfloor = 12$ different multiples of 2 in the list, $\left\lfloor \frac{25.9}{3} \right\rfloor = 8$ different multiples of 3, and $\left\lfloor \frac{25.9}{5} \right\rfloor = 5$ different multiples of 5. If we keep counting multiples of higher primes, we will be just counting the single primes, as there will be no composite numbers left in the list once we count multiples of $\sqrt{\lfloor 25.9 \rfloor} = \sqrt{25} = 5$.

In other words, each composite between $\sqrt{x}$ and x is divisible by a prime less than $\sqrt{x}$. Thus, the optimal choice for z in the sieve is $\sqrt{x}$. Using this value in the sum for $S(\mathcal{A}, P(z))$, $\lfloor x \rfloor - \sum_{p<\sqrt{x}} \left\lfloor \frac{x}{p} \right\rfloor = 25 - (12 + 8 + 5) = 0$. Clearly, too many numbers were deleted, as we can see that there are 6 primes between $\sqrt{x}$ and x. So adding the sum over divisors made up of 2 primes, we have $\left\lfloor \frac{25.9}{2\cdot 3} \right\rfloor = 4$, $\left\lfloor \frac{25.9}{2\cdot 5} \right\rfloor = 2$, and lastly $\left\lfloor \frac{25.9}{3\cdot 5} \right\rfloor = 1$. Thus $\lfloor x \rfloor - \sum_{p<\sqrt{x}} \left\lfloor \frac{x}{p} \right\rfloor + \sum_{p_2<p_1<z} \left\lfloor \frac{x}{p_1 p_2} \right\rfloor = 25-(12+8+5)+(4+2+1) = 7$. The next sum counts multiples of the product of 3 primes, but the product of the smallest 3 primes, $2 \cdot 3 \cdot 5 = 30$, is greater than $x = 25.9$, so each term in the sum will be zero. Thus $S(\mathcal{A}, P(z))$ should be 7, and upon inspection of the list,
$\pi(25.9) - \pi(\sqrt{25.9}) + 1 = 9 - 3 + 1 = 7$, as desired.

For a slightly larger example, which will be useful later, we will do this same calculation for $x = 100.5$. These are the numbers we will need

$$
\begin{array}{lll}
\lfloor x \rfloor = 100 & \lfloor \frac{100.5}{2\cdot 3} \rfloor = 16 & \lfloor \frac{100.5}{2\cdot 3\cdot 5} \rfloor = 3 \\
\lfloor \frac{100.5}{2} \rfloor = 50 & \lfloor \frac{100.5}{2\cdot 5} \rfloor = 10 & \lfloor \frac{100.5}{2\cdot 3\cdot 7} \rfloor = 2 \\
\lfloor \frac{100.5}{3} \rfloor = 33 & \lfloor \frac{100.5}{2\cdot 7} \rfloor = 7 & \lfloor \frac{100.5}{2\cdot 5\cdot 7} \rfloor = 1 \\
\lfloor \frac{100.5}{5} \rfloor = 20 & \lfloor \frac{100.5}{3\cdot 5} \rfloor = 6 & \lfloor \frac{100.5}{3\cdot 5\cdot 7} \rfloor = 0 \\
\lfloor \frac{100.5}{7} \rfloor = 14 & \lfloor \frac{100.5}{3\cdot 7} \rfloor = 4 & \\
 & \lfloor \frac{100.5}{5\cdot 7} \rfloor = 2. &
\end{array}
$$

So we have that, for $\mathcal{A} = \{n \in \mathbb{N} : n < 100.5\}$,

$$\begin{aligned} S(\mathcal{A}, P(z)) &= \lfloor x \rfloor - \sum_{p_1 < z} \left\lfloor \frac{x}{p_1} \right\rfloor + \sum_{p_2 < p_1 < z} \left\lfloor \frac{x}{p_1 p_2} \right\rfloor - \sum_{p_3 < p_2 < p_1 < z} \left\lfloor \frac{x}{p_1 p_2 p_3} \right\rfloor \\ &= 100 - (50 + 33 + 20 + 14) + (16 + 10 + 7 + 6 + 4 + 2) - (3 + 2 + 1) \\ &= 22. \end{aligned}$$

Checking with the correct number of primes in the interval, we have $\pi(100.5) - \pi(\sqrt{100.5}) + 1 = 25 - 4 + 1 = 22$.

This method works well for these small numbers, but does it lead to any theoretical estimates for large values of x? To do this, we need to define the Möbius function $\mu(n)$. $\mu : \mathbb{N} \to \{0, \pm 1\}$ by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^r & \text{if } n = p_1 p_2 \dots p_r \text{ where } p_1, p_2, \dots, p_r \text{ are all distinct primes} \\ 0 & \text{if } p^2 \mid n \text{ for some prime } p. \end{cases}$$

An alternate definition would be to define $\mu(1) = 1$, and at prime powers p^α to be

$$\mu(p^\alpha) = \begin{cases} -1 & \text{if } \alpha = 1 \\ 0 & \text{if } \alpha \geq 2. \end{cases}$$

and then extend the function multiplicatively to all positive integers. A function on the natural numbers is said to be multiplicative if $f(x \cdot y) = f(x) \cdot f(y)$ whenever $(x, y) = 1$.

The Möbius function has a number of useful properties that will come into play.

One such is when the sum of the Möbius function is taken over all divisors of an integer, the result is 1 if the original integer was 1, and 0 otherwise.

The proof of the following lemma can be found in the preliminaries section:

Lemma.

$$\sum_{d|n} \mu(d) = \chi[n = 1].$$

So we get:

$$\begin{aligned} S(\mathcal{A}, P(z)) &= \lfloor x \rfloor - \sum_{p_1 < z} \left\lfloor \frac{x}{p_1} \right\rfloor + \sum_{p_2 < p_1 < z} \left\lfloor \frac{x}{p_1 p_2} \right\rfloor - \sum_{p_3 < p_2 < p_1 < z} \left\lfloor \frac{x}{p_1 p_2 p_3} \right\rfloor + \dots \\ &= \sum_{d|P(z)} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor . \end{aligned}$$

To give a more formal proof of this fact, notice that

$$\begin{aligned} S(\mathcal{A}, P(z)) &= \sum_{n<x} \chi[(n, P(z) = 1] \\ &= \sum_{n<x} \sum_{d|(n,P(z))} \mu(d) \\ &= \sum_{n<x} \sum_{d|P(z)} \mu(d) \chi[d \mid n] \\ &= \sum_{d|P(z)} \mu(d) \sum_{n<x} \chi[d \mid n] \\ &= \sum_{d|P(z)} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor . \end{aligned}$$

Using this, we will get a general upper bound on $S(\mathcal{A}, P(z))$ and then use it to get a (weak) upper bound on $\pi(x)$ in terms of x alone. Note: $\lfloor x \rfloor = x - \{x\} = x + \mathcal{O}(1)$.

