Twin Prime Numbers

Bounded gaps between primes

Universiteit Utrecht

Author: I.F.M.M. Nelen

Supervisor: Prof. Dr. F. Beukers

28 januari 2015

Table of Contents

Twin Prime Numbers

Abstract

In this thesis an in-depth explanation will be given of the proof Maynard gave in his article. In which the steps of the proofs will be expanded. He used a refinement on the GPY sieve to study k-tuples and small gaps between primes. This will show that $\liminf(p_{n+1} - p_n) \leq 600$, and, by assuming the Elliott-Halberstam conjecture, that $\liminf(p_{n+1} - p_n) \leq 12$ and $\liminf(p_{n+2} - p_n) \leq 600$.

Introduction

Prime Numbers

For thousands of years people have been fascinated by numbers. First the positive integers where used to describe quantities. Then these numbers were extended by the negative integers and fractions. But the fascination with positive integers or natural numbers stayed. The building blocks of natural numbers are prime numbers, every natural number greater than two is a prime number or a unique product of prime numbers. The greeks started with examining prime numbers and found there are infinitely many. The proof of this theorem is based on the divisibility of numbers p and $p + 1$. Because there are infinitely many prime numbers one may wonder about the distribution of those prime numbers. For $x > 0$ let $\pi(x)$ denote the number of prime numbers not exceeding x. Because there are infinitely many prime numbers $\pi(x) \to \infty$ as $x \to \infty$. Gauss (1792) and Legendre(1798) proposed the distribution of prime numbers around a number x is asymptotic to $x/\log x$. This means heuristically there is a 1 in $\log x$ chance of a number close to x to be a prime number[9]. It is known as the prime number theorem.

Twin Prime Numbers

When one looks at the first prime numbers one can see some patterns emerge. For example $17, 19$ or $41, 43$ and $107, 109$. These are all prime numbers which differ by exactly 2. One may wonder if there are infintely many of such pairs. This can be phrased as there are infinitely many pairs of natural numbers $(p, p+2)$ with p and $p+2$ both prime. This conjecture is known as the twin prime conjecture. One may wonder if there are more patterns in prime numbers instead of 2 one may take a difference of 4 or 6 between pairs of primes. These prime pairs are called respectively cousin and sexy primes.

Heuristic Approach To Twin Prime Numbers

To estimate the number of twin primes up to a natural number x one can use the distribution of the prime numbers and the prime number theory. This states that a number smaller than x has at least probability $1/\log x$ of being a prime number. This means if we pick two numbers smaller than x the probability of both of them being a prime numbers is at least $1/(\log x)^2$ but only when the event of "p is prime" is independent of the event "p+2 is prime". This isn't true if $p \equiv 1 \ mod \ 3$ and prime then $p+2 \equiv 0 \ mod \ 3$ thus $p+2$ isn't prime.
One needs to correct for this dependence by some correction factor. The probability for an arbitrary number to be divisible by a number q is $1/q$. So the probability for two arbitrary numbers not to be divisible by a number q is $(1-1/q)^2$. For two numbers p and $p+2$ this is different because we need $p \not\equiv 0 \ mod \ q$ and $p \not\equiv -2 \ mod \ q$ which is in $2/q$ of the cases. The ratio between these factors is the correction factor. Thus the correction factor for any number $q > 2$ becomes:

$$\frac{(1-\frac{2}{q})}{(1-\frac{1}{q})^2}$$

For $q = 2$ one has 1 modulo class which is restricted for p. This correction factor becomes:

$$\frac{(1-\frac{1}{2})}{(1-\frac{1}{2})^2} = 2$$

Being divisible by a small prime number is independent of the other small primes. Thus one may multiply the correction factors of the small prime numbers. In fact one may multiply over all prime numbers because when q is large the correction factor converges to 1. This suggests a definition of a twin prime constant of:

$$C = 2 \prod_{\substack{q\ prime \\ q \geq 3}} \frac{(1-\frac{2}{q})}{(1-\frac{1}{q})^2} \approx 1,3203236316$$

This is the total correction factor over all primes q. One may guess the estimate of the number of twin prime pairs smaller than an integer x is:

$$C \frac{x}{(\log x)^2}$$

This would mean by a heuristic approach there would be infinitely many pairs of primes which differ 2. Because the formula for the estimate of the number of pairs under x goes to infinity when x goes to infinity.

Bounding The Gaps

There are two main ways to attempt to prove the twin-prime conjecture. One can try to find the difference between the n^{th} prime p_n and the next prime p_{n+1}. And prove for infinitely many n the difference between them is 2. Or one can use 'GPY method' which takes sets of numbers of the same length and proves at least two of them have to be prime.

