Yet Another Talk about Prime Numbers

And my first experience with Beamer

Gareth Ma

University of Warwick

March 2, 2024

Warwick Imperial Mathematics Ponference (WIMP)

Disclaimer

- I am not good at Beamer
- This talk is not meant to be "formal" or about "difficult maths"
- Instead, it is a taster of some analytic number theory!

Outline

Motivation

Some prime numbers:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, \cdots, 10^9 + 7, \cdots$$

Question 1

How many primes are there $\leq n$?

Motivation

Some prime numbers:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, \cdots, 10^9 + 7, \cdots$$

Question 1

How many primes are there $\leq n$?

Motivation

Some prime numbers:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, \cdots, 10^9 + 7, 10^9 + 9, \cdots$$

Question 2

How many twin primes are there $\leq n$?

Motivation

Some prime numbers:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, \cdots, 10^9 + 7, 10^9 + 9, \cdots$$

Question 2

How many twin primes are there $\leq n$?

Formulation

The Eratosthenes sieve is a very efficient method of *sieving* primes, especially when performed by a computer. How do we phrase it mathematically?

Let us setup some notations:

- We fix N.
- Write $\mathcal{P} := \{2, 3, 5, 7, \cdots\}$ for the primes.
- Write $\mathcal{A} := \{2, 3, 4, \cdots, N-1, N\}$.
- Write $\mathcal{S}_d := \left\{d, 2d, \cdots, \lfloor \frac{N}{d} \rfloor d\right\}$ for the multiples of d not exceeding N.

We include d in $\mathcal{S}_d$ to ease computation.

The *sieving* operation we performed can be expressed as the set

$$\left\{p \in \mathcal{P} : \lfloor \sqrt{N} \rfloor < p \leq N\right\} = \mathcal{A} \setminus \bigcup_{d=2}^{\lfloor \sqrt{N} \rfloor} \mathcal{S}_d$$

Formulation

The Eratosthenes sieve is a very efficient method of *sieving* primes, especially when performed by a computer. How do we phrase it mathematically?

Let us setup some notations:

- We fix N.
- Write $\mathcal{P} \coloneqq \{2, 3, 5, 7, \cdots\}$ for the primes.
- Write $\mathcal{A} \coloneqq \{2, 3, 4, \cdots, N-1, N\}$.
- Write $\mathcal{S}_d \coloneqq \left\{d, 2d, \cdots, \lfloor \frac{N}{d} \rfloor d\right\}$ for the multiples of d not exceeding N.

We include d in $\mathcal{S}_d$ to ease computation.

The *sieving* operation we performed can be expressed as the set

$$\left\{p \in \mathcal{P} : \lfloor \sqrt{N} \rfloor < p \leq N\right\} = \mathcal{A} \setminus \bigcup_{d=2}^{\lfloor \sqrt{N} \rfloor} \mathcal{S}_d$$

Formulation

The Eratosthenes sieve is a very efficient method of *sieving* primes, especially when performed by a computer. How do we phrase it mathematically?

Let us setup some notations:

- We fix N.
- Write $\mathcal{P} := \{2, 3, 5, 7, \cdots\}$ for the primes.
- Write $\mathcal{A} := \{2, 3, 4, \cdots, N-1, N\}$.
- Write $\mathcal{S}_d := \left\{d, 2d, \cdots, \lfloor \frac{N}{d} \rfloor d\right\}$ for the multiples of d not exceeding N.

We include d in $\mathcal{S}_d$ to ease computation.

The *sieving* operation we performed can be expressed as the set

$$\left\{p \in \mathcal{P} : \lfloor \sqrt{N} \rfloor < p \leq N\right\} = \mathcal{A} \setminus \bigcup_{d=2}^{\lfloor \sqrt{N} \rfloor} \mathcal{S}_d$$

Formulation

Notations:

- We fix N.
- Write $\mathcal{P} := \{2, 3, 5, 7, \cdots\}$ for the primes.
- Write $\mathcal{A} := \{2, 3, 4, \cdots, N-1, N\}$.
- Write $\mathcal{S}_d := \left\{d, 2d, \cdots, \lfloor \frac{N}{d} \rfloor d\right\}$ for the multiples of d not exceeding N.

The *sieving* operation we performed can be expressed as the set

$$\left|\left\{p \in \mathcal{P} : \lfloor \sqrt{N} \rfloor < p \leq N\right\}\right| = \left|\mathcal{A} \setminus \bigcup_{d=2}^{\lfloor \sqrt{N} \rfloor} \mathcal{S}_d\right|$$

Formulation

Notations:

- We fix N.
- Write $\mathcal{P} := \{2, 3, 5, 7, \cdots\}$ for the primes.
- Write $\mathcal{A} := \{2, 3, 4, \cdots, N-1, N\}$.
- Write $\mathcal{S}_d := \left\{d, 2d, \cdots, \lfloor \frac{N}{d} \rfloor d\right\}$ for the multiples of d not exceeding N.
- Write $\pi(x) := |\mathcal{P} \cap \mathcal{A}|$ for the number of primes not exceeding N.

