Recent progress on twin prime conjecture (Algebraic Number Theory and Related Topics 2016)

Sono, Keiju

RIMS Kôkyûroku Bessatsu
B77 (2020), 003–028

Recent progress on twin prime conjecture

By

Keiju Sono*

Abstract

In this article, we mainly see several recent results on twin prime conjecture since the works of Goldston, Pintz and Yıldırım (GPY). Among others we focus on the GPY sieve, Zhang's breakthrough and the Maynard-Tao method. We briefly describe the outlines of these new methods.

§ 1. Introduction

It is widely believed that there exist infinitely many prime numbers p for which $p+2$ is also a prime, and this conjecture is called the twin prime conjecture. About one hundred years ago, Hardy and Littlewood [14] generalized this problem and gave some quantitative observation, called the Hardy-Littlewood conjecture. Let $\mathcal{H} = \{h_1, \cdots, h_k\}$ be a set of exactly k distinct non-negative integers. Then, their conjecture asserts that the number of n's below N such that all of $n + h_1, \cdots, n + h_k$ are primes will be asymptotically equal to

$$\frac{N}{\log^k N} \prod_p \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right) \left(1 - \frac{1}{p}\right)^{-k},$$

provided that $\nu_p(\mathcal{H}) < p$ for all primes p, where $\nu_p(\mathcal{H})$ denotes the number of residue classes mod p covered by $\mathcal{H}$ (the set $\mathcal{H}$ satisfying this condition is called admissible). The twin prime conjecture is the case that $k = 2$, $\mathcal{H} = \{0, 2\}$. It might be of some worth

Received March 17, 2017. Revised March 3, 2018.
2010 Mathematics Subject Classification(s): 11N05, 11N35, 11N36
Key Words: *twin prime conjecture, GPY sieve, Maynard-Tao method, bounded gaps between primes.*
Supported by JSPS KAKENHI Grant Number 16K17571
*Ehime University, 790-8577, Dogo-Himata, Matsuyama, Japan.
e-mail: `sono.keiju.jk@ehime-u.ac.jp`

to mention that if $\mathcal{H}$ is not admissible, the set $\{n+h_1, \cdots, n+h_k\}$ contains at least one composite number whenever n is sufficiently large. The Hardy-Littlewood conjecture seems to be far from our reach. However, a number of remarkable progresses toward this conjecture have been established. For example, in 1973, Chen [3] proved that there are infinitely many prime numbers p for which $p+2$ is either a prime number or a product of two prime numbers. In 1983 Heath-Brown [18] proved that the existence of Siegel zeros implies the existence of infinitely many twin primes.

The aim of this article is to give some overview of the results obtained in the 21st century, since the works of Goldston, Pintz, and Yıldırım. Section 2 is devoted to describe their works. Throughout this paper we denote by p_n the n-th prime. The prime number theorem indicates that $p_{n+1} - p_n$ is asymptotically equal to $\log p_n$ on average. This fact implies

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\log p_n} \leq 1.$$

In 2009, Goldston, Pintz and Yıldılım [9] proved

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\log p_n} = 0. \tag{1}$$

The sieve they used in their argument is called the GPY sieve. Moreover, they proved

$$\liminf_{n\to\infty} (p_{n+1} - p_n) < \infty, \tag{2}$$

provided that primes have the level of distribution θ for some $1/2 < \theta < 1$. Regrettably the above assumption is still unproved. However, it is known that the Bombieri-Vinogradov theorem assures that this is valid for $\theta \leq 1/2$. The case $\theta = 1$ is called the Elliott-Halberstam conjecture (EH). Since then they improved their methods and obtained several better results on gaps between primes (see [10], [11], [12]). Their best unconditional result on gaps between consecutive primes is that

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\sqrt{\log p_n}(\log\log p_n)^2} < \infty. \tag{3}$$

We only describe briefly the outline of the proof of (1). There are very nice surveys on GPY sieve, by Hasanalizade [16] and Soundararajan [23]. The author recommends to read these surveys if one wishes to understand completely the argument how the GPY sieve works.

In Section 3, we see the work of Zhang [24]. He found a nice way to avoid some fundamental difficulties resulting from the incompleteness of our understanding for the equidistribution of primes and accomplished to prove so called the bounded gaps between primes. Concretely, he unconditionally proved

$$\liminf_{n\to\infty} (p_{n+1} - p_n) < 7 \times 10^7. \tag{4}$$

This estimate implies that there exist infinitely many consecutive primes for which the gap is at most 7×10^7. His upper bound 7×10^7 has been improved by several experts day by day. In particular, in the paper of DHJ Polymath [20], the right hand side of (4) was replaced by 4680.

In Section 4, we see the work of Maynard [19]. Only a few months after the breakthrough of Zhang, he considered a very nice refinement of the GPY sieve and proved

$$\liminf_{n\to\infty}(p_{n+1} - p_n) \le 600. \tag{5}$$

It is mentioned in his paper that Terence Tao considered a similar method at the same time. They also proved the existence of the bounded gaps between m-consecutive primes for any integer $m \ge 2$ (the gap is dependent only on m). Compared with Zhang's proof, their method, called the Maynard-Tao method, is relatively quite simple, and it is useful to extend to other situations, for example, the distribution of prime ideals in algebraic fields. Section 4 is devoted to describe the sieve considered by Maynard and explain shortly how he obtained the great improvement on small gaps between primes.

§ 2. The works of Goldston, Pintz, Yıldırım

§ 2.1. The main results

As we mentioned in the introduction, we denote the n-th prime number by p_n and for natural number ν, put

$$\Delta_\nu = \liminf_{n\to\infty} \frac{p_{n+\nu} - p_n}{\log p_n}.$$

The prime number theorem means that p_n is asymptotically equal to $n \log n$, and this fact yields the "trivial bound" $\Delta_1 \le 1$. If the twin prime conjecture is true, it follows that $\Delta_1 = 0$, because the twin prime conjecture is equivalent to

$$\liminf_{n\to\infty}(p_{n+1} - p_n) = 2.$$

About one hundred years many mathematicians tried to obtain better upper bound for Δ_1. The detailed history is introduced in [9]. Let us pick up some of them. In 1926, Hardy and Littlewood [15] obtained the first conditional result of the type $\Delta_1 < 1$. Assuming the generalized Riemann hypothesis (GRH) for Dirichlet L-functions, they proved $\Delta_1 \le 2/3$. Later Rankin [22] improved their result and obtained $\Delta_1 \le 3/5$, assuming the GRH. Erdös first obtained an unconditional proof for $\Delta_1 < 1$ in 1940, using Brun's sieve. In 1965, Bombieri and Davenport [2] proved unconditionally that $\Delta_1 \le 1/2$. Their proof depends on the Bombieri-Vinogradov theorem instead of the GRH. The key idea is to consider how the prime numbers are equidistributed in each arithmetic progression. Goldston and Yıldırım [13] improved this and obtained $\Delta_1 \leq 1/4$. Finally, Goldston, Pintz and Yıldırım [9], using so called the GPY sieve, succeeded in proving the following result.

Theorem 2.1 (Goldston, Pintz, Yıldırım ([9], 2009)). *We have*

$$\Delta_1 = \liminf_{n\to\infty} \frac{p_{n+1}-p_n}{\log p_n} = 0. \tag{6}$$

They also obtained a conditional upper bound for $\liminf_{n\to\infty}(p_{n+1}-p_n)$ and a sharp upper bound for Δ_ν $(\nu \geq 2)$ in the same paper. In order to describe the statements, we need to prepare several definitions. Put

$$\theta(n) = \begin{cases} \log n & (n : \text{prime}) \\ 0 & (\text{otherwise}) \end{cases}$$

and consider the sum

$$\Theta(N; q, a) = \sum_{\substack{n \leq N \\ n \equiv a (\bmod q)}} \theta(n).$$

For any fixed q, a such that $(a, q) = 1$, the prime number theorem in arithmetic progressions asserts

$$\Theta(N; q, a) = \frac{N}{\varphi(q)}\left(1 + O_A\left(\frac{1}{\log^A N}\right)\right)$$

holds for any $A > 0$, and the generalized Riemann hypothesis yields

$$\Theta(N; q, a) = \frac{N}{\varphi(q)} + O(N^{\frac{1}{2}}(\log N)^2).$$

We say that primes have level of distribution θ if for any $\epsilon > 0$, the estimate

$$\sum_{q \leq N^{\theta-\epsilon}} \max_{(a,q)=1} \left|\Theta(N; q, a) - \frac{N}{\varphi(q)}\right| \ll_A \frac{N}{\log^A N} \quad (N \to \infty) \tag{7}$$

holds for any $A > 0$. The Bombieri-Vinogradov theorem states that primes have level of distribution 1/2. The Elliott-Halberstam conjecture predicts that primes will have level of distribution 1. However, this conjecture seems to be far from our reach. In fact, currently it has not been proved that primes have level of distribution θ for some $\theta > 1/2$. Next, let $\mathcal{H} = \{h_1, \cdots, h_k\}$ be a set of exactly k distinct non-negative integers. We denote by $\nu_p(\mathcal{H})$ the number of residue classes mod p covered by $\mathcal{H}$. We say that $\mathcal{H}$ is an admissible set if $\nu_p(\mathcal{H}) < p$ holds for any prime p. For example, $\mathcal{H}_1 = \{0, 2\}$ and $\mathcal{H}_2 = \{0, 2, 6\}$ are admissible. For any admissible set $\mathcal{H} = \{h_1, \cdots, h_k\}$, we say that the tuple of k numbers $n + h_1, \cdots, n + h_k$ $(n \in \mathbf{N})$ is an admissible k-tuple. If we assume a hypothesis slightly stronger than the Bombieri-Vinogradov theorem, the GPY sieve yields a very strong consequence on gaps between primes.

