Primes in short intervals

Runbo Li
International Curriculum Center
The High School Affiliated to Renmin University of China

Jul 2025

Ancient number theory

One of the most important topics in analytic number theory is the distribution of prime numbers. In ancient times, people knew that there were infinitely many prime numbers. Let

$$\pi(x) = \sum_{p \leqslant x} 1.$$

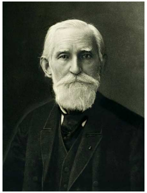
Theorem (Euclid)

$$\pi(x) \to \infty \text{ as } x \to \infty.$$

Euclid constructed a prime number of the form $p_1 p_2 \cdots p_n + 1$ and proved the above theorem by contradiction.

Chebyshev's theorem

Due to the discrete distribution of individual prime numbers, mathematicians began to focus on the distribution of prime counting function $\pi(x)$.

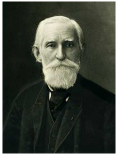
In 1845, Bertrand conjectured the following statement, which was later proved by Chebyshev in 1852.

Bertrand's postulate / Chebyshev's theorem (1852)

For any $x > 1$, there is at least one prime number between x and $2x$. That is,

$$\pi(2x) - \pi(x) > 0.$$

Prime Number Theorem

Chebyshev actually proved the following result.

$$0.92129\frac{x}{\log x} \leqslant \pi(x) \leqslant 1.10555\frac{x}{\log x} \text{ as } x \to \infty,$$

where Gauss and Legendre previously conjectured that

$$\pi(x) \sim \frac{x}{\log x} \text{ as } x \to \infty.$$

By Chebyshev's result, one can easily show that

$$\pi(2x) - \pi(x) \gg \frac{x}{\log x}.$$

Riemann Hypothesis

In 1859, Riemann connected $\pi(x)$ with the zeros of complex function $\zeta(s)$ and put forward his famous hypothesis.

Riemann Hypothesis (RH)

All non–trivial zeros of $\zeta(s)$ lie on the straight line $\mathrm{Re}(s) = \frac{1}{2}$.

As of 2025, RH is still unsolved.

Prime Number Theorem


Using ideas introduced by Riemann, Hadamard and de la Vallée Poussin proved the famous Prime Number Theorem independently in 1896.

Prime Number Theorem (PNT) (Hadamard, 1896; de la Vallée Poussin, 1896)

$$\pi(x) \sim \frac{x}{\log x} \text{ as } x \to \infty.$$

By this theorem, it is easy to prove that

$$\pi(2x) - \pi(x) \sim \frac{x}{\log x}.$$

Primes in short intervals

Can we find primes in intervals shorter than x as $x \to \infty$?

Hoheisel's theorem (1930)

There exists some $\theta < 1$ such that

$$\pi(x + x^{\theta+\varepsilon}) - \pi(x) \sim \frac{x^{\theta+\varepsilon}}{\log x}.$$

Moreover, $\theta = \frac{32999}{33000}$ is acceptable.

Primes in short intervals

Ingham's theorem (1936)

If

$$\zeta\left(\frac{1}{2}+it\right) \ll t^{c},$$

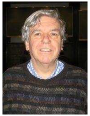
then

$$\pi(x+x^{\theta+\varepsilon})-\pi(x) \sim \frac{x^{\theta+\varepsilon}}{\log x}, \qquad \theta = \frac{1+4c}{2+4c}.$$

Moreover, $c=\frac{1}{6}$ yields

$$\pi(x+x^{\frac{5}{8}+\varepsilon})-\pi(x) \sim \frac{x^{\frac{5}{8}+\varepsilon}}{\log x}.$$

Primes in short intervals, records I

- $\frac{32999}{33000} = 0.9999$, Hoheisel, 1930;
- $\frac{249}{250} = 0.9960$, Heilbronn, 1933;
- $\frac{3}{4} = 0.7500$, Chudakov, 1936;
- $\frac{5}{8} = 0.6250$, Ingham, 1936;
- $\frac{3}{5} = 0.6000$, Montgomery, 1971;
- $\frac{7}{12} = 0.5833$, Huxley, 1972; Ivić, 1979; Heath-Brown, 1988;
- $\frac{17}{30} = 0.5667$, Guth–Maynard, 2025.

Zero-density approach

Let

$$\Lambda(n) = \begin{cases} \log p, & n = p^k, \\ 0, & \text{otherwise.} \end{cases}$$

Because

$$\sum_{n \leqslant x} \Lambda(n) = \sum_{p \leqslant x} \log p + O\left(x^{\frac{1}{2}+\varepsilon}\right),$$

we can study $\sum_{n \leqslant x} \Lambda(n)$ instead of $\pi(x)$. Note that we have

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.$$

Zero-density approach

Perron's formula

Let $a(n) = O(1)$. We have

$$\sum_{n \leqslant x} a(n) = \frac{1}{2\pi i} \int_{1+\varepsilon-i\infty}^{1+\varepsilon+i\infty} \sum_{n=1}^{\infty} \frac{a(n)}{n^s} \frac{x^s}{s} ds + \text{Error}.$$

By Perron's formula we have

$$\sum_{n=1}^{\infty} \Lambda(n) = -\frac{1}{2\pi i} \int_{1+\varepsilon-i\infty}^{1+\varepsilon+i\infty} \sum_{n=1}^{\infty} \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} ds + \text{Error}.$$

By moving the line of integration, we can get the Explicit Formula

$$\sum_{n \leqslant x} \Lambda(n) = x - \sum_{\substack{\rho=\beta+i\gamma \\ |\gamma|<T}} \frac{x^\rho}{\rho} + O\left(\frac{x(\log x)^2}{T}\right).$$

Zero-density approach

Similarly, for the short interval problem we can also get the Explicit Formula

$$\sum_{x-x^\theta<n\leqslant x}\Lambda(n)=x^\theta-\sum_{\substack{\rho=\beta+i\gamma\\|\gamma|<T}}\frac{x^\rho-\left(x-x^\theta\right)^\rho}{\rho}+O\left(\frac{x(\log x)^2}{T}\right).$$

Let $T=x^{1-\theta}(\log x)^3$ and

$$N(\sigma,T)=\#\{\text{zeros of }\zeta(\beta+i\gamma):\beta\geqslant\sigma,\ 0<\gamma\leqslant T\}.$$

Let

$$E(\sigma)=\sum_{\substack{\rho=\beta+i\gamma\\|\gamma|<T\\\sigma\leqslant\beta<\sigma+(\log x)^{-1}}}\frac{x^\rho-\left(x-x^\theta\right)^\rho}{\rho}.$$

We want to show that $E(\sigma)=o\left(x^\theta(\log x)^{-1}\right)$.

