Linear Sieve Talk

Katherine Wimmer

April 10, 2024

1 Introduction

A few weeks ago, Jinoo gave a talk that discussed the Selberg sieve. In the next talk, Robert will prove Chen's theorem which states that every sufficiently large even integer can be written as the sum of a prime and a number that is the product of at most two primes. This proof, however, require some more sophisticated sieve estimates than what the Selberg sieve gives us. Today we will be going over the linear sieve, as well as the Jurkat-Richert theorem which will allow us to get upper and lower bounds for the linear sieve.

2 Set-up for a Generalized Combinatorial Sieve

In order to build up to the eventual functions that will be used in bounding the linear sieve, we must establish some variables, sets and functions. Throughout the talk, certain functions that are integral parts of others can be bounded, thus allowing for bounds on more complex functions. What is about to be initially set up should look very familiar, it is important to reiterate what is going into the bounds in the Jurkat-Richert theorem, as their behavior is important to understanding. First is the arithmetic function:

$$A = \{a(n)\}_{n=1}^{\infty} \quad \text{where} \quad a(n) \geq 0 \quad \text{for all } n$$

We also establish the set P as the sieving range:

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

where z is a the sieving level. z is a real number, $z \geq 2$.
We can now write:

$$S(A, P, z) = \sum_{(n, P(z))=1} a(n)$$

Which counts the number of integers in the set A that are not divisible by any prime $p \in P$ for all $p < z$.

Next we are going to write several multiplicative and arithmetic functions. These functions will build a base for functions to be established later. First we will denote a multiplicative function $g_n(d)$, where for every $n \geq 1$ it holds that for every integer d that is the product of distinct primes $p \in P$:

$$0 \leq g_n(d) \leq 1$$

We also have for every prime $p \in P$. :

$$0 \leq g_n(p) < 1$$

The sieve idea is to reduce the size of the error term by replacing the Möbius function:

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

With carefully constructed arithmetic functions λ^+ and λ^- , where $\lambda^+ = \lambda^- = 1$. Then, for every $m \geq 2$:

$$(1 * \lambda^+)(m) = \sum_{d|m} \lambda^+(d) \geq 0$$

$$(1 * \lambda^-)(m) = \sum_{d|m} \lambda^-(d) \leq 0$$

We will be able to use these arithmetic functions as respective upper and lower bounds. If we take D to be a positive integer, such that $D \in \mathbb{Z}_{>0}$, then:
If $\lambda^+(d) = 0$ for all $d \leq D$, then λ^+ is an upper bound sieve with support level D.
Likewise, if $\lambda^-(d) = 0$ for all $d \geq D$ then λ^- is a lower bound sieve with support level D.

Next we will take the multiplicative function $g_n(d)$ and the $\lambda^\pm$ functions, we can build new functions. We will take P to be our sieving range, where P is a set of primes such that $\lambda^\pm(d) = 0$ whenever d is divisible by a prime that is not in P. We have specified that $\lambda^\pm$ will act as upper and lower bounds with sieving range P and support level D. From these functions we can define:

$$G_n(z, \lambda^\pm) = \sum_{d|P(z)} \lambda^\pm(d) g_n(d)$$

and

$$R^\pm = \sum_{\substack{d|P(z) \\ p<z}} \lambda^\pm(d) r(d)$$

We will define $r(d)$ in the context of A later, but generally is is a remainder term for a given d. And that sets up the necessary functions for Theorem 9.1, which places an upper and lower bound on $S(A, P, z)$:

$$\sum_{n=1}^{\infty} a(n) G_n(z, \lambda^-) + R^- \leq S(A, P, z) \leq \sum_{n=1}^{\infty} a(n) G_n(z, \lambda^+) + R^+$$

This is an important refinement to the basic sieve inequality, for further details on the proof of this inequality, pages 235 - 237 in Nathanson provide the proof in its entirety.

