On the theory of prime producing sieves (part 1)

James Maynard (joint with Kevin Ford)

University of Oxford

Number Theory Web Seminar
October 2024

Introduction

Question

Given a large x and a set $\mathcal{A} \subset [x, 2x]$, how many primes are in $\mathcal{A}$?

Question

Given a large x and a set $\mathcal{A} \subset [x, 2x]$, how many primes are in $\mathcal{A}$?

Many famous problems have this form:

- $\mathcal{A} = [x, 2x]$, Prime Number Theorem/Riemann Hypothesis
- $\mathcal{A} = \{p + 2 : p \text{ prime}\} \cap [x, 2x]$, Twin prime conjecture
- $\mathcal{A} = \{2N - p : p \text{ prime}\} \cap [x, 2x]$, Goldbach conjecture
- $\mathcal{A} = \{n^2 + 1\} \cap [x, 2x]$, Primes of the form $n^2 + 1$
- $\mathcal{A} = \{n^2 + m^4\} \cap [x, 2x]$, Primes of the form $n^2 + m^4$ (Friedlander-Iwaniec Theorem)

We expect there to be many primes (roughly $\#\mathcal{A}/\log x$)

Basic Sieve Methods

Classical sieve methods give a means of studying these problems by understanding $\mathcal{A}$ in arithmetic progressions.

$$\mathcal{A}_d := \Big\{a \in \mathcal{A} : \ d|a\Big\}, \qquad P^-(a) := \min_{p|a} p.$$

Basic Sieve Methods

Classical sieve methods give a means of studying these problems by understanding $\mathcal{A}$ in arithmetic progressions.

$$\mathcal{A}_d := \{a \in \mathcal{A} : \ d|a\}, \qquad P^-(a) := \min_{p|a} p.$$

In each of the cases above, we have

$$\#\mathcal{A}_d \approx \frac{\#\mathcal{A}}{d}$$

when d is small enough.

Basic Sieve Methods

Classical sieve methods give a means of studying these problems by understanding $\mathcal{A}$ in arithmetic progressions.

$$\mathcal{A}_d := \Big\{a \in \mathcal{A} : \ d|a\Big\}, \qquad P^-(a) := \min_{p|a} p.$$

In each of the cases above, we have

$$\#\mathcal{A}_d \approx \frac{\#\mathcal{A}}{d}$$

when d is small enough. So

$$\begin{aligned}\#\{a \in \mathcal{A} : \ P^-(a) > 3\} &= \#\mathcal{A} - \#\mathcal{A}_2 - \#\mathcal{A}_3 + \#\mathcal{A}_6 \\ &\approx \#\mathcal{A}\Big(1 - \frac{1}{2}\Big)\Big(1 - \frac{1}{3}\Big).\end{aligned}$$

Replacing 3 with $x^{1/2}$ gives big divisors and too many terms.

Basic Sieve Methods

Classical sieve methods give a means of studying these problems by understanding $\mathcal{A}$ in arithmetic progressions.

$$\mathcal{A}_d := \left\{a \in \mathcal{A} : \ d|a\right\}, \qquad P^-(a) := \min_{p|a} p.$$

In each of the cases above, we have

$$\#\mathcal{A}_d \approx \frac{\#\mathcal{A}}{d}$$

when d is small enough. So

$$\begin{aligned}\#\{a \in \mathcal{A} : \ P^-(a) > 3\} &= \#\mathcal{A} - \#\mathcal{A}_2 - \#\mathcal{A}_3 + \#\mathcal{A}_6 \\ &\approx \#\mathcal{A}\Big(1 - \frac{1}{2}\Big)\Big(1 - \frac{1}{3}\Big).\end{aligned}$$

Replacing 3 with $x^{1/2}$ gives big divisors and too many terms.

Principle

We can obtain upper/lower bounds by truncating the inclusion-exclusion process using positivity $\#\mathcal{A}_d \geq 0$.