Using this notation, the above equation becomes

$$\begin{aligned}
\sum_{d|P(z)} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor &= \sum_{d|P(z)} \mu(d) \left(\frac{x}{d} - \left\{ \frac{x}{d} \right\} \right) \\
&= \sum_{d|P(z)} \mu(d) \left(\frac{x}{d} + \mathcal{O}(1) \right) \\
&= \sum_{d|P(z)} \mu(d) \frac{x}{d} + \sum_{d|P(z)} \mathcal{O}(1) \\
&= x \sum_{d|P(z)} \frac{\mu(d)}{d} + \mathcal{O}\left(\sum_{d|P(z)} 1 \right).
\end{aligned}$$

Now to state a few lemmas regarding this last equation, which are proven in the preliminaries section:

Lemma.

$$\sum_{d|P(z)} 1 = 2^{\pi(z)}.$$

Lemma. *Möbius inversion formula: Let* $f, g : \mathbb{N} \to \mathbb{C}$*. Then we have*

$$f(n) = \sum_{d|n} g(d) \iff g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d} \right).$$

Lemma.

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

Applying these lemmas, we get

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

The next step is to get an upper bound on $S(\mathcal{A}, P(z))$ by getting a lower bound on the reciprocal of part of the main term:

$$\begin{aligned}
\prod_{p<z}\left(1-\tfrac{1}{p}\right)^{-1} &= \prod_{p<z}\sum_{r=0}^{\infty}\tfrac{1}{p^r} && \text{By the geometric series.}\\
\prod_{p<z}\sum_{r=0}^{\infty}\tfrac{1}{p^r} &> \sum_{n<z}\tfrac{1}{n} && \text{Terms on right are proper subset of terms on left.}\\
\sum_{n<z}\tfrac{1}{n} &> \log(z) && \text{Comparing } \textstyle\int_1^z \frac{1}{x}dx \text{ to the sum.}
\end{aligned}$$

So we get that

$$S(\mathcal{A}, P(z)) < \frac{x}{\log(z)} + \mathcal{O}\left(2^{\pi(z)}\right).$$

Taking $z = \sqrt{x}$ would be ideal, however, for large values of x, by the prime number theorem, $2^{\pi(\sqrt{x})} \sim 2^{\frac{\sqrt{x}}{\log(\sqrt{x})}}$, which is much larger than the main term $\frac{x}{\log(\sqrt{x})}$.

Therefore, to control the error term, we must take $z = \log(x)$. Since $\pi(z) < z$, we have that $2^{\pi(z)} < 2^z = 2^{\log(x)} = x^{\log(2)}$, which is small enough to give us the following upper bound on $S(\mathcal{A}, P(z))$:

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

Thus we can get a bound on $\pi(x)$ using:

$$\begin{aligned}
\pi(x) &= (\pi(x) - \pi(z)) + \pi(z)\\
&\le S(\mathcal{A}, P(z)) + \pi(z)\\
&\le S(\mathcal{A}, P(z)) + z\\
&\ll \frac{x}{\log(\log(x))} + \log(x).
\end{aligned}$$

And therefore

$$\pi(x) \ll \frac{x}{\log(\log(x))}.$$

There is a general form of the sieve of Eratosthenes/Legendre and the proof follows similar arguments to the above, but is much more technical, and for this reason is omitted.

Theorem. *Let $\mathcal{A}$ be any set of natural numbers, and $\mathcal{P}$ be any set of primes. To each of these primes in $\mathcal{P}$, let there be $\omega(p)$ distinguished residue classes modulo p. $\mathcal{A}_p$ is defined to be the set of elements of $\mathcal{A}$ that belong to at least one of these classes. For all squarefree d composed entirely of primes in $\mathcal{P}$, set $\mathcal{A}_1$ to be $\mathcal{A}$, $\mathcal{A}_d = \bigcap_{p|d} \mathcal{A}_p$, and $\omega(d) = \prod_{p|d} \omega(p)$. If $z \in \mathbb{R}^+$, as before, let*

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

Define $S(\mathcal{A}, \mathcal{P}, z) = |\mathcal{A} \setminus \bigcup_{p|P(z)} \mathcal{A}_p|$, and suppose there exists X such that

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

for some R_d.

Now suppose $|R_d| = \mathcal{O}(\omega(d))$ and $\exists \kappa \geq 0$ such that

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

Also assume $\exists y \in \mathbb{R}^+$ such that $\forall d > y$, $|A_d| = 0$.

Then,

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

where

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

So that the theorem is general enough to apply in many situations, rather confusingingly, X is actually a multiplicative function of x. In the application to $\pi(x)$ above, $X(x)$ was just equal to x.

Chapter 4

Selberg's Sieve

Fortunately, the basic ideas of the Sieve of Eratosthenes can be extended to get much better upper bounds on $\pi(x)$. To begin with, we will re-examine the equation

$$S(\mathcal{A}, P(z)) = \sum_{n<x} \chi\left[(n, P(z)) = 1\right] = \sum_{n<x} \sum_{d|(n,P(z))} \mu(d).$$

We will replace the function $\mu(d)$ in the terms in the above sum by a sequence $\{\lambda_d\}$ of real numbers. Now if $\lambda_1 = 1$ and the rest of the λ_d's are arbitrary real numbers, the following will hold: for any fixed $m \in \mathbb{N}$,

$$\sum_{d|m} \mu(d) \leq \left(\sum_{d|m} \lambda_d\right)^2.$$

This fact can be easily verfied by noticing that the left hand side is 1 if $m = 1$ and 0 otherwise, while the right hand side is 1 if $m = 1$ and greater than or equal to zero otherwise. Since the λ_d's were chosen arbitrarily, minimizing equations by carefully choosing values for λ_d will yield valid estimates. So we have

$$\begin{aligned}
S(\mathcal{A}, P(z)) &= \sum_{n<x} \sum_{d|(n,P(z))} \mu(d) \\
&\leq \sum_{n<x} \left(\sum_{d|(n,P(z))} \lambda_d \right)^2 \\
&= \sum_{n<x} \left(\sum_{d_1,d_2|(n,P(z))} \lambda_{d_1}\lambda_{d_2} \right) \\
&= \sum_{n<x} \sum_{d_1,d_2} \lambda_{d_1}\lambda_{d_2} \chi\left[d_1, d_2 \mid (n, P(z))\right] \\
&= \sum_{n<x} \sum_{d_1,d_2} \lambda_{d_1}\lambda_{d_2} \chi\left[d_1 \mid n\right] \chi\left[d_2 \mid n\right] \chi\left[d_1 \mid P(z)\right] \chi\left[d_2 \mid P(z)\right] \\
&= \sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} \sum_{\substack{n<x \\ d_1,d_2|n}} 1.
\end{aligned}$$

Where $[d_1, d_2]$ is the least common multiple of d_1 and d_2. Now if we recall that the number of integers less than x that are divisible by d is

$$\#\{n < x : n \equiv 0(\mod d)\} = \left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} + \mathcal{O}(1),$$

we get

$$S(\mathcal{A}, P(z)) = \underbrace{x \sum_{d_1,d_2|P(z)} \frac{\lambda_{d_1}\lambda_{d_2}}{[d_1, d_2]}}_{\text{Main Term}} + \underbrace{\mathcal{O}\left(\sum_{d_1,d_2|P(z)} |\lambda_{d_1}| \cdot |\lambda_{d_2}| \right)}_{\text{Error Term}}.$$

Further calculations will be simplified by the assumption that $\lambda_d = 0$ for all $d > z$. Note that if in minimizing the main term we get $|\lambda_d| \leq 1$, then the error term will be $\mathcal{O}(z^2)$, which will lead to a more optimal choice of $z = \sqrt{x}$, but this will be clearer further on.