One last characteristics of the arithmetic functions $g_n(d)$ is that they often satisfy one-sided inequalities of the form:

$$\prod_{\substack{p \in P \\ u \leq p < z}} (1 - g_n(p))^{-1} \leq K \left(\frac{\log z}{\log u} \right)^k$$

Where $K > 1$ and $k > 0$ are constants that are independent of n. The inequality holds for all n and $1 < u < z$

We say that this sieve has "dimension" k and in the case of the linear sieve, $k = 1$. Sieving dimension is an assumption on our problem, about what the function $g(p)$ generally looks like. With a linear sieve, $g(p) \approx \frac{1}{p}$ or an equation that looks similar to this. Recall $0 \leq g_n(d) \leq 1$. For the remainder of the talk, we will be discussing inequalities for which $k = 1$.

3 Combinatorial Sieves

We are going to prove some properties of combinatorial sieves which we'll need later. They should be generally conceptually familiar, so results will be stated and the proofs will not be covered in very much depth. We will finalize the replacement of the Möbius function and set up $V(z)$.

We want to reduce the size of the error term that's found in Legendre's formula in a combinatorial sieve. This can be done by replacing the Möbius function with its truncation to a finite set of positive integers.

$$\lambda^{\pm} = \begin{cases} \mu & \text{if } d \in D^{\pm} \quad \text{and} \quad d|P(D) \\ 0 & \text{otherwise} \end{cases}$$

Where $D^{\pm}$ are finite sets of square-free positive integers $d < D$ and $\lambda^{\pm}$ are the lower and upper bounds with sieving range P and support level D. Also, $P(D)$ is the product of all the primes in P are that less than D. These restrictions on D are what enables the reduction in the size of the error term.

We also want to define a function $V(z)$ which will be the probability of not being ruled out by the sieve. We establish P to be a set of primes, $g(d)$ as a multiplicative function $0 \leq g(p) < 1$ for all $p \in P$. $V(z)$ is going to be built from primes in P and the function $g(p)$. It is defined as:

$$V(z) = \prod_{\substack{p \in P \\ p < z}} (1 - g(p)) = \sum_{p|P(z)} \mu(d)g(d)$$

Recall that we were able to previously bound $\prod_{\substack{p \in P \\ p<z}}(1 - g(p))$ from the first section, this then allows up to bound$V(z)$. $V(z)$ is a decreasing function of z, where $0 < V(z) \leq 1$. For all z and $1 \leq w < z$ it follows that $V(z)$ also has the properties:

$$\sum_{\substack{p \in P \\ w \leq p < z}} g(p)V(p) = V(w) - V(z)$$

Since we early put an assumption of linearly on g_n, w know that $V(z)$ is going to be linear as well. This property will be useful going forward since $V(z)$ will be generally predictable. The proof for $V(z)$ can be found on pages 242-244 of Nathanson.

4 Important Lemmas and Theorems

We must prove a series of Lemmas to get to the Jurkat-Richert Theorem. We have already shown that there is an upper and lower bound for $S(A, P, z)$ which is in terms of $G(\lambda^{\pm})$, which is made up of g_n. In Lemma 9.3, we want to express $G(\lambda^{\pm})$ in terms of $V(z)$ and T_n. The main takeaway of Lemma 9.3 is that we can largely express $G(\lambda^{\pm})$ in terms of T_n.

Lets give a brief overview of what makes up $T_n(D, z)$ and what are its properties, though we will go into more depth later in the talk. Lets take $z \geq 2$ and D to be real numbers such that:

$$s = \frac{\log D}{\log z} \geq \begin{cases} 1 & \text{if n is odd} \\ 2 & \text{if n is even} \end{cases}$$

