New bounds on gaps between primes

Andrew V. Sutherland

Massachusetts Institute of Technology

Brandeis-Harvard-MIT-Northeastern Joint Colloquium

October 17, 2013

joint work with
Wouter Castryck, Etienne Fouvry, Gergely Harcos, Emmanuel Kowalski,
Philippe Michel, Paul Nelson, Eytan Paldi, Janos Pintz, Terence Tao, Xiao-Feng Xie
and all the Polymath8 contributors

Bounded gaps between primes

Yitang Zhang

"It is proved that

$$\liminf_{n\to\infty}(p_{n+1}-p_n) < 7\times 10^7,$$

where p_n is the n-th prime.

Our method is a refinement of the recent work of Goldston, Pintz and Yildirim on the small gaps between consecutive primes.
A major ingredient of the proof is a stronger version of the Bombieri-Vinogradov theorem that is applicable when the moduli are free from large prime divisors only, but it is adequate for our purpose."

Annals of Mathematics (to appear).

received April 17, 2013, accepted May 21, 2013.

The distribution of primes

For any fixed tuple of distinct integers $\mathcal{H} = (h_1, \ldots h_k)$, let $\pi_{\mathcal{H}}(x)$ count the integers $n \le x$ such that $n + h_1, \ldots n + h_k$ are all primes.

The prime number theorem says that for any 1-tuple $\mathcal{H}$,

$$\pi_{\mathcal{H}}(x) \sim \int_2^x \frac{dt}{\log t}.$$

The distribution of primes

For any fixed tuple of distinct integers $\mathcal{H} = (h_1, \ldots h_k)$, let $\pi_{\mathcal{H}}(x)$ count the integers $n \le x$ such that $n + h_1, \ldots n + h_k$ are all primes.

The prime number theorem says that for any 1-tuple $\mathcal{H}$,

$$\pi_{\mathcal{H}}(x) \sim \int_2^x \frac{dt}{\log t}.$$

The twin prime number theorem* says that for $\mathcal{H} = (0, 2)$,

$$\pi_{\mathcal{H}}(x) \sim \Pi_2 \int_2^x \frac{dt}{(\log t)^2},$$

where $\Pi_2 = 0.66016181584686957\ldots$ is the twin prime constant.

*Not yet a theorem.

Admissible tuples

We call a k-tuple *admissible* if it does not form a complete set of residues modulo any prime (trivially satisfied for primes $p > k$).

inadmissible: $(0,1),\ (0,2,4),\ (0,2,6,8,12,14)$.
admissible: $(0,2),\ (0,2,6),\ (0,4,6,10,12,16)$.

Conjecture (Hardy-Littlewood 1923)

For any admissible k-tuple $\mathcal{H}$ we have

$$\pi_{\mathcal{H}}(x) \sim c_{\mathcal{H}} \int_2^x \frac{dt}{(\log t)^k},$$

where $c_{\mathcal{H}} > 0$ is an explicitly computable constant.

Counting primes

Prime number theorem:

$$\sum_{p \le x} 1 \sim \frac{x}{\log x}$$

Prime number theorem in arithmetic progressions ($a \perp q$):

$$\sum_{\substack{p \le x \\ p \equiv a \bmod q}} 1 \sim \frac{x}{\phi(q) \log x}$$

Counting primes

Prime number theorem:

$$\sum_{n \le x} \Lambda(n) \sim x$$

Prime number theorem in arithmetic progressions ($a \perp q$):

$$\sum_{\substack{n \le x \\ n \equiv a \bmod q}} \Lambda(n) \sim \frac{x}{\phi(q)}$$

Where

$$\Lambda(n) = \begin{cases} \log p & \text{if } n \text{ is a power of } p, \\ 0 & \text{otherwise.} \end{cases}$$

is the von Mangoldt function.