Theorem 2.2 (Goldston, Pintz, Yıldırım ([9], 2009)). *Suppose that primes have level of distribution $\theta > 1/2$. Then there exist an explicitly calculable constant $C(\theta)$ depending only on θ such that any admissible k-tuple with $k \geq C(\theta)$ contains at least two primes infinitely often. Specially, if $\theta \geq 0.971$, then this is true for $k \geq 6$.*

The above theorem implies that if primes have level of distribution $\theta > 1/2$, then for any admissible set $\mathcal{H} = \{h_1, \cdots, h_k\}$, there are infinitely many natural numbers n such that at least two of $n+h_1, \cdots, n+h_k$ are prime numbers. This yields a conditional upper bound

$$\liminf_{n\to\infty}(p_{n+1} - p_n) \leq \max_{1\leq i<j\leq k} |h_i - h_j|. \tag{8}$$

Among others, since the set $\{0, 4, 6, 10, 12, 16\}$ is admissible, if we assume the Elliott-Halberstam conjecture, we have

$$\liminf_{n\to\infty}(p_{n+1} - p_n) \leq 16.$$

§ 2.2. The basic strategy of GPY

For any set $\mathcal{H} = \{h_1, \cdots, h_k\}$ consisting of distinct k integers, put

$$\mathfrak{S}(\mathcal{H}) := \prod_p \left(1 - \frac{1}{p}\right)^{-k} \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right), \tag{9}$$

where $\nu_p(\mathcal{H})$ denotes the number of the elements of $\mathcal{H}$ modulo p. We note that $\mathfrak{S}(\mathcal{H}) \neq 0$ if and only if $\mathcal{H}$ is admissible. Let $\Lambda(n)$ be the von Mangoldt function defined by $\Lambda(n) = \log p$ if $n = p^m$ (p: prime, $m \in \mathbf{N}$) and otherwise $\Lambda(n) = 0$. A natural approach to the twin prime conjecture is to consider the sum of the function $\Lambda(n)\Lambda(n+2)$. More generally, put

$$\Lambda(n, \mathcal{H}) := \Lambda(n + h_1) \cdots \Lambda(n + h_k). \tag{10}$$

The Hardy-Littlewood conjecture is stated in the form

$$\sum_{n\leq N} \Lambda(n, \mathcal{H}) = N(\mathfrak{S}(\mathcal{H}) + o(1)), \quad N \to \infty. \tag{11}$$

On the other hand, by Möbius inversion formula, the von Mangoldt function is expressed by

$$\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d}. \tag{12}$$

Therefore, $\Lambda(n)$ is approximated by

$$\Lambda_R(n) := \sum_{\substack{d|n \\ d\le R}} \mu(d) \log \frac{R}{d}. \tag{13}$$

Thus $\Lambda(n,\mathcal{H})$ is approximated by

$$\Lambda_R(n+h_1)\cdots\Lambda_R(n+h_k). \tag{14}$$

In fact, the upper bound $\Delta_1 \le 1/4$ in [13] is obtained by investigating the asymptotic behavior of the sum of the above function. We need to detect n for which $\{n+h_1,\cdots,n+h_k\}$ contains at least two primes. One candidate for the tool is the k-th von Mangoldt function defined by

$$\Lambda_{(k)}(n) = \sum_{d|n} \mu(d) \left(\log \frac{n}{d}\right)^k. \tag{15}$$

This function vanishes when n has more than k distinct prime factors. The prime tuple detecting function is given by

$$\Lambda_{(k)}(n,\mathcal{H}) = \frac{1}{k!}\Lambda_{(k)}(P_{\mathcal{H}}(n)), \tag{16}$$

where

$$P_{\mathcal{H}}(n) = \prod_{i=1}^{k}(n+h_i). \tag{17}$$

The function $\Lambda_{(k)}(n,\mathcal{H})$ is approximated by

$$\Lambda_R(n;\mathcal{H}) = \frac{1}{k!} \sum_{\substack{d|P_{\mathcal{H}}(n) \\ d\le R}} \mu(d) \left(\log \frac{R}{d}\right)^k. \tag{18}$$

This function seems to be nice to detect the integer n satisfying the above condition. However, it is mentioned in [9] that this $\Lambda_R(n;\mathcal{H})$ is not adequate to prove $\Delta_1 = 0$. Instead Goldston, Pintz and Yıldırım employed a slightly modified function defined by

$$\Lambda_R(n;\mathcal{H},l) = \frac{1}{(k+l)!} \sum_{\substack{d|P_{\mathcal{H}}(n) \\ d\le R}} \mu(d) \left(\log \frac{R}{d}\right)^{k+l}, \tag{19}$$

where l is some integer satisfying $0 \le l \le k$. To prove Theorem 2.2, they considered the sum

$$\mathcal{S} := \sum_{N<n\le 2N} \left(\sum_{i=1}^{k} \theta(n+h_i) - \log(3N)\right) \Lambda_R(n;\mathcal{H},l)^2. \tag{20}$$

If $\mathcal{S} \to \infty$ as $N \to \infty$, then for any sufficiently large N, the set $\{n+h_1, \cdots, n+h_k\}$ contains at least two primes for some $N < n \le 2N$, and this implies (8). In order to investigate the asymptotic behavior of the right hand side of (20), we need to know two summations $\sum_{N<n\le 2N} \theta(n+h_i)\Lambda_R(n;\mathcal{H},l)^2$ $(h_i \in \mathcal{H})$, $\sum_{N<n\le 2N} \Lambda_R(n;\mathcal{H},l)^2$. The following propositions are the most important parts in [9]. Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are sets of k_1 and k_2 distinct non-negative integers $\le h$, respectively. Let $M := k_1+k_2+l_1+l_2$. Then,

Proposition 2.3 ([9], Proposition 1). *Let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$, $|\mathcal{H}_i| = k_i$ and $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$. If $R \ll N^{\frac{1}{2}}(\log N)^{-4M}$ and $h \le R^C$ for any given constant $C > 0$, then as $R, N \to \infty$, we have*

$$\sum_{n\le N} \Lambda_R(n;\mathcal{H}_1,l_1)\Lambda_R(n;\mathcal{H}_2,l_2) = \binom{l_1+l_2}{l_1}\frac{(\log R)^{r+l_1+l_2}}{(r+l_1+l_2)!}(\mathfrak{S}(\mathcal{H})+o_M(1))N. \tag{21}$$

Proposition 2.4 ([9], Proposition 2). *Let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$, $|\mathcal{H}_i| = k_i$ and $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$, $1 \le h_0 \le h$, and $\mathcal{H}^0 = \mathcal{H} \cup \{h_0\}$. If $R \ll_M N^{\frac{1}{4}}(\log N)^{-B(M)}$ for a sufficiently large positive constant $B(M)$, and $h \le R$, then as $R, N \to \infty$, we have*

$$\begin{aligned}
&\sum_{n\le N} \theta(n+h_0)\Lambda_R(n;\mathcal{H}_1,l_1)\Lambda_R(n;\mathcal{H}_2,l_2) \\
&= \begin{cases}
\binom{l_1+l_2}{l_1}\frac{(\log R)^{r+l_1+l_2}}{(r+l_1+l_2)!}(\mathfrak{S}(\mathcal{H}^0)+o_M(1))N & (h_0 \notin \mathcal{H}) \\
\binom{l_1+l_2+1}{l_1+1}\frac{(\log R)^{r+l_1+l_2+1}}{(r+l_1+l_2+1)!}(\mathfrak{S}(\mathcal{H})+o_M(1))N & (h_0 \in \mathcal{H}_1\backslash\mathcal{H}_2) \\
\binom{l_1+l_2+2}{l_1+1}\frac{(\log R)^{r+l_1+l_2+1}}{(r+l_1+l_2+1)!}(\mathfrak{S}(\mathcal{H})+o_M(1))N & (h_0 \in \mathcal{H}_1\cap\mathcal{H}_2).
\end{cases}
\end{aligned} \tag{22}$$

If primes have level of distribution $\theta > 1/2$, the asymptotics in (22) hold with $R \ll N^{\frac{\theta}{2}-\epsilon}$ and $h \le R^{\epsilon}$ for any fixed $\epsilon > 0$.