Basic Sieve Methods II

Want to choose λ_d for $d \le x^\gamma$ such that

$$\mathbb{1}_{P^-(n) \ge z} = \sum_{\substack{d|n \\ P^+(d) \le z}} \mu(d) \ge \sum_{d|n} \lambda_d.$$

Then

$$\#\Big\{a \in \mathcal{A} : \ P^-(a) \ge z\Big\} \ge \sum_{d \le x^\gamma} \lambda_d \# \mathcal{A}_d.$$

Basic Sieve Methods II

Want to choose λ_d for $d \le x^\gamma$ such that

$$\mathbb{1}_{P^-(n)\ge z} = \sum_{\substack{d|n \\ P^+(d)\le z}} \mu(d) \ge \sum_{d|n} \lambda_d.$$

Then

$$\#\{a \in \mathcal{A} : P^-(a) \ge z\} \ge \sum_{d\le x^\gamma} \lambda_d \#\mathcal{A}_d.$$

Theorem (Linear Sieve)

If

$$\sum_{d<x^\gamma} \left| \#\mathcal{A}_d - \frac{\#\mathcal{A}}{d} \right| \le \frac{\#\mathcal{A}}{(\log x)^A}, \tag{I}$$

then

$$\#\{a \in \mathcal{A} : P^-(a) \ge z\} \ge \left(f\left(\frac{\log x^\gamma}{\log z}\right) + o(1) \right) \#\mathcal{A} \prod_{p\le z} \left(1 - \frac{1}{p}\right)$$

for a continuous increasing function f with $\lim_{s\to\infty} f(s) = 1$.

Gives good results if γ is large enough compared with $\log x / \log z$.

Parity Problem

Unfortunately, this technique alone cannot detect primes.

Example (Parity Problem; Selberg)

Let

$$\mathcal{A}^- := \Big\{ n \in [x, 2x] : \textit{ n has an even number of prime factors} \Big\}.$$

Then

$$\sum_{d < x^{1-\epsilon}} \Big| \#\mathcal{A}_d^- - \frac{\#\mathcal{A}^-}{d} \Big| \le \frac{\#\mathcal{A}}{(\log x)^A},$$

but

$$\#\{p \in \mathcal{A}^-\} = 0.$$

Even best-possible estimates for $\mathcal{A}_d$ cannot produce primes!

Parity Problem

Unfortunately, this technique alone cannot detect primes.

Example (Parity Problem; Selberg)

Let

$$\mathcal{A}^- := \Big\{ n \in [x, 2x] : \text{ } n \textit{ has an even number of prime factors} \Big\}.$$

Then

$$\sum_{d < x^{1-\epsilon}} \Big| \# \mathcal{A}_d^- - \frac{\# \mathcal{A}^-}{d} \Big| \le \frac{\# \mathcal{A}}{(\log x)^A},$$

but

$$\#\{p \in \mathcal{A}^-\} = 0.$$

Even best-possible estimates for $\mathcal{A}_d$ cannot produce primes!

We need to incorporate more arithmetic information to distinguish a set $\mathcal{A}$ of interest from $\mathcal{A}^-$.

Type II sums

We can distinguish between these sets if we can estimate Type II sums:

$$\sum_{n\in[x^\theta,x^{\theta+\nu}]} \sum_{\substack{m \\ mn\in\mathcal{A}}} \alpha_n\beta_m$$

for arbitrary 1-bounded coefficients α_m,β_n.