The Bombieri-Vinogradov theorem

For $a \perp q$, the *discrepancy* in the distribution of $p \equiv a \bmod q$ is

$$\Delta_\Lambda(x;a,q) = \sum_{\substack{x \le n \le 2x \\ n \equiv a \bmod q}} \Lambda(n) - \frac{1}{\phi(q)} \sum_{x \le n \le 2x} \Lambda(n)$$

(reality minus expectation).

Theorem (Bombieri-Vinogradov 1965)

For any real constants c and θ with $\theta < 1/2$:

$$\sum_{q \le x^\theta} \max_{a \perp q} |\Delta_\Lambda(x;a,q)| = O\left(\frac{x}{\log^c x}\right).$$

Elliot-Halberstam conjecture and bounded gaps

Let EH$[\theta]$ denote the bound in the Bombieri-Vinogradov theorem.

Conjecture (Elliott-Halberstam 1968)

EH$[\theta]$ *holds for all* $\theta < 1$.

Let B$[H]$ denote the claim

$$\liminf_{n\to\infty} (p_{n+1} - p_n) \leq H.$$

Theorem (Goldston-Pintz-Yildirim 2009)

For every $\theta > 1/2$ *there is an* H *for which* EH$[\theta]$ *implies* B$[H]$.
In particular, EH$[0.971] \Rightarrow$ B$[16]$.

The GPY lemma

Let DHL$[k, n]$ denote the claim that every admissible k-tuple $\mathcal{H}$ has infinitely many translates $a + \mathcal{H}$ that contain at least n primes.

The GPY lemma

Let $\mathrm{DHL}[k,n]$ denote the claim that every admissible k-tuple $\mathcal{H}$ has infinitely many translates $a+\mathcal{H}$ that contain at least n primes.

Lemma (Goldston-Pintz-Yildirim 2009)

Let $k_0 \geq 2$ and $l_0 \geq 1$ be integers, let $0 < \varpi < 1/4$, and suppose

$$1+4\varpi > \left(1+\frac{1}{2l_0+1}\right)\left(1+\frac{2l_0+1}{k_0}\right).$$

Then $\mathrm{EH}[1/2+2\varpi]$ *implies* $\mathrm{DHL}[k,2]$ *for all* $k \geq k_0$.

The optimal choice of l_0 allows one to make k_0 proportional to ϖ^{-2}.

The GPY lemma

For any integer n and admissible k-tuple $\mathcal{H}$, let

$$\Theta(n+\mathcal{H}) = \sum_{h\in\mathcal{H}} \theta(n+h)$$

where $\theta(n) = \log n$ if n is prime and 0 otherwise.

To prove DHL$[k_0, 2]$ it suffices to show that for any admissible k_0-tuple $\mathcal{H}$ there exists a function $\lambda\colon \mathbb{Z} \to \mathbb{R}$ for which

$$\sum_{x\le n\le 2x} \lambda^2(n)\,\Theta(n+\mathcal{H}) \;>\; \sum_{x\le n\le 2x} \lambda^2(n)\log(3x) \tag{1}$$

holds for all sufficiently large x. GPY show how to construct $\lambda(n)$ so that EH$[1/2+2\varpi]$ and the hypothesis on k_0 imply (1).

Bounded gaps

Let $H(k)$ denote the minimal diameter of an admissible k-tuple.
Then for ϖ and k_0 as in the GPY lemma,

$$\mathsf{EH}[1/2 + 2\varpi] \Rightarrow \mathsf{DHL}[k_0, 2] \Rightarrow \mathsf{B}[H(k_0)].$$

Bounded gaps

Let $H(k)$ denote the minimal diameter of an admissible k-tuple.
Then for ϖ and k_0 as in the GPY lemma,

$$\mathsf{EH}[1/2 + 2\varpi] \Rightarrow \mathsf{DHL}[k_0, 2] \Rightarrow \mathsf{B}[H(k_0)].$$

But proving $\mathsf{EH}[\theta]$ for any $\theta > 1/2$ appears to be out of reach.