By putting $R = N^{\frac{\theta}{2}-\epsilon}$, $l_1 = l_2 = l$ in the above propositions, for any admissible set $\mathcal{H} = \{h_1, \cdots, h_k\}$ consisting of k elements, we have

$$\begin{aligned}
\mathcal{S} &= \sum_{N<n\le 2N}\left(\sum_{i=1}^{k}\theta(n+h_i) - \log(3N)\right)\Lambda_R(n;\mathcal{H},l)^2 \\
&\sim \left\{\frac{k}{k+2l+1}\frac{2l+1}{l+1}(\theta-2\epsilon)\log N - \log 3N\right\} \\
&\qquad \times \frac{1}{(k+2l)!}\binom{2l}{l}\mathfrak{S}(\mathcal{H})N(\log R)^{k+2l}.
\end{aligned} \tag{23}$$

The leading coefficient of (23) becomes positive for any sufficiently large N when

$$\frac{k}{k+2l+1}\frac{2l+1}{l+1}\theta > 1,$$

and this condition is satisfied if $\theta > 1/2$ and $k, l \to \infty$ with $l = o(k)$. In this way the conclusion of the first part of Theorem 2.2 is obtained. To prove the last part of Theorem 2.2, they considered the sum

$$\mathcal{S}' := \sum_{N<n\le 2N}\left(\sum_{i=1}^{k}\theta(n+h_i) - \log 3N\right)\left(\sum_{l=0}^{L} a_l \Lambda_R(n;\mathcal{H},l)\right)^2. \tag{24}$$

Expanding the right hand side of (24) and applying Propositions 2.3, 2.4, we obtain

$$S^* = S^*(N,\mathcal{H},\theta,\mathbf{b}) := \frac{1}{\mathfrak{S}(\mathcal{H})N(\log R)^{k+1}}\mathcal{S}' \sim {}^t\mathbf{b}\mathbf{M}\mathbf{b}, \tag{25}$$

where $\mathbf{b}$ is the column vector ${}^t(b_0, b_1, \cdots, b_L)$ with $b_l = (\log R)^l a_l$, and $\mathbf{M}$ is the square matrix of size $L+1$ given by

$$\mathbf{M} = \left[\binom{i+j}{i}\frac{1}{(k+i+j)!}\left(\frac{k(i+j+2)(i+j+1)}{(i+1)(j+1)(k+i+j+1)} - \frac{2}{\theta}\right)\right]_{0\le i,j\le L}.$$

We need to choose $\mathbf{b}$ so that $S^* > 0$ for a given θ and minimal k. Let λ be an eigenvalue of the matrix $\mathbf{M}$ and $\mathbf{b}$ be an eigenvector of $\mathbf{M}$ corresponding to λ. Then,

$$S^* \sim {}^t\mathbf{b}\lambda\mathbf{b} = \lambda\sum_{i=0}^{L}|b_i|^2.$$

Therefore, if the matrix $\mathbf{M}$ has a positive eigenvalue λ, we have $S^* > 0$, by choosing $\mathbf{b}$ as an eigenvector of $\mathbf{M}$ corresponding to λ. Using *Mathematica*, they investigated when the matrix $\mathbf{M}$ has a positive eigenvector. In particular, by taking $k = 6$, $L = 1$, $b_0 = 1$ and $b_1 = (9\theta - 8)/(2(1-\theta))$, we have

$$S^* \sim -\frac{15\theta^2 - 64\theta + 48}{8!\theta(1-\theta)}.$$

This means that any admissible 6-tuple $n + h_1, \cdots, n + h_6$ contains at least two primes infinitely often if

$$\theta > \frac{4(8-\sqrt{19})}{15} = 0.97096\cdots.$$

This proves the last part of Theorem 2.2.

To prove Theorem 2.1, they considered the sum

$$\tilde{S} := \sum_{N<n\le 2N}\left(\sum_{1\le h_0\le h}\theta(n+h_0) - \nu\log 3N\right)\sum_{\substack{1\le h_1,h_2,\cdots,h_k\le h\\ \text{distinct}}}\Lambda_R(n;\mathcal{H},l)^2, \tag{26}$$

where ν is a positive integer. Using Propositions 2.3, 2.4 with $R = N^{\frac{\theta}{2}-\epsilon}$ and Gallagher's result

$$\sum_{\substack{1 \le h_1, h_2, \cdots, h_k \le h \\ \text{distinct}}} \mathfrak{S}(\mathcal{H}) \sim h^k$$

(see [7]), we find that

$$\tilde{\mathcal{S}} \sim \left\{ \frac{2k}{k+2l+1} \frac{2l+1}{l+1} \left(\frac{\theta}{2} - \epsilon \right) \log N + h - \nu \log 3N \right\} \times \frac{1}{(k+2l)!} \binom{2l}{l} N h^k (\log R)^{k+2l}. \tag{27}$$

This means that some interval $(n, n+h]$ ($N < n \le 2N$) contains at least $\nu + 1$ primes if

$$h > \left(\nu - \frac{2k}{k+2l+1} \frac{2l+1}{l+1} \left(\frac{\theta}{2} - \epsilon \right) \right) \log N.$$

Letting $l = [\sqrt{k}/2]$ and taking k sufficiently large, the above condition becomes

$$h > \left(\nu - 2\theta + 4\epsilon + O\left(\frac{1}{\sqrt{k}} \right) \right) \log N. \tag{28}$$

By letting $\nu = 1$, $\theta = 1/2$, the conclusion of Theorem 2.1 is obtained.

Remark. We also obtain an upper bound

$$\Delta_\nu \le \max\{\nu - 2\theta, 0\}$$

for any $\nu \ge 1$. Hence unconditionally we have $\Delta_\nu \le \nu - 1$, and if the Elliott-Halberstam conjecture is true, we have

$$\Delta_2 := \liminf_{n \to \infty} \frac{p_{n+2} - p_n}{\log p_n} = 0.$$

In [9], Goldston, Pintz, Yıldırım obtained a sharper upper bound

$$\Delta_\nu \le (\sqrt{\nu} - \sqrt{2\theta})^2 \tag{29}$$

for $\nu \ge 2$.

§ 2.3. The outline of the proof of Proposition 2.4

In this subsection we briefly describe the outline of the proof of Proposition 2.4. The proof of Proposition 2.3 is easier, so we omit to describe this. Let

$$\tilde{\mathcal{S}}_R(N; \mathcal{H}_1, \mathcal{H}_2, l_1, l_2, h_0) := \sum_{n=1}^{N} \Lambda_R(n; \mathcal{H}_1, l_1) \Lambda_R(n; \mathcal{H}_2, l_2) \theta(n + h_0). \tag{30}$$

By inserting the definition of Λ_R and expanding the right hand side, we obtain

$$\begin{aligned}\tilde{\mathcal{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,l_1,l_2,h_0) \\ = \frac{1}{(k_1+l_1)!(k_2+l_2)!}\sum_{d,e\le R}\mu(d)\mu(e)\left(\log\frac{R}{d}\right)^{k_1+l_1}\left(\log\frac{R}{e}\right)^{k_2+l_2} \\ \times \sum_{\substack{1\le n\le N\\ d|P_{\mathcal{H}_1}(n),e|P_{\mathcal{H}_2}(n)}}\theta(n+h_0).\end{aligned} \tag{31}$$

Next, write $d=a_1a_{12}$, $e=a_2a_{12}$, where $(d,e)=a_{12}$, so that a_1, a_2, and a_{12} are pairwise relatively prime. Then, the innermost sum becomes

$$\begin{aligned}\sum_{\substack{1\le n\le N\\ d|P_{\mathcal{H}_1}(n),e|P_{\mathcal{H}_2}(n)}}\theta(n+h_0) = \nu^*_{a_1}(\mathcal{H}_1^0)\nu^*_{a_2}(\mathcal{H}_2^0)\overline{\nu}^*_{a_{12}}((\mathcal{H}_1\bar{\cap}\mathcal{H}_2)^0)\frac{N}{\varphi(a_1a_2a_{12})} \\ + O(d_k(a_1a_2a_{12})(\max_b|E(N;a_1a_2a_{12},b)|+h\log N)),\end{aligned} \tag{32}$$

where $\nu^*_{a_i}(\mathcal{H}_i)$ $(i=1,2)$ and $\overline{\nu}^*_{a_{12}}((\mathcal{H}_1\bar{\cap}\mathcal{H}_2)^0)$ are some constants depending only on h_0, a_1, a_2, a_{12} and admissible sets $\mathcal{H}_1$, $\mathcal{H}_2$, and $E(\cdot;\cdot,\cdot)$ in the error term is defined by

$$E(x;q,a) = \sum_{\substack{p\le x\\ p\equiv a(\mathrm{mod}\ q)}}\log p - \frac{x}{\varphi(q)}$$