The *sieving* operation we performed can be expressed as the set

$$\left|\left\{p \in \mathcal{P} : \lfloor \sqrt{N} \rfloor < p \leq N\right\}\right| = \left|\mathcal{A} \setminus \bigcup_{d=2}^{\lfloor \sqrt{N} \rfloor} \mathcal{S}_d\right|$$

Formulation

Notations:

- We fix N.
- Write $\mathcal{P} := \{2, 3, 5, 7, \cdots\}$ for the primes.
- Write $\mathcal{A} := \{2, 3, 4, \cdots, N-1, N\}$.
- Write $\mathcal{S}_d := \left\{d, 2d, \cdots, \lfloor \frac{N}{d} \rfloor d\right\}$ for the multiples of d not exceeding N.
- Write $\pi(x) := |\mathcal{P} \cap \mathcal{A}|$ for the number of primes not exceeding N.

The *sieving* operation we performed can be expressed as the set

$$\pi(N) - \pi(\lfloor \sqrt{N} \rfloor) = |\mathcal{A}| - \left| \bigcup_{d=2}^{\lfloor \sqrt{N} \rfloor} \mathcal{S}_d \right|$$

Formulation

Notations:

- We fix N.
- Write $\mathcal{P} := \{2, 3, 5, 7, \cdots\}$ for the primes.
- Write $\mathcal{A} := \{2, 3, 4, \cdots, N-1, N\}$.
- Write $\mathcal{S}_d := \left\{d, 2d, \cdots, \lfloor \frac{N}{d} \rfloor d\right\}$ for the multiples of d not exceeding N.
- Write $\pi(x) := |\mathcal{P} \cap \mathcal{A}|$ for the number of primes not exceeding N.
- Write $\mathcal{P}' := \mathcal{P} \cap [2, \lfloor \sqrt{N} \rfloor]$ for the primes not exceeding $\lfloor \sqrt{N} \rfloor$.

The *sieving* operation we performed can be expressed as the set

$$\pi(N) - \pi(\lfloor \sqrt{N} \rfloor) = |\mathcal{A}| - \left| \bigcup_{d \in \mathcal{P}'} \mathcal{S}_d \right|$$

Inclusion-Exclusion 1

Before we can continue, let us recall some set theory! *faints*

Recall the inclusion-exclusion principle:

$$\left|\bigcup_{i=1}^{m} A_i\right| = \sum_{1\le i\le m} |A_i| - \sum_{1\le i<j\le m} |A_i \cap A_j| + \sum_{1\le i<j<k\le m} |A_i \cap A_j \cap A_k| - \cdots + (-1)^m |A_1 \cap \cdots \cap A_m|$$

In full generality, we have

Set Theoretic Inclusion-Exclusion

$$\left|\bigcup_{i=1}^{m} A_i\right| = \sum_{\mathcal{I}\subset[m]} (-1)^{|\mathcal{I}|+1} \left|\bigcap_{i\in\mathcal{I}} A_i\right|$$

Inclusion-Exclusion 1

Before we can continue, let us recall some set theory! *faints*
Recall the inclusion-exclusion principle:

$$\left|\bigcup_{i=1}^{m} A_i\right| = \sum_{1\le i\le m} |A_i| - \sum_{1\le i<j\le m} |A_i \cap A_j| + \sum_{1\le i<j<k\le m} |A_i \cap A_j \cap A_k| - \cdots + (-1)^m |A_1 \cap \cdots \cap A_m|$$

In full generality, we have

Set Theoretic Inclusion-Exclusion

$$\left|\bigcup_{i=1}^{m} A_i\right| = \sum_{\mathcal{I}\subset[m]} (-1)^{|\mathcal{I}|+1} \left|\bigcap_{i\in\mathcal{I}} A_i\right|$$

Inclusion-Exclusion 1

Before we can continue, let us recall some set theory! *faints*
Recall the inclusion-exclusion principle:

$$\left|\bigcup_{i=1}^{m} A_i\right| = \sum_{1\le i\le m} |A_i| - \sum_{1\le i<j\le m} |A_i \cap A_j| + \sum_{1\le i<j<k\le m} |A_i \cap A_j \cap A_k| - \cdots + (-1)^m |A_1 \cap \cdots \cap A_m|$$

In full generality, we have

Set Theoretic Inclusion-Exclusion

$$\left|\bigcup_{i=1}^{m} A_i\right| = \sum_{\mathcal{I}\subset[m]} (-1)^{|\mathcal{I}|+1} \left|\bigcap_{i\in\mathcal{I}} A_i\right|$$

Inclusion-Exclusion 2

$$\left|\bigcup_{d=2}^{\lfloor\sqrt{N}\rfloor} \mathcal{S}_d\right| = |\mathcal{S}_2| + |\mathcal{S}_3| + |\mathcal{S}_5| - |\mathcal{S}_6| + |\mathcal{S}_7| - |\mathcal{S}_{10}| + |\mathcal{S}_{11}| + \cdots$$

$$|\mathcal{A}| - \left|\bigcup_{d=2}^{\lfloor\sqrt{N}\rfloor} \mathcal{S}_d\right| = |\mathcal{S}_1| - |\mathcal{S}_2| - |\mathcal{S}_3| - |\mathcal{S}_5| + |\mathcal{S}_6| - |\mathcal{S}_7| + |\mathcal{S}_{10}| - |\mathcal{S}_{11}| + \cdots$$

