Lecture 3: Small Prime Gaps: Tuple Approximations, Selberg Sieve, and Almost Primes

Daniel Goldston

Summary of Last Lecture: Where we were in April 2003:

Good News: We had learned how to compute (see below in 10 pages)

$\mathcal{S}_k(N, \boldsymbol{k}, \boldsymbol{a})$

$$= \sum_{n=1}^{N} \Lambda_R(n+k_1)^{a_1}\Lambda_R(n+k_2)^{a_2}\cdots\Lambda_R(n+k_r)^{a_r}$$

and

$\tilde{\mathcal{S}}_k(N, \boldsymbol{k}, \boldsymbol{a})$

$$= \sum_{n=1}^{N} \Lambda_R(n+k_1)^{a_1}\cdots\Lambda_R(n+k_{r-1})^{a_{r-1}}\Lambda(n+k_r)$$

where $\boldsymbol{k} = (k_1, k_2, \ldots, k_r)$ and $\boldsymbol{a} = (a_1, a_2, \ldots a_r)$, the k_i's are distinct integers, $a_i \geq 1$ and $\sum_{i=1}^{r} a_i = k$. In the mixed correlation we assume that $r \geq 2$ and take $a_r = 1$.

Theorem 1 *Given $k \geq 1$, $\max_i |j_i| \leq R$ and $R \geq 2$. Then*

$$\mathcal{S}_k(N, \boldsymbol{j}, \boldsymbol{a}) = \Big(\mathcal{C}_k(\boldsymbol{a})\mathfrak{S}(\boldsymbol{j}) + o_k(1)\Big) N(\log R)^{k-r} + O(R^k).$$

For $N^{\epsilon} \ll R \ll N^{\frac{1}{2(k-1)}}$,

$$\tilde{\mathcal{S}}_k(N, \boldsymbol{j}, \boldsymbol{a}) = \Big(\mathcal{C}_k(\boldsymbol{a})\mathfrak{S}(\boldsymbol{j}) + o(1)\Big) N(\log R)^{k-r}.$$

The $\mathcal{C}_k(\boldsymbol{a})$ are rational numbers, and Denoting $\mathcal{C}_k(k)$ as $\mathcal{C}_k$. With $\boldsymbol{a} = (a_1, a_2, \ldots, a_r)$

$$\mathcal{C}_k(\boldsymbol{a}) = \prod_{i=1}^{r} \mathcal{C}_{a_i}.$$

Here

$$\sum_{n \leq N} \Lambda_R(n)^k \sim \mathcal{C}_k N(\log R)^{k-1}$$

Currently (2006) we only know the values of the first six correlation constants:

$$\mathcal{C}_1 = 1, \quad \mathcal{C}_2 = 1, \quad \mathcal{C}_3 = \frac{3}{4}, \quad \mathcal{C}_4 = \frac{3}{4},$$
$$\mathcal{C}_5 = \frac{11065}{2^{14}} = .675\ldots,$$
$$\mathcal{C}_6 = \frac{11460578803}{2^{34}} = .667\ldots\ .$$

Bad News:

1. Not knowing $\mathcal{C}_k$ means we can not use our formulas for more than 7th moment approximations — A BIG PROBLEM.

2. We knew that $\mathcal{C}_k \to \infty$, actually $\mathcal{C}_k \gg k^k$, so even if we can compute them our approximations are getting rapidly increasingly lousy.

3. Our new improved approximation fails in all important cases to converge.

4. Everyone knows our proof bombed.

More bad news later that year

1. There seemed to be no way to fix the new approximation.

2. The numerical evidence that suggested there was a better approximation because multiple truncations approximations improved the gap results, while correct, disappear as the number of moments increase and seem to disappear entirely.

Sadder but Wiser

Actually in understanding our proof before discovering it was wrong, Granville and Sound found a much better way to approach the problem, which we now use.

Let $\mathcal{H} = \{h_1, h_2, \ldots, h_k\}$ be distinct integers.