As noted by Goldston, Pintz, and Yildirim,

"*However, any improvement in the level of distribution beyond 1/2 probably lies very deep, and even the GRH does not help.*"

A weaker version of EH$[\theta]$

Let MPZ$[\varpi, \delta]$ denote the claim that for any fixed c,

$$\sum_q |\Delta_\Lambda(x; a, q)| = O\left(\frac{x}{\log^c x}\right),$$

where q varies over x^δ-smooth squarefree integers up to $x^{1/2+2\varpi}$ and a is a fixed x^δ-coarse integer (depending on x but not q).*

A weaker version of EH$[\theta]$

Let MPZ$[\varpi, \delta]$ denote the claim that for any fixed c,

$$\sum_q |\Delta_\Lambda(x; a, q)| = O\left(\frac{x}{\log^c x}\right),$$

where q varies over x^δ-smooth squarefree integers up to $x^{1/2+2\varpi}$ and a is a fixed x^δ-coarse integer (depending on x but not q).*

Theorem (Zhang 2013)

1. *For $\varpi, \delta > 0$ there exists k_0 such that*

$$\text{MPZ}[\varpi, \delta] \Rightarrow \text{DHL}[k_0, 2].^\dagger$$

2. MPZ$[\varpi, \varpi]$ *holds for* $\varpi \leq 1/1168$.

*Zhang imposes an additional constraint on a that we don't use.

†This statement was essentially proved independently by Motohashi and Pintz.

Zhang's proof

The proof that MPZ$[\varpi, \delta] \Rightarrow$ DHL$[k_0, 2]$ follows along the lines of the GPY proof that EH$[\theta] \Rightarrow$ DHL$[k_0, 2]$.

To establish MPZ$[\varpi, \delta]$, Zhang uses the Heath-Brown identity and a combinatorial lemma to reduce the problem to estimating three types of sums (Types I, II, and III).

The most delicate part of the proof involves the Type III sums. Here Zhang relies on a result of Birch and Bombieri that depends on Deligne's proof of the Weil conjectures.*

*Specifically, the Riemann Hypothesis (RH) for varieties over finite fields.

Zhang's result

Using $\varpi = \delta = 1/1168$, Zhang shows that $\mathrm{DHL}[k_0, 2]$ holds for

$$k_0 \ \geq \ 3.5 \times 10^6.$$

He then notes that any k-tuple of primes $p > k$ is admissible, and therefore $\pi(7 \times 10^7) - \pi(3.5 \times 10^6) > 3.5 \times 10^6$ implies

$$\liminf_{n\to\infty} \, (p_{n+1} - p_n) < 7 \times 10^7. \qquad (1.5)$$

"*This result is, of course, not optimal. The condition* $k_0 \geq 3.5 \times 10^6$ *is also crude and there are certain ways to relax it. To replace the right side of (1.5) by a value as small as possible is an open problem that will not be discussed in this paper.*"

Two weeks later…

Shortly after Zhang's paper appeared it was noted that

$$\begin{aligned} p_{\pi(k_0)+k_0} - p_{k_0+1} &= 63\,374\,611 - 3\,500\,017 \\ &= 59\,874\,594. \end{aligned}$$

So using the k_0-tuple suggested by Zhang actually yields

$$\liminf\, (p_{n+1} - p_n) < 6 \times 10^7.$$

Less trivial improvements to both $H(k_0)$ and k_0 soon followed.
On June 4th the Polymath8 project was officially launched.

What is a polymath project?

According to wikipedia:

"*The Polymath Project is a collaboration among mathematicians to solve important and difficult mathematical problems by coordinating many mathematicians to communicate with each other on finding the best route to the solution.*"

The first polymath project was initiated by Tim Gowers in 2009 to find a combinatorial proof of the density Hales-Jewett theorem.

The project was a success and led to two papers published under the pseudonym D.H.J. Polymath.

Polymath8: Bounded gaps between primes

Primary goals:

1. Improving the bound on gaps between primes.
2. Understanding and clarifying Zhang's argument.

Three natural sub-projects for addressing the first goal:

1. Improving upper bounds on $H = H(k_0)$.
2. Minimizing k_0 for which MPZ(ϖ, δ) implies DHL$(k_0, 2)$.
3. Maximizing ϖ for which MPZ(ϖ, δ) holds.