Inclusion-Exclusion 2

$$|\mathcal{A}| - \left| \bigcup_{d=2}^{\lfloor \sqrt{N} \rfloor} \mathcal{S}_d \right| = |\mathcal{S}_1| - |\mathcal{S}_2| - |\mathcal{S}_3| - |\mathcal{S}_5| + |\mathcal{S}_6| - |\mathcal{S}_7| + |\mathcal{S}_{10}| - |\mathcal{S}_{11}| + \cdots$$

The coefficient in front of each $\mathcal{S}_d$ can be determined explicitly: it is 0 if d is divisible by a square (i.e. not *squarefree*), 1 if d has even number of prime divisors, and -1 otherwise. This is the **Möbius** function, denoted $\mu(d)$:

$$\mu(d) = \begin{cases} 1 & \text{if } d = p_1 p_2 \cdots p_{2n}, p_1 < p_2 < \cdots < p_{2n} \\ -1 & \text{if } d = p_1 p_2 \cdots p_{2n+1}, p_1 < p_2 < \cdots < p_{2n+1} \\ 0 & \text{otherwise} \end{cases}$$

An important property we need later is that if m, n are coprime, then $\mu(m)\mu(n) = \mu(mn)$.

Inclusion-Exclusion

Recall that $\mathcal{S}_d$ consists of the multiples of d not exceeding N, so $|\mathcal{S}_d| = \lfloor \frac{N}{d} \rfloor$.
Also, $\mathcal{S}_d$ appears in the sum if and only if the prime factors of d is a subset of $\mathcal{P}'$. If we let $\mathscr{P} := \prod_{p \in \mathcal{P}'} p$, then we want $d \mid \mathscr{P}$. Therefore, we can write the expression compactly:

Number Theoretic Inclusion-Exclusion

$$|\mathcal{A}| - \left| \bigcup_{d=2}^{\lfloor \sqrt{N} \rfloor} \mathcal{S}_d \right| = \sum_{d \mid \mathscr{P}} \mu(d) |\mathcal{S}_d| = \sum_{d \mid \mathscr{P}} \mu(d) \lfloor \frac{N}{d} \rfloor$$

Estimation

Going back to the very beginning, we wrote

$$
\begin{aligned}
\pi(N) - \pi(\lfloor\sqrt{N}\rfloor) &= |\mathcal{A}| - \left|\bigcup_{d \in \mathcal{P}'} \mathcal{S}_d\right| \\
&= \sum_{d \mid \mathscr{P}} \mu(d) \lfloor \frac{N}{d} \rfloor \\
&= \sum_{d \mid \mathscr{P}} \mu(d) \left(\frac{N}{d} + O(1)\right) \\
&= \underbrace{N \sum_{d \mid \mathscr{P}} \frac{\mu(d)}{d}}_{\text{main term}} + \underbrace{\sum_{d \mid \mathscr{P}} \mu(d) O(1)}_{\text{error term}}
\end{aligned}
$$

Estimation

Going back to the very beginning, we wrote

$$
\begin{aligned}
\pi(N) - \pi(\lfloor\sqrt{N}\rfloor) &= |\mathcal{A}| - \left|\bigcup_{d\in\mathcal{P}'} \mathcal{S}_d\right| \\
&= \sum_{d|\mathscr{P}} \mu(d) \lfloor \frac{N}{d} \rfloor \\
&= \sum_{d|\mathscr{P}} \mu(d) \left(\frac{N}{d} + O(1)\right) \\
&= \underbrace{N \sum_{d|\mathscr{P}} \frac{\mu(d)}{d}}_{\text{main term}} + \underbrace{\sum_{d|\mathscr{P}} \mu(d) O(1)}_{\text{error term}}
\end{aligned}
$$

Estimation(Error)

$$
\begin{aligned}
\pi(N) - \pi(\lfloor\sqrt{N}\rfloor) &= \overbrace{N\sum_{d\mid\mathscr{P}} \frac{\mu(d)}{d}}^{\text{main term}} + \overbrace{\sum_{d\mid\mathscr{P}} \mu(d) O(1)}^{\text{error term}} \\
&= N\sum_{d\mid\mathscr{P}} \frac{\mu(d)}{d} + O\left(\sum_{d\mid\mathscr{P}} 1\right) \\
&= N\sum_{d\mid\mathscr{P}} \frac{\mu(d)}{d} + O\left(2^{|\mathcal{P}'|}\right) \\
&= N\sum_{d\mid\mathscr{P}} \frac{\mu(d)}{d} + O\left(2^{\pi(\lfloor\sqrt{N}\rfloor)}\right) \\
&= N\sum_{d\mid\mathscr{P}} \frac{\mu(d)}{d} + O\left(2^{\sqrt{N}}\right)
\end{aligned}
$$

Estimation(Main)

$$\pi(N) - \pi(\lfloor\sqrt{N}\rfloor) = N \sum_{d \mid \mathscr{P}} \frac{\mu(d)}{d} + O\left(2^{\sqrt{N}}\right)$$
$$= N \prod_{p \le \lfloor\sqrt{N}\rfloor} \left(\frac{\mu(1)}{1} + \frac{\mu(p)}{p}\right) + O\left(2^{\sqrt{N}}\right)$$

This is an example of an *Euler product*. For a simpler example, consider the equality

$$(1 + f(m))(1 + f(n)) = 1 + f(m) + f(n) + f(mn) = \sum_{d \mid mn} f(d)$$

When f is a multiplicative function.

