Primitive and Geometric Progression Free Sets Without Large Gaps

Nathan McNew
Towson University

CNTA XV
Laval University
July 11th, 2018

Gaps between prime numbers

Problem (Cryptography)

Need a random 100 digit prime number.

Gaps between prime numbers

Problem (Cryptography)

Need a random 100 digit prime number.

Solution: Pick a random 100 digit integer, check if it is prime. If not, add 1 and repeat until you find one.

Gaps between prime numbers

Problem (Cryptography)

Need a random 100 digit prime number.

Solution: Pick a random 100 digit integer, check if it is prime. If not, add 1 and repeat until you find one.

Worst case scenario: How far do we have to go to find one?

Gaps between prime numbers

Problem (Cryptography)

Need a random 100 digit prime number.

Solution: Pick a random 100 digit integer, check if it is prime. If not, add 1 and repeat until you find one.

Worst case scenario: How far do we have to go to find one?

In practice: At most 7384 (largest known to 10^{100})

Gaps between prime numbers

Problem (Cryptography)

Need a random 100 digit prime number.

Solution: Pick a random 100 digit integer, check if it is prime. If not, add 1 and repeat until you find one.

Worst case scenario: How far do we have to go to find one?

In practice: At most 7384 (largest known to 10^{100}) 116 on average.

Gaps between prime numbers

Problem (Cryptography)

Need a random 100 digit prime number.

Solution: Pick a random 100 digit integer, check if it is prime. If not, add 1 and repeat until you find one.

Worst case scenario: How far do we have to go to find one?

In practice: At most 7384 (largest known to 10^{100}) 116 on average.
In theory:

Gaps between prime numbers

Problem (Cryptography)

Need a random 100 digit prime number.

Solution: Pick a random 100 digit integer, check if it is prime. If not, add 1 and repeat until you find one.

Worst case scenario: How far do we have to go to find one?

In practice: At most 7384 (largest known to 10^{100}) 116 on average.
In theory: At most 10^{53}. (Best we can prove, assuming RH).

Gaps between prime numbers

How far apart can prime numbers be?

Gaps between prime numbers

How far apart can prime numbers be?

Let p_n be the nth prime number.

Gaps between prime numbers

How far apart can prime numbers be?

Let p_n be the nth prime number. $p_n - p_{n-1}$ is the nth prime gap.

Gaps between prime numbers

How far apart can prime numbers be?

Let p_n be the nth prime number. $p_n - p_{n-1}$ is the nth prime gap.

Theorem (Baker, Harman, Pintz)

$$p_n - p_{n-1} \ll p_n^{0.525}$$

Gaps between prime numbers

How far apart can prime numbers be?

Let p_n be the nth prime number. $p_n - p_{n-1}$ is the nth prime gap.

Theorem (Baker, Harman, Pintz)

$$p_n - p_{n-1} \ll p_n^{0.525}$$

Theorem (Cramér)

Assuming the Riemann Hypothesis

$$p_n - p_{n-1} \ll \sqrt{p_n} \log p_n$$

Gaps between prime numbers

How far apart can prime numbers be?

Let p_n be the nth prime number. $p_n - p_{n-1}$ is the nth prime gap.

Theorem (Baker, Harman, Pintz)

$$p_n - p_{n-1} \ll p_n^{0.525}$$

Theorem (Cramér)

Assuming the Riemann Hypothesis

$$p_n - p_{n-1} \ll \sqrt{p_n} \log p_n$$

Conjecture (Cramér)

$$p_n - p_{n-1} \ll \log^2 p_n$$

Gaps between prime numbers

How far apart can prime numbers be?

Gaps between prime numbers

How far apart can prime numbers be?

"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Gaps between prime numbers

How far apart can prime numbers be?

"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Recent Developments!

Gaps between prime numbers

How far apart can prime numbers be?
"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Recent Developments!
$p_n - p_{n-1} \leq 70,000,000$ infinitely often (Zhang 2013)

Gaps between prime numbers

How far apart can prime numbers be?

"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Recent Developments!

$p_n - p_{n-1} \le 70,000,000$ infinitely often (Zhang 2013)

$p_n - p_{n-1} \le 246$ infinitely often (Maynard, Tao, Polymath 2015)

Large gaps:

Gaps between prime numbers

How far apart can prime numbers be?

"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Recent Developments!

$p_n - p_{n-1} \leq 70,000,000$ infinitely often (Zhang 2013)

$p_n - p_{n-1} \leq 246$ infinitely often (Maynard, Tao, Polymath 2015)

Large gaps:

$p_n - p_{n-1} > c\frac{\log p_n \log\log p_n \log\log\log\log p_n}{(\log\log\log p_n)^2}$ Infinitely often (Rankin, 1931)

Gaps between prime numbers

How far apart can prime numbers be?

"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Recent Developments!

$p_n - p_{n-1} \leq 70,000,000$ infinitely often (Zhang 2013)

$p_n - p_{n-1} \leq 246$ infinitely often (Maynard, Tao, Polymath 2015)

Large gaps:

$p_n - p_{n-1} > c\frac{\log p_n \log\log p_n \log\log\log\log p_n}{(\log\log\log p_n)^2}$ Infinitely often (Rankin, 1931)

$c = e^{\gamma}$ (Rankin, 1968)

Gaps between prime numbers

How far apart can prime numbers be?

