UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI MATEMATICA
CORSO DI LAUREA IN MATEMATICA

TESI DI LAUREA
MASTER THESIS

BOUNDED GAPS BETWEEN PRIMES

RELATORE: PROF. A. LANGUASCO

LAUREANDO: MOHAMED TAOUFIQ DAMIR

ANNO ACCADEMICO 2014/2015

Acknowledgment

First and foremost, I express my gratitude to my supervisor, Prof. A.Languasco for giving me the opportunity to work on a topic that fascinated me, and for his guidance, precious advice, and engagement through the learning process of this work. This thesis would not exist without the support of ALGANT program, and I am deeply indebted to my professors in the University of Bordeaux and the University of Padova for influencing my mathematical perspective. I would like to thank the persons who made my last two years in Bordeaux and Padova, and specially L.Jakobsson for the support and sharing the long days of hard work. Lastly and more importantly, I want to express my deepest thanks to my parents and my family for making me the person I am today.

Introduction

The positive integers were undoubtedly the first objects investigated by the early mathematics; more than 2000 years ago and exactly 300 B.C, Euclid found out that the primes are the "bones" of the integers skeleton, or, in our modern language, he proved the first form of the fundamental theorem of arithmetics, namely the fact that every integer greater than one is either a prime or a product of primes. Few decades after Euclid, Eratosthenes gave an algorithm to determine the primes less than a given bound; by this Eratosthenes initiated a new research activity known as Sieve Theory. The Sieve Theory is a family of methods (i.e. sieves) that estimate the size of a given set of integers. One of the most asked questions in sieve theory is to estimate the number of primes satisfying some propriety in an interval of arbitrary length, for example counting the primes given by a polynomial expression up to some real number x.

It is easy to find a polynomial that contains infinitely many primes, starting from the fact that the polynomial $2x+1$ contains all the odd numbers for integer values of x, then definitely it contains all the odd primes. Indeed it was believed that one could find a polynomial P with integer coefficients such that $P(x)$ is prime for all integer x. Unfortunately in 1752 Goldbach proved that such a polynomial doesn't exist.

Introducing new analytic tools, 85 years after Goldbach, Dirichlet proved that the polynomial $ax+b$ contains infinitely many primes for a and b being co-prime. In 1904 Dickson conjectured that there are infinitely many positive integers x such that the polynomials a_ix+b_i are all primes for $0 \leq i \leq k$. Until writing these lines Dickson conjecture remains unsolved, even some of its special cases seems to be very hard to prove, for example taking $k=1$, $(a_0, b_0) = (1, 0)$, and $(a_1, b_1) = (1, 2)$ we find the twin primes conjecture; furthermore the special case $a_i = 1$ for $0 \leq i \leq k$ of the Dickson conjecture remain unsolved. The last special case, with an additional condition on the b_i's (to be defined later), is called the k-tuples conjecture.

Remark that if we relax the condition "the polynomials $x + b_i$ are all primes for $0 \leq i \leq k$" in the k-tuples conjecture and we replace it by "at least two of the polynomials $x + b_i$ are primes for some $0 \leq i \leq k$", then this implies that there exist infinitely many pairs of successive primes of difference H where $H = \max |b_i - b_j|$ for $i \neq j$. In other words we will prove the existence of bounded gaps between infinitely many successive primes.

In 2005, investigating this relaxed version of the k-tuples conjecture, Goldston, Pintz and Yildirim designed a sieve method (i.e. the GPY sieve) to deal with some problems related to the gaps between primes. Indeed they conditionally proved for the first time the existence of a finite gap $H \leq 16$ Surprisingly Yitang Zhang in 2013 proved unconditionally that $H \leq 70\ 000\ 000$ introducing more advanced analytic machinery. Two years later a further breakthrough was obtained by James Maynard (independently found also by Terence Tao) improving unconditionally the bound to $H \leq 600$ using more simpler arguments based on the ideas of Selberg and developing a "multidimensional" GPY sieve.

Organization of the thesis

In the first section, we will give an introduction of sieve theory, starting by estimating the integers given by a polynomial expression, then we will give the description of a sieve method developed by Selberg, and we will illustrate it by an application on the twin prime conjecture. In the rest of the thesis we will be mostly interested in the bounded gaps between primes. In the second section we will discus the Goldston, Pintz and Yildirim sieve, then we will give the complete conditional proof on the existence of a bounded gap between infinitely many consecutive primes. In the last section we will present Maynard's work and his unconditional proof of the bounded gaps. Hence, the main results in the present thesis are

Theorem 0.0.1. *(Goldston, Pintz and Yildirim* (2005)*) We have*

$$\Delta_1 = \lim_{n\to\infty} \inf \frac{p_{n+1} - p_n}{\log p_n} = 0.$$

Theorem 0.0.2. *(Goldston, Pintz and Yildirim* (2009)*)*

Assume the primes have level of distribution $\theta \geq 1/2$, then there exist an explicitly calculable constant $C(\theta)$, such that any admissible $k-$tuple with $k \geq C(\theta)$ contains at least two primes *infinitely often. In particular, we have*

$$\liminf(p_{n+1} - p_n) \leq 20.$$

Theorem 0.0.3. *(Maynard (2013)) We have*

$$\lim_{n\to\infty} \inf p_{n+1} - p_n = 600$$

Theorem 0.0.4. *(Maynard (2013)) Assuming the Elliott-Halberstam conjecture, we have*

$$\lim_{n\to\infty} \inf p_{n+1} - p_n = 12$$

—-

Contents

Chapter 1

Selberg sieve

1.1 Definitions and notations

Throughout this document, we will use with or without subscripts n, m, a and b for positive integers, p and q for prime numbers, and x, y for real numbers.

We denote by

1. $\omega(n)$ the number of prime divisors of$n$.
2. (a,b) is the greater commun divisor of a and b.
3. $a \equiv b \pmod{m}$ if it exists m such that $a = mn + b$.
4. $p|n \quad p$ divides n.
5. $p \nmid n \quad p$ doesn't divide n.
6. $\lfloor x \rfloor$ is the integer part of x.
7. $\{x\}$ is the fractional part of x.
8. $|A|$ is the cardinality of the set A.
9. $\varphi(n)$ is Euler's totient function, which is the number of positive integers less than or equal to n that are relatively prime to n.
10. $\pi(x)$ is the number of primes $p \leq x$, where x a real number.
11. $\pi(x, m, n)$ is the number of primes $p \leq x$, where $p \equiv n \pmod{m}$.

Recall the following definition of the Landau symbols.

Definition 1.1.1. *Let f, g two real functions*

1) If $g(x) > 0$ for all $x \geq a$, with $a \in \mathbb{R}$, we write

$$f(x) = O(g(x)) \qquad \text{(to be read as } f \text{ is big-oh of } g\text{)},$$

to mean that the quotient $\frac{f(x)}{g(x)}$ is bounded for $x \geq a$, that is, there exists a constant $M > 0$ such that

$$|f(x)| \leq Mg(x). \qquad \text{for every} x \geq a$$

2) We write

$$f(x) = o(g(x)) \qquad \text{(to be read as } f \text{ is little-oh of } g\text{)},$$

to mean that f is asymptotically dominated by g, and that is

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0.$$

3) We write

$$f(x) \sim g(x) \qquad \text{(to be read as } f \text{ has the same order of } g\text{)},$$

to mean that f is asymptotically equal to g, and that is

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1.$$

We will also use Vinogradov symbols $\ll$ and $\gg$. If $f(x) = O(g(x))$, it's equivalent to write $f(x) \ll g(x)$ or $g(x) \gg f(x)$.

In the next section we will present an important sieve method carrying by Atle Selberg. In order to do that, we will give an assortment of notations and definitions frequently used in sieve theory. [15]

1.2 Numbers given by polynomial expression

Let

$$A = \{h(n); n \leq x\} \quad \text{and} \quad A_d = \{a \in A; a \equiv 0 \pmod{d}\}.$$

Let further $h(n) \in \mathbb{Z}[X]$ and

$$\rho(d) = \#\{0 \leq n \leq d-1 : \quad h(n) \equiv 0 \pmod{d}\},$$

which denotes the number of solutions of $h(n) \equiv 0 \pmod{d}\}$.

Let P be the set of primes and $p \in P$; we denote

$$P(y) = \prod_{p \leq y} p.$$

To estimate $|A_d|$ we consider each residue class $\pmod{d}$ separately, so we get

$$|A_d| = \sum_{\substack{1 \leq l \leq d \\ h(l) \equiv 0 \pmod{d}}} \sum_{\substack{n \leq x \\ n \equiv l \pmod{d}\}}} 1 = \sum_{\substack{1 \leq l \leq d \\ h(l) \equiv 0 \pmod{d}}} (\frac{x}{d} + O(1)).$$

Hence

$$|A_d| = x\frac{\rho(d)}{d} + O(\rho(d)). \tag{1.1}$$

Now we define

$$S(A, P(y)) = |\{n \in A; 1 \leq n \leq x; (n, P(y)) = 1\}|.$$

In other words $S(A, P(y))$ is the number of integers in A not divisible by any prime less than y.

We can re-write $S(A, P(y))$ using a property of Möbius function μ:

$$\sum_{a \in A} \mu(d) = \begin{cases} 1 \text{ if } (a, P) = 1 \\ 0 \text{ otherwise .} \end{cases}$$

Hence

$$S(A, P) = \sum_{a \in A} \sum_{d | (A, P)} \mu(d) = \sum_{a \in A} \sum_{\substack{d | a \\ d | P}} \mu(d).$$

Interchanging the order of summation, we get the Legendre identity

$$S(A, P) = \sum_{d | P} \mu(d) |A_d|.$$

From (1.1) we get

$$S(A,P) = x\sum_{d|P}\frac{\mu(d)}{d}\rho(d) + O(\sum_{d|P}\rho(d)). \tag{1.2}$$

Recalling the following identity

$$\sum_{d|P}\frac{\mu(d)}{d}\rho(d) = \prod_{p\leq y}(1-\frac{\rho(d)}{p}),$$

we get

$$S(A,P(y)) = x\prod_{p\leq y}(1-\frac{\rho(d)}{p}) + O(\sum_{d|P}\rho(d)).$$

Our problem now is to get a suitable approximations to $\mu(d)$, based on the ideas developed by Atle Selberg [15].

1.3 Selberg sieve

From the previous manipulations we could estimate $S(A,P(y))$ the number of integers n less than a given bound x, for which $h(n)$ is not divisible by any prime (or a product of primes) less than a given y, with $h(n)\in\mathbb{Z}[X]$.

Let us write again as in (1.1)

$$|A_d| = x\frac{\rho(d)}{d} + R(d),$$

with $R(d)\leq\rho(d)$. So to get an upper bound on $S(A,P(y))$, we need a multiplicative function $\lambda(d)$, such that

$$\sum_{d|(n,P(y))}\mu(d) \leq \sum_{d|(n,P(y))}\lambda(d).$$

If the above inequality holds, (1.2) becomes

$$S(A,P) = x\sum_{d|P}\frac{\lambda(d)\rho(d)}{d} + \sum_{d|P(y)}\lambda(d)R(d). \tag{1.3}$$

Selberg's idea is to let Φ be a multiplicative function and define λ as

$$\sum_{d|(n,P(y))}\lambda(d) = (\sum_{d|(n,P(y))}\Phi(d))^2.$$

The right hand side in the previous equation is always greater than or equal to zero, and it is equal to one for $(n, P(y)) = 1$, so the equality (1.3) holds. Recalling that Φ is multiplicative, we define

$$\lambda(d) = \sum_{\substack{d_1,d_2|P(y)\\ d=[d_1,d_2]}} \Phi(d_1)\Phi(d_2).$$

It remains to find the optimal Φ, which means to get a main term in (1.3) as small as possible. Selberg sets for all $d|P(y)$

$$f(d) = \frac{d}{\rho(d)}. \tag{1.4}$$

For all $k|P(y)$, we have

$$g(k) = f(k)\prod_{p|k}(1 - \frac{\rho(p)}{p}) = f(k)\prod_{p|k}(1 - \frac{1}{f(p)}).$$

This way he proved the following theorem

Theorem 1.3.1. *Let*

$$Q = \sum_{d|P(y)} \frac{1}{g(d)} = \frac{\rho(d)}{d}\prod_{p|k}\left(1 - \frac{\rho(p)}{p}\right)^{-1}.$$

Let Φ be a multiplicative function, with $\Phi(p) = 0$ if $p|P(y)$, and

$$\lambda(d) = \sum_{\substack{d_1,d_2|P(y)\\ d=[d_1,d_2]}} \Phi(d_1)\Phi(d_2).$$

If $\lambda(d) = 0$ if $d \nmid P(y)$, then

$$\sum_{d|P} \frac{\lambda(d)\rho(d)}{d} \geq \frac{1}{Q}.$$

Moreover the previous inequality becomes an equality for

$$\Phi(d) = \frac{d\mu(d)}{Q\rho(d)}\sum_{t|d}\frac{1}{g(t)},$$

for all $d|P(y)$.

Proof. We have that if f is multiplicative and all the k's are square-free. Hence $d|k$ we have $(d, k/d) = 1$, and this implies

$$f(k) = f(d)f(k/d).$$

Then we get

$$g(k) = f(k)\prod_{p|k}(1 - \frac{1}{f(k)}) = \sum_{d|k} \mu(d)\frac{f(k)}{f(d)} = \sum_{d|k} \mu(d) f(k/d).$$

So $g = \mu * f$. Using the Möbius inversion formula we obtain

$$f(k) = \sum_{d|k} g(d).$$

We know also that $(d_1, d_2)[d_1, d_2] = d_1 d_2$, so

$$f([d_1, d_2]) = \frac{f(d_1) f(d_2)}{f((d_1, d_2))}.$$

Hence we can now with that write

$$\frac{1}{f([d_1, d_2])} = \frac{1}{f(d_1) f(d_2)} \sum_{d|(d_1,d_2)} g(d).$$

Now we have

$$\begin{aligned}
\sum_{d|P} \frac{\lambda(d)\rho(d)}{d} &= \sum_{d|P} \frac{\lambda(d)}{f(d)} = \sum_{d_1,d_2|P(y)} \frac{\Phi(d_1)\Phi(d_2)}{f([d_1, d_2])} \\
&= \sum_{d_1,d_2|P(y)} \frac{\Phi(d_1)\Phi(d_2)}{f(d_1) f(d_2)} \sum_{t|(d_1,d_2)} g(t) \\
&= \sum_{t|P(y)} g(t) \sum_{\substack{d_1|P(y) \\ d_2|P(y) \\ t|d_1 \\ t|d_2}} \frac{\Phi(d_1)\Phi(d_2)}{f(d_1) f(d_2)} \\
&= \sum_{t|P(y)} g(t) (\sum_{d|P(y)} \frac{\Phi(d)}{f(d)})^2
\end{aligned}$$

Now we introduce a change of variable

$$y(t) = \sum_{d|P(y)} \frac{\Phi(d)}{f(d)}.$$

From the equality (1.8) we have

$$\sum_{d|P} \frac{\lambda(d)\rho(d)}{d} = \sum_{t|P(y)} g(t) y(t)^2.$$

We see that

$$\sum_{d|t|P(y)} \mu(t/d) \sum_{d|t|P} \frac{\Phi(d)}{f(d)} = \frac{\Phi(d)}{f(d)}.$$

Then we get

$$\Phi(d) = f(d) \sum_{t|P(y)} \mu(t/d) y(t). \tag{1.5}$$

Remark 1.3.2. *One of the important steps in Selberg's combinatorial technique is to make an invertible (to be defined later) change of variable. We will see in Chapter III that Maynard's argument is a variation of this technique in higher dimension.*

For $d = 1$

$$\sum_{t|P(y)} \mu(t/d) y(t) = 1.$$

Then

$$\begin{aligned}
\sum_{d|P} \frac{\lambda(d)\rho(d)}{d} &= \sum_{t|P(y)} g(t) y(t)^2 \\
&= \sum_{t|P(y)} g(t) y(t)^2 - \frac{2}{Q} \sum_{t|P(y)} \mu(t/d) y(t) + \frac{1}{Q^2} \sum_{d|P(y)} \frac{\mu(t)^2}{g(d)} + \frac{1}{Q} \\
&= \sum_{t|P(y)} \frac{1}{g(t)} (g(t) y(t) - \frac{\mu(t)}{Q})^2 + \frac{1}{Q}
\end{aligned}$$

So finally we get

$$\sum_{d|P} \frac{\lambda(d)\rho(d)}{d} \geq \frac{1}{Q}.$$

The above inequality becomes an equality if

$$y(t) = \frac{\mu(t)}{Q g(t)}. \tag{1.6}$$

Hence from (1.5) and (1.6) we get

$$\Phi(d) = \frac{f(d)}{Q} \sum_{d|t|P(y)} \mu(t/d) \frac{\mu(t)}{g(t)} = \frac{f(d)\mu(d)}{Q} \sum_{d|t|P(y)} \frac{1}{g(t)}.$$

□

Now we are ready to prove the following theorem.

Theorem 1.3.3. *Under the assumptions of Theorem (1.3.1), we have*

$$S(a, P(y)) \leq \frac{x}{Q} + y^2 \prod_{p|P(y)} (1 - \frac{\rho(p)}{p})^{-2}.$$

Proof. We already proved in (1.3) that

$$S(A, P(y)) \leq x \sum_{d|P(y)} \frac{\lambda(d)\rho(d)}{d} + \sum_{d|P(y)} \lambda(d)R(d).$$

Recalling that

$$\sum_{d|P(y)} \frac{\lambda(d)\rho(d)}{d} = \frac{1}{Q},$$

it remains to estimate the error term

$$E = \sum_{d|P(y)} \lambda(d)R(d) = \sum_{d|P(y)} \Phi(d_1)\Phi(d_2)R[d_1d_2].$$

The optimal Φ by Selberg sieve is

$$\Phi(d) = \frac{f(d)\mu(d)}{Q} \sum_{t|d|P} \frac{1}{g(t)}.$$

Hence

$$|\Phi(d)| = \frac{f(d)}{Q} \sum_{t|d|P} \frac{1}{g(t)} \leq \frac{f(d)}{g(d)Q} \sum_{k|P(y)} \frac{1}{g(k)} \leq \frac{f(d)}{g(d)}.$$

since

$$|R([d_1d_2])| \leq \frac{[d_1d_2]}{f([d_1d_2])} \leq \frac{d_1}{f(d_1)} \frac{d_2}{f(d_2)}.$$

Finally

$$E \leq \sum_{d_1,d_2|P(y)} \frac{d_1d_2}{g(d_1)g(d_2)} (\sum_{d|P(y)} \frac{d}{g(d)})^2.$$

So we have

$$E \leq y^2Q^2.$$

$\square$

1.4 Application : twin primes

The twin primes conjecture states that there are infinitely many primes p such that $p+2$ is prime. One can think about a non trivial lower bound on the number of pairs of twin primes p, $p+2$ up to a given x, so if this bound involves a term that goes to infinity for x going to infinity then we will be done, but until writing these lines this seems to be far out of reach with current techniques. In this section we will give an upper bound on the number of twin pairs up to x.

We denote by $\pi_2(x)$ the cardinality of the following subset of primes p

$$\pi_2(x) = |\{p, p+2 \leq x | p, p+2 \quad \text{primes}\}|.$$

Using the prime number theorem "PNT" we can give a heuristic to estimate $\pi_2(x)$, as from the PNT the chance to pick randomly a pair of primes p, and $p+2$ in an interval of length x is $\frac{1}{\log x}$, if we assume that these two events are independent, we can expect that $\pi_2(x) \sim \frac{1}{\log x} . \frac{1}{\log x} = \frac{1}{(\log x)^2}$.

Clearly this is obviously false if we look to the trivial parity dependence between n and $n+2$ (if n even $\Rightarrow n+2$ even). For example to get a "correction" factor on the last non dependence, the probability that a random n is even is $1/2$, so the probability to choose independently two integers non divisible by 2 is $(1-\frac{1}{2})^2$. Then the correction factor for the divisibility by 2 is $\frac{(1-\frac{1}{2})}{(1-\frac{1}{2})^2} = 2$.

With the same argument the probability that a prime q does not divide p or $p+2$ is $(1-\frac{1}{p})^2$, and we need p and $p+2$ to be non-zero modulo q, so p could be in the $q-2$ residues classes mod q, then correction factor for the primes q grater than 2 is $\frac{(1-\frac{2}{q})}{(1-\frac{1}{q})^2}$. Using the Chinese Reminder Theorem we can expect to multiply the correction factors over all the primes. Indeed we should just do it for finitely many primes, but here it will not really affect the result as the expected error coming from large primes is very small. Finally we define the twin prime constant C_2, as

$$C_2 = 2 \prod_{q \geq 3} \frac{(1-\frac{2}{q})}{(1-\frac{1}{q})^2} \approx 1.32032363...$$

Now we can conjecture that

$$\pi_2(x) \sim C_2 \frac{x}{(\log x)^2},$$

as $x \to \infty$. From Selberg sieve we will prove the following result

Proposition 1.4.1. *With the same notations above we have*

$$\pi_2(x) \ll \frac{x}{(\log x)^2} + x^{1/3}.$$

Proof. We set

$$h(n) = n(n+2).$$

If $p < x$ is a twin prime then $p \le x^{1/3}$ or $h(p)$ has no prime factors less than $x^{1/3}$.