Admissible Tuples and Sets

We are interested in sets which when you add n all of them could, in theory, be prime. If you take the tuple $0, 2, 4$ then it is known when you add a arbitrary integer n. One of the $n, n+2, n+4$ is always divisible by 3. This is a restriction on the tuple. Thus when one needs to find prime numbers this tuple is not used.

Definition 1. *A k-tuple $\mathcal{H} = h_1, ..., h_k$ is called an admissible set when there is no integer q such that $q | \prod_{i=1}^{k}(n + h_i)$ for all $n \in \mathbb{Z}$*

The tuples used here are all tuples of natural numbers but this is not needed. One may take $k-tuples$ of lineair forms such as $m, m+n, m+4n$ for two natural numbers m, n. The next conjecture isn't proven for any $k > 1$. It is proven for any admissible set of linear forms provided no two satisfy a linear equation over the integers. Unfortunatly most of the questions mathematicians are interested in do not satisfy these conditions. Such as the twin prime conjecture which satisfies the linear equation $q - p = 2$.

Conjecture 1. *(Prime k-tuples conjecture).*
Let $\mathcal{H} = \{h_1, ..., h_k\}$ be admissible. Then there are infinitely many integers n such that all of $n + h_1, ..., n + h_k$ are prime.

If one wants to prove the twin prime conjecture the prime k-tuples conjecture has to be proven for $k = 2$ and $h_2 = h_1 + 2$. When a bound needs to be proven a less strict conjecture is needed.

Conjecture 2. *Let $\mathcal{H} = \{h_1, ..., h_k\}$ be admissilbe. Then there are infinitely many integers n such that at least two of $n + h_1, ..., n + h_k$ are prime.*

When this conjecture was proven for some $\mathcal{H}$ we have $\liminf_n p_{n+1} - p_n \leq \max_{i,j} |h_i - h_j|$. The breaktrough of finding a bounded gaps between primes is made by Zhang. Who in his paper proved $\liminf_n p_{n+1} - p_n = B$ for a $B < \infty$ and even gave an upper bound of $B \leq 70000000$. As well in his paper he stated "This result is not optimal ... to replace the [upper bound of B] by a value as small as possible is an open problem that will not be discussed in this paper"[10]. The mathematical community worked together in the Polymath8 project to show this bound could indeed be improved. This project features different mathematicians all over the world who with modern technology try to make a breaktrough in a certain mathematical problem. In Polymath8 new ways to calculate the bounds were found. The results followed each other in a span of weeks.[1] In less than a year the bound has been lowered to less than 600. This step was a major breaktrough in the project because a different method was used to find this bound. All other improvements where made by small improvements in the proof by Zhang. This improvement uses a different method. In my thesis I will give the proof of Maynard [6] in a extended and simplified way. In the next chapter an outline of the proof will be given.

[1] For the complete timeline look at http://michaelnielsen.org/polymath1/index.php?title=Timeline_of_prime_gap_bounds

Outline of Maynard's Proof

The proof of Maynard is based on the 'GPY method' named after Goldston, Pintz and Yildrim[2]. This method uses the distribution of primes to study prime tuples and small gaps between primes. Given $\theta > 0$ we say the primes have "level of distribution θ' if, for any $A > 0$

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

where $\pi(x; q, a)$ is the number of primes up to x of the form $a \mod q$. We have the following results.

Theorem 1. *(Bombieri-Vinogradov Theorem)*
The primes have a level of distribution θ for any $\theta < 1/2$.

Conjecture 3. *(Elliott-Halberstam Conjecture)*
The primes have a level of distribution θ for any $\theta < 1$.

Now the basic idea behind the approach of Maynard is when $\mathcal{H} = \{h_1, ..., h_k\}$ is a fixed admissible set one considers the sum

$$S(N, \rho) = \sum_{N \leq n \leq 2N} (\sum_{i=1}^{k} \chi_{\mathbb{P}}(n + h_i) - \rho) w_n \tag{2}$$

where $\chi_{\mathbb{P}}$ is the characteristic function of the primes. Thus $\chi_{\mathbb{P}}(p) = 1$ when p is prime and 0 otherwise. $\rho > 0$ and w_n are non negative weights. If one can show $S(N, \rho) > 0$. Then at least one term in (2) has a positive contribution. This means there exists some integer $n \in [N, 2N]$ such that at least $\lfloor \rho + 1 \rfloor$ of the $n + h_i$ are prime. Thus if $S(N, \rho) > 0$ for all large N, then there are infinitely many integers n for which at least $\lfloor \rho + 1 \rfloor$ of the $n + h_i$ are prime. Thus there are infinitely many bounded intervals containing $\lfloor \rho + 1 \rfloor$ primes. The hard part in this is choosing the weights such that this happens. The difference between these sieves and the GPY sieve is in the λ's. In the standard Sielberg sieve the following weights are chosen.