Let $\nu_{\mathcal{H}}(p)$ denote the number of distinct residue classes $\pmod p$ the numbers $h \in \mathcal{H}$ fall into, and extend this definition to $\nu_{\mathcal{H}}(d)$ for squarefree integers d by multiplicativity. Define the singular series

$$\mathfrak{S}(\mathcal{H}) = \prod_p \left(1 - \frac{1}{p}\right)^{-k} \left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right)$$

If $\mathfrak{S}(\mathcal{H}) \neq 0$ then $\mathcal{H}$ is called *admissible*. Thus $\mathcal{H}$ is admissible if and only if $\nu_p(\mathcal{H}) < p$ for all p.

Prime Tuple Conjecture: All the components of

$$(n + h_1, n + h_2, \ldots, n + h_k)$$

are prime for infinitely many n whenever $\mathcal{H}$ is admissible.

Define

$$\Lambda(n; \mathcal{H}) = \Lambda(n + h_1)\Lambda(n + h_2) \cdots \Lambda(n + h_k).$$

Then Hardy and Littlewood conjectured that for $\mathcal{H}$ admissible,

$$\sum_{n \le N} \Lambda(n; \mathcal{H}) = N\Big(\mathfrak{S}(\mathcal{H}) + o(1)\Big), \quad \text{as} \quad N \to \infty.$$

Our goal is to construct a prime tuple approximation function $\Lambda_R(n;\mathcal{H})$ for $\mathcal{H}$ and then detect if there are primes in $\mathcal{H}$ by showing

$$\sum_{n=N+1}^{2N}\Big(\sum_{h_i\in\mathcal{H}}\Lambda(n+h_i)-\log 3N\Big)\Lambda_R(n;\mathcal{H})^2$$

is positive. If this fails for a single admissible tuple, we may try this for the union of a set of many admissible tuples.

Approximating Prime Tuples

Going back to our original

$$\Lambda_R(n) = \sum_{\substack{d|n \\ d\leq R}} \mu(d) \log \frac{R}{d},$$

our approximation for $\Lambda(n;\mathcal{H})$ is

$$\Lambda_R(n;\mathcal{H}) = \Lambda_R(n+h_1)\Lambda_R(n+h_2)\cdots\Lambda_R(n+h_k).$$

We now prove: For $\mathcal{H}_1$ and $\mathcal{H}_2$ are both sets of distinct positive integers $\leq h$, $|\mathcal{H}_1| = k_1$, $|\mathcal{H}_2| = k_2$, and let $k = k_1 + k_2$.

Proposition 1 *Let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$, and $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$. If $R = o(N^{\frac{1}{k}})$ and $h \leq R^A$ for any large constant $A > 0$, then we have for $R, N \to \infty$,*

$$\sum_{n\leq N} \Lambda_R(n;\mathcal{H}_1)\Lambda_R(n;\mathcal{H}_2) = N\Big(\mathfrak{S}(\mathcal{H}) + o_k(1)\Big)(\log R)^r.$$

Proposition 2 *Let* $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$, , $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$, *and* $1 \le h_0 \le h$. *Let* $\mathcal{H}_0 = \mathcal{H} \cup \{h_0\}$, *and* $r_0 = r$ *if* $h_0 \notin \mathcal{H}$ *and* $r_0 = r + 1$ *if* $h_0 \in \mathcal{H}$. *If* $R \ll_k N^{\frac{1}{2k}}(\log N)^{-B(k)}$ *for a sufficiently large positive constant* $B(k)$, *and* $h \le R^{\frac{1}{2k}}$, *then we have for* $R, N \to \infty$,

$$\sum_{n \le N} \Lambda_R(n; \mathcal{H}_1)\Lambda_R(n; \mathcal{H}_2)\Lambda(n + h_0) = N\Big(\mathfrak{S}(\mathcal{H}_0) + o_k(1)\Big)(\log R)^{r_0}.$$

Assuming the Elliott-Halberstam conjecture, then this holds for $R \ll_k N^{\frac{1}{k}-\epsilon}$ *with any* $\epsilon > 0$.