If we take $y = x^{1/3}$, we get

$$\pi_2(x) \le S(A, P(y)) + y.$$

By Selberg sieve we get

$$S(A, P(y)) \le x\Big(\sum_{d|P(y)} \frac{\rho(d)}{d} \prod_{p|d} \big(1 - \frac{\rho(p)}{p}\big)^{-1}\Big)^{-1} + y \prod_{p|d} \big(1 - \frac{\rho(p)}{p}\big)^{-2}.$$

First we remark that for $p = 2$, $\rho(p) = 1$, and $\rho(p) = 2$ otherwise.

We will now estimate $\prod_{p \le x^{1/3}} \big(1 - \frac{\rho(p)}{p}\big)^{-1}$. We remark that If $p > 5$ then

$$\big(1 - \frac{2}{p}\big)^{-1} \le \big(1 - \frac{1}{p}\big)^{-1}\big(1 - \frac{2}{p^2}\big)^{-1}.$$

But it's known that

$$\prod_p \big(1 - \frac{2}{p^2}\big)^{-1} \le \prod_p \big(1 - \frac{1}{p^2}\big)^{-1} = \zeta(2) = \frac{\pi^2}{6}.$$

Then

$$\prod_p \big(1 - \frac{2}{p^2}\big)^{-1} < \infty$$

and by Mertens estimate

$$\prod_{p \le x^{1/3}} \big(1 - \frac{1}{p^2}\big)^{-1} = e^{\gamma} \log(x^{1/3}) + O(1).$$

Finally we get

$$\prod_{p \le x^{1/3}} \big(1 - \frac{\rho(p)}{p}\big)^{-1} \ll (\log x)^2.$$

It remains to estimate the contribution of $\sum_{d\leq x^{1/3}} \frac{\rho(d)}{d}$. We will use the fact that

$$\rho(d) \leq d(d).$$

Letting $d(n)$ be the number of divisors of n, in fact if writing $n = p_1^{\alpha_1}...p_n^{\alpha_n}$. We obtain that

$$\rho(n) = 2^{\alpha_1}...2^{\alpha_n}, \quad \text{and} \quad d(n) = (\alpha_1+1)...(\alpha_n+1).$$

From the multiplicativity of $d(.)$, and $\rho(.)$, we finally obtain

$$\sum_{d\leq x^{1/3}} \frac{\rho(d)}{d} \geq \sum_{d\leq x^{1/3}} \frac{d(d)}{d} \geq \Big(\sum_{d\leq x^{1/3}} \frac{1}{d}\Big)^2 \gg (\log x)^2.$$

So we can finally write that

$$\pi_2(x) \ll \frac{x}{\log(x)^2}.$$

□

Indeed we are still far from proving the twin primes conjecture, and that leads number theorists to investigate a more general situation of the twin prime conjecture, in the next section we will present some of the spectacular results on this problem proved by Goldston, Pintz and Yildirim.

Chapter 2

The work of Goldston, Pintz and Yildirim

2.1 Primes in tuples

"This conjecture (2.1.2) is extremely difficult (containing the twin prime conjecture, for instance, as a special case), and in fact there is no explicitly known example of an admissible k-tuple with $k \geq 2$ for which we can verify this conjecture"

Terence Tao

Let us define a $k-$tuple $\mathcal{H} = (h_1, ..., h_k)$ as a collection of increasing positive integers. Our aim here is to study the case when set $n+\mathcal{H} = \{n+h_1, ..., n+h_k\}$, the translates of $\mathcal{H}$, consists entirely of primes. Obviously the case $\mathcal{H} = (0)$ is Euclid's theorem, and $\mathcal{H} = (0, 2)$ is the twin primes conjecture, to study the general case we should add another condition on the $k-$tuple, which is the admissibility.

Definition 2.1.1. *We said that a $k-$tuple $\mathcal{H} = (h_1, ..., h_k)$ is admissible if the h_i with $1 \leq i \leq k$ avoid at least one congruence class mod every prime.*

In fact if the h_i covers all the congruence classes modulo some prime p, at least one of the elements of $n + \mathcal{H} = \{n + h_1 ... n + h_k\}$ will be divisible by p for every n.

Now we can state the so called Hardy-Littlewood conjecture.

Conjecture 2.1.2. *If $\mathcal{H}$ is an admissible $k-$tuple, then there exists infinitely many translates of $\mathcal{H}$ that consist entirely of primes.*

As in the heuristic discussion on the twin primes conjecture up to a given x the probability to pick a k-tuple of primes is $\frac{1}{(\log x)^k}$, if we assume the independence between the k events, and similarly we construct a correction factor

$$\mathcal{G}(\mathcal{H}) = \prod_p \frac{\left(1 - \frac{\rho(p)}{p}\right)}{(1 - \frac{1}{p})^k}. \tag{2.1}$$

With $h(n) = \prod_{1 \leq i \leq k}(n + h_i)$ we define $\rho(p)$ as the number of solutions of

$$h(n) \equiv 0 \pmod{p}.$$

In the following we will call the quantity $\mathcal{G}(\mathcal{H})$ in (2.1) the singular series of this problem.

Assuming $\mathcal{H}$ admissible implies that $\rho(p) < p$ and for large p we have $\rho(p) = k$ so that gives the non-vanishing of $\mathcal{G}(\mathcal{H})$.

We then have the quantitative form of the k-tuples conjecture.

Conjecture 2.1.3. *Let $\mathcal{H} = (h_1, ..., h_k)$ be an admissible k-tuple then*

$$|\{n \leq x, n+h_1, ..., n+h_k \textit{ such that } n+h_1, ..., n+h_k \textit{ are all primes } \}| \sim \mathcal{G}(\mathcal{H}) \frac{x}{(\log x)^k} \qquad x \to \infty.$$

To work on the GPY method, we will need an assortment of tools on the primes in arithmetic progressions.

2.2 Primes in arithmetic progressions

As we mentioned in the introduction, Dirichlet in 1837, using the theory of L-functions, proved that that the polynomial $ax + b$ contains infinitely many primes for a and b co-primes, on showing that the series

$$\sum_{\substack{p \equiv a \pmod{m} \\ (a,m)=1}} \frac{1}{p^s},$$

is divergent for $s \to 1^+$. 59 years after Dirichlet, Hadamard and de la Valee-Poussin proved the Prime Number Theorem, namely

$$\pi(x) \sim \frac{x}{\log x},$$

as $x \to \infty$.

Taking

$$\vartheta(x) = \sum_{\substack{p \text{ prime} \\ p \leq x}} \log p \quad \text{and} \quad \vartheta(x; a, m) = \sum_{\substack{p \text{ prime} \\ p \equiv a \pmod{m} \\ p \leq x}} \log p,$$

one can prove that the prime number theorem is equivalent to $\vartheta(x) \sim x$ as $x \longrightarrow \infty$, so we can expect that for $x \to \infty$

$$\vartheta(x; a, m) \sim \frac{x}{\phi(m)}.$$

If this will hold for all $(a, m) = 1$ and $m \to \infty$, we will say that the primes are (more or less) equi-distributed amongst the arithmetic progression $a \pmod{m}$.

One of the best known results on the distribution of primes is the Siegel-Walfisz theorem, which gives the equi-distribution once $x \geq e^{m^\epsilon}$. In general this is a limitation in applying this theorem as we need x to be very large comparing to m. We can state Siegel-Walfisz in the following form.

Theorem 2.2.1. *For some $c > 0$, and for all $(a, m) = 1$ we have*

$$\vartheta(x; a, m) = \frac{x}{\phi(m)} + O\Big(x \exp(-c\sqrt{\log x})\Big),$$

for $x \geq e^{m^\epsilon}$.

Assuming the Generalized Riemann Hypothesis (GRH) one can prove that (see Corollary 13.8 in [11])

$$\vartheta(x; a, m) = \frac{x}{\phi(m)} + O\Big(x^{1/2}(\log x)^2\Big),$$

for $x \geq e^{m^\epsilon}$.

But in many applications we don't need that the equi-distribution holds for all a and m, but just for m up to some Q. The best unconditional result in this context is the Bombieri-Vinogradov theorem.

Theorem 2.2.2. *(Bombieri-Vinogradov)*

For all $A > 0$ there exists a constant $B = B(A)$, such that

$$\sum_{m \leq Q} \max_{\substack{a \bmod m \\ (a,m)=1}} |\vartheta(x; a, m) - \frac{x}{\phi(m)}| \ll_A \frac{x}{(\log x)^A},$$

where $Q = x^{1/2}/(\log x)^B$.

We have the following definition.

Definition 2.2.3. *Using the same notation as in (2.2.2) we say that the primes have a level of distribution* θ *if* $Q = x^{\theta-\epsilon}$ *for all* $\epsilon > 0$.

By the Bombieri-Vinogradov theorem, the primes have a level of distribution $\theta = 1/2$. Eliott-Halberstam (1968) conjectured that $\theta = 1$.

Conjecture 2.2.4. *(Eliott-Halberstam) The primes have a level of distribution* $\theta = 1$.

Remark 2.2.5. *In April* 2013, *Yitang Zhang [21] proved a weaker version of Elliott-Halberstam conjecture when restricting to one particular residue class, considering* m *to be squarefree and* $y-$*smooth integer (that is, when all the prime factors of* q *are less than* y*), and he proved that*

$$\sum_{\substack{m \leq Q \\ (a,m)=1 \\ m \text{ is } y\text{-smooth} \\ m \text{ squarefree}}} |\vartheta(x; a, m) - \frac{x}{\phi(m)}| \ll_A \frac{x}{(\log x)^A}.$$

For $Q = x^{1/2+\eta}$ *and* $y = x^{\delta}$, *for all* $\eta, \delta > 0$. *And that was the key estimate in his breakthrough on the bounded gaps between primes, namely the fact that there exist a calculable constant* B, *such that there exist infinitely many pairs of primes which differ by no more than* B, *and he even showed that we can take* $B = 70\ 000\ 000$.

2.3 The work of Goldston Pintz and Yildirim

Indeed to prove the bounded gaps between infinitely many consecutive primes, it's sufficient to find a suitable $k-$tuple $\mathcal{H} = (h_1, ..., h_k)$, such that $n + \mathcal{H}$ contains at least two primes for infinitely many values of n. In other words we have

$$\lim_{n\to\infty} inf(p_{n+1} - p_n) \leq |h_k - h_1|. \tag{2.2}$$

For a long time equation (2.2) seemed to be out of reach, and here raised the problem of proving the existence of infinitely many "short" intervals containing consecutive primes. The Prime Number Theorem tells us that in average the length of such intervals is $C \log x$ where C is a positive constant. In order to be allowed to choose any "small" positive constant C, one should prove that $\Delta = 0$ where

$$\Delta = \lim_{n\to\infty} \inf \frac{p_{n+1} - p_n}{\log p_n}.$$

The first result on the topic is due to Hardy and Littlewood (1926): in fact they proved that $\Delta \leq 3/2$ with a conditional proof assuming GRH. This last bound was improved by many specialists : Erdos showed that $\Delta \leq 1-c$ with c a calculable constant, Bombieri and Davenport $\Delta \leq \frac{2+\sqrt{3}}{8}$, Huxley $\Delta \leq 0.4394$, and Maier $\Delta \leq 0.2484$.

In 2005 Goldston, Pintz and Yildirim developed a new sieve method to prove the following theorem.

Theorem 2.3.1. *Let*

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

We have

$$\Delta_\nu = \max(\nu - 2\theta, 0),$$

Where θ is the level of distribution of primes. Taking $\theta = \frac{1}{2}$, and $\nu = 1$ it follows that $\Delta = 0$.

Theorem 2.3.2. *Assume the primes have level of distribution $\theta \geq 1/2$, then there exist an explicitly calculable constant $C(\theta)$, such that any admissible $k-$tuple with $k \geq C(\theta)$ contains at least two primes infinitely often.*

Specifically, if $\theta > 20/21$, then this is true for $k \geq 7$, and, since the 7-tuple $(n, n+2, n+6, n+8, n+12, n+18, n+20)$ is admissible then, the following corollary is an immediate consequence of the previous theorem.

Corollary 2.3.3. *The Elliott-Halberstam conjecture implies that*

$$\lim\inf(p_{n+1} - p_n) \leq 20.$$

Let us define the following function

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

Letting $\mathcal{H} = (h_1, ..., h_k)$ be an admissible $k-$tuple, to count the primes in the translates of $\mathcal{H}$ we consider the following sum

$$S = \sum_{x<n\leq 2x} \Big(\sum_{i=1}^{k} (\theta(n+h_i) - \log(3x)) \Big) \mathcal{W}_n, \tag{2.3}$$

where $\mathcal{W}_n$ is a non-negative weight.

From the simple fact that $x < n \leq 2x$ implies that $n + h_i < 2x + h_k < 2x + x$ for any $1 \leq i \leq k$ and large x, we obtain that $\theta(n+h_i) < \log(3x)$ for $i \in \{1, \ldots, k\}$.

If we prove that $S > 0$, then there exist at least two different elements h_i and h_j in $\mathcal{H}$, such that $n + h_i$ and $n + h_j$ are primes. The first difficulty in this method comes from choosing a suitable $\mathcal{W}_n$ to evaluate (2.3).

2.4 On the weight $\mathcal{W}_n$

Our aim in this section is to find a positive weight which is sensitive to the prime k-tuples. The first propriety we need is that $\mathcal{W}_n$ has to vanish on the integers that have more than k prime factors. Moreover from (2.3) we see that to get $S > 0$ we should consider $\mathcal{W}_n$ that maximizes the quantity

$$\kappa = \frac{1}{\log(3x)} \frac{A_\theta(n)}{A(n)},$$

where

$$A(n) = \sum_{x<n\leq 2x} \mathcal{W}_n \qquad \text{and} \qquad A_\theta(n) = \sum_{x<n\leq 2x} (\sum_{i=1}^{k} \theta(n+h_i) \mathcal{W}_n.$$

We define

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

where $\Lambda(n)$ the von Mangoldt function defined as

$$\Lambda(n) = \begin{cases} \log(n) & \text{if } n = p^m \text{ for } m \geq 1, p \text{ prime} \\ 0 & \text{otherwise.} \end{cases}$$

We recall the following result on von Mangoldt function.

Proposition 2.4.1. *For $n \geq 1$ we have*

$$\log n = \sum_{d|n} \Lambda(d).$$

The above proposition follows naturally from the fundamental theorem of arithmetic. We will now use some results on the convolution of two arithmetic functions ([19]).

Definition 2.4.2. *Let f and g be two arithmetic functions (i.e a real or complex valued function defined on the set of positive integers). We define $f * g$, the Dirichlet convolution of f and g, by*

$$(f * g)(n) = \sum_{d|n} f(d)g(\frac{n}{d}).$$

Remark 2.4.3.

- *The set of arithmetic functions forms a commutative ring under the usual addition and Dirichlet convolution.*

- *the function* $\log$ *defines a derivative on the ring R of arithmetic functions as an endomorphism of the additive group of R satisfying the Leibniz rule* $\log(n).(f * g)(n) = f * (\log .g)(n) + g * (\log .f)(n)$

We can write 2.4.1 as $\Lambda = \mu * \log$, so we get the following identity as a direct application of Möbius inversion formula, namely the fact that if $g = f * 1$ then $f = g * \mu$ with μ the Möbius function:

$$\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d}. \tag{2.4}$$

By definition, the von Mangoldt function detects only prime powers, and from (2.4) we can study directly $\Lambda(n)$ as a sum of arithmetic functions.

We define the generalized Mangold function for a positive integer k as

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

Note that

$$\Lambda_0(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let's investigate the case $k = 2$.

Proposition 2.4.4. *For all positive integers n we have*

$$\Lambda_2(n) = \Lambda(n)\log(n) + \Lambda * \Lambda(n).$$

Proof. We have

$$\begin{aligned}\log^2 &= \log.\log\\ &= \log\times(1*\Lambda)\\ &= 1*(\log\times\Lambda)+(\log\times 1)*\Lambda \qquad \text{we used the fact that log is a derivative}\\ &= 1*(\Lambda\times\log)+\log*\Lambda.\end{aligned}$$

Then finally we get

$$\Lambda_2=\mu*\log^2=\mu*1*(\Lambda\times\log)+\mu*\log*\Lambda=\Lambda\times\log+\Lambda*\Lambda.$$

This proves the proposition. □

From (2.4.4), it is easy to see that

$$\Lambda_2(n)=\begin{cases}(2m-1)(\log p)^2 & \text{if } n=p^m\\ 2\log p\log q & \text{if } n=p^aq^b \text{ for } p\neq q\\ 0 & \text{otherwise .}\end{cases}$$

This shows that $\Lambda_2(n)$ is non-zero on the integers n that has at most two prime factors. Indeed we have the general recurrent relation.

Proposition 2.4.5. *For all positive integer k we have*

$$\Lambda_{k+1}=\Lambda_k\log+\Lambda*\Lambda_k.$$

Proof. By definition

$$\begin{aligned}\Lambda_{k+1} &= \mu*(\log^k.\log)\\ &= (\mu*\log^k)\log+(-\mu\log)*1*\Lambda_k \qquad \text{from } \log(n/d)=\log(n)-\log(d)\\ &= \Lambda_k\log+\Lambda*\Lambda_k, \qquad \text{from } \Lambda=(-\mu\log)*1.\end{aligned}$$

□

Proposition 2.4.6. *Let k be a positive integer then*

$$\Lambda_k(n)=\begin{cases}(\alpha^k+(\alpha-1)^k)(\log p)^k & \text{if } n=p^\alpha,\\ k!(\log p_1)(\log p_2)...(\log p_k) & \text{if } n \text{ has } k \text{ distinct prime factors } p_i,\\ 0 & n \text{ has more than } k \text{ distinct prime factors .}\end{cases}$$

Proof. • If $n = p^\alpha$ then by definition

$$\Lambda_k(n) = \mu(1)(\log p^\alpha)^k + \mu(p)(\log p^{\alpha-1})^k = (\alpha^k - (1-\alpha)^k)(\log p)^k.$$

• Let $n = \prod_{i=1}^m p_i^{\alpha_i}$ We assume that

$$\Lambda_k(n) = \begin{cases} k!(\log p_1)(\log p_2)...(\log p_k) & \text{if } n \text{ has } k \text{ distinct prime factors } p_i \\ 0 & n \text{ has more than } k \text{ factors} \end{cases} \tag{2.5}$$

which is true for $k = 0$. Now we prove by induction that (2.5) holds for $k+1$ too, i.e.,

$$\Lambda_{k+1}(n) = (k+1)!(\log p_1)(\log p_2)...(\log p_{k+1})$$

From (2.4.5) and the inductive hypothesis we have

$$\begin{aligned} \Lambda_{k+1}(n) &= \Lambda * \Lambda_k(n) \\ &= \sum_{d|n} \Lambda(d)\Lambda_k(n/d) \\ &= (\log p_1)k! \prod_{i\neq 1} \log(p_i) + (\log p_2)k! \prod_{i\neq 2} \log(p_i) + ... + (\log p_{k+1})k! \prod_{i\neq k+1} \log(p_i) \\ &= (k+1)!(\log p_1)(\log p_2)...(\log p_{k+1}). \end{aligned}$$

□

Then the generalized Mangolt function is non zero (i.e., supported) on the integers that has at most k prime factors.

Assume now that $h(n)$ has r prime factors $p_1, p_2, ..., p_r$ with $r < k$. Then there exists an element $n + h_j \in \{n + h_1, \dots, n + h_k\}$ such that for all $p_i^{\alpha_i} || n + h_j$ there exists some other element $n + h_{j'} \in \{n + h_1 ... n + h_k\}$ with $p_i^{\alpha_i} || n + h_j$, hence

$$p^{\alpha\alpha_i} | n + h_j - n + h_{j'} = h_j - h_{j'}.$$

This holds for all the prime factors of $n + h_j$, and that implies that $n + h_j | \prod_{\substack{1\leq i\leq k \\ i\neq j}} (h_j - h_i)$. Then in this case $n < n + h_j \leq h_k^{k-1}$. From this argument and (2.4.6) we conclude that for $n > h_k^{k-1}$ if $\Lambda_k(h(n)) \neq 0$. So, $h(n)$ has exactly k distinct factors.

We define the truncated divisor sum as

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

and

$$\Lambda_k(n;\mathcal{H}) = \frac{1}{k!}\Lambda_k(h(n)),$$

where $\{h_1,\dots,h_k\}$ is an admissible k-tuple, and

$$h(n) = (n+h_1)(n+h_2)...(n+h_k).$$

We can approximate $\Lambda_k(n;\mathcal{H})$ by the truncated sum

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

in which we divide by $k!$ to simplify the estimates.

