Additive Number Theory Seminar, MATH 3952
Columbia University
Spring 2024
Jinoo Kim

Talk 14

Selberg's sieve and Goldbach's conjecture

1 The Goldbach Conjecture

Goldbach's conjecture is one of the most famous unsolved problems in number theory. In 1742, Christian Goldbach wrote to Leonhard Euler and proposed that every positive even integer can be written as the sum of two primes.

This is still unproven but has been shown to hold for all integers less than 4×10^{18}.

In this talk we will introduce another sieve which can be used to prove a major result on the way to the Goldbach conjecture - Schnirelmann's Theorem - which bounds the number of primes required to sum in order to yield any integer greater than one.

2 The Selberg Sieve

When Schnirelmann originally proved his theorem, he actually used Brun's sieve, which we are all very familiar with from last week. The Selberg Sieve, however, is a different combinatorial sieve and can also be used to prove Schnirelmann's Theorem.

Selberg's Sieve was developed about 30 years after Brun's Sieve. Nathanson claims it is "more elegant," which may be true, but we will also see that it has notable advantages in terms of 'accuracy.'

Before showing the sieve, consider the following intermediate lemma.

Lemma 1. *Let $a_1, ..., a_n$ be positive real numbers and $b_1, ..., b_n$ be any real numbers. The minimum value of the quadratic form*

$$Q(y_1, ..., y_n) = a_1 y_1^2 + ... + a_n y_n^2$$

subject to the linear constraint

$$b_1 y_1 + ... + b_n y_n = 1 \tag{1}$$

is

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

and this value is attained if and only if

$$y_i = \frac{mb_i}{a_i}$$

for all $i = 1, ..., n$.

The proof is fairly uninteresting and follows naturally from Cauchy-Schwartz inequality (See Nathanson §7.2).

The basic intuition behind Selberg's sieve is to replace the difficult and random behavior of the Möbius function and replace it with something that is slightly more 'arbitrary,' for lack of a better term. This takes on the form of a series $\{\lambda_d\}$ subject to a few simple constraints.

Theorem 1. *Let* A *be a finite sequence of integers, and let* $|A|$ *denote the number of terms of the sequence. Let* $\mathcal{P}$ *be a set of primes. For any real number* $z \geq 2$*, let*

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

The "sieving function"

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

denotes the number of terms of the sequence A *that are not divisible by any prime* $p \in \mathcal{P}$ *such that* $p < z$*. For every square-free positive integer* d*, let* $|A_d|$ *denote the number of terms of the sequence* A *that are divisible by* d*. Let* $g(k)$ *be a multiplicative function such that*

$$0 < g(p) < 1 \text{ for all } p \in \mathcal{P}$$

and let $g_1(m)$ *be a completely multiplicative function such that* $g_1(p) = g(p)$ *for all* $p \in \mathcal{P}$*. Define the "remainder term"* $r(d)$ *and the function* $G(z)$ *by*

$$r(d) = |A_d| - g(d)|A|$$

and

$$G(z) = \sum_{\substack{m<z \\ p|m \implies p\in\mathcal{P}}} g_1(m)$$

Then

$$S(A, \mathcal{P}, z) \leq \frac{|A|}{G(z)} + \sum_{\substack{d<z^2 \\ d|P(z)}} 3^{\omega(d)}|r(d)|, \tag{2}$$

where $\omega(d)$ *is the number of distinct prime divisors of* d*.*

• **Pf.** g is multiplicative, meaning that $g(mn) = g(m)g(n)$ whenever m and n are relatively prime positive integers. It then follows that $g([m,n])g((m,n)) = g(m)g(n)$ (Proof in Appendix §A.3), where $[m,n]$ is the LCM.

Let $\{\lambda_d\}$ be any sequence of real numbers such that

$$\lambda_1 = 1$$

and

$$\lambda_d = 0 \text{ for any } d > z.$$

Since

$$\left(\sum_{d|(a,P(z))} \lambda(d)\right)^2 \geq 0$$

for all non-negative integers a (because something squared is always non-negative) and

$$\left(\sum_{d|(a,P(z))} \lambda(d)\right)^2 = 1 \text{ if (a, P(z)) = 1,}$$

(because if a is coprime with the product of all primes less than z then the only non-zero element of the sum will be when $d = 1$), we have the following:

$$\begin{aligned} S(A,\mathcal{P},z) &= \sum_{\substack{a\in A \\ (a,P(z))=1}} 1 \\ &\leq \sum_{a\in A}\left(\sum_{d|(a,P(z))}\lambda(d)\right)^2 \\ &= |A|Q + R, \end{aligned}$$

where

$$Q = \sum_{\substack{d_1,d_2<z \\ d_1,d_2|P(z)}} \frac{1}{g((d_1,d_2))} g(d_1)\lambda(d_1)g(d_2)\lambda(d_2) \tag{3}$$

and

$$R = \sum_{\substack{d_1,d_2<z \\ d_1,d_2|P(z)}} \lambda(d_1)\lambda(d_2)r([d_1,d_2]).$$

So we have now obtained a simplified expression for the sieving function S, in terms of three variables $|A|, Q$, and R. In this next part, we will be dealing with the Q and attempting to give it some bound we can work with.