Zero-density approach

Note that

$$\frac{x^\rho - \left(x - x^\theta\right)^\rho}{\rho} = \int_{x-x^\theta}^{x} u^{\rho-1} du \ll x^\theta x^{\mathrm{Re}(\rho)-1},$$

we have

$$E(\sigma) \ll x^\theta x^{\sigma-1} N(\sigma, T).$$

Thus, by Vinogradov zero-free region and bounds of the types

$$N(\sigma, T) \ll T^{A(1-\sigma)}(\log T)^B \quad \text{or} \quad N(\sigma, T) \ll T^{A(1-\sigma)+\varepsilon},$$

we only need

$$(1-\sigma)(A(1-\theta)-1) < 0 \quad \text{or} \quad \theta > 1 - \frac{1}{A}.$$

Huxley: $A = \frac{12}{5} \implies \theta > \frac{7}{12}$. Guth–Maynard: $A = \frac{30}{13} \implies \theta > \frac{17}{30}$.

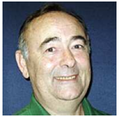
Primes in short intervals

One can get shorter intervals if we don't require an asymptotic formula. Using sieve methods, Iwaniec and Jutila got in 1979 that


Theorem (Iwaniec–Jutila, 1979)

$$\pi(x + x^{\frac{13}{23}+\varepsilon}) - \pi(x) \gg \frac{x^{\frac{13}{23}+\varepsilon}}{\log x}.$$

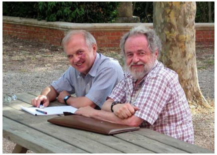
Primes in short intervals, records II

- $\frac{13}{23} = 0.5652$, Iwaniec–Jutila, 1979;
- $\frac{5}{9} = 0.5556$, Iwaniec–Jutila, 1979;
- $\frac{11}{20} = 0.5500$, Heath-Brown–Iwaniec, 1979;
- $\frac{17}{31} = 0.5484$, Pintz, 1981; Iwaniec (Unpublished);
- $\frac{23}{42} = 0.5476$, Iwaniec–Pintz, 1984;
- $\frac{1051}{1920} = 0.5474$, Mozzochi, 1986;
- $\frac{35}{64} = 0.5469$, Lou–Yao (Unpublished), 1985;
- $\frac{6}{11} = 0.5455$, Lou–Yao, 1992;
- $\frac{7}{13} = 0.5385$, Lou–Yao, 1992;
- $\frac{107}{200} = 0.5350$, Baker–Harman, 1996;
- $\frac{21}{40} = 0.5250$, Baker–Harman–Pintz, 2001;
- $\frac{13}{25} = 0.5200$, L. (preprint), 2025.

Primes in short intervals: New proofs

Without using too many deep results, Motohashi and Friedlander and Iwaniec gave simplified proofs of the existence of primes in short intervals.

Theorem (Motohashi, 1983)

We have

$$\pi(x + x^{0.56}) - \pi(x) \gg x^{0.56}(\log x)^{-1}.$$

Theorem (Friedlander–Iwaniec, 2010)

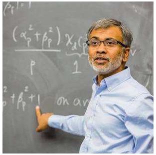
We have

$$\pi(x + x^{0.58}) - \pi(x) \gg x^{0.58}(\log x)^{-1}.$$

Primes in short intervals: New proofs

In 2019, Granville, Harper and Soundararajan gave a new proof of Hoheisel's theorem with an asymptotic formula.





Theorem (Granville–Harper–Soundararajan, 2019)

For some $\delta > 0$, we have

$$\pi(x + x^{1-\delta}) - \pi(x) \sim x^{1-\delta}(\log x)^{-1}.$$

Primes in short intervals: New proofs

In 2024, Matomäki, Merikoski and Teräväinen gave a pure elementary proof of Hoheisel's theorem.

Theorem (Matomäki–Merikoski–Teräväinen, 2024)

For some $\delta > 0$, we have

$$\pi(x + x^{\frac{39}{40}}) - \pi(x) \gg x^{\frac{39}{40}} (\log x)^{-1}.$$

Sieve approach

Let

$$\mathcal{A} = \{a : x - x^{\theta} < a \leqslant x\}, \quad \mathcal{A}_d = \{a : ad \in \mathcal{A}\}, \quad S(\mathcal{A}, z) = \sum_{\substack{a \in \mathcal{A} \\ (a, P(z)) = 1}} 1.$$

Then by a simple observation, we can find that

$$\pi(x) - \pi(x - x^{\theta}) = S\left(\mathcal{A}, x^{\frac{1}{2}}\right).$$

We have another useful tool:

Buchstab's identity

For any $w \leqslant z$, we have

$$S(\mathcal{A}, z) = S(\mathcal{A}, w) - \sum_{w \leqslant p < z} S(\mathcal{A}_p, p).$$

Sieve approach

Iwaniec and Jutila used the following decomposition:

Sieve decomposition (Iwaniec–Jutila $\frac{13}{23}$, Motohashi 0.56)

For some $v \geqslant u \geqslant 2$, we have

$$S\left(\mathcal{A}, x^{\frac{1}{2}}\right) = S\left(\mathcal{A}, x^{\frac{1}{v}}\right) - \sum_{x^{\frac{1}{v}} \leqslant p < x^{\frac{1}{u}}} S\left(\mathcal{A}_p, p\right) - \sum_{x^{\frac{1}{u}} \leqslant p < x^{\frac{1}{2}}} S\left(\mathcal{A}_p, p\right).$$

They also used two important devices: weighted zero-density estimate and mean values of Dirichlet polynomials.

Weighted zero-density estimate

Let

$$M(s) = \sum_{m \sim M} a_m m^{-s}, \quad N(s) = \sum_{n \sim N} b_n n^{-s}, \quad R(s) = \sum_{r \sim R} c_r r^{-s}, \quad K(s) = \sum_{k \sim K} k^{-s},$$

where a_m, b_n and c_r are divisor–bounded. We want to get estimates of the type

Weighted zero-density estimate

$$\sum_{\substack{\rho = \beta + i\gamma \\ \beta \geqslant \sigma, \ |\gamma| < T}} |M(\rho) N(\rho)| \ll x^{1-\sigma} (\log x)^c.$$

Note that by a variant of the Explicit formula above, this type of estimates lead to an asymptotic formula for sums of the form

$$\sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} \left(\pi\left(\frac{x}{p_1 \cdots p_n} \right) - \pi\left(\frac{x - x^\theta}{p_1 \cdots p_n} \right) \right).$$

Mean values of Dirichlet polynomials

Using Iwaniec's linear sieve, one need to estimate the "error term"

$$\sum_{\substack{m\sim M\\ n\sim N}} a_m b_n \left(\left[\frac{x}{mn} \right] - \left[\frac{x-x^{\theta}}{mn} \right] - \frac{x^{\theta}}{mn} \right)$$

in order to bound sums like

$$S\left(\mathcal{A}, z\right) \quad \text{and} \quad \sum_{p\sim P} S\left(\mathcal{A}_p, z\right).$$

This can be estimated by using classical mean and large value results of Dirichlet polynomials and power moments of zeta function.