We recall that our aim is to prove that $S > 0$ for

$$S = \sum_{x<n\leq 2x} \Big(\sum_{i=1}^{k} (\theta(n+h_i) - \log(3x)) \Big) \mathcal{W}_n. \tag{2.6}$$

Inspired from Selberg's work we take

$$\mathcal{W}_n = \Big(\sum_{\substack{d\leq R \\ d|h(n)}} \lambda(d) \Big)^2,$$

and we look for a function λ which maximizes the quantity

$$\kappa = \frac{1}{\log 3x} \frac{A_\theta(n)}{A(n)}.$$

where

$$A(n) = \sum_{x<n\leq 2x} \mathcal{W}_n \qquad \text{and} \qquad A_\theta(n) = \sum_{x<n\leq 2x} (\sum_{i=1}^{k} \theta(n+h_i)\mathcal{W}.$$

In [17] Soundrarajan showed that for more general family of weights and, in particular for $\Lambda_R(n;\mathcal{H})$, we can not unconditionally achieve the bounded gaps between primes. Indeed he proved that $\kappa < 1$ if the level of distribution of primes is $\frac{1}{2}$.

The idea behind the success of GPY sieve is modifying the classical k-tuples detecting weights.

Letting $\nu(n)$ be the number of distinct prime factors of n, we have $\nu(h(n)) = k + l$, where $0 \leq l < k$. Remark that if $l = k$, then $\mathcal{H}$ will not be admissible. By the pigeon-hole principle, we can conclude that there exist $k-l$ primes among $n+h_1, \ldots, n+h_k$. Summing up, we define

$$\Lambda_R(n;\mathcal{H},l) = \frac{1}{(k+l)!} \sum_{\substack{d|h(n) \\ d\leq R}} \mu(d)\big(\log \frac{R}{d}\big)^{k+l},$$

and to prove $\Delta = 0$ our $k-$tuples detecting weight will be

$$\mathcal{W}_n = \Lambda_R(n;\mathcal{H},l)^2.$$

2.4.7 Outline of the GPY method

In order to motivate the next sections we will discuss the general setting of the GPY method, taking for instance the weight $\mathcal{W}_n$ to be

$$\mathcal{W}_n = \Big(\sum_{\substack{d\leq R \\ d|h(n)}} \lambda(d)\Big)^2,$$

for some $R > 0$ and

$$h(n) = (n+h_1)(n+h_2)...(n+h_k),$$

where $\lambda(.)$ is a multiplicative function non zero only on the positive square-free integers less than R.

Arguing analogously to Selberg's method, and expanding (2.6), we obtain

$$\begin{aligned} S &= \sum_{x<n\leq 2x} \sum_{i=1}^{k} (\theta(n+h_i)\Big(\sum_{\substack{d|h(n) \\ d\leq R}} \lambda(d)\Big)^2 - \log(3x) \sum_{x<n\leq 2x} \Big(\sum_{\substack{d|h(n) \\ d\leq R}} \lambda(d)\Big)^2 \\ &= \sum_{\substack{d_1,d_2\leq R \\ D=[d_1,d_2]}} \lambda(d_1)\lambda(d_2)\Big(\sum_{i=1}^{k} \sum_{\substack{x<n\leq 2x \\ D|h(n)}} (\theta(n+h_i) - \log(3x) \sum_{\substack{x<n\leq 2x \\ D|h(n)}} 1\Big). \end{aligned}$$

Then

$$\log(3x) \sum_{\substack{x<n\leq 2x \\ D|h(n)}} 1 = \log(3x)\Big(x\frac{\rho(D)}{D} + O(\rho(D))\Big).$$

By definition of D, we have $D = [d_1, d_2] \leq d_1 d_2 \leq R^2$, and hence we can choose $R \leq x^{1/2+o(1)}$ to be able to have a good final estimate for the error term.

We evaluate the sum on $\theta(n + h_i)$ analogously except that here we insert an additional condition, which is $(D, n + h_i) = 1$ as θ is zero whenever $n + h_i$ is not a prime.

Letting

$$\rho_i^*(D) = \{m \in \mathbb{Z}/D\mathbb{Z}, \text{ such that } D|h(n) \text{ and } (D, m + h_i) = 1\},$$

we have then

$$\sum_{i=1}^{k} \sum_{\substack{x<n\leq 2x \\ D|h(n)}} \theta(n + h_i) = \sum_{i=1}^{k} \sum_{m\in\rho^*(D)} \sum_{\substack{x<n\leq 2x \\ n\equiv m \pmod{D}}} \theta(n + h_i).$$

Remarking that $|\rho_i^*(p)| = \rho(p) - 1$ for p prime, we have now to evaluate a sum on primes over an arithmetic progression.

By the Siegel-Walfisz theorem, we can get the estimate

$$\sum_{\substack{x<n\leq 2x \\ n\equiv m \pmod{D}}} \theta(n + h_i) \sim \frac{x}{\phi(D)},$$

as $x \to \infty$.

As we mentioned before the best available result that can allow us to control the error term is the Bombieri-Vinogradov theorem. To get an unconditional result we should assume that $D < x^{1/2-o(1)}$.

Remark 2.4.8. *Again from $D = [d_1, d_2] \leq d_1 d_2 \leq R^2$, the condition $D < x^{1/2-o(1)}$ forces us to choose $R < x^{1/4-o(1)}$, and we will prove that exceeding the barrier $\frac{1}{4}$ implies the bounded gaps between primes.*

So an application of the Bombieri-Vinogradov theorem gives us

$$\sum_{i=1}^{k} \sum_{\substack{x<n\leq 2x \\ D|h(n)}} \theta(n + h_i) \ll xk \frac{|\rho^*(D)|}{\phi(D)}.$$

Hence S has a main term M_s of the form

$$M_s = x\Big(k \sum_{\substack{d_1,d_2\leq R \\ D=[d_1,d_2]}} \lambda(d_1)\lambda(d_2) \frac{|\rho^*(D)|}{\phi(D)} - \log(3x) \sum_{\substack{d_1,d_2\leq R \\ D=[d_1,d_2]}} \lambda(d_1)\lambda(d_2) \frac{\rho(D)}{D}\Big). \tag{2.7}$$

Remark 2.4.9. *The term between the parenthesis in M_s is negative if $\lambda(.)$ is positive; it means that we can't choose $\lambda(.)$ to be positive for all d.*

Till now we used just combinatorial arguments to deal with our sums but to estimate the two sums in M_s it will be more efficient to introduce another technique. As we expect that these sums will involve the Möbius function which produces a lot of different terms of opposite signs which makes the combinatorial arguments more complicated. Surprisingly Selberg introduced a combinatorial argument to deal with such sums, but Goldston, Pintz and Yildirim transformed the them into suitable integrals using the following discontinuous factor known as Perron's formula ([2]):

$$\frac{1}{2i\pi}\int_{Re(s)=2}\frac{y^s}{s}ds = \begin{cases} 1 & \text{if } y > 1, \\ 1/2 & \text{if } y = 1, \\ 0 & \text{if } 0 < y < 1. \end{cases}$$

Let

$$M_{1,s} = k\sum_{\substack{d_1,d_2\leq R\\ D=[d_1,d_2]}}\lambda(d_1)\lambda(d_2)\frac{|\rho(D)^*|}{\phi(D)},$$

and

$$M_{2,s} = \sum_{\substack{d_1,d_2\leq R\\ D=[d_1,d_2]}}\lambda(d_1)\lambda(d_2)\frac{\rho(D)}{D}.$$

We will just study $M_{2,s}$ since $M_{1,s}$ can be studied in a similar way. To evaluate this sum with the condition $d < R$ we take $y = R/d$ in Perron's formula thus obtaining

$$M_{2,s} = \frac{1}{(2i\pi)^2}\int_{\substack{Re(s_1)=2\\ Re(s_2)=2}}\Big(\sum_{\substack{d_1,d_2\geq 1\\ D=[d_1,d_2]}}\frac{\lambda(d_1)\lambda(d_2)}{d_1^{s_1}d_2^{s_2}}\frac{\rho(D)}{D}\Big)R^{s_1+s_2}\frac{ds_1}{s_1}\frac{ds_2}{s_2}. \tag{2.8}$$

Obviously, the evaluation of the previous integral depends on the nature of the function $\lambda(.)$. In the next section we will show that the best choice for λ is $\lambda(d) = \frac{1}{(k+l)!}\mu(d)\big(\log\frac{R}{d}\big)^{k+l}$. To motivate the next sections we will assume for instance that $\lambda(d) = \mu(d)$, then we can write

$$\sum_{\substack{d_1,d_2\geq 1\\ D=[d_1,d_2]}}\frac{\lambda(d_1)\lambda(d_2)}{d_1^{s_1}d_2^{s_2}}\frac{\rho(D)}{D} = \sum_{\substack{d_1,d_2\geq 1\\ D=[d_1,d_2]}}\frac{\mu(d_1)\mu(d_2)}{d_1^{s_1}d_2^{s_2}}\frac{\rho(D)}{D}.$$

Clearly the function $\frac{\mu(d_1)\mu(d_2)}{d_1^{s_1}d_2^{s_2}}\frac{\rho(D)}{D}$ is multiplicative and supported on the square-free integers. Hence

$$\sum_{\substack{d_1,d_2\geq 1\\ D=[d_1,d_2]}} \frac{\mu(d_1)\mu(d_2)}{d_1^{s_1}d_2^{s_2}}\frac{\rho(D)}{D} = \prod_{p \text{ prime}} \Big(1 - \frac{\rho(p)}{p^{s_1+1}} - \frac{\rho(p)}{p^{s_2+1}} + \frac{\rho(p)}{p^{s_1+s_2+1}}\Big).$$

Using the Euler product formula for the Riemann zeta-function, we multiply by a factor $F(s_1, s_2) = 1$. To evaluate the integral using our understanding of the Riemann zeta function, we consider

$$F(s_1,s_2) = \frac{\zeta(1+s_1+s_2)^k}{\zeta(1+s_1)^k\zeta(1+s_2)^k}\prod_{p \text{ prime}}\Big(1-\frac{1}{p^{1+s_1+s_2}}\Big)^k\Big(1-\frac{1}{p^{1+s_1}}\Big)^{-k}\Big(1-\frac{1}{p^{1+s_2}}\Big)^{-k}.$$

Then, finally, we obtain

$$M_{2,s} = \frac{1}{(2i\pi)^2}\int_{\substack{Re(s_1)=2\\ Re(s_2)=2}} \frac{\zeta(1+s_1+s_2)^k}{\zeta(1+s_1)^k\zeta(1+s_2)^k} G(s_1,s_2) R^{s_1+s_2}\frac{ds_1}{s_1}\frac{ds_2}{s_2},$$

where

$$G(s_1,s_2) = \prod_{p \text{ prime}}\Big(1-\frac{1}{p^{1+s_1+s_2}}\Big)^k\Big(1-\frac{1}{p^{1+s_1}}\Big)^{-k}\Big(1-\frac{1}{p^{1+s_2}}\Big)^{-k}\Big(1 - \frac{\rho(p)}{p^{s_1+1}} - \frac{\rho(p)}{p^{s_2+1}} + \frac{\rho(p)}{p^{s_1+s_2+1}}\Big).$$

Remark that under the hypothesis of admissibility, we have

$$G(0,0) = \mathcal{G}(\mathcal{H}) \neq 0.$$

We can easily see that $\rho(p) = k$ if $p > h_k$, and that $\rho(p) < k$ if there exists some $h_j < h_i$ such that $h_j \equiv h_i \pmod p$. This holds if $p|h_j - h_i$ for $i < j$, then we get also $\rho(p) = k$ if $p \nmid V$ where

$$V = \prod_{1\leq i<j\leq k} |h_j - h_i|$$

Then in order to evaluate $G(s_1, s_2)$, we will fix an upper bound U for $\log V$: using the trivial bound $V \leq h_k^{k^2}$, we can choose

$$U = Ck^2\log(2h_k). \tag{2.9}$$

Remark 2.4.10. *We see that V is the absolute value of a Vandermonde's determinant, so one can use Hadamard's inequality to bound V, namely $V \leq k^{k/2}h_k^k$.*

We take

$$G(s_1,s_2) \ll \prod_{p \text{ prime}} \Big(1-\frac{1}{p^{1+s_1}}\Big)^{-k}\Big(1-\frac{1}{p^{1+s_2}}\Big)^{-k}\Big(1-\frac{\rho(p)}{p^{s_1+1}}\Big)\Big(1-\frac{\rho(p)}{p^{s_2+1}}\Big).$$

Then we evaluate the products separately, taking for $s_i=\sigma_i+it_i$, $-1/4\le\sigma_i\le 1$, $t_i\in\mathbb{R}$ and $\delta_i=\max(-\sigma_i,0)$ for $i\in\{1,2\}$. Recalling that $\rho(p)$ is at most k, we obtain

$$\begin{aligned}
|\prod_{p\le U}\Big(1-\frac{\rho(p)}{p^{1+s_1}}\Big)| &\le \prod_{p\le U}\Big(1+\frac{\rho(p)}{p^{1-\delta}}\Big) = \exp\Big(\sum_{p\le U}\log\Big(1+\frac{k}{p^{1-\delta}}\Big)\Big)\\
&\le \exp\Big(\sum_{p\le U}\frac{k}{p^{1-\delta}}\Big) \quad (\text{ from } \log(x+1)\le x \text{ for } x\ge 0)\\
&\le \exp\Big(kU^{\delta}\sum_{p\le U}\frac{1}{p}\Big) \ll \exp(kU^{\delta}\log\log U) \quad (\text{ by Mertens's estimate})\\
&\ll \exp(kU^{\delta}\log\log x) \qquad (U\le x).
\end{aligned}$$

By the inequality $(1-x)^{-1}\le 1+3x$ for $0\le x\le 2/3$, we have $(1-\frac{1}{p^{1-\delta}})\le(1+\frac{3}{p^{1-\delta}})$, so that, arguing analogously to the previous case, we get

$$\Big|\prod_{p\le U}\Big(1-\frac{1}{p^{1+s_1}}\Big)^{-k}\Big| \le \Big(\prod_{p\le U}\Big(1-\frac{1}{p^{1-\delta}}\Big)^{-1}\Big)^{k} \ll \exp(3kU^{\delta}\log\log x).$$

Similarly we prove that exists $\beta_0>0$ such that

$$\prod_{p\le U}\Big(1-\frac{1}{p^{1+s_2}}\Big)^{-k}\Big(1-\frac{\rho(p)}{p^{s_2+1}}\Big) \ll \exp(\beta_0 kU^{\delta}\log\log x).$$

Then we find for $p\le U$ that

$$G(s_1,s_2) \ll \exp(\beta_1 kU^{\delta_1+\delta_2}\log\log x), \tag{2.10}$$

where $\beta_1>0$ is an absolute constant.

Now we prove that the upper bound in (2.12) holds also for the case $p>U$. Remark first that in this case we have $|\frac{k}{p^{s+1}}|\le\frac{k}{U^{1-\delta}}\le\frac{1}{2}$. Then we consider the case in which $p|V$ but the primes p will be replaced by the integers between U and $2U$ as we have just few divisors $p|V$.

For the case $p \nmid V$ we have $\rho(p) = k$ and we will use the Taylor formula for the logarithm. Hence

$$\begin{aligned}
\prod_{\substack{p\geq U\\ p|V}}\Big(1-\frac{1}{p^{1+s_1}}\Big)^{-k}\Big(1-\frac{\rho(p)}{p^{s_1+1}}\Big) &\leq \prod_{\substack{p\geq U\\ p|V}}\Big(1-\frac{3}{p^{1-\delta}}\Big)^{k}\Big(1+\frac{k}{p^{1-\delta}}\Big)\\
&\leq \exp\Big(\sum_{\substack{p\geq U\\ p|V}}\frac{4k}{p^{1-\delta}}\Big)\leq \exp\Big(4k\sum_{U<n\leq 2U}\frac{1}{n^{1-\delta}}\Big)\\
&\leq \exp\Big(4k(2U)^{\delta}\sum_{U<n\leq 2U}\frac{1}{n}\Big)\leq \exp(4kU^{\delta})
\end{aligned}$$

and

$$\begin{aligned}
\Big|\prod_{\substack{p\geq U\\ p\nmid V}}\Big(1-\frac{1}{p^{1+s_1}}\Big)^{-k}\Big(1-\frac{\rho(p)}{p^{s_1+1}}\Big)\Big| &= \Big|\exp\Big(\sum_{\substack{p\geq U\\ p\nmid V}}\Big(-\sum_{m=1}^{\infty}\frac{1}{m}\Big(\frac{k}{p^{1+s_1}}\Big)^{m}+k\sum_{m=1}^{\infty}\frac{1}{m}\Big(\frac{1}{p^{1+s_1}}\Big)^{m}\Big)\Big)\Big|\\
&\leq \exp\Big(\sum_{\substack{p\geq U\\ p\nmid V}}\sum_{m=2}^{\infty}\frac{2}{m}\Big(\frac{k}{p^{1-\delta}}\Big)^{m}\Big)\\
&\leq \exp\Big(\sum_{\substack{p\geq U\\ p\nmid V}}\sum_{m=2}^{\infty}\Big(\frac{k}{p^{1-\delta}}\Big)^{m}\Big)\leq \exp\Big(2k^2\sum_{n>U}\frac{1}{n^{2-2\delta}}\Big)\\
&\leq \exp\Big(\frac{4k^2U^{\delta}}{U^{1-\delta}}\Big)\leq \exp(2kU^{\delta}).
\end{aligned}$$

Then for $p > U$ we have

$$G(s_1,s_2) \ll \exp(\beta_2 kU^{\delta}), \tag{2.11}$$

where $\beta_2 > 0$ is an absolute constant.

Finally, from (2.12) and (2.11) we conclude that

$$G(s_1,s_2) \ll \exp(CkU^{\delta}\log\log x), \tag{2.12}$$

for a suitable absolute positive constant C.

Then integrating over some carefully chosen contours, Goldston, Pintz and Yildirim proved that the main contribution in $M_{2,s}$ and $M_{1,s}$ comes from the pole at $s_1 = s_2 = 0$, but in order to do that we should first give the final form of the weight $\mathcal{W}_n$.

2.4.11 The average of sifting functions

In this section we will prove that with $\lambda(d) = \frac{1}{(k+l)!}\mu(d)\big(\log\frac{R}{d}\big)^{k+l}$. We will get $\kappa > 1$, recalling that

$$\kappa = \frac{1}{\log 3x}\frac{A_\theta(n)}{A(n)},$$

where

$$A(n) = \sum_{x<n\leq 2x}\mathcal{W}_n^2 \qquad \text{and} \qquad A_\theta(n) = \sum_{x<n\leq 2x}(\sum_{i=1}^{k}\theta(n+h_i)\mathcal{W}_n^2.$$

In order to do that we will prove the following propositions and we will see that they give accurate approximations to the sifting functions $A(n)$ and $A_\theta(n)$. This will supply $\kappa > 1$.

Proposition 2.4.12. *Let $\mathcal{H}$ be an admissible $k-$tuple, with $|\mathcal{H}| = k$, $M = 2(k+l)$. If $R \ll x^{1/2}(\log x)^{-4M}$ and $h \leq R^C$ for all positive C, then*

$$\sum_{n\leq x}\Lambda_R(n;\mathcal{H},l)^2 = \binom{2l}{l}\frac{(\log R)^{k+2l}}{(k+2l)!}\big(\mathcal{G}(\mathcal{H}) + o_M(1)\big)x,$$

holds as R, $x \to \infty$.

Proposition 2.4.13. *Let $1 \leq h_0 \leq h$, and $\mathcal{H}^0 = \mathcal{H}\cup h_0$. If $R \ll_M x^{1/4}(\log x)^{-B(M)}$ and $h \leq R$, then*

$$\sum_{n\leq x}\Lambda_R(n;\mathcal{H},l)^2\theta(n+h_0) = \begin{cases} \binom{2l}{l}\frac{(\log R)^{k+2l}}{(k+2l)!}\big(\mathcal{G}(\mathcal{H}^0) + o_M(1)\big)x & \textit{if } h_0 \textit{ not in } \mathcal{H} \\ \binom{2l+1}{l+1}\frac{(\log R)^{k+2l+1}}{k+2l+1)!}\big(\mathcal{G}(\mathcal{H} + o_M(1)\big)x & \textit{if } h_0 \in \mathcal{H}_1\backslash\mathcal{H}_2 \end{cases}$$

holds as R, $x \to \infty$.

We recall the following facts about Riemann zeta function ([2]) for $s = \sigma + it$, there exists a constant $\overline{c} > 0$, such that $\zeta(s) \neq 0$ in the region

$$\sigma \geq 1 - \frac{4\overline{c}}{\log(|t|+3)}. \tag{2.13}$$

We have also for all $t \in \mathbb{R}$ that

$$\zeta(s) - \frac{1}{s-1} \ll \log(|t|+3). \tag{2.14}$$

$$\frac{1}{\zeta(s)} \ll \log(|t|+3), \text{ and } \quad \frac{\zeta'}{\zeta}(s) - \frac{1}{s-1} \ll \log(|t|+3). \tag{2.15}$$

We choose $\overline{c} = 10^{-2}$ and define $\mathcal{L}$ to be the contour given by

$$s = -\frac{\overline{c}}{\log(|t|+3)} + it. \tag{2.16}$$

To prove the Propositions 2.4.12 and 2.4.13 we will need a couple of lemmas from [3].

