PATTERNS OF PRIMES IN ARITHMETIC PROGRESSIONS

JÁNOS PINTZ
Rényi Institute of the
Hungarian Academy of Sciences

CIRM, Dec. 13, 2016

1. Patterns of primes

Notation: p_n the n^{th} prime, $\mathcal{P} = \{p_i\}_{i=1}^{\infty}$, $d_n = p_{n+1} - p_n$.

Abbreviation: i.o. = infinitely often, $\mathbb{Z}^+ = \{1, 2, \dots\}$

Twin Prime Conjecture $\{n, n+2\} \in \mathcal{P}^2$ i.o. $\Longleftrightarrow d_\nu = 2$ i.o.

Polignac Conjecture (1849) $2 \mid h \longrightarrow d_n = h$ i.o.

Definition

$\mathcal{H} = \mathcal{H}_k = \{h_i\}_{i=1}^{k}$, $0 \le h_1 < h_2 < h_k$ is admissible if the number of residue classes covered by $\mathcal{H} \bmod p$, $\nu_p(\mathcal{H}) < p$ for every prime p.

Dickson Conjecture (1904). *$\mathcal{H}_k$ admissible $\Longrightarrow \{n + h_i\}_{i=1}^{k} \in \mathcal{P}^k$ i.o.*

Hardy–Littlewood Conjecture (1923). $\mathcal{H} = \mathcal{H}_k$ *admissible* $\Longrightarrow$

$$\sum_{\substack{n \le x \\ \{n+h_i\} \in \mathcal{P}^k}} 1 \sim \frac{x}{\log^k x} \sigma(\mathcal{H}),$$

$$\sigma(\mathcal{H}) = \prod_p \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right) \left(1 - \frac{1}{p}\right)^{-k}$$

Remark 1. HL conjecture implies:

Strong HL conjecture: $\mathcal{H} = \mathcal{H}_k$ admissible $\Longrightarrow$

$$\sum_{\substack{n \le x \\ \{n+h_i\} \in \mathcal{P}^k \\ n+h \notin \mathcal{P} \text{ if } h \in [0,h_k],\ h \ne h_i}} 1 \sim \frac{x}{(\log x)^k} \sigma(\mathcal{H})$$

Proof. By Selberg's upper bound sieve

$$\sum_{\substack{n \le x \\ \{n+h_i\} \in \mathcal{P}^k \\ n+h \in \mathcal{P}}} 1 \ll \frac{x}{(\log x)^{k+1}}.$$

Remark 2. Dickson Conjecture $\not\Longrightarrow$ Strong Dickson Conjecture.

2. Primes in Arithmetic Progressions (AP)

Conjecture (Lagrange, Waring, Erdős–Turán (1936)). *The primes contain k-term AP's for every k.*

Conjecture (Erdős–Turán (1936)). *If $A \subset \mathbb{Z}^+$ has positive upper density then A contains k-term AP's for every k.*

Roth (1953): This is true for $k = 3$.

Szemerédi (1975) This is true for every k.

3. History before 2000 (2004)

Erdős (1940) $d_n < (1 - c_0) \log n$ i.o. $c_0 > 0$ fix.

Bombieri–Davenport (1966) $d_n < (\log n)/2$ i.o.

H. Maier (1988) $d_n < (\log n)/4$ i.o.

Van der Corput (1939) $\mathcal{P}$ contains infinitely many 3-term AP's.

Heath-Brown (1984) There are infinitely many pairs n, d such that $n, n + d, n + 2d \in \mathcal{P}$, $n + 3d \in \mathcal{P}_2$.

Definition

$n = P_k$-number if it has at most k prime factors (ALMOST PRIMES).

4. Results after 2000

Green–Tao Theorem (2004–2008): *$\mathcal{P}$ contains k-term AP's for every k.*

Goldston–Pintz–Yıldırım (2005–2009): $\liminf_{n\to\infty} d_n/\log n = 0$.

GPY (2006–2010): $\liminf_{n\to\infty} d_n/(\log n)^c = 0$ if $c > \dfrac{1}{2}$.