Notice there are no correlation coefficients because no Λ_R is raised to a power greater than 2. The key idea is that with moments you consider

$$\left(\sum_{1 \le h_1 \le h} \Lambda_R(n + h_1) \right)^k$$

$$= \sum_{1 \le h_1, h_2, \ldots h_k \le h} \Lambda_R(n + h_1) \Lambda_R(n + h_2) \cdots \Lambda_R(n + h_k)$$

But this includes all the nondistinct ways the h_i's occur.

SOLUTION

$$\left(\sum_{1\le h_1\le h} \Lambda_R(n+h_1)\right)^k$$
$$= \sum_{1\le h_1,h_2,\ldots h_k\le h} \Lambda_R(n+h_1)\Lambda_R(n+h_2)\cdots\Lambda_R(n+h_k)$$

while obviously you should just use

$$\sum_{\substack{1\le h_1,h_2,\ldots,h_k\le h\\ \text{distinct}}} \Lambda_R(n+h_1)\Lambda_R(n+h_2)\cdots\Lambda_R(n+h_k)$$

Using

$$\sum_{\substack{1\le h_1,h_2,\ldots,h_k\le h\\ \text{distinct}}} \Lambda_R(n;\mathcal{H})^2$$

to detect all tuples in $[1,h]$ gives in two lines

$$\Delta = \liminf_{n\to\infty}\left(\frac{p_{n+1}-p_n}{\log p_n}\right) < 3/4$$

because Prop 2 gives you $k\log R$ extra from hitting your tuple approximation, and $R = N^{1/4k}$.

In effect you detect 1/4 of a prime in your tuple by this approximation.

Next, we can optimize by considering not just a k-tuple approximation, but a linear combination of all tuple approximations:

$$a_0 + \sum_{j=1}^{k} a_j \Bigg(\sum_{\substack{1 \le h_1, h_2, \ldots, h_j \le h \\ \text{distinct}}} \Lambda_R(n; \mathcal{H}_j) \Bigg).$$

Since we now have (in effect) Poisson moments, the result turns out to be the least zero asymptotics of a certain Laguerre polynomial as found by Rubinstein. That our method actually has this solution was proved by Bombieri and Percy Deift. The result is

$$\Delta = \liminf_{n \to \infty} \left(\frac{p_{n+1} - p_n}{\log p_n} \right) < 1/4.$$

Trying to do better pays off

We actually prove both Propositions with the error term $o_k(1)$ replaced by a series of lower order terms, which however are not needed in any of our applications.

Granville having a preprint and knowing about this work was asked by Ben Green about a sieve bound he couldn't find in the literature.

Approximating Prime Tuples - New Idea

This idea came out of a paper of Heath-Brown (1997).

Instead of the tuple $(n+h_1, n+h_2, \ldots, n+h_k)$ consider the polynomial

$$\mathcal{P}(n, \mathcal{H}) = (n+h_1)(n+h_2)\ldots(n+h_k)$$

Then tuple is prime tuple when $\mathcal{P}$ has k prime factors or $\leq k$ prime factors. The generalized von Mangoldt function

$$\Lambda_k(n) = \sum_{d|n} \mu(d)(\log \frac{n}{d})^k$$

is non-zero on P_k's, but vanishes otherwise, so

$$\Lambda_k(\mathcal{P}(n, \mathcal{H}))$$

detects prime tuples, and we can approximate this with

$$\Lambda_R(n; \mathcal{H}) = \frac{1}{k!} \sum_{\substack{d|P_{\mathcal{H}}(n) \\ d \leq R}} \mu(d) \left(\log \frac{R}{d}\right)^k.$$

By chance this is extremely close to the failed approximation of 2003.

However, this approximation fails to prove any of the results above, giving $\Delta \leq .135\ldots$.

What Makes Everything Work

This idea is actually what Heath-Brown (1997) was working on.