Lemma 2.4.14. *For $R \geq C$, $k \geq 2$, $B \leq Ck$,*

$$\int_{\mathcal{L}} (\log(|s|+3))^B |\frac{R^s}{s^k} ds| \ll_C C_1^k R^{-c_2} + e^{-\sqrt{\overline{c}\log(\frac{R}{2})}}, \tag{2.17}$$

where C_1, c_2 depends on C. Moreover if for a sufficiently small c_3, we have $k \leq c_3 \log(R)$, then

$$\int_{\mathcal{L}} (\log(|s|+3))^B |\frac{R^s}{s^k} ds| \ll_C e^{-\sqrt{\overline{c}\log(\frac{R}{2})}}. \tag{2.18}$$

Lemma 2.4.15. *Let q be a square-free integer and define $d_m(q) = m^{\omega(q)}$, for all positive real m. For $x \geq 1$ we have*

$$\sideset{}{^\flat}\sum_{q \leq x} \frac{d_m(q)}{q} \leq (\lceil m \rceil + \log x)^{\lceil m \rceil} \leq (m+1+\log x)^{m+1}. \tag{2.19}$$

$$\sideset{}{^\flat}\sum_{q \leq x} d_m(q) \leq x(\lceil m \rceil + \log x)^{\lceil m \rceil} \leq x(m+1+\log x)^{m+1}, \tag{2.20}$$

where the sum $\sum^\flat$ indicate the sum over the square free integers.

Proof. (of the Proposition 2.4.12) We have for $D = [d_1, d_2]$ that

$$\sum_{n \leq x} \Lambda_R(n; \mathcal{H}, l)^2 = \frac{x}{(k+l)!^2} \sum_{d_1, d_2} \mu(d_1)\mu(d_2) \frac{\rho(D)}{D} \Big(\log \frac{R}{d_1} \Big)^{k+l} \Big(\log \frac{R}{d_2} \Big)^{k+l} + O(M), \tag{2.21}$$

where

$$M = \frac{x}{(k+l)!^2} \sum_{d_1, d_2} \mu(d_1)\mu(d_2) \Big(\log \frac{R}{d_1} \Big)^{k+l} \Big(\log \frac{R}{d_2} \Big)^{k+l} \rho(D).$$

A direct estimate gives

$$\begin{aligned} M &\ll (\log R)^{2(k+l)} \sum_{d_1, d_2 \leq R} k^{\omega(D)} \quad \text{from } \rho(d) \leq k^{\omega(d)} \\ &\ll (\log R)^{4k} \sum_{r \leq R^2} (3k)^{\omega(r)} \quad \text{using lemma (2.4.15)} \\ &\ll R^2 (\log R)^{7k} \\ &\ll x^{1-\epsilon} \quad \text{imposing } R < x^{1/2-\epsilon}. \end{aligned}$$

Now we consider $\Lambda_R(n;\mathcal{H},l)^2 = \sum_{\substack{d\leq R\\ d|h(n)}} \lambda(d)$, with $\lambda(d) = \frac{1}{(k+l)!}\log(\frac{R}{d})^{(k+l)}$ and we proceed as in section 2.4.7. So using Perron's formula, the main term in (3.5) becomes

$$\mathcal{M} = \frac{1}{(2i\pi)^2}\int_{\substack{Re(s_1)=c\\ Re(s_2)=c}} \frac{\zeta(1+s_1+s_2)^k}{\zeta(1+s_1)^k\zeta(1+s_2)^k} G(s_1,s_2)\frac{R^{s_1+s_2}}{(s_1s_2)^{l+k+1}}ds_1ds_2, \tag{2.22}$$

where

$$G(s_1,s_2) = \prod_{p \text{ prime}} \Big(1-\frac{1}{p^{1+s_1+s_2}}\Big)^k\Big(1-\frac{1}{p^{1+s_1}}\Big)^{-k}\Big(1-\frac{1}{p^{1+s_2}}\Big)^{-k}\Big(1-\frac{\rho(p)}{p^{s_1+1}}-\frac{\rho(p)}{p^{s_2+1}}+\frac{\rho(p)}{p^{s_1+s_2+1}}\Big).$$

Then we have

$$\mathcal{M} = \frac{1}{(2i\pi)^2}\int_{\substack{Re(s_1)=c\\ Re(s_2)=c}} \frac{D(s_1,s_2)}{(s_1+s_2)^k}\frac{R^{s_1+s_2}}{(s_1s_2)^{l+1}}ds_1ds_2, \tag{2.23}$$

where

$$D(s_1,s_2) = G(s_1,s_2)\frac{(\zeta(1+s_1+s_2)(s_1+s_2))^k}{(s_1\zeta(1+s_1)^k(s_2\zeta(1+s_2))^k}.$$

We have for $\delta_i = -\min(\sigma_i, 0)$

$$G(s_1,s_2) \ll \exp(CMU^{\delta_1+\delta_2}\log\log U). \tag{2.24}$$

With $U = CM^2\log(2h_k)$, for some $M > k$.

Assuming that s_1, s_2 and s_1+s_2 are on the right of the contour $\mathcal{L}$ and using 2.4.11 and Lemma 2.4.14, we find

$$D(s_1,s_2) \ll (\log(|t_1|+3)^{2k}(\log|t_2|+3))^{2k}\exp(CMU^{\delta_1+\delta_2}\log\log U). \tag{2.25}$$

Letting

$$V = \exp(\sqrt{\log R}),$$

we define the following contours:

$$L'_j = \{4^{-j}c(\log V)^{-1} + it : t \in \mathbb{R}\},$$

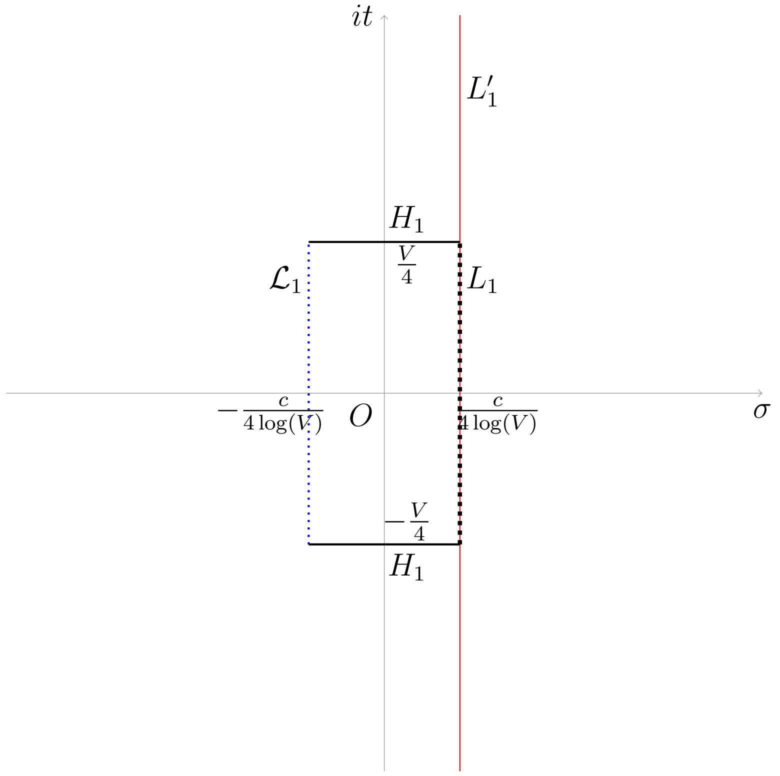
$$L_j = \{4^{-j}c(\log V)^{-1} + it : |t| \leq 4^{-j}V\},$$

$$\mathcal{L}_j = \{-4^{-j}c(\log V)^{-1} + it : |t| \leq 4^{-j}V\},$$

$$H_j = \{\sigma_j \pm i4^{-j}V : |\sigma_j| \leq 4^{-j}c(\log V)^{-1}\},$$

where $c > 0$ is a sufficiently small constant and $j = 1$ or 2.

For the case $j = 1$ we illustrate the contours by the following figure



From the estimate (2.25) we see that the integrand in $\mathcal{M}$ vanishes if $|t_1| \to \infty$ or $|t_2| \to \infty$ for s_1 and s_2 on the right of L_2'. Our first aim is to truncate L_j' to reach the contour L_j then we will prove that the error term generated by this operation is $O(\exp(-c\sqrt{\log R}))$. Hence to replace the s_j-contours over L_j with $\mathcal{L}_j$, we will consider the rectangle $L_jH_j\mathcal{L}_j$ then, finally, we will get a main term in the form of integrals over contours containing the poles at $s_1 = s_2 = 0$ and $s_1 = -s_2$. We will finish the proof by using the Cauchy theorem and observing that the contribution of the pole $s_1 = -s_2$ is negligible.

Let us start by estimating the error term. Indeed there are two error terms (similar up to a constant): one from truncating L_1' and the other from truncating L_2'. For s_1 and s_2 on the right of L_2 we have

$$\frac{R^{s_1+s_2}}{(s_1+s_2)^k} \ll (\log V)^k R^{\frac{5c}{16\log V}}. \tag{2.26}$$

We have also

$$\int_{L_2'\setminus L_2} \frac{(\log|s_1|+3)^{2k}}{|s_1|^{l+1}} ds_1 \ll \int_V^\infty \frac{(\log t+3)^{2k}}{t^{l+1}} dt \ll \frac{(\log V)^{2k}}{V} \tag{2.27}$$

$$\begin{aligned}\int_{L_1} \frac{(\log|s_2|+3)^{2k}}{|s_2|^{l+1}} ds_2 &\ll \int_{\frac{c}{4\log V}}^{\frac{c}{4\log V}+i} \frac{(\log|s_2|+3)^{2k}}{|s_2|^{l+1}} ds_2 + \int_{\frac{c}{4\log V}+i}^{\frac{c}{4\log V}+i\infty} \frac{(\log|s_2|+3)^{2k}}{|s_2|^{l+1}} ds_2 \\ &\ll (\log V)^{l+1}.\end{aligned}$$

Then recalling $V=\exp(\sqrt{\log R})$

$$\begin{aligned}\int_{L_1}\int_{L_2'\setminus L_2} \frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}} ds_1 ds_2 &\ll (\log U)^{CMU^{\delta_1+\delta_2}} (\log V)^{2k+l+1} \frac{R^{\frac{5c}{16\log V}}}{V} \\ &\ll \frac{(\log V)^{C_2M}}{V^{1-\frac{5c}{16}}} \ll \exp(-c\sqrt{\log R}).\end{aligned}$$

Consequently we get

$$\begin{aligned}\mathcal{M} &= \frac{1}{(2\pi i)^2}\int_{L_2'}\int_{L_1'} \frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}} ds_1 ds_2 \\ &= \frac{1}{(2\pi i)^2}\Big\{\int_{L_2}\int_{L_1} + \int_{L_2'\setminus L_2}\int_{L_1} + \int_{L_2}\int_{L_1'\setminus L_1}\Big\} \frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}} ds_1 ds_2 \\ &= \frac{1}{(2\pi i)^2}\int_{L_2}\int_{L_1} \frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}} ds_1 ds_2 + O(\exp(-c\sqrt{\log R}).\end{aligned}$$

Now we shift the L_i contours to $\mathcal{L}_i$, and we consider the rectangle $L_jH_j\mathcal{L}_j$ which contains a pole at $s_1=s_2=0$ of order $l+1$ and a pole at $s_1=-s_2$ of order k. We denote

$$\mathcal{C}_i = H_i \cup \mathcal{L}_i \cup L_i.$$

Then we get

$$\int_{L_1} \frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}} ds_1 = \int_{\mathcal{C}_1} \frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{k+1}} ds_1 - \int_{H_1\cup\mathcal{L}_1} \frac{H(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}} ds_1.$$

Consequently for

$$I_1 = \int_{\mathcal{C}_1} \frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{k+1}} ds_1.$$

We have

$$I_1 = 2i\pi\Big(Res_{s_1=-s_2}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{k+1}}\Big) + \Big(Res_{s_1=0}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{k+1}}\Big)\Big). \tag{2.28}$$

Arguing analogously on L_2 we obtain

$$\begin{aligned}
\mathcal{M} &= Res_{s_2=0}Res_{s_1=0}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{k+1}}\Big)\\
&+ \frac{1}{2i\pi}\int_{H_2\cup\mathcal{L}_2} Res_{s_1=0}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}}\Big)ds_2\\
&+ \frac{1}{2i\pi}\int_{H_1\cup\mathcal{L}_1} Res_{s_2=0}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}}\Big)ds_1\\
&+ \frac{1}{2i\pi}\int_{L_2} Res_{s_1=-s_2}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}}\Big)ds_2\\
&+ \frac{1}{(2i\pi)^2}\int_{H_2\cup\mathcal{L}_2}\int_{H_1\cup\mathcal{L}_1}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}}\Big)ds_2ds_2\\
&+ O(e^{-c\sqrt{\log x}})\\
&= J_0+J_1+J_2+J_3+J_4+O(e^{-c\sqrt{\log x}}),
\end{aligned}$$

say. We can write the residue in J_0 as

$$J_0=\frac{1}{(2i\pi)^2}\int_{\Gamma_1}\int_{\Gamma_2}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{l+1}}\Big)ds_2ds_2,$$

where

$$\Gamma_1=\{|s_1|=r:r>0\} \text{ and } \Gamma_2=\{|s_1|=2r:r>0\}.$$

If we choose $s_1=s$ and $s_2=\lambda s$, and Γ_3 is the circle $|\lambda|=2$, then J_0 becomes

$$J_0=\frac{1}{(2i\pi)^2}\int_{\Gamma_1}\int_{\Gamma_3}\Big(\frac{D(s,s\lambda)R^{s(1+\lambda)}}{(1+\lambda)^k\lambda^{l+1}s^{2l+k+1}}\Big)dsd\lambda.$$

For a fixed λ the integrand has a pole at $s=0$ of order $2l+k+1$. Recalling the formula

$$Res_{s=0}\Big(\frac{D(s,\lambda s)R^{s(1+\lambda)}}{s^{2l+k+1}}\Big)=\frac{1}{(2l+k)!}\big(\frac{\partial^{2l+k}}{\partial s}\big)_{s=0}(D(s,\lambda s)R^{s(1+\lambda)}),$$

we get a main term

$$\frac{2i\pi D(0,0)(\log R)^{k+2l}}{(k+2l)!}.$$

By the well-known Cauchy estimate, we have

$$|f^{(j)}(z_0)| \leq \max_{|z-z_0|=\eta} |f(z)| \frac{j!}{\eta^j},$$

where $f^{(j)}$ is the j-th derivative of f, and f analytic on $|z - z_0| \leq \eta$.

Choosing z_0 on the right of $\mathcal{L}$, and

$$\eta = \frac{1}{C \log U \log T},$$

for $T = |s_1| + |s_2| + 3$ we get that the partial derivatives of $D(s_1, s_2)$ satisfies (2.25). In fact we have

$$\begin{aligned} \frac{\partial^m}{\partial^m s_2} \frac{\partial^j}{\partial^j s_2} D(s_1, s_2) &\leq j!m!(C \log U \log T)^{j+m} \max_{\substack{|s_1'-s_1| \leq \eta \\ |s_2'-s_2| \leq \eta}} |D(s_1, s_2)| \qquad (2.29) \\ &\ll \exp(CMU^{\delta_1+\delta_2} \log \log U). \qquad (2.30) \end{aligned}$$

Hence

$$J_0 = \frac{D(0,0)(\log R)^{k+2l}}{2i\pi(k+2l)!} \int_{\Gamma_3} \frac{(1+\lambda)^{2l}}{\lambda^{l+1}} d\lambda + O((\log U)^{k+l-1} (\log \log U)^C).$$

By Newton's formula we have

$$\int_{\Gamma_3} \frac{(1+\lambda)^{2l}}{\lambda^{l+1}} d\lambda = Res_{\lambda=0} \frac{(1+\lambda)^{2l}}{\lambda^{l+1}} = \binom{2l}{l}.$$

Finally we can write

$$J_0 = \binom{2l}{l} \frac{D(0,0)(\log R)^{k+2l}}{2i\pi(k+2l)!} + O((\log U)^{k+l-1} (\log \log U)^C). \qquad (2.31)$$

Now to evaluate the integral J_3, we should calculate the residue at $s_1 = -s_2$ of order k. By the residue formula we have

$$Res_{s_1=-s_2} \Big(\frac{D(s_1, s_2) R^{s_1+s_2}}{(s_1+s_2)^k (s_1 s_2)^{l+1}} \Big) = \lim_{s_1 \to -s_2} \frac{1}{(k-1)!} \frac{\partial^{k-1}}{\partial^{k-1} s_1} \frac{D(s_1, s_2) R^{s_1+s_2}}{(s_1 s_2)^{l+1}}.$$

Then by Leibniz rule we get

$$\lim_{s_1 \to -s_2} \frac{\partial^{k-1}}{\partial^{k-1} s_1} \frac{D(s_1, s_2) R^{s_1+s_2}}{(s_1 s_2)^{l+1}} = \sum_{i=0}^{k-1} (\log R)^{k-1-i} \mathcal{B}_i(s_2),$$

where

$$\mathcal{B}_i(s_2) = \binom{k-1}{i} \sum_{j=0}^{i} \binom{i}{j} \frac{\partial^{i-j}}{\partial^{i-j} s_1} D(s_1, s_2)\Big|_{s_1=-s_2} \frac{(-1)^j (l+1)...(l+j)}{(-1)^{l+j+1} s_2^{2l+j+2}}.$$

Hence we can finally write

$$Res_{s_1=-s_2}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k (s_1 s_2)^{l+1}}\Big) = \frac{1}{(k-1)!}\sum_{i=0}^{k-1} (\log R)^{k-1-i} \mathcal{B}_i(s_2).$$

Summing up we get

$$J_3 = \frac{1}{2i\pi(k-1)!}\sum_{i=0}^{k-1} (\log R)^{k-1-i} \int_{L_2} \mathcal{B}_i(s_2) ds_2.$$

Using again the Cauchy estimate and arguing analogously we obtain the estimaate

$$\mathcal{B}_i(s_2) \ll \exp(CMU^{\delta_2} \log\log U) \frac{\log(|t_2|+3)^{4M}}{|t_2|^{4k+2} \max(1, |t_2|^i)},$$

which holds for s_2 on the right of $\mathcal{L}$. Now we can finally say that the contribution from J_3 in $\mathcal{M}$ is in fact an error term. For J_1, J_2 and J_4 we repeat the same argument in $(29)-(31)$, and this completes the proof of the Proposition 2.4.12.

□

Proof. (of the Proposition 2.4.13) The proof is very similar to the one we used for proving the Proposition 2.4.12; we just need to translate k in $k-1$ and l in $l+1$. Furthermore, instead of $G(s_1, s_2)$ we use

$$\begin{aligned} G^*(s_1,s_2) &= \textstyle\prod_{p \text{ prime}} \Big(1 - \frac{1}{p^{1+s_1+s_2}}\Big)^{k-1} \Big(1 - \frac{1}{p^{1+s_1}}\Big)^{-(k-1)} \Big(1 - \frac{1}{p^{1+s_2}}\Big)^{-(k-1)} \\ &\qquad \Big(1 - \frac{\rho(p)-1}{p-1}\big(\frac{1}{p^{s_1+1}} + \frac{1}{p^{s_2+1}} - \frac{1}{p^{s_1+s_2+1}}\big)\Big). \end{aligned}$$

Remarking that

$$G^*(0,0) = G(0,0),$$

□

and using the previous propositions, by the Eliott-Halberstam conjecture, Goldston Pintz and Yildirim proved the following theorem.

Theorem 2.4.16. *Assume the primes have a level of distribution $\theta \geq 1/2$. Then there exists an explicit constant $C(\theta)$, such that any admissible $k-$tuple with $k \geq C(\theta)$ contains at least two primes infinitely often.*

Moreover, if $\theta > 20/21$, then Theorem 2.4.16 holds for $k \geq 7$, and the 7-tuple $(n, n+2, n+6, n+8, n+12, n+18, n+20)$ is admissible. Then the following corollary is an immediate consequence of the previous theorem.

Corollary 2.4.17. *The Elliott-Halberstam conjecture implies that*

$$\lim_{n\to\infty} \inf(p_{n+1} - p_n) \leq 20.$$

Proof. (of Theorem 2.4.16)

From Proposition 2.4.12, we get

$$\sum_{n\leq x} \Lambda_R(n;\mathcal{H},l)^2 \sim \frac{1}{(k+2l)!}\binom{2l}{l}\mathcal{G}(\mathcal{H})x(\log R)^{k+2l},$$

as $x \to \infty$, and for all positive ϵ, $h_i \in \mathcal{H}$, and $R \ll x^{\frac{\theta}{2}-\epsilon}$.

