Chen's Theorem

Robert Baxter
Columbia University

Spring 2024

1 Introduction

In today's talk, we will be proving one of the most famous results in additive prime number theory: Chen's theorem. This theorem gets us closer to solving Goldbach's conjecture, which we looked at a few weeks ago in Jinoo's talk. Goldbach's conjecture suggests that every even number is the sum of two primes.

Chen's theorem tells us that any large even number can be broken down into a prime and another number that's almost prime—meaning it's either a prime itself or made by multiplying two primes together.

We'll go through some basic sieve methods that help us understand how primes are spread out, and we'll discuss important terms like the representation function $r(N)$ and the singular series $\mathfrak{S}(N)$. These will be key ideas in proving Chen's theorem.

In the end, we'll see how all parts of this theorem fit together. By understanding Chen's theorem, we learn more about prime numbers and take a step forward in solving the Goldbach Conjecture.

2 Primes and Almost Primes

Theorem 1 (Chen's Theorem)**.** *Every sufficiently large even integer can be written as the sum of an odd prime and a number that is either prime or the product of two primes.*

An integer that is the product of at most r not necessarily distinct prime numbers is called an *almost prime of order* r, denoted P_r. Chen's theorem can be expressed as follows:

$$N = p + P_2$$

for every sufficiently large even integer N. We shall demonstrate not only that every large even integer N has at least one representation as the sum of a prime and an almost prime of order two but that there are, in fact, multiple such representations.

Theorem 2 (Chen)**.** *Let $r(N)$ denote the number of representations of N in the form*

$$N = p + n,$$

where p is an odd prime and n is the product of at most two primes. Then

$$r(N) \gg \mathfrak{S}(N) \frac{2N}{(\log N)^2}.$$

where,

$$\mathfrak{S}(N) = \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{p|N} \frac{p-1}{p-2}$$

3 Weights

In this next section we're going to talk about Weights, which in Sieve theory, helps us focus on numbers with certain prime factorization properties. In the context of Chen's theorem, we assign a weight to every positive integer to better understand the distribution of almost primes. These weights will play an important role in our sieving process.

Weights are going to help us get a pretty good estimate on the amount of prime and almost prime numbers there are up to an even integer N.

Let N be an even integer, $N \geq 4^8$. We begin by assigning a weight $w(n)$ to every positive integer n. Let

$$z = N^{1/8} \tag{1}$$

and

$$y = N^{1/3}. \tag{2}$$

Then $z \geq 4$. We define

$$w(n) = 1 - \frac{1}{2} \sum_{z \leq q < n} k - \frac{1}{2} \sum_{\substack{p_1 p_2 p_3 = n \\ z \leq p_1 < y \leq p_2 \leq p_3}} \tag{3}$$

This equation assigns a weight to each integer and sifts out all the less desirable values. Our goal is to use this to identify which integers are prime and almost prime. To help, we can express these three sums as sieving functions. Here is a lower bound for $r(N)$ in terms of sieving functions.

Theorem 3. *For every N, we have*

$$r(N) > S(A, \mathcal{P}, z) - \frac{1}{2} \sum_{z \leq q < y} S(A_q, \mathcal{P}, z) - \frac{1}{2} S(B, \mathcal{P}, y) - 2N^{7/8} - N^{1/3}.$$

4 Sieve Notation and its Implications

In applying the linear sieve to estimate the three sieving functions, we choose the multiplicative function:

$$g(d) = g_n(d) = \frac{1}{\varphi(d)}$$

Theorem 10.3 Let N be an even positive integer, and let

$$V(z) = \prod_{p | P(z)} (1 - g(p)) = \prod_{\substack{p<z \\ (p,N=1)}} \left(1 - \frac{1}{p-1}\right)$$

Then

$$V(z) = \mathfrak{S}(N) \frac{e^{-\gamma}}{logz} \left(1 + O\left(\frac{1}{\log N}\right)\right),$$

.