$$w_n = (\sum_{\substack{d|\prod_{i=1}^k (n+h_i)\\ d<R}} \lambda_d)^2, \quad \lambda_d = \mu(d)(\log R/D)^k \tag{3}$$

But Goldston, Pintz and Yildrim used different weights of the form

$$\lambda_d = \mu(d)F(\log R/D) \tag{4}$$

for a suitable smooth function F and μ the Möbius function. They chose $F(x) = x^{k+l}$ for suitable $l \in \mathbb{N}$, which has been shown to be essentially optimal when k is large. This will prove the existence of bounded gaps between primes when the level of distribution of primes $\theta > 1/2$ but not when $\theta < 1/2$. Thus the existence of bounded gaps between primes then relies on the the Elliott-Halberstam Conjecture (conjecture 3) which has not been proven. Thus a weaker condition is needed to prove the existence of bounded gaps. Zhang used in his paper a modified form of the distribution of primes. Maynard uses in his paper a different weight.
Maynard's weights aren't only based on a $d|\prod_{i+1}^k (n+h_i)$ but uses different $d_i|n+h_i\forall i$. Which gives sieve weights of the form

$$w_n = (\sum_{d_i|n+h_i\forall i} \lambda_{d_1,\dots,d_k})^2 \tag{5}$$

In Maynard's proof he choses his weights to be zero unless n lies in a fixed residue class $v_0 (\mod W)$, where $W = \prod_{p\leq D_0} p$ the product of all primes numbers smaller than D_0. Thus will remove some minor complications when dealing with small prime factors. He choses D_0 to be $D_0 = \log\log\log N$ to be sure that $W \ll (\log\log N)^2$ by the prime number theorem.
To estimate the sum (2) it is split in two parts.

$$S_1 = \sum_{\substack{N\leq n\leq 2N\\ n\equiv v_0 (\mod W)}} \left(\sum_{d_i|n+h_i\forall i} \lambda_{d_1,\dots,d_k} \right)^2 \tag{6}$$

$$S_2 = \sum_{\substack{N\leq n\leq 2N\\ n\equiv v_0 (\mod W)}} \left(\sum_{i=1}^k \chi_{\mathbb{P}}(n+h_i) \right) \left(\sum_{d_i|n+h_i\forall i} \lambda_{d_1,\dots,d_k} \right)^2 \tag{7}$$

So $S = S_2 - \rho S_1$ with a few minor resrictions added. In the same way as Zhang's proof the first part is rewriting both sums untill they are more

or less smooth functions. In the sum itself you have some functions which can't be specifically calculated. By rewriting both sums, and changing the variables the sum will be easier to handle. This will be the first part of his proof.
The second part will be proving the desired result will be achieved when the sum is calculated. The next part is finding the length of the tuple in which two prime numbers will be found. To finish the proof a tuple is given and the bound will be set at 600.

Estimating The Sums

We will use a proposition to evaluate the sums.

Proposition 1. *Let the primes have exponent of distribution $\theta > 0$, and let $R = N^{\theta/2-\delta}$ for some small fixed $\delta > 0$. Let $\lambda_{d_1,...,d_k}$ be defined in terms of a fixed piecewise differentiable function F by*

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

whenever $(\prod_{i=1}^{k} d_i, W) = 1$ and let $\lambda_{d_1,...,d_k} = 0$ otherwise. Let F be supported on $\mathcal{R} = \{(x_1, ..., x_k) \in [0,1]^k : \sum_{i=1}^{k} x_i \leq 1\}$. This means if $\lambda_{d_1,...,d_k} \neq 0$ then $d_1...d_k \leq R$ and $(d_i, d_j) = 1$ when $i \neq j$. Then we have

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

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

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

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

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

Technical Lemmas

The following Lemmas are used troughout the proofs

Lemma 1.

$$\sum_{d<R} \tau_k(d) < R(\log R)^{k-1} \tag{8}$$

Proof.