Examining the part of the main term that is controlled by the choice of λ_d's, we want to minimize

$$\sum_{d_1,d_2<z} \frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]}.$$

We will use the fact that the product of the least common multiple and the greatest common divisor of two numbers is the same as the product of the two numbers: i.e., $[d_1,d_2](d_1,d_2) = d_1 \cdot d_2$. Also, we will use the summation function for the sum over the divisors of a number of Euler's totient function, just as we did in the Eratosthenes/Legendre sieve: $n = \sum_{d|n} \phi(d)$.

So we have:

$$\begin{aligned}
\sum_{d_1,d_2<z} \frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} &= \sum_{d_1,d_2<z} \frac{\lambda_{d_1}\lambda_{d_2}}{d_1 \cdot d_2}(d_1,d_2) \\
&= \sum_{d_1,d_2<z} \frac{\lambda_{d_1}\lambda_{d_2}}{d_1 \cdot d_2} \sum_{\delta|(d_1,d_2)} \phi(\delta) \\
&= \sum_{\delta<z} \phi(\delta) \sum_{\substack{d_1,d_2<z \\ \delta|(d_1,d_2)}} \frac{\lambda_{d_1}\lambda_{d_2}}{d_1 \cdot d_2} \\
&= \sum_{\delta<z} \phi(\delta) \left(\sum_{\substack{d<z \\ \delta|d}} \frac{\lambda_d}{d} \right)^2 \\
&= \sum_{\delta<z} \phi(\delta) u_\delta^2
\end{aligned}$$

where we define the function $u_\delta := \sum_{\substack{d<z \\ \delta|d}} \frac{\lambda_d}{d}$. This is the new sum to be minimized subject to $\lambda_1 = 1$ and $\forall d > z, \lambda_d = 0$. Notice that this implies $u_\delta = 0$ for all $\delta > z$.

Using the dual Möbius inversion formula on the formula for u_δ, we get

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

This implies that $\lambda_1 = 1 = \sum_{\delta<z} \mu(\delta)u_\delta$.

The dual Möbius inversion formula states that if $\mathcal{D}$ is a divisor-closed subset of $\mathbb{N}$, and $f, g : \mathbb{N} \to \mathbb{C}$, then

$$f(n) = \sum_{\substack{n|d \\ d\in\mathcal{D}}} g(d)$$

holds if and only if

$$g(n) = \sum_{\substack{n|d \\ d\in\mathcal{D}}} \mu\left(\frac{d}{n}\right) f(d),$$

assuming all the series are absolutely convergent. A set $\mathcal{D}$ is divisor-closed if and only if for all $x \in \mathcal{D}$, if $d \mid x$, then $d \in \mathcal{D}$. See Preliminaries section for proof.

Thus we have,

$$\sum_{\delta<z} \phi(\delta)u_\delta^2 = \sum_{\delta<z} \phi(\delta)\left(u_\delta - \frac{\mu(\delta)}{\phi(\delta)V(z)}\right)^2 + \frac{1}{V(z)},$$

where

$$V(z) = \sum_{d<z} \frac{\mu^2(d)}{\phi(d)} = \sum_{\substack{d<z \\ d \text{ squarefree}}} \frac{1}{\phi(d)}.$$

The minimum of $\sum_{\delta<z} \phi(\delta)u_\delta^2$ will thus be $\frac{1}{V(z)}$, which will occur when $u_\delta = \frac{\mu(\delta)}{\phi(\delta)V(z)}$. An alternative way to see this is by an application of the Cauchy-Schwarz Inequality,

$$\left|\sum_{i=1}^{n} a_i b_i\right| = \left(\sum_{i=1}^{n} |a_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} |b_i|^2\right)^{\frac{1}{2}},$$

to the right hand side of the equation $1 = \sum\limits_{\delta<z} \mu(\delta)u_\delta$. So we have

$$\begin{aligned} 1 &= \sum_{\delta<z} \mu(\delta)u_\delta \frac{\sqrt{\phi(\delta)}}{\sqrt{\phi(\delta)}} \\ &\leq \left(\sum_{\delta<z} \frac{\mu^2(\delta)}{\phi(\delta)}\right)^{\frac{1}{2}} \left(\sum_{\delta<z} \phi(\delta)u_\delta^2\right)^{\frac{1}{2}}. \end{aligned}$$

This implies the same minimum bound stated above

$$\sum_{\delta<z} \phi(\delta)u_\delta^2 \geq \frac{1}{\sum\limits_{\delta<z} \frac{\mu^2(\delta)}{\phi(\delta)}} = \frac{1}{V(z)},$$

which occurs with the same value of u_δ as above which was

$$u_\delta = \frac{\mu(\delta)}{\phi(\delta)V(z)}.$$

Therefore, the optimal choice of λ_δ is

$$\lambda_\delta = \delta \sum_{\substack{d<z \\ \delta|d}} \frac{\mu(d/\delta)\mu(d)}{\phi(d)V(z)}.$$

There will be no terms in the sum for $\delta > z$, and if we plug in $\delta = 1$, we get

$$\lambda_1 = \frac{1}{V(z)} \cdot \sum_{d<z} \frac{\mu^2(d)}{\phi(d)} = 1$$

by the definition of $V(z)$. Thus we get

$$S(\mathcal{A}, P(z)) \leq \frac{x}{V(z)} + \mathcal{O}\left(\sum_{d_1,d_2|P(z)} |\lambda_{d_1}| \cdot |\lambda_{d_2}|\right).$$

To analyze the error term and get it to be the desired $\mathcal{O}(z^2)$, we need to establish that $|\lambda_\delta| \leq 1$ for all δ. If we multiply both sides of $\lambda_\delta = \delta \sum_{\substack{d<z \\ \delta|d}} \frac{\mu(d/\delta)\mu(d)}{\phi(d)V(z)}$ by $V(z)$, we get

$$\begin{aligned} V(z)\lambda_\delta &= \delta \sum_{\substack{d<z \\ \delta|d}} \frac{\mu(d/\delta)\mu(d)}{\phi(d)} \\ &= \delta \sum_{t<\frac{z}{\delta}} \frac{\mu(t)\mu(\delta t)}{\phi(\delta t)} \\ &= \delta \sum_{\substack{t<\frac{z}{\delta} \\ (t,\delta)=1}} \frac{\mu^2(t)\mu(\delta)}{\phi(\delta)\phi(t)} \\ &= \mu(\delta) \prod_{p|\delta} \left(1 + \frac{1}{p-1}\right) \sum_{\substack{t<\frac{z}{\delta} \\ (t,\delta)=1}} \frac{\mu^2(t)}{\phi(t)}. \end{aligned}$$