for $(a,q)=1$. The contribution of the error term in (32) to (31) is bounded by $O(N(\log N)^{-A})$, using our assumption that primes have level of distribution θ. On the other hand, the contribution of the main term of (32) to (31) is

$$\begin{aligned}&\frac{N}{(k_1+l_1)!(k_2+l_2)!} \\ &\times\sum_{a_1a_{12}\le R,a_2a_{12}\le R}\frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2\nu^*_{a_1}(\mathcal{H}_1^0)\nu^*_{a_2}(\mathcal{H}_2^0)\overline{\nu}^*_{a_{12}}((\mathcal{H}_1\bar{\cap}\mathcal{H}_2)^0)}{\varphi(a_1a_2a_{12})} \\ &\times\left(\log\frac{R}{a_1a_{12}}\right)^{k_1+l_1}\left(\log\frac{R}{a_2a_{12}}\right)^{k_2+l_2} \\ &=: N\tilde{\mathcal{S}}_R(\mathcal{H}_1,\mathcal{H}_2,l_1,l_2,h_0),\end{aligned} \tag{33}$$

say. Using Perron's formula, we find that

$$\tilde{\mathcal{S}}_R(\mathcal{H}_1,\mathcal{H}_2,l_1,l_2,h_0) = \frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)}F(s_1,s_2)\frac{R^{s_1}}{s_1^{k_1+l_1+1}}\frac{R^{s_2}}{s_2^{k_2+l_2+1}}ds_1ds_2, \tag{34}$$

where

$$
\begin{aligned}
&F(s_1, s_2) \\
(35)\qquad &= \sum_{a_1,a_2,a_{12}} \frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2 \nu^*_{a_1}(\mathcal{H}_1^0)\nu^*_{a_2}(\mathcal{H}_2^0)\overline{\nu}^*_{a_{12}}((\mathcal{H}_1 \overline{\cap} \mathcal{H}_2)^0)}{\varphi(a_1)a_1^{s_1}\varphi(a_2)a_2^{s_2}\varphi(a_{12})a_{12}^{s_1+s_2}} \\
&= \prod_p \left(1 - \frac{\nu_p^*(\mathcal{H}_1^0)}{(p-1)p^{s_1}} - \frac{\nu_p^*(\mathcal{H}_2^0)}{(p-1)p^{s_2}} + \frac{\overline{\nu}_p^*((\mathcal{H}_1 \overline{\cap} \mathcal{H}_2)^0)}{(p-1)p^{s_1+s_2}}\right).
\end{aligned}
$$

The function $F(s_1, s_2)$ above depends on h_0 and $\mathcal{H}_1$, $\mathcal{H}_2$. Here we recall $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$. We need to consider the following four cases.

1) Suppose $h_0 \notin \mathcal{H}$ $(= \mathcal{H}_1 \cup \mathcal{H}_2)$. In this case, write

$$F(s_1, s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2) \frac{\zeta(1+s_1+s_2)^r}{\zeta(1+s_1)^{k_1}\zeta(1+s_2)^{k_2}}.$$

2) Suppose $h_0 \in \mathcal{H}_1 \backslash \mathcal{H}_2$. In this case, write

$$F(s_1, s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2) \frac{\zeta(1+s_1+s_2)^r}{\zeta(1+s_1)^{k_1-1}\zeta(1+s_2)^{k_2}}.$$

3) Suppose $h_0 \in \mathcal{H}_2 \backslash \mathcal{H}_1$. In this case, write

$$F(s_1, s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2) \frac{\zeta(1+s_1+s_2)^r}{\zeta(1+s_1)^{k_1}\zeta(1+s_2)^{k_2-1}}.$$

4) Suppose $h_0 \in \mathcal{H}_1 \cap \mathcal{H}_2$. In this case, write

$$F(s_1, s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2) \frac{\zeta(1+s_1+s_2)^{r-1}}{\zeta(1+s_1)^{k_1-1}\zeta(1+s_2)^{k_2-1}}.$$

Then, in each case the function G is analytic and uniformly bounded for $\sigma_i > -c$ $(i = 1, 2)$ with any $c < 1/4$. Further, by investigating the Euler product of $F(s_1, s_2)$, we obtain

$$G_{\mathcal{H}_1,\mathcal{H}_2}(0, 0) = \mathfrak{S}(\mathcal{H}^0).$$

Using these properties and moving the paths of integral in (34) appropriately (this process requires some knowledge for the zero-free region of the Riemann zeta-function), we obtain the asymptotic formula for $\tilde{\mathcal{S}}_R$ in (33), by Cauchy's residue theorem. In this way the asymptotic formulas in Proposition 2.4 are obtained.

§ 3. The breakthrough of Zhang

§ 3.1. The result

As we have seen in Section 2, Goldston, Pintz and Yıldırım [9] proved that if the primes have level of distribution $\theta = 1/2 + \varpi$ for an arbitrary small $\varpi > 0$, then

$$\liminf_{n\to\infty}(p_{n+1} - p_n) < \infty \qquad (36)$$

holds. Since the result $\theta = 1/2$ is known (the Bombieri-Vinogradov theorem), it is said in [9] that the estimate (36) is within a hair's breadth. In 2013, Zhang [24] finally proved the following result, called the bounded gaps between primes:

Theorem 3.1 (Zhang, [24]). *Suppose that $\mathcal{H} = \{h_1, \cdots, h_{k_0}\}$ is admissible with $k_0 \geq 3.5 \times 10^6$. Then the k_0-tuple*

$$\{n + h_1, n + h_2, \cdots, n + h_{k_0}\} \tag{37}$$

contains at least two primes infinitely often. Consequently we have

$$\liminf_{n\to\infty}(p_{n+1} - p_n) < 7 \times 10^7. \tag{38}$$

The upper bound (38) follows from the fact that the set $\mathcal{H}$ is admissible if it consists of k_0 distinct primes, each of which is greater than k_0, and the inequality

$$\pi(7 \times 10^7) - \pi(3.5 \times 10^6) > 3.5 \times 10^6,$$

where $\pi(x)$ denotes the number of primes below x.

§ 3.2. Sketch of the proof of Theorem 3.1

Hereafter we let $k_0 = 3.5 \times 10^6$. Let $\mathcal{H} = \{h_1, \cdots, h_{k_0}\}$ be an admissible set consisting of k_0 elements. For any prime p, let $\nu_p = \nu_p(\mathcal{H})$ be the number of elements of $\mathcal{H}$ modulo p, and put

$$\mathfrak{S} = \prod_p \left(1 - \frac{\nu_p}{p}\right)\left(1 - \frac{1}{p}\right)^{-k_0}.$$

Recall that the function $\theta(n)$ was defined by

$$\theta(n) = \begin{cases} \log n & (n : \text{prime}) \\ 0 & (\text{otherwise}). \end{cases}$$

For any sufficiently large x and coprime integers c, d, put

$$\Delta(\theta; d, c) = \sum_{\substack{x<n\leq 2x \\ n\equiv c (\text{mod } d)}} \theta(n) - \frac{1}{\varphi(d)} \sum_{\substack{x<n\leq 2x \\ (n,d)=1}} \theta(n).$$

Next, put

$$P(n) = \prod_{j=1}^{k_0}(n + h_j)$$

and for $1 \le i \le k_0$, define the set $\mathcal{C}_i(d)$ by

$$\mathcal{C}_i(d) = \{c \,|\, 1 \le c \le d, (c, d) = 1, P(c - h_i) \equiv 0 \pmod{d}\}.$$

We put

$$D = x^{\frac{1}{4}+\varpi}$$

for $\varpi = 1/1168$, and define the function $g(y)$ by

$$g(y) = \frac{1}{(k_0 + l_0)!}\left(\log \frac{D}{y}\right)^{k_0+l_0} \qquad (\text{if } y < D),$$

and

$$g(y) = 0 \quad (\text{if } y \ge D),$$

where $l_0 = 180$. Write

$$D_1 = x^{\varpi}, \quad \mathcal{P} = \prod_{p<D_1} p.$$

Then Zhang's weight for the sieve is defined by

$$\lambda(n) = \sum_{d|(P(n),\mathcal{P})} \mu(d)g(d). \tag{39}$$

We evaluate and compare the sums

$$S_1 = \sum_{x<n\le 2x} \lambda(n)^2 \tag{40}$$

and

$$S_2 = \sum_{x<n\le 2x} \left(\sum_{i=1}^{k_0} \theta(n + h_i)\right) \lambda(n)^2. \tag{41}$$

It follows from the argument in Section 2 that if

$$S_2 - (\log 3x)S_1 > 0, \tag{42}$$

there is an integer $n \in (x, 2x]$ such that at least two of $n + h_1, \cdots, n + h_{k_0}$ are primes. To obtain the asymptotic behavior of S_2, we need to evaluate the sum

$$\mathcal{E} = \sum_{1\le i\le k_0} \sum_{\substack{d<D^2 \\ d|\mathcal{P}}} |\mu(d)|\tau_3(d)\tau_{k_0-1}(d) \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta; d, c)|.$$

Zhang's breakthrough was accomplished by proving the following result:

Lemma 3.2 ([24], Theorem 2). *For $1 \le i \le k_0$, we have*

$$\sum_{\substack{d<D^2 \\ d|\mathcal{P}}} \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta;d,c)| \ll_A x\mathcal{L}^{-A} \tag{43}$$

for any $A > 0$, where $\mathcal{L} = \log x$.