$$\sum_{d<R} \tau_k(d) = \sum_{\substack{d_1,\ldots,d_k \\ d_1\ldots d_k<R}} 1 = R \sum_{\substack{d_1,\ldots,d_k \\ \prod_i d_i<R}} \frac{1}{R}$$

$$< R \sum_{\substack{d_1,\ldots,d_k \\ d_i<R}} \frac{1}{d_i} = R\left(\sum_{d<R} \frac{1}{d}\right)^k < R(\log R)^{k-1} \tag{9}$$

□

Lemma 2. *(Generalized Möbius inversion)*
If $A_{d_1,\ldots,d_k}$ *with support in* $d_i \in \mathbb{N}$, $\prod_i d_i < R$.
Define

$$B_{r_1,\ldots,r_k} = \sum_{r_i|d_i} A_{d_1,\ldots,d_k} \tag{10}$$

then

$$A_{d_1,\ldots,d_k} = \sum_{r_i|d_i} \prod_{i=1}^{k} \mu(r_i) B_{r_1,\ldots,r_k} \tag{11}$$

Lemma 3.

$$\sum_{a=1}^{R} \frac{\mu(a)^2}{\varphi(a)} \ll \log R \tag{12}$$

Proof.

$$\log \prod_{\substack{p\le R \\ p\ prime}} (1+\frac{1}{p-1}) = \sum_{p\le R} \log(1+\frac{1}{p-1}) \le \sum_{p\le R} \frac{1}{p-1} < c+\log\log R \tag{13}$$

$$\sum_{a=1}^{R} \frac{\mu(a)^2}{\varphi(a)} < \prod_{\substack{p\le R \\ p\ prime}} (1+\frac{1}{p-1}) \ll \log R \tag{14}$$

Lemma 4.

$$\sum_{\substack{(u,W)=1\\u>D_0}} \frac{\mu(u)^2}{\varphi(u)^2} \ll \frac{1}{D_0} \tag{15}$$

Proof.

$$\sum_{\substack{(u,W)=1\\u>D_0}} \frac{\mu(u)^2}{\varphi(u)^2} = \prod_{p>D_0}\left(1+\frac{1}{(p-1)^2}\right) \tag{16}$$

If one looks at the logarithm of the right hand side we get

$$\sum_{p>D_0} \log\left(1+\frac{1}{(p-1)^2}\right) < \sum_{p>D_0} \frac{2}{(p-1)^2} < \sum_{n>D_0} \frac{2}{(n-1)^2} < \frac{2}{D_0} \tag{17}$$

□

Changing the Variables

To prove this proposition a change of variables will be introduced. This approach is based on the elementary combinatorial ideas of Selberg. [7]
We assume that the primes have a fixed level of distribution θ, and $R = N^{\theta/2-\epsilon}$. The weight $\lambda_{d_1,\ldots,d_k}$ is restricted to tuples with $d = \prod_{i=1}^k d_i$ and $d < R$, $(d, W) = 1$ and $\mu(d)^2 = 1$. This implies all the d_i are pairwise coprime and square-free.

Proposition 2. *Let*

$$y_{r_1,\ldots,r_k} = \prod_{i=1}^{k} \mu(r_i)\varphi(r_i) \sum_{\substack{d_1,\ldots,d_k\\r_i|d_i\forall i}} \frac{\lambda_{d_1,\ldots,d_k}}{\prod_{i=1}^k d_i} \tag{18}$$

Let $y_{max} = \sup_{r_1,\ldots,r_k} |y_{r_1,\ldots,r_k}|$ *then*

$$S_1 = \frac{N}{W} \sum_{r_1,\ldots r_k} \frac{y^2_{r_1,\ldots,r_k}}{\prod_{i=1}^k \varphi(r_i)} + O(\frac{y^2_{max}\varphi(W)^k N(LogR)^k}{W^k D_0}) \tag{19}$$

Proof. Expand out the square in S_1 and then swap the order of summation

$$S_1 = \sum_{\substack{N\leqslant n<2N \\ n\equiv v_0 mod W}} \left(\sum_{d_i|n+h_i \forall i} \lambda_{d_1,\ldots,d_k} \right)^2 = \sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k}} \lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k} \sum_{\substack{N\leqslant n<2N \\ n\equiv v_0 mod W \\ [d_i,e_i]|n+h_i\forall i}} 1 \tag{20}$$