Define $\mathcal{D}$ as the set of all positive divisors of $P(z)$ that are strictly less than z.

$$\mathcal{D} = \{k|P(z) : 1 \leq k < z\}.$$

By the properties of $P(z)$, every $k \in \mathcal{D}$ is square-free. Thus, k is the product of a series of distinct primes $p_1, p_2, ...p_n$. Since $g(k) = g(p_1)g(p_2)...g(p_n)$ and each $g(p_i)$ is between 0 and 1, $0 < g(k) \leq 1$.

We then define the function $f(k)$ supported for $k \in \mathcal{D}$:

$$f(k) = \frac{1}{g(k)} \prod_{p|k} (1 - g(p))$$

By the properties of our function g, we have that $f(k) > 0$ and $f(k_1 k_2) = f(k_1)f(k_2)$ if $k_1, k_2 \in \mathcal{D}$ and k_1, k_2 are coprime.

We can then use Möbius inversion, a technique familiar to us from previous talks, to say:

$$\frac{1}{g(k)} = \sum_{d|k} f(d)$$

With this fact and doing some algebraic manipulation on (3), we get that

$$\begin{aligned} Q &= \sum_{\substack{d_1,d_2<z \\ d_1,d_2|P(z)}} \frac{1}{g((d_1,d_2))} g(d_1)\lambda(d_1)g(d_2)\lambda(d_2) \\ &= \sum_{k\in\mathcal{D}} f(k) y_k^2 \end{aligned}$$

where

$$y_k = \sum_{\substack{d\in\mathcal{D} \\ k|d}} g(d)\lambda(d)$$

Thus, Q is a quadratic form in y_k. We can further use Möbius inversion to evaluate the inside of y_k:

$$g(d)\lambda(d) = \sum_{\substack{k\in\mathcal{D} \\ d|k}} \mu\left(\frac{k}{d}\right) y_k = \mu(d) \sum_{\substack{k\in\mathcal{D} \\ d|k}} \mu(k) y_k. \tag{4}$$

In particular for $d = 1$ we get

$$\sum_{k\in\mathcal{D}} \mu(k) y_k = 1 \tag{5}$$

Define $F(z) = \sum_{k\in\mathcal{D}} \frac{1}{f(k)}$. By **Lemma 1**, the minimum value of quadratic form Q subject to linear constraint (5) is

$$\frac{1}{F(z)}$$

and this minimum is attained when

$$y_k = \frac{\mu(k)}{F(z)f(k)}$$

Insert these values of y_k into (4) to compute $\lambda(d)$:

$$\begin{aligned}\lambda(d) &= \frac{\mu(d)}{g(d)} \sum_{\substack{k\in\mathcal{D}\\ d|k}} \mu(k) y_k \\ &= \frac{\mu(d) F_d(z)}{f(d)g(d)F(z)},\end{aligned}$$

where

$$F_d(z) = \sum_{\substack{l<z/d\\ dl|P(z)}} \frac{1}{f(l)}$$

and it follows that

$$|\lambda(d)| = \frac{F_d(z)}{f(d)g(d)F(z)} \leq 1.$$

So we have a minimum value, i.e. a lower bound on Q which is $\frac{1}{F(z)}$:

$$S(A,\mathcal{P},z) \leq \frac{|A|}{F(z)} + R$$

Now, we deal with R.

Taking from an exercise in the chapter, we use without proof that for any square-free integer d there are exactly $3^{\omega(d)}$ pairs of positive integers d_1, d_2 such that the LCM of d_1 and d_2 is d.