To detect *SOME* primes in tuples, you only need to show that $P(n, \mathcal{H})$ has less than $k+\ell$ prime factors, for some $\ell < k$. Thus we should approximate with $\Lambda_{k+\ell}$. Hence define

$$\Lambda_R(n; \mathcal{H}, \ell) = \frac{1}{(k+\ell)!} \sum_{\substack{d \mid P(n,\mathcal{H}) \\ d \le R}} \mu(d) \left(\log \frac{R}{d}\right)^{k+\ell}.$$

We prove

Theorem 1. For $R \le N^{1/2}/(\log N)^{2k}$ and $h \le R^{\epsilon}$ with $R, N \to \infty$ we have

$$\sum_{n \le N} \Lambda_R(n; \mathcal{H}, \ell)^2 = \binom{2\ell}{\ell} \frac{(\log R)^{k+2\ell}}{(k+2\ell)!} (\mathfrak{S}(\mathcal{H}) + o(1)) N.$$

Theorem 2.

If the Bombieri-Vinogradov theorem holds with $Q = N^{\vartheta-\epsilon}$,, then for $R \leq N^{\vartheta/2-\epsilon}$ and $h \leq R^{\epsilon}$ with $R, N \to \infty$ then if $h_0 \notin \mathcal{H}$

$$\sum_{n \leq N} \Lambda_R(n; \mathcal{H}, \ell)^2 \Lambda(n + h_0) =$$

$$\binom{2\ell}{\ell} \frac{(\log R)^{k+2\ell}}{(k+2\ell)!} (\mathfrak{S}(\mathcal{H}) + o(1)) N,$$

while if $h_0 \in \mathcal{H}$ we make the replacement $\ell \to \ell + 1$ and $k \to k - 1$.

Actually Soundararajan found that you can handle

$$\Lambda_R(n;\mathcal{H},P) = \sum_{\substack{d|P(n,\mathcal{H})\\ d\le R}} \mu(d)P(\frac{R}{d}).$$

Let P be a polynomial vanishing to order k at 0. Then

$$\sum_{n\le N} \Lambda_R(n;\mathcal{H},P)^2$$

is

$$\begin{aligned}&\sim \mathfrak{S}(\mathcal{H})\frac{(\log x)^{-k}}{(k-1)!}\int_0^1 y^{k-1}P^{(k)}(1-y)^2dy\,N\\ &\sim \mathfrak{S}(\mathcal{H})I(k)N.\end{aligned}$$

and

$$\sum_{n\le N} \Lambda(n+h_0)\Lambda_R(n;\mathcal{H},P)^2$$

is

$$\sim \mathfrak{S}(\mathcal{H})I(k-1)N \quad \text{if} \quad h_0 \in \mathcal{H}$$

and

$$\sim \mathfrak{S}(\mathcal{H}\cup\{h_0\})I(k)N \quad \text{if} \quad h_0 \notin \mathcal{H}$$

Goldston-Pintz-Yildirim New Result

Theorem. $\Delta = 0$, i. e.

$$\liminf_{n\to\infty}\left(\frac{p_{n+1}-p_n}{\log p_n}\right) = 0.$$

Theorem. (Preliminary) For some $C > 0$,

$$p_{n+1}-p_n < C(\log p_n)^{1/2}(\log\log p_n)^2 \quad \text{infinitely often.}$$

Main limitation: We can not yet prove the same results for $p_{n+r} - p_n$ for $r \geq 2$.

G-P-Y Conditional Results

Let $\theta(N; m, a)$ denote the number of primes $\leq N$ which are $\equiv a \pmod{m}$. The Bombieri-Vinogradov theorem states that for any $\epsilon > 0$ and $A > 0$ we have

$$\sum_{m \leq Q} \max_{\substack{a \\ (a,m)=1}} \left| \theta(N; m, a) - \frac{N}{\phi(m)} \right| \ll \frac{N}{(\log N)^A}$$

for $Q = N^{1/2}/(\log N)^{B(A)}$.

If this holds for $Q = N^{\vartheta - \epsilon}$ we say the primes have *level of distribution* ϑ.

BV implies $\vartheta = 1/2$ is true.

Elliott-Halberstam conjecture $\vartheta = 1$ is true.

