Gaps between primes

James Maynard

University of Oxford

ICM Rio de Janeiro, August 2018
Let’s try to understand the gaps between primes.
Let’s try to understand the gaps between primes.

**Theorem (Prime number theorem)**

\[
\#\{\text{primes} \leq x\} \approx \frac{x}{\log x}.
\]

This means that for \( p_n \leq x \), the **average** gap \( p_{n+1} - p_n \approx \log x \), so the primes get sparser.
Let’s try to understand the gaps between primes.

Theorem (Prime number theorem)

$$\#\{\text{primes } \leq x\} \approx \frac{x}{\log x}.$$ 

This means that for $p_n \leq x$, the average gap $p_{n+1} - p_n \approx \log x$, so the primes get sparser.

Question

*Are prime gaps always this big?*
Introduction II

Question

How small can prime gaps be?

(2, 3) is the only pair of primes which differ by 1. (One of \(n\) and \(n+1\) is a multiple of 2 for every integer \(n\)). There are lots of pairs of primes which differ by 2: (3, 5), (5, 7), (11, 13), ..., (1031, 1033), ..., (1000037, 1000039), ..., (1000000007, 1000000009), ...

Conjecture (Twin prime conjecture)

There are infinitely many pairs of primes \((p, p')\) which differ by 2.

This is one of the oldest problems in mathematics, and is very much open!

James Maynard

Gaps between primes
Introduction II

Question

How small can prime gaps be?

- (2, 3) is the only pair of primes which differ by 1. (One of $n$ and $n + 1$ is a multiple of 2 for every integer $n$).
- There are lots of pairs of primes which differ by 2: (3, 5), (5, 7), (11, 13), ..., (1031, 1033), ..., (1000037, 1000039), ..., (1000000007, 1000000009), ...
Question

How small can prime gaps be?

- $(2, 3)$ is the only pair of primes which differ by 1. (One of $n$ and $n + 1$ is a multiple of 2 for every integer $n$).
- There are lots of pairs of primes which differ by 2: $(3, 5), (5, 7), (11, 13), \ldots, (1031, 1033), \ldots, (1000037, 1000039), \ldots, (1000000007, 1000000009), \ldots$

Conjecture (Twin prime conjecture)

There are infinitely many pairs of primes $(p, p')$ which differ by 2.

This is one of the oldest problems in mathematics, and is very much open!
What about more general gaps?

There is at most one pair of primes \((p_1, p_2)\) with \(p_1 - p_2 = h\) if \(h\) is odd.

There are lots of pairs which differ by 2, 4, 6, 8, \ldots.

Conjecture (De Polignac's conjecture)
For every even \(h\), there are infinitely many pairs \((p_1, p_2)\) of primes such that \(p_1 - p_2 = h\).

More generally, we can look for triples (or more) of primes. \((2, 3, 5)\), \((2, 3, 7)\), \((2, 5, 7)\), \((3, 5, 7)\) are the only triples contained in an interval of length 5. (At least one of \(n, n+2, n+4\) is a multiple of 3.)

There are lots of triples of primes in an interval of length 6.
Question

What about more general gaps?

- There is at most one pair of primes \((p_1, p_2)\) with \(p_1 - p_2 = h\) if \(h\) is odd.
- There are lots of pairs which differ by 2, 4, 6, 8, . . . .
Question

*What about more general gaps?*

- There is at most one pair of primes \((p_1, p_2)\) with \(p_1 - p_2 = h\) if \(h\) is odd.
- There are lots of pairs which differ by 2, 4, 6, 8, \ldots.

Conjecture (De Polignac's conjecture)

*For every even \(h\), there are infinitely many pairs \((p_1, p_2)\) of primes such that \(p_1 - p_2 = h\).*
Question

What about more general gaps?

- There is at most one pair of primes \((p_1, p_2)\) with \(p_1 - p_2 = h\) if \(h\) is odd.
- There are lots of pairs which differ by 2, 4, 6, 8, \ldots.

Conjecture (De Polignac's conjecture)

For every even \(h\), there are infinitely many pairs \((p_1, p_2)\) of primes such that \(p_1 - p_2 = h\).

More generally, we can look for triples (or more) of primes.

- \((2, 3, 5), (2, 3, 7), (2, 5, 7), (3, 5, 7)\) are the only triples contained in an interval of length 5. (At least one of \(n, n + 2, n + 4\) is a multiple of 3.)
- There are lots of triples of primes in an interval of length 6.
Prime $k$-tuples conjecture

We can look for general patterns $L_1(n), \ldots, L_k(n)$ of linear functions of primes.
Prime $k$-tuples conjecture

We can look for general patterns $L_1(n), \ldots, L_k(n)$ of linear functions of primes.