Zhang (2013–2014): $\mathcal{H}_k$ admissible, $k > 3.5 \cdot 10^6 \implies n + \mathcal{H}_k$ contains at least 2 primes i.o.

Maynard (2013–2015): This is true for $k \geq 105$.

Polymath (2014): This is true for $k \geq 50$.

Maynard (2013–2015), Tao unpublished: $\mathcal{H}_k$ admissible
$\Longrightarrow n + \mathcal{H}_k$ contains at least $\left(\frac{1}{4} + o(1)\right) \log k$ primes i.o.

5. Patterns of primes in arithmetic progressions

A common generalization of the Green–Tao and Maynard–Tao theorem is

Theorem 1 (J. P. 2017)

Let $m > 0$ *and* $\mathcal{A} = \{a_1, \ldots, a_r\}$ *be a set of* r *distinct integers with* r *sufficiently large depending on* m*. Let* $N(\mathcal{A})$ *denote the number of integer* m*-tuples* $\{h_1, \ldots, h_m\} \subseteq \mathcal{A}$ *such that there exist for every* ℓ *infinitely many* ℓ*-term arithmetic progressions of integers* $\{n_i\}_{i=1}^{\ell}$ *where* $n_i + h_j$ *is the* j^{th} *prime following* n_i *prime for each pair* i, j*. Then*

$$N(\mathcal{A}) \gg_m \#\big\{(h_1, \ldots, h_m) \in \mathcal{A}\big\} \gg_m |A|^m = r^m. \tag{5.1}$$

Theorem 1 will follow by the application of Maynard's method from the weaker

Theorem 2 (J. P. 2017)

Let m be a positive integer, $\mathcal{H} = \{h_1, \dots, h_k\}$ be an admissible set of k distinct non-negative integers $h_i \leqslant H$, $k = \lceil Cm^2 e^{4m} \rceil$ with a sufficiently large absolute constant C. Then there exists an m-element subset

$$\{h'_1, h'_2, \dots, h'_m\} \subseteq \mathcal{H} \tag{5.2}$$

such that for every positive integer ℓ we have infinitely many ℓ-element non-trivial arithmetic progressions of integers n_i such that $n_i + h'_j \in \mathcal{P}$ for $1 \leqslant i \leqslant \ell$, $1 \leqslant j \leqslant m$, further $n_i + h'_j$ is always the j^{th} prime following n_i.

In fact we prove a stronger result, namely

Theorem 3 (J. P. 2017)

There is some C, such that for all k_0 and all $k > Ck_0^2 e^{4k_0}$ there is some $c > 0$, such that for all admissible tuples $\{h_1, \ldots, h_k\}$ the number $N(x)$ of integers $n \leq x$, such that $n + h_i$ is n^c-pseudo prime, and among these k integers there are at least k_0 primes, satisfies $N(x) \gg \frac{x}{\log^k x}$. These $N(x)$ integers $n \leq x$ contain an m-term AP if $x > C_0(m)$.

6. Structure of the proof of the Maynard–Tao theorem

(i) Key parameters: $\mathcal{H} = \{h_1, \ldots, h_k\}$ given $(0 \le h_1 < h_2 < \ldots < h_k)$

N large, we look for primes of the form $n + h_i$ with $n \in [N, 2N)$

$R = N^{\theta/2-\varepsilon}$, where θ is a level of distribution of primes:

$$\text{(6.1)} \qquad \sum_{q \leqslant x^\theta} \max_{\substack{a \\ (a,q)=1}} \left| \pi(x, q, a) - \frac{\pi(x)}{\varphi(q)} \right| \ll_A \frac{x}{(\log x)^A}$$

holds for any $A > 0$ where the $\ll$ symbol of Vinogradov means that $f(x) = O(g(x))$ is abbreviated by $f(x) \ll g(x)$.