The inner sum runs over all n such that $N \leq n < 2N$ with $n \equiv v_0 \mod W$ and $n+h_i \equiv 0 \mod [d_i < e_i]$ for all i. The $W, [d_i, e_i]$ are pairwise coprime. If not it would mean $([d_i, e_i], [d_j, e_j]) = d > 1$ thus $d|d_i, d_j$ but the product over the d_i is square free. This is a contradiction so they are pairwise coprime. Because they are coprime there exists, by the chinese remainder theorem, a residue class $a \mod q$ with $q = W\prod_{i=1}^{k}[d_i, e_i]$ such that the summation runs over $n \equiv a \mod q$ and $N \leq n < 2N$
There is N/q times a n modulo this residue class in an interval with lenght N. This calculation has an error which approaches 1 when $N \to \infty$. The inner sum becomes $N/q + O(1)$ when the integers $W, [d_i, e_i]$ are pairwise coprime, when they are not the sum is zero. This restriction on the integers $W, [d_i, e_i]$ will be denoted by $\sum'$. This gives

$$S_1 = \frac{N}{W} \sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k}}{}' \frac{\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}}{\prod_{i=1}^{k}[d_i, e_i]} + O(\sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k}}{}' |\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}|) \tag{21}$$

To ease notation we put $y_{max} = \sup_{r_1,\ldots,r_k} |y_{r_1,\ldots,r_k}|$. The assumptions on λ state it is only non zero when $\prod_{i=1}^{k} d_i < R$, thus the errorterm contributes

$$\sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k}} |\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}| \leqslant \sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k}} (\lambda_{max})^2 \ll \lambda_{max}^2 (\sum_{d<R} \tau_k(d))^2 \tag{22}$$

Where $\tau_k(d)$ means in how many ways d can be written as a product of k integers.. This is bigger than the number of ways d can be written with k divisors which are square free. This can be estimated in order of magnitude by:

$$\lambda_{max}^2 (\sum_{d<R} \tau_k(d))^2 \ll \lambda_{max}^2 R^2 (Log R)^{2k} \tag{23}$$

In this estimation Lemma 1 is used to estimate the sum. In the main sum we remove the dependencies between the d_i and the e_j variables. We use the identity:

$$\frac{1}{[d_i,e_i]} = \frac{1}{d_i e_i}(e_i,d_i) = \frac{1}{d_i e_i}\sum_{u_i|d_i,e_i}\varphi(u_i)$$

This means the least common multiple can be rewritten to the product divided by the greatest common divisor. This is a combination of all u_i dividing the greatest common divisor and the equality $\sum_{n|d}\varphi(n) = d$. Thus the sum is the greatest common divisor.

The main term becomes

$$S_1 = \frac{N}{W}\sum_{u_1,\ldots,u_k}\left(\prod_{i=1}^{k}\varphi(u_i)\right)\sideset{}{'}\sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k \\ u_i|d_i,e_i\forall i}}\frac{\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}}{\prod_{i=1}^{k}d_i\prod_{i=1}^{k}e_i} \tag{24}$$

Now recall the requirement on the summation which is that $W,[d_1,e_1],\ldots,[d_k,e_k]$ are alle pairwise coprime. The weight λ is only supported on integers $d_1,\ldots,d_k$ which are coprime with W. Thus the requirement of W being coprime with the least common multiples can be dropped. Similarly the requirements of $(d_i,d_j) = 1$ for all $i \neq j$ and $(e_i,e_j) = 1$ for all $i \neq j$ may be dropped. The only restriction left from the pairwise coprimality of $W,[d_1,e_1],\ldots,[d_k,e_k]$ is $(d_i,e_j) = 1$ for all $i \neq j$.

To be certain the requirement will be held without a requirement in the sum, multiply the main term with $\sum_{s_{i,j}|d_i,e_j}\mu(s_{i,j})$. This works by the equality $\sum_{d|n}\mu(d)$ is 1 when $n = 1$ and 0 otherwise. So the sum for a pair (i,j) will not be empty only when $(d_i,e_j) = 1$. This applies to all i,j with $i \neq j$. The main term becomes

$$S_1 = \frac{N}{W}\sum_{u_1,\ldots,u_k}\left(\prod_{i=1}^{k}\varphi(u_i)\right)\sum_{s_{1,2},\ldots,s_{k,k-1}}\left(\prod_{\substack{1\le i,j\le k \\ i\neq j}}\mu(s_{i,j})\right)\sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k \\ u_i|d_i,e_i\forall i \\ s_{i,j}|d_i,e_j\forall i\neq j}}\frac{\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}}{\prod_{i=1}^{k}d_i\prod_{i=1}^{k}e_i} \tag{25}$$

