

# Gaps between primes: The story so far



1785

The University  
of Georgia

Paul Pollack

University of Georgia  
Number Theory Seminar

September 24, 2014

# PART I: (MOSTLY) PREHISTORY

# PART I: (MOSTLY) PREHISTORY ( $>$ 2 years ago)

## 300 BCE

---

Let  $p_n$  denote the  $n$ th prime number, so  $p_1 = 2$ ,  $p_2 = 3$ ,  $p_3 = 5$ ,  
.....

### Theorem (Euclid)

*There are infinitely many primes. In other words, if  $\pi(x) := \#\{p \leq x : p \text{ prime}\}$ , then  $\pi(x) \rightarrow \infty$  as  $x \rightarrow \infty$ .*

There are at this point seemingly infinitely many proofs of this theorem. Euclid's theorem suggests there might be something to be gained by studying the sequence of gaps

$$d_n := p_{n+1} - p_n.$$



## Twin primes

---

The sequence  $\{d_n\}$  begins

1, 2, 2, 4, 2, 4, 2, 4, 6, 2, 6, 4, 2, 4, 6, 6, 2, 6, 4, 2, 6, 4, 6, 8, ...

(OEIS A001223)

It was noticed many moons ago that  $d_n = 2$  appears to appear infinitely often.

# Twin primes

---

The sequence  $\{d_n\}$  begins

1, 2, 2, 4, 2, 4, 2, 4, 6, 2, 6, 4, 2, 4, 6, 6, 2, 6, 4, 2, 6, 4, 6, 8, ...

(OEIS A001223)

It was noticed many moons ago that  $d_n = 2$  appears to appear infinitely often.

**Recall:** A pair of prime numbers  $\{p, p + 2\}$  is called a *twin prime pair*.



**Twin prime pairs:**  $\{3, 5\}$ ,  $\{5, 7\}$ ,  $\{11, 13\}$ ,  $\{17, 19\}$ ,  $\{29, 31\}$ ,  $\{41, 43\}$ ,  $\{59, 61\}$ ,  $\{71, 73\}$ , ...

**Twin prime conjecture:** There are infinitely many twin prime pairs.

## What is a *large* prime gap?

---

It is clear that the smallest gap size we can hope to achieve infinitely often is  $d_n = 2$ ; i.e.,

$$\liminf_{n \rightarrow \infty} d_n \geq 2.$$

What is the largest gap size we can hope to see infinitely often?

### Theorem (Folk L. Ore)

*For each  $N$ , there is a pair of consecutive primes separated by a distance of at least  $N$ . In other words,  $\limsup_{n \rightarrow \infty} d_n = \infty$ .*

### Proof.

We can explicitly construct a large prime-free zone. Indeed, as long as  $N \geq 2$ , there are no primes among the  $N - 1$  consecutive integers  $N! + 2, \dots, N! + N$ . So if  $p_n$  is the largest prime  $\leq N! + 1$ , then  $p_{n+1} \geq N! + N + 1$ , and  $d_n \geq N$ .

## Large and large. What is large?

---



How large a gap did we just construct? Let  $X = N! + N$ . Then there are no primes in  $(X - (N - 1), X]$ .

How big is  $N$  vs.  $X$ ? Notice that

$$\log X \approx \log N! = \sum_{j=1}^N \log j \approx \int_1^N \log t \, dt \approx N \log N.$$

Consequently,

$$\log \log X \approx \log N \log N \approx \log N,$$

and so

$$N = \frac{N \log N}{\log N} \approx \frac{\log X}{\log \log X}.$$



## What is large? Ctd.

---

Dotting  $i$ 's and crossing  $t$ 's yields a proof of the following proposition:

### Proposition

*Let  $\epsilon > 0$ . Along a sequence of  $X \rightarrow \infty$ , the interval  $(X - (1 - \epsilon) \frac{\log X}{\log \log X}, X]$  is prime-free.*

So we get a gap near  $X$  of size  $> (1 - \epsilon) \frac{\log X}{\log \log X}$ .

## What is large? Ctd.

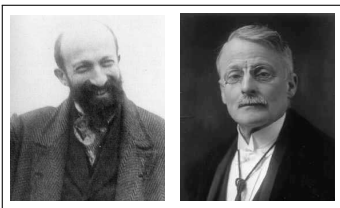
---

Dotting  $i$ 's and crossing  $t$ 's yields a proof of the following proposition:

### Proposition

Let  $\epsilon > 0$ . Along a sequence of  $X \rightarrow \infty$ , the interval  $(X - (1 - \epsilon) \frac{\log X}{\log \log X}, X]$  is prime-free.

So we get a gap near  $X$  of size  $> (1 - \epsilon) \frac{\log X}{\log \log X}$ .



Theorem (Prime number theorem; Hadamard and de la Vallée Poussin, 1899)

As  $X \rightarrow \infty$ , we have  
 $\pi(X) \sim X / \log X$ .

## Calibrating our expectations

---

As a consequence of the prime number theorem, one can show that as  $X \rightarrow \infty$ ,

$$\sum_{p_n \leq X} d_n = p_{\pi(X)+1} - 2 \sim X,$$

and thus

$$\frac{1}{\pi(X)} \sum_{p_n \leq X} d_n \sim \frac{X}{X/\log X} = \log X.$$

**Consequence:** Thus, looking at primes up to  $X$ , the average distance to the next prime is  $\approx \log X$ . So the “large” gap of size  $\approx \log X / \log \log X$  was actually a small gap.