Proof. Let

$$W(z) = \prod_{2<p<z} \left(1 - \frac{1}{p-1}\right).$$

Then,

$$\frac{V(z)}{W(z)} = \prod_{\substack{p>2 \\ p|N}} \frac{p-1}{p-2} \prod_{\substack{p>z \\ p|N}} \left(1 - \frac{1}{p-1}\right).$$

Since $1 - x > e^{-2x}$ for $0 < x < \frac{\log(2)}{2}$ and $1 - x < e^{-x}$ for all x, we have

$$\prod_{\substack{p>z \\ p|N}} \left(1 - \frac{1}{p-1}\right) > \prod_{\substack{p>z \\ p|N}} \exp\left(-\frac{2}{p-1}\right)$$

Which simplifies to,

$$> 1 - \frac{8 \log N}{N^{1/8}}.$$

Thus,

$$\frac{V(z)}{W(z)} = \prod_{\substack{p>2 \\ p|N}} \frac{p-1}{p-2} \left(1 + O\left(\frac{\log N}{N^{1/8}}\right)\right).$$

To estimate W(z), we use Merten's Formula from chapter 6 to obtain,

$$-2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \frac{e^{-\gamma}}{logz} \left(1 + O\left(\frac{1}{\log N}\right)\right).$$

Therefore,

$$V(z) = \frac{V(z)}{W(z)} W(z)$$

$$= \mathfrak{S}(N) \frac{e^{-\gamma}}{logz} \left(1 + O\left(\frac{1}{\log N}\right)\right).$$

5 Sieve Estimates

Theorem 10.4 (A Lower Bound For $S(A, \mathcal{P}, z)$)

$$S(A, \mathcal{P}, z) > \left(\frac{e^{\gamma} \log 3}{2} + O(\varepsilon)\right) \frac{NV(z)}{\log N}.$$

Proof (Condensed). Applying the linear sieve, we start with the prime number theorem to approximate $|A|$, the size of the sieving set, considering contributions from primes that do not divide N:

$$|A| = \frac{N}{\log N} + O\left(\frac{N}{\log^2 N}\right).$$

For the error term $r(d)$, we capture the deviation from the expected distribution of primes:

$$r(d) = |A_d| - \frac{|A|}{\phi(d)} + O(\log N).$$

We then use the Jurkat–Richert theorem to handle the main term, where X and $V(z)$ are defined earlier in the text, and the function $f(s)$ governs the linear sieve's effectiveness:

$$X = V(z) \frac{N}{\log N} \left(1 + O\left(\frac{1}{\log N}\right)\right),$$

$$f(s) \geq \frac{2e^\gamma \log s - e^\gamma \log 3}{s}.$$

Finally, the remainder term R is bounded using the Bombieri–Vinogradov theorem to address the error sum over $r(d)$:

$$R \ll \frac{N}{(\log N)^3}.$$

Combining these estimates, we conclude that:

$$S(A, \mathcal{P}, z) > f(s)X - R,$$

which after substituting the estimates for $f(s)$ and R, yields the lower bound claimed in the theorem,

$$S(A, \mathcal{P}, z) > \left(\frac{e^\gamma \log 3}{2} + O(\varepsilon) \right) \frac{NV(z)}{\log N}.$$

Theorem 10.5 (An Upper Bound For $S(A_q, \mathcal{P}, z)$)

$$\sum_{z \leq q < y} S(A_q, \mathcal{P}, z) < \left(\frac{e^\gamma \log 6}{2} + O(\varepsilon) \right) \frac{NV(z)}{\log N}.$$

Proof (Condensed). To establish an upper bound for the sum of the sieving function $S(A_q, \mathcal{P}, z)$, we aggregate the contributions from each prime q within the interval $[z, y)$. The analysis begins by considering the Jurkat–Richert theorem to handle the main term and applying the Bombieri–Vinogradov theorem to control the error terms.

The main term involves the product of the individual sieving functions, each associated with a prime q, and is influenced by the distribution of primes in arithmetic progressions. This term is bounded by a function $f(s)$ that accounts for the effectiveness of the linear sieve, given by

$$f(s) \approx \frac{e^\gamma \log 6}{2} + O\left(\frac{\log \log N}{\log N} \right),$$

where s is a parameter derived from D and z, representing the sieving level.

For the remainder term, denoted R, the Bombieri–Vinogradov theorem provides a way to estimate the cumulative error across all relevant primes. By bounding the sum of the error terms $r(d)$, we assure that the contribution from the remainder is negligible compared to the main term, leading to the overall upper bound for the sieving function sum:

$$R \ll \frac{N}{(\log N)^3}.$$

Incorporating the bounds for both the main term and the remainder, we conclude that the sum of the sieving functions is constrained by the estimated upper bound as stated in the theorem.