The estimate (43) may be regarded as an extension of the Bombieri-Vinogradov theorem, since $D^2 = x^{1/2+2\varpi} > x^{1/2}$ ($\varpi = 1/1168$), although the sum is restricted by suitable conditions. We will briefly see how the estimate (43) is obtained in the next subsection.

§ 3.3. A brief overview of the proof of Lemma 3.2

The basic idea of Zhang is to extend the range of the Bombieri-Vinogradov theorem, and in some sense he accomplished this by proving Lemma 3.2. He reduced the problem to evaluating the sum of $|\Delta(\gamma;d,c)|$, where γ represents certain Dirichlet convolutions and $\Delta(\gamma;d,c)$ is defined by replacing $\theta(n)$ with $\gamma(n)$ in the definition of $\Delta(\theta;d,c)$. He considered three types of convolutions. Write

$$x_1 = x^{\frac{3}{8}+8\varpi},\ x_2 = x^{\frac{1}{2}-4\varpi},\ \eta = 1+\mathcal{L}^{-2A}.$$

In the first two types, the function γ is of the form $\gamma = \alpha * \beta$, where $(\alpha*\beta)(n) = \sum_{d|n} \alpha(n/d)\beta(d)$, and the functions α, β satisfy the following conditions:

(A1) $\alpha = (\alpha(n))$ is supported on $[M, \eta^{j_1}M)$, $j_1 \le 19$, $\alpha(m) \ll \tau_{j_1}(m)\mathcal{L}$.

(A2) $\beta = (\beta(n))$ is supported on $[N, \eta^{j_2}N)$, $j_2 \le 19$, $\beta(n) \ll \tau_{j_2}(n)\mathcal{L}$, $x_1 < N \le 2x^{\frac{1}{2}}$. Further, for q, r and a satisfying $(a,r)=1$, the Siegel-Walfisz type estimate

$$\sum_{\substack{n\equiv a \pmod r \\ (n,q)=1}} \beta(n) - \frac{1}{\varphi(r)} \sum_{(n,qr)=1} \beta(n) \ll \tau_{20}(q)N\mathcal{L}^{-200A} \tag{44}$$

holds.

(A3) $j_1 + j_2 \le 20,\ [MN, \eta^{20}MN) \subset [x, 2x)$.

We say that γ is of Type I if $x_1 < N \le x_2$, and that γ is of Type II if $x_2 < N < 2x^{\frac{1}{2}}$. We say that γ is of Type III if it is of the form $\gamma = \alpha * \aleph_{N_1} * \aleph_{N_2} * \aleph_{N_3}$, where $\aleph_N$ denotes the characteristic function of the set $[N, \eta N)\cap\mathbf{Z}$, and α satisfies (A1) with $j_1 \le 17$, and the following conditions are satisfied:

(A4) $N_3 \le N_2 \le N_1,\ MN_1 \le x.$

(A5) $[MN_1N_2N_3, \eta^{20}MN_1N_2N_3) \subset [x, 2x).$

Write

$$D_0 = \exp\left(\mathcal{L}^{\frac{1}{k_0}}\right),\ D_1 = x^{\varpi},\ D_2 = x^{\frac{1}{2}-\epsilon}$$

and put

$$P_0 = \prod_{p \le D_0} p.$$

To prove Lemma 3.2, we first observe that the Bombieri-Vinogradov theorem indicates that the contribution of the terms with $d \le D_2$ is at most $O(x\mathcal{L}^{-A})$. Further, the function θ may be replaced with the von Mangoldt function Λ defined in the previous section. We also find that we may impose the constraint $(d, P_0) < D_1$. Hence the proof of Lemma 3.2 is reduced to prove

$$\sum_{\substack{D_2<d<D^2 \\ d|\mathcal{P} \\ (d,P_0)<D_1}} \sum_{c \in \mathcal{C}_i(d)} |\Delta(\Lambda; d, c)| \ll x\mathcal{L}^{-A}. \tag{45}$$

Zhang decomposed the von Mangoldt function by

$$\Lambda(n) = \sum_{j=1}^{10} (-1)^{j-1} \binom{10}{j} \sum_{\substack{M_j, \cdots, M_1 \\ N_j, \cdots, N_1}} (\mu \aleph_{M_j}) * \cdots * (\mu \aleph_{M_1}) * (\aleph_{N_j}) * \cdots * (\aleph_{N_1} \log)(n) \tag{46}$$

for $x < n \le 2x$, where $M_j, \cdots, M_1, N_j, \cdots, N_1 \ge 1$ run over the powers of η satisfying

$$M_t \le x^{\frac{1}{10}},$$

$$[M_j \cdots M_1 N_j \cdots N_1, \eta^{20} M_j \cdots M_1 N_j \cdots N_1) \cap [x, 2x) \ne \emptyset.$$

This decomposition follows from the Heath-Brown identity (see [17]). After several combinatorial arguments, we find that the proof of (45) is reduced to showing

$$\sum_{\substack{D_2<d<D^2 \\ d|\mathcal{P} \\ (d,P_0)<D_1}} \sum_{c \in \mathcal{C}_i(d)} |\Delta(\gamma; d, c)| \ll x\mathcal{L}^{-A} \tag{47}$$

for γ of Types I, II, III. Note that if $d|\mathcal{P}$ and d is not too small, then d can be factored as

$$d = rq \tag{48}$$

with the range of r flexibly chosen. In Type I and II estimates, Zhang combined the dispersion method invented by Bombieri, Friedlander and Iwaniec [1] with the factorization (48). In addition, a variant of Weil's bound for Kloosterman sums is employed. The Type III estimate essentially relies on the result of Birch and Bombieri, which is introduced in the appendix of the paper of Friedlander and Iwaniec [6]. This result follows from the Deligne's proof of the Weil Conjecture [4].

Remark. After the work of Zhang, the value of the gap between primes is improved day by day by several mathematicians. Among others, the POLYMATH 8 project was initiated by Terence Tao to understand and improve the techniques of Zhang. In the first version of the paper [20], which was published in October 2013 (Zhang's work was announced in May 2013), Tao et al. obtained

$$\liminf_{n\to\infty}(p_{n+1}-p_n) \leq 4680.$$

However, at that time James Maynard [19] obtained significantly better upper bound

$$\liminf_{n\to\infty}(p_{n+1}-p_n) \leq 600. \tag{49}$$

In the next section, we will see how the upper bound (49) was obtained.

§ 4. The work of Maynard

§ 4.1. The results

We have defined the concept of the level of distribution of primes in Section 2. In this section, we employ a slightly different definition for this terminology. Let $\pi(x)$ be the number of primes p below x, and $\pi(x;q,a)$ be the number of primes p below x satisfying $p \equiv a(\bmod\ q)$. Given $\theta > 0$, we say that the primes have level of distribution θ if

$$\sum_{q\leq x^\theta} \max_{(a,q)=1} \left|\pi(x;q,a) - \frac{\pi(x)}{\varphi(q)}\right| \ll_A \frac{x}{(\log x)^A} \tag{50}$$

for any $A > 0$. The Bombieri-Vinogradov theorem implies that the primes have level of distribution θ for every $\theta < 1/2$, and the Elliott-Halberstam conjecture indicates that this will be extended to every $\theta < 1$.

We would like to see three of Maynard's results in [19].

Theorem 4.1 (Maynard [19], Theorem 1.3)**.** *We have*

$$\liminf_{n\to\infty}(p_{n+1}-p_n) \leq 600. \tag{51}$$

On the other hand, assuming the Elliott-Halberstam conjecture, rather stronger results can be obtained:

Theorem 4.2 (Maynard [19], Theorem 1.4). *Assume that the primes have level of distribution θ for every $\theta < 1$. Then*

$$\liminf_{n\to\infty}(p_{n+1} - p_n) \le 12, \tag{52}$$

$$\liminf_{n\to\infty}(p_{n+2} - p_n) \le 600. \tag{53}$$

He also obtained a result on gaps between m consecutive primes for general m.

Theorem 4.3 (Maynard [19], Theorem 1.1). *Let $m \in \mathbf{N}$. We have*

$$\liminf_{n\to\infty}(p_{n+m} - p_n) \ll m^3 e^{4m}. \tag{54}$$

Remark. It is mentioned in the paper of Maynard [19] that Terence Tao (private communication with Maynard) had independently proven a (slightly weaker) bound like Theorem 4.3 by a similar method. Hence the method (resp. sieve) to prove Theorem 4.3 is called the Maynard-Tao method (resp. sieve).