## Calibrating our expectations

---

As a consequence of the prime number theorem, one can show that as  $X \rightarrow \infty$ ,

$$\sum_{p_n \leq X} d_n = p_{\pi(X)+1} - 2 \sim X,$$

and thus

$$\frac{1}{\pi(X)} \sum_{p_n \leq X} d_n \sim \frac{X}{X/\log X} = \log X.$$

**Consequence:** Thus, looking at primes up to  $X$ , the average distance to the next prime is  $\approx \log X$ . So the “large” gap of size  $\approx \log X / \log \log X$  was actually a small gap.

### Remark

Most primes  $p \leq X$  exceed  $X/(\log X)^2$  (say), and so  $\log X \approx \log p$ . So the gap from  $p$  to the next prime is  $\approx \log p$  on average.

## Asking the right question

---

All of this suggests that instead of asking about  $\liminf d_n$  and  $\limsup d_n$ , it makes most sense to ask about the *normalized* prime gaps, defined by

$$\frac{d_n}{\log p_n}.$$

From the average value statement on the preceding slide, one has that

$$\limsup \frac{d_n}{\log p_n} \geq 1 \quad \text{and} \quad \liminf \frac{d_n}{\log p_n} \leq 1.$$

### Questions

Is  $\limsup \frac{d_n}{\log p_n} = \infty$ ? is  $\liminf \frac{d_n}{\log p_n} = 0$ ?

## Large gaps

---

Backlund (1929) got  $\limsup \frac{d_n}{\log p_n} \geq 2$  and Brauer–Zeitiz (1930) got  $\limsup \frac{d_n}{\log p_n} \geq 4$ .

**Theorem (Westzynthius, 1931)**

Yes,  $\limsup \frac{d_n}{\log p_n} = \infty$ . *In fact,*

$$\limsup \frac{d_n}{\log n \cdot \log_3 n / \log_4 n} > 0.$$

**Theorem (Ricci, 1934)**

$$\limsup \frac{d_n}{\log n \cdot \log_3 n} > 0.$$

## Large gaps

---

Theorem (Erdős, 1935)

$$\limsup \frac{d_n}{\log n \cdot \log_2 n / (\log_3 n)^2} > 0.$$

Theorem (Rankin, 1938)

$$\limsup \frac{d_n}{\log n \cdot \log_2 n \cdot \log_4 n / (\log_3 n)^2} > 0.$$

In 1963, Rankin showed that the lim sup was  $\geq e^\gamma$ . In 1990, Maier and Pomerance (GO DAWGS) improved this by a factor of  $1.31256 \dots$ . In 1996, Pintz got up to  $2e^\gamma$ .

All good things . . .



Erdős Problem (\$10000 prize)

Prove that

$$\limsup \frac{d_n}{\log n \cdot \log_2 n \cdot \log_4 n / (\log_3 n)^2} = \infty.$$

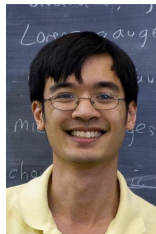
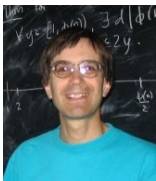
**Breaking news:** Two solutions to Erdős's problem have recently been announced. The first is by Kevin Ford, Ben Green, Sergei Konyagin, and Terence Tao — their preprint appeared on the arXiv on August 20. The second solutions is by James Maynard, and appeared on the arXiv on August 21. The work was done independently and the methods of proof are different.



# Cast of characters

---

Ford, Green, Konyagin, Tao, and Maynard



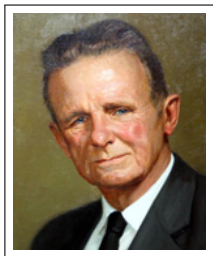
## What's the truth? I mean, really...

---

In 1936, Harald Cramér suggested a statistical model for the primes:

We flip an infinite sequence of biased coins  $C_2, C_3, C_4, \dots$ , where  $C_n$  comes up heads with probability  $1/\log n$ .

Each infinite sequence of coin flips gives us a set of **random primes**, namely those  $n$  for which  $C_n$  comes up heads.



### Theorem

Fix  $\epsilon > 0$ . With probability 1, the random primes satisfy  $\tilde{\pi}(x) = \int_2^x \frac{dt}{\log t} + O(x^{1/2+\epsilon})$ . Here  $\tilde{\pi}(x)$  is the counting function.

This is consistent with our expectations of the actual sequence of primes, assuming the Riemann Hypothesis.

## Theorem (Cramér)

*With probability 1, we have*

$$\limsup_{n \rightarrow \infty} \frac{\tilde{d}_n}{(\log n)^2} = 1,$$

*where  $\tilde{d}_n$  is the analogue of  $d_n$  for our sequence of random primes.*

Cramér conjectured that the same should be true for actual primes.

## Theorem (Cramér)

With probability 1, we have

$$\limsup_{n \rightarrow \infty} \frac{\tilde{d}_n}{(\log n)^2} = 1,$$

where  $\tilde{d}_n$  is the analogue of  $d_n$  for our sequence of random primes.

Cramér conjectured that the same should be true for actual primes.



However, work of Maier on the distribution of primes in short intervals shows that Cramér's model is not always reliable.

Conjecture (Granville,  $\approx$  1995?)

$$\limsup_{n \rightarrow \infty} \frac{\tilde{d}_n}{(\log n)^2} \geq 2e^{-\gamma}.$$

Most people seem to believe that  $d_n = O((\log n)^{2+\epsilon})$ .

Even the Riemann Hypothesis would not be of much help here. On RH, one can show (also noted by Cramér) that

$$d_n = O(p_n^{1/2} \log p_n).$$

This RH-conditional result is not much better than we can already show unconditionally. We know (Baker–Harman–Pintz, 2001) that

$$d_n \ll p_n^{\frac{1}{2} + \frac{1}{40}}.$$

## Small gaps

---

The first person to prove a nontrivial result about small gaps between primes was Viggo Brun, circa 1915.

