Small and large gaps in the primes

Terence Tao

University of California, Los Angeles

Latinos in the Mathematical Sciences Conference
Apr 9, 2015

Denote the sequence $2, 3, 5, 7, 11, \ldots$ of primes as $p_1, p_2, p_3, p_4, \ldots$. We can then consider the prime gaps $p_{n+1} - p_n$:

$$1, 2, 2, 2, 4, 2, 4, 2, 2, 2, 6, \ldots$$

We consider two very classical questions in analytic number theory:

- As one continues this sequence, how small can the prime gaps get?
- As one continues this sequence, how large can the prime gaps get?

There has been recent progress on both problems. In an unintuitive twist, progress on the first problem has led to progress on the second!

- Clearly, the prime gaps $p_{n+1} - p_n$ are all even once $n > 1$, since all primes after 2 are odd. So the prime gaps $p_{n+1} - p_n$ have to be at least 2 for $n > 1$.
- One of the oldest open problems in analytic number theory (going back at least to de Polignac in 1849) is the

Twin prime conjecture

We have $p_{n+1} - p_n = 2$ infinitely often.

- This conjecture remains unsolved (and for good reason - there is an important obstruction to solving it, known as the parity problem). But we have many partial results...

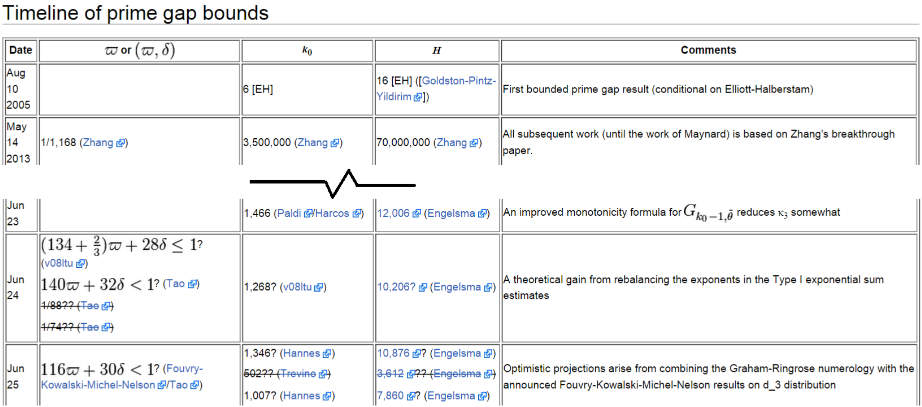
- Let X be a large number, and consider the least prime gap $p_{n+1} - p_n$ with $p_n, p_{n+1} \in [X, 2X]$.
- The prime number theorem tells us that there are $(1 + o(1))\frac{X}{\log X}$ primes in $[X, 2X]$. From the pigeonhole principle, this implies that one has $p_{n+1} - p_n \leq (1 + o(1)) \log X$ for some prime gap in $[X, 2X]$.
- This pigeonhole bound was steadily improved over the years, though not all the way to 2...

In $[X, 2X]$ for large X, one can make $p_{n+1} - p_n$ less than ...

- $(\frac{2}{3} + o(1)) \log X$ assuming the Generalised Riemann Hypothesis (GRH) (Hardy-Littlewood, 1926)
- $(\frac{3}{5} + o(1)) \log X$ assuming GRH (Rankin, 1940)
- $(1 - c + o(1)) \log X$ for some $c > 0$ (Erdős, 1940)
- $(\frac{15}{16} + o(1)) \log X$ (Ricci, 1954)
- $(0.4665 + o(1)) \log X$ (Bombieri-Davenport, 1965)
- $(0.4571 + o(1)) \log X$ (Pil'tai 1972)
- $(0.4542 + o(1)) \log X$ (Uchiyama 1975)
- $(0.4425 + o(1)) \log X$ (Huxley 1975)
- $(0.4393 + o(1)) \log X$ (Huxley 1984)
- $(0.2484 + o(1)) \log X$ (Maier 1988)

Then, multiple breakthroughs!