§ 4.2. The sieve of Maynard and Tao

Theorems 4.1-4.3 are proved by the same sieve. In this subsection we shortly see how Maynard and Tao tried to evaluate the gaps between primes. Throughout this section we assume that the admissible set $\mathcal{H} = \{h_1, \cdots, h_k\}$ is bounded, that is, there exists a constant $C = C_k > 0$ which depends only on k such that $0 \le h_i \le C_k$ $(i = 1, \cdots, k)$ holds. Let N be a sufficiently large positive integer and put

$$D_0 = \log\log\log N, \quad W = \prod_{p \le D_0} p \ll (\log\log N)^2.$$

We denote by $\chi_{\mathbf{P}}$ the indicator function of the set of primes $\mathbf{P}$. Let $\mathcal{H} = \{h_1, \cdots, h_k\}$ be an admissible set with k elements. The admissibility of the set $\mathcal{H}$ allows us to choose an integer ν_0 such that $\nu_0 + h_i$ is coprime to W for all $i = 1, \cdots, k$. Next, let $F : \mathbf{R}^k \to \mathbf{R}$ be a smooth function supported on $\mathcal{R}_k = \{(x_1, \cdots, x_k) \in \mathbf{R}^k \mid x_1, \cdots, x_k \ge 0, \sum_{i=1}^k x_i \le 1\}$. We assume that the primes have level of distribution $\theta > 0$ and put $R = N^{\frac{\theta}{2} - \delta}$, where $\delta > 0$ is a small fixed number. We put

$$\lambda_{d_1,\cdots,d_k} = \left(\prod_{i=1}^k \mu(d_i) d_i\right) \sum_{\substack{r_1,\cdots,r_k \\ d_i \mid r_i (\forall i) \\ (r_i, W)=1 (\forall i)}} \frac{\mu(\prod_{i=1}^k r_i)^2}{\prod_{i=1}^k \varphi(r_i)} F\left(\frac{\log r_1}{\log R}, \cdots, \frac{\log r_k}{\log R}\right)$$

whenever $(d_i, W) = 1$ for all $i = 1, \cdots, k$, and otherwise put $\lambda_{d_1,\cdots,d_k} = 0$. The weight of the sieve of Maynard is defined by

$$w_n = w_n(\mathcal{H}) := \left(\sum_{d_i \mid n+h_i \ (\forall i)} \lambda_{d_1,\cdots,d_k}\right)^2.$$

We put

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

Theorems 4.1-4.3 are obtained by making use of the following asymptotic result:

Proposition 4.4. *We have*

$$S_1 = \frac{(1+o(1))\varphi(W)^k N (\log R)^k}{W^{k+1}} I_k(F),$$

$$S_2 = \frac{(1+o(1))\varphi(W)^k N (\log R)^{k+1}}{W^{k+1} \log N} \sum_{m=1}^{k} J_k^{(m)}(F)$$

provided $I_k(F) \neq 0$ *and* $J_k^{(m)}(F) \neq 0$ *for each* m*, where*

$$I_k(F) = \int_0^1 \cdots \int_0^1 F(t_1, \cdots, t_k)^2 dt_1 \cdots dt_k,$$

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

We consider the sum

$$S = S_2 - \rho S_1 = \sum_{\substack{N \le n < 2N \\ n \equiv \nu_0 \pmod W}} \left(\sum_{i=1}^{k} \chi_{\mathbf{P}}(n+h_i) - \rho \right) w_n$$

for $\rho \ge 1$. If one can show $S > 0$ for all large N, then at least $\lfloor \rho + 1 \rfloor$ of the $n + h_i$ are prime infinitely often. Let $\mathcal{S}_k$ denote the set of Riemann-integrable functions $F : [0,1]^k \to \mathbf{R}$ supported on $\mathcal{R}_k$ with $I_k(F) \neq 0$, $J_k^{(m)}(F) \neq 0$ $(m = 1, \cdots, k)$. We put

$$M_k = \sup_{F \in \mathcal{S}_k} \frac{\sum_{m=1}^{k} J_k^{(m)}(F)}{I_k(F)}, \quad r_k = \lceil \frac{\theta M_k}{2} \rceil.$$

The definition of M_k allows us to choose $F_0 \in \mathcal{S}_k$ such that

$$\sum_{m=1}^{k} J_k^{(m)}(F_0) > (M_k - \delta) I_k(F_0) > 0,$$

where $\delta > 0$ is a fixed small constant in the definition of R (recall $R := N^{\frac{\theta}{2} - \delta}$). Since F_0 is Riemann-integrable, there is a smooth function F_1 supported on $\mathcal{R}_k$ such that

$$\sum_{m=1}^{k} J_k^{(m)}(F_1) > (M_k - 2\delta) I_k(F_1) > 0.$$

Applying Proposition 4.4 with $F = F_1$, we obtain

$$\begin{aligned} S &= \frac{\varphi(W)^k N(\log R)^k}{W^{k+1}} \left(\frac{\log R}{\log N} \sum_{j=1}^{k} J_k^{(m)}(F_1) - \rho I_k(F_1) + o(1) \right) \\ &\geq \frac{\varphi(W)^k N(\log R)^k}{W^{k+1}} I_k(F_1) \left(\left(\frac{\theta}{2} - \delta \right) (M_k - 2\delta) - \rho + o(1) \right). \end{aligned}$$

The right hand side becomes positive if $\rho = \theta M_k/2 - \epsilon$ and $\delta > 0$ is sufficiently small. When $\epsilon > 0$ is suitably small, we have $\lfloor \rho + 1 \rfloor = \lceil \theta M_k/2 \rceil = r_k$. Hence the admissible k-tuple $n+h_1, \cdots, n+h_k$ contains at least r_k prime numbers infinitely often. In particular, we have

$$\liminf_{n\to\infty}(p_{n+r_k-1} - p_n) \leq \max_{1\leq i<j\leq k} |h_i - h_j|. \tag{55}$$

Theorems 4.1-4.3 are obtained by combining the argument above and the lower bounds

(a) $M_5 > 2$,

(b) $M_{105} > 4$,

(c) For any sufficiently large k, $M_k > \log k - 2\log\log k - 2$.

Maynard proved (a), (b) by numerical computations, using *Mathematica*. The lower bound (c) is obtained by choosing a smooth function $F(t_1, \cdots, t_k)$ of k variables precisely. We omit to describe this in detail. Theorem 4.1 follows from (55) and the fact that there exists an admissible set $\mathcal{H}_1 = \{0, 10, 12, 24, \cdots, 600\}$ consisting of 105 elements whose the smallest element is 0 and the largest one is 600. He found this admissible set by *Mathematica*, and the elements are listed at the bottom of page 390 of his paper [19]. Theorem 4.2 follows from (55), (a) and (b). To prove (53), we use the admissible set $\mathcal{H}_1$ above. To prove (52), we use an admissible set $\mathcal{H}_2 = \{0, 2, 6, 8, 12\}$. To prove Theorem 4.3, we use the facts that $\mathcal{H} = \{p_{\pi(k)+1}, p_{\pi(k)+2}, \cdots, p_{\pi(k)+k}\}$ is an admissible set consisting of precisely k elements and p_n is asymptotically equal to $n \log n$.

§ 4.3. How the equidistribution of primes works ?

The most important part of Maynard's work in [19] is, the author thinks, the arguments to prove Proposition 4.4. One of the key processes is the translations in the weight by

$$y_{r_1,\cdots,r_k} = \left(\prod_{i=1}^{k} \mu(r_i)\varphi(r_i)\right) \sum_{\substack{d_1,\cdots,d_k \\ r_i|d_i\ (\forall i)}} \frac{\lambda_{d_1,\cdots,d_k}}{\prod_{i=1}^{k} d_i}, \tag{56}$$

$$y^{(m)}_{r_1,\cdots,r_k} = \left(\prod_{i=1}^{k} \mu(r_i)g(r_i)\right) \sum_{\substack{d_1,\cdots,d_k \\ r_i|d_i\ (\forall i) \\ d_m=1}} \frac{\lambda_{d_1,\cdots,d_k}}{\prod_{i=1}^{k} \varphi(d_i)} \quad (m = 1,\cdots,k), \tag{57}$$

where g is the totally multiplicative function defined by $g(p) = p-2$ for any $p \in \mathbf{P}$. The inverse relation of (56) is given by

$$\lambda_{d_1,\cdots,d_k} = \left(\prod_{i=1}^{k} \mu(d_i)d_i\right) \sum_{\substack{r_1,\cdots,r_k \\ d_i|r_i\ (\forall i)}} \frac{y_{r_1,\cdots,r_k}}{\prod_{i=1}^{k} \varphi(r_i)}, \tag{58}$$

if $\prod_{i=1}^{k} d_i$ is square-free. Put

$$y_{\max} = \sup_{r_1,\cdots,r_k} |y_{r_1,\cdots,r_k}|, \quad y^{(m)}_{\max} = \sup_{r_1,\cdots,r_k} |y^{(m)}_{r_1,\cdots,r_k}|.$$