Theorem 10.6 (An Upper Bound For $S(B, \mathcal{P}, y)$)

$$S(B, \mathcal{P}, y) < \left(\frac{ce^{\gamma} + O(\varepsilon)}{2} \right) \frac{NV(z)}{\log N} + O\left(\frac{e^{-N}}{(\log N)^3} \right).$$

Proof (Condensed). For the sieving function $S(B, \mathcal{P}, y)$, we begin by partitioning the primes into disjoint intervals using a parameter ℓ of the form $\ell = z(1 + e^k)$. This lets us form subsets $B^{(\ell)}$ with properties that facilitate the sieving process.

We proceed by bounding the cardinality of B using the prime number theorem and express $|B|$ in terms of the subsets $B^{(\ell)}$:

$$|B| \leq \sum_{\ell} |B^{(\ell)}|,$$

where the error term $r^{(\ell)}(d)$ is small due to the Bombieri–Vinogradov theorem:

$$r^{(\ell)}(d) \ll \frac{N}{(\log N)^3}.$$

Next, we apply the Jurkat–Richert theorem, which provides an upper bound on $S(B^{(\ell)}, \mathcal{P}, y)$ involving the function $g(d) = g^{(\ell)}(d) - \frac{1}{\phi(d)}$ with support level D:

$$S(B, \mathcal{P}, y) \leq \sum_{\ell} S(B^{(\ell)}, \mathcal{P}, y),$$

$$S(B^{(\ell)}, \mathcal{P}, y) < F(s) \cdot e^{e^{\gamma} + O(\varepsilon)} |B^{(\ell)}| V(z) + O\left(\frac{N}{(\log N)^4} \right),$$

where s is a parameter defined earlier, and $F(s)$ is a sieving function from the Jurkat–Richert theorem.

Incorporating these estimates, and after simplifying the expressions using properties of logarithmic and Riemann–Stieltjes integrals, we arrive at the stated upper bound for $S(B, \mathcal{P}, y)$.

6 Proving Chen's Theorem

Bringing together everything have looked at thus far, we can finally prove Chen's theorem.

Proof. (Chen's Theorem) It follows from the formula for $V(z)$ in Theorem 10.3 that

$$\frac{NV(z)}{\log N} = \mathfrak{S}(N)\left(\frac{8e^{-\gamma}N}{(\log N)^2}\right)\left(1 + O\left(\frac{1}{\log N}\right)\right).$$

Theorem 10.2 gives a lower bound for $r(N)$ in terms of three sieving functions. Using the estimates for the three sieving functions we looked at before, we obtain

$$\begin{aligned} r(N) &> S(A,P,z) - \frac{1}{2}\sum_{q\leq y} S(A_q,P,z) - \frac{1}{2}S(B,P,y) - 2N^{7/8} - N^{1/3} \\ &> (2\log 3 - \log 6 - c - O(\epsilon))\,\frac{e^{\gamma}NV(z)}{4\log N} \\ &\quad + O\left(\frac{e^{-\gamma}N}{(\log N)^3}\right) - 2N^{7/8} - N^{1/3} \\ &> (2\log 3 - \log 6 - c - O(\epsilon))\,\mathfrak{S}(N)\frac{2N}{(\log N)^2}\left(1 + O\left(\frac{1}{\log N}\right)\right) \\ &\quad + O\left(\frac{e^{-\gamma}N}{(\log N)^3}\right) - 2N^{7/8} - N^{1/3}. \end{aligned}$$

Since

$$2\log 3 - \log 6 - c = 0.042\ldots > 0,$$

we can choose ϵ such that $0 < \epsilon < 1/200$ and

$$2\log 3 - \log 6 - c - O(\epsilon) > 0.$$

For this fixed value of ϵ, we have

$$O\left(\frac{e^{-\gamma}N}{(\log N)^3}\right) = O\left(\frac{N}{(\log N)^3}\right).$$

Then

$$r(N) \gg \mathfrak{S}(N)\frac{2N}{(\log N)^2}.$$

This completes the proof of Chen's theorem.