- $o(\log X)$ (Goldston-Pintz-Yıldırım, 2005)
- $C \log^{1/2} X (\log\log X)^2$ (Goldston-Pintz-Yıldırım, 2009)
- $\log^{1/3+o(1)} X$ (Pintz, 2013, unpublished)
- $70,000,000$ (Yitang Zhang, 21 May 2013)
- $4,680$ (Polymath8a, 27 July 2013)
- 600 (Maynard, 19 Nov 2013)
- 246 (Polymath8b, 14 Apr 2014)



Timeline of prime gap bounds

Date	ϖ or (ϖ, δ)	k_0	H	Comments
Aug 10 2005		6 [EH]	16 [EH] ([Goldston-Pintz-Yildirim])	First bounded prime gap result (conditional on Elliott-Halberstam)
May 14 2013	1/1,168 (Zhang)	3,500,000 (Zhang)	70,000,000 (Zhang)	All subsequent work (until the work of Maynard) is based on Zhang's breakthrough paper.
Jun 23		1,466 (Paldi/Harcos)	12,006 (Engelsma)	An improved monotonicity formula for $G_{k_0-1,\bar\theta}$ reduces κ_3 somewhat
Jun 24	$(134+\frac{2}{3})\varpi + 28\delta \leq 1$? (v08ltu) $140\varpi + 32\delta < 1$? (Tao) ~~1/88?? (Tao)~~ ~~1/74?? (Tao)~~	1,268? (v08ltu)	10,206? (Engelsma)	A theoretical gain from rebalancing the exponents in the Type I exponential sum estimates
Jun 25	$116\varpi + 30\delta < 1$? (Fouvry-Kowalski-Michel-Nelson/Tao)	1,346? (Hannes) ~~502?? (Trevino)~~ 1,007? (Hannes)	10,876? (Engelsma) ~~3,612??~~ ~~(Engelsma)~~ 7,860? (Engelsma)	Optimistic projections arise from combining the Graham-Ringrose numerology with the announced Fouvry-Kowalski-Michel-Nelson results on d_3 distribution

Since the work of Goldston-Pintz-Yıldırım, the results on small prime gaps have proceeded by a sieve theory argument. To explain the method, we return to the trivial pigeonhole bound of $p_{n+1} - p_n \leq (1 + o(1)) \log X$, and prove it a different way:

1. Pick a natural number n uniformly at random from $[X, 2X]$.
2. By the prime number theorem, n will be prime with probability $\frac{1+o(1)}{\log X}$.
3. Similarly, $n + 1$ will be prime with probability $\frac{1+o(1)}{\log X}$, as will $n + 2$, $n + 3$, etc.
4. Thus, for some $k = (1 + o(1)) \log X$, one can make the probabilities that each of $n, n + 1, \ldots, n + k$ are prime add up to more than 1. The pigeonhole principle then shows that with positive probability, at least two of $n, n + 1, \ldots, n + k$ are prime, giving a prime gap of at most $k = (1 + o(1)) \log X$.

The deceptively simple strategy of Goldston, Pintz, and Yıldırım modifies the above argument in one important detail:

1. Pick a natural number n non-uniformly at random from $[X, 2X]$.
2. If the probability distribution for n is chosen well, then $n + h_1$ will be prime with a large probability, for certain h_1. Similarly for $n + h_2, \ldots, n + h_k$.
3. If one can make the probabilities that each of $n + h_1, \ldots, n + h_k$ are prime sum up to more than 1, then with positive probability, at least two of the $n + h_1, \ldots, n + h_k$ are prime, giving a prime gap of size at most $\mathrm{diam}(h_1, \ldots, h_k)$.

- The difficult parts of the Goldston-Pintz-Yıldırım argument are then to (a) select the right choice of probability distribution for n, and (b) to compute the probability that $n + h_j$ is prime for various h_j.
- One wants a probability distribution which gives each of the $n + h_1, \ldots, n + h_k$ a high chance of being prime. To do this, Goldston, Pintz, and Yıldırım applied a standard sieve (the Selberg sieve) to the polynomial $P(n) := (n + h_1) \ldots (n + h_k)$ to almost eliminate those n for which $P(n)$ had too many prime factors. The probability density function for n took the form

$$c \left(\sum_{d | P(n)} \mu(d) F(\frac{\log d}{\log R}) \right)^2$$

where R was a small power of X, F was a suitable cutoff function, and c was a normalising constant.