Taking the absolute value of both sides, we get

$$|V(z)| \cdot |\lambda_\delta| \leq |V(z)|.$$

Thus, $\forall \delta, |\lambda_\delta| \leq 1$ as desired, and we have that as $z, x \to \infty$,

$$\mathcal{O}\left(\sum_{d_1,d_2|P(z)} |\lambda_{d_1}| \cdot |\lambda_{d_2}|\right) = \mathcal{O}(z^2).$$

Recall that, $\forall z \leq x$, we have

$$\begin{aligned} \pi(x) &\leq S(\mathcal{A}, P(z)) + \pi(z) \\ &\leq S(\mathcal{A}, P(z)) + z. \end{aligned}$$

We now just need to find a lower bound on $V(z)$ and to choose z such that the error term does not get too big. Since $V(z) = \sum_{d<z} \frac{\mu^2(d)}{\phi(d)}$, we have that

$$\begin{aligned}
V(z) &\geq \sum_{d<z} \frac{\mu^2(d)}{d} \\
&= \sum_{d<z} \frac{1}{d} - \sum_{\substack{d<z \\ d \text{ not} \\ \text{squarefree}}} \frac{1}{d}.
\end{aligned}$$

Examining the two parts of this last expression separately, we can see first that

$$\sum_{d<z} \frac{1}{d} = \log(z) + \mathcal{O}(1).$$

Also, we can see that

$$\sum_{\substack{d<z \\ d \text{ not} \\ \text{squarefree}}} \frac{1}{d} \leq \frac{1}{4} \sum_{d<\frac{z}{4}} \frac{1}{d}.$$

Combining these two give us:

$$V(z) \gg \log(z).$$

We can conclude that $\pi(x) \ll \frac{x}{\log z} + z^2$, and thus, choosing $z = \left(\frac{x}{\log(x)}\right)^{\frac{1}{2}}$, we get

$$\begin{aligned}
\pi(x) &\ll \frac{2x}{\log\left(\frac{x}{\log(x)}\right)} + \frac{x}{\log(x)} \\
&= \frac{2x}{\log(x) - \log(\log(x))} + \frac{x}{\log(x)}.
\end{aligned}$$

From this, we can recover Chebycheff's upper estimate:

$$\pi(x) \ll \frac{x}{\log(x)}.$$

An almost identical proof works for the general statement of Selberg's Sieve, but again is more technical than enlightening, and is thus omitted. The general statement goes as follows:

Theorem. *Given $\mathcal{A} = \{a_n\}_1^{\lfloor x \rfloor} \subseteq \mathbb{N}$ a finite set such that $|\mathcal{A}| = \lfloor x \rfloor$. For all primes $p \in \mathcal{P}$, define $A_p = \{a_n \in \mathcal{A} : p \mid a_n\}$, and for d squarefree, $A_d = \{a_n \in \mathcal{A} : d \mid a_n\} = \bigcap\limits_{p|d} A_p$, and let $A_1 = \mathcal{A}$. Define*

$$S(\mathcal{A}, P(z)) = |\{a_n \in N : (a_n, P(z)) = 1\}| = |\mathcal{A} \setminus \bigcup_{p|P(z)} A_p|.$$

Now assume that, for any squarefree d divisible only by primes in $\mathcal{P}$,

$$|A_d| = \frac{x}{f(d)} + R_d$$

where f is a multiplicative function chosen so that

$$V(z) = \sum_{\substack{d<z \\ d|P(z)}} \frac{\mu^2(d)}{f_1(d)}$$

can be bounded from below, where f_1 is defined by $f(n) = \sum\limits_{d|n} f_1(d)$, i.e.,

$$f_1(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right),$$

and $R_d \in \mathbb{R}$ is a remainder term.

Then

$$S(\mathcal{A}, P(z)) \leq \frac{x}{V(z)} + \mathcal{O}\left(\sum_{\substack{d_1,d_2<z \\ d_1,d_2|P(z)}} |R_{[d_1,d_2]}|\right).$$

While it may be hard to see what this theorem is saying, it will become clear in the applications. The key step is finding the multiplicative function f that provides the best lower bound for $V(z)$, so the reciprocal is bounded from above. The next step is to get an upper bound on R_d, and thus for the whole error term in the main formula. Putting these pieces together will yield an upper bound on $S(\mathcal{A}, P(z))$, as desired.

There is another lemma that will be useful in the proof of the upper bound on the number of twin primes less than a number:

Lemma. *With the above assumptions of the general Selberg Sieve, define $\tilde{f}$ such that $\forall p \in \mathcal{P}$, $\tilde{f}(p) = f(p)$ and if $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$, then $\tilde{f}(n) = f(p_1)^{\alpha_1} \dots f(p_r)^{\alpha_r}$. Then we have:*

$$V(z) \geq \sum_{\substack{d<z \\ (d,P(z))\neq 1}} \frac{1}{\tilde{f}(d)}.$$

Proof. By definition, we have that $\tilde{f}(n) = \prod_{i=1}^{r} \tilde{f}(p_i)^{\alpha_i} = \prod_{i=1}^{r} f(p_i)^{\alpha_i}$. Recall that $V(z) = \sum_{\substack{d<z \\ d|P(z)}} \frac{\mu^2(d)}{f_1(d)}$ where $f_1(d) = \sum_{c|d} \mu(c) f\left(\frac{d}{c}\right)$. If d is squarefree, write $d = p_1 p_2 \dots p_s$, and

$$\begin{aligned}
f_1(d) &= \mu(1) f\left(\frac{d}{1}\right) + \mu(p_1) f\left(\frac{d}{p_1}\right) + \dots + \mu(p_i p_j) f\left(\frac{d}{p_i p_j}\right) + \dots + \mu(d) f\left(\frac{d}{d}\right) \\
&= f(d) - f\left(\frac{d}{p_1}\right) - \dots + f\left(\frac{d}{p_1 p_2}\right) + \dots + (-1)^{\nu(d)} f(1) \\
&= f(p_1 \dots p_s) - f(p_2 \dots p_s) - \dots + f(p_3 \dots p_s) + \dots + (-1)^{\nu(d)} \\
&= f(p_1) \dots f(p_s) - f(p_2) \dots f(p_s) - \dots + f(p_3) \dots f(p_s) + \dots + (-1)^{\nu(d)} \\
&= \prod_{p|d} (f(p) - 1).
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{1}{f_1(d)} &= \prod_{p|d} \frac{1}{f(p)-1} = \prod_{p|d} \frac{1/f(p)}{1-1/f(p)} = \prod_{p|d} \sum_{k\geq 1} \left(\frac{1}{f(p)}\right)^k \\
&= \left(\frac{1}{f(p_1)} + \left(\frac{1}{f(p_1)}\right)^2 + \dots\right) \dots \left(\frac{1}{f(p_s)} + \left(\frac{1}{f(p_s)}\right)^2 + \dots\right) \\
&= \sum_{n\in D_d} \frac{1}{\tilde{f}(n)}
\end{aligned}$$

where

$$D_d = \{n \in \mathbb{N} : n = p_1^{\alpha_1} \dots p_s^{\alpha_s}, \alpha_i \geq 1 \forall i = 1, \dots, s\} = \{n \in \mathbb{N} : p \mid n \iff p \mid d\}.$$