"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Recent Developments!

$p_n - p_{n-1} \le 70,000,000$ infinitely often (Zhang 2013)

$p_n - p_{n-1} \le 246$ infinitely often (Maynard, Tao, Polymath 2015)

Large gaps:

$p_n - p_{n-1} > c \frac{\log p_n \log\log p_n \log\log\log\log p_n}{(\log\log\log p_n)^2}$ Infinitely often (Rankin, 1931)

$c = e^{\gamma}$ (Rankin, 1968) $c = 2.3$ (Maier, Pomerance, 1989)

Gaps between prime numbers

How far apart can prime numbers be?

"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Recent Developments!

$p_n - p_{n-1} \leq 70,000,000$ infinitely often (Zhang 2013)

$p_n - p_{n-1} \leq 246$ infinitely often (Maynard, Tao, Polymath 2015)

Large gaps:

$p_n - p_{n-1} > c \frac{\log p_n \log\log p_n \log\log\log\log p_n}{(\log\log\log p_n)^2}$ Infinitely often (Rankin, 1931)

$c = e^{\gamma}$ (Rankin, 1968) $c = 2.3$ (Maier, Pomerance, 1989) $c = 2e^{\gamma}$ (Pintz 1999)

Gaps between prime numbers

How far apart can prime numbers be?

"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Recent Developments!

$p_n - p_{n-1} \leq 70,000,000$ infinitely often (Zhang 2013)

$p_n - p_{n-1} \leq 246$ infinitely often (Maynard, Tao, Polymath 2015)

Large gaps:

$p_n - p_{n-1} > c\frac{\log p_n \log\log p_n \log\log\log\log p_n}{(\log\log\log p_n)^2}$ Infinitely often (Rankin, 1931)

$c = e^\gamma$ (Rankin, 1968) $c = 2.3$ (Maier, Pomerance, 1989) $c = 2e^\gamma$ (Pintz 1999)

Erdős \$10,000 prize: Show c can be arbitrarily large.

Gaps between prime numbers

How far apart can prime numbers be?
"On average" $p_n - p_{n-1}$ is $\log p_n$ (Prime number theorem)

Recent Developments!
$p_n - p_{n-1} \leq 70,000,000$ infinitely often (Zhang 2013)
$p_n - p_{n-1} \leq 246$ infinitely often (Maynard, Tao, Polymath 2015)

Large gaps:
$p_n - p_{n-1} > c\frac{\log p_n \log\log p_n \log\log\log\log p_n}{(\log\log\log p_n)^2}$ Infinitely often (Rankin, 1931)
$c = e^\gamma$ (Rankin, 1968) $c = 2.3$ (Maier, Pomerance, 1989) $c = 2e^\gamma$ (Pintz 1999)
Erdős \$10,000 prize: Show c can be arbitrarily large.

$p_n - p_{n-1} \gg \frac{\log p_n \log\log p_n \log\log\log\log p_n}{\log\log\log p_n}$ infinitely often
(Ford, Green, Konyagin, Maynard, Tao, 2016)

Gaps between prime numbers

Summary: The largest gap between primes up to x has size between

$$\frac{\log x \log\log x \log\log\log\log x}{\log\log\log x} \quad \text{and} \quad x^{0.525}$$

but we think the truth is $\log^2 x$.

Primitive Sets

Definition

A set $S \subset \mathbb{N}$ is **primitive** if no element of the set divides another: if $m, n \in S$ are distinct then $m \nmid n$.

Primitive Sets

Definition

A set $S \subset \mathbb{N}$ is **primitive** if no element of the set divides another: if $m, n \in S$ are distinct then $m \nmid n$.

Examples:

- Prime numbers $\mathcal{P} = \{2, 3, 5, \ldots\}$.

Primitive Sets

Definition

A set $S \subset \mathbb{N}$ is **primitive** if no element of the set divides another: if $m, n \in S$ are distinct then $m \nmid n$.

Examples:

- Prime numbers $\mathcal{P} = \{2, 3, 5, \ldots\}$.
- $\{m, m+1, m+2, \ldots 2m-1\}$.

Primitive Sets

Definition

A set $S \subset \mathbb{N}$ is **primitive** if no element of the set divides another: if $m, n \in S$ are distinct then $m \nmid n$.

Examples:

- Prime numbers $\mathcal{P} = \{2, 3, 5, \ldots\}$.
- $\{m, m+1, m+2, \ldots 2m-1\}$.
- $\mathcal{P}_k = \{n \in \mathbb{N} \nmid \Omega(n) = k\}$ for any $k \geq 1$.

Primitive Sets

Definition

A set $S \subset \mathbb{N}$ is **primitive** if no element of the set divides another: if $m, n \in S$ are distinct then $m \nmid n$.

Examples:

- Prime numbers $\mathcal{P} = \{2, 3, 5, \ldots\}$.
- $\{m, m+1, m+2, \ldots 2m-1\}$.
- $\mathcal{P}_k = \{n \in \mathbb{N} \nmid \Omega(n) = k\}$ for any $k \geq 1$. Here $\Omega(n)$ is the total number of prime factors of n, counted with multiplicity.

Primitive Sets

Definition

A set $S \subset \mathbb{N}$ is **primitive** if no element of the set divides another: if $m, n \in S$ are distinct then $m \nmid n$.