- To compute the probabilities that $n+h_1, \dots, n+h_k$ were prime, Goldston, Pintz, and Yıldırım used standard tools in analytic number theory, and in particular the Bombieri-Vinogradov theorem on the number of primes in arithmetic progressions.
- In 2013, Yitang Zhang managed (after many difficult arguments, using deep tools such as Deligne's work on the Riemann hypothesis over finite fields) to prove a slight strengthening of the Bombieri-Vinogradov theorem which allowed him to obtain a bounded gap result $p_{n+1} - p_n \leq 70,000,000$. These methods were then pushed further by the Polymath8 project to $p_{n+1} - p_n \leq 4,680$.

Later in 2013, James Maynard (and also myself) came up with a different way to improve upon the Goldston-Pintz-Yıldırım argument, namely to replace the one-dimensional Selberg sieve weight

$$c\left(\sum_{d|P(n)} \mu(d) F(\frac{\log d}{\log R})\right)^2$$

with a multidimensional Selberg sieve weight

$$c\left(\sum_{d_1|n+h_1,\dots,d_k|n+h_k} \mu(d_1)\dots\mu(d_k) F(\frac{\log d_1}{\log R},\dots,\frac{\log d_k}{\log R})\right)^2.$$

This, combined with an optimisation of the multidimensional cutoff F, allows one to avoid the difficult arguments of Zhang to give better bounds on prime gaps (with the current record being $p_{n+1} - p_n \leq 246$).

In fact, the argument is so efficient that it can force significantly more than two of the $n + h_1, \ldots, n + h_k$ to be prime; for k large enough, it can force about $\log k$ of the quantities $n + h_1, \ldots, n + h_k$ to be prime. This strengthening has many applications; for instance, it allows one to also bound larger gaps such as $p_{n+2} - p_n$ or more generally $p_{n+m} - p_n$, not just $p_{n+1} - p_n$.

A variant of Maynard's arguments are also useful in analysing large gaps between primes, which we turn to next.

- For any large X, let $G(X)$ denote the largest prime gap $p_{n+1} - p_n$ in $[1, X]$. How does $G(X)$ grow in X?
- The prime number theorem and the pigeonhole principle give $G(X) \geq (1 + o(1)) \log X$.
- In 1920, Cramér showed that $G(X) \leq CX^{1/2} \log X$ assuming the Riemann hypothesis, and in 1936 conjectured that $G(X)$ is comparable to $\log^2 X$ using a probablistic model now known as the Cramér random model. This model was tweaked by Granville in 1995 (who predicted that $G(X)$ was at least $(2e^{-\gamma} - o(1)) \log^2 X = (1.229\dots - o(1)) \log^2 X$), but Cramér's conjecture is still widely believed (up to multiplicative constants).
- Without the Riemann hypothesis, the best known upper bound on $G(X)$ is $G(X) \leq CX^{0.525}$ (Baker-Harman-Pintz 2001).
- What about lower bounds?

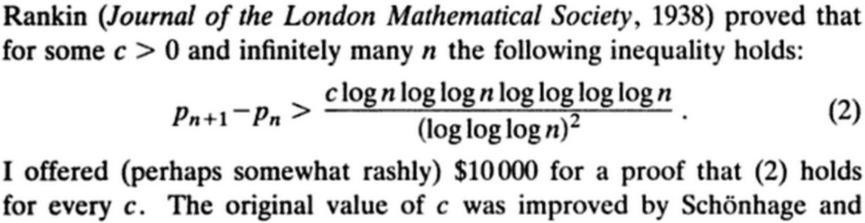
$G(X)$ is at least...