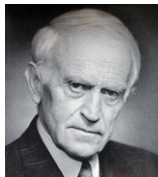
Let  $\pi_2(x) = \#\{p \leq x : p, p + 2 \text{ prime}\}$ .

## Small gaps

---

The first person to prove a nontrivial result about small gaps between primes was Viggo Brun, circa 1915.

Let  $\pi_2(x) = \#\{p \leq x : p, p + 2 \text{ prime}\}$ .



### Theorem (Brun, 1919)

*For all large enough values of  $x$ , we have*

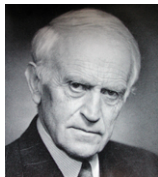
$$\pi_2(x) < 100 \frac{x}{(\log x)^2}.$$

## Small gaps

---

The first person to prove a nontrivial result about small gaps between primes was Viggo Brun, circa 1915.

Let  $\pi_2(x) = \#\{p \leq x : p, p + 2 \text{ prime}\}$ .



### Theorem (Brun, 1919)

*For all large enough values of  $x$ , we have*

$$\pi_2(x) < 100 \frac{x}{(\log x)^2}.$$

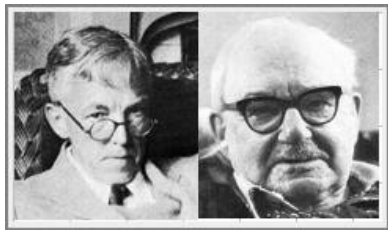
As a corollary, Brun deduced that the sum

$$\left(\frac{1}{3} + \frac{1}{5}\right) + \left(\frac{1}{5} + \frac{1}{7}\right) + \left(\frac{1}{11} + \frac{1}{13}\right) + \cdots < \infty.$$

We do not even know the left-hand sum to a single decimal place!



To appreciate Brun's result, one should recall what we expect for the true behavior of  $\pi_2(x)$ .



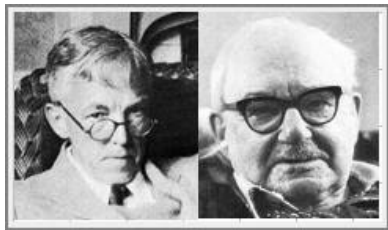
Conjecture (Hardy & Littlewood, 1923)

As  $x \rightarrow \infty$ , we have

$$\pi_2(x) \sim \mathfrak{S} \frac{x}{(\log x)^2}, \quad \text{where}$$

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

To appreciate Brun's result, one should recall what we expect for the true behavior of  $\pi_2(x)$ .



Conjecture (Hardy & Littlewood, 1923)

As  $x \rightarrow \infty$ , we have

$$\pi_2(x) \sim \mathfrak{S} \frac{x}{(\log x)^2}, \quad \text{where}$$

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

Brun's method is important precisely because — in this and many related problems — it gives an *upper bound* that is correct up to a constant factor.

## Small normalized prime gaps

---

For a start, we would like to know that

$$\liminf_{n \rightarrow \infty} \frac{d_n}{\log p_n} < 1.$$

It would seem that Brun is of no help to us, since he gives us an **upper bound** on how often gaps occur, rather than a **lower bound**. This turns out to be mistaken.

### Proposition (Erdős)

Fix  $\epsilon > 0$ . Then

$$\frac{\#\{x/2 < p_n \leq x : (1 - \epsilon) \log p_n < d_n < (1 + \epsilon) \log p_n\}}{\#\{x/2 < p_n \leq x\}} \leq K\epsilon$$

for some absolute constant  $K$ .

## Theorem (Erdős, 1940)

$$\liminf \frac{d_n}{\log p_n} < 1.$$

### Proof.

Let us suppose that from some point on,  $d_n > (1 - \epsilon) \log p_n$ . We will show that if  $\epsilon$  is chosen sufficiently small, we reach a contradiction. Hence,  $\liminf \frac{d_n}{\log p_n} \leq 1 - \epsilon$ .

## Theorem (Erdős, 1940)

$$\liminf \frac{d_n}{\log p_n} < 1.$$

### Proof.

Let us suppose that from some point on,  $d_n > (1 - \epsilon) \log p_n$ . We will show that if  $\epsilon$  is chosen sufficiently small, we reach a contradiction. Hence,  $\liminf \frac{d_n}{\log p_n} \leq 1 - \epsilon$ .

What is the average of  $d_n$  over primes  $p_n \in (X/2, X]$ ? PNT implies it is  $\sim \log X$ .

## Theorem (Erdős, 1940)

$$\liminf \frac{d_n}{\log p_n} < 1.$$

### Proof.

Let us suppose that from some point on,  $d_n > (1 - \epsilon) \log p_n$ . We will show that if  $\epsilon$  is chosen sufficiently small, we reach a contradiction. Hence,  $\liminf \frac{d_n}{\log p_n} \leq 1 - \epsilon$ .

What is the average of  $d_n$  over primes  $p_n \in (X/2, X]$ ? PNT implies it is  $\sim \log X$ .

By assumption, there are only two kinds of primes like this, those with  $d_n / \log p_n \in (1 - \epsilon, 1 + \epsilon)$  and those with  $d_n / \log p_n \in (1 + \epsilon, \infty)$ .

By assumption, there are only two kinds of  $p_n$ , those with  $d_n / \log p_n \in (1 - \epsilon, 1 + \epsilon)$  and those with  $d_n / \log p_n \in (1 + \epsilon, \infty)$ .