Examples:

- Prime numbers $\mathcal{P} = \{2, 3, 5, \ldots\}$.
- $\{m, m+1, m+2, \ldots 2m-1\}$.
- $\mathcal{P}_k = \{n \in \mathbb{N} \nmid \Omega(n) = k\}$ for any $k \geq 1$. Here $\Omega(n)$ is the total number of prime factors of n, counted with multiplicity.
  For example, $\mathcal{P}_2 = \{4, 6, 9, 10, 14, 15, 21, 22, \ldots\}$.

Large primitive sets

Theorem (Erdős)

If S is a primitive set then

$$\sum_{n \in S} \frac{1}{n \log n} < \infty.$$

Large primitive sets

Theorem (Erdős)

If S is a primitive set then

$$\sum_{n \in S} \frac{1}{n \log n} < \infty.$$

Conjecture (Erdős)

If S is a primitive set then

$$\sum_{n \in S} \frac{1}{n \log n} \leq \sum_{p \in \mathcal{P}} \frac{1}{p \log p} = 1.63\ldots.$$

Large primitive sets

Nevertheless, there exist primitive sets much "larger" than the primes.

Large primitive sets

Nevertheless, there exist primitive sets much "larger" than the primes.

Theorem (Ahlswede, Khatchatrian, Sárközy, 1999)

For every $\epsilon > 0$ there exists a primitive set whose counting function is asymptotic to $\frac{x}{(\log\log x)^{1+\epsilon}}$.

Large primitive sets

Nevertheless, there exist primitive sets much "larger" than the primes.

Theorem (Ahlswede, Khatchatrian, Sárközy, 1999)

For every $\epsilon > 0$ there exists a primitive set whose counting function is asymptotic to $\frac{x}{(\log\log x)^{1+\epsilon}}$.

(**Note:** This is (almost) best possible:

Large primitive sets

Nevertheless, there exist primitive sets much "larger" than the primes.

Theorem (Ahlswede, Khatchatrian, Sárközy, 1999)

For every $\epsilon > 0$ there exists a primitive set whose counting function is asymptotic to $\frac{x}{(\log\log x)^{1+\epsilon}}$.

(**Note:** This is (almost) best possible: if the counting function of S is at least $\frac{x}{\log\log x}$, then $\displaystyle\sum_{n\in S}\frac{1}{n\log n}$ diverges.)

Large primitive sets

Nevertheless, there exist primitive sets much "larger" than the primes.

Theorem (Ahlswede, Khatchatrian, Sárközy, 1999)

For every $\epsilon > 0$ there exists a primitive set whose counting function is asymptotic to $\frac{x}{(\log\log x)^{1+\epsilon}}$.

(**Note:** This is (almost) best possible: if the counting function of S is at least $\frac{x}{\log\log x}$, then $\displaystyle\sum_{n\in S}\frac{1}{n\log n}$ diverges.)

The average size of the gaps in such a set is $(\log\log x)^{1+\epsilon}$, much smaller than the primes $(\log x)$.

Large primitive sets

Nevertheless, there exist primitive sets much "larger" than the primes.

Theorem (Ahlswede, Khatchatrian, Sárközy, 1999)

For every $\epsilon > 0$ there exists a primitive set whose counting function is asymptotic to $\frac{x}{(\log\log x)^{1+\epsilon}}$.

(**Note:** This is (almost) best possible: if the counting function of S is at least $\frac{x}{\log\log x}$, then $\displaystyle\sum_{n\in S}\frac{1}{n\log n}$ diverges.)

The average size of the gaps in such a set is $(\log\log x)^{1+\epsilon}$, much smaller than the primes $(\log x)$.

Due to the nature of the construction of these sets, it isn't possible to get good upper bounds for the largest gap in such a set.

Primitive sets without large gaps

Question

Do there exist primitive sets which (provably) have no large gaps?

Primitive sets without large gaps

Question

Do there exist primitive sets which (provably) have no large gaps?

Theorem (He, 2015)

There exists a primitive set of integers $s_1, s_2, s_3 \ldots$ such that

$$s_n - s_{n-1} < 2\sqrt{s_n}.$$

Primitive sets without large gaps

Question

Do there exist primitive sets which (provably) have no large gaps?

Theorem (He, 2015)

There exists a primitive set of integers $s_1, s_2, s_3 \dots$ such that

$$s_n - s_{n-1} < 2\sqrt{s_n}.$$

Theorem (M., 2017)

For every $\epsilon > 0$ there exists a primitive set $a_1, a_2, a_3, \dots,$ such that

Primitive sets without large gaps

Question

Do there exist primitive sets which (provably) have no large gaps?

Theorem (He, 2015)

There exists a primitive set of integers $s_1, s_2, s_3 \ldots$ such that

$$s_n - s_{n-1} < 2\sqrt{s_n}.$$

Theorem (M., 2017)

For every $\epsilon > 0$ there exists a primitive set $a_1, a_2, a_3, \ldots,$ such that

$$a_n - a_{n-1} \leq \exp\left(\sqrt{(2+\epsilon)\log a_n \log\log a_n}\right).$$

Lemmas

Theorem (Linear Sieve of Rosser and Iwaniec)

There exists a positive constant c such that every interval $[x - cy, x]$ contains at least $\frac{y}{\log^2 y}$ integers free of prime factors smaller than $\sqrt{y}$.