Estimation(Main)

$$
\begin{aligned}
\pi(N) - \pi(\lfloor \sqrt{N} \rfloor) &= N \sum_{d \mid \mathscr{P}} \frac{\mu(d)}{d} + O\left(2^{\sqrt{N}}\right) \\
&= N \prod_{p \le \lfloor \sqrt{N} \rfloor} \left(\frac{\mu(1)}{1} + \frac{\mu(p)}{p} \right) + O\left(2^{\sqrt{N}}\right) \\
&= N \prod_{p \le \lfloor \sqrt{N} \rfloor} \left(1 - \frac{1}{p} \right) + O\left(2^{\sqrt{N}}\right) \\
&< \frac{N}{\log N} + O\left(2^{\sqrt{N}}\right)
\end{aligned}
$$

Where the last line follows from either Merten's bound or through

$$
\prod_{p \le \lfloor \sqrt{N} \rfloor} \left(1 - \frac{1}{p} \right)^{-1} = \prod_{p \le \lfloor \sqrt{N} \rfloor} \sum_{r=0}^{\infty} p^{-r} \ge \sum_{n \le \lfloor \sqrt{N} \rfloor} n^{-1} > \log\left(\lfloor \sqrt{N} \rfloor \right)
$$

Estimation(Main)

$$\pi(N) - \pi(\lfloor\sqrt{N}\rfloor) < \frac{N}{\log N} + O\left(2^{\sqrt{N}}\right)$$

Problem

Error term is larger than the main term.

Solution

Redo the computations with a smaller *sieving parameter*, i.e. replace $\lfloor\sqrt{N}\rfloor$ with some carefully chosen z.

Estimation(Speedrun ver.)

Writing $\mathscr{P}_z := \prod_{p\in\mathcal{P}|[2,z]} p$,

$$\begin{aligned}\pi(N) - \pi(z) &= \sum_{d|\mathscr{P}_z} \mu(d)\lfloor\frac{N}{d}\rfloor \\ &= N\prod_{p\leq z}\left(1-\frac{1}{p}\right) + O(2^{\pi(z)}) \\ &< \frac{N}{\log z} + O(2^z)\end{aligned}$$

Taking $z = \log N$ and noting that $2^z = N^{\log 2} = O\left(\frac{N}{\log\log N}\right)$, we get

Estimation of $\pi(N)$

$$\pi(N) = O\left(\frac{N}{\log\log N}\right)$$

Why is the result wrong?

Some of you may know about the Prime Number Theorem, which states that $\pi(N) \sim \frac{N}{\log N}$, meaning our derived asymptotic is not tight.

The reason lies in the unpredictable behaviour of $\mu(d)$. Indeed, there are several results on the "pseudorandomness" of $\mu(d)$ [4]. Even its prefix sum $\sum_{d=1}^{N} \mu(d)$ isn't well understood for a long time.

Instead, Viggo Brun suggested replacing $\mu(d)$ with better-behaved functions λ_d that allow us to control the error term better.

Why is the result wrong?

Some of you may know about the Prime Number Theorem, which states that $\pi(N) \sim \frac{N}{\log N}$, meaning our derived asymptotic is not tight.

The reason lies in the unpredictable behaviour of $\mu(d)$. Indeed, there are several results on the "pseudorandomness" of $\mu(d)$ [4]. Even its prefix sum $\sum_{d=1}^{N} \mu(d)$ isn't well understood for a long time.

Instead, Viggo Brun suggested replacing $\mu(d)$ with better-behaved functions λ_d that allow us to control the error term better.

Why is the result wrong?

Some of you may know about the Prime Number Theorem, which states that $\pi(N) \sim \frac{N}{\log N}$, meaning our derived asymptotic is not tight.

The reason lies in the unpredictable behaviour of $\mu(d)$. Indeed, there are several results on the "pseudorandomness" of $\mu(d)$ [4]. Even its prefix sum $\sum_{d=1}^{N} \mu(d)$ isn't well understood for a long time.

Instead, Viggo Brun suggested replacing $\mu(d)$ with better-behaved functions λ_d that allow us to control the error term better.

Brun's Pure Sieve

Brun's idea is to *truncate* the inclusion-exclusion process in order to reduce the number of terms.

Let us write $\omega(d)$ for the number of prime divisors of d. We can group the terms in the number theoretic inclusion-exclusion by $\omega(d)$, corresponding to the size of the subsets $\mathcal{I}$, as follows.

$$\sum_{d \mid \mathscr{P}_z} \mu(d) \lfloor \frac{N}{d} \rfloor = \sum_{k=0}^{|\mathscr{P}_z|} \sum_{\substack{d \mid \mathscr{P}_z \\ \omega(d)=k}} \mu(d) \lfloor \frac{N}{d} \rfloor = \sum_{k=0}^{|\mathscr{P}_z|} (-1)^k \sum_{\substack{d \mid \mathscr{P}_z \\ \omega(d)=k}} \lfloor \frac{N}{d} \rfloor$$

Brun's Pure Sieve

Brun observes that if we truncate the sum to $k \le r$ for some threshold r, this will actually provide a lowerbound and an upperbound, depending purely on the sign of r! This is not trivial but intuitively plausible. Instead of $\pi(N) - \pi(z) = \sum_{d \mid \mathscr{P}_z} \mu(d) \lfloor \frac{N}{d} \rfloor$, we now have

$$\begin{cases} \pi(N) - \pi(z) \ge \sum_{k=0}^{r} \sum_{d \mid \mathscr{P}_z} \mu(d) \lfloor \frac{N}{d} \rfloor & \text{if } r \text{ is odd} \\ \pi(N) - \pi(z) \le \sum_{k=0}^{r} \sum_{d \mid \mathscr{P}_z} \mu(d) \lfloor \frac{N}{d} \rfloor & \text{if } r \text{ is even} \end{cases}$$

Less terms yields better control on the error term. In particular, you can get $\pi(N) \sim O\left(\frac{N}{\log N}\right)$, asymptotically correct up to an unspecified constant.