So for $d_1 \neq d_2$ both squarefree, we have that $D_{d_1} \cap D_{d_2} = \emptyset$ and $\forall e < z, e \in D_d$ for some D_d, namely the D_d such that $d = \prod_{p|e} p$.

Thus, since d is squarefree which means $\mu^2(d) = 1$,

$$\begin{aligned}
V(z) &= \sum_{\substack{d<z \\ d|P(z)}} \frac{1}{f_1(d)} \\
&= \sum_{\substack{d<z \\ d|P(z)}} \sum_{n\in D_d} \frac{1}{\tilde{f}(n)} \\
&\geq \sum_{\substack{d<z \\ d|P(z)}} \sum_{n\in D_d} \frac{1}{\tilde{f}(n)} \chi[n<z] \\
&= \sum_{\substack{d<z \\ (d,P(z))\neq 1}} \frac{1}{\tilde{f}(d)}.
\end{aligned}$$

$\square$

Chapter 5

Applications of Selberg's Sieve

5.1 Bounding the Number of Primes in an Interval

In this section, we will work through a simple application of Selberg's Sieve to bound from above the number of primes in an interval.

Theorem. *Let $H \geq 2$, then*

$$\pi(x; H) \leq \frac{2H}{\log(H)} + \mathcal{O}\left(\frac{H}{\log^2(H)}\right).$$

Proof. We have $\pi(x; H) = \pi(x + H) - \pi(x) = |\{p : x < p \leq x + H, p \text{ prime}\}|$. So the interval we're working with is $(x, x + H] \bigcap \mathbb{N}$, and this will be the set we take for $\mathcal{A}$.

If p is a prime, $p \mid P(z)$ or $(p, P(z)) = 1$, and thus we have

$$\pi(x; H) \leq |\{n : x < n \leq x + H, (n, P(z)) = 1\}| + \pi(z).$$

Then

$$\begin{aligned}|A_d| &= \left\lfloor \frac{x+H}{d} \right\rfloor - \left\lfloor \frac{x}{d} \right\rfloor \\ &= \frac{x+H}{d} - \frac{x}{d} + \mathcal{O}(1) \\ &= \frac{H}{d} + \mathcal{O}(1).\end{aligned}$$

And thus by Selberg's Sieve we get

$$S(\mathcal{A}, P(z)) \leq \frac{H}{V(z)} + \mathcal{O}\Bigg(\sum_{\substack{d_1, d_2 < z \\ d_1, d_2 | P(z)}} 1 \Bigg) = \frac{H}{V(z)} + \mathcal{O}(z^2).$$

So we can choose $f(d) = d$, and thus

$$f_1(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right) = n \sum_{d|n} \frac{\mu(d)}{d} = \phi(n).$$

Thus we get

$$\begin{aligned}V(z) &= \sum_{d<z} \frac{\mu^2(d)}{\phi(d)} \\ &= \sum_{d<z} \frac{\mu^2(d)}{d} \prod_{p|d} \left(1 - \frac{1}{p}\right)^{-1} \\ &= \sum_{d<z} \frac{\mu^2(d)}{d} \sum_{\substack{m=1 \\ p|m \Rightarrow p|d}}^{\infty} \frac{1}{m} \\ &\geq \sum_{n<z} \frac{1}{n} \\ &> \log(z).\end{aligned}$$

And thus we get

$$S(\mathcal{A}, P(z)) \leq \frac{2H}{\log(z^2)} + \mathcal{O}(z^2).$$

Choosing $z = \frac{\sqrt{H}}{\log(H)}$ yields

$$S(A, P(z)) \leq \frac{2H}{\log(H) - 2\log(\log(H))} + \mathcal{O}\left(\frac{H}{\log^2(H)}\right).$$

Since

$$\begin{aligned}
\pi(x; H) &\leq S(\mathcal{A}, P(z)) + \pi(z) \\
&\leq S(\mathcal{A}, P(z)) + z \\
&\leq \frac{2H}{\log(H) - 2\log(\log(H))} + \mathcal{O}\left(\frac{H}{\log^2(H)}\right) + \frac{\sqrt{H}}{\log(H)},
\end{aligned}$$

we conclude that

$$\pi(x; H) \leq \frac{2H}{\log(H)} + \mathcal{O}\left(\frac{H}{\log^2(H)}\right).$$

□

5.2 Bounding the Number of Twin Primes

In this section, we will discuss bounding the number of twin primes less than x from above. The famous Twin Prime Conjecture would be solved if one could prove any positive bound on this number from below that grows to infinity as x does.

To frame the problem of getting an upper bound on the number of twin primes in the setup of Selberg's Sieve, first we need to define $\mathcal{A}$. In this case, we want to look at integers n and those 2 larger, so we can define $\mathcal{A}$ as $\{n(n+2) : n < x\}$, and take all the primes as our $\mathcal{P}$. So, for example, if $x = 10$, $\mathcal{A} = \{3, 8, 15, 24, 35, 48, 63, 80, 99\}$. Note that we are not looking for primes in this set, just for elements of this set that are divisible by exactly 2 primes. So the main goal is to sift all numbers out of this set that are divisible by primes less than z and be left only with composite numbers with large prime factors, and hopefully take z large enough that the estimate is accurate. To accomplish this, we to count the elements of $\mathcal{A}$ that are relatively prime to $P(z)$. So we define $A_d = \{a_n \in \mathcal{A} : d \mid a_n\} = \{n(n+2) : n < x, d \mid (n(n+2))\}$, and want to consider $\mathcal{A} \setminus \bigcup\limits_{p|P(z)} A_p$.

