On Zhang's proof of the bounded gap conjecture

p_n: nth prime, nth prime gap: $p_{n+1}-p_n$ (all even except $p_2-p_1=3-2=1$)

Conjectures on small prime gaps:

1. twin prime conjecture: $\liminf_{n\to\infty}(p_{n+1}-p_n)=2$? ($\Leftrightarrow \exists\infty$ many n: $p_{n+1}-p_n=2$)

2. de Polignac's conjecture: $\forall k\ \exists\infty$ many n: $p_{n+1}-p_n=2k$.

3. bounded gap conjecture: $\exists C>0$: $\liminf_{n\to\infty}(p_{n+1}-p_n)\le C$

$\Leftrightarrow \exists C>0\ \exists\infty$ many n: $p_{n+1}-p_n\le C$.

(1. and 2. open, 3. solved by Zhang. Progress started ≈ 2007, this was the work by:)

GPY [Goldston, Pintz, Yildirim, 2009]:

(a) $\liminf_{n\to\infty}\frac{p_{n+1}-p_n}{\sqrt{\log p_n}\cdot(\log\log p_n)^2}<\infty$

(b) Elliott-Halberstam-Conj. $\Rightarrow \liminf_{n\to\infty}(p_{n+1}-p_n)\le 16$.

[Y. Zhang 2013+]: Conj. 3 is true with $C=7\cdot 10^7$.

[J. Pintz, 9.6.2013]: " $C=2.530.338$.

Polymath-Project, 13.6.2013: " $C=248.910$.

Thm. 1 (Zhang): Let $\mathcal{H}=\{h_1,h_2,\dots,h_k\}\subseteq\mathbb{N}_0$ be admissible, $k\ge 3{,}5\cdot 10^6$. Then $\exists\infty$ many n: $\{n+h_1,\dots,n+h_k\}$ contains at least 2 primes.

$\hookrightarrow$ [Conj. 3 since $\pi(7\cdot 10^7)-\pi(3{,}5\cdot 10^6)>3{,}5\cdot 10^6$.] [HL-Conj. $\Rightarrow$ all prime for ∞ many n]

Def.: $\mathcal{H}$ adm. $:\Leftrightarrow \forall p$: $\nu_p(\mathcal{H})<p$, $\nu_p(\mathcal{H}):=\#\{h_i \bmod p \mid 1\le i\le k\}$.

GPY-Idea: Let $\theta(n):=\begin{cases}\log n, & n \text{ prime},\\ 0, & \text{else},\end{cases}$ let $x>0$ be large, $\mathcal{L}:=\log x$.

[PNT: $\sum_{n\le x}\theta(n)\sim x$]

Consider $f:\mathbb{N}\cap[x,2x[\,\to\mathbb{R}_{>0}$, $S_1:=\sum_{n\sim x}f(n)$, $S_2:=\sum_{n\sim x}\left(\sum_{j=1}^{k}\theta(n+h_j)\right)f(n)$.

[$n\sim x :\Leftrightarrow x\le n<2x$]

Goal: Show $S_2>S_1\cdot\log(3x)$ for a function f (dep. on x) and all large x.

Then $\sum_{j=1}^{k}\theta(n+h_j)>\log(3x)$ for some $n\sim x$, so $\exists i\ne j$: $n+h_i$, $n+h_j$ both prime.

[one $\theta(n+h_j)>0 \Rightarrow \sum=\log(n+h_j)<\log(2x+x)=\log(3x)$ if $x\ge h_j$.]

GPY-Sieve: $f(n) := \lambda(n)^2$, $\lambda(n) = \sum_{\substack{d \le D \\ d \mid P(n)}} \mu(d) \left(\log\left(\frac{D}{d}\right)\right)^{k+\ell}$, $\ell \approx \sqrt{k}$, $D = x^{\vartheta}$, $P(n) := \prod_{j=1}^{k} (n + h_j)$.