Bonus: You can obtain a weakened Theorem (explained later):

$$\#\{n \le N : \omega(n), \omega(n+2) \le 7\} \gg \frac{N}{\log^2 N}$$

Brun's Pure Sieve

Brun observes that if we truncate the sum to $k \leq r$ for some threshold r, this will actually provide a lowerbound and an upperbound, depending purely on the sign of r! This is not trivial but intuitively plausible. Instead of $\pi(N) - \pi(z) = \sum_{d|\mathscr{P}_z} \mu(d)\lfloor \frac{N}{d} \rfloor$, we now have

$$\begin{cases} \pi(N) - \pi(z) \geq \sum_{k=0}^{r} \sum_{d|\mathscr{P}_z} \mu(d)\lfloor \frac{N}{d} \rfloor & \text{if } r \text{ is odd} \\ \pi(N) - \pi(z) \leq \sum_{k=0}^{r} \sum_{d|\mathscr{P}_z} \mu(d)\lfloor \frac{N}{d} \rfloor & \text{if } r \text{ is even} \end{cases}$$

Less terms yields better control on the error term. In particular, you can get $\pi(N) \sim O\left(\frac{N}{\log N}\right)$, asymptotically correct up to an unspecified constant.

Bonus: You can obtain a weakened Theorem (explained later):

$$\#\{n \leq N : \omega(n), \omega(n+2) \leq 7\} \gg \frac{N}{\log^2 N}$$

Brun's Pure Sieve

I will briefly mention some further optimisations.

- (Brun) We replace $\mu(d)$ with the truncated $\mu_r(d)$.
- (Brun?) We can replace $\mu(d)$ with a weighted version $\mu(d)g(d)$.
- (Selberg Λ^2) We can replace $\mu(d)$ with a sequence λ_d, such that the **quadratic form** $\sum_{d_1,d_2 \leq z} \frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]}$ is minimised.
- (Goldston, Pintz, Yildirim (GPY)) We can use special weights motivated by the Selberg weights.
- (Maynard) We can use **multidimensional** sieves! James used this to improve the bound on gaps between primes.

State of the Art

Sieve theory is *much* more general than what I demonstrated.

For example, one can tackle the twin prime conjecture by replacing the sieving sets $\mathcal{S}_d = d\mathbb{N}$ with $\mathcal{S}_p := \{n : p \mid n(n+2)\}$. Combined with the Brun's sieve, one gets the result mentioned before, where

Twin not-primes

$$\#\{n \leq N : \omega(n), \omega(n+2) \leq 7\} \gg \frac{N}{\log^2 N}$$

The constant 7 appears because using notation before, we set $z > N^{1/8}$, meaning only prime divisors $> z > N^{1/8}$ remains, of which there can only be 7 of.

Even though we don't know how to prove the twin prime conjecture, the more general question of "bounded gaps between primes", which asks "what is the smallest prime gap that appears infinitely often", has been tackled using the GPY and Maynard sieves.

State of the Art

Sieve theory is *much* more general than what I demonstrated.

For example, one can tackle the twin prime conjecture by replacing the sieving sets $\mathcal{S}_d = d\mathbb{N}$ with $\mathcal{S}_p := \{n : p \mid n(n+2)\}$. Combined with the Brun's sieve, one gets the result mentioned before, where

Twin not-primes

$$\#\{n \le N : \omega(n), \omega(n+2) \le 7\} \gg \frac{N}{\log^2 N}$$

The constant 7 appears because using notation before, we set $z > N^{1/8}$, meaning only prime divisors $> z > N^{1/8}$ remains, of which there can only be 7 of.

Even though we don't know how to prove the twin prime conjecture, the more general question of "bounded gaps between primes", which asks "what is the smallest prime gap that appears infinitely often", has been tackled using the GPY and Maynard sieves.

State of the Art

Sieve theory is *much* more general than what I demonstrated.

For example, one can tackle the twin prime conjecture by replacing the sieving sets $\mathcal{S}_d = d\mathbb{N}$ with $\mathcal{S}_p := \{n : p \mid n(n+2)\}$. Combined with the Brun's sieve, one gets the result mentioned before, where

Twin not-primes

$$\#\{n \le N : \omega(n), \omega(n+2) \le 7\} \gg \frac{N}{\log^2 N}$$

The constant 7 appears because using notation before, we set $z > N^{1/8}$, meaning only prime divisors $> z > N^{1/8}$ remains, of which there can only be 7 of.

Even though we don't know how to prove the twin prime conjecture, the more general question of "bounded gaps between primes", which asks "what is the smallest prime gap that appears infinitely often", has been tackled using the GPY and Maynard sieves.