Now we will get an estimate on $|A_d|$, which will lead to the desired function f. To do this, we define $\rho(d) = |\{n(\mod d) : n(n+2) \equiv 0 \mod d\}| = |\{n(\mod d) : d \mid n(n+2)\}|$. If we evaluate this function at a prime, we get

$$\rho(p) = |\{0, p-2(\mod p)\}| = \begin{cases} 1 & \text{if } p = 2 \\ 2 & \text{otherwise.} \end{cases}$$

We thus get that

$$|A_d| = \left\lfloor \frac{x\rho(d)}{d} \right\rfloor = \frac{x\rho(d)}{d} + \mathcal{O}(\rho(d)).$$

Thus, in the general formula for Selberg's Sieve, we take $f(d) = \frac{d}{\rho(d)}$, and $R_d = \rho(d)$. Using this, we can get an upper bound on the error term $\sum\limits_{\substack{d_1, d_2 < z \\ d_1, d_2 | P(z)}} |R_{[d_1, d_2]}|$. By noticing that for d squarefree, we have that $R_d = \rho(d) \leq 2^{\nu(d)}$, so we get

$$\begin{aligned}
\sum_{\substack{d_1,d_2<z \\ d_1,d_2|P(z)}} |R_{[d_1,d_2]}| &= \sum_{\substack{d_1,d_2<z \\ d_1,d_2 \text{ squarefree}}} \rho([d_1,d_2]) \\
&\leq \sum_{\substack{d_1,d_2<z \\ d_1,d_2 \text{ squarefree}}} 2^{\nu([d_1,d_2])} \\
&\leq \sum_{\substack{d_1,d_2<z \\ d_1,d_2 \text{ squarefree}}} 2^{\nu(d_1)}2^{\nu(d_2)} \\
&= \left(\sum_{\substack{d<z \\ d \text{ squarefree}}} 2^{\nu(d)} \right)^2
\end{aligned}$$

since $\nu([d_1,d_2]) \leq \nu(d_1 d_2) \leq \nu(d_1) + \nu(d_2)$.

Since d is assumed squarefree, $2^{\nu(d)} = \tau(d)$, where $\tau(d)$ is the number of divisors of d, and so

$$\begin{aligned}
\sum_{\substack{d<z \\ d \text{ squarefree}}} 2^{\nu(d)} &= \sum_{\substack{d<z \\ d \text{ squarefree}}} \tau(d) \\
&= \sum_{a \cdot b < z} 1 \\
&= \sum_{a<z} \left\lfloor \frac{z}{a} \right\rfloor \\
&= z \sum_{a<z} \frac{1}{a} + \mathcal{O}(z) \\
&= z \log(z) + \mathcal{O}(z).
\end{aligned}$$

Now we can state the bound on the error term:

$$\sum_{\substack{d_1,d_2<z\\ d_1,d_2|P(z)}} |R_{[d_1,d_2]}| \ll (z\log(z))^2.$$

The next step is to get a lower bound on $V(z)$. Define

$$\tilde{\rho}(n) = \rho(p_1)^{\alpha_1}\ldots\rho(p_r)^{\alpha_r}, \text{ where } n = p_1^{\alpha_1}\ldots p_r^{\alpha_r}.$$

So, like $\tilde{f}$ in the lemma on the bound on $V(z)$ in the section on Selberg's Sieve, $\tilde{\rho}$ is just the totally multiplicative extension of $\rho(n)$ restricted to the primes. It follows from this definition that

$$\tilde{\rho}(n) = \prod_{p_i|n} \rho(p_i)^{\alpha_i} = \prod_{\substack{p_i|n\\ p_i\neq 2}} 2^{\alpha_i} = 2^{\Omega_{\mathrm{odd}}(n)},$$

where $\Omega_{\mathrm{odd}}(n) = \sum_{\substack{p_i|n\\ p_i\neq 2}} \alpha_i$.

Using this, by the lemma on the bound on $V(z)$ in the section on Selberg's Sieve, we get

$$\begin{aligned} V(z) \geq \sum_{\substack{d<z\\ p|d\Rightarrow p|P(z)}} \frac{1}{\tilde{f}(d)} &= \sum_{\substack{d<z\\ p|d\Rightarrow p|P(z)}} \frac{1}{(f(p_1))^{\alpha_1}\ldots(f(p_r))^{\alpha_r}} \\ &= \sum_{\substack{d<z\\ p|d\Rightarrow p|P(z)}} \frac{1}{(\frac{p_1}{\rho(p_1)})^{\alpha_1}\ldots(\frac{p_r}{\rho(p_r)})^{\alpha_r}} = \sum_{\substack{d<z\\ p|d\Rightarrow p|P(z)}} \frac{\tilde{\rho}(d)}{d} \\ &= \sum_{\substack{d<z\\ p|d\Rightarrow p|P(z)}} \frac{2^{\Omega_{\mathrm{odd}}(d)}}{d}. \end{aligned}$$

If $p = 2$, then $2^{\Omega_{\text{odd}}(p^k)} = 1$, and $\tau_{\text{odd}}(p^k) = 1$. Now if p is an odd prime, then $2^{\Omega_{\text{odd}}(p^k)} = 2^k$, and $\tau_{\text{odd}}(p^k) = k + 1$. Thus, $2^{\Omega_{\text{odd}}(p^k)} \geq \tau_{\text{odd}}(p^k)$ for all p, and so we can substitute into the above inequality to get

$$V(z) \geq \sum_{\substack{d<z \\ p|d \Rightarrow p|P(z)}} \frac{2^{\Omega_{\text{odd}}(d)}}{d} \geq \sum_{\substack{d<z \\ p|d \Rightarrow p|P(z)}} \frac{\tau_{\text{odd}}(d)}{d}.$$

Now

$$\begin{aligned}
\sum_{n<z} \tau_{\text{odd}}(n) &= \sum_{n<z} \sum_{d|n} \chi[d \text{ odd}] \\
&= \sum_{n} \sum_{d} \chi[d \text{ odd}] \chi[n < z] \chi[d \mid n] \\
&= \sum_{\substack{d<z \\ d \text{ odd}}} \sum_{\substack{n<z \\ d|n}} 1 \\
&= \sum_{\substack{d<z \\ d \text{ odd}}} \left\lfloor \frac{z}{d} \right\rfloor \\
&= \sum_{\substack{d<z \\ d \text{ odd}}} \frac{z}{d} + \mathcal{O}(1) \\
&= z \sum_{\substack{d<z \\ d \text{ odd}}} \frac{1}{d} + \mathcal{O}(z).
\end{aligned}$$

If we break $\sum\limits_{\substack{d<z \\ d \text{ odd}}} \frac{1}{d}$ into the difference of two sums, i.e., $\sum\limits_{d<z} \frac{1}{d} - \sum\limits_{c<\frac{z}{2}} \frac{1}{2c}$ we see

$$\begin{aligned} z \sum_{\substack{d<z \\ d \text{ odd}}} \frac{1}{d} + \mathcal{O}(z) &= z \left(\sum_{d<z} \frac{1}{d} - \sum_{c<\frac{z}{2}} \frac{1}{2c} \right) + \mathcal{O}(z) \\ &= z \left(\log(z) + \mathcal{O}(1) - \frac{1}{2} \log\left(\frac{z}{2}\right) + \mathcal{O}(1) \right) + \mathcal{O}(z) \\ &= \frac{z}{2} \log(z) + \mathcal{O}(z). \end{aligned}$$