Remark. $\theta = 1/2$ admissible: Bombieri–Vinogradov theorem (1965). $\theta > 0$ Rényi (1947)

$$\text{(6.2)} \qquad W = \prod_{p \le D_0} p \qquad D_0 = C^*(k) \text{ suitably large}$$

$$\text{(6.3)} \quad n \equiv \nu_0 (\mathrm{mod}\ W) \ \ (\nu_0 + h_i, W) = 1 \ \text{ for } i = 1, \ldots, k.$$

(ii) We weight the numbers $n \equiv \nu_0 \pmod{W}$, $n \in [N, 2N)$ by w_n, so that $w_n \geq 0$ and on average w_n would be large if we have many primes among $\{n + h_i\}_{i=1}^{k}$,

$$w_n = \Bigg(\sum_{d_i | n + h_i \ \forall i} \lambda_{d_1, \ldots, d_k} \Bigg)^2 . \tag{6.4}$$

$$\lambda_{d_1, \ldots, d_k} = \left(\prod_{i=1}^{k} \mu(d_i) d_i \right) \sum_{\substack{r_1, \ldots, r_k \\ d_i | r_i \ \forall i \\ (r_i, W) = 1}} \frac{\mu\left(\prod_{i=1}^{k} r_i \right)^2}{\prod_{i=1}^{k} \varphi(r_i)} y_{r_1, \ldots, r_k} \tag{6.5}$$

whenever $\left(\prod_{i=1}^{k} d_i, W \right) = 1$ and $\lambda_{d_1, \ldots, d_r} = 0$ otherwise.

$$(6.6) \qquad y_{r_1,\ldots,r_k} = F\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R}\right)$$

where F is piecewise differentiable, real, F and F' bounded, supported on

$$(6.7) \qquad R_k = \left\{(x_1, \ldots, x_k) \in [0,1]^k : \sum_{i=1}^{k} x_i \leqslant 1\right\}.$$

(iii) Let $\chi_{\mathcal{P}}(n)$ denote the characteristic function of $\mathcal{P}$,
(6.8)

$$S_1 := \sum_{\substack{n \\ N \leqslant n < 2N \\ n \equiv \nu_0 \pmod{W}}} w_n, \quad S_2 := \sum_{\substack{n \\ N \leqslant n < 2N \\ n \equiv \nu_0 \pmod{W}}} \left(w_n \sum_{i=1}^{k} \chi_{\mathcal{P}}(n+h_i)\right),$$

If we succeed to choose F, thereby $\lambda_{\mathbf{d}}$ and w_n in such a way that $\left(r_k \in \mathbb{Z}^+\right)$

$$(6.9) \qquad S_2 > S_1, \quad \text{or} \quad S_2 > (r_k - 1)S_1 \quad \text{resp.}$$

we obtain at lest two, or k_0 primes, resp. among $n + h_1, \ldots, n + h_k \Longrightarrow$ bounded gaps between primes or even k_0 primes in bounded intervals i.o.

(iv) First step towards this: evaluation of S_1 and S_2.

Proposition 1. *We have as $N \to \infty$*

$$S_1 = \frac{\left(1 + O\left(\frac{1}{D_0}\right)\right)\varphi(W)^k N(\log R)^k}{W^{k+1}} I_k(F), \tag{6.10}$$

$$S_2 = \frac{\left(1 + O\left(\frac{1}{D_0}\right)\right)\varphi(W)^k N(\log R)^{k+1}}{W^{k+1}} \sum_{j=1}^{k} J_k^{(j)}(F), \tag{6.11}$$

$$I_k(F) = \int_0^1 \dots \int_0^1 F(t_1, \dots, t_k)^2 dt_1 \dots dt_k, \tag{6.12}$$

(6.13)

$$J_k^{(j)}(F) = \int_0^1 \dots \int_0^1 \left(\int_0^1 F(t_1, \dots, t_k) dt_j\right)^2 dt_1 \dots dt_{j-1} dt_{j+1} \dots dt_k.$$

After this we immediately obtain

Corollary. *If the* sup *is taken with F_k as before and*

$$(6.14) \qquad M_k = \sup \frac{\sum_{j=1}^{k} J_k^{(j)}(F)}{I_k(F)}, \quad r_k = \left\lceil \frac{\theta M_k}{2} \right\rceil$$

and let $\mathcal{H}$ be a fixed admissible sequence $\mathcal{H} = \{h_1, \ldots, h_k\}$ of size k. Then there are infinitely many integers n such that at least r_k of the $n + h_i$ $(1 \leqslant i \leqslant k)$ are simultaneously primes.