We also have from Proposition 2.4.13 that

$$\sum_{n\leq x} \Lambda_R(n;\mathcal{H},l)^2\theta(n+h_i) \sim \frac{1}{(k+2l+1)!}\binom{2l+2}{l+1}\mathcal{G}(\mathcal{H})x(\log R)^{k+2l+1},$$

as $x \to \infty$. Choosing $R = x^{\frac{\theta}{2}-\epsilon}$, we obtain for S defined in (2.6), that

$$S \sim \frac{k}{(k+2l+1)!}\binom{2l+2}{l+1}\mathcal{G}(\mathcal{H})x(\log R)^{k+2l+1} - \log 3x\frac{1}{(k+2l)!}\binom{2l}{l}\mathcal{G}(\mathcal{H})x(\log R)^{k+2l},$$

as $x \to \infty$. So we can write that

$$S \sim \Big(\frac{2k}{k+2l+1}\frac{2l+1}{l+1}\log R - \log 3x\Big)\frac{1}{(k+2l)!}\binom{2l}{l}\mathcal{G}(\mathcal{H})x(\log R)^{k+2l},$$

as $x \to \infty$.

As we mentioned before, if $S > 0$ then there exists $n \in [x+1, 2x]$ such that at least two of the integers $n + h_1, \ldots, n + h_k$ will be primes. This follows from the condition

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

Since for any $\theta > 1/2$ and $k, l \to \infty$ with $l = o(k)$, we have that

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

And this proves Theorem 2.4.16.

For the Corollary it is sufficient to take $l = 1$, $k = 7$ and $\theta = 20/21$ to get that

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

□

Using the previous results we prove the following theorem.

Theorem 2.4.18. *Letting*

$$\Delta_\nu = \lim_{n\to\infty} \inf \frac{p_{n+\nu} - p_n}{\log p_n},$$

we have

$$\Delta_\nu = \max(\nu - 2\theta, 0),$$

where θ is the level of distribution of primes.

We will need a result of Gallagher [18] (for a simpler proof see also [12]).

Lemma 2.4.19. *For $h \longrightarrow \infty$, we have*

$$\sum_{\substack{1\le h_1,\dots h_k \le h \\ distinct}} \mathcal{G}(\mathcal{H}_k) \sim h^k.$$

Proof. (Of Theorem 2.4.18)

We use the same idea already applied in (2.4.16), we just modify the sum S by considering

$$\overline{\mathcal{S}} = \sum_{n=x+1}^{2x} \Big(\sum_{i=1}^{k} \theta(n+h_i) - \nu \log 3x \Big) \Lambda(n; \mathcal{H}, l)^2,$$

where υ is positive. From Proposition 2.4.13 we have if h_0 is not in $\mathcal{H}_k$ that

$$\sum_{n\le x} \Lambda_R(n; \mathcal{H}_k, l)^2 \upsilon(n+h_0) \sim \frac{1}{(k+2l)!}\binom{2l}{l}\mathcal{G}(\mathcal{H}_k \cup \{h_0\}) x (\log R)^{k+2l},$$

as $x \to \infty$.

Choosing now $R = x^{\frac{\theta}{2}-\epsilon}$, we get

$$\begin{aligned}\overline{\mathcal{S}} \sim \sum_{\substack{1\leq h_1,\dots h_k\leq h\\ \text{distinct}}} &\Big(\frac{k}{(k+2l+1)!}\binom{2l+2}{l+1}\mathcal{G}(\mathcal{H}_k)x(\log R)^{k+2l+1}\\ &+\sum_{\substack{1\leq h_0\leq h\\ h_0\leq h, 1\leq i\leq k}} \frac{1}{(k+2l)!}\binom{2l}{l}\mathcal{G}(\mathcal{H}_k\cup\{h_0\})x(\log R)^{k+2l}\\ &-\nu\log 3x\frac{1}{(k+2l)!}\binom{2l}{l}\mathcal{G}(\mathcal{H}_k)x(\log R)^{k+2l}\Big),\end{aligned}$$

as $x\to\infty$. Assuming

$$h > \Big(\nu - \frac{2k}{k+2l+1}\frac{2l+1}{l+1}\Big(\frac{\theta}{2}-\epsilon\Big)\Big)\log x,$$

we obtain

$$\overline{\mathcal{S}} \sim \Big(\frac{2k}{k+2l+1}\frac{2l+1}{l+1}\log R + h - \nu\log 3x\Big)\frac{1}{(k+2l)!}\binom{2l}{l}xh^k(\log R)^{k+2l},$$

as $x\to\infty$. Hence, we have at least $\nu+1$ primes in the interval $(n, n+h]$, for $N < n \leq 2N$. Letting k be large and choosing $l = [\sqrt{k}/2]$ we get

$$h > \Big(\nu - 2\theta + 4\epsilon + O\Big(\frac{1}{\sqrt{k}}\Big)\Big)\log x.$$

This proves Theorem 2.4.18. □

By the Bombieri-Vinogradov Theorem we have $\theta = 1/2$; so the following corollary is a trivial consequence of the previous theorem.

Corollary 2.4.20. *We have*

$$\Delta_1 = \lim_{n\to\infty}\inf\frac{p_{n+1}-p_n}{\log p_n} = 0.$$

Chapter 3

Higher dimensional analysis

3.1 The basic setting

Our aim in this section is to present the Maynard's work on the bounded gaps between primes [13], namely the following theorem

Theorem 3.1.1. *We have*

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

Assuming the Elliot-Halberstam conjecture we will prove the following result

Theorem 3.1.2. *Assume that the primes have a level of distribution $\theta < 1$. then*

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

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

Definitions and notations

In the following, we will perform many multi-dimensional summations, so in order to simplify the notations, we denote

$$\sum_{a_1\geq 1}\sum_{a_2\geq 1}\cdots\sum_{a_k\geq 1} \text{ by } \sum_{a_1,\ldots,a_k},$$

and

$$\sum_{\substack{a_1\geq 1\\ a_1|b_1}}\sum_{\substack{a_2\geq 1\\ a_2|b_2}}\cdots\sum_{\substack{a_k\geq 1\\ a_k|b_k}} \text{ by } \sum_{\substack{a_1,\ldots,a_k\\ a_i|b_i}}.$$

For the one dimensional sums we will use the usual notation $\sum_{m|n}\lambda(m)$ to denote the sum of the values of λ in all the divisors of n. But for the multi-dimensional sums, we will use $\sum_{\substack{d_1,\dots,d_k \\ a_i|d_i}}\lambda_{d_1,\dots,d_k}$ to denote the restriction of $\sum_{d_1,\dots,d_k}\lambda_{d_1,\dots,d_k}$ to d_i divisible by a_i. We will also denote by

- $\mathcal{R}_k$ is the simplex $\{(x_1,\dots,x_k)\}\in[0,1]^k : \sum_{i=1}^k x_i \le 1\}$.
- $\mathcal{S}_k$ is the set of real valued Riemann integrable functions supported on $\mathcal{R}_k$.
- $\alpha_{a_1,\dots,a_k}$ is the real sequence indexed by $(a_1,\dots,a_k)\in\mathbb{Z}^k_{\ge 0}$, where $\mathbb{Z}_{\ge 0}$ is the set of positive integers.

For an admissible k-tuple $\mathcal{H}=\{h_1,\dots,h_k\}$, we define the sum

$$S(x,\rho)=\sum_{x\le n<2x}\Big(\sum_{i=1}^{k}\chi(n+h_i)-\rho\Big)\mathcal{W}'_n, \tag{3.1}$$

where $\rho >$. Here we consider the characteristic function of the primes

$$\chi(n)=\begin{cases}1 \text{ if } n \text{ prime}\\ 0 \text{ otherwise} .\end{cases}$$

The key idea in Maynard's improvement is to consider a multi-dimensional weight by taking

$$\mathcal{W}'_n=\Big(\sum_{\substack{d_1,\dots,d_k\\ d_i|n+h_i}}\lambda_{d_1,\dots,d_k}\Big)^2,$$

where $(d_i,d_j)=1$ for $i\neq j$.

We take $D=\prod_{p\le D_0}p$, where $D_0=\log\log\log x$, and by the Prime Number Theorem we have $D\ll(\log\log x)^2$. Recall that the role of D is eliminating the contribution coming from the primes less than D_0, we remark that the optimal choice of D_0 is not important in our context.

3.2 Maynard's combinatorial approach

In the rest of the thesis, we assume that

$$\boxed{d\le R,\ (d,D)=1,\ \text{and } d \text{ square-free},} \tag{3.2}$$

where, $(d_1,\dots,d_k)$ is the support of $\lambda_{d_1,\dots,d_k}$, and $d=\prod_i^k d_i$.

We remark that, if $S(x,\rho) > 0$, then by the positivity of $\mathcal{W}'_n$, we have that there exists at least one element $n_0 \in [x, 2x[$ such that $\sum_{i=1}^{k} \chi(n_0 + h_i) > \rho$, that implies that, at least $\lfloor \rho + 1 \rfloor$ of the $n_0 + h_i$ are primes, where $1 \le i \le k$. If that's true for any large x, then there exists infinitely many integers n for which at least $\lfloor \rho + 1 \rfloor$ of the $n + h_i$ are primes, where $1 \le i \le k$. Taking $0 \le h_1 \le h_2 \le \cdots \le h_k$, we obtain $\liminf_{n\to\infty} p_{n+1} - p_n \le h_k - h_1$.

In the same fashion as in the GPY method, we write

$$S(x,\rho) = S_2 - \rho S_1,$$

where

$$S_1 = \sum_{\substack{x<n\le 2x \\ n\equiv m \pmod D}} \mathcal{W}'_n, \tag{3.3}$$

and

$$S_2 = \sum_{\substack{x<n\le 2x \\ n\equiv m \pmod D}} \Big(\sum_{i=1}^{k} \chi(n+h_i)\Big)\mathbb{W}'_n. \tag{3.4}$$

Maynard proved the following proposition on estimating S_1 and S_2.

Proposition 3.2.1. *Assume that the primes have a level of distribution $\theta > 0$, and let $R = x^{\theta/2-\delta}$ for a fixed $\delta > 0$. Then*

$$\lambda_{d_1,\dots,d_k} = \Big(\prod_{i=1}^{k} \mu(d_i) d_i\Big) \sum_{\substack{r_1,\dots,r_k \\ d_i | r_i \\ (r_i, D)=1}} \frac{\mu\Big(\prod_{i=1}^{k} r_i\Big)^2}{\prod_{i=1}^{k} \phi(r_i)} F\Big(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\Big),$$

whenever $(\prod_{i=1}^{k} d_i, D) = 1$, and $\lambda_{d_1,\dots,d_k} = 0$ otherwise, where F is a fixed smooth function. Moreover, if F is supported on $\mathcal{R}_k$, then we have

$$S_1 = \frac{(1+o(1))\phi(D)^k x (\log R)^k}{D^{k+1} \log x} I_k(F),$$

$$S_2 = \frac{(1+o(1))\phi(D)^k x (\log R)^{k+1}}{D^{k+1}} \sum_{m=1}^{k} J_k^{(m)}(F),$$

as $x \to \infty$, where $I_k(F) \ne 0$, $J_k^{(m)}(F) \ne 0$, and

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

$$J_k^{(m)}(F) = \int_0^1 \dots \int_0^1 \Big(\int_0^1 F(t_1, \dots, t_k) dt_m\Big)^2 dt_1 \dots dt_{m-1} dt_{m+1} dt_k,$$

We will start by proving the following lemma, as a first step toward proving the Proposition 3.2.1.

Lemma 3.2.2. *Let S_1 be as in 3.2.1, and assuming that $\lambda_{d_1,\ldots,d_k}$ is a sequence of real numbers supported on $(d_1,\ldots,d_k)$ satisfying the conditions in* 3.2*. We have*

$$S_1 = \frac{x}{D}\sum_{\substack{d_1,\ldots,d_k\\ e_1,\ldots,e_k\\ (e_i,d_i)=1\forall i\neq j}}\frac{\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}}{\prod_{i=1}^{k}[d_i,e_i]} + O\Big(\lambda_{\max}^2 R^2(\log R)^{2k}\Big),$$

where $\lambda_{\max} = \sup_{d_1,\ldots,d_k}|\lambda_{d_1,\ldots,d_k}|$.

Proof. We have

$$S_1 = \sum_{\substack{x\le n<2x\\ n\equiv m \pmod D}}\left(\sum_{d_i|n+h_i}\lambda_{d_1,\ldots,d_k}\right)^2 = \sum_{\substack{d_1,\ldots,d_k\\ e_1,\ldots,e_k}}\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}\sum_{\substack{x\le n<2x\\ n\equiv m \pmod D\\ [d_i,e_i]|n+h_i}}1.$$

If D, $[d_1,e_1],\ldots,[d_i,e_i]$ are pairwise coprime, then using the Chinese Remainder Theorem, we can write the sum over a residue class modulo $D\prod_{i=1}^{2}[d_i,e_i]$. From the conditions $(d_i,d_j)=1$ and $(D,d_i)=1$ for all i and j, we conclude that $(D,[d_i,e_i])=1$. It remains the case where $([d_i,e_i],[d_j,e_j])>1$ for some $1\le i,j\le k$. Again from the condition $(d_i,d_j)=1$, that is satisfied if only there exist a prime p dividing d_i and e_j for $1\le i,j\le k$. But in this case we will have $p|n+h_i-n+h_j$, that implies $p|h_i-h_j$. If we choose $D_0>\max|h_i-h_j|$, this will contradict the fact that $(D,d_i)=1$, and that justifies our choice of $D_0=\log\log\log x$ for x sufficiently large. Hence we have

$$\sum_{\substack{x\le n<2x\\ n\equiv m \pmod D\\ [d_i,e_i]|n+h_i}}1 = \frac{x}{D\prod_{i=1}^{2}[d_i,e_i]} + O(1).$$

We recall that $\lambda_{d_1,\ldots,d_k}$ is supported on the d_i's such that $(d_i,d_j)=1$ and $(D,d_i)=1$ for all i and j. Hence, we have

$$S_1 = \sum_{\substack{d_1,\ldots,d_k\\ e_1,\ldots,e_k\\ (d_i,e_i)=1\forall i\neq j}}\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}\frac{x}{D\prod_{i=1}^{k}[d_i,e_i]} + O\Big(\sum_{\substack{d_1,\ldots,d_k\\ e_1,\ldots,e_k\\ (d_i,e_i)=1\forall i\neq j}}|\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}|\Big).$$

It is easy to see that

$$O\Big(\sum_{\substack{d_1,\ldots,d_k\\ e_1,\ldots,e_k\\ (d_i,e_i)=1\forall i\neq j}}|\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}|\Big) \ll \lambda_{max}^2\Big(\sum_{n<R}\tau_k(n)\Big)^2,$$

where $\tau_k(n)$ is the number of ways expressing n as a product of k factors, and we have the following lemma.

Lemma 3.2.3. *We have*

$$\sum_{n<R} \tau_k(n) \ll R(\log R)^{k-1},$$

as $R \to \infty$

Proof. We will show the result by induction. We see that the result is trivial for $k = 1$, then assuming the estimate for $k-1$ we prove the estimate for k. We have

$$\tau_k(n) = \sum_{d|n} \tau_k - 1(n/d).$$

Hence

$$\begin{aligned} \sum_{n<R} \tau_k(n) &= \sum_{n<R} \sum_{d|n} \tau_{k-1}(n/d) \leq \sum_{d<R} \sum_{m<R/d} \tau_{k-1}(m) \\ &\ll \sum_{d \leq R} \frac{R}{d} \log \left(\frac{R}{d}\right)^{k-2} \ll R(\log R)^{k-2} \sum_{d \leq R} \frac{1}{d} \ll R(\log R)^{k-1} \end{aligned}$$

□

For our purpose, it will be sufficient to take

$$\sum_{n<R} \tau_k(n) \ll R(\log R)^k.$$

Hence

$$S_1 = \frac{x}{D} \sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k \\ (e_i,d_i)=1 \forall i \neq j}} \frac{\lambda_{d_1,\ldots,d_k} \lambda_{e_1,\ldots,e_k}}{\prod_{i=1}^k [d_i, e_i]} + O\Big(\lambda_{\max}^2 R^2 (\log R)^{2k}\Big). \tag{3.5}$$

□

Recalling that the main term in S_1 depends in the condition $(e_i, d_i) = 1, \forall i \neq j$, we can remove this condition multiplying by the well known discontinuous factor

$$\sum_{s_{i,j}|e_i,d_i} \mu(s_{i,j}) = \begin{cases} 1 \text{ if } (e_i, d_i) = 1 \\ 0 \text{ otherwise .} \end{cases}$$

Replacing $\frac{1}{[d_i,e_i]}$ in (3.5) by $\frac{1}{d_1e_1}\sum_{u_i|d_i,e_i}\phi(u_i)$, we get a main term of the form

$$\begin{aligned} M_1 &= \frac{x}{D}\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k\\ (e_i,d_i)=1\forall i\neq j}} \frac{\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k [d_i,e_i]} \\ &= \frac{x}{D}\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k}} \Big(\prod_{i\neq j}\sum_{s_{i,j}|e_i,d_i}\mu(s_{i,j})\Big)\frac{\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k d_i\prod_{i=1}^k e_i}\prod_{i=1}^k\Big(\sum_{u_i|d_i,e_i}\phi(u_i)\Big) \\ &= \frac{x}{D}\sum_{u_1,\dots,u_k}\prod_{i=1}^k\phi(u_i)\Big(\prod_{i\neq j}\sum_{s_{i,j}}\mu(s_{i,j})\Big)\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k\\ u_i|d_i,e_i\\ s_{i,j}|d_i,e_j}}\frac{\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k d_i\prod_{i=1}^k e_i} \end{aligned}$$

Remarking that with the assumptions $u_i|d_i,e_i$ and $s_{i,j}|d_i,e_j$, we have $(s_{i,j},u_i)=1$ (resp. $(s_{i,j},u_j)=1$), from $(e_i,e_j)=1$ (resp. $(d_i,d_j)=1$), otherwise $\lambda_{e_1,\dots,e_k}$ (resp. $\lambda_{d_1,\dots,d_k}$) will be zero. Hence, $s_{i,j}$ is coprime to u_j and u_i. Furthermore, from $s_{i,j}|d_i,e_j$ and $(d_i,d_j)(e_i,e_j)=1$ for all $i\neq j$, we get $(s_{i,j},\prod_{i'\neq i}s_{i',j})=1$, and $(s_{i,j},\prod_{j'\neq j}s_{i,j'})=1$. Now, we take

$$a_j = u_j\prod_{i\neq j}s_{j,i}, \text{ and } b_j = u_j\prod_{i\neq j}s_{i,j}. \tag{3.6}$$

This implies that

$$\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k\\ u_i|d_i,e_i\\ s_{i,j}|d_i,e_j}}\frac{\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k d_i\prod_{i=1}^k e_i} = \sum_{\substack{d_1,\dots,d_k\\ a_i|d_i}}\frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^k d_i}\sum_{\substack{e_1,\dots,e_k\\ b_i|d_i}}\frac{\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k e_i}.$$

Hence

$$M_1 = \frac{x}{D}\sum_{u_1,\dots,u_k}\prod_{i=1}^k\phi(u_i)\Big(\prod_{i\neq j}\sum_{s_{i,j}}\mu(s_{i,j})\Big)\sum_{\substack{d_1,\dots,d_k\\ a_i|d_i}}\frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^k d_i}\sum_{\substack{e_1,\dots,e_k\\ b_i|d_i}}\frac{\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k e_i}. \tag{3.7}$$

Obviously, a meaningful simplification of M_1, could be done by replacing the sums

$$\sum_{\substack{e_1,\dots,e_k\\ b_i|d_i}}\frac{\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k e_i}, \text{ and } \sum_{\substack{d_1,\dots,d_k\\ a_i|d_i}}\frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^k d_i},$$

by a quantity satisfying the conditions (3.2).

In order to do that we will need the following lemma, which can be seen as a multi-dimensional Möbius inversion formula.