**Definition (admissibility)**

\{L_1, \ldots, L_k\} is **admissible** if $\prod_i L_i(n)$ has no fixed prime divisor.

**Conjecture (Prime $k$-tuple conjecture)**

Let $\{L_1, \ldots, L_k\}$ be admissible. Then there are infinitely many integers $n$ such that all of $L_1(n), \ldots, L_k(n)$ are primes.
Consequences of $k$-tuples

This has a huge number of other consequences!

- $L_1(n) = n$, $L_2(n) = n + 2$: Twin prime conjecture.
- $L_i(n) = n + h_i$: $m$ primes in intervals of length $\approx m \log m$.
- $L_1(n) = n$, $L_2(n) = 2n - 1$: there are infinitely many Sophie Germain primes. (Weak FLT and Artin’s conjecture on primitive roots.)
- $L_j(n) = n + j \times k!$: $k$-term arithmetic progressions of primes
- $L_1(n) = n$, $L_2(n) = 2N - n$: Goldbach’s conjecture.
- $L_i(n) = n + h_iq$: Residue class containing many small primes.
- ...
Theorem (M., Tao)

Let \( \{L_1, \ldots, L_k\} \) be admissible. Then there are infinitely many integers \( n \) such that \( (1/4 + o(1)) \log k \) of \( L_1(n), \ldots, L_k(n) \) are primes.
Theorem (M., Tao)

Let \{L_1, \ldots, L_k\} be admissible. Then there are infinitely many integers \(n\) such that 
\[
(1/4 + o(1)) \log k \text{ of } L_1(n), \ldots, L_k(n) \text{ are primes.}
\]

Zhang proved that 2 of the \(L_i\) are simultaneously prime for \(k\) large enough by using a different method.

Both this result and Zhang's are based on earlier ideas of Goldston, Pintz and Yıldırım.

Independently discovered by Tao.

The underlying method is flexible and can work with subsets, uniformity etc.
This automatically gives bounded gaps between primes.

**Theorem (Zhang)**

*We have*

\[
\liminf_n (p_{n+1} - p_n) \leq 70\,000\,000.
\]

**Theorem (M.)**

*We have*

1. \[
\liminf_n (p_{n+m} - p_n) \leq m^3 e^{4m+8} \text{ for all } m \in \mathbb{N}.
\]
2. \[
\liminf_n (p_{n+1} - p_n) \leq 600.
\]

**Theorem (Polymath 8b)**

*We have*

\[
\liminf_n (p_{n+1} - p_n) \leq 246.
\]
We can view these results as an application of the probabilistic method.

1. Choose $n \in [X, 2X]$ randomly (according to some probability measure).
2. Calculate the expected number of $L_i(n)$ which are prime.
3. If this expectation is $> m$ for all large $X$, then there are infinitely many $n$ such that at least $m$ of the $L_i(n)$ are prime.
We can view these results as an application of the probabilistic method.

1. Choose \( n \in [X, 2X] \) randomly (according to some probability measure).
2. Calculate the expected number of \( L_i(n) \) which are prime.
3. If this expectation is \( > m \) for all large \( X \), then there are infinitely many \( n \) such that at least \( m \) of the \( L_i(n) \) are prime.

**Question**

*How to choose the probability measure to be concentrated on \( n \) when ‘many’ of the \( L_i(n) \) prime?*
The choice of probability measure is inspired by sieve methods.

One way to view sieve methods is the study of ‘almost-primes’.

\[
\text{Integers} \quad \overset{Easy}{\Rightarrow} \quad \text{‘Almost Primes’ (Weighted Integers)} \quad \overset{Moderate}{\Rightarrow} \quad \text{Primes} \quad \overset{Difficult}{\Rightarrow}
\]
The choice of probability measure is inspired by sieve methods.

One way to view sieve methods is the study of ‘almost-primes’.

The primes have positive density in the almost-primes

We can solve additive problems for almost-primes if we know solutions in arithmetic progressions

We understand integers and primes in arithmetic progressions, so we can calculate the expected number of $L_i$ which are prime.

This expectation depends on the definition of ‘almost-prime’.
Question

How do we choose the weights $w_n$ which define almost-primes?

We choose $w_n$ to mimic ‘Selberg sieve’ weights.