If $d_1, d_2 < z$, then $d < z^2$ because the LCM of d_1 and d_2 is less than z^2. If d_1 and d_2 divide $P(z)$, then d is a square-free number that also divides $P(z)$. Therefore,

$$\begin{aligned}|R| &= \left| \sum_{\substack{d_1,d_2<z\\ d_1,d_2|P(z)}} \lambda(d_1)\lambda(d_2) r([d_1,d_2]) \right| \\ &\leq \sum_{\substack{d_1,d_2<z\\ d_1,d_2|P(z)}} |r([d_1,d_2])| \\ &\leq \sum_{\substack{d<z^2\\ d|P(z)}} 3^{\omega(d)} |r_d|.\end{aligned}$$

Thus, we have dealt with R:

$$S(A,\mathcal{P},z) \leq \frac{|A|}{F(z)} + \sum_{\substack{d<z^2\\ d|P(z)}} 3^{\omega(d)} |r_d|$$

Finally, we must show that $F(z) \geq G(z)$ in order to complete the proof. We will leave the details of this to the reader and the full proof can be found in Nathanson §7.2

3 Elementary Application to Goldbach Conjecture

Theorem 2. *Let N be an even integer, and let $r(N)$ denote the number of representations of N as the sum of two primes. Then*

$$r(N) << \frac{N}{(log\ N)^2} \prod_{p|N} (1 + \frac{1}{p}),$$

where the implied constant is absolute.

Let

$$a_n = n(N - n)$$

Then

$$A = \{a_n\}_{n=1}^{N}$$

is a finite sequence of integers with $|A| = N$ terms. Let $\mathcal{P}$ be the set of all prime numbers. Let

$$2 < z \leq \sqrt{N}.$$

The sieving function $S(A, \mathcal{P}, z$ denotes the number of terms of the sequence A that are divisible by no prime $p < z$. If

$$\sqrt{N} < n < N - \sqrt{N},$$

and if $a_n \equiv 0 \pmod{p}$ for some prime $p < z$, then either n or $N - n$ is composite. This implies that

$$r(N) \leq 2\sqrt{N} + S(A, \mathcal{P}, z).$$

We get the $2\sqrt{N}$ for the possible n which do not satisfy $\sqrt{N} < n < N - \sqrt{N}$ (i.e. the n which fall outside of that inequality). We get $S(A, \mathcal{P}, z)$ because if a_n is "counted" by the sieve, that means it is not divisible by any primes $p < z \leq \sqrt{N}$ and thus n and $(N - n)$ are not divisible by any such primes. This means n and $N - n$ could both be prime and thus a valid representation of N as the sum of two primes.

Now we can use the Selberg sieve. First we define our multiplicative function $g(p)$:

$$g(p) = \begin{cases} 2/p & \text{if } p \text{ does not divide } N \\ 1/p & \text{if } p \text{ divides } N \end{cases}$$

We see that this function satisfies the condition required by the Selberg sieve of $0 < g(p) < 1$ (since $g(2) = 1/2$)

Further, $a_n \equiv 0 \pmod{p}$ if and only if $n \equiv 0$ or $n \equiv N \pmod{p}$. Notice that if $p \nmid N$, then $N \not\equiv 0 \pmod{p}$ and if $p|n$, then $N \equiv 0$.

$$d = p_1 ... p_k q_1 ... q_l$$

be a squarefree integer, where p_i divide N and q_i do not. Then

$$g(d) = \frac{2^l}{d}$$

a_n is divisible by d if and only if a_n is divisible by every prime p_i that divides d. Now, notice that since each q_j does not divide N, we have two distinct congruence classes (0 mod p or N mod p). Thus, modulo d (i.e. within the set of numbers less than d), we have two choices of congruence class for each p_j (there are l such choices). Basic combinatorics tells us that there are 2^l distinct congruence classes that emerge.

Therefore,

$$|A_d| = |A|g(d) + r(d),$$

where A_d is the set of a_n which are divisible by d and

$$|r(d)| \leq 2^l \leq 2^{\omega(d)}. \tag{6}$$

We can now start to fill in parts of our Selberg sieve. Here we obtain a bound on $G(z)$: Let

$$m = \prod_{i=1}^{k} p_i^{r_i} \prod_{j=1}^{l} q_j^{s_j}$$

where the primes p_i divide N and the primes q_j do not divide N. Then

$$g(m) = \prod_{i=1}^{k} \left(\frac{1}{p_i}\right)^{r_i} \prod_{j=1}^{l} \left(\frac{2}{q_j}\right)^{s_j} = \frac{2^{s_1+\ldots s_l}}{m}.$$

Let $d_N(m)$ denote the number of positive divisors of m that are relatively prime to N. Then

$$d_N(m) = d\left(\prod_{j=1}^{l} q_j^{s_j}\right) = \left(\prod_{j=1}^{l}(s_j+1)\right) \leq \prod_{j=1}^{m} 2^{s_j} = 2^{s_1+\ldots s_l}$$

Therefore,

$$g(m) \geq \frac{d_N(m)}{m},$$

and so

$$G(z) \geq \sum_{m<z} \frac{d_N(m)}{m}.$$

We are then able to simplify:

$$\prod_{p|N}\left(1-\frac{1}{p}\right)^{-1} G(z) \geq \sum_{m<z}\frac{d_N(m)}{m}\sum_{\substack{t=1\\ p|t \implies p|N}}^{\infty}\frac{1}{t}$$
$$\geq \sum_{w<z}\frac{1}{w}\sum_{\substack{m|w\\ p|(w/m) \implies p|N}} d_N(m).$$