Lemmas

Theorem (Linear Sieve of Rosser and Iwaniec)

There exists a positive constant c such that every interval $[x - cy, x]$ contains at least $\frac{y}{\log^2 y}$ integers free of prime factors smaller than $\sqrt{y}$.

Lemma

Any interval $[x - cy, x]$ contains a subset of integers which are pairwise coprime, of size at least $\frac{c'\sqrt{y}}{\log y}$ for some positive constant c'.

Lemmas

Theorem (Linear Sieve of Rosser and Iwaniec)

There exists a positive constant c such that every interval $[x-cy, x]$ contains at least $\frac{y}{\log^2 y}$ integers free of prime factors smaller than $\sqrt{y}$.

Lemma

Any interval $[x-cy, x]$ contains a subset of integers which are pairwise coprime, of size at least $\frac{c'\sqrt{y}}{\log y}$ for some positive constant c'.

Proof:

Lemmas

Theorem (Linear Sieve of Rosser and Iwaniec)

There exists a positive constant c such that every interval $[x - cy, x]$ contains at least $\frac{y}{\log^2 y}$ integers free of prime factors smaller than $\sqrt{y}$.

Lemma

Any interval $[x - cy, x]$ contains a subset of integers which are pairwise coprime, of size at least $\frac{c'\sqrt{y}}{\log y}$ for some positive constant c'.

Proof: By theorem, the interval has a large subset without prime factors greater than $\sqrt{y}$.

Lemmas

Theorem (Linear Sieve of Rosser and Iwaniec)

There exists a positive constant c such that every interval $[x - cy, x]$ contains at least $\frac{y}{\log^2 y}$ integers free of prime factors smaller than $\sqrt{y}$.

Lemma

Any interval $[x - cy, x]$ contains a subset of integers which are pairwise coprime, of size at least $\frac{c'\sqrt{y}}{\log y}$ for some positive constant c'.

Proof: By theorem, the interval has a large subset without prime factors greater than $\sqrt{y}$. Each such integer can have at most $\frac{2\log x}{\log y}$ prime factors.

Lemmas

Theorem (Linear Sieve of Rosser and Iwaniec)

There exists a positive constant c such that every interval $[x - cy, x]$ contains at least $\frac{y}{\log^2 y}$ integers free of prime factors smaller than $\sqrt{y}$.

Lemma

Any interval $[x - cy, x]$ contains a subset of integers which are pairwise coprime, of size at least $\frac{c'\sqrt{y}}{\log y}$ for some positive constant c'.

Proof: By theorem, the interval has a large subset without prime factors greater than $\sqrt{y}$. Each such integer can have at most $\frac{2\log x}{\log y}$ prime factors. Construct a graph on these numbers, connecting those vertices with a prime factor in common.

Lemmas

Theorem (Linear Sieve of Rosser and Iwaniec)

There exists a positive constant c such that every interval $[x - cy, x]$ contains at least $\frac{y}{\log^2 y}$ integers free of prime factors smaller than $\sqrt{y}$.

Lemma

Any interval $[x - cy, x]$ contains a subset of integers which are pairwise coprime, of size at least $\frac{c'\sqrt{y}}{\log y}$ for some positive constant c'.

Proof: By theorem, the interval has a large subset without prime factors greater than $\sqrt{y}$. Each such integer can have at most $\frac{2\log x}{\log y}$ prime factors. Construct a graph on these numbers, connecting those vertices with a prime factor in common. Turan's graph theorem implies this graph has a large independent vertex set.

The probabilistic method

Proof Sketch: (Main theorem)

The probabilistic method

Proof Sketch: (Main theorem) Construct a primitive set probabilistically, show that it almost surely has no large gaps.

The probabilistic method

Proof Sketch: (Main theorem) Construct a primitive set probabilistically, show that it almost surely has no large gaps.

Fix $\epsilon > 0$.

The probabilistic method

Proof Sketch: (Main theorem) Construct a primitive set probabilistically, show that it almost surely has no large gaps.

Fix $\epsilon > 0$. For each prime number p_i choose an integer valued random variable X_i with distribution

$$P(X_i = m) = \frac{C_\epsilon}{m^{1+\epsilon/4}}$$

with $C_\epsilon = \frac{1}{\zeta(1+\epsilon/4)}$ chosen to normalize the distribution.

The probabilistic method

Proof Sketch: (Main theorem) Construct a primitive set probabilistically, show that it almost surely has no large gaps.

Fix $\epsilon > 0$. For each prime number p_i choose an integer valued random variable X_i with distribution

$$P(X_i = m) = \frac{C_\epsilon}{m^{1+\epsilon/4}}$$

with $C_\epsilon = \frac{1}{\zeta(1+\epsilon/4)}$ chosen to normalize the distribution. Let

$$Q = \{n \geq 2 \mid p_i | n \rightarrow \Omega(n) = X_i\}$$

The probabilistic method

Proof Sketch: (Main theorem) Construct a primitive set probabilistically, show that it almost surely has no large gaps.

Fix $\epsilon > 0$. For each prime number p_i choose an integer valued random variable X_i with distribution

$$P(X_i = m) = \frac{C_\epsilon}{m^{1+\epsilon/4}}$$

with $C_\epsilon = \frac{1}{\zeta(1+\epsilon/4)}$ chosen to normalize the distribution. Let

$$Q = \{n \geq 2 \mid p_i | n \to \Omega(n) = X_i\}$$

consist of precisely those integers, n, where the total number of prime factors of n agrees with the variable X_i for **all** its prime divisors, p_i.