Polymath8 web page.

Improved bounds on prime gaps

ϖ, δ constraint	k_0	H	comment
$\varpi = \delta = 1/1168$	3 500 000	70 000 000	Zhang's paper
$\varpi = \delta = 1/1168$	3 500 000	55 233 504	Optimize H
$\varpi = \delta = 1/1168$	341 640	4 597 926	Optimize k_0
$\varpi = \delta = 1/1168$	34 429	386 344	Make $k_0 \propto \varpi^{-3/2}$
$828\varpi + 172\delta < 1$	22 949	248 816	Allow $\varpi \neq \delta$
$280\varpi + 80\delta < 3$	873	6712	Optimize ϖ, δ constraint
$280\varpi + 80\delta < 3$	720	5414	Make k_0 less sensitive to δ
$600\varpi + 180\delta < 7$	632	4680	Further optimize ϖ, δ
$1080\varpi + 330\delta < 13$	603	4422	Subject to verification

Without relying on Deligne's theorems (just RH for curves):

$168\varpi + 48\delta < 1$	1783	14 950

A detailed timeline of improvements can be found here.

Improving the dependence of k_0 on ϖ

In both Zhang's work and the GPY theorem $k_0 \propto \varpi^{-2}$.
But this can be improved to $k_0 \propto \varpi^{-3/2}$.

Theorem (D.H.J. Polymath 2013)

Let $k_0 \geq 2$ *and* $0 < \varpi < 1/4$ *and* $0 < \delta < 1/4 + \varpi$ *satisfy*

$$(1 + 4\varpi)(1 - \kappa) > \frac{j_{k_0-2}^2}{k_0(k_0 - 1)},$$

where j_k *is the first positive zero of the Bessel function* J_k *of the first kind and* $\kappa = \kappa(\varpi, \delta, k_0)$ *is an explicit error term.*

Then $\mathrm{MPZ}[\varpi, \delta] \Rightarrow \mathrm{DHL}[k_0, 2]$.
Moreover, $\mathrm{EH}[1/2 + 2\varpi] \Rightarrow \mathrm{DHL}[k_0, 2]$ *with* $\kappa = 0$.*

*The second statement was independently proved by Farkas, Pintz, and Revesz.

Reducing the dependence of k_0 on δ

In order to maximize ϖ we would like to make δ as small as possible, but decreasing δ increases the error term $\kappa = \kappa(\varpi, \delta, k_0)$.

Using a stronger form of MPZ$[\varpi, \delta]$ where the moduli are allowed to range over values that are not necessarily x^δ-smooth but satisfy a weaker “dense-divisibility” requirement (with respect to x^δ), we are able to improve the dependence of κ on δ.

This allows us to make δ very small (thereby increasing ϖ), while still keeping κ nearly negligible.

Narrow admissible tuples

The final step to proving an explicit bound on prime gaps is constructing an admissible k-tuple with small diameter.
This gives us an upper bound on $H(k)$, hence on the gap.

This problem has been studied in relation to the *second* conjecture of Hardy and Littlewood (the first is the prime tuples conjecture).

Conjecture (Hardy-Littlewood 1923)

For all $x, y \geq 2$ *we have* $\pi(x+y) \leq \pi(x) + \pi(y)$.

Narrow admissible tuples

The final step to proving an explicit bound on prime gaps is constructing an admissible k-tuple with small diameter. This gives us an upper bound on $H(k)$, hence on the gap.

This problem has been studied in relation to the *second* conjecture of Hardy and Littlewood (the first is the prime tuples conjecture).

Conjecture (Hardy-Littlewood 1923)

For all $x, y \geq 2$ *we have* $\pi(x+y) \leq \pi(x) + \pi(y)$.

The two Hardy-Littlewood conjectures together imply

$$H(k) \geq p_k.$$

But must this be true?