State of the Art

Finally, let's talk about the Goldbach conjecture! Most progress on it comes in the form of "$m + n$ Theorems":

$m + n$ Theorem

All even integers ≥ 4 can be written as the sum of an integer with at most m prime divisors, and another integer with at most n prime divisors.

- Brun (1920): $9 + 9$ Theorem
- Buchstab: $5 + 5$ Theorem
- Buchstab, Selberg: $4 + 4$ Theorem
- Kuhn's weighted sieve: $2 + 3$ Theorem
- Using GRH: $1 + 7$ Theorem
- Chen JingRun (1966 in Chinese, 1973 in English): $1 + 2$ Theorem

As usual, these mysterious numbers come from the choice of the sieving parameter z: if we sieve away multiples of primes $\leq (z \geq N^{1/10})$ from $\mathcal{A}$, then the remaining numbers will all have prime divisors at least $N^{1/10}$, meaning there can at most be 9 divisors. This is how Brun's result is achieved.

State of the Art

Finally, let's talk about the Goldbach conjecture! Most progress on it comes in the form of "$m + n$ Theorems":

$m + n$ Theorem

All even integers ≥ 4 can be written as the sum of an integer with at most m prime divisors, and another integer with at most n prime divisors.

- Brun (1920): $9 + 9$ Theorem
- Buchstab: $5 + 5$ Theorem
- Buchstab, Selberg: $4 + 4$ Theorem
- Kuhn's weighted sieve: $2 + 3$ Theorem
- Using GRH: $1 + 7$ Theorem
- Chen JingRun (1966 in Chinese, 1973 in English): $1 + 2$ Theorem

As usual, these mysterious numbers come from the choice of the sieving parameter z: if we sieve away multiples of primes $\leq (z \geq N^{1/10})$ from $\mathcal{A}$, then the remaining numbers will all have prime divisors at least $N^{1/10}$, meaning there can at most be 9 divisors. This is how Brun's result is achieved.

Chen JingRun's seminal paper's abstract:

"In this paper we shall prove that every sufficiently large even integer is a sum of a prime and a product of at most 2 primes. The method used is simple without any complicated numerical calculations."

References

[1] A.C. Cojocaru, M.R. Murty, and London Mathematical Society. *An Introduction to Sieve Methods and Their Applications*. An Introduction to Sieve Methods and Their Applications. Cambridge University Press, 2005. ISBN: 9780521848169. URL: https://books.google.co.uk/books?id=1swo9Yf3d2YC.

[2] D. A. Goldston, J. Pintz, and C. Y. Yildirim. *Primes in Tuples I*. 2005. arXiv: math/0508185 [math.NT].

[3] George Greaves. *Sieves in Number Theory*. Springer Berlin Heidelberg, 2001. ISBN: 9783662046586. DOI: 10.1007/978-3-662-04658-6. URL: http://dx.doi.org/10.1007/978-3-662-04658-6.

[4] Ben Green and Terence Tao. "The Möbius function is strongly orthogonal to nilsequences". In: *Annals of Mathematics* (2012), pp. 541–566.

[5] Zi Hao Liu. "Sieve Method (2) - Inclusion-Exclusion and the Eratosthenes Sieve (Translated)". In: (2022). URL: https://zhuanlan.zhihu.com/p/385530517.

Any questions?

Prerequisites 1

We need a theorem.

Cauchy's Theorem #3971

If $f(z) = \sum_{n \geq 0} f_n z^n$, then

$$f_n = \frac{1}{2\pi i} \oint_{\mathcal{C}} f(z) z^{-n-1} \, \mathrm{d}z$$

Where $\mathcal{C}$ is a counterclockwise contour around the origin.

Prerequisites 2

We need a method.

Generating Functions

Let $A, B \subset \mathbb{N}$ be (multi-)sets, and $A + B := \{a + b : a \in A, b \in B\}$ be their sum-set. Define the indicator generating function $f_a(x) := \sum_{a \in A} x^a$ and $f_b(x) := \sum_{b \in B} x^b$. Then,

$$f_a(x) f_b(x) = \sum_{t \in A+B} c_t x^t$$

Where c_t is the number of ways that t can be represented as $a + b$, where $a \in A$ and $b \in B$. We write $[x^t]\,(f_a(x) f_b(x))$ to indicate the coefficient of x^t in $f_a(x) f_b(x)$.

Waring's Problem

To motivate, let we recall two classical theorems.

Sum of Four Squares Theorem

All nonnegative integers n can be written as sum of at most four squares. In fact, if we denote by $r_4(n)$ the number of ways to write n as sum of four squares, then

$$r_4(n) = 8 \sum_{\substack{m \mid n \\ 4 \nmid m}} m$$

Sum of Two Squares Theorem

A positive integer $n > 1$ can be written as sum of two squares if and only if its prime factorisation does not contain any p^e where $p \equiv 1 \pmod 4$ and e is odd. The value $r_2(n)$ relates to the factorisation in $\mathbb{Z}[i]$.

Waring's Problem

As the prerequisites slides suggested, we can tackle these two problems by the method of generating function.

For example, the value $r_4(n)$ can be computed as $[x^n]\left(\theta^4(x)\right)$, where

$$\theta(x) \coloneqq \sum_{m \in \mathbb{Z}} x^{m^2} = 1 + 2x + 2x^4 + 2x^9 + \cdots$$

This begs for modular forms, but we will see now that it doesn't help.