1 Standard choice: Gives expectation $\approx \frac{1}{2}$. Fails to prove bounded gaps.

2 GPY choice: Gives expectation $\approx 1 - \epsilon$. Just fails to prove bounded gaps.

3 Zhang: Proves stronger result about primes in arithmetic progressions, gets expectation $1 + \epsilon$!

4 New choice: Get expectation $(\frac{1}{4} + o(1)) \log k$. Gives $m$ primes if $k > C\epsilon^{-4} m$.
Optimised GPY method

Question

*How do we choose the weights $w_n$ which define almost-primes?*

We choose $w_n$ to mimic ‘Selberg sieve’ weights.

1. **Standard choice:** Gives expectation $\approx 1/2$. Fails to prove bounded gaps.
Optimised GPY method

Question

How do we choose the weights \( w_n \) which define almost-primes?

We choose \( w_n \) to mimic ‘Selberg sieve’ weights.

1. **Standard choice**: Gives expectation \( \approx 1/2 \). Fails to prove bounded gaps.
2. **GPY choice**: Gives expectation \( \approx 1 - \epsilon \). Just fails to prove bounded gaps.

James Maynard
Gaps between primes
Question

How do we choose the weights $w_n$ which define almost-primes?

We choose $w_n$ to mimic ‘Selberg sieve’ weights.

1. **Standard choice**: Gives expectation $\approx 1/2$. Fails to prove bounded gaps.
2. **GPY choice**: Gives expectation $\approx 1 - \epsilon$. Just fails to prove bounded gaps.
3. **Zhang**: Proves stronger result about primes in arithmetic progressions, gets expectation $1 + \epsilon$!
**Question**

*How do we choose the weights $w_n$ which define almost-primes?*

We choose $w_n$ to mimic ‘Selberg sieve’ weights.

1. **Standard choice**: Gives expectation $\approx 1/2$. Fails to prove bounded gaps.
2. **GPY choice**: Gives expectation $\approx 1 - \epsilon$. Just fails to prove bounded gaps.
3. **Zhang**: Proves stronger result about primes in arithmetic progressions, gets expectation $1 + \epsilon!$
4. **New choice**: Get expectation $(1/4 + o(1)) \log k$. Gives $m$ primes if $k > C \epsilon^{(4+\epsilon)m}$. 
Overview

Primes in A.P.s

Bombieri-Vinogradov theorem

Sieve Method

Modified GPY sieve

Optimization problem

Choice of smooth weight

Small gaps between primes

Combinatorial problem

Dense admissible sets

Figure: Outline of steps to prove small gaps between primes
What about large gaps? How big is $G(X) = \sup_{p_n \leq X} (p_{n+1} - p_n)$?

$G(X) \geq (1 + o(1)) \log X$ by prime number theorem.
Question

What about large gaps? How big is \( G(X) = \sup_{p_n \leq X} (p_{n+1} - p_n) \)?

\[ G(X) \geq (1 + o(1)) \log X \] by prime number theorem.

Theorem (Rankin, 1938)

\[
G(X) \gg \frac{(\log X)(\log \log X)(\log \log \log \log X)}{(\log \log \log X)^2}
\]

Question (Erdős, $10,000)

Can we improve on Rankin’s result by an arbitrarily large constant?
Large gaps between primes

**Question**

*What about large gaps? How big is* $G(X) = \sup_{p_n \leq X} (p_{n+1} - p_n)$?

*G(X) ≥ (1 + o(1)) log X* by prime number theorem.

**Theorem (Rankin, 1938)**

$$G(X) \gg \frac{(\log X)(\log \log X)(\log \log \log \log X)}{(\log \log \log X)^2}$$

**Question (Erdős, $10,000)**

*Can we improve on Rankin’s result by an arbitrarily large constant?*

**Theorem (Ford-Green-Konyagin-M.-Tao, 2015)**

$$G(X) \gg \frac{(\log X)(\log \log X)(\log \log \log \log X)}{\log \log \log X}$$
Take the integers between 1 and \( y \):

\[
1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad \cdots \quad \cdots \quad y
\]

Step 1: Cross out all integers which are congruent to \( a \mod 2 \):

Step 2: Cross out all integers which are congruent to \( a \mod 3 \):

Continue until all integers have been crossed out.

Choosing \( U \equiv -a \pmod{p} \) for all \( p \) gives \( U + n \) composite for \( n \leq y \).

Question: What is the minimal number of steps?

If the minimal number of steps is small, then we can find long gaps between primes.
Take the integers between 1 and $y$:

Step 1: Cross out all integers which are congruent to $a_2 \pmod{2}$:

Step 2: Cross out all integers which are congruent to $a_3 \pmod{3}$:

Continue until all integers have been crossed out.