- $(1 + o(1)) \log X$ (prime number theorem + pigeonhole principle)
- $(2 + o(1)) \log X$ (Backlund, 1929)
- $(4 + o(1)) \log X$ (Brauer-Zeitz, 1930)
- $c \log X \frac{\log \log \log X}{\log \log \log \log X}$ (Westzynthius, 1931)
- $c \log X \frac{\log \log X}{(\log \log \log X)^2}$ (Erdős, 1935)
- $c \log X \frac{\log \log X \log \log \log \log X}{(\log \log \log X)^2}$ (Rankin, 1938)

One has $G(X) \geq (c - o(1)) \log X \frac{\log\log X \log\log\log\log X}{(\log\log\log X)^2}$ with

- $c = 1/3$ (Rankin, 1938)
- $c = \frac{1}{2}e^\gamma = 0.8905\ldots$ (Schönhage, 1963)
- $c = e^\gamma = 1.781\ldots$ (Ricci, 1952; Rankin, 1963)
- $c = 1.31256e^\gamma = 2.336\ldots$ (Maier-Pomerance, 1990)
- $c = 2e^\gamma = 3.562\ldots$ (Pintz, 1997)

In 1979, Erdős offered \$10,000 for a proof that c could be taken arbitrarily large.



Rankin (*Journal of the London Mathematical Society*, 1938) proved that for some $c > 0$ and infinitely many n the following inequality holds:

$$p_{n+1} - p_n > \frac{c \log n \log\log n \log\log\log\log n}{(\log\log\log n)^2} \,. \qquad (2)$$

I offered (perhaps somewhat rashly) \$10 000 for a proof that (2) holds for every c. The original value of c was improved by Schönhage and

This was finally shown by Ford-Green-Konyagin-T. (20 Aug 2014) and independently by Maynard (21 Aug 2014).

The best lower bound currently is

$$G(X) \geq (c - o(1)) \log X \frac{\log \log X \log \log \log \log X}{\log \log \log X}$$

for some $c > 0$ (improving upon Rankin's bound by a factor of $\log \log \log X$) (Ford-Green-Konyagin-Maynard-T., 16 Dec 2014). To continue Erdős's prize, I offer \$10,000 for a published proof that c can be taken to be arbitrarily large.

- All the lower bounds for $G(X)$ come from variants of the same basic construction. The first observation is that finding a large prime gap is the same thing as finding a long consecutive string of composite numbers.
- An easy construction for a long string of composite numbers is $n! + 2, n! + 3, \dots, n! + n$, because every element of $\{2, \dots, n\}$ is divisible by a prime less than or equal to n. Using Stirling's formula, this gives $G(X) \geq (1 + o(1))\frac{\log X}{\log\log X}$ - weaker than the pigeonhole principle argument!
- One can do better by using the primorial $n\# := \prod_{p \leq n} p$ instead of the factorial $n! := \prod_{j \leq n} j$. This (and the prime number theorem) recovers the pigeonhole bound $G(X) \geq (1 + o(1)) \log X$.

- As noted previously, the main reason the above construction works is because every element of $\{2, \ldots, n\}$ was divisible by at least one prime $p \leq n$. In other words, the residue classes 0 mod p for $p \leq n$ cover $\{2, \ldots, n\}$.
- More generally, if one can cover $\{1, \ldots, y\}$ with residue classes a_p mod p, one for each prime $p \leq x$, then one can (roughly speaking) show that $G(e^x) \geq y$, thanks to the Chinese remainder theorem and the prime number theorem.
- For instance, to prove Rankin's result, one needs to cover $\{1, \ldots, y\}$ by residue classes modulo primes $p \leq x$ with

$$y \asymp x \frac{\log x \log \log \log x}{(\log \log x)^2}.$$

One can think of the problem via the following "shooting" metaphor. We have a row $\{1, \ldots, y\}$ of ducks. As "ammunition", we have one rifle for each prime $p \leq x$, which can be aimed to knock out one residue class modulo p of our choosing, but each prime p can only be shot once. Our objective is to use these rifles to shoot all the ducks.

One can find efficient “shooting strategies” by deploying a variety of tricks and observations.

The first observation is that it suffices to shoot most of the ducks, rather than all of the ducks, provided one uses slightly fewer primes. More precisely, if one can eliminate all but $\frac{x}{2 \log x}$ or so of the elements of $\{1, \ldots, y\}$ using residue classes mod p for $p \leq x/2$, this is enough, because one can use each of the remaining primes $x/2 < p \leq x$ to remove one of the surviving elements, and the prime number theorem guarantees that there are enough such remaining primes to remove all the survivors.