Type II sums

We can distinguish between these sets if we can estimate Type II sums:

$$\sum_{n\in[x^\theta,x^{\theta+\nu}]}\ \sum_{\substack{m\\ mn\in\mathcal{A}}}\alpha_n\beta_m$$

for arbitrary 1-bounded coefficients α_m,β_n. For example

Lemma (Vaughan's Identity)

If $\theta,\nu,\gamma\in[0,1]$ are such that $\gamma+\nu>1$ and

$$\sum_{d<x^\gamma}\Big|\#\mathcal{A}_d-\frac{\#\mathcal{A}}{d}\Big|\le\frac{\#\mathcal{A}}{(\log x)^A},\qquad (\text{'Type I' up to }\gamma)\qquad \text{(I)}$$

$$\sum_{\substack{n,m\\ n\in(2x^\theta,x^{\theta+\nu}]\\ mn\in[x,2x]}}\alpha_n\beta_m\Big(\mathbb{1}_{nm\in\mathcal{A}}-\frac{\#\mathcal{A}}{x}\Big)\le\frac{\#\mathcal{A}}{(\log x)^A},\qquad (\text{'Type II' }[\theta,\theta+\nu])\qquad \text{(II)}$$

Then

$$\#\{p\in\mathcal{A}\}=(1+o(1))\frac{\#\mathcal{A}}{\log x}.$$

Examples from the Iterature

Many results from the literature establish a Type I estimate (value of γ) and a Type II estimate ($[\theta, \theta + \nu]$) and use this to deduce something about primes in our set.

γ (Type I)	$[\theta, \theta + \nu]$ (Type II)	Result
3/4	[1/4, 3/4]	$p = x^2 + y^4$ (Friedlander-Iwaniec)
2/3	[1/3, 2/3]	$p = x^3 + 2y^3$ (Heath-Brown)
19/28	[9/28, 10/28]	$\{\alpha p + \beta\} < p^{-9/28}$ (Jia)
16/25	[0.36, 0.425]	p missing a digit (M.)
5/6	[1/6, 7/24]	$p = x^2 + (y^3 + z^3)^2$ (Merikoski)
1/2	[0, 1/3]	$x^2 \equiv a \pmod p$, $x/p \in \mathcal{I}$ (DFI)

Examples from the Iterature

Many results from the literature establish a Type I estimate (value of γ) and a Type II estimate ($[\theta, \theta+\nu]$) and use this to deduce something about primes in our set.

γ (Type I)	$[\theta, \theta+\nu]$ (Type II)	Result
3/4	[1/4, 3/4]	$p = x^2 + y^4$ (Friedlander-Iwaniec)
2/3	[1/3, 2/3]	$p = x^3 + 2y^3$ (Heath-Brown)
19/28	[9/28, 10/28]	$\{\alpha p + \beta\} < p^{-9/28}$ (Jia)
16/25	[0.36, 0.425]	p missing a digit (M.)
5/6	[1/6, 7/24]	$p = x^2 + (y^3 + z^3)^2$ (Merikoski)
1/2	[0, 1/3]	$x^2 \equiv a \pmod p$, $x/p \in \mathcal{I}$ (DFI)

Question

For which values of γ, θ, ν do we obtain an asymptotic estimate? For which values do we have a non-trivial lower bound? What are limiting examples?

Going further

Summary so far:

- Vaughan's identity gives an asymptotic if $\gamma + \nu > 1$.
- Sometimes we can get an asymptotic even when $\gamma + \nu < 1$ (e.g. $\gamma = 1/2$, $\theta = 0$, $\nu = 1/3$, DFI case)
- Even when we can't get an asymptotic, sometimes we can still get a lower bound of the right order of magnitude (Harman's sieve)
- This whole process is poorly understood; no real limiting examples like Selberg's
- Many different papers establish Type I/II estimates, then there is a tedious computation to check this produces primes.

Going further

Summary so far:

- Vaughan's identity gives an asymptotic if $\gamma + \nu > 1$.
- Sometimes we can get an asymptotic even when $\gamma + \nu < 1$ (e.g. $\gamma = 1/2$, $\theta = 0$, $\nu = 1/3$, DFI case)
- Even when we can't get an asymptotic, sometimes we can still get a lower bound of the right order of magnitude (Harman's sieve)
- This whole process is poorly understood; no real limiting examples like Selberg's
- Many different papers establish Type I/II estimates, then there is a tedious computation to check this produces primes.