Lemma 3.2.4. *Let $\beta_{d_1,\ldots,d_k}$ and $\alpha_{a_1,\ldots,a_k}$ be two sequences of real numbers supported on a finite number of k-tuples of integers $(d_1,\ldots,d_k)$ and $(a_1,\ldots,a_k)$. If*

$$\alpha_{a_1,\ldots,a_k} = \sum_{\substack{d_1,\ldots,d_k \\ a_i|d_i}} \beta_{d_1,\ldots,d_k},$$

then

$$\beta_{d_1,\ldots,d_k} = \sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i}} \prod_{i=1}^{k} \mu(\frac{d_i}{a_i})\alpha_{a_1,\ldots,a_k},$$

Proof. We have

$$\begin{aligned}
\sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i}} \prod_{i=1}^{k} \mu(\frac{a_i}{d_i})\alpha_{a_1,\ldots,a_k} &= \sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i}} \prod_{i=1}^{k} \mu(\frac{a_i}{d_i}) \sum_{\substack{c_1,\ldots,c_k \\ c_i:a_i|c_i}} \beta_{c_1,\ldots,c_k} \\
&= \sum_{c_1,\ldots,c_k} \beta_{c_1,\ldots,c_k} \sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i|c_i}} \prod_{i=1}^{k} \mu(\frac{a_i}{d_i}).
\end{aligned}$$

We put $m_i = \frac{a_i}{d_i}$. Hence

$$\sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i|e_i}} \prod_{i=1}^{k} \mu(\frac{a_i}{d_i}) = \sum_{\substack{m_1,\ldots,m_k \\ m_i|\frac{e_i}{d_i}}} \prod_{i=1}^{k} \mu(m_i) = \prod_{i=1}^{k} \sum_{\substack{m_1,\ldots,m_k \\ m_i|\frac{e_i}{d_i}}} \mu(m_i).$$

We know that

$$\sum_{m_i|\frac{e_i}{d_i}} \mu(m_i) = \begin{cases} 1 \text{ if } e_i = d_i \\ 0 \text{ otherwise} \end{cases}.$$

Hence

$$\sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i}} \prod_{i=1}^{k} \mu(\frac{a_i}{d_i})\alpha_{a_1,\ldots,a_k} = \beta_{d_1,\ldots,d_k}.$$

□

Corollary 3.2.5. *Let $\beta_{d_1,\ldots,d_k}$ and $\alpha_{a_1,\ldots,a_k}$ be two sequences of real numbers supported on finitely many k-tuples of square-free integers $(d_1,\ldots,d_k)$ and $(\alpha_{a_1,\ldots,a_k})$, then*

$$\alpha_{a_1,\ldots,a_k} = \prod_{i=1}^{k} \mu(a_i) \sum_{\substack{d_1,\ldots,d_k \\ a_i|d_i}} \beta_{d_1,\ldots,d_k},$$

if and only if

$$\beta_{d_1,\ldots,d_k} = \prod_{i=1}^{k} \mu(d_i) \sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i}} \alpha_{a_1,\ldots,a_k}.$$

Proof. By symmetry, it's sufficient to prove just one implication. Assuming that

$$\alpha_{a_1,\ldots,a_k} = \prod_{i=1}^{k} \mu(a_i) \sum_{\substack{d_1,\ldots,d_k \\ a_i|d_i}} \beta_{d_1,\ldots,d_k},$$

We apply the lemma 3.2.4 to $\beta_{d_1,\ldots,d_k}$ and $\frac{\alpha_{a_1,\ldots,a_k}}{\prod_{i=1}^{k}\mu(a_i)}$, Hence

$$\beta_{d_1,\ldots,d_k} = \sum_{\substack{d_1,\ldots,d_k \\ d_i|a_i}} \prod_{i=1}^{k} \mu(\frac{a_i}{d_i})\mu(a_i) \sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i}} \alpha_{a_1,\ldots,a_k}.$$

Using the fact that a_i and d_i are square-free, we get $\mu(\frac{a_i}{d_i})\mu(a_i) = \mu(d_i)$. Hence

$$\alpha_{a_1,\ldots,a_k} = \sum_{\substack{d_1,\ldots,d_k \\ a_i|d_i}} \beta_{d_1,\ldots,d_k},$$

Then

$$\beta_{d_1,\ldots,d_k} = \sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i}} \prod_{i=1}^{k} \mu(\frac{d_i}{a_i})\alpha_{a_1,\ldots,a_k} = \sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i}} \prod_{i=1}^{k} \mu(d_i)\alpha_{a_1,\ldots,a_k}.$$

Hence

$$\beta_{d_1,\ldots,d_k} = \prod_{i=1}^{k} \mu(d_i) \sum_{\substack{a_1,\ldots,a_k \\ d_i|a_i}} \alpha_{a_1,\ldots,a_k}.$$

□

Now we set

$$y_{a_1,\ldots,a_k} = \Big(\prod_{i=1}^{k} \mu(a_i)\phi(a_i)\Big) \sum_{\substack{d_1,\ldots,d_k \\ a_i|d_i}} \frac{\lambda_{d_1,\ldots,d_k}}{\prod_{i=1}^{k} d_i},$$

Then, applying the corollary 3.2.5, with

$$\beta_{d_1,\ldots,d_k} = \frac{\lambda_{d_1,\ldots,d_k}}{\prod_{i=1}^{k} d_i},$$

$$\alpha_{a_1,\ldots,a_k} = \frac{y_{a_1,\ldots,a_k}}{\prod_{i=1}^{k} \phi(a_i)},$$

We obtain

$$\frac{\lambda_{d_1,\ldots,d_k}}{\prod_{i=1}^k d_i} = \prod_{i=1}^k \mu(d_i) \sum_{a_i|d_i} \frac{y_{a_1,\ldots,a_k}}{\prod_{i=1}^k \phi(a_i)}.$$

Hence

$$\lambda_{d_1,\ldots,d_k} = \Big(\prod_{i=1}^k \mu(d_i) d_i\Big) \sum_{d_i|a_i} \frac{y_{a_1,\ldots,a_k}}{\prod_{i=1}^k \phi(a_i)}. \tag{3.8}$$

Remark 3.2.6. *With any choice of $y_{a_1,\ldots,a_k}$ satisfying (3.2), we can get a suitable wight function $\lambda_{d_1,\ldots,d_k}$. We comment that the conditions (3.2) will be satisfied if we take $y_{a_1,\ldots,a_k} = F(\frac{a_1}{\log R}, \ldots, \frac{a_1}{\log R})$, where F is a smooth function supported on $\mathcal{R}_k$, a_i square-free for all $1 \leq i \leq k$, and $\prod_{i=1}^k a_i < R$.*

The change of variable above, gives

$$M_1 = \frac{x}{D} \sum_{u_1,\ldots,u_k} \prod_{i=1}^k \phi(u_i) \Big(\prod_{i \neq j} \sum_{s_{i,j}} \mu(s_{i,j})\Big) \prod_{i=1}^k \frac{\mu(a_i)}{\phi(a_i)} \frac{\mu(b_i)}{\phi(b_i)} y_{a_1,\ldots,a_k} y_{b_1,\ldots,b_k}.$$

Recalling that from (3.6), we have

$$\phi(a_i) = \phi(u_i) \prod_{j \neq i} \phi(s_{i,j}),\ \mu(a_i) = \mu(u_i) \prod_{j \neq i} \mu(s_{i,j}),$$

$$\phi(b_i) = \phi(u_i) \prod_{j \neq i} \phi(s_{j,i}) \text{ and } \mu(b_i) = \mu(u_i) \prod_{j \neq i} \mu(s_{j,i}).$$

Then

$$\frac{\mu(a_i)}{\phi(a_i)} \frac{\mu(b_i)}{\phi(b_i)} = \frac{\mu(u_i)^2 \prod_{j \neq i} \mu(s_{j,i})^2}{\phi(u_i)^2 \prod_{j \neq i} \phi(s_{i,j})^2}$$

Hence

$$M_1 = \frac{x}{D} \sum_{u_1,\ldots,u_k} \Big(\prod_{i=1}^k \frac{\mu(u_i)^2}{\phi(u_i)}\Big) \Big(\prod_{i \neq j} \sum_{s_{i,j}} \frac{\mu(s_{i,j})}{\phi(s_{i,j})^2}\Big) y_{a_1,\ldots,a_k} y_{b_1,\ldots,b_k}.$$

Now we split the sum over $s_{i,j}$ to a sum over $s_{i,j} = 1$ for all $1 \leq i, j, \leq k$, and a sum over $s_{i,j} > 1$. By hypothesis, we have, if $s_{i,j} = 1$ then $a_i = b_i = u_i$. So, we can write

$$M_1 = \frac{x}{D} \sum_{u_1,\ldots,u_k} \Big(\prod_{i=1}^k \frac{1}{\phi(u_i)}\Big) y_{u_1,\ldots,u_k}^2 + E_M,$$

where E_M is the term coming from the contribution of $s_{i,j} > 1$. Remarking that the condition $s_{i,j} | d_i, e_j$ implies that $(s_{i,j}, D) = 1$ and $s_{i,j} < R$, we deduce that the contribution of $s_{i,j} > 1$

comes from $D_0 < s_{i,j} \leq R$. We have also that, from $(e_i, e_j) = 1$ and $(d_i, d_j) = 1$ for all $i \neq j$, we have $(s_{i,j}, s_{i,k}) = 1$ for all $k \neq j$, and $(s_{i,j}, s_{l,j}) = 1$ for all $l \neq i$. Additionally, we have the product over $i \neq j$ have $k^2 - k$ elements, then factoring by $\Big(\sum_{\substack{D_0 < s_{i,j} \leq R \\ (s_{i,j}, D)=1}} \frac{\mu(s_{i,j})^2}{\phi(s_{i,j})^2}\Big)$, we get

$$\Big(\prod_{i \neq j} \sum_{s_{i,j}} \frac{\mu(s_{i,j})}{\phi(s_{i,j})^2}\Big) \ll \Big(\sum_{\substack{D_0 < s_{i,j} \leq R \\ (s_{i,j}, D)=1}} \frac{\mu(s_{i,j})^2}{\phi(s_{i,j})^2}\Big)\Big(\sum_{\substack{1 \leq m \leq R \\ (m,D)=1}} \frac{\mu(m)^2}{\phi(m)^2}\Big)^{k^2-k-1}$$

Hence

$$\begin{aligned} E_M &\ll \frac{x}{D} \sum_{u_1,\ldots,u_k} \Big(\prod_{i=1}^{k} \frac{\mu(u_i)^2}{\phi(u_i)}\Big)\Big(\sum_{\substack{D_0 < s_{i,j} \leq R \\ (s_{i,j}, D)=1}} \frac{\mu(s_{i,j})^2}{\phi(s_{i,j})^2}\Big)\Big(\sum_{\substack{1 \leq m \leq R \\ (m,D)=1}} \frac{\mu(m)^2}{\phi(m)^2}\Big)^{k^2-k-1} y_{a_1,\ldots,a_k} y_{b_1,\ldots,b_k} \\ &\ll y_{\max}^2 \frac{x}{D}\Big(\sum_{\substack{1 \leq u \leq R \\ (u,D)=1}} \frac{\mu(u_i)^2}{\phi(u_i)}\Big)^k \Big(\sum_{\substack{D_0 < s_{i,j} \leq R \\ (s_{i,j}, D)=1}} \frac{\mu(s_{i,j})^2}{\phi(s_{i,j})^2})\Big)\Big(\sum_{\substack{1 \leq m \leq R \\ (m,D)=1}} \frac{\mu(m)^2}{\phi(m)^2}\Big)^{k^2-k-1}. \end{aligned}$$

In order to continue our analysis, we will need the following lemma.

Lemma 3.2.7. *Let x be a large positive real number.*

1. *We have*

$$\sum_{n \leq x} \frac{\mu(n)^2}{\phi(n)} \ll \log x.$$

2. *Let D be a square-free integer, then we have*

$$\sum_{\substack{n \leq x \\ (n,D)=1}} \frac{\mu(n)^2}{\phi(n)} \ll \frac{\phi(D)}{D} \log x.$$

Proof. 1. For the first estimate we apply the Abel's summation formula to $\sum_{n \leq x} \frac{\mu(n)^2 n}{\phi(n)} \frac{1}{n}$, remarking that

$$\sum_{n \leq x} \frac{\mu(n)^2}{\phi(n) n} \leq \zeta(2).$$

We find

$$\sum_{n \leq x} \frac{\mu(n)^2 n}{\phi(n)} \leq x \sum_{n \leq x} \frac{\mu(n)^2}{\phi(n) n} + O\Big(\sum_{n \leq x} \frac{1}{\phi(n)}\Big) \ll x.$$

Hence

$$\sum_{n \leq x} \frac{\mu(n)^2}{\phi(n)} \ll \frac{x}{x} + \int_1^x t \times \frac{1}{t^2} dt \ll \log x.$$

2. It is easy to see that

$$\Big(\sum_{\substack{n\le x\\(n,D)=1}}\frac{\mu(n)^2}{\phi(n)}\Big)\Big(\sum_{d|D}\frac{\mu(d)^2}{\phi(d)}\Big)\le\sum_{n\le Dx}\frac{\mu(n)^2}{\phi(n)}.$$

For D square-free, we have

$$\sum_{d|D}\frac{\mu(d)^2}{\phi(d)}=\frac{\mu(D)^2D}{\phi(D)}=\frac{D}{\phi(D)},$$

then the result follows from the first estimate.

□

Remarking that

$$\sum_{\substack{D_0<s_{i,j}\le R\\(s_{i,j},D)=1}}\frac{\mu(s_{i,j})^2}{\phi(s_{i,j})^2}\ll\frac{1}{D_0},$$

and

$$\Big(\sum_{\substack{1\le m\le R\\(m,D)=1}}\frac{\mu(m)^2}{\phi(m)^2}\Big)^{k^2-k-1}\ll 1.$$

Then, by the Lemma 3.2.7, we have

$$E_M\ll y_{\max}^2\frac{x}{DD_0}\Big(\frac{\phi(D)(\log R)}{D}\Big)^k.$$

Now we can write

$$S_1=\frac{x}{D}\sum_{u_1,\dots,u_k}\frac{y^2_{u_1,\dots,u_k}}{\prod_{i=1}^k\phi(u_i)}+O\Big(x\frac{y^2_{\max}\phi(D)^k(\log R)^k}{D^{k+1}D_0}\Big)+O\Big(\lambda^2_{max}R^2(\log R)^{2k}\Big). \tag{3.9}$$

It remains to compare the error terms above. Recall that

$$\lambda_{d_1,\dots,d_k}=\Big(\prod_{i=1}^k\mu(d_i)d_i\Big)\sum_{a_i|d_i}\frac{y_{a_1,\dots,a_k}}{\prod_{i=1}^k\phi(a_i)}.$$

Hence

$$\lambda_{\max}\quad\le\quad y_{\max}\sup_{a_1,\dots,a_k}\Big(\prod_{i=1}^k d_i\Big)\sum_{\substack{a_1,\dots,a_k\\d_i|a_i}}\frac{\mu(a_i)^2}{\phi(a_i)}$$

$$\begin{aligned}
&\le y_{\max}\prod_{i=1}^{k}\frac{d_i}{\phi(d_i)}\Big(\sum_{n\le R/\prod d_i}\frac{\mu(n)^2\tau_k(n)}{\phi(n)}\Big)\\
&\le y_{\max}\Big(\prod_{i=1}^{k}\sum_{l|d_i}\frac{\mu(l)^2}{\phi(l)}\Big)\Big(\sum_{n\le R/\prod d_i}\frac{\mu(n)^2\tau_k(n)}{\phi(n)}\Big) \qquad \Big(\text{because } \frac{d_i}{\phi(d_i)}=\sum_{l|d_i}\frac{\mu(l)^2}{\phi(l)}\Big)\\
&\le y_{\max}\sum_{n\le R}\frac{\mu(n)^2\tau_k(n)}{\phi(n)}\le y_{\max}\Big(\sum_{n\le R}\frac{\mu(n)^2\tau_k(n)}{\phi(n)}\Big)^k \ll y_{\max}(\log R)^k.
\end{aligned}$$

Combining the two error terms in (3.9), we get

$$S_1=\frac{x}{D}\sum_{u_1,\dots,u_k}\Big(\prod_{i=1}^{k}\frac{1}{\phi(u_i)}\Big)y^2_{u_1,\dots,u_k}+O\Big(x\frac{y^2_{\max}\phi(D)^k(\log R)^k}{D^{k+1}D_0}+y^2_{\max}R^2(\log R)^{4k}\Big). \tag{3.10}$$

By hypothesis, we have $R=x^{1/2-\delta}$ and $D\ll(\log\log x)^2$, and hence

$$S_1=\frac{x}{D}\sum_{u_1,\dots,u_k}\frac{y^2_{u_1,\dots,u_k}}{\prod_{i=1}^{k}\phi(u_i)}+O\Big(x\frac{y^2_{\max}\phi(D)^k(\log R)^k}{D^{k+1}D_0}\Big) \tag{3.11}$$

3.3 Sums of multiplicative functions

Now taking $y_{a_1,\dots,a_k}=F(\frac{a_1}{\log R},\dots,\frac{a_1}{\log R})$, where F is a smooth function supported on $\mathcal{R}_k$, a_i square-free integers for all $1\le i\le k$, and $\prod_{i=1}^{k}a_i<R$. By the Corollary 3.2.5, F defines a weight function $\lambda_{d_1,\dots,d_k}$ satisfying the conditions (3.2). Hence, the main term in S_1 becomes

$$\mathcal{M}_1=\frac{x}{D}\sum_{u_1,\dots,u_k}\frac{F(\frac{u_1}{\log R},\dots,\frac{u_1}{\log R})^2}{\prod_{i=1}^{k}\phi(u_i)}.$$

We have that u_i is a square-free for all i, hence we can also write

$$\mathcal{M}_1=\frac{x}{D}\sum_{u_1,\dots,u_k}\prod_{i=1}^{k}\frac{\mu(u_i)}{\phi(u_i)}F(\frac{u_1}{\log R},\dots,\frac{u_1}{\log R})^2.$$

Our main goal in this section is to get an accurate estimate of the main term in S_1. Indeed, using the lemma 3.2.7, we can get

$$\mathcal{M}_1\ll y^2_{\max}\frac{x}{D}\Big(\sum_{(u\le R(u,D)=1)}\frac{\mu(u)^2}{\phi(u)}\Big)^k\ll y^2_{\max}\frac{x}{D}\Big(\frac{\phi(D)(\log R)}{D}\Big)^k.$$

We see that we gained a factor D_0 comparing with the error term. Recall that our main goal is to maximize the ratio $\frac{S_2}{\rho S_1}$, but unfortunately, this rough estimate will not serve our purpose.

In order to get a better estimate, we will introduce some tools from Analytic Number Theory [10].

Definition 3.3.1. *Let f be an arithmetic function, the Dirichlet series associated to f, is formally defined by*

$$D_f(s) = \sum_{n \geq 1} \frac{f(n)}{n^s},$$

where s is a complex variable. The von Mangolt function Λ_f associated to f is defined by the formal equality

$$-\frac{D_f'(s)}{D_f(s)} = \sum_{n \geq 1} \frac{\Lambda_f(n)}{n^s}.$$

Remark 3.3.2. *1. For $f(n) = 1$ for all $n \geq 1$, we have $D_f(s) = \zeta(s)$ and $\Lambda_f(n) = \Lambda(n)$.*

2. If f is a multiplicative function and $Re(s) > 1$, then $D_f(s)$ has an Euler product

$$D_f(s) = \prod_p \Big(1 + \frac{f(p)}{p^s} + \frac{f(p)}{p^{2s}} + \dots\Big),$$

and

$$\Lambda_f(n) = \Lambda(n) f(n).$$

Lemma 3.3.3. *Let f be a multiplicative function satisfying*

$$\sum_{n \leq x} \Lambda_f(n) = \kappa \log x + O(1),$$

and

$$\sum_{n \leq x} |f(n)| \ll (\log x)^k,$$

where $\kappa > -1/2$. Then

$$\sum_{n \leq x} f(n) = \frac{\mathcal{G}_f}{\Gamma(\kappa + 1)} (\log x)^k + O(\log x)^{|\kappa| - 1},$$

where

$$\mathcal{G}_f = \prod_p (1 - \frac{1}{p})^k (1 + f(p) + f(p^2 + \dots),$$

and Γ the Euler Gamma function.

Proof. For the proof see [[10],Theorem 1.1]

□

The following corollary follows from the Lemma above.

Corollary 3.3.4. *Let f be a multiplicative function satisfying the conditions in (3.3.3), and F be a smooth function on $[0,1]$, with*

$$F_{\max} = \sup_{x\in[0,1]} |F(x)| + |F'(x)|.$$

Then

$$\sum_{n\le x} f(n)F\big(\frac{\log n}{\log x}\big) = \frac{\mathcal{G}_f \log^{\kappa} x}{\Gamma(\kappa)} \int_0^1 t^{\kappa-1}F(t)dt + O(F_{\max}(\log x)^{\kappa-1}).$$

We establish the corollary using the summation formula and replacing $\sum_{n\le x} f(n)$ by its value in the Lemma (3.3.3), for the error term, we use an integration by parts with a change of variable $u = x^t$.

Lemma 3.3.5. *Let F be a smooth function on $[0,1]$, with*

$$F_{\max} = \sup_{x\in[0,1]} |F(x)| + |F'(x)|.$$

Then

$$\sum_{\substack{n\le x\\(n,D)=1}} \frac{\mu(n)^2}{\phi(n)} F\big(\frac{\log n}{\log x}\big)^2 = \frac{\phi(D)\log x}{D}\int_0^1 F(t)^2 dt + O(F_{max}),$$

where D is a square-free integer.