Waring's Problem

As the prerequisites slides suggested, we can tackle these two problems by the method of generating function.

For example, the value $r_4(n)$ can be computed as $[x^n]\left(\theta^4(x)\right)$, where

$$\theta(x) := \sum_{m \in \mathbb{Z}} x^{m^2} = 1 + 2x + 2x^4 + 2x^9 + \cdots$$

This begs for modular forms, but we will see now that it doesn't help.

Waring's Problem

As the prerequisites slides suggested, we can tackle these two problems by the method of generating function.

For example, the value $r_4(n)$ can be computed as $[x^n]\left(\theta^4(x)\right)$, where

$$\theta(x) := \sum_{m\in\mathbb{Z}} x^{m^2} = 1 + 2x + 2x^4 + 2x^9 + \cdots$$

This begs for modular forms, but we will see now that it doesn't help.

Waring's Problem

Instead, let's consider the problem in its full generality.

Sum of s k^{th} Powers

Given integer N and positive integers k and s, determine the number of solutions $(x_1, \cdots, x_s) \in \mathbb{N}^s$[1] to the equation

$$x_1^k + \cdots + x_s^k = N$$

We denote the quantity above by $r_{k,s}(N)$. [2]

[2] We only consider nonnegative integer solutions, as the rest can be obtained by symmetry.

Waring's Problem

Just like before, we can directly write down $r_{k,s}(N) = [z^N]f^s(z)$, where

$$f(z) := \sum_{n=0}^{\infty} z^{n^k}$$

Applying Cauchy's Theorem gives

$$r_{k,s}(N) = \frac{1}{2\pi i} \oint_{\mathcal{C}} f^s(z) z^{-N-1} \,\mathrm{d}z$$

Let's use a short hand $e(t)$ for $\exp(2\pi i t)$. We can parametrise the unit circle $\mathcal{C}$ by $z = e(t)$ with $t \in [0,1]$. Then, $\mathrm{d}z = 2\pi i e(t)\mathrm{d}t$, giving

$$r_{k,s}(N) = \int_0^1 f^s(e(t))e(t)^{-N-1} \cdot e(t)\mathrm{d}t = \int_0^1 \left(\sum_{n=0}^{\lfloor N^{1/s} \rfloor} e\left(tn^k\right) \right)^s e(-Nt)\mathrm{d}t$$

Waring's Problem

Just like before, we can directly write down $r_{k,s}(N) = [z^N]f^s(z)$, where

$$f(z) := \sum_{n=0}^{\infty} z^{n^k}$$

Applying Cauchy's Theorem gives

$$r_{k,s}(N) = \frac{1}{2\pi i} \oint_{\mathcal{C}} f^s(z) z^{-N-1} \, \mathrm{d}z$$

Let's use a short hand $e(t)$ for $\exp(2\pi i t)$. We can parametrise the unit circle $\mathcal{C}$ by $z = e(t)$ with $t \in [0, 1]$. Then, $\mathrm{d}z = 2\pi i e(t)\mathrm{d}t$, giving

$$r_{k,s}(N) = \int_0^1 f^s(e(t)) e(t)^{-N-1} \cdot e(t) \mathrm{d}t = \int_0^1 \left(\sum_{n=0}^{\lfloor N^{1/s} \rfloor} e\left(tn^k\right) \right)^s e(-Nt) \mathrm{d}t$$

Waring's Problem

Just like before, we can directly write down $r_{k,s}(N) = [z^N]f^s(z)$, where

$$f(z) := \sum_{n=0}^{\infty} z^{n^k}$$

Applying Cauchy's Theorem gives

$$r_{k,s}(N) = \frac{1}{2\pi i} \oint_{\mathcal{C}} f^s(z) z^{-N-1} \, \mathrm{d}z$$

Let's use a short hand $e(t)$ for $\exp(2\pi i t)$. We can parametrise the unit circle $\mathcal{C}$ by $z = e(t)$ with $t \in [0, 1]$. Then, $\mathrm{d}z = 2\pi i e(t) \mathrm{d}t$, giving

$$r_{k,s}(N) = \int_0^1 f^s(e(t)) e(t)^{-N-1} \cdot e(t) \mathrm{d}t = \int_0^1 \left(\sum_{n=0}^{\lfloor N^{1/s} \rfloor} e\left(tn^k\right) \right)^s e(-Nt) \mathrm{d}t$$

Algebraic Manipulation

Let us look closer at the *exponential sum*

$$g(t) = \sum_{n=0}^{\infty} e(tn^k), t \in [0, 1]$$

Suppose that t is (close to) a rational number $t = a/q$. Then,

$$g(t) = \sum_{n=0}^{\lfloor N^{1/s} \rfloor} e\left(\frac{an^k}{q}\right) = \sum_{n=0}^{\lfloor N^{1/s} \rfloor} e\left(\frac{an^k \pmod q}{q}\right)$$

The numerator is not uniform at all! For example, if $k = 2$, then $n^2 \pmod q$ only takes values on half of $\{0, 1, \cdots, q-1\}$. This is especially significant when q is small, where the sum might have large values compared to when q is large. This motivates us to split the previous integral into when q is small, and the rest.