Proposition 2. *$M_{105} > 4$ and $M_k > \log k - 2 \log\log k - 2$ for $k > k_0$.*

7. A stronger version of the Maynard–Tao theorem

Theorem 3 gives a stronger form in 3 aspects:

(i) all numbers $n + h_i$ are almost primes, having all prime factors greater than $n^{c(k)}$;

(ii) the number of such n's is at least $\dfrac{c'(k)N}{\log^k N}$, the true order of magnitude of $n \in [N, 2N)$ with all $n + h_i$ being n^c-almost primes;

(iii) for every ℓ we have (infinitely) many ℓ-term AP's with the same prime pattern.

Remark: properties (i) and (ii) are interesting in themselves, but crucial in many applications, in particular in showing (iii).

Let $P^{-}(n)$ denote the smallest prime factor of n.

The following Lemma shows that the contribution of n's to S_1 with at least one prime $p \mid \prod(n+h_i)$, $p < n^{c_1(k)}$ is negligible if $c_1(k)$ is suitably small $\left(R = N^{\theta/2-\varepsilon}\right)$.

Lemma 1. *We have*

$$(7.1) \qquad S_1^{-} = \sum_{\substack{N \leqslant n < 2N \\ n \equiv \nu_0 \pmod W \\ P^{-}\left(\prod_{i=1}^{k}(n+h_i)\right) < n^{c_1(k)}}} w_n \ll_{k,H} \frac{c_1(k)\log N}{\log R} S_1.$$

Corollary. *We get immediately property* (i)*. Further*

$$w_n \ll \lambda_{\max}^2 \ll y_{\max}^2 (\log R)^{2k} \ll (\log R)^{2k} \tag{7.2}$$

since $\prod (n + h_i)$ *has in this case just a bounded number of prime factors, so the sum over the divisors can be substituted by the largest term (apart from a factor depending on* k*). So we get*

$$S_1^* := \sum_{\substack{n \\ N \leqslant n < 2N \\ n \equiv \nu_0 \pmod W \\ P^-\left(\prod_{i=1}^{k}(n+h_i)\right) > n^{c_1(k)} \\ \#\{i; n+h_i \in \mathcal{P}\} \geqslant r_k}} 1 \geq \frac{S_1(1 + O(c_1(k))}{(\log R)^{2k}}, \tag{7.3}$$

by which we obtained property (ii).

8. Green–Tao theorem: structure of proof

Original Szemerédi theorem. If $\mathcal{A} \subseteq \mathbb{Z}_N$ has a positive density then $\mathcal{A}$ contains m-term AP for every m, if $N > C(m)$.

Relative Szemerédi theorem (Green–Tao). If $\mathcal{A} \subseteq \mathbb{Z}_N$ is a *pseudorandom* set, $\mathcal{B} \subseteq \mathcal{A}$ has a positive relative density within $\mathcal{A}$, i.e. with a measure $\nu(n)$ obeying the linear forms condition

$$\lim_{N\to\infty} \frac{\sum\limits_{n\le N, n\in\mathcal{B}} \nu(n)}{\sum\limits_{n\le N, n\in\mathcal{A}} \nu(n)} = \delta > 0 \tag{8.1}$$

then $\mathcal{B}$ contains m-term AP for every m, if $N > C(\delta, m)$.

Definition. A set $\mathcal{A} \subseteq \mathbb{Z}_N$ is a *pseudorandom* set if there is a measure $\nu : \mathbb{Z}_N \to R^+$ which satisfies the *linear forms condition* if the following holds.