Sieve approach

In 1979, Heath-Brown and Iwaniec used another sieve decomposition together with the above tools to obtain $\frac{11}{20}$.

Sieve decomposition (Heath-Brown–Iwaniec $\frac{11}{20}$, Pintz $\frac{17}{31}$)

For some $z^{\frac{1}{2}} \leqslant D \leqslant z^4$, we have

$$S\left(\mathcal{A}, x^{\frac{1}{2}}\right) = \left(S(\mathcal{A}, z) - \sum_{\left(\frac{D}{p_1}\right)^{\frac{1}{3}} \leqslant p_2 < p_1 < z} S\left(\mathcal{A}_{p_1 p_2}, p_2\right) \right) + \sum_{\left(\frac{D}{p_1}\right)^{\frac{1}{3}} \leqslant p_2 < p_1 < z} S\left(\mathcal{A}_{p_1 p_2}, p_2\right)$$

$$- \sum_{z \leqslant p_1 < D^{\frac{1}{2}}} S\left(\mathcal{A}_{p_1}, p_1\right) - \sum_{D^{\frac{1}{2}} \leqslant p_1 < x^{\frac{1}{2}}} S\left(\mathcal{A}_{p_1}, \left(\frac{D}{p_1}\right)^{\frac{1}{3}}\right) + \sum_{\substack{D^{\frac{1}{2}} \leqslant p_1 < x^{\frac{1}{2}} \\ \left(\frac{D}{p_1}\right)^{\frac{1}{3}} \leqslant p_2 < p_1}} S\left(\mathcal{A}_{p_1 p_2}, p_2\right).$$

Sieve approach

In their work ($\frac{11}{20}$), Heath-Brown and Iwaniec only used the fourth power moment of zeta function. Pintz ($\frac{17}{31}$) inserted a deep result of Deshouillers and Iwaniec:

Deshouillers–Iwaniec's Theorem (1982)

We have

$$\int_{T_0}^{T} \left| M\left(\frac{1}{2}+it\right)^2 K\left(\frac{1}{2}+it\right)^4 \right| \ll T^{1+\varepsilon} + M^2 T^{\frac{1}{2}+\varepsilon} + M^{\frac{5}{4}} \left(T \min\left(K, \frac{T}{K}\right)\right)^{\frac{1}{2}}.$$

This can be seen as an approximation of the sixth power moment of zeta function.

Sieve approach

Using another delicate sieve decomposition, Iwaniec and Pintz in 1984 got $\frac{23}{42}$.

Sieve decomposition (Iwaniec–Pintz $\frac{23}{42}$, Mozzochi $\frac{1051}{1920}$)

For $\frac{1051}{1920} < \theta \leqslant \frac{23}{42}$, we have

$$S\left(\mathcal{A}, x^{\frac{1}{2}}\right) \geqslant \left(S\left(\mathcal{A}, x^{7-12\theta}\right) - \sum_{\substack{p_2 < p_1 < x^{7-12\theta} \\ p_1 p_2^3 \geqslant x^{\frac{12\theta-2}{5}}}} S\left(\mathcal{A}_{p_1 p_2}, p_2\right) \right) + \sum_{\substack{p_2 < p_1 \\ x^{\frac{8-8\theta}{5}} < p_1 p_2^2 < x^{\frac{13\theta-3}{5}}}} S\left(\mathcal{A}_{p_1 p_2}, p_2\right) - \sum_{x^{7-12\theta} \leqslant p_1 < x^{\frac{6\theta-1}{5}}} S\left(\mathcal{A}_{p_1}, p_1\right)$$

Sieve approach

Sieve decomposition (Iwaniec–Pintz $\frac{23}{42}$, Mozzochi $\frac{1051}{1920}$)

$$
\begin{aligned}
&-\sum_{x^{\frac{6\theta-1}{5}}\leqslant p_1<x^{\frac{8\theta-1}{7}}} S\left(\mathcal{A}_{p_1}, \min\left(\frac{x^{\frac{4\theta+1}{5}}}{p_1}, x^{\frac{20\theta-9}{11}}\right)\right) - \sum_{x^{\frac{8\theta-1}{7}}\leqslant p_1<x^{\frac{1}{2}}} S\left(\mathcal{A}_{p_1}, \left(\frac{x^{\frac{12\theta-2}{5}}}{p_1}\right)^{\frac{1}{3}}\right)\\
&+\sum_{\substack{x^{\frac{6\theta-1}{5}}\leqslant p_1<x^{\frac{1}{2}}\\ p_1p_2\leqslant x^{\frac{3\theta+2}{5}}\\ p_1p_2^{-1}\leqslant x^{\frac{8\theta-3}{5}}}} S\left(\mathcal{A}_{p_1p_2}, p_2\right).
\end{aligned}
$$

Vaughan's identity

While working on the Bombieri–Vinogradov theorem, Vaughan introduced a finite approximation to $-\frac{\zeta'(s)}{\zeta(s)}$. Note that

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)} = F(s) - \zeta(s)F(s)G(s) - \zeta'(s)G(s)$$
$$+ \left(-\frac{\zeta'(s)}{\zeta(s)} - F(s)\right)(1 - \zeta(s)G(s)),$$

$$F(s) = \sum_{m \leqslant U} \Lambda(n)n^{-s}, \quad G(s) = \sum_{d \leqslant V} \mu(d)d^{-s}$$

and all functions of the form n^{-s} are linearly independent, we have the following

Vaughan's identity

Vaughan's identity

$$\Lambda(n) = a_1(n) + a_2(n) + a_3(n) + a_4(n),$$

where

$$a_1(n) = \begin{cases} \Lambda(n), & n \leqslant U, \\ 0, & n > U, \end{cases} \qquad a_2(n) = -\sum_{\substack{mdr=n \\ m\leqslant U,\ d\leqslant V}} \Lambda(n)\mu(d),$$

$$a_3(n) = \sum_{\substack{dh\leqslant n \\ d\leqslant V}} \mu(d) \log h, \qquad a_4(n) = \sum_{\substack{mk=n \\ m>U,\ k>1}} \Lambda(m) \sum_{\substack{d|k \\ d\leqslant V}} \mu(d).$$

Vaughan's identity

This identity helps us break $\sum_{n\sim N}\Lambda(n)f(n)$ into sums (taking $U = V = x^{\beta}$ for some $0 < \beta < \frac{1}{2}$)

$$\sum_{\substack{m\leqslant M\\ mn\leqslant x}} a_m f(mn), \quad M \leqslant \max(x^{1-\beta}, x^{2\beta})$$

and

$$\sum_{\substack{m\sim K\\ mn\leqslant x}} a_m b_n f(mn), \quad x^{\beta} \leqslant K \leqslant x^{1-\beta}.$$