By assumption, there are only two kinds of  $p_n$ , those with  $d_n/\log p_n \in (1 - \epsilon, 1 + \epsilon)$  and those with  $d_n/\log p_n \in (1 + \epsilon, \infty)$ .

Let  $\rho$  be the proportion of primes  $p_n \in (X/2, X]$  which belong to the former class. Erdős says  $\rho \leq K\epsilon$ . Since  $\log p_n \gtrsim \log X$  for primes  $p_n \in (X/2, X]$ , these primes make a contribution to the average of

$$\gtrsim \rho(1 - \epsilon) \log X.$$



By assumption, there are only two kinds of  $p_n$ , those with  $d_n/\log p_n \in (1 - \epsilon, 1 + \epsilon)$  and those with  $d_n/\log p_n \in (1 + \epsilon, \infty)$ .

Let  $\rho$  be the proportion of primes  $p_n \in (X/2, X]$  which belong to the former class. Erdős says  $\rho \leq K\epsilon$ . Since  $\log p_n \gtrsim \log X$  for primes  $p_n \in (X/2, X]$ , these primes make a contribution to the average of

$$\gtrsim \rho(1 - \epsilon) \log X.$$

The proportion of primes of the second class is  $1 - \rho$ . the contribution to the average from these primes is

$$\gtrsim (1 - \rho) \cdot (1 + \epsilon) \log X.$$

Now we add.

We find that the average value of  $d_n$ , taken along primes  $p_n \in (X/2, X]$ , is

$$\begin{aligned} &\gtrsim (\rho(1 - \epsilon) + (1 - \rho)(1 + \epsilon)) \log X \\ &= (1 + (1 - 2\rho)\epsilon) \log X \\ &\geq (1 + (1 - 2K\epsilon)\epsilon) \log X. \end{aligned}$$

If we suppose we had initially fixed  $\epsilon < \frac{1}{2K}$ , the constant in front of  $\log X$  is  $> 1$ .

But this contradicts that the average value is  $\sim \log X$ .

As shown by Ricci (1954), Erdős's method leads to

$$\liminf \frac{d_n}{\log p_n} \leq \frac{15}{16}.$$

## Landmark results

---



Theorem (Bombieri and Davenport, 1966)

$$\liminf \frac{d_n}{\log p_n} \leq 0.46650 \dots$$

Theorem (Maier, 1988)

$$\liminf \frac{d_n}{\log p_n} \leq 0.2486 \dots$$



## A new hope

---



(YPG)

Theorem (Goldston, Pintz,  
and Yıldırım, 2005)

$$\liminf \frac{d_n}{\log p_n} = 0.$$

BUT WAIT, THERE'S MORE!

## A new hope

---



(YPG)

Theorem (Goldston, Pintz,  
and Yıldırım, 2005)

$$\liminf \frac{d_n}{\log p_n} = 0.$$

BUT WAIT, THERE'S MORE!

**Any** improvement in the level of distribution of the primes would imply that  $\liminf d_n < \infty$  — i.e., infinitely many pairs of primes that lie in a bounded length interval.

## An aside: Primes in arithmetic progressions

---

For your favorite integer  $q$ , one can ask: How are the primes distributed among the residue classes modulo  $q$ ?

In other words, instead of studying  $\pi(x)$ , one can study

$$\pi(x; q, a) := \#\{p \leq x : p \equiv a \pmod{q}\}.$$

This is only interesting when  $\gcd(a, q) = 1$ , since otherwise there is at most one prime  $p \equiv a \pmod{q}$ .

### Theorem (Dirichlet, 1837)

*As long as  $\gcd(a, q) = 1$ , the function  $\pi(x; q, a) \rightarrow \infty$  as  $x \rightarrow \infty$ .*



As soon as the prime number theorem was proved, it was recognized that its proof could be modified to prove that primes were equidistributed in arithmetic progressions.

Specifically, fix  $a$  and  $q$  with  $\gcd(a, q) = 1$ . Then as  $x \rightarrow \infty$ ,

$$\pi(x; q, a) \sim \frac{\pi(x)}{\phi(q)},$$

as  $x \rightarrow \infty$ . So in the long run, each progression gets its “fair share” of primes.

As soon as the prime number theorem was proved, it was recognized that its proof could be modified to prove that primes were equidistributed in arithmetic progressions.

Specifically, fix  $a$  and  $q$  with  $\gcd(a, q) = 1$ . Then as  $x \rightarrow \infty$ ,

$$\pi(x; q, a) \sim \frac{\pi(x)}{\phi(q)},$$

as  $x \rightarrow \infty$ . So in the long run, each progression gets its “fair share” of primes.

But what if  $q$  isn't fixed? Is  $\pi(x; q, a) \sim \frac{\pi(x)}{\phi(q)}$  if  $q \approx \log x$ ?  $q \approx x^{1/2}$ ?  $q \approx x^{0.99}$ ?



Define  $E(x; q, a) = \pi(x; q, a) - \frac{1}{\phi(q)}\pi(x)$ .

### Theorem (Siegel–Walfisz)

Fix  $A > 0$ . Whenever  $q \leq (\log x)^A$ , we have

$$E(x; q, a) = O_A(x \exp(-c\sqrt{\log x})).$$

Here  $c$  is a certain absolute positive constant.

So the asymptotic  $\pi(x; q, a) \sim \frac{\pi(x)}{\phi(q)}$  holds if  $q \leq (\log x)^A$ .

### Theorem (assuming GRH for Dirichlet $L$ -functions)

For all  $q \leq x$ , we have

$$E(x; q, a) = O(x^{1/2} \log x).$$

This gives the asymptotic for  $q \leq x^{\frac{1}{2}-\epsilon}$ .