Proof. We apply the corollary 3.3.4 to $f(n) = \frac{\mu(n)^2}{\phi(n)}$ supported on n which is co-prime to D. From the estimate (3.2.7), we have

$$\sum_{n\le x} |f(n)| \ll (\log x),$$

and

$$\sum_{n\le x} \Lambda_f(n) = \sum_{n\le x} \Lambda(n) f(n) = \sum_{p\le x} \frac{\log p}{p-1} \ll \log x + O(1).$$

Recalling that f is supported on square-free integers, we get

$$\mathcal{G}_f = \prod_{p\nmid D} \big(1-\frac{1}{p}\big)^{-1} \prod_{p} \big(1-\frac{1}{p}\big) = \prod_{p\mid D} \frac{p-1}{p} = \frac{\phi(D)}{D}.$$

□

With k applications of Lemma 3.3.5, we obtain

$$\sum_{u_1,\dots,u_k}\prod_{i=1}^{k}\frac{\mu(u_i)}{\phi(u_i)}F\Big(\frac{\log u_1}{\log x},\dots,\frac{\log u_k}{\log x}\Big)^2=\Big(\frac{\phi(D)\log x}{D}\Big)^k\int_0^1\dots\int_0^1 F(t_1,\dots.t_k)dt_1\dots dt_k, \tag{3.12}$$

where

$$F_{\max}=\sup_{(t_1,\dots,t_k)\in\mathcal{R}_k}\Big(|F(t_1,\dots,t_k)|+\sum_{i=1}^{k}|\frac{\partial F}{\partial t_i}(t_1,\dots,t_k)|\Big).$$

Remark 3.3.6. *To get the estimate (3.12), Maynard used the Lemma 3 in ([7]).*

Let now

$$y_{u_1,\dots,u_k}=F\Big(\frac{\log u_1}{\log x},\dots,\frac{\log u_k}{\log x}\Big)$$

Recalling that

$$S_1=\frac{x}{D}\sum_{u_1,\dots,u_k}\frac{y^2_{u_1,\dots,u_k}}{\prod_{i=1}^k\phi(u_i)}+O\Big(x\frac{y^2_{\max}\phi(D)^k(\log R)^k}{D^{k+1}D_0}\Big)$$

by substitution, S_1 becomes

$$S_1=\frac{x}{D}\sum_{\substack{u_1,\dots,u_k\\(u_i,u_j)=1\\(u_i,D)=1}}\frac{y^2_{u_1,\dots,u_k}}{\prod_{i=1}^k\phi(u_i)}+O\Big(x\frac{y^2_{\max}\phi(D)^k(\log R)^k}{D^{k+1}D_0}\Big)$$

To apply the estimate (3.12), we should definitely discard the condition $(u_i,u_j)=1$ from the sum in the main term above. To do so, we multiply again by the discontinuous factor

$$\sum_{s'_{i,j}|u_i,u_i}\mu(s'_{i,j}=\begin{cases}1 \text{ if } (u_i,u_j)=1\\ 0 \text{ otherwise .}\end{cases} \tag{3.13}$$

Then we proceed as in 3.3.5, recalling that this cancellations will cost an error term of size $\frac{F^2_{\max}\phi(D)^k x(\log R)^k}{D^{k+1}D_0}$. Hence

$$S_1=x\frac{\phi(D)^k(\log R)^k}{D^{k+1}}\int_0^1\dots\int_0^1F(t_1,\dots.t_k)^2dt_1\dots dt_k+O\Big(\frac{F^2_{\max}\phi(D)^kx(\log R)^k}{D^{k+1}D_0}\Big),$$

This completes the proof of the first part of Proposition 3.1.1.

Recalling that

$$S_2=\sum_{\substack{x<n\le 2x\\ n\equiv m \pmod D}}\Big(\sum_{i=1}^{k}\chi(n+h_i)\Big)\Big(\sum_{\substack{d_1\dots,d_k\\ d_i|n+h_i}}\lambda_{d_1,\dots,d_k}\Big)^2,$$

we write

$$S_2 = \sum_{l=1}^{k} S_2^{(l)},$$

where

$$S_2^{(l)} = \sum_{\substack{x<n\leq 2x \\ n\equiv m \pmod D}} \chi(n+h_l)\Big(\sum_{\substack{d_1\ldots,d_k \\ d_i|n+h_i}} \lambda_{d_1,\ldots,d_k}\Big)^2.$$

Hence

$$S_2^{(l)} = \sum_{\substack{d_1\ldots,d_k \\ e_1,\ldots,e_k}} \lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k} \sum_{\substack{x<n\leq 2x \\ n\equiv m \pmod D \\ [d_i,e_i]|n+h_i}} \chi(n+h_l).$$

We can restrict again our sum to the elements e_i and d_i such that $(e_1, d_i) = 1$ and $(d_i, D) = 1$; additionally, the condition $\chi(n + h_l) \neq 0$ and $[d_i, e_i]|n + h_l$, implies $d_i = e_i = 1$

By the Chinese Remainder Theorem, we can see that the inner sum counts the primes in the arithmetic progression $m \pmod{D\prod_{i=1}^{k}[d_i, e_i]}$. That justifies that the accurate choice of m is to take $(m, D\prod_{i=1}^{k}[d_i, e_i]) = 1$. Hence, we write

$$\sum_{\substack{x<n\leq 2x \\ n\equiv m \pmod D \\ [d_i,e_i]|n+h_i}} \chi(n+h_l) = \frac{X_x}{\phi(D\prod_{i=1}^{k}[d_i,e_i])} + O(E(x, D\prod_{i=1}^{k}[d_i,e_i])),$$

where

$$E(x,D) = \sup_{(m,q)=1}\max_{m\leq D}\Big|\sum_{\substack{x<n\leq 2x \\ n\equiv m \pmod D}} \chi(n) - \frac{1}{\phi(D)}\sum_{x<n\leq 2x}\chi(n)\Big|,$$

and

$$X_x = \sum_{x<n\leq 2x} \chi(n).$$

Hence

$$S_2^{(l)} = \frac{X_x}{\phi(D)} \sum_{\substack{d_1\ldots,d_k \\ e_1\ldots,e_k \\ (d_i,e_i)=1 \\ d_l=e_l=1}} \frac{\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}}{\prod_{i=1}^{k}\phi([d_i,e_i])} + O\Big(\sum_{\substack{d_1\ldots,d_k \\ e_1,\ldots,e_k}} \lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k} E(x,q)\Big),$$

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

By hypothesis, we have that q is a square-free integer, and from $\prod_i d_i < R$, we deduce $q < DR^2$. Then using the fact that the number of ways of writing $D\prod_{i=1}^k [d_i, e_i] = r$, where r is a square-free integer is bounded by $\tau_k(r)$. Hence

$$\sum_{\substack{d_1\dots,d_k\\ e_1,\dots,e_k}} \lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k} E(x,q) \ll \lambda_{\max}^2 \sum_{r<DR^2} \mu(r)^2\tau_k(r)E(x,r).$$

We prove the following proposition.

Proposition 3.3.7. *Let*

$$y^{(m)}_{r_1,\dots,r_k} = \Big(\prod_{i=1}^k \mu(r_i)g(r_i)\Big) \sum_{s_{i,j}|d_i,e_i} \mu(s_{i,j})\Big) \sum_{\substack{d_1\dots,d_k\\ d_m=1\\ r_i|d_i}} \frac{\lambda'_{d_1,\dots,d_k}}{\prod_{i=1}^k \phi(d_i)},$$

where g is the totally multiplicative function defined on primes by $g(p) = p-2$. Let $y^{(m)}_{\max} = \sup_{r_1,\dots,r_k} |y^{(m)}_{r_1,\dots,r_k}|$. Then for any fixed $A>0$, we have

$$S_2^{(l)} = \frac{x}{\phi(D)\log x} \sum_{\substack{u_1\dots,u_k\\ u_l=1}} \frac{(y^{(l)}_{\max})^2}{\prod_{i=1}^k g(u_i)} + O\Big(x\frac{y^{(l)}_{\max}\phi(D)^{k-2}(\log R)^{k-2}}{D^{k-1}D_0}\Big).$$

Proof. By the Prime Number Theorem in arithmetic progressions, we have $E(x,D) \ll \frac{x}{\phi(D)}$, but as we mentioned in the previous chapter, assuming that the primes have a level of distribution θ, we have$\sum_{q\le x^{\theta-\delta}} E(x,D) \ll \frac{x}{(\log x)^A}$, for any fixed $A>0$. Thus using the Cauchy-Shwarz inequality, and recalling that $\lambda_{max} \ll y^2_{\max}(\log R)^{2k}$, we obtain

$$\begin{aligned}
\sum_{\substack{d_1\dots,d_k\\ e_1,\dots,e_k}} \lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k} E(x,q) &\ll \lambda_{\max}^2 \sum_{r<DR^2} \mu(r)^2\tau_k(r)E(x,r)^{1/2}E(x,r)^{1/2}\\
&\ll \lambda_{\max}^2\Big(\sum_{r<DR^2} \mu(r)^2\tau_k(r)^2\frac{x}{\phi(r)}\Big)^{1/2}\Big(\sum_{r<DR^2}\mu(r)^2E(x,r)\Big)^{1/2}\\
&\ll x^{1/2}(\log x)^{6k}\frac{x^{1/2}}{(\log x)^{1/2}} \ll y^2_{\max}\frac{x}{(\log x)^A}.
\end{aligned}$$

Now we will study the main term in $S_2^{(l)}$. First, we drop the condition $(e_i,d_j)=1$ using the same factor defined in (3.13). Recall that in the estimation of the main term in S_1, we replaced

$\frac{1}{[d_i,e_i]}$ by $\frac{1}{d_1 e_1}\sum_{u_i|d_i,e_i}\phi(u_i)$. Using the same technique, we define g to be an arithmetic function satisfying

$$\frac{1}{\phi((d_i,e_i))}=\frac{1}{\phi(d_1)\phi(e_1)}\sum_{u_i|d_i,e_i} g(u_i).$$

It's not hard to prove that for d_i and e_i square-free, we have

$$\frac{\phi(d_i)\phi(e_i)}{\phi([d_i,e_i])}=\phi((d_i,e_i)).$$

By Möbius inversion formula, we find that

$$g((d_i,e_i))=\sum_{r|d_i,e_i}\mu(r)\phi((d_i,e_i)/r).$$

Hence, g is multiplicative and $g(p)=p-2$.

This gives a main term of

$$M_2^{(l)}=\frac{X_x}{\phi(D)}\sum_{\substack{d_1\dots,d_k\\ e_1\dots,e_k\\ d_l=e_l=1}}\Big(\prod_{i\neq j}\sum_{s_{i,j}|d_i,e_i}\mu(s_{i,j})\Big)\frac{\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k\phi(d_i)\phi(e_i)}\sum_{u_i|d_i,e_j} g(u_i).$$

Using the same change of variable above

$$a_j=u_j\prod_{i\neq j}s_{j,i},\text{ and } b_j=u_j\prod_{i\neq j}s_{i,j}$$

we obtain

$$M_2^{(l)}=\frac{X_x}{\phi(D)}\sum_{\substack{u_1\dots,u_k\\ u_l=1}}\prod_{i=1}^k g(u_i)\Big(\prod_{\substack{i\neq j\\ i,j\neq l}}\sum_{s_{i,j}|d_i,e_i}\mu(s_{i,j})\Big)\sum_{\substack{d_1\dots,d_k\\ d_l=1\\ a_i|d_i}}\frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^k\phi(d_i)}\sum_{\substack{e_1\dots,e_k\\ e_l=1\\ b_i|e_i}}\frac{\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k\phi(e_i)}.$$

Now we make the following invertible change of variable

$$y^{(m)}_{r_1,\dots,r_k}=\Big(\prod_{i=1}^k\mu(r_i)g(r_i)\Big)\sum_{\substack{d_1\dots,d_k\\ d_m=1\\ r_i|d_i}}\frac{\lambda'_{d_1,\dots,d_k}}{\prod_{i=1}^k\phi(d_i)}.$$

Remark 3.3.8. *We denoted $\lambda_{d_1,\dots,d_k}$ by $\lambda'_{d_1,\dots,d_k}$ to avoid the confusion.*

Substituting this change of variable into $M_2^{(l)}$, we obtain

$$M_2^{(l)} = \frac{X_x}{\phi(D)} \sum_{\substack{u_1\dots,u_k \\ u_l=1}} \prod_{i=1}^{k} \frac{\mu(u_i)^2}{g(u_i)} \Big(\prod_{\substack{i\neq j \\ i,j\neq l}} \sum_{s_{i,j}|d_i,e_i} \frac{\mu(s_{i,j})}{g(s_{i,j})^2} \Big) y^{(l)}_{a_1,\dots,a_k} y^{(l)}_{b_1,\dots,b_k}.$$

As we have seen before, splitting the sum on $s_{i,j}$ to $s_{i,j}=1$ and $s_{i,j}>D_0$, $M_2^{(l)}$ becomes

$$M_2^{(l)} = \frac{X_x}{\phi(D)} \sum_{\substack{u_1\dots,u_k \\ u_l=1}} \Big(\prod_{i=1}^{k} \frac{\mu(u_i)^2}{g(u_i)} \Big) (y^l_{\max})^2 + O\Big(x \frac{y^l_{\max}\phi(D)^{k-2}(\log R)^{k-2}}{D^{k-1}D_0} \Big).$$

(Remark that here we have $k-2$ instead of $k-1$, this shift of exponents arises naturally from the condition $u_l=1$).

Now, we can write

$$S_2^{(l)} = \frac{X_x}{\phi(D)} \sum_{\substack{u_1\dots,u_k \\ u_l=1}} \frac{(y^{(l)}_{\max})^2}{\prod_{i=1}^k g(u_i)} + O\Big(x \frac{y^{(l)}_{\max}\phi(D)^{k-2}(\log R)^{k-2}}{D^{k-1}D_0} \Big) + O\big(x \frac{(\lambda'_{\max})^2}{(\log x)^A} \big).$$

Using the inversion formula, and with the same argument we used to compare $\lambda_{\max}$ and $y_{\max}$, we prove that

$$(\lambda'_{\max})^2 \ll (y^{(l)}_{\max})^2 (\log R)^{2k}.$$

Hence

$$S_2^{(l)} = \frac{X_x}{\phi(D)} \sum_{\substack{u_1\dots,u_k \\ u_l=1}} \frac{(y^{(l)}_{\max})^2}{\prod_{i=1}^k g(u_i)} + O\Big(x \frac{y^{(l)}_{\max}\phi(D)^{k-2}(\log R)^{k-2}}{D^{k-1}D_0} \Big).$$

By the Prime Number Theorem

$$X_x = \frac{x}{\log x} + O(x/(\log x)^2).$$

The contribution of the error term of X_x in $S_2^{(l)}$, is of size

$$\begin{aligned} O\Big(x \frac{(y^{(l)}_{\max})^2}{\phi(D)(\log x)^2} \Big(\sum_{\substack{u<R \\ (u,D)=1}} \frac{\mu(u)^2}{g(u)} \Big)^{k-1} \Big) &= O\Big(x \frac{(y^{(l)}_{\max})^2}{\phi(D)(\log x)^2} \Big(\frac{\phi(D)\log x}{D} \Big)^{k-1} \Big) \\ &= O\Big(x \frac{(y^{(l)}_{\max})^2 \phi(D)^{k-2}(\log x)^{k-3}}{D^{k-2}} \Big). \end{aligned}$$

Hence

$$S_2^{(l)} = \frac{x}{\phi(D)\log x} \sum_{\substack{u_1\dots,u_k \\ u_l=1}} \frac{(y^{(l)}_{\max})^2}{\prod_{i=1}^k g(u_i)} + O\Big(x \frac{y^{(l)}_{\max}\phi(D)^{k-2}(\log R)^{k-2}}{D^{k-1}D_0} \Big).$$

This completes the proof of Proposition 3.3.7. $\square$

The following lemma give a relation between $y^{(m)}_{r_1,\dots,r_k}$ and $y_{r_1,\dots,r_k}$.

Lemma 3.3.9. *If $r_m = 1$, then*

$$y^{(m)}_{r_1,\dots,r_k} = \sum_{a_m} \frac{y_{r_1,\dots,r_{m-1},a_m,r_{m+1},\dots,r_k}}{\phi(a_m)} + O\Big(\frac{y_{\max}\phi(D)\log R}{DD_0}\Big).$$

Proof. Recalling that

$$y^{(m)}_{r_1,\dots,r_k} = \Big(\prod_{i=1}^{k} \mu(r_i)g(r_i)\Big) \sum_{\substack{d_1\dots,d_k \\ d_m=1 \\ r_i|d_i}} \frac{\lambda'_{d_1,\dots,d_k}}{\prod_{i=1}^{k}\phi(d_i)},$$

and

$$\lambda_{d_1,\dots,d_k} = \Big(\prod_{i=1}^{k} \mu(d_i)d_i\Big) \sum_{a_i|d_i} \frac{y_{a_1,\dots,a_k}}{\prod_{i=1}^{k}\phi(a_i)},$$

we obtain

$$\begin{aligned} y^{(m)}_{r_1,\dots,r_k} &= \Big(\prod_{i=1}^{k} \mu(r_i)g(r_i)\Big)\Big) \sum_{\substack{d_1\dots,d_k \\ d_m=1 \\ r_i|d_i,d_i|a_i}} \prod_{i=1}^{k} \frac{\mu(d_i)d_i}{\phi(d_i)} \\ &= \Big(\prod_{i=1}^{k} \mu(r_i)g(r_i)\Big)\Big) \sum_{\substack{d_1\dots,d_k \\ d_m=1 \\ r_i|d_i,d_i|a_i}} \prod_{i\neq m} \frac{\mu(a_i)a_i}{\phi(a_i)}. \end{aligned}$$

Considering the support of $y_{r_1,\dots,r_k}$, only the terms a_i, such that $(a_i, D) = 1$ contribute to the above sum. Then, again splitting the sum over a_i to $a_i = r_i$ and $a_i > D_0 r_j$, by Lemma **??** and the same argument used in 3.3.7, we find

$$y^{(m)}_{r_1,\dots,r_k} = \Big(\prod_{i=1}^{k} \frac{g(r_i)r_i}{\phi(r_i)^2}\Big) \sum_{a_m} \frac{y_{r_1,\dots,r_{m-1},a_m,r_{m+1},\dots,r_k}}{\phi(a_m)} + O\Big(\frac{y_{\max}\phi(D)\log R}{DD_0}\Big).$$

But we have that $g(p) = p - 2$, hence $\frac{g(p)p}{\phi(p)^2} = 1 + O(\frac{1}{p^2})$, and for any prime p such that $p|r_i$, we have $p > D_0$. Hence

$$\prod_{i=1}^{k} \frac{g(r_i)r_i}{\phi(r_i)^2} = 1 + O(\frac{1}{D_0}).$$

$\square$

Without loss of generality, we take $l = k$. Recalling that

$$y_{r_1,\dots,r_k} = F\big(\frac{\log r_1}{\log x}, \dots, \frac{\log r_k}{\log x}\big),$$

from Lemma 3.3.9, we obtain

$$y^{(k)}_{r_1,\dots,r_{k-1},1} = \sum_{\substack{a_k \le R \\ (a_k, W\prod r_i)=1}} \frac{\mu(a_k)^2}{\phi(a_k)} F\left(\frac{\log r_1}{\log R}, \cdots, \frac{\log r_{k-1}}{\log R}, \frac{\log a_k}{\log R}\right) + O\left(\frac{F_{\max}\phi(D)\log R}{DD_0}\right).$$

We remark that the main term in $y^{(k)}_{r_1,\dots,r_{k-1},1}$ is a sum of multiplicative functions, then applying 3.2.5, with $\kappa = 1$, and $f(n) = \frac{\mu(n)^2}{\phi(n)}$, we obtain

$$y^{(k)}_{r_1,\dots,r_{k-1},1} = \left(\prod_{i=1}^{k-1} \frac{\phi(r_i)}{r_i}\right) \frac{\phi(D)\log R}{D} \int_0^1 F(r_1, \dots, r_{k-1}, t)dt + O\left(\frac{F_{\max}\phi(D)\log R}{DD_0}\right).$$

Using again

$$y_{\max} \ll \frac{\phi(D)F_{\max}\log R}{D}, \tag{3.14}$$

we get

$$S_2^{=} \frac{x}{\phi(W)\log x} \sum_{\substack{r_1,\dots,r_{k-1} \\ (r_i,r_j)=1\forall i,j \\ (r_i,D)=1\forall i}} \left(\prod_{i=1}^{k-1} \frac{\mu(r_i)^2}{g(r_i)}\right) (y^{(k)}_{r_1,\dots,r_{k-1},1})^2 + O\left(x\frac{(F_{\max}^2\phi(D)^k(\log R)^k}{D^{k+1}D_0}\right).$$

We drop the condition $(r_i, r_j) = 1$ using the same argument we used to find (3.7).