Heath-Brown's identity

In 1982, Heath-Brown produced what he called a generalized Vaughan identity by using the following formula, which is valid for all $k \in \mathbb{N}$ and any function $M(s)$:

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{1 \leqslant j \leqslant k} (-1)^{j-1} \binom{k}{j} \zeta(s)^{j-1} \zeta'(s) M(s)^j + \zeta(s)^{-1} (1 - \zeta(s)M(s))^k \zeta'(s).$$

Heath-Brown used this to give another proof of Huxley's $\frac{7}{12}$ with an asymptotic formula. Let

$$M(s) = \sum_{m \leqslant M} \mu(m) m^{-s},$$

this implies an identity

Heath-Brown's identity

Heath-Brown's identity

$$\Lambda(n) = \sum_{1 \leqslant j \leqslant k} (-1)^{j-1} \binom{k}{j} a_j(n),$$

where

$$a_j(n) = \sum_{\substack{n = r_1 \cdots r_{2j} \\ i > j \Rightarrow r_i \leqslant x^{\frac{1}{k}}}} (\log r_1) \mu(r_{j+1}) \cdots \mu(r_{2j}).$$

One can use Heath-Brown's identity to construct several identities that do not follow from Vaughan's identity.

Heath-Brown's identity

Heath-Brown's identity

Suppose that $u \leqslant N^{\frac{1}{10}}$, then

$$\sum_{n \sim N} \Lambda(n) f(n)$$

can be written as $\ll (\log x)^5$ sums of the forms

$$\sum_{\substack{m \leqslant M \\ n \sim N}} a_m f(mn), \quad M \ll Nu$$

and

$$\sum_{\substack{m \sim M \\ n \sim N}} a_m b_n f(mn), \quad u^2 \leqslant M \ll N^{\frac{1}{3}}.$$

Heath-Brown's identity

Heath-Brown's identity has the advantage that more flexible sums are produced. However, the disadvantage persists that if one makes a problem harder, the method collapses. There is no "grey area" between an asymptotic formula and no result at all. Heath-Brown produced another identity that can be applied to remove this disadvantage.

Heath-Brown–Linnik identity

For $z > x^{\frac{1}{k}}$, we have

$$S\left(\mathcal{A}, x^{\frac{1}{2}}\right) = \sum_{1 \leqslant j \leqslant k} \frac{(-1)^{j-1}}{j} S\left(\mathcal{A}^k, z\right) + O\left(x^{\frac{1}{2}}\right),$$

where $\mathcal{A}^k = \{n_1 \cdots n_k \in \mathcal{A}\}$.

In 1988, he used this identity with $k = 7$ to prove $\frac{7}{12} - \varepsilon$ with an asymptotic formula.

Heath-Brown's identity

Heath-Brown (unpublished) used this identity with $k = 7$ to prove

$$0.99\frac{x^{\frac{11}{20}+\varepsilon}}{\log x} \leqslant \pi(x) - \pi(x - x^{\frac{11}{20}+\varepsilon}) \leqslant 1.01\frac{x^{\frac{11}{20}+\varepsilon}}{\log x}.$$

Lou and Yao (1992) used this identity with $k = 7$ to prove

$$0.969\frac{x^{\frac{6}{11}+\varepsilon}}{\log x} \leqslant \pi(x) - \pi(x - x^{\frac{6}{11}+\varepsilon}) \leqslant 1.031\frac{x^{\frac{6}{11}+\varepsilon}}{\log x}$$

and

$$\pi(x) - \pi(x - x^{\frac{7}{13}+\varepsilon}) \gg \frac{x^{\frac{7}{13}+\varepsilon}}{\log x}.$$

Weighted zero-density estimate

In 1996, Baker and Harman used a stronger version of the weighted zero-density estimate:

Weighted zero-density estimate, stronger version

$$\sum_{\substack{\rho=\beta+i\gamma \\ |\gamma|<T}} x^{\beta-1} \left|M(\rho)N(\rho)\right| \ll (\log x)^{-A}.$$

Using this estimate and a truncated Perron's formula, they got

$$\sum_{\substack{mnr\in\mathcal{A} \\ m\sim M \\ n\sim N}} a_m b_n \Lambda(r) - x^{\theta} \sum_{\substack{mnr\in\mathcal{A} \\ m\sim M \\ n\sim N}} \frac{a_m b_n}{mn} \leqslant x^{\theta} \sum_{\substack{\rho=\beta+i\gamma \\ 0\leqslant\beta\leqslant 1 \\ |\gamma|<T}} x^{\beta-1} \left|M(\rho)N(\rho)\right| + O\left(x^{\theta-\varepsilon}\right) \ll x^{\theta}(\log x)^{-A}.$$

Sieve approach

By using a truncated Perron's formula to remove the dependencies between variables, they obtained an asymptotic formula for sums of the form

$$\sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} S\left(\mathcal{A}_{p_1 \cdots p_n}, p_n\right).$$

Sieve approach

The most important observation of Baker and Harman is that we can use Buchstab's identity in this way:

$$\sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} S\left(\mathcal{A}_{p_1 \cdots p_n}, z\right) = \sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} S\left(\mathcal{A}_{p_1 \cdots p_n}, x^{\varepsilon}\right) - \sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n \\ x^{\varepsilon} \leqslant p_{i+1} < z}} S\left(\mathcal{A}_{p_1 \cdots p_{n+1}}, p_{n+1}\right).$$

The estimate of the first sum on the right-hand side using Iwaniec's linear sieve is asymptotic. This means that if we can find $z = x^{\delta}$ with $\delta > 0$ as large as possible such that the second sum on the right-hand side has an asymptotic formula, then we can obtain an asymptotic formula for the sum on the left-hand side. This estimate is better than the bounds we get using only Iwaniec's linear sieve.

Sieve approach

Suppose that we want to give a lower bound for $\sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} S\left(\mathcal{A}_{p_1 \cdots p_n}, z\right)$.