Assume $(u_i,s_{i,j}) \neq 1$ this means $s_{i,j} > 1$ and $s_{i,j}|d_i,e_i,e_j$. Thus $(e_i,e_j) \geq s_{i,j} > 1$ and $\lambda_{e_1,\ldots,e_k} = 0$ unless $(e_i,e_j) = 1$. So there may be a restriction to $s_{i,j}$ by only summing over the ones which are coprime to u_i and u_j. Because when this is not true $\lambda_{e_1,...,e_k}$ does not support the give tuple. In a similar way the restriction may be expanded by demanding $s_{i,j}$ to be coprime with $s_{i,a}$ and $s_{b,j}$ for all $a \neq j$ and $b \neq i$. The summation over these will be further noted by $\sum^*$. To make our sum S_1 more straightforward we will introduce a change of variables. Let

$$y_{r_1,...,r_k} = (\prod_{i=1}^{k} \mu(r_i)\varphi(r_i)) \sum_{\substack{d_1,...,d_k \\ r_i|d_i \forall i}} \frac{\lambda_{d_1,...,d_k}}{\prod_{i=1}^{k} d_i} \tag{26}$$

This change of variables has to be invertible. When it is not invertible the sum will not run over all possible tuples, for $d_1, ..., d_k$ square free. We will prove our change of variables is invertible. This follows by using Lemma 2

$$\sum_{\substack{r_1,...,r_k \\ d_i|r_i \forall i}} \frac{y_{r_1,...,r_k}}{\prod_{i=1}^{k} \varphi(r_i)} = \frac{\lambda_{d_1,...,d_k}}{\prod_{i=1}^{k} \mu(d_i) d_i} \tag{27}$$

This means any choice of $y_{r_1,...,r_k}$ supported on $r_1, ..., r_k$ when the product $\prod_{i=1}^{k} r_i = r$ is square-free, $r < R$ and $(r, W) = 1$ will give a suitable choise of $\lambda_{d_{1]},...,d_k}$. By changing our variable we need to re-estimate the λ_{max}. Let $y_{max} = sup_{r_1,...,r_k}|y_{r_1,...,r_k}|$. Since $d/\varphi(d) = \sum_{e|d} 1/\varphi(d)$ for d square-free. Take $r' = \prod_{i=1}^{k} r_i/d_i$ then by the change of variables and the definition of λ_{max}

$$\lambda_{max} \leq \sup_{\substack{d_1,...,d_k \\ \prod_{i=1}^{k} d_i \text{ square-free}}} y_{max}(\prod_{i=1}^{k} d_i) \sum_{\substack{r_1,...,r_k \\ d_i r_i \forall i \\ \prod_{i=1}^{k} r_i < R \\ \prod_{i=1}^{k} r_i \text{ square-free}}} (\prod_{i=1}^{k} \frac{\mu(r_1)^2}{\varphi(r_i)}) \tag{28}$$

by changing the variable from d_i to r' the limits of the sum change $\prod_{i=1}^{k} r_i < R \to r' < R/\prod_{i=1}^{k} d_i$ and because $d_i|r_i$ the other limit changes $\prod_{i=1}^{k} r_i$ square-free $\to (r', \prod_{i=1}^{k} d_i) = 1$. When r' is used instead of a sum over $r_1, .., r_k$ there need to be corrected by the number of ways the $r_1, ..., r_i$ can form a certain product. This correction is smaller than the number of ways the product can be written by all it's divisors. So we can estimate this term true multiplying by $\tau_k(r')$.

$$\lambda_{max} \leq y_{max} \sup_{\substack{d_1,\dots,d_k \\ \prod_{i=1}^k d_1 \text{ square-free}}} (\prod_{i=1}^{k} \frac{d_i}{\varphi(d_i)}) \sum_{\substack{r'<R/\prod_{i=1}^k d_i \\ (r',\prod_{i=1}^k d_i)=1}} \frac{\mu(r')^2\tau_k(r')}{\varphi(r')} \tag{29}$$

$$\leq y_{max} \sup_{d_1,\dots,d_k} (\sum_{d|\prod_{i=1}^k d_i} \frac{\mu(d)}{\varphi(d)}) \sum_{\substack{r'<R/\prod_{i=1}^k d_i \\ (r',\prod_{i=1}^k d_i)=1}} \frac{\mu(r')^2\tau_k(r')}{\varphi(r')} \tag{30}$$

The product we can rewrite by using the equality stated and rewriting $1 = \mu(d)^2$ to remove the requirement of needing $\prod_{i=1}^k d_i$ to be square-free. If we let $u = dr'$ and by using the fact that $\tau_k(dr') \geq \tau_k(r')$ both sums can be combined an the estimate becomes

$$\leq y_{max} \sum_{u<R} \frac{\mu(u)^2\tau_k(u)}{\varphi(u)} \ll y_{max}(logR)^k \tag{31}$$

This is by combining (Lemma 1) and Lemma 3.