Where n stipulates the properties of D and z. And we have β such that:

$$\beta \leq \frac{\log D}{\log z} = s$$

The function $T_n(D, z)$ is a sum over integers $p_1 \ldots p_n$ that has the properties:

$$T_n(D, z) = 0 \quad \text{for} \quad n \leq s - \beta$$

Lemma 9.3 takes $G(z, \lambda^{\pm})$ which is (as previously stated with definitions):

$$G(z, \lambda^{\pm}) = \sum_{d|P(z)} \lambda^{\pm}(d)g_n(d)$$

From there we can write $G_n(z, \lambda^{\pm})$ as an expression of $V(z)$ and $T_n(D, z)$:

$$G(z, \lambda^{+}) = V(z) + \sum_{\substack{n=1 \\ n \equiv 1 (mod 2)}}^{\infty} T_n(D, z)$$

and similarly:

$$G(z, \lambda^{-}) = V(z) - \sum_{\substack{n=1 \\ n \equiv 0 (mod 2)}}^{\infty} T_n(D, z)$$

Then we can write that when $n \geq 2$:

$$T_1(D, z) = V(D^{\frac{1}{3}}) - V(z)$$

For when n is even or odd and $s \geq 3$ then:

$$T_n(D, z) = \sum_{\substack{p \in P \\ p < z}} g(p) T_{n-1} \left(\frac{D}{p}, p \right)$$

and if n is off and $1 \leq s \leq 3$ then:

$$T_n(D, z) = \sum_{\substack{p \in P \\ p < D^{\frac{1}{3}}}} g(p) T_{n-1} \left(\frac{D}{p}, p \right)$$

We also want to define functions $F(s)$ and $f(s)$. These both rely on $f_n(s)$ which is a sequence of continuous functions which is defined by a recursive relation. It is created by some very complicated definitions using integrals. Such as:

$$s f_n(s) = \int_s^{\infty} f_{n-1}$$

Which is a volume integral dependent on s, we can go backwards using integration starting with some known values of f_1 and potentially f_2 . This talk will not explore this function past this general understanding. We introduce $F(s)$ which is continuous and differentiable for $s \geq 1$:

$$F(s) = 1 + \sum_{\substack{n=1 \\ n \equiv 1 (mod 2)}}^{\infty} f_n(s) \quad \text{and} \quad F(s) = 1 + O(e^{-s})$$

The function $f(s)$ is also continuous and differentiable for $s \geq 2$:

$$f(s) = 1 - \sum_{\substack{n=2 \\ n \equiv 0 (mod 2)}}^{\infty} f_n(s) \quad \text{and} \quad f(s) = 1 + O(e^{-s})$$

We will not go over the proof but it can also be said that:

$$\sum_{n=1}^{\infty} f_n(s) \ll e^{-s}$$

Theorem 9.5 will show that T_n can be bounded by f_n and $V(z)$, such that $T_n \leq f_n$ and $V(z)$. This is important because before we were able to express $G(\lambda^{\pm})$ in terms of T_n. This comes will full circle in Theorem 9.6, where we are able to bound $G(\lambda^{\pm})$ using f_n, such that $G(\lambda^{\pm}) \leq f_n$ and $V(z)$. As for Theorem 9.5, lets recall from the first section of the talk that:

$$\prod_{\substack{p \in P \\ u \leq p < z}} (1 - g_n(p))^{-1} \leq K \frac{\log z}{\log u} \quad \text{where} \quad 1 < K < 1 + \frac{1}{200}$$

and the function $V(z)$:

$$V(z) = \prod_{\substack{p \in P \\ p < z}} (1 - g(p))$$

We can then bound $T_n(D, z)$ with the inequality:

$$T_n(D, z) < V(z) \left(f_n(s) + (K - 1) \left(\frac{99}{100} \right)^n e^{10-s} \right)$$

Since we have been able to bound $T_n(D, z)$, we can go ahead with Theorem 9.6 and bound G_n using the functions $F(s)$, $f(s)$ and $V(z)$. We will say that $K = 1 + \epsilon$ satisfies the hypotheses of Theorem 9.5 (such that $\epsilon < \frac{1}{200}$. We then can say that, for $G_n(z, \lambda^{\pm})$ there is the upper bound :

$$G(z, \lambda^+) < V(z) \left(F(s) + \epsilon e^{14-s} \right)$$

and the lower bound:

$$G(z, \lambda^-) > V(z) \left(f(s) - \epsilon e^{14-s} \right)$$

Thus we have bound G_n by the functions of $F(s)$,$f(s)$ and $V(z)$, utilizing the properties of T_n.