Recall that S_1 and S_2 were defined by

$$S_1 = \sum_{\substack{N\le n<2N \\ n\equiv\nu_0 \ (\mathrm{mod}\ W)}} \left(\sum_{\substack{d_1,\cdots,d_k \\ d_i|n+h_i\ (\forall i)}} \lambda_{d_1,\cdots,d_k}\right)^2,$$

$$S_2 = \sum_{m=1}^{k} \left\{ \sum_{\substack{N\le n<2N \\ n\equiv\nu_0 \ (\mathrm{mod}\ W)}} \chi_{\mathbf{P}}(n+h_m) \left(\sum_{\substack{d_1,\cdots,d_k \\ d_i|n+h_i\ (\forall i)}} \lambda_{d_1,\cdots,d_k}\right)^2 \right\} =: \sum_{m=1}^{k} S_2^{(m)}.$$

The first step of Maynard is to obtain the following new expressions for S_1 and $S_2^{(m)}$:

$$S_1 = \frac{N}{W} \sum_{r_1,\cdots,r_k} \frac{y^2_{r_1,\cdots,r_k}}{\prod_{i=1}^{k} \varphi(r_i)} + O\left(\frac{y^2_{\max}\varphi(W)^k N(\log R)^k}{W^{k+1}D_0}\right), \tag{59}$$

$$\begin{aligned} S_2^{(m)} =& \frac{N}{\varphi(W)\log N} \sum_{r_1,\cdots,r_k} \frac{(y^{(m)}_{r_1,\cdots,r_k})^2}{\prod_{i=1}^{k} g(r_i)} \\ &+ O\left(\frac{(y^{(m)}_{\max})^2\varphi(W)^{k-2}N(\log R)^{k-2}}{W^{k-1}D_0}\right) + O_A\left(\frac{y^2_{\max}N}{(\log N)^A}\right) \quad (\forall A>0). \end{aligned} \tag{60}$$

We omit to describe the detailed arguments as usual, but the most important point is the adaption of our assumption on level of distribution of primes. The influence can be seen in the second error term in (60). Let us take a look at this argument. First, by expanding the weight and exchanging the order of summation, we have

$$S_2^{(m)} = \sum_{\substack{d_1,\cdots,d_k \\ e_1,\cdots,e_k}} \lambda_{d_1,\cdots,d_k}\lambda_{e_1,\cdots,e_k} \sum_{\substack{N\le n<2N \\ n\equiv \nu_0 \pmod W \\ [d_i,e_i]|n+h_i\ (\forall i)}} \chi_{\mathbf{P}}(n+h_m).$$

If $W, [d_1,e_1],\cdots,[d_k,e_k]$ are pairwise coprime, the inner sum above can be written as a sum over a single residue class modulo $q := W\prod_{i=1}^k [d_i,e_i]$. We note that we may impose the condition $d_m = e_m = 1$, because otherwise the condition $[d_m,e_m]|n+h_m$ implies $\chi_{\mathbf{P}}(n+h_m)=0$. We also note that if either one pair of $W, [d_1,e_1],\cdots,[d_k,e_k]$ share a common factor, then the contribution of the inner sum is zero. This fact requires some consideration. If a prime p divides both W and $[d_i,e_i]$, then either $(W,d_i)>1$ or $(W,e_i)>1$ holds. In each case we have $\lambda_{d_1,\cdots,d_k}\lambda_{e_1,\cdots,e_k}=0$. If a prime p divides both $[d_i,e_i]$ and $[d_j,e_j]$ $(i\neq j)$, then at least one of the following is true:
1) $p|(d_i,d_j)$, 2) $p|(e_i,e_j)$, 3) $p|(d_i,e_j)$, 4) $p|(d_j,e_i)$.
If 1) or 2) is valid, then $\lambda_{d_1,\cdots,d_k}\lambda_{e_1,\cdots,e_k}=0$. If 3) or 4) is valid, then the condition of the inner sum indicates $n+h_i\equiv 0 \pmod p$, $n+h_j\equiv 0 \pmod p$. This implies that $h_i\equiv h_j \pmod p$, which means that $|h_i-h_j|$ is a positive multiple of the prime p. However, if $\lambda_{d_1,\cdots,d_k}\lambda_{e_1,\cdots,e_k}\neq 0$, then all of $d_1,\cdots,d_k,e_1,\cdots,e_k$ are coprime to $W=\prod_{p\le D_0} p$. Hence the prime p which divides either one of $d_1,\cdots,d_k,e_1,\cdots,e_k$ is greater than $D_0=\log\log\log N$. These facts contradict to our assumption that the admissible set $\mathcal{H}=\{h_1,\cdots,h_k\}$ is bounded. Consequently we have
(61)

$$S_2^{(m)} = \frac{X_N}{\varphi(W)} \sideset{}{'}\sum_{\substack{d_1,\cdots,d_k \\ e_1,\cdots,e_k \\ d_m=e_m=1}} \frac{\lambda_{d_1,\cdots,d_k}\lambda_{e_1,\cdots,e_k}}{\prod_{i=1}^k \varphi([d_i,e_i])} + O\left(\sum_{\substack{d_1,\cdots,d_k \\ e_1,\cdots,e_k}} |\lambda_{d_1,\cdots,d_k}\lambda_{e_1,\cdots,e_k}| E(N,q)\right),$$

where $q = W\prod_{i=1}^k [d_i,e_i]$,

$$X_N = \sum_{N\le n<2N} \chi_{\mathbf{P}}(n) = \frac{N}{\log N} + O\left(\frac{N}{\log^2 N}\right),$$

$$E(N,q) = 1 + \sup_{(a,q)=1} \left| \sum_{\substack{N\le n<2N \\ n\equiv a \pmod q}} \chi_{\mathbf{P}}(n) - \frac{1}{\varphi(q)} \sum_{N\le n<2N} \chi_{\mathbf{P}}(n) \right|,$$

and the first sum in (61) is over $d_1, \cdots, d_k, e_1, \cdots, e_k$ satisfying the condition that $W, [d_1, e_1], \cdots, [d_k, e_k]$ are pairwise coprime. It follows from the above argument that we only need to deal with square-free q. Moreover, if $d_1, \cdots, d_k$ and $e_1, \cdots, e_k$ are in the support of λ, then $q < WR^2$. Given an integer r, there are at most $\tau_{3k}(r)$ choices of $d_1, \cdots, d_k, e_1, \cdots, e_k$ for which $W\prod_{i=1}^k [d_i, e_i] = r$. Moreover, with a brief computation using (58), we find that $\lambda_{\max} \ll y_{\max}(\log R)^k$. Consequently we have

$$\sum_{\substack{d_1,\cdots,d_k \\ e_1,\cdots,e_k}} |\lambda_{d_1,\cdots,d_k}\lambda_{e_1,\cdots,e_k}|\, E(N,q) \ll y_{\max}^2(\log R)^{2k} \sum_{r<WR^2} \mu^2(r)\tau_{3k}(r)E(N,r).$$

Using the Cauchy-Schwarz inequality and the trivial bound $E(N,q) \ll N/\varphi(q)$, the right hand side is bounded by

$$\ll y_{\max}^2(\log R)^{2k} \left(\sum_{r<WR^2} \mu^2(r)\tau_{3k}^2(r)\frac{N}{\varphi(r)}\right)^{\frac{1}{2}} \left(\sum_{r<WR^2} \mu^2(r)E(N,r)\right)^{\frac{1}{2}}.$$

The first sum is bounded by

$$\sum_{r<WR^2} \mu^2(r)\tau_{3k}^2(r)\frac{N}{\varphi(r)} \ll N(\log N)^{c_k} \quad (\exists c_k > 0).$$

On the other hand, since we have assumed that primes have level of distribution θ, we have

$$\sum_{r<WR^2} \mu^2(r)E(N,r) \ll_B \frac{N}{(\log N)^B}$$

for any $B > 0$. This is because

$$WR^2 \ll N^{\theta-2\delta}(\log\log N)^2 \ll N^{\theta-\delta}.$$

Combining these results we obtain

$$\sum_{\substack{d_1,\cdots,d_k \\ e_1,\cdots,e_k}} |\lambda_{d_1,\cdots,d_k}\lambda_{e_1,\cdots,e_k}|\, E(N,q) \ll_A \frac{Ny_{\max}^2}{(\log N)^A}$$

for any $A > 0$, by choosing $B = 2A + c_k + 4k$.