The probabilistic method

Proof Sketch: (Main theorem) Construct a primitive set probabilistically, show that it almost surely has no large gaps.

Fix $\epsilon > 0$. For each prime number p_i choose an integer valued random variable X_i with distribution

$$P(X_i = m) = \frac{C_\epsilon}{m^{1+\epsilon/4}}$$

with $C_\epsilon = \frac{1}{\zeta(1+\epsilon/4)}$ chosen to normalize the distribution. Let

$$Q = \{n \geq 2 \mid p_i | n \to \Omega(n) = X_i\}$$

consist of precisely those integers, n, where the total number of prime factors of n agrees with the variable X_i for **all** its prime divisors, p_i.

Note Q is primitive: If $a \in Q$, and $a|b$ with $b > a$, then $\Omega(a) < \Omega(b)$, but any prime dividing a also divides b, and so b cannot be in Q.

The probabilistic method

Let $y = \exp\left(\sqrt{(2+\epsilon)\log x \log\log x}\right)$.

The probabilistic method

Let $y = \exp\left(\sqrt{(2+\epsilon)\log x \log\log x}\right)$.

We expect every interval $I_x = [x - cy, x]$ to contain an element of Q:

The probabilistic method

Let $y = \exp\left(\sqrt{(2+\epsilon)\log x \log\log x}\right)$.

We expect every interval $I_x = [x - cy, x]$ to contain an element of Q:

By lemma I_x has a subset S of size at least $\frac{c'\sqrt{y}}{\log y}$ which is pairwise coprime and the integers in S have at most $\frac{2\log x}{\log y}$ prime factors.

The probabilistic method

Let $y = \exp\left(\sqrt{(2+\epsilon)\log x \log\log x}\right)$.

We expect every interval $I_x = [x - cy, x]$ to contain an element of Q:

By lemma I_x has a subset S of size at least $\frac{c'\sqrt{y}}{\log y}$ which is pairwise coprime and the integers in S have at most $\frac{2\log x}{\log y}$ prime factors.

Suppose $n \in S$, then the probability that $n \in Q$ is

$$P(n \in Q) = \prod_{p_i | n} P(X_i = \Omega(n))$$

The probabilistic method

Let $y = \exp\left(\sqrt{(2+\epsilon)\log x \log\log x}\right)$.

We expect every interval $I_x = [x - cy, x]$ to contain an element of Q:

By lemma I_x has a subset S of size at least $\frac{c'\sqrt{y}}{\log y}$ which is pairwise coprime and the integers in S have at most $\frac{2\log x}{\log y}$ prime factors.

Suppose $n \in S$, then the probability that $n \in Q$ is

$$P(n \in Q) = \prod_{p_i|n} P(X_i = \Omega(n)) = \prod_{p_i|n} \frac{C_\epsilon}{\Omega(n)^{1+\epsilon/4}}$$

The probabilistic method

Let $y = \exp\left(\sqrt{(2+\epsilon)\log x \log\log x}\right)$.

We expect every interval $I_x = [x - cy, x]$ to contain an element of Q:

By lemma I_x has a subset S of size at least $\frac{c'\sqrt{y}}{\log y}$ which is pairwise coprime and the integers in S have at most $\frac{2\log x}{\log y}$ prime factors.

Suppose $n \in S$, then the probability that $n \in Q$ is

$$P(n \in Q) = \prod_{p_i | n} P(X_i = \Omega(n)) = \prod_{p_i | n} \frac{C_\epsilon}{\Omega(n)^{1+\epsilon/4}}$$

$$\geq \left(\frac{C_\epsilon}{\left(\frac{2\log x}{\log y}\right)^{1+\epsilon/4}} \right)^{\frac{2\log x}{\log y}}$$

The probabilistic method

Let $y = \exp\left(\sqrt{(2+\epsilon)\log x \log\log x}\right)$.

We expect every interval $I_x = [x - cy, x]$ to contain an element of Q:

By lemma I_x has a subset S of size at least $\frac{c'\sqrt{y}}{\log y}$ which is pairwise coprime and the integers in S have at most $\frac{2\log x}{\log y}$ prime factors.

Suppose $n \in S$, then the probability that $n \in Q$ is

$$P(n \in Q) = \prod_{p_i \mid n} P(X_i = \Omega(n)) = \prod_{p_i \mid n} \frac{C_\epsilon}{\Omega(n)^{1+\epsilon/4}}$$

$$\geq \left(\frac{C_\epsilon}{\left(\frac{2\log x}{\log y}\right)^{1+\epsilon/4}}\right)^{\frac{2\log x}{\log y}} = \exp\left(-(2 + \tfrac{\epsilon}{2} + o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)$$

as $x \to \infty$.

The probabilistic method

Because the elements of S are pairwise coprime, their probabilities of being included in Q are independent.