## Level of distribution

---

We say the primes have **level of distribution**  $\theta$  if for each  $\epsilon > 0$  and each  $B > 0$ ,

$$\sum_{q \leq x^{\theta - \epsilon}} \max_{\substack{a \bmod q \\ \gcd(a, q) = 1}} |E(x; q, a)| \ll_B \frac{x}{(\log x)^B}.$$

If GRH holds, the primes have level of distribution  $\theta = \frac{1}{2}$ .

## Level of distribution

---

We say the primes have **level of distribution**  $\theta$  if for each  $\epsilon > 0$  and each  $B > 0$ ,

$$\sum_{q \leq x^{\theta - \epsilon}} \max_{\substack{a \bmod q \\ \gcd(a, q) = 1}} |E(x; q, a)| \ll_B \frac{x}{(\log x)^B}.$$

If GRH holds, the primes have level of distribution  $\theta = \frac{1}{2}$ .

**Theorem (Bombieri–A. I. Vinogradov, 1965)**

*The primes have level of distribution  $\theta = \frac{1}{2}$ .*

$\theta = \frac{1}{2}$  was the threshold value for GPY: If we only knew  $\theta < \frac{1}{2}$ , we could not prove  $\liminf \frac{d_n}{\log p_n} = 0$ . If  $\theta > \frac{1}{2}$ , we get  $\liminf d_n < \infty$ .

## Level of distribution

---

We say the primes have **level of distribution**  $\theta$  if for each  $\epsilon > 0$  and each  $B > 0$ ,

$$\sum_{q \leq x^{\theta-\epsilon}} \max_{\substack{a \bmod q \\ \gcd(a,q)=1}} |E(x; q, a)| \ll_B \frac{x}{(\log x)^B}.$$

If GRH holds, the primes have level of distribution  $\theta = \frac{1}{2}$ .

**Theorem (Bombieri–A. I. Vinogradov, 1965)**

*The primes have level of distribution  $\theta = \frac{1}{2}$ .*

$\theta = \frac{1}{2}$  was the threshold value for GPY: If we only knew  $\theta < \frac{1}{2}$ , we could not prove  $\liminf \frac{d_n}{\log p_n} = 0$ . If  $\theta > \frac{1}{2}$ , we get  $\liminf d_n < \infty$ .

**Conjecture (Elliott–Halberstam, 1968)**

*Can take  $\theta = 1$ .*

# PART II: Zhang, Maynard, and Tao (OH MY!)



## Theorem (Y. Zhang, April 2013)

*We can take  $\theta > \frac{1}{2}$  if we restrict to “smooth” moduli  $q$  (numbers  $q$  with only small prime factors).*

## Corollary

$$\liminf_{n \rightarrow \infty} d_n < 70 \cdot 10^6.$$

## Theorem (Maynard)

We have  $\liminf d_n \leq 600$ .

*BUT WAIT, THERE'S MORE!*

For each  $k$ , define the  $k$ th order gap  $d_n^{(k)} := p_{n+k} - p_n$ . We always have

$$\liminf_{n \rightarrow \infty} d_n^{(k)} < \infty.$$



Similar results were discovered concurrently by Terry Tao.

**Polymath 8b:**  $\liminf_{n \rightarrow \infty} d_n \leq 246$ .

## The story behind the story

---

GPY, Zhang, and Maynard do not study  $d_n$  directly. Rather, they study a variant of the twin prime conjecture due to Hardy and Littlewood, called the *k-tuples conjecture*.

### Problem

*Let  $\mathcal{H}$  be a set of  $k$  integers, say  $a_1, \dots, a_k$ . Under what conditions on  $\mathcal{H}$  do we expect that  $n + a_1, \dots, n + a_k$  are simultaneously prime infinitely often?*



## The story behind the story

---

GPY, Zhang, and Maynard do not study  $d_n$  directly. Rather, they study a variant of the twin prime conjecture due to Hardy and Littlewood, called the *k-tuples conjecture*.

### Problem

*Let  $\mathcal{H}$  be a set of  $k$  integers, say  $a_1, \dots, a_k$ . Under what conditions on  $\mathcal{H}$  do we expect that  $n + a_1, \dots, n + a_k$  are simultaneously prime infinitely often?*

We need to rule out examples like  $n, n + 1$ , one of which is always even, or  $n, n + 2, n + 4$ , one of which is always a multiple of 3.

### Definition

We say  $\mathcal{H}$  is admissible if  $\#\mathcal{H} \bmod p < p$  for all primes  $p$ .

## Conjecture (*k*-tuples conjecture)

*If  $\mathcal{H}$  is admissible, then there are infinitely many  $n$  with all  $n + h_i$  simultaneously prime.*

## Conjecture ( $k$ -tuples conjecture)

If  $\mathcal{H}$  is admissible, then there are infinitely many  $n$  with all  $n + h_i$  simultaneously prime.

The theorems of GPY, Maynard, and Zhang about prime gaps are really corollaries of their results towards the  $k$ -tuples conjecture.

## Theorem (Zhang)

There is a constant  $k_0$  so that if  $k \geq k_0$ , and  $\mathcal{H}$  is an admissible  $k$ -tuple, then infinitely often **at least two** of  $n + h_1, \dots, n + h_k$  are prime.

## Theorem (Maynard)

Fix  $m \geq 2$ . There is a constant  $k_0(m)$  so that if  $k \geq k_0(m)$ , and  $\mathcal{H}$  is an admissible  $k$ -tuple, then infinitely often **at least  $m$**  of  $n + h_1, \dots, n + h_k$  are prime.