Our work aims to introduce a general framework for studying these questions.

New approach (Ford-M.)

- We replace iterative approaches with a direct approach, deploying *all* Type I and Type II information at once. This gives general sieve bounds and a means of constructing limiting examples.
- This (essentially) reduces matters to purely combinatorial problems, which are more tractable (but still difficult!).
- In various cases we can determine precisely the best possible bounds for primes given (γ, θ, ν) and limiting sets.
- We hope that this will lead to a simple practical procedure which will produce close-to-optimal bounds for a wide variety of parameters (γ, θ, ν). In progress!

New approach (Ford-M.)

- We replace iterative approaches with a direct approach, deploying *all* Type I and Type II information at once. This gives general sieve bounds and a means of constructing limiting examples.
- This (essentially) reduces matters to purely combinatorial problems, which are more tractable (but still difficult!).
- In various cases we can determine precisely the best possible bounds for primes given (γ, θ, ν) and limiting sets.
- We hope that this will lead to a simple practical procedure which will produce close-to-optimal bounds for a wide variety of parameters (γ, θ, ν). In progress!

Theorem (Ford-M.)

For any $\theta, \gamma < 1$ there is a $\nu > 0$ and a set $\mathcal{A} \subseteq [x, 2x]$ satisfying (I) and (II) but containing no primes.

Example 1: $\theta = 0$, $\nu = \gamma < 1/2$

Example 1: $\theta = 0$, $\nu = \gamma < 1/2$

By inclusion-exclusion on the smallest prime factor:

$$\#\{p \in \mathcal{A}\} = \#\{a \in \mathcal{A} :\ P^-(a) > x^{1/2}\}$$

$$= \underbrace{\#\{a \in \mathcal{A} :\ P^-(a) > x^{\epsilon}\}}_{\text{Asymptotic with Type I}} - \underbrace{\sum_{x^\epsilon < p \le x^\gamma} \#\Big\{a \in \mathcal{A} :\ P^-(a) = p\Big\}}_{\text{Asymptotic with Type II}}$$

$$- \sum_{x^\gamma \le p \le x^{1/2}} \#\Big\{a \in \mathcal{A} :\ P^-(a) = p\Big\}.$$

Example 1: $\theta = 0$, $\nu = \gamma < 1/2$

By inclusion-exclusion on the smallest prime factor:

$$\#\{p \in \mathcal{A}\} = \#\{a \in \mathcal{A} : P^-(a) > x^{1/2}\}$$
$$= \underbrace{\#\{a \in \mathcal{A} : P^-(a) > x^{\epsilon}\}}_{\text{Asymptotic with Type I}} - \underbrace{\sum_{x^{\epsilon} < p \le x^{\gamma}} \#\Big\{a \in \mathcal{A} : P^-(a) = p\Big\}}_{\text{Asymptotic with Type II}}$$
$$- \sum_{x^{\gamma} \le p \le x^{1/2}} \#\Big\{a \in \mathcal{A} : P^-(a) = p\Big\}.$$

Therefore

$$\#\{p \in \mathcal{A}\} + \#\{p_1 p_2 \in \mathcal{A} : x^{\gamma} \le p_1 \le p_2\} = (\text{Expected asymptotic}).$$

Example 1: $\theta = 0$, $\nu = \gamma < 1/2$

Therefore

$$\#\{p \in \mathcal{A}\} + \#\{p_1 p_2 \in \mathcal{A} : \ x^\gamma \le p_1 \le p_2\} = (\text{Expected asymptotic}).$$

Example 1: $\theta = 0$, $\nu = \gamma < 1/2$

Therefore

$$\#\{p \in \mathcal{A}\} + \#\{p_1 p_2 \in \mathcal{A} : \ x^\gamma \le p_1 \le p_2\} = (\text{Expected asymptotic}).$$