The probabilistic method

Because the elements of S are pairwise coprime, their probabilities of being included in Q are independent.
Thus the probability that no element of S is included in Q is

$$\prod_{n \in S} (1 - P(n \in Q))$$

The probabilistic method

Because the elements of S are pairwise coprime, their probabilities of being included in Q are independent.
Thus the probability that no element of S is included in Q is

$$\prod_{n\in S}(1-P(n\in Q))$$
$$\leq \prod_{n\in S}\left(1-\exp\left(-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right)$$

The probabilistic method

Because the elements of S are pairwise coprime, their probabilities of being included in Q are independent.
Thus the probability that no element of S is included in Q is

$$\prod_{n\in S}(1-P(n\in Q))$$

$$\leq \prod_{n\in S}\left(1-\exp\left(-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right)$$

$$\leq \left(1-\exp\left(-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right)^{\frac{c'\sqrt{y}}{\log y}}$$

The probabilistic method

Because the elements of S are pairwise coprime, their probabilities of being included in Q are independent.

Thus the probability that no element of S is included in Q is

$$\begin{aligned}
&\prod_{n\in S}(1-P(n\in Q))\\
&\quad\le \prod_{n\in S}\left(1-\exp\left(-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right)\\
&\quad\le \left(1-\exp\left(-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right)^{\frac{c'\sqrt{y}}{\log y}}\\
&\quad\le \exp\left(-\frac{c'\sqrt{y}}{\log y}\times\exp\left(-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right)
\end{aligned}$$

The probabilistic method

Because the elements of S are pairwise coprime, their probabilities of being included in Q are independent.
Thus the probability that no element of S is included in Q is

$$\begin{aligned}
&\prod_{n\in S}(1-P(n\in Q))\\
&\quad\le \prod_{n\in S}\left(1-\exp\left(-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right)\\
&\quad\le \left(1-\exp\left(-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right)^{\frac{c'\sqrt{y}}{\log y}}\\
&\quad\le \exp\left(-\frac{c'\sqrt{y}}{\log y}\times\exp\left(-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right)\\
&\quad= \exp\left(-\exp\left(\tfrac{1}{2}\log y-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right).
\end{aligned}$$

The probabalistic method

$$\prod_{n\in S}(1-P(n\in Q)) \leq \exp\left(-\exp\left(\tfrac{1}{2}\log y-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right).$$

The probabalistic method

$$\prod_{n\in S}(1 - P(n \in Q))$$

$$\leq \exp\left(-\exp\left(\tfrac{1}{2}\log y - (2 + \tfrac{\epsilon}{2} + o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right).$$

Using that $y = \exp\left(\sqrt{(2+\epsilon)\log x \log\log x}\right)$ The inner exponent is

$$\tfrac{1}{2}\log y - (2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)$$

The probabalistic method

$$\prod_{n\in S}(1-P(n\in Q))$$
$$\leq \exp\left(-\exp\left(\tfrac{1}{2}\log y-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right).$$

Using that $y=\exp\left(\sqrt{(2+\epsilon)\log x\log\log x}\right)$ The inner exponent is

$$\tfrac{1}{2}\log y-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)$$
$$=\tfrac{1}{2}\sqrt{(2+\epsilon)\log x\log\log x}-\frac{(2+\frac{\epsilon}{2}+o(1))\sqrt{\log x}}{\sqrt{(2+\epsilon)\log\log x}}\log\left(\sqrt{\log x}\right)$$

The probabalistic method

$$\prod_{n\in S}(1-P(n\in Q))$$
$$\leq \exp\left(-\exp\left(\tfrac{1}{2}\log y-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)\right)\right).$$

Using that $y=\exp\left(\sqrt{(2+\epsilon)\log x\log\log x}\right)$ The inner exponent is

$$\tfrac{1}{2}\log y-(2+\tfrac{\epsilon}{2}+o(1))\frac{\log x}{\log y}\log\left(\frac{\log x}{\log y}\right)$$
$$=\tfrac{1}{2}\sqrt{(2+\epsilon)\log x\log\log x}-\frac{(2+\frac{\epsilon}{2}+o(1))\sqrt{\log x}}{\sqrt{(2+\epsilon)\log\log x}}\log\left(\sqrt{\log x}\right)$$
$$=\frac{\epsilon}{4\sqrt{2+\epsilon}}\sqrt{\log x\log\log x}=C'_\epsilon\sqrt{\log x\log\log x}$$

The probabilistic method

The probability that no integer in $[x - cy, x]$ is included in Q is at most $\exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right)$.

The probabilistic method

The probability that no integer in $[x - cy, x]$ is included in Q is at most $\exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right)$.

$$\sum_{x=1}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < \infty,$$

The probabilistic method

The probability that no integer in $[x - cy, x]$ is included in Q is at most $\exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right)$.

$$\sum_{x=1}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < \infty,$$

so there exists an N such that

$$\sum_{x=N}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < 1.$$

The probabilistic method

The probability that no integer in $[x - cy, x]$ is included in Q is at most $\exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right)$.

$$\sum_{x=1}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < \infty,$$

so there exists an N such that

$$\sum_{x=N}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < 1.$$

Starting the sequence at N and using linearity of expectation, the expected number of intervals of this form avoiding Q is at most

$$\sum_{x>N} P([x-cy, x] \cap Q = \emptyset)$$

The probabilistic method

The probability that no integer in $[x - cy, x]$ is included in Q is at most $\exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right)$.