Maynard showed that one could take  $k_0(2) = 105$ .

Hence, if we fix an admissible 105-tuple, say  $h_1 < \dots < h_k$ , then infinitely often at least two of  $n + h_1, \dots, n + h_k$  are prime, and so infinitely often

$$p_{j+1} - p_j \leq h_k - h_1.$$

Maynard showed that one could take  $k_0(2) = 105$ .

Hence, if we fix an admissible 105-tuple, say  $h_1 < \dots < h_k$ , then infinitely often at least two of  $n + h_1, \dots, n + h_k$  are prime, and so infinitely often

$$p_{j+1} - p_j \leq h_k - h_1.$$

There is an example of such an  $\mathcal{H}$  with  $h_k - h_1 = 600$ ;

$$\mathcal{H} = \{0, 10, 12, 24, 28, \dots, 598, 600\}.$$

Hence,

$$\liminf_{n \rightarrow \infty} d_n \leq 600.$$

**Polymath8b:** Can take  $k_0(2) = 50$ .

Choosing a “narrow” admissible 50-tuple gives  $\liminf d_n \leq 246$ .

Working with  $k_0(m)$  instead of  $k_0(2)$ , one proves in an entirely analogous way that

$$\liminf_{n \rightarrow \infty} (\rho_{n+m} - \rho_n) < \infty.$$

In fact, the LHS is  $\exp(O(m))$ .

Working with  $k_0(m)$  instead of  $k_0(2)$ , one proves in an entirely analogous way that

$$\liminf_{n \rightarrow \infty} (\rho_{n+m} - \rho_n) < \infty.$$

In fact, the LHS is  $\exp(O(m))$ .

As promised, I will say nothing about Maynard's proof except the following: Before Maynard, the level of distribution  $\theta = \frac{1}{2}$  was thought of as a threshold. Maynard showed that this phenomenon is illusory. All of Maynard's arguments go through (up to changes in constants) with **any**  $\theta > 0$ .

# BIG DOINGS WITH SMALL GAPS



## What have you done for me lately?

---

This progress on the  $k$ -tuples conjecture is certainly interesting even if thought of only in terms of its consequences for prime gaps.

## What have you done for me lately?

---

This progress on the  $k$ -tuples conjecture is certainly interesting even if thought of only in terms of its consequences for prime gaps.

BUT WAIT, THERE'S MORE.

## What have you done for me lately?

---

This progress on the  $k$ -tuples conjecture is certainly interesting even if thought of only in terms of its consequences for prime gaps.

BUT WAIT, THERE'S MORE.

The  $k$ -tuples conjecture (and a variant due to Dickson allowing leading coefficients) has traditionally been a “working hypothesis” in a number of investigations in elementary number theory. Many theorems have been proved conditionally on the truth of this conjecture.



The work of Maynard–Tao opens the door towards the **unconditional** resolution of some of these problems.

# Erdős, Turán, and the local behavior of prime gaps

## Question

Are there infinitely many  $n$  with  $d_n < d_{n+1}$ ?

Yes, trivially, because otherwise  $\{d_n\}$  would be bounded.

# Erdős, Turán, and the local behavior of prime gaps

## Question

Are there infinitely many  $n$  with  $d_n < d_{n+1}$ ?

Yes, trivially, because otherwise  $\{d_n\}$  would be bounded.

## Question

Are there infinitely many  $n$  with  $d_n > d_{n+1}$ ?

Yes (E&T, 1948), but no longer so trivial.

## Question (Erdős and Turán, 1948)

Are there infinitely many  $n$  with  $d_n > d_{n+1} > d_{n+2}$ ? What about  $d_n < d_{n+1} < d_{n+2}$ ?

Yes to both, if (a variant of) the prime  $k$ -tuples conjecture is true.

Yes to both, as a consequence of Maynard's theorem.



Theorem (Banks, Freiberg, Turnage-Butterbaugh, 2013)

*For every  $k$ , one can find infinitely many  $n$  with  $d_n < d_{n+1} < \cdots < d_{n+k}$ , and infinitely many  $n$  with  $d_n > d_{n+1} > \cdots > d_{n+k}$ .*

## Lemma (Banks–Freiberg–Turnage-Butterbaugh)

Fix  $m \geq 2$ . There is a constant  $k_0(m)$  so that if  $k \geq k_0(m)$ , and  $\mathcal{H}$  is an admissible  $k$ -tuple, then infinitely often **at least**  $m$  of  $n + h_1, \dots, n + h_k$  are prime, and the primes among  $n + h_1, \dots, n + h_k$  are consecutive in the sequence of all primes.

Apply the Lemma with a given large  $m$ ,  $k = k_0(m)$ , and  $\mathcal{H} = \{2, 2^2, \dots, 2^{k_0(m)}\}$ . (It is easy to check admissibility of  $\mathcal{H}$ .)

## Lemma (Banks–Freiberg–Turnage-Butterbaugh)

Fix  $m \geq 2$ . There is a constant  $k_0(m)$  so that if  $k \geq k_0(m)$ , and  $\mathcal{H}$  is an admissible  $k$ -tuple, then infinitely often **at least**  $m$  of  $n + h_1, \dots, n + h_k$  are prime, and the primes among  $n + h_1, \dots, n + h_k$  are consecutive in the sequence of all primes.

Apply the Lemma with a given large  $m$ ,  $k = k_0(m)$ , and  $\mathcal{H} = \{2, 2^2, \dots, 2^{k_0(m)}\}$ . (It is easy to check admissibility of  $\mathcal{H}$ .)