(Refinement of Selberg's sieve)

Def.: For $\gamma : \mathbb{N} \to \mathbb{R}_{>0}$ let $\Delta(\gamma; d, c) := \sum_{\substack{n \sim x \\ n \equiv c(d)}} \gamma(n) - \frac{1}{\varphi(d)} \sum_{\substack{n \sim x \\ n \equiv c(d)}} \gamma(n)$ for $(c,d) = 1$.

Bombieri-Vinogradov (1967): $\sum_{d \sim x^{\vartheta}} \max_{(c,d)=1} |\Delta(\theta; d, c)| = o(x)$ with $\vartheta = \frac{1}{2} - \varepsilon$ [often serves as a replacement of GRH]

Elliott-Halberstam-Conjecture: " holds true with $\vartheta = 1 - \varepsilon$ [ϑ is called exponent of distribution]

GPY showed: $S_2 - S_1 \log(3x) \underset{(\geq)}{=} (k J_2^* - J_1^* \mathcal{L}) x + O(\mathcal{E}_1) + O(\mathcal{E}_2)$, where

$\mathcal{E}_1 := x \mathcal{L}^{k+2\ell}$ and $\mathcal{E}_2 := \sum_{\ell \le k} \sum_{d < D^2} \overbrace{\mu^2(d) \tau_3(d) \tau_{k-1}(d)}^{\text{harmless factors}} \cdot \sum_{c, d \mid P(c-h_i)} |\Delta(\theta; d, c)|$

If $D = x^{1/4 + \omega}$, then $k J_2^* - J_1^* \mathcal{L} >> \mathcal{L}^{k+2\ell+1}$, so $\mathcal{E}_1$ ok.

Also $\mathcal{E}_2$ ok if $\vartheta = \frac{1}{2} + 2\omega$ would work in BV's Theorem. $\rightsquigarrow$ EH-Conj. provides the result

Zhang's idea: truncate d in $\lambda(n)$, set $\lambda(n) := \sum_{\substack{d \le D \\ d \mid (P(n), \mathcal{P})}} \mu(d) \cdot \left(\log\left(\frac{D}{d}\right)\right)^{k+\ell} \cdot \frac{1}{(k+\ell)!}$, $\mathcal{P} := \prod_{p < x^{\omega}} p$.

(take smooth d)

Observations: 1) main term $(k J_2^* - J_1^* \mathcal{L}) x$ still good compared to $\mathcal{E}_1$

(Contributions from d with large prime factors are relatively small.)

2) $\mathcal{E}$-estimate works: Thm. 2 (Zhang): $\sum_{\substack{d < D^2 \\ d \mid \mathcal{P}}} \sum_{c, d \mid P(c-h_i)} |\Delta(\theta; d, c)| \ll \frac{x}{\mathcal{L}^A}$, $D = x^{\frac{1}{4} + \omega}$ for a certain $\omega > 0$.

keywords:

[Variant of dispersion method [Linnik, for bilinear forms with exponentials] from BFI [Bombieri, Friedlander, Iwaniec 1987]:

Factor the d with $x^{1/2 - \varepsilon} < d < D^2$, $d \mid \mathcal{P}$, $\rightsquigarrow \sum_d = \sum_r \sum_q$ $\rightsquigarrow$ char. factn. of these d is well-factorable

as $d = rq$ with $\frac{R}{x^{\omega}} < r < R$ for any R

- Start with Heath-Brown's identity $\rightsquigarrow$ Split in sums of type I, II, III
- Use a truncated Poisson-identity (Fourier-Trick)

and get exponential sums to be bounded $\rightsquigarrow$ some nontrivial, but standard expo-Σ-estimates suffice in some cases

$\rightsquigarrow$ Zhang also uses bounds for incomplete Kloosterman sums

[Kloosterman sum: $S(a, b; c) := \sum_{\substack{x \bmod c \\ (x,c)=1}} e\left(\frac{ax + b\bar{x}}{c}\right)$, $x\bar{x} \equiv 1 (c)$]]