Let (L_{ij}), $1 \le i \le \ell$, $1 \le j \le t$ rational numbers with all numerators and denominators at most L_0, $b_i \in \mathbb{Z}_N$, $\ell \le \ell_0$, $m \le m_0$. Let $\psi_i(\mathbf{x}) = \sum_{j=1}^{t} L_{i,j} x_j + b_i$, where the t-tuples $(L_{ij})_{1 \le j \le t} \in \mathbb{Q}^t$ are non-zero and no t-tuple is a rational multiple of another. Then

$$(8.2) \quad \mathbb{E}\Big(\nu(\psi_1(\mathbf{x})) \ldots \nu(\psi_m(\mathbf{x})) \mid \mathbf{x} \in \mathbb{Z}_N^t\Big) = 1 + o_{L_0,\ell_0,m_0}(1).$$

Remark 1. The primes up to N form a set of density $1/\log N \to 0$ as $N \to \infty$. Therefore we cannot use the Szemerédi theorem.

Step 1. To formulate and show a generalization of the Szemerédi theorem where the set $\mathbb{Z}_N = [1, 2, \ldots, N]$ can be substituted by some *sparse set* satisfying some regularity condition like (8.2). This result is called Relative Szemerédi Theorem.

Remark 2. Another condition, the *correlation condition* in the original work of Green and Tao could be avoided by a different proof of Conlon–Fox–Zhao (2015).

Step 2. To find a suitable *pseudo-random* set $\mathcal{A}$ where the set $\mathcal{P}$ of the primes can be embedded as a subset of positive density. This was proved in an unpublished manuscript of Goldston and Yıldırım (2003). This set $\mathcal{A}$ is the set of almost primes; the measure (μ is the Möbius function, $c > 0$ small)

$$\nu(n) = \left(\sum_{d|n, d \le R} \mu(d) \left(1 - \frac{\log d}{\log R}\right)\right)^2 \quad R = N^c. \tag{8.3}$$

9. Combination of the methods of Green–Tao and Maynard–Tao

Difficulty: the original Maynard–Tao method produces directly (without using any further ideas) only at least

$$N^{c/\log\log N} \tag{9.1}$$

integers $n \in [N, 2N)$ with at least $k_0 = \frac{1}{4}(1 + o(1)) \log k$ primes among $\{n + h_i\}_{i=s}^{k}$. The expected number of n's with this property is

$$c_2(k_0)\frac{N}{(\log N)^{k_0}}, \tag{9.2}$$

which is much more.

Hope: By Theorem 1 (cf. 7 (i)–(ii)) we obtain

$$(9.3) \qquad \frac{c_3(k)N}{\log^k N}$$

such numbers, which is still less than (9.2).

Further idea: if we require additionally that all $n + h_i$'s should be almost primes, i.e. $P^-(n + h_i) > n^{c_1(k)}$, then we obtain also $c(k)N/\log^k N$ numbers $n \in [N, 2N)$ which is already the true order of magnitude of such n's.

Solution: instead of embedding primes into the set of almost primes we embed the set of n's, $n \in [N, 2N)$ with

at least k_0 primes among $\{n + h_i\}_{i=1}^{k}$ and

$$P^{-}\left(\prod_{i=1}^{k}(n + h_i)\right) > n^{c_1(k)} \tag{9.4}$$

into the set of n's, $n \in [N, 2N)$ with (9.4).

Remark: in some sense we embed the set of almost prime k-tuples with at least k_0 primes into the set of all almost prime k-tuples.

Lemma 2. *Let k be an arbitrary positive integer and $\mathcal{H} = \{h_1, \ldots, h_k\}$ be an admissible k-tuple. If the set $\mathcal{N}(\mathcal{H})$ satisfies with constants $c_1(k)$, $c_2(k)$*

$$\mathcal{N}(\mathcal{H}) \subseteq \left\{ n; P^{-}\left(\prod_{i=1}^{k}(n+h_i)\right) \geqslant n^{c_1(k)} \right\} \tag{9.5}$$

and

$$\#\{n \leqslant X, n \in \mathcal{N}(\mathcal{H})\} \geqslant \frac{c_2(k)X}{\log^k X} \tag{9.6}$$

for $X > X_0$, then $N(\mathcal{H})$ contains ℓ-term arithmetic progressions for every ℓ.