**Key idea:** The distance from a power of 2 to a smaller power of 2 is  $<$  the distance to any larger power of 2.

We get  $m$  consecutive primes  $q_1 < q_2 < \dots < q_m$  from among  $\{n + 2, n + 2^2, \dots, n + 2^{k_0}\}$ .

Then  $q_{j+1} - q_j \leq q_{j+2} - q_{j+1}$  for all  $j$ .

This gives increasing gaps. For decreasing gaps, replace  $\mathcal{H}$  with



## Shiu strings

---

### Theorem (D. K. L. Shiu, 2000)

*Fix a coprime progression  $a \pmod q$ . For every  $m$ , one can find  $m$  consecutive primes  $p_n \equiv p_{n+1} \equiv \cdots \equiv p_{n+m-1} \equiv a \pmod q$ .*

### Theorem (Banks, Freiberg, Turnage-Butterbaugh)

*Shiu's theorem is still true fourteen years later. Moreover, it remains true even if one restricts  $p_n, \dots, p_{n+m-1}$  to lie in a bounded length interval. ("Bounded" means bounded in terms of  $q$  and  $m$ .)*

## Lemma (Banks, Freiberg, Turnage-Butterbaugh)

Let  $m \geq 2$ . Let  $\mathcal{H}$  be an admissible  $k$ -tuple, say  $\mathcal{H} = \{h_1, \dots, h_k\}$ . Let  $g$  be a natural number with  $(g, h_1 \cdots h_k) = 1$ .

If  $k \geq k_0(m)$ , and  $\mathcal{H}$  is an admissible  $k$ -tuple, then infinitely often at least  $m$  of  $gn + h_1, \dots, gn + h_k$  are prime, and the primes among  $gn + h_1, \dots, gn + h_k$  are consecutive.

### Proof of modified Shiu's theorem.

Given  $m$ , let  $k = k_0(m)$ , and take integers  $h_1, \dots, h_k$  all congruent to  $a \pmod q$  for which  $\{h_1, \dots, h_k\}$  is admissible. (It is an easy exercise to show you can choose such an  $\mathcal{H}$ .) Apply the lemma to  $qn + h_1, \dots, qn + h_k$ . We get  $\geq m$  consecutive primes, all  $\equiv a \pmod q$ .

## Some questions of Sierpiński

---

Let  $s(n)$  denote the sum of the decimal digits of  $n$ . For example,  $s(2014) = 2 + 1 + 4 = 7$ . We can observe that

$$s(1442173) = s(1442191) = s(1442209) = s(1442227).$$



### Questions (Sierpiński, 1961)

Given  $m$ , are there infinitely many  $m$ -tuples of consecutive primes  $p_n, \dots, p_{n+m-1}$  with

$$s(p_n) = s(p_{n+1}) = \dots = s(p_{n+m-1})?$$



**Answer (Thompson and P.): Yes.**

## Back to normalized prime gaps

---

Recall that the  $n$ th normalized prime gap was defined by  $\frac{p_{n+1} - p_n}{\log p_n}$ . At this point in the story, we know that 0 and  $\infty$  are limit points of the sequence of normalized prime gaps. Let  $\mathbf{L}$  denote the set of limit points.

**Erdős and Ricci** (mid 50s):  $\mu(\mathbf{L}) > 0$ .

**Hildebrand and Maier** (1988):  $\mu(\mathbf{L} \cap [1, T]) > cT$  for large  $T$ .

**Theorem (Banks–Freiberg–Maynard, 2014)**

*Given any 50 distinct real numbers  $\beta_1 < \dots < \beta_{50}$ , some  $\beta_j - \beta_i$  belongs to  $\mathbf{L}$ .*

**Corollary**

*More than 2% of the nonnegative reals belong to  $\mathbf{L}$ .*

## Maynard meets Chebotarev (spoiler: they get along)

---

Suppose that  $K/\mathbb{Q}$  is a Galois extension. For each prime rational  $p$  unramified in  $K$ , there is a well-defined Frobenius conjugacy class  $\text{Frob}_p \subset \text{Gal}(K/\mathbb{Q})$ .

Let  $C$  be a union of conjugacy classes contained in  $\text{Gal}(K/\mathbb{Q})$ . We can consider the set

$$\{p : \text{Frob}_p \in C\}.$$

A set of primes that arises in this fashion is called a **Chebotarev set**.



### Theorem (Chebotarev)

*Every Chebotarev set of primes is infinite. In fact, the proportion of primes belonging to this set is precisely  $\frac{\#C}{\#G}$ , where  $G = \text{Gal}(K/\mathbb{Q})$ .*

## Theorem (J. Thorner, 2013)

*Every Chebotarev set of primes has the bounded gaps property.*

## Theorem (J. Thorner, 2013)

*Every Chebotarev set of primes has the bounded gaps property.*

This has many splendid consequences. For example, applying Thorner's theorem to the ring class field of  $\mathbb{Z}[\sqrt{-4n}]$ , one finds:

### Corollary

*Let  $n$  be a positive integer. Then the set of primes  $p$  of the form  $x^2 + ny^2$  has the bounded gaps property.*

Key to Thorner's proof is that Chebotarev sets of primes are still well distributed in progressions (after excluding certain well-understood moduli); one has an analogue of the Bombieri–Vinogradov theorem due to Murty and Murty.

# A different number field generalization



One of these is not a real person!

Theorem (Castillo, Hall, Lemke Oliver, Pollack, Thompson, 2014)

*Let  $K/\mathbb{Q}$  be a number field. There are infinitely many pairs of prime elements  $\alpha_1, \alpha_2 \in \mathcal{O}_K$  with*

$$\max_{v \text{ infinite}} |\alpha_1 - \alpha_2|_v \leq C_K.$$

*Here  $C_K$  depends only on the number of complex embeddings of  $K$ .*



## Primitive roots

---

In Art. 315–317 of his *Disquisitiones*, Gauss studies the decimal expansion of the fractions  $a/p$  for  $1 < a < p$ .