To contain fewer primes than expected, we must have more products of two primes

Example 1: $\theta = 0$, $\nu = \gamma < 1/2$

Therefore

$$\#\{p \in \mathcal{A}\} + \#\{p_1 p_2 \in \mathcal{A} : \; x^\gamma \leq p_1 \leq p_2\} = (\text{Expected asymptotic}).$$

To contain fewer primes than expected, we must have more products of two primes

Construct $\mathcal{A}$ randomly by including $a \in [x, 2x]$ with probability

$$\mathbb{P}(a \in \mathcal{A}) = \begin{cases} 0, & a \text{ prime}, \\ K, & a = p_1 p_2 \text{ with } x^\gamma < p_1 \leq p_2, \\ \alpha, & \text{otherwise.} \end{cases}$$

Choose K such that $\mathbb{E}(\#\mathcal{A}) = \alpha$ (possible if α is sufficiently small in terms of $1/2 - \gamma$).

Then (with high probability) $\mathcal{A}$ **satifies Type I and Type II estimates, but doesn't contain primes.**

Example 2: $\theta = 1/4 + \delta$, $\nu = 1 - 3\theta$, $\gamma = 1 - \theta$

Harman studied the 1-parameter family $(\gamma, \theta, \nu) = (1 - \theta, \theta, 1 - 3\theta)$.

Example 2: $\theta = 1/4 + \delta$, $\nu = 1 - 3\theta$, $\gamma = 1 - \theta$

Harman studied the 1-parameter family $(\gamma, \theta, \nu) = (1 - \theta, \theta, 1 - 3\theta)$. He shows that when $\theta \approx 1/4 + \delta$ (by inclusion-exclusion as before)

$$\#\{p \in \mathcal{A}\} - \#\left\{pq_1q_2 \in \mathcal{A} : \substack{p \le x^{1/2} \\ q_1 \le q_2 \le x^\theta}\right\} - \#\left\{pq_1q_2 \in \mathcal{A} : \substack{q_2 q_1^2 \ge x^{1-\theta} \\ q_1 \le q_2 \le x^\theta}\right\}$$
$$= (\text{Expected asymptotic})$$

Using positivity, this gives a non-trivial lower bound for $\#\{p \in \mathcal{A}\}$. This lower bound is sharp iff $\mathcal{A}$ has no such products of 3 primes.

Example 2: $\theta = 1/4 + \delta$, $\nu = 1 - 3\theta$, $\gamma = 1 - \theta$

Harman studied the 1-parameter family $(\gamma, \theta, \nu) = (1 - \theta, \theta, 1 - 3\theta)$. He shows that when $\theta \approx 1/4 + \delta$ (by inclusion-exclusion as before)

$$\#\{p \in \mathcal{A}\} - \#\Big\{pq_1q_2 \in \mathcal{A} : \substack{p \le x^{1/2} \\ q_1 \le q_2 \le x^\theta}\Big\} - \#\Big\{pq_1q_2 \in \mathcal{A} : \substack{q_2 q_1^2 \ge x^{1-\theta} \\ q_1 \le q_2 \le x^\theta}\Big\}$$
$$= \text{(Expected asymptotic)}$$

Using positivity, this gives a non-trivial lower bound for $\#\{p \in \mathcal{A}\}$. This lower bound is sharp iff $\mathcal{A}$ has no such products of 3 primes.

Fact (Harman bound not optimal)

For every set $\mathcal{A}$ satisfying (I) and (II)

$$\#\Big\{pq_1q_2 \in \mathcal{A} : \substack{q_2 q_1^2 > x^{1-\theta} \\ q_1 \le q_2 \le x^\theta}\Big\} \gg \frac{\#\mathcal{A}}{\log x}$$