Main idea of the proof: We use the measure

$$(9.7) \qquad \nu(n) := \begin{cases} \left(\frac{\varphi(W)}{W}\right)^k \prod_{i=1}^{k} \frac{\Lambda_R^2(Wn+\nu_0+h_i)}{\log R}, & n \in [N, 2N) \\ 0 & \text{otherwise} \end{cases}$$

with $R = N^{c_1(k)}$ and

$$(9.8) \qquad \Lambda_R(u) = \sum_{d \le R, d|u} \mu(d) \log \frac{R}{d}.$$

Remark. If $P^-\left(\prod_{i=1}^{k}(u + h_i)\right) > N^{c_1(k)} = R$, then $\Lambda_R(u + h_i) = \log R$ (the single term in the sum is that with $d = 1$) and $\nu(u) = (\varphi(W)/W)^k \log^k R$ does not depend on u.

The pseudorandomness of the measure ν can be proved by a generalization of the original Goldston–Yıldırım method. The original GY method is exactly the case $k = 1$. The possible methods are either

(i) analytic number theoretical (using the zeta-function) or

(ii) Fourier series or

(iii) real elementary.

Remark. The proof that we obtain *consecutive* primes by this procedure follows from the fact that the number of $n \in [N, 2N)$ obtained is at least $c(k)N/\log^k N$. If any of the numbers $n+h$, $0 \le h \le h_k$, $h \ne h_i$ $(i = 1, 2, \ldots, k)$ were additionally prime then by Selberg's upper bound sieve we would find at most $c'(k)N/\log^{k+1} N$ such numbers (cf. the estimate in 1.) since all $n + h_i$ $(i = 1, 2, \ldots, k)$ are almost primes (similarly to the case of the Strong HL conjecture). So here we also need both properties 7 (i) and 7 (ii).

10. Sketch of the proof of Lemma 2

The proof is essentially the same for an arbitrary k as for the simplest case $k = 1$. So let $k = 1$. We choose a prime $p < N^{c_1(k)}$ and try to evaluate

$$S_p^* = \sum_{\substack{N \le n < 2N, p \mid n+h, n \equiv \nu_0(W) \\ [d,e] \mid n+h}} \lambda_d \lambda_e \tag{10.1}$$

Distinguishing the cases

$$(10.2) \qquad p \nmid [d,e] \Longrightarrow \sim \frac{N}{pW} \frac{\lambda_d \lambda_e}{[d,e]}$$

$$(10.3) \qquad d = d'p, \quad p \nmid e \Longrightarrow \sim \frac{N}{pW} \frac{\lambda_d \lambda_e}{[d',e]} \text{ (or reversed)}$$

$$(10.4) \qquad d = d'p, \quad e = e'p \Longrightarrow \sim \frac{N}{pW} \frac{\lambda_d \lambda_e}{[d',e']}$$

we obtain in all cases an asymptotic of type

$$S_p^* = \frac{N}{pW} \sum \frac{\lambda_d \lambda_e}{[d, e, p]/p} + O(R^{2+\varepsilon}). \tag{10.5}$$

Lemma (Selberg, Coll. Works 1991, Greaves 2000)

$$T_p := \sum \frac{\lambda_d \lambda_e}{[d, e, p]/p} = \sum_{\substack{r \\ p \nmid r}} \frac{\mu^2(r)}{\varphi(r)} (y_r - y_{rp})^2. \tag{10.6}$$

However, by the definition of y_r and F we have
(10.7)

$$(y_r - y_{rp})^2 = \left(F\left(\frac{\log r}{\log R}\right) - F\left(\frac{\log r + \log p}{\log R}\right)\right)^2 \ll_F \frac{\log p}{\log R},$$

(10.8) $$T_p \ll \frac{\log p}{\log R} \cdot \log R \frac{\varphi(W)}{W} \Longrightarrow S_p^* \ll \frac{\log p}{p} \cdot \frac{N}{W} \cdot \frac{\varphi(W)}{W},$$

(10.9)

$$S^* = \sum_{p < N^{c_1(k)}} S_p^* \ll \frac{N}{W} c_1(k) \log N \cdot \frac{\varphi(W)}{W} \ll_{k,\theta} c_1(k) S_1$$

which is negligible if $c_1(k)$ is sufficiently small. ■