5 The Jurkat-Richert Theorem

This theorem allows us to bound original generalized sieve

$$S(A, P, z)$$

using the past the functions f_n and $V(z)$ since we have used them to bound $G(z, \lambda^{\pm})$ which we used to bound $S(A, P, z)$ when we did our set up for a generalized sieve. Recall:

$$\sum_{n=1}^{\infty} a(n) G_n(z, \lambda^-) + R^- \leq S(A, P, z) \leq \sum_{n=1}^{\infty} a(n) G_n(z, \lambda^+) + R^+$$

As a few reminders, we will go over some of the definitions of the functions we will be using for the Jurkat-Richert. Starting with the cardinality of the set A:

$$|A| = \sum_{n=1}^{\infty} a(n) < \infty \quad \text{and that} \quad S(A, P, z) = \sum_{\substack{n=1 \\ (n, P(z))=1}}^{\infty} a(n)$$

We then define $r(d)$ in the context of the cardinally of A where all the numbers are divisible by a certain d:

$$|A_d| = \sum_{\substack{n=1 \\ d|n}}^{\infty} a(n) = \sum_{a(n)}^{\infty} a(n) g_n(d) + r(d)$$

Next we can define $\mathcal{Q}$ to be a finite subset of P and let Q be the product of the primes in $\mathcal{Q}$. Such that:

$$Q = \prod_{p \in \mathcal{Q}} p$$

Supposed that for some ϵ satisfying $0 < \epsilon < \frac{1}{200}$ then:

$$\prod_{\substack{p \in P/\mathcal{Q} \\ u \leq p < z}} (1 - g_n(p))^{-1} < (1 + \epsilon) \frac{\log z}{\log u}$$

which holds for all n and $1 < u < z$. Then we define X as:

$$X = \sum_{n=1}^{\infty} a(n) \prod_{p|P(z)} (1 - g_n(p))$$

which can be written more concisely as:

$$X = V(z)|A|$$

and R as:

$$R = \sum_{\substack{d|P(z) \\ d<DQ}} |r(d)|$$

Then we can bound $S(A, P, z)$ using the functions that have been defined throughout the talk. For any $D \geq z$ there is an upper bound:

$$S(A, P, z) < \left(F(s) + \epsilon e^{14-s}\right) X + R$$

and for any $D \geq z^2$ there is the lower bound:

$$S(A, P, z) > \left(f(s) - \epsilon e^{14-s}\right) X - R$$

In the following talk, Robert will apply this theorem, with the Vinogradov theorem that Julie talked about last week, in the pursuit of proving Chen's Theorem.

6 Extra Information on $F(s)$ and $f(s)$

The last section of the chapter on the linear sieve in Nathanson goes over the computation of the initial values of $F(s)$ and $f(s)$, but for the sake of time and brevity, these will only be discussed if there is extra time. Since these functions are of large importance in the Jurkat-Richert theorem, we can state the facts that are discovered in the last section, but the proofs will not be included.

$$F(s) = 1 + \sum_{\substack{n=1 \\ n \equiv 1 (mod 2)}}^{\infty} f_n(s) \quad \text{for} \quad s \geq 1$$

and

$$f(s) = 1 - \sum_{\substack{n=2 \\ n \equiv 0 (mod 2)}}^{\infty} f_n(s) \quad \text{for} \quad s \geq 2$$

We can also state that:

$$F(s) = \frac{2e^{\gamma}}{s} \quad \text{for} \quad 1 \leq s \leq 3$$

and

$$f(s) = \frac{2e^{\gamma} \log(s-1)}{s} \quad \text{for} \quad 2 \leq s \leq 4$$

Where γ is Euler's constant. We define $f(s) = 0$ for $1 \leq s \leq 2$.