Using Iwaniec's linear sieve directly, we only have

$$\sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} S\left(\mathcal{A}_{p_1 \cdots p_n}, z\right) \geqslant (1 + o(1)) \frac{x^\theta}{\log x} e^{-\gamma} f(u).$$

Using the above procedure, we can get

$$\sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} S\left(\mathcal{A}_{p_1 \cdots p_n}, z\right) = (1 + o(1)) \frac{x^\theta}{\log x} \omega(u).$$

Note that

$$\omega(u) = \frac{e^{-\gamma}(f(u) + F(u))}{2}.$$

Sieve approach

In 2001, Baker, Harman and Pintz (BHP) developed a new method of estimating $\sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} S\left(\mathcal{A}_{p_1 \cdots p_n}, p_n\right)$. They used mean value results of Dirichlet polynomials instead of weighted zero-density estimates. Specifically, they proved that

Theorem (BHP, 2001)

If

$$\int_{T_0}^{T} \left| M\left(\frac{1}{2} + it\right) \right| \ll x^{\frac{1}{2}} (\log x)^{-A},$$

then

$$\sum_{m \in \mathcal{A}} a_m = (1 + o(1)) \frac{x^\theta}{X} \sum_{x - X < m \leqslant x} a_m,$$

where $X = x \exp\left(-3 \log x\right)^{\frac{1}{3}}$.

Sieve approach

Using the above Theorem, one can easily show the two relations between mean value results and asymptotic formulas:

$$\int_{T_0}^{T} |MNR| \ll x^{\frac{1}{2}} (\log x)^{-A} \implies \sum_{\substack{mnr \in \mathcal{A} \\ m \sim M \\ n \sim N}} a_m b_n c_r \tag{A}$$

and

$$\int_{T_0}^{T} |MNK| \ll x^{\frac{1}{2}} (\log x)^{-A} \implies \sum_{\substack{m \sim M \\ n \sim N}} a_m b_n S\left(\mathcal{A}_{mn}, \exp\left(\frac{\log x}{\log\log x}\right)\right). \tag{B}$$

Thus, one only need to find longer ranges of M and N such that (A) or (B) holds.

Sieve approach

BHP used Watt's Theorem together with Hölder's inequality to get more type (B) estimates.

Watt's Theorem (1995)

We have

$$\int_{T_0}^{T} \left| M\left(\frac{1}{2}+it\right)^2 K\left(\frac{1}{2}+it\right)^4 \right| \ll T^{1+\varepsilon} + M^2 T^{\frac{1}{2}+\varepsilon}.$$

Watt's Theorem improves Deshouillers–Iwaniec's Theorem.

Sieve approach

$$\sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} S\left(\mathcal{A}_{p_1 \cdots p_n}, z\right) = \sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n}} S\left(\mathcal{A}_{p_1 \cdots p_n}, \exp\left(\frac{\log x}{\log\log x}\right)\right) - \sum_{\substack{p_i \sim P_i \\ 1 \leqslant i \leqslant n \\ p_{i+1} < z}} S\left(\mathcal{A}_{p_1 \cdots p_{n+1}}, p_{n+1}\right),$$

$$\begin{aligned}
S\left(\mathcal{A}, x^{\frac{1}{2}}\right) &= S(\mathcal{A}, z) - \sum_{z \leqslant p_1 < x^{\frac{1}{2}}} S\left(\mathcal{A}_{p_1}, z\right) + \sum_{z \leqslant p_2 < p_1 < x^{\frac{1}{2}}} S\left(\mathcal{A}_{p_1 p_2}, p_2\right) \\
&= S(\mathcal{A}, z) - \sum_{z \leqslant p_1 < x^{\frac{1}{2}}} S\left(\mathcal{A}_{p_1}, z\right) + \sum_{z \leqslant p_2 < p_1 < x^{\frac{1}{2}}} S\left(\mathcal{A}_{p_1 p_2}, z\right) \\
&\quad - \sum_{z \leqslant p_3 < p_2 < p_1 < x^{\frac{1}{2}}} S\left(\mathcal{A}_{p_1 p_2 p_3}, z\right) + \sum_{z \leqslant p_4 < p_3 < p_2 < p_1 < x^{\frac{1}{2}}} S\left(\mathcal{A}_{p_1 p_2 p_3 p_4}, p_4\right) \\
&= \cdots
\end{aligned}$$

Sieve approach

- BHP (1996);

$$0.9953\frac{x^{\frac{11}{20}+\varepsilon}}{\log x} \leqslant \pi(x) - \pi(x - x^{\frac{11}{20}+\varepsilon}) \leqslant 1.0001\frac{x^{\frac{11}{20}+\varepsilon}}{\log x}.$$

- BHP (2001), 0.525;
- L. (preprint, 2025), 0.52;
  1. More type (A) estimates with 5 or more variables (the lower bound of the sum of all variables decreases as the number of variables increases);
  2. An optimized sieve argument (one can get 0.523 with BHP's original argument).
- Possible refinements:
  1. More type (B) estimates obtained by Hölder's inequality and higher power means of zeta function
$$\int_{T_0}^{T} \left|K^A\right| \ll T^{1+\frac{A-4}{8}+\varepsilon} \text{ for } 4 \leqslant A \leqslant 12;$$
  2. Careful discussions on asymptotic regions and calculations.

Legendre's conjecture

Legendre's conjecture

For any $x > 1$, there is at least one prime number between x^2 and $(x+1)^2$.

Legendre's conjecture ($x \to \infty$)

We have

$$\pi(x + x^{\frac{1}{2}}) - \pi(x) > 0.$$

Cramér's conjecture

Cramér's conjecture (1937)

The interval

$$[x, x + f(x) \log^2 x]$$

contains primes for some $f(x) \to 1$ as $x \to \infty$.

Lindelöf Hypothesis

Lindelöf Hypothesis (LH)

For any $\varepsilon > 0$, we have
$$\zeta\left(\frac{1}{2} + it\right) \ll t^{\varepsilon}.$$

Clearly, we have RH $\Rightarrow$ LH.

Primes in short intervals

Under RH

We have

$$\pi(x + x^{\frac{1}{2}} \log x) - \pi(x) \gg x^{\frac{1}{2}}.$$

Under LH

We have

$$\pi(x + x^{\frac{1}{2}+\varepsilon}) - \pi(x) \sim x^{\frac{1}{2}+\varepsilon}(\log x)^{-1}.$$

Unconditional

We have

$$\pi(x + x^{\frac{17}{30}+\varepsilon}) - \pi(x) \sim x^{\frac{17}{30}+\varepsilon}(\log x)^{-1},$$

$$\pi(x + x^{0.5195}) - \pi(x) \gg x^{0.5195}(\log x)^{-1}.$$

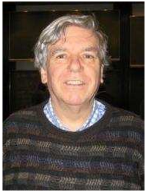
Primes in almost all short intervals

In 1943, Selberg obtained the following two results.

Theorem (Selberg, 1943)

1. Under RH, Cramér's interval contains primes for almost all x if $f(x) \to \infty$ as $x \to \infty$.
2. The interval
$$[x, x + x^{\frac{19}{77}+\varepsilon}]$$
contains $\sim \frac{x^{\frac{19}{77}+\varepsilon}}{\log x}$ primes for almost all x.