$$\sum_{x=1}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < \infty,$$

so there exists an N such that

$$\sum_{x=N}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < 1.$$

Starting the sequence at N and using linearity of expectation, the expected number of intervals of this form avoiding Q is at most

$$\sum_{x>N} P([x-cy, x] \cap Q = \emptyset) \leq \sum_{x>N} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right)$$

The probabilistic method

The probability that no integer in $[x - cy, x]$ is included in Q is at most $\exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right)$.

$$\sum_{x=1}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < \infty,$$

so there exists an N such that

$$\sum_{x=N}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < 1.$$

Starting the sequence at N and using linearity of expectation, the expected number of intervals of this form avoiding Q is at most

$$\sum_{x>N} P([x-cy, x]\cap Q = \emptyset) \leq \sum_{x>N} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < 1.$$

The probabilistic method

The probability that no integer in $[x - cy, x]$ is included in Q is at most $\exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right)$.

$$\sum_{x=1}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < \infty,$$

so there exists an N such that

$$\sum_{x=N}^{\infty} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < 1.$$

Starting the sequence at N and using linearity of expectation, the expected number of intervals of this form avoiding Q is at most

$$\sum_{x>N} P([x-cy, x]\cap Q = \emptyset) \leq \sum_{x>N} \exp\left(-\exp\left(C'_\epsilon\sqrt{\log x \log\log x}\right)\right) < 1.$$

Thus there exists a set, Q satisfying the properties of the theorem.

Summary

Summary

Asymptotically the best possible upper bound for the gaps between consecutive terms in a primitive set lies between

$$\log\log x \quad \text{and} \quad \exp\left(\sqrt{(2+\epsilon)\log x \log\log x}\right).$$

Geometric Progression Free Sets

Geometric Progression Free Sets

Question

Does there exist a set with bounded gaps (syndetic) that does not contain a geometric progression?

Geometric Progression Free Sets

Question

Does there exist a set with bounded gaps (syndetic) that does not contain a geometric progression?

The squarefree integers avoid 3-term-geometric progressions (a, ar, ar^2), but have unbounded gaps.

Geometric Progression Free Sets

Question

Does there exist a set with bounded gaps (syndetic) that does not contain a geometric progression?

The squarefree integers avoid 3-term-geometric progressions (a, ar, ar^2), but have unbounded gaps.

Theorem (Filaseta, Trifonov, 1992)

The squarefree integers $s_1, s_2, \ldots$ satisfy

$$s_n - s_{n-1} \ll s_n^{1/5} \log s_n.$$

Geometric Progression Free Sets

Question

Does there exist a set with bounded gaps (syndetic) that does not contain a geometric progression?

The squarefree integers avoid 3-term-geometric progressions (a, ar, ar^2), but have unbounded gaps.

Theorem (Filaseta, Trifonov, 1992)

The squarefree integers $s_1, s_2, \ldots$ satisfy

$$s_n - s_{n-1} \ll s_n^{1/5} \log s_n.$$

We can find larger sets that avoid progressions, but can't say much about the gaps of these sets.

Geometric Progression Free Sets

Theorem (He, 2015)

For each $\epsilon > 0$ there exists a set $\{t_1, t_2, \ldots\}$ avoiding 6-term geometric progressions (rational ratio) with gaps satisfying

Geometric Progression Free Sets

Theorem (He, 2015)

For each $\epsilon > 0$ there exists a set $\{t_1, t_2, \ldots\}$ avoiding 6-term geometric progressions (rational ratio) with gaps satisfying

$$t_n - t_{n-1} \ll_\epsilon \exp\left(\left(\tfrac{5}{6}\log 2 + \epsilon\right)\frac{\log t_n}{\log\log t_n}\right).$$

Geometric Progression Free Sets

Theorem (He, 2015)

For each $\epsilon > 0$ *there exists a set* $\{t_1, t_2, \ldots\}$ *avoiding 6-term geometric progressions (rational ratio) with gaps satisfying*

$$t_n - t_{n-1} \ll_\epsilon \exp\left(\left(\tfrac{5}{6}\log 2 + \epsilon\right)\frac{\log t_n}{\log\log t_n}\right).$$

There exists a set $\{u_1, u_2, \ldots\}$ *avoiding 3-term geometric progressions with* **integer** *ratio with* $u_n - u_{n-1} \ll_\epsilon u_n^\epsilon$*.*

Geometric Progression Free Sets

Theorem (He, 2015)

For each $\epsilon > 0$ there exists a set $\{t_1, t_2, \ldots\}$ avoiding 6-term geometric progressions (rational ratio) with gaps satisfying

$$t_n - t_{n-1} \ll_\epsilon \exp\left(\left(\frac{5}{6}\log 2 + \epsilon\right)\frac{\log t_n}{\log\log t_n}\right).$$

There exists a set $\{u_1, u_2, \ldots\}$ avoiding 3-term geometric progressions with **integer** *ratio with $u_n - u_{n-1} \ll_\epsilon u_n^\epsilon$.*

Theorem (M. 2017)

For each $\epsilon > 0$ there exists a set $\{v_1, v_2, \ldots\}$ avoiding 3-term geometric progressions (rational ratio) with gaps satisfying

Geometric Progression Free Sets

Theorem (He, 2015)

For each $\epsilon > 0$ there exists a set $\{t_1, t_2, \ldots\}$ avoiding 6-term geometric progressions (rational ratio) with gaps satisfying

$$t_n - t_{n-1} \ll_\epsilon \exp\left(\left(\tfrac{5}{6}\log 2 + \epsilon\right)\frac{\log t_n}{\log\log t_n}\right).$$