By substituting the change of variables and with the estimation of the error term, we obtain

$$S_1 = \frac{N}{W} \sum_{u_1,\dots,u_k} (\prod_{i=1}^{k} \varphi(u_i)) \sum\nolimits^*_{s_{1,2},\dots,s_{k,k-1}} (\prod_{\substack{1\leq i,j\leq k \\ i\neq j}} \mu(s_{i,j}))(\prod_{i=1}^{k} \frac{\mu(a_i\mu(b_i))}{\varphi(a_i)\varphi(b_i)})y_{a_1,\dots,a_k}y_{b_1,\dots,b_k}$$
$$+O(y_{max}^2 R^2 (LogR)^{4k}) \tag{32}$$

where $a_i = u_i \prod_{i\neq j} s_{i,j}$ and $b_j = u_j \prod_{i\neq j} s_{i,j}$ The functions φ and μ are multiplicative and all factors are pairwise coprime thus $\mu(a_i) = \mu(u_i)\prod_{i\neq j}\mu(s_{i,j})$ the same for $\mu(b_i), \varphi(a_i)$ and $\varphi(b_i)$ but only when all $s_{i,j}$ are coprime with all other terms in the a_i and b_i. All terms with these not square free do not contribute to the sum and $\mu(s_{i,j})^3 = \mu(s_{i,j})$ hence the sum can be rewritten to as

$$S_1 = \frac{N}{W} \sum_{u_1,\dots,u_k} (\prod_{i=1}^{k} \frac{\mu(u_i)^2}{\varphi(u_i)}) \sum\nolimits^*_{s_{1,2},\dots,s_{k,k-1}} \prod_{\substack{1\leq i,j\leq k \\ i\neq j}} \frac{\mu(s_{i,j})^2}{\varphi(s_{i,j})^2} y_{a_1,\dots,a_k}y_{b_1,\dots,b_k}$$
$$+O(y_{max}^2 R^2 (LogR)^{4k}) \tag{33}$$

Again there is no contribution from $s_{i,j}$ with $(s_{i,j}, W) \neq 1$ because of the restricted support of y. Thus only $s_{i,j} = 1$ and $s_{i,j} > D_0$ need to be considered. When $s_{i,j} > D_0$ the contribution to the sum is

$$\ll \frac{y_{max}^2 N}{W} \left(\sum_{\substack{u<R \\ (u,W)=1}} \frac{\mu(u)^2}{\varphi(u)} \right)^k \left(\sum_{s_{i,j} > D_0} \frac{\mu(s_{i,j})^2}{\varphi(s_{i,j})^2} \right) \left(\sum_{s \geq 1} \frac{\mu(s)^2}{\varphi(s)^2} \right)^{k^2-k-1} \tag{34}$$

The first sum can be rewritten in much the same way as in Lemma 3 to $\varphi(W)^k (LogR)^k / W^k$. Because one needs to correct for the terms u which are not coprime with W this correction term is $\varphi(W)/W$. Thus the ratio between W and the numbers which are coprime to it. The last sum is convergent thus can be rewritten as a contstant. For the middle sum we use Lemma 4.

If we combine the Lemma with the other estimations then the equation of (34) is estimated by

$$\ll \frac{y_{max}^2 \varphi(W)^k N (LogR)^k}{W^{k+1} D_0} \tag{35}$$

When $s_{i,j} = 1$ and all other terms are in our error term we get

$$S_1 = \frac{N}{W} \sum_{u_1,\ldots,u_k} \frac{y_{u_1,\ldots,u_k}^2}{\prod_{i=1}^k \varphi(u_i)} + O\left(\frac{y_{max}^2 \varphi(W)^k N (LogR)^k}{W^{k+1} D_0} + y_{max}^2 R^2 (LogR)^{4k}\right) \tag{36}$$

To prove the lemma it now suffices to show the first error term dominates the second. Recall $R^2 = N^{\theta - 2\delta} \leq N^{1-2\delta}$, $W \ll N^\delta$ and $R \gg (LogR)^{3k}$. Thus $N/WD_0 \gg R^2 (LogR)^{3k}$ hence $N/WD_0 \gg N^{1-2\delta} \geq R^2$ and the first term dominates. This ends the proof and

$$S_1 = \frac{N}{W} \sum_{u_1,\ldots u_k} \frac{y_{u_1,\ldots,u_k}^2}{\prod_{i=1}^k \varphi(u_i)} + O\left(\frac{y_{max}^2 \varphi(W)^k N (LogR)^k}{W^k D_0}\right) \tag{37}$$