Primes in almost all short intervals, records I

- $\frac{19}{77} = 0.2468$, Selberg, 1943; ($\theta_1 = 2\theta_0 - 1$)
- $\frac{1}{5} = 0.2000$, Montgomery, 1971; ($\frac{3}{5} \Leftrightarrow \frac{1}{5}$)
- $\frac{1}{6} = 0.1667$, Huxley, 1972; ($\frac{7}{12} \Leftrightarrow \frac{1}{6}$)
- $\frac{1}{7.5} = 0.1333$, Guth–Maynard, 2025. ($\frac{17}{30} \Leftrightarrow \frac{1}{7.5}$)

Primes in almost all short intervals

Using sieve methods, Harman got in 1982 that

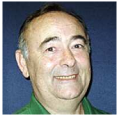
Theorem (Harman, 1982)

The interval

$$[x, x + x^{\frac{1}{10}+\varepsilon}]$$

contains $\gg \frac{x^{\frac{1}{10}+\varepsilon}}{\log x}$ primes for almost all x.

Primes in almost all short intervals, records II

- $\frac{1}{10} = 0.1000$, Harman, 1982;
- $\frac{14}{159} = 0.0881$, Lou–Yao (Unpublished), 1985;
- $\frac{1}{12} = 0.0833$, Harman, 1983; Heath-Brown, 1984;
- $\frac{1}{13} = 0.0769$, Jia, 1995;
- $\frac{17}{227} = 0.0749$, Lou–Yao (Unpublished), 1985;
- $\frac{1}{13.5} = 0.0740$, H. Li, 1995;
- $\frac{1}{14} = 0.0714$, Jia, 1995; Watt, 1995;
- $\frac{1}{15} = 0.0667$, H. Li, 1997;
- $\frac{1}{16} = 0.0625$, Baker–Harman–Pintz, 1997;
- $\frac{1}{18} = 0.0556$, Wong, 1996; Jia, 1996; Harman, 2007;
- $\frac{1}{20} = 0.0500$, Jia, 1996;
- $\frac{1}{21.5} = 0.0476$, L. (preprint), 2024.

A weak Legendre's conjecture

Legendre's conjecture

The interval

$$[x, x + x^{\frac{1}{2}}]$$

contains an integer with a prime factor larger than x^{1}.

Conjecture $LPF(\theta)$

The interval

$$[x, x + x^{\frac{1}{2}}]$$

contains an integer with a prime factor larger than x^{θ} for some $\theta > 0$.

Largest prime factors of integers in short intervals

In 1969, Ramachandra got the first result in this direction.

Theorem (Ramachandra, 1969)

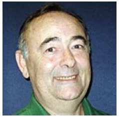
The interval

$$[x, x + x^{\frac{1}{2}}]$$

contains an integer with a prime factor larger than $x^{0.576}$.

Largest prime factors of integers in short intervals



- 0.576, Ramachandra, 1969;
- 0.625, Ramachandra, 1970;
- 0.662, Graham, 1981;
- 0.675225, Zhu, 1987;
- 0.692, Jia, 1986;
- 0.7, Baker, 1986;
- 0.71, Jia, 1989;
- 0.723, Jia, 1993; H.-Q. Liu, 1993;
- 0.728, Jia, 1996;
- 0.732, Baker–Harman, 1995;
- 0.738, H.-Q. Liu–Wu, 1999;
- 0.74, Harman, 2007;
- 0.7428, Baker–Harman, 2009;
- 0.7437, L. (preprint), 2025.

Largest prime factors of integers in short intervals

If we increase the interval length to $x^{\frac{1}{2}+\varepsilon}$, then we can get better results since many powerful analytic tools, such as the estimation of Dirichlet polynomials, can be used. In 1973, Jutila obtained

Theorem (Jutila, 1973)

The interval

$$[x, x + x^{\frac{1}{2}+\varepsilon}]$$

contains an integer with a prime factor larger than $x^{\frac{2}{3}-\varepsilon}$.

Largest prime factors of integers in short intervals

- $\frac{2}{3} = 0.6666$, Jutila, 1973;
- $\frac{73}{100} = 0.7300$, Balog, 1980;
- $\frac{193}{250} = 0.7720$, Balog, 1984;
- $\frac{41}{50} = 0.8200$, Balog–Harman–Pintz, 1983;
- $\frac{11}{12} = 0.9166$, Heath-Brown, 1996;
- $\frac{17}{18} = 0.9444$, Heath-Brown–Jia, 1998;
- $\frac{19}{20} = 0.9500$, Harman, 2007;
- $\frac{24}{25} = 0.9600$, Haugland, 1998;
- $\frac{25}{26} = 0.9615$, Jia–M.-C. Liu, 2000;
- $\frac{51}{53} = 0.9622$, L. (preprint), 2024.

Largest prime factors of integers in short intervals

In 1983, Balog, Harman and Pintz proved a result with "medium" interval lengths.

Theorem (Balog–Harman–Pintz, 1983)

The interval

$$[x, x + x^{\frac{1}{2}}(\log x)^A]$$

contains an integer with a prime factor larger than $x^{0.712-\varepsilon}$.

Largest prime factors of integers in short intervals

- 0.7120, Balog–Harman–Pintz, 1983;
- $\frac{5}{6} = 0.8333$, Lou, 1984;
- $\frac{18}{19} = 0.9473$, Merikoski, 2021; ($A < 1.39$)
- $\frac{37}{39} = 0.9487$, L. (Unpublished); ($A < 1.39$)

Almost–primes in short intervals

Instead of considering the size of prime factors, one can also consider the number of prime factors. We define the "Almost–primes" P_r and E_r as

Definition (Almost–primes)

An integer n is a P_r if n has at most r prime factors counted with multiplicity.
An integer n is an E_r if n has exactly r prime factors counted with multiplicity.

Of course, short–interval results for P_r are easier to obtain than corresponding results for E_r.

Almost–primes in short intervals

Theorem (Brun, 1920)

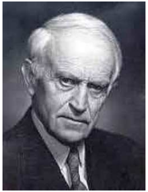
The interval $[x, x + x^{\frac{1}{2}}]$ contains a P_{11}. $LPF(\frac{1}{11})$ is true.

Theorem (Wang, 1957)

The interval $[x, x + x^{\frac{1}{2}}]$ contains a P_3. $LPF(\frac{1}{3})$ is true.

Theorem (Wang, 1959)

The interval $[x, x + x^{\frac{10}{17}}]$ contains a P_2.
The interval $[x, x + x^{\frac{20}{49}}]$ contains a P_3.