The next observation is that the “sieve of Eratosthenes” strategy - namely, choosing the residue class 0 mod p for each p - does a reasonably good job of eliminating most of the ducks. Indeed, if one removes 0 mod p for all $p \leq x/4$ (say), then the only survivors from $\{1, \ldots, y\}$ are the number 1, together with the primes between $x/4$ and y.

But one can do better than the sieve of Eratosthenes, because some of the residue classes 0 mod p are almost redundant. Suppose one only removes 0 mod p for $p \leq y/x$ and $z < p \leq x/4$, where z is a parameter between y/x and $x/4$ to be chosen later. Of course, the primes between $x/4$ and y still survive; but the only other remaining survivors are those numbers which are z-smooth - that is to say, they only have prime factors less than or equal to z.

If one chooses z correctly, the number of such additional survivors is negligible. It turns out that the optimal choice of z is

$$z := x^{c \log\log\log x / \log\log x}$$

for some small $c > 0$. This choice arises from the asymptotics of smooth numbers, and is the main reason for all the weird logarithmic factors in the large gap results.

After using this modified sieve of Eratosthenes, there are two remaining "rounds of ammunition" remaining: the medium-sized primes p between y/x and z, and the large primes between $x/4$ and $x/2$.

For the medium-sized primes p, the most efficient thing to do is simply to "shoot randomly", selecting residue classes a_p mod p uniformly at random. This achieves a fair amount of cutdown in the number of survivors, and leads to Rankin's bound $G(X) \geq (c - o(1)) \log X \frac{\log\log X \log\log\log\log X}{(\log\log\log X)^2}$ for some small c.

For the large primes $x/4 < p \leq x/2$, the simplest thing to do is to use each prime p to shoot one of the surviving ducks. This is (more or less) Rankin's original argument.

The subsequent improvements in the constant c come (roughly speaking) from trying to select residue classes a_p mod p that each cover two survivors rather than one.

To make c larger, it is thus natural (given that the survivors consist mostly of primes) to try to look for residue classes a_p mod p that contain many primes in $\{1, \ldots, y\}$ - in particular, more than two.

We know of two ways to do this. One way, worked out by Ford, Green, Konyagin, and myself, builds upon the earlier work of Green and myself in which we located long arithmetic progressions $a, a+r, a+2r, \ldots, a+(k-1)r$ that consisted entirely of primes. It turns out that one can modify these results to find long progressions of primes in which r itself is a small multiple of a large prime p. The residue class a mod p is then guaranteed to contain lots of primes, and this turns out to be enough to make the c in Rankin's bound arbitrarily large.

The other way, discovered by Maynard, is to modify the small prime gaps arguments. These arguments produce many numbers n for which many of the numbers $n + h_1, \ldots, n + h_k$ are prime. It turns out that a variant of these arguments shows that, for any large prime p, one can find many n for which many of the $n + h_1 p, n + h_2 p, \ldots, n + h_k p$ are prime. This produces residue classes n mod p that are guaranteed to hit many primes in $\{1, \ldots, y\}$, and gives an alternate approach to making the c in Rankin's bound arbitrarily large.

The above strategies do not quite reach the current record lower bound

$$G(X) \geq (c - o(1)) \log X \frac{\log \log X \log \log \log \log X}{\log \log \log X}$$

because of a lack of “coordination” between the residue classes a_p mod p for large p; each of these residue classes may cover a lot of surviving primes, but many of the primes are covered multiple times by different residue classes, leading to a loss of efficiency.

The problem is a special case of a hypergraph covering problem: given a collection of subsets of a large set V, what is the most efficient way to cover most of V using as few subsets as possible?

The hypergraph covering problem has been extensively studied in the combinatorics literature. An efficient covering algorithm was developed by Pippenger and Spencer in 1989, using a method known as the semi-random method or the Rödl nibble. The rough idea is to select a small number of the residue classes a_p mod p randomly (this is a single “nibble”), but then remove from consideration all further residue classes which intersect the residue classes just chosen. Then take a further nibble, selecting a few more residue classes at random from the remaining pool of available classes. Iterating this process eliminates almost all of the losses coming from overlapping residue classes, and leads to the final bound

$$G(X) \geq (c - o(1)) \log X \frac{\log\log X \log\log\log\log X}{\log\log\log X}$$

which appears to be the limit of known methods.

Thanks for listening!