He notes that if 10 generates the group of units modulo  $p$ , then  $a \equiv 10^j \pmod{p}$  for some  $j$ , and so

$$\frac{a}{p} \equiv 10^j \cdot \frac{1}{p} \quad \text{in } \mathbb{R}/\mathbb{Z}.$$

The expansion of the right-hand side is trivial to compute if one has already has the expansion of  $\frac{1}{p}$ .

### Example

$p = 7$ . Then  $3 \equiv 10 \pmod{7}$ , and so  $\frac{3}{7} \equiv 10 \cdot \frac{1}{7}$  in  $\mathbb{R}/\mathbb{Z}$ .  
Indeed,  $\frac{1}{7} = 0.\overline{142857}$  and  $\frac{3}{7} = 0.\overline{428571}$ .

## Question

Is 10 a primitive root mod  $p$  for infinitely many primes  $p$ ?

Numerical experiments support the guess that 10 is a primitive root for roughly  $3/8$  of the primes.

## Question

Is 10 a primitive root mod  $p$  for infinitely many primes  $p$ ?

Numerical experiments support the guess that 10 is a primitive root for roughly  $3/8$  of the primes.

Hasse's diary claims that the following conjecture was proposed on September 27, 1927:



## Conjecture (Artin)

*Yes. In fact, the proportion of primes  $p$  for which 10 is a primitive root exists and equals  $\prod_p \left(1 - \frac{1}{p(p-1)}\right) \approx 0.373956$ .*

*Moreover, an analogous holds with 10 replaced by any fixed integer  $g \neq -1$  and not a square.*

## Artin's heuristic

---

10 is a primitive root mod  $p$  iff there is no prime  $\ell$  for which the following holds:

$$\ell \mid p - 1 \quad \text{and} \quad 10^{\frac{p-1}{\ell}} \equiv 1 \pmod{p}.$$

For  $p \nmid 10\ell$ , the previous condition is equivalent (Dedekind–Kummer) to saying

$$p \text{ splits completely in } \mathbb{Q}(\zeta_\ell, 2^{1/\ell});$$

for each  $\ell$ , this determines a Chebotarev set of primes with density  $\frac{1}{\ell(\ell-1)}$ . Assuming independence (corresponding to linear disjointness) suggests a density of

$$\prod_{\ell} \left( 1 - \frac{1}{\ell(\ell-1)} \right).$$

## Theorem (Hooley, 1967)

*Assume the Generalized Riemann Hypothesis for Dedekind zeta functions. Then Artin's conjecture holds, even in quantitative form.*

## Theorem (P., 2014)

*Assume the Generalized Riemann Hypothesis for Dedekind zeta functions. For each fixed  $g \neq -1$  and not a square, the set of primes possessing  $g$  as a primitive root has the bounded gaps property. Moreover, the primes in the bounded length interval can be taken to be consecutive.*

**Forthcoming work (L. Troupe, 2014<sup>+</sup>):** Variant for polynomials over finite fields (unconditional!).

## Cyclic reductions of elliptic curves

---

Fix an elliptic curve  $E/\mathbb{Q}$ . For each prime  $p$  of good reduction, we can reduce  $E \bmod p$  to get a finite abelian group  $E(\mathbb{F}_p)$ . This is not necessarily cyclic; in general,

$$E(\mathbb{F}_p) \cong \mathbb{Z}/d_p\mathbb{Z} \oplus \mathbb{Z}/e_p\mathbb{Z},$$

where  $d_p \mid e_p$ .

One can study  $d_p$  and  $e_p$  statistically, as  $p$  varies.

### Question

Is  $d_p = 1$  for infinitely many  $p$ ?

**NO** if  $E$  has full rational 2-torsion, since then



$$E(\mathbb{F}_p) \supset \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}.$$

Modifying Hooley's conditional proof of Artin's conjecture, Serre (1976) showed:

## Theorem

*Assume GRH. Assume  $E$  has an irrational 2-torsion point. Then  $E(\mathbb{F}_p)$  is cyclic infinitely often, and in fact this occurs for a positive proportion of primes  $p$ .*

If  $E$  has CM, the GRH assumption can be removed.

(Ram Murty , 1979 and a simpler proof by Cojocaru , 2001.)

Without assuming GRH, one can still get infinitely many such  $p$  in general, but one does not get the asymptotic formula (Gupta and Ram Murty, 1990).



### Theorem (Baker and P., 2014)

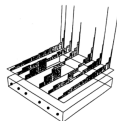
*Assume GRH. If  $E$  has an irrational 2-torsion point, then the set of  $p$  for which  $E(\mathbb{F}_p)$  is cyclic has the bounded gaps property.*

*Moreover, the primes here can be taken to be consecutive. If  $E$  has CM, the assumption of GRH can be removed.*





ONLY PROBLEMS, NOT SOLUTIONS!



## Only problems, not solutions!

---

- Are there infinitely many pairs (or triples, quadruples, etc.) of consecutive primes  $p$  belonging to a given Chebotarev set? When  $K/\mathbb{Q}$  is abelian, the answer is yes by Shiu.
- Are there infinitely many pairs of primes  $p$  and  $q$  with  $p - q$  bounded and  $p, q$  incongruent mod 4?
- Are there infinitely many pairs of primes  $p, q$  with  $3p - 2q$  bounded?



Thank you very much!