Incompatibility of the Hardy-Littlewood conjectures

Theorem (Hensley-Richards 1972)

The Hardy-Littlewood conjectures are incompatible.

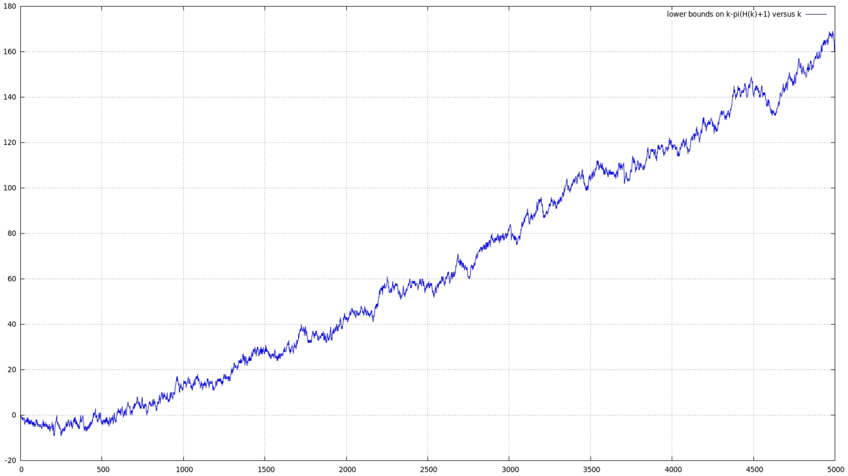
The proof uses an admissible tuple constructed by sieving an interval $[-x, x]$ of residue classes $0 \bmod p$ for increasing primes p, stopping as soon as the middle k survivors form an admissible k-tuple. They show that, asymptotically, one can make $2x < p_k$.

Theorem (Engelsma 2005)

The Hardy-Littlewood conjectures are incompatible for $k = 447$.

The proof is a 447-tuple with diameter $3158 < p_{447} = 3163$.
In the process of searching for this example, Engelsma obtained tight bounds on $H(k)$ for $k \leq 342$.

Incompatibility of the Hardy-Littlewood conjectures



Asymptotic bounds on $H(k)$

The Hensley-Richards results imply that

$$H(k) \le \big(1 + o(1)\big) k \log k.$$

Brun and Titchmarsh proved that for all sufficiently large x we have

$$\pi(x+y) - \pi(x) \le \big(1 + o(1)\big) \frac{2y}{\log y}.$$

Their proof applies to any admissible tuple, yielding the lower bound

$$H(k) \ge (1/2 - o(1)) k \log k.$$

Asymptotic bounds on $H(k)$

The Hensley-Richards results imply that

$$H(k) \le \big(1 + o(1)\big) k \log k.$$

Brun and Titchmarsh proved that for all sufficiently large x we have

$$\pi(x+y) - \pi(x) \le \big(1 + o(1)\big) \frac{2y}{\log y}.$$

Their proof applies to any admissible tuple, yielding the lower bound

$$H(k) \ge (1/2 - o(1)) k \log k.$$

We actually expect that $H(k) = k \log k + O(k)$. Numerically,

$$H(k) \le k \log k + k$$

appears to hold for all sufficiently large k.

Sieving for admissible tuples

We can explicitly construct admissible k-tuples by sieving an interval of integers of one residue class modulo each prime $p \leq k$.
But this is overkill, one can typically terminate the sieve early.

Some examples:

1. Eratosthenes: sieve $[2, x]$ at $0 \bmod p$ until the first k survivors are admissible.
2. Hensley-Richards: sieve $[-x, x]$ at $0 \bmod p$ until the middle k survivors are admissible.
3. Schinzel: sieve $[2, x]$ at $1 \bmod p$ for $p \leq y$ and $0 \bmod p$ for $p > y$ until the first k survivors are admissible.
4. Greedy: sieve $[0, x]$ of a minimally occupied residue class $a \bmod p$ until the first k survivors are admissible.