Therefore, we have that

$$\begin{aligned} V(z) &\geq \sum_{d<z} \frac{\tau_{\text{odd}}(d)}{d} \\ &= \frac{1}{z} \sum_{d<z} \tau_{\text{odd}}(d) + \int_1^z \frac{\sum\limits_{d<z} \tau_{\text{odd}}(d)}{t^2} dt \\ &= \mathcal{O}(\log(z)) + \int_1^z \frac{\frac{1}{2} t \log(t) + \mathcal{O}(t)}{t^2} dt \\ &= \frac{1}{4} \log^2(z) + \mathcal{O}(\log(z)). \end{aligned}$$

Thus we get that

$$V(z) \gg \log^2(z).$$

Combining this with the above upper bound on the error term, we get that

$$\begin{aligned} S(\mathcal{A}, P(z)) &\leq \frac{x}{V(z)} + \mathcal{O}\left(\sum_{\substack{d_1, d_2 < z \\ d_1, d_2 | P(z)}} |R_{[d_1, d_2]}| \right) \\ &\ll \frac{x}{\log^2(z)} + \mathcal{O}((z \log(z))^2) \end{aligned}$$

and conclude, taking $z = x^{\frac{1}{4}}$, that

$$\begin{aligned} S(\mathcal{A}, P(z)) &\ll \frac{x}{\log^2(x)} + \mathcal{O}\left(\frac{x^{\frac{1}{2}}\log^2(x)}{2}\right) \\ &\ll \frac{x}{\log^2(x)}. \end{aligned}$$

Recall that $S(\mathcal{A}, P(z)) = |\{n(n+2) : (n(n+2), P(z)) = 1, n < x\}|$. Thus we get a bound on $\pi_2(x)$:

$$\begin{aligned} \pi_2(x) &= (\pi_2(x) - \pi_2(z)) + \pi_2(z) \\ &\leq S(\mathcal{A}, P(z)) + \pi_2(z) \\ &\leq S(\mathcal{A}, P(z)) + z \\ &\ll \frac{x}{\log^2(x)} + x^{\frac{1}{4}}. \end{aligned}$$

And therefore

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

Chapter 6

Sum of Reciprocals

6.1 Divergence of the Sum of the Reciprocals of the Primes

Theorem. $\sum_p \frac{1}{p}$ *diverges.*

Proof. Fix $j \in \mathbb{N}$, and let $2, 3, 5, \ldots, p_j$ be the first j primes. Define:

$$N(x) = \#\{n \in \mathbb{N} \ : \ n \leq x \text{ and } \forall p_k > p_j, (n, p_k) = 1\}.$$

Thus $N(x)$ is the number of positive integers less than or equal to x that are not divisible by any prime greater than p_j.

We will first prove a lemma: $N(x) \leq 2^j \sqrt{x}$. Let n be a number in the set that $N(x)$ is counting. If all the non-zero even powers of primes in the factorization of n are lumped into one term n_1, then n can be written as $n_1^2 m$, where m is square-free.

Since all of the prime factors of n are less than or equal to p_j, $m = 2^{b_1}3^{b_2}\dots p_j^{b_j}$, where $b_i \in \{0,1\}$. Since there are 2^j subsets of a set of size j, there are 2^j possible different values of m. The largest n_1 can be in relation to n is $\sqrt{n}$, so because $n \leq x$, we have that $n_1 \leq \sqrt{n} \leq \sqrt{x}$. Thus there are at most $\sqrt{x}$ possible values of n_1. Therefore $N(x) \leq 2^j\sqrt{x}$.

Now to finish the proof of the theorem. Suppose $\sum\limits_p \frac{1}{p}$ converges. Then we can pick j such that $\sum\limits_{p>p_j} \frac{1}{p} < \frac{1}{2}$. Notice that

$$\#\{n \leq x : p \mid n\} = \left\lfloor \frac{x}{p} \right\rfloor \leq \frac{x}{p}.$$

Hence,

$$x - N(x) = \#\{n \leq x : p_i \mid n \text{ for some } i > j\} \leq \sum_{i>j} \frac{x}{p_i}.$$

By assumption, this is less than $\frac{x}{2}$ and we get, by the lemma:

$$\frac{x}{2} < N(x) \leq 2^j\sqrt{x}.$$

This implies that $\sqrt{x} < 2^{j+1}$, and so $x < 2^{2j+2}$. But this is clearly not true for all $x \geq 2^{2j+2}$, a contradiction, and therefore $\sum\limits_p \frac{1}{p}$ diverges. $\square$

6.2 Brun's Theorem

Recall that a twin prime is a prime number p such that $p+2$ is also prime, such as $p = 29$. We will now prove the convergence of the sum of the reciprocals of the twin primes. In 1915, using sieves, Viggo Brun showed that

$$\pi_2(x) = \text{ The number of twin primes less than } x \ll \frac{x(\log\log(x))^2}{(\log(x))^2}.$$

In the Applications of Selberg's Sieve section, we showed the better bound

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

which implies that $\exists c \in \mathbb{R}$ such that for all x large enough, $\pi_2(x) < \frac{cx}{\log^2(x)}$.

To prove $\sum_{\text{twin primes}} \frac{1}{p}$ converges, we will get an estimate on the growth of $\sum_{\substack{p \text{ a twin} \\ p<x}} \frac{1}{p}$ as $x \to \infty$ using a lemma, whose proof can be found in the preliminaries section:

Lemma. *Partial Summation.*

If $\{a_n\}_1^\infty \subseteq \mathbb{C}$ and if $f : \mathbb{N} \to \mathbb{C}$ is continuously differentiable and we define:

$$S(x) = \sum_{1 \le n \le x} a_n,$$

then, $\forall A, B \in \mathbb{N}$ with $A < B$, we have:

$$\sum_{A<n\le B} a_n f(n) = f(B)S(B) - f(A)S(A) - \int_A^B S(x)f'(x)dx.$$

We will apply this lemma with $a_n = \chi[n \text{ is a twin prime}]$, $f(n) = \frac{1}{n}$, and $S(x) = \pi_2(x)$. So $\chi[n \text{ is a twin prime}] = \pi_2(n) - \pi_2(n-1)$.