Choosing $U \equiv -a_p \pmod{p}$ gives $U + n$ composite for $n \leq y$.

Question: What is the minimal number of steps?

If the minimal number of steps is small, then we can find long gaps between primes.
Take the integers between 1 and $y$:

Step 1: Cross out all integers which are congruent to $a_2 \pmod{2}$:

Step 2: Cross out all integers which are congruent to $a_3 \pmod{3}$:

Choosing $U \equiv -a \pmod{p}$ for all $p$ gives $U + n$ composite for $n \leq y$.

Question: What is the minimal number of steps? If the minimal number of steps is small, then we can find long gaps between primes.
Take the integers between 1 and \( y \):

1 2 3 4 5 6 7 8 9 10  \hspace{1cm} y

Step 1: Cross out all integers which are congruent to \( a_2 \pmod{2} \):

1 2 3 4 5 6 7 8 9 10  \hspace{1cm} y

Step 2: Cross out all integers which are congruent to \( a_3 \pmod{3} \):

1 2 3 4 5 6 7 8 9 10  \hspace{1cm} y

Continue until all integers have been crossed out.

Choosing \( U \equiv -a_p \pmod{p} \) \( \forall p \) gives \( U + n \) composite for \( n \leq y \).
Take the integers between 1 and $y$:

1 2 3 4 5 6 7 8 9 10

Step 1: Cross out all integers which are congruent to $a_2 \ (\text{mod } 2)$:

1 2 3 4 5 6 7 8 9 10

Step 2: Cross out all integers which are congruent to $a_3 \ (\text{mod } 3)$:

1 2 3 4 5 6 7 8 9 10

Continue until all integers have been crossed out.

Choosing $U \equiv -a_p \ (\text{mod } p)$ \forall p gives $U + n$ composite for $n \leq y$.

**Question**

*What is the minimal number of steps?*

If the minimal number of steps is small, then we can find long gaps between primes.
Idea behind large gaps

Erdős-Rankin method: Every integer \( n \leq y \) satisfies \( n \equiv a_p \pmod{p} \) for some \( p \leq (y \log_2 y)/(\log y \cdot \log_3 y) \) where

\[
a_p = \begin{cases} 
0, & p \text{ ‘medium’}, \\
\text{(chosen greedily)} & p \text{ ‘small’ and ‘large’}. 
\end{cases}
\]
Erdős-Rankin method: Every integer $n \leq y$ satisfies $n \equiv a_p \pmod{p}$ for some $p \leq (y \log_2 y)/(\log y \cdot \log_3 y)$ where
\[
a_p = \begin{cases} 
0, & \text{p ‘medium’}, \\
(\text{chosen greedily}) & \text{p ‘small’ and ‘large’}.
\end{cases}
\]

We improve the final stage for $p$ ‘large’. This requires us to find a residue class $a_p \pmod{p}$ for $p \approx x$ which contains many primes $\leq x(\log x)/\log_2 x$. 
Erdős-Rankin method: Every integer $n \leq y$ satisfies $n \equiv a_p \pmod{p}$ for some $p \leq (y \log_2 y)/(\log y \cdot \log_3 y)$ where

$$a_p = \begin{cases} 0, & \text{p ‘medium’,} \\ \text{(chosen greedily)}, & \text{p ‘small’ and ‘large’}. \end{cases}$$

We improve the final stage for $p$ ‘large’. This requires us to find a residue class $a_p \pmod{p}$ for $p \approx x$ which contains many primes $\leq x(\log x)/\log_2 x$.

But this we solved using weak prime $k$-tuples with $L_i(n) = n + h_ip$!
Thank you for listening.
Recall that the quality of our ‘almost-primes’ depends on results about primes in arithmetic progressions. What if we assume the most optimistic conjectures of this type?
Recall that the quality of our ‘almost-primes’ depends on results about primes in arithmetic progressions. What if we assume the most optimistic conjectures of this type?

**Theorem (Polymath 8b)**

*Assume ‘GEH’. Then we have,*

$$\liminf_{n} (p_{n+1} - p_{n}) \leq 6.$$  

Unfortunately, this is a hard limit of our methods.
This has a second consequence:

Theorem (Polymath 8b)

Assume ‘GEH’. Then at least one of the following is true.

1. There are infinitely many twin primes.
2. Every large even number is within 2 of a number which is the sum of two primes.

Of course we expect both to be true!
Thank you for listening.