There exists a set $\{u_1, u_2, \ldots\}$ avoiding 3-term geometric progressions with **integer** *ratio with $u_n - u_{n-1} \ll_\epsilon u_n^\epsilon$.*

Theorem (M. 2017)

For each $\epsilon > 0$ there exists a set $\{v_1, v_2, \ldots\}$ avoiding 3-term geometric progressions (rational ratio) with gaps satisfying

$$v_n - v_{n-1} \ll_\epsilon \exp\left(2\sqrt{(\log 2 + \epsilon)\log v_n}\right).$$

Geometric Progression Free Sets

Theorem (M. 2017)

For each $\epsilon > 0$ there exists a set $\{v_1, v_2, \ldots\}$ avoiding 3-term geometric progressions (rational ratio) with gaps satisfying

$$v_n - v_{n-1} \ll_\epsilon \exp\left(2\sqrt{(\log 2 + \epsilon)\log v_n}\right).$$

Construction:

Geometric Progression Free Sets

Theorem (M. 2017)

For each $\epsilon > 0$ *there exists a set* $\{v_1, v_2, \ldots\}$ *avoiding 3-term geometric progressions (rational ratio) with gaps satisfying*

$$v_n - v_{n-1} \ll_\epsilon \exp\left(2\sqrt{(\log 2 + \epsilon)\log v_n}\right).$$

Construction: A probabilistic analogue of the squarefree numbers.

Geometric Progression Free Sets

Theorem (M. 2017)

For each $\epsilon > 0$ *there exists a set* $\{v_1, v_2, \ldots\}$ *avoiding 3-term geometric progressions (rational ratio) with gaps satisfying*

$$v_n - v_{n-1} \ll_\epsilon \exp\left(2\sqrt{(\log 2 + \epsilon)\log v_n}\right).$$

Construction: A probabilistic analogue of the squarefree numbers.

For each prime p_i choose an integer valued random variable X_i with

$$P(X_i = m) = \frac{1}{2^m}.$$

Geometric Progression Free Sets

Theorem (M. 2017)

For each $\epsilon > 0$ there exists a set $\{v_1, v_2, \ldots\}$ avoiding 3-term geometric progressions (rational ratio) with gaps satisfying

$$v_n - v_{n-1} \ll_\epsilon \exp\left(2\sqrt{(\log 2 + \epsilon)\log v_n}\right).$$

Construction: A probabilistic analogue of the squarefree numbers. For each prime p_i choose an integer valued random variable X_i with

$$P(X_i = m) = \frac{1}{2^m}.$$

Let $V = \{n \geq 2 : p_i | n \to p_i^{X_i} || n\}$.

Geometric Progression Free Sets

Theorem (M. 2017)

For each $\epsilon > 0$ there exists a set $\{v_1, v_2, \ldots\}$ avoiding 3-term geometric progressions (rational ratio) with gaps satisfying

$$v_n - v_{n-1} \ll_\epsilon \exp\left(2\sqrt{(\log 2 + \epsilon)\log v_n}\right).$$

Construction: A probabilistic analogue of the squarefree numbers. For each prime p_i choose an integer valued random variable X_i with

$$P(X_i = m) = \frac{1}{2^m}.$$

Let $V = \{n \geq 2 : p_i | n \to p_i^{X_i} || n\}$. V contains the integers divisible only by primes p_i appearing to the corresponding power X_i.

Geometric Progression Free Sets

Theorem (M. 2017)

For each $\epsilon > 0$ there exists a set $\{v_1, v_2, \ldots\}$ avoiding 3-term geometric progressions (rational ratio) with gaps satisfying

$$v_n - v_{n-1} \ll_\epsilon \exp\left(2\sqrt{(\log 2 + \epsilon)\log v_n}\right).$$

Construction: A probabilistic analogue of the squarefree numbers. For each prime p_i choose an integer valued random variable X_i with

$$P(X_i = m) = \frac{1}{2^m}.$$

Let $V = \{n \geq 2 : p_i | n \to p_i^{X_i} || n\}$. V contains the integers divisible only by primes p_i appearing to the corresponding power X_i.

Note V avoids 3-term progressions: If (a, ar, ar^2) is a progression, a prime in r appears to different powers in at least two of these terms.

Summary

Summary

Asymptotically the best upper bound for gaps between consecutive terms in a 3-term geometric progression free set is between

Summary

Asymptotically the best upper bound for gaps between consecutive terms in a 3-term geometric progression free set is between

$$3 \quad \text{and} \quad \exp\left(2\sqrt{(\log 2 + \epsilon)\log x}\right).$$

Pairwise coprime sets

Pairwise coprime sets

Note that while the set we constructed is primitive, it isn't a pairwise coprime set.

Pairwise coprime sets

Note that while the set we constructed is primitive, it isn't a pairwise coprime set.

Clearly the primes are pairwise coprime, and any pairwise coprime set is primitive.

Pairwise coprime sets

Note that while the set we constructed is primitive, it isn't a pairwise coprime set.

Clearly the primes are pairwise coprime, and any pairwise coprime set is primitive.

Can we find a pairwise coprime set with gaps that are provably smaller than what is known for the primes?

Thank you!