Thus we have

$$\begin{aligned}
\sum_{\substack{p \text{ a twin} \\ p<x}} \frac{1}{p} &= \sum_{2\le n\le x} \frac{1}{n}\chi[\text{n is a twin prime}] \\
&= \frac{1}{x}\pi_2(x) - \frac{1}{2}\pi_2(2) - \int_2^x \frac{\pi_2(y)}{-y^2}dy \\
&= \frac{\pi_2(x)}{x} + \int_2^x \frac{\pi_2(y)}{y^2}dy.
\end{aligned}$$

Applying the bound we got from Selberg's Sieve, we get:

$$\frac{\pi_2(x)}{x} + \int_2^x \frac{\pi_2(y)}{y^2}dy < \frac{c}{\log^2(x)} + \int_2^x \frac{dy}{y\log^2(y)}.$$

The first term $\frac{c}{\log^2(x)}$ clearly goes to 0 as $x \to \infty$. Evaluating the integral, we get

$$\int_2^x \frac{dy}{y\log^2(y)} = \frac{-1}{\log(y)}\bigg|_2^x = \frac{1}{\log(2)} - \frac{1}{\log(x)} \to \frac{1}{\log(2)} \text{ as } x \to \infty.$$

Therefore, we get

$$\sum_{\substack{\text{p a twin} \\ p\le x}} \frac{1}{p} < 2c\left(\frac{1}{\log(2)} - \frac{1}{\log(x)}\right) \to \frac{2c}{\log(2)} \text{ as } x \to \infty,$$

and thus

$$\sum_{\substack{\text{p a twin} \\ \text{prime}}} \frac{1}{p} < \infty.$$

Chapter 7

Conclusion

While the Sieve of Eratothenes/Legendre can recover a weak upper bound for the number of primes less than a given number, the more sophisticated Selberg's Sieve is needed to recover the Chebycheff's upper bound. The convergence of the sum of the reciprocals of the twin primes is usually proved in the literature using Brun's estimate, as that is historically where the proof originates. However the proof shown in this work is using the slightly better bound from Selberg's Sieve.

There are methods of using sieves to get lower bounds for growth rates of sequences, which is how the positive prime gap results were reached. Additionally, rather than choosing an optimal set of λ_d's, it turns out that a less than optimal set is used to prove Zhang's Theorem.

There are some conjectures regarding these tools that are yet to be settled, but to discuss them, first we need to define what an admissible set is.

Definition. Admissible Set: A set of natural numbers $H = \{h_1, h_2, \ldots, h_k\}$ is called admissible if $\forall p$, at least one of the residue classes modulo p is missed by every one of the h_i's.

For example, the sets $\{0,2\},\{0,4\},\{0,4,6\}$ are all admissible, but the set $\{0,2,4\}$ is not admissible since it fails the criterion for $p=3$. So the Twin Prime Conjecture is saying that the admissible set $\{0,2\}$ has infinitely many prime translates on the natural numbers. That is, there exist infinitely many n such that $n+0$ and $n+2$ are both prime. A strong generalization of this conjecture is the Admissible Set Conjecture:

Conjecture. *There exist infinitely many k-tuples, $(n+h_1, n+h_2, \ldots, n+h_k)$, consisting entirely of primes if and only if the set $H = \{h_1, h_2, \ldots, h_k\}$ is an admissible set.*

So Zhang's Theorem is a weaker form of this theorem. In 2013, James Maynard and Terrance Tao proved the related case for any number of primes which is known as the Maynard-Tao Theorem:

Theorem. $\forall m \geq 2, \exists k \in \mathbb{N}$ *such that if* $\{h_1, h_2, \ldots, h_k\}$ *is an admissible set, then there exists infinitely many n for which at least m entries of the k-tuple $(n+h_1, n+h_2, \ldots, n+h_k)$ are all prime.*

Assuming a strong form of the Elliot-Halberstam Conjecture, Maynard was also able to show:

Conjecture. $\liminf_n (p_{n+1}-p_n) \leq 12$ *and* $\liminf_n (p_{n+2}-p_n) \leq 600$.

These bounds are considered the limit of the current sieve methods. To state the Elliot-Halberstam Conjecture, first we need to define the function

$$\Theta(x) = \sum_{\substack{p \text{ prime} \\ p \leq x}} \log(p).$$

It can be shown that the Prime Number Theorem is equivalent to showing that $\Theta(x) \sim x$ as $x \to \infty$. The Prime Number Theorem for Arithmetic Progressions can then be stated as

$$\Theta(x;q,a) = \sum_{\substack{p \text{ prime} \\ p \leq x \\ p \equiv a (\mod q)}} \log(p) \sim \frac{x}{\phi(x)} \text{ as } x \to \infty.$$

The Bombieri-Vinogradov Theorem can be thought of as a Prime Number Theorem for Arithmetic Progressions on average and is stated:

Theorem. $\forall A > 0, \exists B(A)$ *such that*

$$\sum_{q \leq Q} \max_{\substack{a (\mod q) \\ (a,q)=1}} \left| \Theta(x;q,a) - \frac{x}{\phi(q)} \right| \ll_A \frac{x}{(\log(x))^A}$$

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

Finally, the Elliot-Halberstam Conjecture appears similar but has eluded proof and is stated as:

Conjecture. $\forall A > 0$ *and* ϵ *such that* $0 < \epsilon < \frac{1}{2}$,

$$\sum_{q \leq Q} \max_{\substack{a (\mod q) \\ (a,q)=1}} \left| \Theta(x;q,a) - \frac{x}{\phi(q)} \right| \ll \frac{x}{(\log(x))^A}$$

where $Q = x^{\frac{1}{2}+\epsilon}$.

Bibliography

[1] T. Apostol, *Introduction to Analytic Number Theory*, Springer Verlag, New York, NY, 2010.

[2] V. Brun. Über das Goldbachsche Gesetz und die Anzahl der Primzahlpaare. Archiv for Math. og Naturvif. B34 (1915), 19 pages.

[3] A. C. Cojocaru and M. Ram Murty, *An Introduction to Sieve Methods and their Applications*, London Mathematical Society Student Texts, 66, Cambridge University Press, Cambridge, 2006.

[4] H. Davenport, *Multiplicative Number Theory* Springer Verlag, New York, NY, 2000.

[5] A. Granville. Primes in Intervals of Bounded Length. Bull. Amer. Math. Soc. (N.S.) 52 (2015), No. 2, 171–222.

[6] G. Greaves, *Sieves in Number Theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas], 43, Springer-Verlag, Berlin, 2001.

[7] H. Halberstam and H.-E. Richert, *Sieve Methods*, Dover, Mineola, NY, 2011.

[8] H. Halberstam and K. F. Roth, *Sequences*, Springer Verlag, New York, NY, 1983.

[9] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., Oxford University Press, Oxford, 1979.

[10] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory* Springer Verlag, New York, NY, 1983.

[11] H. Iwaniec and E. Kowalski, *Analytic Number Thoery*, Colloq. Publications, vol. 53, Amer. Math. Soc., Providence, RI, 2004.

[12] J. Friedlander and H. Iwaniec, *Opera de Cribro*, Colloq. Publications, vol. 57, Amer. Math. Soc., Providence, RI, 2010.

[13] M. Ram Murty, *Problems in Analytic Number Theory*, Springer Verlag, New York, NY, 2001.

[14] H. Rademacher, *Lectures on Elementary Number Theory*, Blaisdell Publishing Co., New York, NY, 1964.

[15] Z. Rudnick: Sieve Theory and its applications, `http://www.math.tau.ac.il/~rudnick/courses/sieves2015.html`