Hence

$$S_2 = \frac{x}{\phi(W)\log x} \sum_{\substack{r_1,\dots,r_{k-1} \\ (r_i,D)=1\forall i}} \left(\prod_{i=1}^{k-1} \frac{\mu(r_i)^2}{g(r_i)}\right) (y^{(k)}_{r_1,\dots,r_{k-1},1})^2 + O\left(x\frac{(F_{\max}^2\phi(D)^k(\log R)^k}{D^{k+1}D_0}\right).$$

Then the result follows by $k - 1$ application of 3.2.5, with $\kappa = 1$, and $f(n) = \frac{\mu(n)^2\phi(n)^2}{g(n)n^2}$, if $(n, D\prod r_i) = 1$, 0 otherwise.

Remark 3.3.10. *Let q be a real quadratic form, i.e., q—is a homogeneous polynomial of degree 2 in k variables with coefficients in $\mathbb{R}$. Hence we can write*

$$q(x_1, \dots, x_n) = \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}x_ix_j.$$

It is well known that, we can write $q(x_1, \ldots, x_n) = v^t M v$, where $v = (x_1, \ldots, v_k)$ and $M = (a_i j)_{1 \leq i,j \leq k}$ a symmetric matrix. The theory of quadratic forms tells us that any real quadratic form over a field of characteristic different than 2 is diagonalizable, that means that we can write q in the following form

$$q(y_1, \ldots, y_k) = b_1 y_1^2 + b_2 y_2^2 + \ldots + b_k y_k^2.$$

Indeed, we can see S_1 and S_2 as quadratic forms in $\lambda_{d_1,\ldots,d_k}$, and Proposition 3.2.1 gave the diagonal form of S_1 and S_2. The same concept applies to the sums in the GPY method. This serve our purpose, which is maximizing the ratio of S_1 and S_2 to get a positive S, because the diagonal form is easier to manipulate and it's a sum of positive terms.

3.4 An optimization problem

We recall that if we show that $S(x, \rho) = S_2 - \rho S_1 > 0$, then there exist infinitely many n such that at least $\lfloor \rho + 1 \rfloor$ of the $n + h_i$ are all prime, for $1 \leq i \leq k$. We define

$$M_k = \sup_{F \in \mathcal{S}_k} \frac{\sum_{m=1}^{k} J_k^m(F)}{I_k(F)}, \tag{3.15}$$

where J_k^m and $I_k(F)$ as in 3.2.1 and

$$r_k = \lceil \frac{\theta M_k}{2} \rceil,$$

where θ is the level of distribution of primes. By hypothesis, there exist a function F_0 such that $\sum_{m=1}^{k} J_k^m(F_0) > (M_k - 2\delta) I_k(F_0)$, for a small $\delta > 0$. Hence, from Proposition 3.2.1, and taking $R = x^{\theta/2-\delta}$, we get

$$\begin{aligned} S(x,\rho) &= x \frac{\phi(D)^k (\log R)^k}{D^{k+1}} \Big(\frac{\log R}{\log x} \sum_{m=1}^{k} J_k^m(F_0) - \rho I_k(F_0) + o(1) \Big) \\ &\geq x \frac{\phi(D)^k (\log R)^k I_k(F_0)}{D^{k+1}} \Big((\theta/2 - \delta)(M_k - \delta) + o(1) \Big), \end{aligned}$$

as $x \to \infty$.

We see that to prove that $S(x, \rho) > 0$. Indeed if we take $\rho = \theta M_k/2 - \epsilon$, and choosing δ such that $1 + \epsilon/2\delta > M_k$, we will get the desired result.

That means that, to prove the unconditional bounded gap between primes (taking $\theta = 1/2$) it's enough to prove that $M_k > 4$, and assuming Elliott-Halberstam conjecture we need just $M_k > 2$. With this, our problem turns out to be a pure optimization problem.

3.4.1 Optimizing the ratio of two quadratic forms

We can see easily that $\sum_{m=1}^{k} J_k^m(F)$ and $I_k(F)$ are symmetric (i.e., if any of the variables are interchanged the value remains the same), then we can find some symmetric function $F_{\max}$ defined on $\mathcal{R}_k$, such that

$$M_k = \sup_{F \in \mathcal{S}_k} \frac{\sum_{m=1}^{k} J_k^m(F_{\max)}}{I_k(F_{\max)}}.$$

We have that if $F_{\max}$ is a symmetric and continuous function on a compact, then we can approximate it with a symmetric polynomial P. Such polynomials are polynomial expressions in the jth power-sum polynomials

$$P_j = \sum_{i=1}^{k} t_i^j.$$

Indeed Goldston, Pintz, and Yildirim argument is equivalent to choose $F = P_1$ in the actual setting. With a basic integral calculation, we can prove that this choice we cannot get $M_k > 4$. The key idea in Maynard's lower bound of M_k, is to consider symmetric polynomials in P_1 and P_2. In particular he focused focused on the polynomials of the form $(1 - P_1)^a P_2^b$. The following lemmas give formulas to calculate the integrals above.

Lemma 3.4.2. *We have*

$$\int_{\mathcal{R}_k} \left(1 - \sum_{i=1}^{k} t_i\right)^a \prod_{i=1}^{k} t_i^{b_i} dt_1 \dots dt_k = \frac{a! \prod_{i=1}^{k} b_i!}{(k + a + \sum_{i=1}^{k} b_i)!},$$

where a and b_i for $1 \le i \le k$ are positive integers.

Proof. We consider the following integral

$$\int_0^{1-\sum_2^k t_i} \left(1 - \sum_{i=1}^{k} t_i\right)^a \prod_{i=1}^{k} t_i^{b_i} dt_1.$$

Then, we insert a change of variables $u = \frac{t_1}{1-\sum_{i=2}^{k} t_i}$. Hence, using the beta function integral (i.e. $\int_0^1 t^a(1-t)^b = \frac{a!b!}{(a+b+1)!}$) we find

$$\int_0^{1-\sum_2^k t_i} \left(1 - \sum_{i=1}^{k} t_i\right)^a \prod_{i=1}^{k} t_i^{b_i} dt_1 \quad = \quad \prod_{i=2}^{k} t_i^{b_i} \left(1 - \sum_{i=2}^{k} t_i\right)^{a+b_1+1} \int_0^1 (1-u)^a u^{b_1} du$$

$$= \frac{a!b_1!}{(a+b_1+1)!}\prod_{i=2}^{k} t_i^{b_i}\left(1-\sum_{i=2}^{k} t_i\right)^{a+b_1+1}.$$

And the result follows by induction. □

Lemma 3.4.3. *Let P_j denotes the j^{th} symmetric power polynomial in $t_1,\ldots,t_k$, then*

$$\int_{\mathcal{R}_k}(1-P_1)^a P_2^b dt_1\ldots dt_k = \frac{a!}{k+a+2b!}\sum_{b_1+\ldots+b_k=b}\frac{b!}{b_1!\ldots b_k!}\prod_{i=1}^{k}(2b_i)!.$$

Proof. By the multinomial theorem,

$$P_2^b = \sum_{b_1+\ldots+b_k=b}\frac{b!}{b_1!\ldots b_k!}\prod_{i=1}^{k} t^{2b_i}.$$

Using Lemma 3.4.2, we have

$$\begin{aligned}\int_{R_k}(1-P_1)^a P_2^b dt_1\ldots dt_k &= \int_{R_k}(1-P_1)^a\sum_{b_1+\ldots+b_k=b}\frac{b!}{b_1!\ldots b_k!}\prod_{i=1}^{k} t^{2b_i}dt_1\ldots dt_k\\ &= \sum_{b_1+\ldots+b_k=b}\frac{b!}{b_1!\ldots b_k!}\int_{R_k}(1-P_1)^a\prod_{i=1}^{k} t^{2b_i}dt_1\ldots dt_k\\ &= a!\sum_{b_1+\ldots+b_k=b}\frac{b!}{b_1!\ldots b_k!}\frac{\prod_{i=1}^{k}(2b_i)!}{(k+a+\sum_{i=1}^{k}2b_i)!}.\end{aligned}$$

□

Taking

$$F = \sum_{i=1}^{k} a_i b_i (1-P_1)^{\alpha_i}P_2^{\beta_i},$$

we find

$$I_k(F) = \sum_{i,j=1}^{k} a_i b_i \int_{R_k}(1-P_1)^{\alpha_i+\alpha_j}P_2^{\beta_i+\beta_j}.$$

Then we apply the previous lemma. For $J_k^m(F)$ we apply 3.4.2.

Indeed we expressed $J_k^m(F)$ and $I_k(F)$ as positive definite quadratic forms. Hence, we can write

$$J_k^m(F) = v^T M_2^{(m)} v \quad , \text{ and } I_k(F) = v^T M_1 v,$$

where M_1 and $M_2^{(m)}$ are symmetric positive definite matrices. To get a lower bound for M_k, we should maximize the ratios $\frac{v^T M_2^{(m)}}{v^T M_1 v}$. We recall the definition of the norm of a matrix M

$$\|M\| = \sup_{\|v\|=1} \|Mv\|.$$

From linear Algebra, we can write

$$M = ADB^T,$$

where $D = diag(\lambda_1, \dots, \lambda_k)$, and A, D are two orthonormal matrices (i.e., $\|A\| = \|B\| = 1$). Taking $v = (v_1, \dots, v_k)$, we get

$$\|M\| = \sup_{\|v\|=1} \|Dv\|.$$

By the usual inner product, we have that

$$\|Dv\|^2 = \sum_{i=1}^{k} \lambda_i^2 v_i^2.$$

We want to find $\sup \|Dv\|$, under the constraint $\sum_i v^2 = 1$. This holds, when $\lambda_l = \max |\lambda_i|$, $v_l = 1$, and $v_i = 0$ for $i \neq l$. Hence

$$\|M\| = \lambda_{\max}.$$

By hypothesis, we have that M_1 is a symmetric positive definite matrix, so M_1 defines an inner product

$$< a, b >_{M_1} = a^T M_1 b.$$

Using this product, we find

$$\|M_1^{-1}(M_2^{(m)})v\| = v^T M_1 M_1^{-1}(M_2^{(m)})v = v^T (M_2^{(m)})v.$$

Hence, $v^T(M_2^{(m)})v$ is maximal when v is the eigenvector corresponding the maximal eigenvalue of $M_1^{-1}(M_2^{(m)})$.

Theorem 3.4.4. *We have*

$$\liminf_{n\to\infty} p_{n+1} - p_n = 600$$

Proof. Taking F to be a linear combination of symmetric polynomials in the form $(1 - P_1)^a P_2^b$. To work in a finite dimensional vector space, we impose a bound on a and b. We let $a+2b \leq 11$, so finding F is equivalent to work in a vector space of dimension 42 spanned by $(1-P_1)^a P_2^b$. we use the integration formulas in 3.4.2 to calculate M_1 and $M_2^{(m)}$, and then we calculate the maximal eigenvalue λ of $M_1^{-1}M_2^{(m)}$. This process is feasible by computer. indeed Maynard was able to get that for $k = 105$ and $a + 2b \leq 11$

$$M_1 05 \geq 4.0020679 > 4$$

. To complete the proof, we should find a 105-tuple. By an exhaustive search we can prove that the minimal diameter of an admissible 105-tuple is 600, explicitly the following 105-tuple ([13]) $\mathcal{H} = \{0, 10, 12, 24, 28, 30, 34, 42, 48, 52, 54, 64, 70, 72, 78, 82, 90, 94, 100, 112, 114,$
$118, 120, 124, 132, 138, 148, 154, 168, 174, 178, 180, 184, 190, 192, 202, 204, 208, 220,$
$222, 232, 234, 250, 252, 258, 262, 264, 268, 280, 288, 294, 300, 310, 322, 324, 328,$
$330, 334, 342, 352, 358, 360, 364, 372, 378, 384, 390, 394, 400, 402, 408, 412, 418, 420,$
$430, 432, 442, 444, 450, 454, 462, 468, 472, 478, 484, 490, 492, 498, 504, 510, 528, 532,$
$534, 538, 544, 558, 562, 570, 574, 580, 582, 588, 594, 598, 600\}$, is what we want to prove.

□

Theorem 3.4.5. *Assuming the Elliott-Halberstam conjecture, we have*

$$\liminf_{n\to\infty} p_{n+1} - p_n = 12.$$

Proof. Taking $k = 5$, and

$$F = (1-P_1)P_2 + 7/10(1-P_1)^2 + \frac{1}{14}P_2 - 3/14(1-P_1)$$

Maynard showed that

$$\frac{\sum_{m=1}^{k} J_k^m(F)}{I_k(F)} = \frac{1417255}{708216} = 2.00116 > 2.$$

The minimal 5-tuple is $\{0, 2, 6, 8, 12\}$. □

We saw that to get the bounded gap between primes, Maynard translated the problem to an optimization problem, and he used an optimization method based on basic linear Algebra, so it was believed that the bound in Theorem 3.1.1 could be improved combining more sophisticated tools from Optimization Theory and Analytic Number Theory. Few months after

Maynard's breakthrough, Tao administered an on-line collaboration project called Polymath8b [?] to improve Maynard's result, and they succeeded to prove that

$$\liminf_{n\to\infty} p_{n+1} - p_n = 246,$$

and assuming Elliott-Halberstam conjecture, they showed that

$$\liminf_{n\to\infty} p_{n+1} - p_n = 6.$$

We will not prove this results here, but we will present an upper bound for M_k proved by Polymath8b.

Proposition 3.4.6. *Let M_k be as defined in 3.15, we have*

$$M_k \leq \frac{k\log k}{k-1}$$

Proof. By the Cauchy-Schwarz inequality we have

$$\left(\int_{t_k=0}^{1-t_1-\dots-t_{k-1}} F(t_1,\dots,t_k)dt_k\right)^2 \leq \int_{t_k=0}^{1-t_1-\ldots-t_{k-1}} \frac{dt_k}{1-t_1-\ldots-t_{k-1}+(k-1)t_k}$$
$$\times \int_{t_k=0}^{1-t_1-\dots-t_{k-1}} F(t_1,\dots,t_k)^2(1-t_k-\ldots-t_{k-1}+(k-1)t_k)dt_k$$

Let $u = C + (k-1)t$ we have

$$\int_0^C \frac{dt}{C+(k-1)t} = \frac{1}{k-1}\int_C^{kC}\frac{du}{u} = \frac{\log k}{k-1}.$$

Taking $C = 1 - t_1 - \ldots - t_k$, $t = t_k$, and , we find by induction that

$$\sum_{m=1}^{k} J_m^k(F) \leq \frac{\log k}{k-1}\int_{\mathcal{R}_k} F(t_1,\dots,t_k)^2\sum_{j=1}^{k}(1-t_1-\ldots-t_k+kt_j)dt_k\ldots dt_1.$$

Recalling that on $\mathcal{R}_k$, we have that $\sum_{i=1}^k t_i \leq 1$, we get

$$\frac{\log k}{k-1}\int_{\mathcal{R}_k} F(t_1,\dots,t_k)^2\sum_{j=1}^{k}(1-t_1-\ldots-t_k+kt_j)dt_k\ldots dt_1 \quad\leq\quad \frac{k\log k}{k-1}\int_{\mathcal{R}_k} F(t_1,\dots,t_k)^2 dt_1\ldots dt_k.$$

Hence

$$M_k \leq \frac{k\log k}{k-1}.$$

□

This upper bound turns to be very meaningful. Maynard obtained that

$$M_{105} \geq 4.0020679 \text{ and } M_5 \geq 2.00116.$$

With the upper bound in 3.4.6, we find

$$M_{105} \leq 4.698709 \text{ and } M_5 \geq 2.011797.$$

That means that Maynard's bounds were optimal.

Using the upper bound in 3.4.6, we get $M_{50} \leq 3.99$, we have seen that for an unconditional gap we need M_k to be greater than 4, that means that we still far from proving the twin prime conjecture, even assuming the Elliot-Halberstam conjecture, we need $M_k > 2$, and we have $M_2 = 1.3862 < 2$ (remark that the only admissible 2-tuple is $\{0, 2\}$). Using Maynard's method its even impossible to prove unconditionally that there exist infinitely many gaps of length 240, as we have $M_49 \leq 3.97$, and for $k = 49$, the shortest admissible $k-$tuple is of diameter 240. To hope to improve Maynard's unconditional gap using a variant of his method, one should start from $k = 51$ as $M_{50} \leq 3097$ and $M_{51} \leq 4.01$. We have also $M_{52} \leq 4.02$ and $M_{53} \leq 4.0466$. Indeed, Polymath8b proved that $M_{53} > 4$.

3.5 Conclusion

To summarize, we have seen that Maynard's improvment of the GPY mathod was to consider weights in a higher dimension, keeping the same basic setting on detecting the prime k-tuples, considering a weighted sum say S, and proving that S is strictly positive. The first step in both methods was to express S as a difference of two quadratic forms S_1 and S_2. The main difficulty here lies in expressing S_1 and S_2 in a simple from, in other words diagonalizing the quadratic forms S_1 and S_2, and this could be achieved making the accurate change of basis (variable).

Estimating sums of multiplicative functions is another difficulty that arises during the diagonalization process, In this context, Goldston, Pintz, and Yildirim's favorite method was transforming this sums to integrals using Perron's formula, then treating this integrals using the theory of Riemann zeta function and the Dirichlet series associated to a multiplicative function. Maynard introduced a relatively simple combinatorial argument, where he generalized a sort of Möbius inversion formula in a higher dimension to estimate sums of multi-dimensional multiplicative functions. Historically speaking that was few months after Y.Zhang approach [21] and the complicated machinery he used to get a bounded gap, so coming up with an elegant combinatorial sieve setting that improves considerably Zhang's bound was surprising.

We should also mention that an analytic approach to estimate the multiplicative functions in Maynard's work is available. Indeed, Tao introduced a multi-dimensional analytic method in [6] based on Fourier Analysis, with this he also obtained (unpublished work) the theorem 3.1.1 with a weaker bound.

Few weeks after publishing Maynard's paper, many results were derived from their work. To name just few, for example, Pintz ([14]) proved that there exists a positive integer B such that there are infinitely many arithmetic progressions of primes $p_n, \ldots, p_{n+k}$ such that $p_n + B, \ldots, p_{n+k} + B$ are all primes. Thorner [16] Proved that, in a Galois extension K of $\mathbb{Q}$, there exists infinitely many pairs of prime ideals of the ring of integers of K whose norms are distinct primes that differ by a positive constant B. and until writing this lines mathematical journals still receiving applications of Maynard's method. Unfortunately, Maynard and even Polymath8b's methods are inadequate to prove the twin prime conjecture, and that's an open problem that we leave for the reader.

References

[1] A. Granville, Primes in intervals of bounded lenght. Bull. Amer. Math. Soc. 52 (2015), 171-222.

[2] A. Karatsuba, Complex Analysis in Number Theory, CRC Press, 1994.

[3] D. A. Goldston, J. Pintz, and C. Y. Yildirim, "Primes in tuples. I", Ann. of Math. (2) 170(2) (2009),824-856.

[4] D. A. Goldston, J. Pintz, and C. Y. Yildirim, "Primes in tuples. II" Acta Math. 204(1) (2010),6-25.

[5] D. A. Goldston, J. Pintz, and C. Y. Yildirim, "The path to recent progresson small gaps between primes", Clay Math. Proc. 7 (2007),135-138.

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

[7] D. A. Goldston, S. W. Graham, J. Pintz, and C. Y. Yildirim, Small gaps between products of two primes, Proc. Lond. Math. Soc. 98 (2009).

[8] H. Davenport, Multiplicative number theory, 3rd ed Springer-Verlag, New York, (2000).

[9] H. Halberstam and H.- E. Rihert, Sieve Methods, Academic Press, (1974).

[10] H. Iwaniec and E. Kowalski, Analytic number theory, AMS Colloquium Publications, 53 (2004).

[11] H. Montgomery and R. Vaughan, Multiplicative Number Theory I, classical theory, Cambridge University Press (2006).

[12] K. Ford, Simple proof of Gallagher's singulare series sum estimate, http://arxiv.org/abs/1108.3861v1.

[13] J.Maynard, Small gaps between primes. Ann. of Math. 181 (2015),386-412.

[14] J. Pintz, Polignac Numbers, Conjectures of Erdos on Gaps between Primes Arithmetic Progressions in Primes, and the Bounded Gap Conjecture

[15] J. M De Koninck and F. Luca, Analytic Number Theory: Exploring the Anatomy of Integers. (Graduate Studies in Mathematics, Vol 134). AMS (2012).

[16] J. Thorner, âBounded gaps between primes in Chebotarev setsâ,Res. Math. Sci.(2014).

[17] K. Soundararajan, Small gaps between prime numbers: the work of Goldston-Pintz-Yildirim, Bull. Amer. Math. Soc. 44 (2007).

[18] P. X. Gallagher, On the distribution of primes in short intervals, Mathematika 23 (1976) 4-9, Corrigendum, Mathematika 28 (1981) 86.

[19] T. Apostol, Introduction to Analytic Number Theory, Springer-Verlag 1976.

[20] Y. Motohashi, "The twin prime conjecture", http://www.arxiv.org/abs/1401.6614v2, 12-16 .

[21] Y. Zhang, Bounded gaps between primes, Ann. of Math. 1121-1174 Vol 179 (2014).