Algebraic Manipulation

Let us look closer at the *exponential sum*

$$g(t) = \sum_{n=0}^{\infty} e(tn^k), t \in [0,1]$$

Suppose that t is (close to) a rational number $t = a/q$. Then,

$$g(t) = \sum_{n=0}^{\lfloor N^{1/s} \rfloor} e\left(\frac{an^k}{q}\right) = \sum_{n=0}^{\lfloor N^{1/s} \rfloor} e\left(\frac{an^k \pmod q}{q}\right)$$

The numerator is not uniform at all! For example, if $k = 2$, then $n^2 \pmod q$ only takes values on half of $\{0, 1, \cdots, q-1\}$. This is especially significant when q is small, where the sum might have large values compared to when q is large. This motivates us to split the previous integral into when q is small, and the rest.

Algebraic Manipulation

Let us look closer at the *exponential sum*

$$g(t) = \sum_{n=0}^{\infty} e(tn^k), t \in [0, 1]$$

Suppose that t is (close to) a rational number $t = a/q$. Then,

$$g(t) = \sum_{n=0}^{\lfloor N^{1/s} \rfloor} e\left(\frac{an^k}{q}\right) = \sum_{n=0}^{\lfloor N^{1/s} \rfloor} e\left(\frac{an^k \pmod q}{q}\right)$$

The numerator is not uniform at all! For example, if $k = 2$, then $n^2 \pmod q$ only takes values on half of $\{0, 1, \cdots, q-1\}$. This is especially significant when q is small, where the sum might have large values compared to when q is large. This motivates us to split the previous integral into when q is small, and the rest.

Defining Arcs

Our definitions depends on a constant $P = N^{\upsilon}$ that is optimised later.

Major & Minor Arcs

For every natural number with $1 \leq a < q \leq P$, define the *major arc* around a/q to be

$$\mathfrak{M}(q,a) := \left\{ \alpha \in \mathbb{R} : \left| \alpha - \frac{a}{q} \right| \leq PN^{-k} \right\}$$

The major arcs are then defined as $\mathfrak{M} = \bigcup_{q \leq P} \bigcup^{q}_{\substack{a=1 \\ (a,q)=1}} \mathfrak{M}(q,a)$, and the minor arcs are defined as

$$\mathfrak{m} = [0,1] \setminus \mathfrak{M}^3$$

4

[4] In actual computation, we shift the interval to $(PN^{-k}, 1 + PN^{-k})$ instead.

Now, our original integral is transformed into the sum of two:

$$r_{k,s}(N) = \int_{\mathfrak{M}} g^s(t)e(-Nt)\mathrm{d}t + \int_{\mathfrak{m}} g^s(t)e(-Nt)\mathrm{d}t$$

Using various techniques, one can prove the following two bounds

Major arc estimate

For all sufficiently large N and $s \geq 2^k + 1$,

$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha N)\,\mathrm{d}\alpha \gg N^{s/k-1}$$

and

Minor arc estimate

For all sufficiently large N and $s \geq 2^k + 1$,

$$\int_{\mathfrak{m}} f(\alpha)^s e(-\alpha N)\,\mathrm{d}\alpha = o\left(N^{s/k-1}\right)$$

Now, our original integral is transformed into the sum of two:

$$r_{k,s}(N) = \int_{\mathfrak{M}} g^s(t) e(-Nt) \mathrm{d}t + \int_{\mathfrak{m}} g^s(t) e(-Nt) \mathrm{d}t$$

Using various techniques, one can prove the following two bounds

Major arc estimate

For all sufficiently large N and $s \geq 2^k + 1$,

$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha N) \, \mathrm{d}\alpha \gg N^{s/k-1}$$

and

Minor arc estimate

For all sufficiently large N and $s \geq 2^k + 1$,

$$\int_{\mathfrak{m}} f(\alpha)^s e(-\alpha N) \, \mathrm{d}\alpha = o\left(N^{s/k-1}\right)$$

Applications & Further Improvements

The *Goldbach conjecture* has also been attacked by the circle method. More specifically, the circle method has been applied to the *weak* Goldbach conjecture, which asks whether every odd integer ≥ 7 can be written as the sum of 3 primes.
Vinogradov proved an asymptotic version of the result in 1937, and Helfgott proved the full version in 2013.

References

[1] A.C. Cojocaru, M.R. Murty, and London Mathematical Society. *An Introduction to Sieve Methods and Their Applications*. An Introduction to Sieve Methods and Their Applications. Cambridge University Press, 2005. ISBN: 9780521848169. URL: https://books.google.co.uk/books?id=1swo9Yf3d2YC.

[2] D. A. Goldston, J. Pintz, and C. Y. Yildirim. *Primes in Tuples I*. 2005. arXiv: math/0508185 [math.NT].

[3] George Greaves. *Sieves in Number Theory*. Springer Berlin Heidelberg, 2001. ISBN: 9783662046586. DOI: 10.1007/978-3-662-04658-6. URL: http://dx.doi.org/10.1007/978-3-662-04658-6.

[4] Ben Green and Terence Tao. "The Möbius function is strongly orthogonal to nilsequences". In: *Annals of Mathematics* (2012), pp. 541–566.

[5] Zi Hao Liu. "Sieve Method (2) - Inclusion-Exclusion and the Eratosthenes Sieve (Translated)". In: (2022). URL: https://zhuanlan.zhihu.com/p/385530517.

Any questions?