Example 2: $\theta = 1/4 + \delta$, $\nu = 1 - 3\theta$, $\gamma = 1 - \theta$

This follows from the following **symmetry relation**:

$$\#\{p_1 p_2 \in \mathcal{A}\} + \#\Big\{p_1 q_1 q_2 \in \mathcal{A} : \substack{p_1 \le x^{1/2} \\ q_1 \le q_2 \le x^{\theta}}\Big\} = \text{(Expected asymptotic)}$$

$$\#\{p_1 p_2 \in \mathcal{A}\} + \#\Big\{p_2 q_1 q_2 \in \mathcal{A} : \substack{p_2 \ge x^{1/2} \\ q_1 \le q_2 \le x^{\theta}}\Big\} = \text{(Expected asymptotic)}$$

so more products $p_1 q_1 q_2$ (large prime $< x^{1/2}$) implies more products $p_2 q_1 q_2$ (large prime $> x^{1/2}$).

Example 2: $\theta = 1/4 + \delta$, $\nu = 1 - 3\theta$, $\gamma = 1 - \theta$

This follows from the following **symmetry relation**:

$$\#\{p_1 p_2 \in \mathcal{A}\} + \#\Big\{p_1 q_1 q_2 \in \mathcal{A} : \substack{p_1 \le x^{1/2} \\ q_1 \le q_2 \le x^\theta}\Big\} = \text{(Expected asymptotic)}$$

$$\#\{p_1 p_2 \in \mathcal{A}\} + \#\Big\{p_2 q_1 q_2 \in \mathcal{A} : \substack{p_2 \ge x^{1/2} \\ q_1 \le q_2 \le x^\theta}\Big\} = \text{(Expected asymptotic)}$$

so more products $p_1 q_1 q_2$ (large prime $< x^{1/2}$) implies more products $p_2 q_1 q_2$ (large prime $> x^{1/2}$).

This leads to

$$\#\{p \in \mathcal{A}\} - 2\#\Big\{p_1 q_1 q_2 \in \mathcal{A} : \substack{p_1 \le x^{1/2} \\ q_1 \le q_2 \le x^\theta}\Big\} = \text{(Expected asymptotic)}$$

and an improved lower bound.

Example 2: $\theta = 1/4 + \delta$, $\nu = 1 - 3\theta$, $\gamma = 1 - \theta$

This leads to

$$\#\{p \in \mathcal{A}\} - 2\#\left\{p_1 q_1 q_2 \in \mathcal{A} : \substack{p_1 \le x^{1/2} \\ q_1 \le q_2 \le x^{\theta}}\right\} = (\text{Expected asymptotic})$$

and an improved lower bound.

Example 2: $\theta = 1/4 + \delta$, $\nu = 1 - 3\theta$, $\gamma = 1 - \theta$

This leads to

$$\#\{p \in \mathcal{A}\} - 2\#\Big\{p_1 q_1 q_2 \in \mathcal{A} : \substack{p_1 \le x^{1/2} \\ q_1 \le q_2 \le x^{\theta}}\Big\} = (\text{Expected asymptotic})$$

and an improved lower bound. A consequence of our work is

Theorem (Ford-M.)

This improved lower bound is optimal; for any set $\mathcal{A}$ satifying (I) and (II) we have

$$\#\{p \in \mathcal{A}\} \ge \left(1 - 2\int_{1-2\theta}^{1/2} \frac{\log\left(\frac{\theta}{1-\theta-\alpha}\right)}{\alpha(1-\theta)} d\alpha + o(1)\right)\frac{\#\mathcal{A}}{\log x},$$

and there are explicit examples of sets $\mathcal{A}$ where this is achieved.

The examples construct sets satisfying (I) and (II) but with no products $p_1 q_1 q_2$ (and many products $q_1 q_2 q_3 q_4$ with $q_i \approx x^{1/4}$). This ultimately reduces to solving a Volterra integral equation.