§ 4.4. The Euler-Maclaurin type argument

Recall that $\lambda_{d_1,\cdots,d_k}$ and $y_{r_1,\cdots,r_k}$ are related by

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

whenever $\prod_{i=1}^k d_i$ is square-free. Maynard constructed the sieve by

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

if $r := \prod_{i=1}^k r_i$ satisfies $(r, W) = 1$, $\mu^2(r) = 1$, and otherwise $y_{r_1,\cdots,r_k} = 0$, where $F : \mathbf{R}^k \to \mathbf{R}$ is a smooth function supported on $\mathcal{R}_k = \{(x_1, \cdots, x_k) \mid x_1, \cdots, x_k \geq 0, \sum_{i=1}^k x_i \leq 1\}$. To compute the arithmetic sums in (59) and (60), the following lemma is adapted:

Lemma 4.5. *Let $A_1, A_2, L > 0$. Let γ be a multiplicative function satisfying*

$$0 \leq \frac{\gamma(p)}{p} \leq 1 - A_1, \quad -L \leq \sum_{w \leq p \leq z} \frac{\gamma(p) \log p}{p} - \log \frac{z}{w} \leq A_2$$

for any $2 \leq w \leq z$. Let g be the totally multiplicative function defined on primes by $g(p) = \gamma(p)/(p - \gamma(p))$. Finally, let $G : [0, 1] \to \mathbf{R}$ be smooth, and let $G_{\max} = \sup_{t \in [0,1]}(|G(t)| + |G'(t)|)$. Then

$$\sum_{d<z} \mu^2(d) g(d) G\left(\frac{\log d}{\log z}\right) = \mathfrak{S} \log z \int_0^1 G(x) dx + O_{A_1,A_2}(\mathfrak{S} L G_{\max}),$$

where

$$\mathfrak{S} = \prod_p \left(1 - \frac{\gamma(p)}{p}\right)^{-1} \left(1 - \frac{1}{p}\right).$$

The implied constant in the O-term is independent of G and L.

The above result is obtained by putting $\kappa = 1$ in Lemma 4 of the paper of Goldston, Graham, Pintz and Yıldırım [8]. Put

$$F_{\max} = \sup_{(t_1,\cdots,t_k) \in [0,1]^k} \left\{ |F(t_1, \cdots, t_k)| + \sum_{i=1}^k \left| \frac{\partial F}{\partial t_i}(t_1, \cdots, t_k) \right| \right\}.$$

Then S_1 becomes

$$S_1 = \frac{N}{W} \sum_{r_1,\cdots,r_k}^{*} \left(\prod_{i=1}^k \frac{\mu^2(r_i)}{\varphi(r_i)} \right) F\left(\frac{\log r_1}{\log R}, \cdots, \frac{\log r_k}{\log R}\right)^2 + O\left(\frac{F_{\max}^2 \varphi(W)^k N (\log R)^k}{W^{k+1} D_0}\right),$$

where the symbol $*$ means that two conditions $(r_i, r_j) = 1$ $(\forall i \neq j)$, $(r_i, W) = 1$ $(\forall i)$ are imposed on $r_1, \cdots, r_k$. We easily find that we can remove the former condition at the cost of an error of the size

$$O\left(\frac{F_{\max}^2 \varphi(W)^k N (\log R)^k}{W^{k+1} D_0}\right).$$

This is because if r_i and r_j have a common prime divisor, it is greater than D_0. Hence it suffices to compute the sum

$$\sum_{\substack{r_1,\cdots,r_k \\ (r_i,W)=1\ (\forall i)}} \left(\prod_{i=1}^{k} \frac{\mu^2(r_i)}{\varphi(r_i)}\right) F\left(\frac{\log r_1}{\log R},\cdots,\frac{\log r_k}{\log R}\right)^2.$$

By applying Lemma 4.5 with

$$\gamma(p) = \begin{cases} 1 & ((p,W)=1) \\ 0 & (\text{otherwise}) \end{cases}$$

to each variable, we find that the sum above becomes

$$\frac{\varphi(W)^k(\log R)^k}{W^k} I_k(F) + O\left(\frac{F_{\max}^2 \varphi(W)^k \log D_0 (\log R)^{k-1}}{W^k}\right),$$

where

$$I_k(F) = \int_0^1 \cdots \int_0^1 F(t_1,\cdots,t_k)^2 dt_1\cdots dt_k.$$

In this way the asymptotic formula (59) is obtained. On the other hand, (60) is obtained by rewriting

$$y^{(m)}_{r_1,\cdots,r_k} = \sum_{a_m} \frac{y_{r_1,\cdots,r_{m-1},a_m,r_{m+1},\cdots,r_k}}{\varphi(a_m)} + O\left(\frac{y_{\max}\varphi(W)\log R}{WD_0}\right) \quad (r_m = 1)$$

and applying Lemma 4.5 to each variable.

Remark. Currently (March 2018) the best results on small gaps between primes are due to DHJ Polymath ([21]), administered by Terence Tao. Put

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

Unconditionally they proved $H_1 \le 246$, $H_2 \le 398130$, $H_3 \le 24797814$, $\cdots$ and generally

$$H_m \le Cm\exp\left(\left(4 - \frac{28}{157}\right)m\right)$$

for all $m \ge 1$ and an absolute (and effective) constant C. They also proved $H_1 \le 6$, $H_2 \le 252$ under the assumption of so called the generalized Elliott-Halberstam conjecture. These great improvements are achieved by establishing a further generalization of the multidimensional Selberg sieve in this section and obtaining efficient numerical methods for solving a certain multidimensional variational problem to find a good test function.

§ 5. Acknowledgements

The author sincerely expresses his gratitude to Professor Yasuo Ohno for giving him the opportunities to talk at the RIMS workshop "Algebraic Number Theory and Related Topics" in 2016 and submit this article to Kôkyûroku Bessatsu. He also thanks to Professor Takashi Taniguchi for suggesting him to talk at that workshop. Finally, he expresses his gratitude to the referee for reading this article very carefully and giving him many valuable comments and suggestions.

References

[1] E. Bombieri, J. B. Friedlander, H. Iwaniec. *Primes in arithmetic progressions to large moduli*, Acta Math. **156** (1986), 203-251.
[2] E. Bombieri, H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. Ser. A **293** (1966), 1-18.
[3] J.-R. Chen, *On the representation of a large even integer as the sum of a prime and the product of at most two primes*, Sci. Sinica **16** (1973), 157-176.
[4] P. Deligne, *La conjecture de Weil I*, Publ. Math. IHES. **43** (1974), 273-307.
[5] P. Erdös *The difference of consecutive primes*, Duke Math. J. **6** (1940), 438-441.
[6] J. B. Friedlander, H. Iwaniec, *Incomplete Kloosterman sums and a divisor problem*, Ann of Math. (2) **121** (1985), 319-350.
[7] P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), 4-9.
[8] D. A. Goldston, S. W. Graham, J. Pintz, C. Y. Yıldırım, *Small gaps between products of two primes*, Proc. London Math. Soc. **88** (2009), 741-774.
[9] D. A. Goldston, J. Pintz, C. Y. Yıldırım, *Primes in tuples I*, Ann. of Math. (2) **170** (2009), No. 2, 819-862.
[10] D. A. Goldston, J. Pintz, C. Y. Yıldırım, *Primes in tuples II*, Acta. Math. **204** (2010), No. 1, 1-47.
[11] D. A. Goldston, J. Pintz, C. Y. Yıldırım, *Primes in tuples III. On the difference* $p_{n+\nu}-p_n$, Funct. Approx. Comment. Math. **35** (2006), 79-89.
[12] D. A. Goldston, J. Pintz, C. Y. Yıldırım, *Primes in tuples IV: Density of small gaps between consecutive primes*, Acta Arith. **160** (2013), No. 1, 37-53.
[13] D. A. Goldston, C. Y. Yıldırım, *Higher correlations of divisor sums related to primes II I: Small gaps between primes*, Proc. London Math. Soc. **95** (2007), 653-686.
[14] G. H. Hardy, J. E. Littlewood, *Some problems of 'Partitio Numerorum' III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), No. 1, 1-70.
[15] G. H. Hardy, J. E. Littlewood, unpublished manuscript, see [22].
[16] E. Hasanalizade, *The Goldston-Pintz-Yıldırım sieve and some applications*, Master's Thesis in Mathematics at KTH (2012).
[17] D. R. Heath-Brown, *Prime numbers in short intervals and a generalized Vaughan identity*, Canad. J. Math. **34** (1982), 1365-1377.
[18] D. R. Heath-Brown, *Prime twins and Siegel zeros*, Proc. London Math. Soc. (1983) (3) **47** (2), 193-224.
[19] J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), No. 1, 383-413.

[20] D. H. J. Polymath, W. Castryck, E. Fouvry, G. Harcos, E. Kowalski, P. Michel, P. Nelson, E. Paldi, J. Pintz, A. V. Sutherland, T. Tao, X. Xiao-Feng, *New equidistribution estimates of Zhang type*, Algebra Number Theory **8** (2014), No. 9, 2067-2199.

[21] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci. **1** (2014), 1:12.

[22] R. A. Rankin, *The difference between consecutive prime numbers II*, Proc. Cambridge Pilos. Soc. **36** (1940), 255-266.

[23] K. Soundararajan, *Small gaps between prime numbers : the work of Gondston-Pintz-Yıldırım*, Bull. Amer. Math. Soc. (N.S.) **44** (2007), No. 1, 1-18.

[24] Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), No. 3, 1121-1174.