Theorem. If the primes have level of distribution ϑ for a value of $\vartheta > 1/2$, then

$$p_{n+1} - p_n \leq C(\vartheta) \quad \text{infinitely often.}$$

In particular, if $\vartheta \geq .971$

$$p_{n+1} - p_n \leq 16 \quad \text{infinitely often.}$$

More generally: **Weak form of Hardy-Littlewood prime tuple conjecture.** If $\vartheta > 1/2$, then every admissible k-tuple with $k \geq c(\vartheta)$ contains at least two primes infinitely often.

ϑ	k	L	$h(k)$
1	6	1	16
.95	7	1	20
.90	8	2	26
.85	10	2	32
.80	12	2	42
.75	16	2	60
.70	22	4	90
.65	35	4	158
.60	65	6	336
.55	193	9	1204

Limitation: So far, the method only produces two primes in large enough admissible tuples, never three or more.

Goldston-Motohashi-Pintz-Yildirim Result

We prove $\Delta = 0$ and in an 8 page paper.

Goldston-Sid Graham-Pintz-Yildirim Results

1) Second proof of above results, using sieve methods, no contour integrals or zero-free regions beyond what is needed for BV theorem.

2) Let P_k be a number with $\leq k$ distinct prime factors, let E_k be a number with exactly k distinct prime factors.

Chen: $p + 2 = P_2$ infinitely often.

Also $2n = p + P_2$

Theorem(Preliminary) Every admissible k-tuple with $k = k(r)$ sufficiently large will contain at least r E_2 numbers infinitely often.

Also if q_n is the n-th E_2 number, then

$$q_{n+1} - q_n \leq 6 \quad \text{infinitely often}$$

Sketch of Proof

For $\ell \geq 0$ and $R = N^{\vartheta/2-\epsilon/4}$, we have

$$\begin{aligned}
\mathcal{S} := & \sum_{n=N+1}^{2N} \left(\sum_{i=1}^{k} \Lambda(n+h_i) - \log 3N \right) \Lambda_R(n;\mathcal{H},\ell)^2 \\
&= k\binom{2\ell+2}{\ell+1}\frac{(\log R)^{k+2\ell+1}}{(k+2\ell+1)!}(\mathfrak{S}(\mathcal{H})+o(1))N \\
&\qquad - \log 3N\binom{2\ell}{\ell}\frac{(\log R)^{k+2\ell}}{(k+2\ell)!}(\mathfrak{S}(\mathcal{H})+o(1))N \\
&= \left(\frac{2k}{k+2\ell+1}\frac{2\ell+1}{\ell+1}\log R - \{1+o(1)\}\log 3N\right)\mathcal{M} \\
&\geq \left(\frac{k}{k+2\ell+1}\frac{\ell+1/2}{\ell+1}\cdot 2\vartheta - 1 - \epsilon\right)(\log 3N)\mathcal{M},
\end{aligned}$$

where

$$\mathcal{M} = \frac{(\log R)^{k+2\ell}}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H})N.$$

The term inside the brackets is greater than a positive constant (for ϵ sufficiently small) provided $\frac{k}{k+2\ell+1}\frac{\ell+1/2}{\ell+1}\cdot 2\vartheta > 1$. Evidently one can choose such k and ℓ for any $\vartheta > \frac{1}{2}$; in particular one can take $\ell = 1$ and $k = 7$ when $\vartheta > 20/21$.

We just fail to prove the result unconditionally (as we had to take $\theta > 1/2$), so the question is as to how we can "win an ϵ". To do this we consider instead

$$\sum_{n=N+1}^{2N} \left(\sum_{1 \le h_i \le h} \Lambda(n + h_i) - \log 3N \right) \times \sum_{\substack{1 \le h_1, h_2, \ldots, h_k \\ \text{distinct}}} \Lambda_R(n; \mathcal{H}, \ell)^2$$

and use a result the Gallagher for the average of the singular series. This gives an addition factor of h in the previous factor, hence with $\vartheta = 1/2$ this expression becomes positive if $h > \epsilon \log N$.

Alternatively we can just throw $\epsilon \log N$ more primes into the sum on the left.