Combinatoial decompositions

Lemma (Linnik's identity)

$$t(n) := \frac{\Lambda(n)}{\log n} = \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sum_{\substack{n=d_1\cdots d_j \\ 2\le d_i\,\forall i}} 1$$

Combinatoial decompositions

Lemma (Linnik's identity)

$$t(n) := \frac{\Lambda(n)}{\log n} = \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sum_{\substack{n=d_1\cdots d_j \\ 2\le d_i\ \forall i}} 1$$

Let $t_y(n)$ be the truncation

$$t_y(n) := \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sum_{\substack{n=d_1\cdots d_j \\ 2\le d_i\le y\ \forall i}} 1.$$

- $t_y(n) = 0$ if n has a prime factor bigger than y.
- $t(n)$ and $t_y(n)$ differ by 'long' integer variables.

Reduce to the region $\mathcal{R}$

Let $w_n := \mathbb{1}_{n \in \mathcal{A}} - \frac{\#\mathcal{A}}{x}, \quad y := x^{1-\gamma}$

Reduce to the region $\mathcal{R}$

Let $w_n := \mathbb{1}_{n\in\mathcal{A}} - \frac{\#\mathcal{A}}{x}$, $\quad y := x^{1-\gamma}$

$$\sum_p w_p \approx \sum_n w_n t(n) \qquad \text{(Linnik identity, ignore prime powers)}$$

Reduce to the region $\mathcal{R}$

Let $w_n := \mathbb{1}_{n\in\mathcal{A}} - \frac{\#\mathcal{A}}{x}$, $\quad y := x^{1-\gamma}$

$$\sum_p w_p \approx \sum_n w_n t(n) \qquad \text{(Linnik identity, ignore prime powers)}$$

$$\approx \sum_n w_n t_y(n) \qquad \text{(long integer variables negligible by Type I)}$$

Reduce to the region $\mathcal{R}$

Let $w_n := \mathbb{1}_{n\in\mathcal{A}} - \frac{\#\mathcal{A}}{x}, \quad y := x^{1-\gamma}$

$$\sum_p w_p \approx \sum_n w_n t(n) \qquad \text{(Linnik identity, ignore prime powers)}$$

$$\approx \sum_n w_n t_y(n) \qquad \text{(long integer variables negligible by Type I)}$$

$$\approx \sum_{n\in U} w_n t_y(n) \qquad \text{(numbers with a divisor in } [x^\theta, x^{\theta+\nu}] \text{ negligble by Type II)}$$

where

$$U = \Big\{ n \in [x, 2x] : \underbrace{x^{1-\gamma} - smooth}_{Type\ I}, \underbrace{\text{no divisor in } [x^\theta, x^{\theta+\nu}]}_{Type\ II} \Big\}.$$

Reduce to the region $\mathcal{R}$

Let $w_n := \mathbb{1}_{n\in\mathcal{A}} - \frac{\#\mathcal{A}}{x}$, $\quad y := x^{1-\gamma}$

$$\sum_p w_p \approx \sum_n w_n t(n) \qquad \text{(Linnik identity, ignore prime powers)}$$

$$\approx \sum_n w_n t_y(n) \qquad \text{(long integer variables negligible by Type I)}$$

$$\approx \sum_{n\in U} w_n t_y(n) \qquad \text{(numbers with a divisor in } [x^\theta, x^{\theta+\nu}] \text{ negligble by Type II)}$$

where

$$U = \Big\{ n \in [x, 2x] : \underbrace{x^{1-\gamma} - \mathit{smooth}}_{\textit{Type I}}, \underbrace{\text{no divisor in } [x^\theta, x^{\theta+\nu}]}_{\textit{Type II}} \Big\}.$$

If $n = p_1 \cdots p_r \in U$ then $\mathbf{v}(n) = \left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_r}{\log n} \right) \in \mathcal{R}$, a union of polytopes.

Consequneces

$$\#\{p \in \mathcal{A}\} \approx \frac{\#\mathcal{A}}{\log x} + \sum_{n \in U} w_n t_y(n)$$

Consequneces

$$\#\{p \in \mathcal{A}\} \approx \frac{\#\mathcal{A}}{\log x} + \sum_{n \in U} w_n t_y(n)$$

- If U (or $\mathcal{R}$) is empty, we get an asymptotic.
- We can reverse this process; we choose values for w_n when $n \in U$, and use similar arguments to extend w_n uniquely to all n with (I) and (II) holding.
- This reversal process gives a way of constructing example with unusually many or unusually few primes.