Shifting the sieve interval slightly often yields better results.

Discipline versus greed

All of the structured approaches are demonstrably sub-optimal, and the greedy approach is usually worse!

However, there is a hybrid approach works that remarkably well.

Let $w = k \log k + k$, and for even integers s in $[-w, w]$:

1. Sieve $[s, s+w]$ at $1 \bmod 2$ and $0 \bmod p$ for primes $p \le \sqrt{w}$.
2. For increasing primes $p > \sqrt{w}$ sieve a minimally occupied residue class mod p until the tuple $\mathcal{H}$ of survivors is admissible.
3. If $|\mathcal{H}| \ne k$, adjust the sieving interval and repeat until $|\mathcal{H}| = k$.

Output an $\mathcal{H}$ with minimal diameter among those constructed.

Discipline versus greed

All of the structured approaches are demonstrably sub-optimal, and the greedy approach is usually worse!

However, there is a hybrid approach works that remarkably well.

Let $w = k \log k + k$, and for even integers s in $[-w, w]$:

1. Sieve $[s, s+w]$ at $1 \bmod 2$ and $0 \bmod p$ for primes $p \leq \sqrt{w}$.
2. For increasing primes $p > \sqrt{w}$ sieve a minimally occupied residue class mod p until the tuple $\mathcal{H}$ of survivors is admissible.
3. If $|\mathcal{H}| \neq k$, adjust the sieving interval and repeat until $|\mathcal{H}| = k$.

Output an $\mathcal{H}$ with minimal diameter among those constructed.

This algorithm is not optimal, but it typically gets within one percent of the best known results (including cases where $H(k)$ is known).

(demo)

Sieve comparison

k	632	1783	34 429	341 640	3 500 000
k primes past k	5028	16 174	420 878	5 005 362	59 874 594
Eratosthenes	4860	15 620	411 946	4 923 060	59 093 364
H-R	4918	15 756	402 790	4 802 222	57 554 086
Shifted H-R	4876	15 470	401 700	4 788 240	57 480 832
Shifted Schinzel	4868	15 484	399 248	4 740 846	56 789 070
Shifted hybrid	4710	15 036	388 076	4 603 276	55 233 744
Best known	4680	14 950	386 344	4 597 926	55 233 504
$\lfloor k \log k + k \rfloor$	4707	15 130	394 096	4 694 650	56 238 957

Further optimizations

Given a narrow admissible tuple, there a variety of combinatorial optimization methods that we can apply to try and improve it. These include local search and perturbation methods.

The technique that we have found most useful can be described as a type of *genetic algorithm*.

Further optimizations

Given a narrow admissible tuple, there a variety of combinatorial optimization methods that we can apply to try and improve it. These include local search and perturbation methods.

The technique that we have found most useful can be described as a type of *genetic algorithm*.

Given a k-tuple $\mathcal{H}$, we generate a new tuple $\mathcal{H}'$ by sieving the same interval of the same residue classes for $p \leq \sqrt{k \log k}$, and randomly choosing a nearly minimally occupied class for $p > \sqrt{k \log k}$.

The set $\mathcal{H} \cup \mathcal{H}'$ contains an admissible k-tuple (namely, $\mathcal{H}$), but if we sieve this set by greedily choosing residue classes as required, we may obtain a k-tuple $\mathcal{H}''$ that is actually narrower than $\mathcal{H}$.

Database of admissible tuples

We have established an online database of admissible tuples.

It includes at least one example of an admissible k-tuple of least known diameter for $2 \le k \le 5000$ and is open for submission.

For $k \le 342$ it contains optimal tuples contributed by Engelsma. For many $k > 342$ we have tuples narrower than those obtained by Engelsma, and in every case we are able to match his results.

Finding better lower bounds for $H(k)$ remains an open problem.

For $k = 632$ we have $H(632) \ge 4276$, but we suspect that in fact the upper bound $H(632) \le 4680$ is tight.

(admissible 632-tuple of diamter 4680)