Let

$$w = \prod_{i=1}^{k} p_i^{u_i}\prod_{j=1}^{l} q_j^{v_j}$$

and

$$m = \prod_{i=1}^{k} p_i^{r_i}\prod_{j=1}^{l} q_j^{s_j}$$

where the p_i divide N and q_j do not. Since m divides w, it follows that $0 \leq r_i \leq u_i$ for all i, $0 \leq s_j \leq v_j$ for all j, and

$$\frac{w}{m} = \prod_{i=1}^{k} p_i^{u_i-r_i}\prod_{j=1}^{l} q_j^{v_j-s_j}$$

Since every prime divisor of w/m divides N, it follows that no prime q_j divides w/m, and so $s_j = v_j$ for all j. Therefore,

$$m = \prod_{i=1}^{k} p_i^{r_i}\prod_{j=1}^{l} q_j^{v_j}$$

and

$$d_N(m) = \prod_{j=1}^{l}(v_j+1).$$

For each integer w, the number of such divisors m is

$$\prod_{i=1}^{l}(u_i+1)$$

It follows that for every positive integer $w < z$, we have

$$\sum_{\substack{m|w\\ p|(w/m) \implies p|N}} d_N(m) = d(w)$$

where $d(w)$ counts the number of positive divisors of w. Let $z = N^{1/8}$ and we can use the theorem from page 313 in Nathanson

$$\sum_{n \leq x} \frac{d(n)}{n} = \frac{1}{2}(\log\ x)^2 + O(\log\ x).$$

to obtain the following:

$$\prod_{p|N} \left(1 - \frac{1}{p}\right)^{-1} G(z) \geq \sum_{w<z} \frac{d(w)}{w} \gg (\log\ z)^2 \gg (\log\ N)^2.$$

Equivalently,

$$\begin{aligned} \frac{|A|}{G(z)} &\ll \frac{N}{(\log\ N)^2} \prod_{p|N} \left(1 - \frac{1}{p}\right)^{-1} \\ &= \frac{N}{(\log\ N)^2} \prod_{p|N} \left(1 - \frac{1}{p^2}\right)^{-1} \prod_{p|N} \left(1 + \frac{1}{p}\right) \\ &\ll \frac{N}{(\log\ N)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right) \end{aligned}$$

All that remains is to deal with the remainder term of the Selberg sieve:

$$\sum_{\substack{d<z^2 \\ d|P(z)}} 3^{\omega(d)} |r(d)| \leq \sum_{\substack{d<z^2 \\ d|P(z)}} 3^{\omega(d)} 2^{\omega(d)} \leq \sum_{d<z^2} 6^{\omega(d)}.$$

The first step follows from (6) and the second step is fairly obvious - even if every $d < z^2$ divided $P(z)$, $\sum_{\substack{d<z^2 \\ d|P(z)}} 3^{\omega(d)} 2^{\omega(d)}$ would still be at most $\sum_{d<z^2} 6^{\omega(d)}$.

Since

$$2^{\omega(d)} \leq d$$

and

$$6^{\omega(d)} = (2^{\omega(d)})^{\log\ 6/\ \log\ 2} \leq d^{\log\ 6/\ \log\ 2} < z^{2\log\ 6/\ \log\ 2},$$

it follows that

$$R \leq \sum_{d<z^2} z^{2\log\ 6/\ \log\ 2} < z^{2+2\log\ 6/\ \log\ 2} < z^{7.2} = N^{9/10}$$

since $z = N^{1/8}$
It follows further that

$$S(A, \mathcal{P}, z) \ll \frac{N}{(\log\ N)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right) + N^{9/10} \ll \frac{N}{(\log\ N)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right)$$

So we have shown that

$$r(N) \leq 2\sqrt{N} + S(A, \mathcal{P}, z) \ll \frac{N}{(\log\ N)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right)$$

4 Application to Density of Twin Primes

Theorem 3. *Let N be a positive even integer, and let $\pi_N(x)$ denote the number of primes p up to x such that $p+N$ is also prime. Then*

$$\pi_N(x) << \frac{x}{(log\ x)^2} \prod_{p|N} (1 + \frac{1}{p}),$$

where the implied constant is absolute

The proof is almost identical to that of **Theorem 2**. In the case where $N = 2$, we obtain an interesting result as it applies to twin primes:

Theorem 4. *Let $\pi_2(x)$ denote the number of twin primes up to x. Then*

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

Notice that this is a stronger result than that of Brun's sieve which had the bound

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