Consequneces

$$\#\{p \in \mathcal{A}\} \approx \frac{\#\mathcal{A}}{\log x} + \sum_{n \in U} w_n t_y(n)$$

- If U (or $\mathcal{R}$) is empty, we get an asymptotic.
- We can reverse this process; we choose values for w_n when $n \in U$, and use similar arguments to extend w_n uniquely to all n with (I) and (II) holding.
- This reversal process gives a way of constructing example with unusually many or unusually few primes.

Theorem (Ford-M.)

*Let $\mathcal{A}$ satisfy (I) and (II) for some $\gamma, \theta, \nu \in [0,1]$. Then we have an asymptotic estimate for $\#\{p \in \mathcal{A}\}$ **if and only if** the following hold:*

- *For all integers $n > \lfloor 1/(1-\gamma) \rfloor$, $\exists a \in \mathbb{N}$ with $\frac{a}{n} \in [\theta, \theta+\nu]$.*
- *There is a positive integer h with $h(1-\gamma) \in [\theta, \theta+\nu] \cup [1-\theta-\nu, 1-\theta]$.*

Sieve bounds

We complement our constructions with an improved sieve setup.

Sieve bounds

We complement our constructions with an improved sieve setup.

The functions $t_y(n)$ have complicated signs, so instead we introduce a sieve $H(n) = \sum_{d|n} \lambda_d$ but instead of requiring $H^-(n) \leq \mathbb{1}(n \text{ prime})$ we **only require this for integers n which can be non-trivially decomposed into terms from U**; call this V.

Sieve bounds

We complement our constructions with an improved sieve setup.

The functions $t_y(n)$ have complicated signs, so instead we introduce a sieve $H(n) = \sum_{d|n} \lambda_d$ but instead of requiring $H^-(n) \le \mathbb{1}(n \text{ prime})$ we **only require this for integers** n **which can be non-trivially decomposed into terms from** U; call this V.

$$\#\{p \in \mathcal{A}\} \ge \sum_{n \in \mathcal{A} \cap V} H^-(n) = \frac{\#\mathcal{A}}{x} \sum_{n \in V} H^-(n) + \sum_{n \in V} w_n H^-(n)$$

$$= \frac{\#\mathcal{A}}{x} \sum_{n \in V} H^-(n) + \underbrace{\sum_{n} w_n H^-(n)}_{\approx 0 \text{ by Type I}} - \underbrace{\sum_{n \notin V} w_n H^-(n)}_{\approx 0 \text{ by decomposition}}$$

Sieve bounds

We complement our constructions with an improved sieve setup.

The functions $t_y(n)$ have complicated signs, so instead we introduce a sieve $H(n) = \sum_{d|n} \lambda_d$ but instead of requiring $H^-(n) \le \mathbb{1}(n \text{ prime})$ we **only require this for integers n which can be non-trivially decomposed into terms from** U; call this V.

$$\#\{p \in \mathcal{A}\} \ge \sum_{n \in \mathcal{A} \cap V} H^-(n) = \frac{\#\mathcal{A}}{x} \sum_{n \in V} H^-(n) + \sum_{n \in V} w_n H^-(n)$$
$$= \frac{\#\mathcal{A}}{x} \sum_{n \in V} H^-(n) + \underbrace{\sum_{n} w_n H^-(n)}_{\approx 0 \text{ by Type I}} - \underbrace{\sum_{n \notin V} w_n H^-(n)}_{\approx 0 \text{ by decomposition}}$$

This gives a sieve lower bound which can take into account all available information.

In many situations we can reduce further and fairly simple (piecewise constant) choices of λ_d give optimal bounds.

Questions

Thanks for listening!