Almost–primes in short intervals

- $\frac{10}{17} = 0.5882$, Wang, 1959;
- $\frac{14}{25} = 0.5600$, Jurkat–Richert, 1965;
- $\frac{6}{11} = 0.5454$, Richert, 1969;
- $\frac{1}{2} = 0.5000$ ($LPF(\frac{1}{2})$ is true), Chen, 1975;
- 0.4856, Laborde, 1978;
- 0.4770, Chen, 1979;
- 0.4550, Halberstam–Heath-Brown–Richert, 1981;
- 0.4500, Iwaniec–Laborde, 1981;
- 0.4476, Halberstam–Richert, 1985;
- $\frac{63}{142} = 0.4436$, Fouvry, 1990;
- 0.4400, Wu, 1992;
- 0.4382, H. Li, 1994;
- 0.4378, Cao, 1995;
- 0.4360, H.-Q. Liu, 1996;
- 0.43596, Sargos–Wu, 2000;
- $\frac{101}{232} = 0.43535$, Wu, 2010.

Almost–primes in short intervals


Theorem (Matomäki–Teräväinen, 2023)

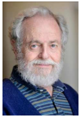
The interval $[x, x + x^{\frac{1}{2}}(\log x)^{1.55}]$ contains an E_3.

Almost–primes in almost all short intervals




Author	Form	Length	Year
Wolke	E_2	$(\log x)^{5000000}$	1979
Harman	E_2	$(\log x)^{7+\varepsilon}$	1979
Bourgain	E_2	$(\log x)^{6.86}$	2000
Teräväinen	E_2	$(\log x)^{3.51+\varepsilon}$	2016
Matomäki–Teräväinen	E_2	$(\log x)^{2.1+\varepsilon}$	2023

Almost–primes in almost all short intervals

Author	Form	Length	Year
Heath-Brown	P_2	$x^{\frac{1}{11}}$	1978
Heath-Brown	P_3	$(\log x)^{35+\varepsilon}$	1978
Friedlander	P_4	$(\log x)^5$	1982
Motohashi	P_2	x^{ε}	Unpublished
Mikawa	P_2	$h(x)(\log x)^5$	1989
Matomäki	P_2	$h(x) \log x$	2022
Teräväinen	E_3	$(\log x)(\log \log x)^{6+\varepsilon}$	2016
Teräväinen	E_k	$(\log x)(\log_{k-1} x)^{C_k+\varepsilon}$	2016

Mean square gap between primes

In 1943, Selberg proved the following result under RH.

Theorem (Selberg, 1943)

Under RH, we have

$$\sum_{p_n \leqslant x} (p_{n+1} - p_n)^2 \ll x(\log x)^3.$$

Mean square gap between primes

In 1978, Heath-Brown obtained a weaker bound of Selberg's mean square gap unconditionally.

Theorem (Heath-Brown, 1978)

We have

$$\sum_{p_n \leqslant x} (p_{n+1} - p_n)^2 \ll x^{\frac{4}{3}+\varepsilon}.$$

Mean square gap between primes

- 1 (on RH), Selberg, 1943;
- $\frac{4}{3} = 1.3333$, Heath-Brown, 1978;
- $\frac{1413}{1067} = 1.3242$, Heath-Brown, 1979;
- $\frac{7}{6} = 1.1666$ (on LH), Heath-Brown, 1979;
- $\frac{23}{18} = 1.2777$, Heath-Brown, 1979;
- 1 (on LH), Yu, 1996;
- $\frac{5}{4} = 1.25$, Peck, 1996; Maynard, 2012;
- $\frac{123}{100} = 1.23$, Stadlmann, 2022.

Large differences between primes

In the same paper, Selberg also considered a variant of the mean square gap.

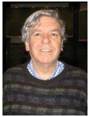
Theorem (Selberg, 1943)

Under RH, we have

$$\sum_{\substack{p_n \leqslant x \\ p_{n+1}-p_n \geqslant x^{\frac{1}{2}+\varepsilon}}} (p_{n+1}-p_n) \ll x^{\frac{1}{2}+\varepsilon}.$$

We call this $LD(\frac{1}{2}+\varepsilon, \frac{1}{2})$.

Large differences between primes

- $LD(\frac{1}{2}+\varepsilon, \frac{1}{2})$ (on RH), Selberg, 1943;
- $LD(\frac{1}{2}, \frac{29}{30}=0.9666)$, Wolke, 1975;
- $LD(\frac{1}{2}+\varepsilon, \frac{85}{98}=0.8673)$, Cook, 1979;
- $LD(\frac{1}{2}+\varepsilon, \frac{1759}{2134}=0.8242)$, Huxley, 1980;
- $LD(\frac{1}{2}+\varepsilon, \frac{3}{4})$ (on LH), Huxley, 1980;
- $LD(\frac{1}{2}, \frac{215}{266}=0.8082)$, Ivić, 1979;
- $LD(\frac{1}{2}, \frac{3}{4})$, Heath-Brown, 1979;
- $LD(\frac{1}{2}+\varepsilon, \frac{5}{8})$, Heath-Brown, 1979;
- $LD(\frac{1}{2}, \frac{25}{36}=0.6944)$, Peck, 1998;
- $LD(\frac{1}{2}, \frac{2}{3})$, Matomäki, 2007;
- $LD(\frac{1}{2}-\delta, \frac{2}{3}+5\delta)$, Islam, 2015 $(0 \leqslant \delta \leqslant \frac{1}{6}\sqrt{327}-3=0.01385)$;
- $LD(\frac{1}{2}, \frac{3}{5})$, Heath-Brown, 2021;
- $LD(\frac{1}{2}, 0.57)$, Järviniemi, 2022;
- $LD(0.45, 0.63)$, Järviniemi, 2022;

Primes in short intervals: Explicit results

Primes in short intervals: Explicit version 1

For all $x \geqslant N_0$, we have

$$\pi(x + x^{\theta}) - \pi(x) > 0.$$

Author	θ	N_0	**Year**
Caldwell–Cheng	$\frac{2}{3}$	1 (on RH)	2005
Dudek	$\frac{2}{3}$	$\exp(\exp(33.217))$	2014
Mattner	$\frac{2}{3}$	$\exp(\exp(33.1981))$	2017
Cully-Hugill	$\frac{2}{3}$	$\exp(\exp(32.892))$	2021
Mossinghoff–Trudgian–Yang	$\frac{2}{3}$	$\exp(\exp(32.76))$	2024
Cully-Hugill	$\frac{2}{3}$	$\exp(\exp(32.537))$	2023

Primes in short intervals: Explicit results

Author	θ	N_0	Year
Caldwell–Cheng	$\frac{2}{3}$	1 (on RH)	2005
Dudek	$1 - \frac{1}{5\cdot 10^9}$	1	2014
Mattner	$1 - \frac{1}{1.5\cdot 10^6}$	1	2017
Cully-Hugill	$1 - \frac{1}{296}$	1	2021
Cully-Hugill	$1 - \frac{1}{180}$	1	2021
Cully-Hugill	$1 - \frac{1}{155}$	1	2023
Dudek–Johnston	$\frac{1}{2}$ (P_4)	1	2025

Primes in short intervals: Explicit results

Legendre's conjecture

We have
$$\pi(x + x^{\frac{1}{2}}) - \pi(x) > 0.$$

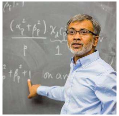
Primes in short intervals: Under RH / LH

We have
$$\pi(x + x^{\frac{1}{2}+\varepsilon}) - \pi(x) \sim x^{\frac{1}{2}+\varepsilon}(\log x)^{-1}.$$

Primes in short intervals: Explicit version 2 (Under RH)

For all $x \geqslant N_0$, we have
$$\pi(x + cx^{\frac{1}{2}} \log x) - \pi(x) > 0.$$

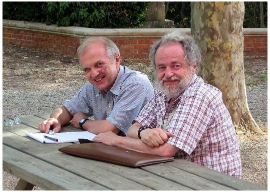
Primes in short intervals: Explicit results

Author	c	N_0	**Year**
von Koch	$c_0 \log x$	$< \infty$ (on RH)	1901
Schoenfeld	$\frac{1}{4\pi} \log x$	599 (on RH)	1976
Cramér	$< \infty$	sufficiently large (on RH)	1920
Goldston	5	sufficiently large (on RH)	1983
Ramaré–Saouter	$\frac{8}{5} = 1.6$	2 (on RH)	2003
Dudek	$\frac{4}{\pi} = 1.2732$	2 (on RH)	2015
Dudek–Grenié–Molteni	1.2204	2 (on RH)	2016
Dudek–Grenié–Molteni	$1 + \frac{4}{\log x}$	2 (on RH)	2016
Carneiro–Milinovich–Soundararajan	$\frac{22}{25} = 0.88$	4 (on RH)	2019

Exceptional characters and primes in short intervals

In 2001, Friedlander and Iwaniec first proved an asymptotic formula for the number of primes in intervals shorter than $x^{\frac{1}{2}}$ under the existence of exceptional characters.



Theorem (Friedlander–Iwaniec, 2001)

We have

$$\pi(x + x^{\frac{39}{79}}) - \pi(x) \gg x^{\frac{39}{79}} (\log x)^{-1} \left(1 + O\left(L(1,\chi)(\log x)^A\right)\right).$$

In 2024, L. (preprint) improved the exponent $\frac{39}{79} = 0.4937$ to 0.4923.

Upper bounds

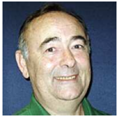
In 1973, Montgomery and Vaughan considered the upper bounds for the number of primes in short intervals.

Theorem (Montgomery–Vaughan, 1973)

For any $0 < \theta < 1$, we have

$$\pi(x + x^\theta) - \pi(x) \leqslant \frac{2}{\theta} \frac{x^\theta}{\log x}.$$

Upper bounds

- $\frac{2}{\theta}$ $(0 < \theta < 1)$, Montgomery–Vaughan, 1973;
- $\frac{18}{15\theta-2}$ $(\frac{1}{3} < \theta < 1)$, Iwaniec, 1982;
- $\frac{4}{\theta+1}$ $(\frac{1}{2} < \theta < 1)$, Iwaniec, 1982;
- $\frac{22}{100\theta-45}$ $(\frac{6}{11} < \theta < \frac{11}{20})$, Lou–Yao, 1989;
- 1.031 $(\frac{6}{11} < \theta < 1)$, Lou–Yao, 1992;
- 1.0001 $(\frac{11}{20} < \theta < 1)$, Baker–Harman–Pintz, 1997;
- 1 $(\frac{17}{30} < \theta < 1)$, Guth–Maynard, 2025.

Upper bounds

Theorem (L. (preprint), 2025)

For any $0.52 < \theta \leqslant 0.535$, we have

$$\pi(x + x^\theta) - \pi(x) \leqslant C(\theta) \frac{x^\theta}{\log x},$$

where

$$C(\theta) \leqslant \begin{cases} 2.7626, & 0.52 < \theta \leqslant 0.521, \\ 2.6484, & 0.521 < \theta \leqslant 0.522, \\ 2.5630, & 0.522 < \theta \leqslant 0.523, \\ 2.4597, & 0.523 < \theta \leqslant 0.524, \\ 2.3759, & 0.524 < \theta \leqslant 0.535. \end{cases}$$

Exceptional sets in PNT in short intervals

We also want to know how frequently an asymptotic formula in PNT in short intervals "does not hold".

Definition ($E(\theta)$)

For any $0 < \theta < 1$, let $E(\theta)$ denote the least exponent such that

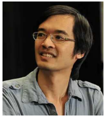
$$\pi(x + x^{\theta}) - \pi(x) \sim x^{\theta}(\log x)^{-1}$$

holds for all $x \in [X, 2X]$ except for a set of measure $O(X^{E(\theta)+\varepsilon})$.

Note that we have the following simple relations:

$$E(\theta) = -\infty,\ \theta > \frac{17}{30}; \quad E(\theta) \geqslant 0,\ \theta \leqslant \frac{17}{30}; \quad E(\theta) < 1,\ \theta \geqslant \frac{1}{21.5}.$$

Exceptional sets in PNT in short intervals

- $E(\theta) \leqslant 1 - \theta$ for $0 < \theta \leqslant \frac{1}{2}$ (on RH), Bazzanella–Perelli, 2000;
- $E(\theta) \leqslant \frac{3(1-\theta)}{2}$ for $\frac{1}{2} < \theta \leqslant \frac{11}{21}$, Bazzanella, 2000;
- $E(\theta) \leqslant \frac{47-42\theta}{35}$ for $\frac{11}{21} < \theta \leqslant \frac{23}{42}$, Bazzanella, 2000;
- $E(\theta) \leqslant \frac{36\theta^2-96\theta+55}{39-36\theta}$ for $\frac{23}{42} < \theta \leqslant \frac{17}{30}$, Bazzanella, 2000;
- $E(\frac{1}{2}) \leqslant \frac{3}{5}$, Heath-Brown, 2021;
- various bounds for $E(\theta)$, Gafni–Tao, 2025.

Bounded gaps between primes

Let

$$H_m = \liminf_{n\to\infty}(p_{n+m} - p_n).$$

Then we have the following bounds:

- $H_1 \leqslant 246$, Polymath8b, 2014;
- $H_2 \leqslant 396504$, Stadlmann, 2025;
- $H_3 \leqslant 24407016$, Stadlmann, 2025;
- $H_4 \leqslant 1391051532$, Stadlmann, 2025;
- $H_5 \leqslant 77510685234$, Stadlmann, 2025;
- $H_m \ll e^{3.8075m}$, Stadlmann, 2025.

Thank you!