□

Rewriting the second sum

In a similar way we estimate $S_2 = \sum_{m=1}^{k} S_2^{(m)}$ where

$$S_2^{(m)} = \sum_{\substack{N \le n < 2N \\ n \equiv v_0 (modW)}} \chi_{\mathbb{P}}(n+h_m) (\sum_{\substack{d1,...,d_k \\ d_i | n+h_i \forall i}} \lambda_{d_1,...,d_k})^2 \tag{38}$$

In the next lemma we estimate $S_2^{(m)}$ in a similar way to S_1

Proposition 3. *Let*

$$y^{(m)}_{r_1,...,r_k} = \prod_{i=1}^{k} \mu(r_i) g(r_i) \sum_{\substack{d_1,..,d_k \\ r_i | d_i \forall i \\ d_m = 1}} \frac{\lambda_{d_1,...,d_k}}{\prod_{i=1}^{k} \varphi(d_i)} \tag{39}$$

where g is the totally multiplicative function defined on primes by $g(p) = p-2$. Let $y^{(m)}_{max} = \sup_{r_1,...,r_k} |y^{(m)}_{r_1,...,r_k}|$ then

$$S_2^{(m)} = \frac{N}{\varphi(W) LogN} \sum_{r_1,...r_k} \frac{(y^{(m)}_{r_1,...,r_k})^2}{\sum_{i=1}^{k} g(r_i)} + O(\frac{(y^{(m)}_{max})^2 \varphi(W) N (LogR)^{k-2}}{W^{k-1} D_0}) + O(\frac{y^2_{max} N}{(LogN)^A}) \tag{40}$$

*Proof.*By interchanging the order of summation and expanding out the squares we get a resemblence with S_1 in equation (20).

$$S_2^{(m)} = \sum_{\substack{d_1,...,d_k \\ e_1,...,e_k}} \lambda_{d_1,...d_k} \lambda_{e_1,...e_k} \sum_{\substack{N \le n < 2N \\ n \equiv v_0 \mod W \\ [e_i, d_i] | n+h_i}} \chi_p(n+h_m) \tag{41}$$

The inner sum is over $k+1$ residue classes which can be written as a sum over a single residue. This residue class is $q = W \prod_{i=1}^{k} [d_i, e_i]$ with $W, [d_1, e_1], ..., [d_k, e_k]$ are pairwise coprime. Because of the χ-prime function the inner sum for S_2^m is only non zero when $d_m = e_m = 1$. By the same reason the inner sum will contribute $X_N / \varphi(q)$ and an error term. Where

$$X_N = \sum_{N \le n < 2N} \chi_{\mathbb{P}}(n) \tag{42}$$

It means we have at least the sum of prime numbers between N and $2N$ divide by the quantity of numbers which are coprime by q. Let

$$E(N,q) = \sup_{(a,q)=1} |\sum_{\substack{N\leq n<2N \\ n\equiv a(mod q)}} \chi_{\mathbb{P}}(n) - \frac{1}{\varphi(q)} \sum_{N\leq n<2N} \chi_{\mathbb{P}}(n)| \tag{43}$$

Thus the size of the contribution to the error term becomes

$$O(E(N,q)) \tag{44}$$

It's easy to see by summing the two we get the desired result. Thus by the same notation of the restriction the sum becomes

$$S_2^{(m)} = \frac{X_N}{\varphi(W)} \sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k \\ e_m=d_m=1}}' \frac{\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}}{\prod_{i=1}^k \varphi([d_i,e_i])} + O(\sum_{\substack{d_1,\ldots,d_k \\ e_1,\ldots,e_k}} |\lambda_{d_1,\ldots,d_k}\lambda_{e_1,\ldots,e_k}| E(N,q)) \tag{45}$$

In a similar way to the first sum we estimate the error term. The support of $\lambda_{d_1,\ldots,d_k}$ restricts q to be only square free and smaller than R^2W. Given a square free integer r there are only $\tau_{3k}(r)$ possibilities of $d_1,\ldots,d_k,e_1,\ldots,e_k$ for which $W\prod_{i=1}^k [d_i,e_i] = r$. Recall $\lambda_{max} \ll y_{max}(LogR)^k$. Hence the error term becomes

$$\ll y_{max}^2 (LogR)^{2k} \sum_{r<R^2W} \mu(r)^2 E(N,r)\tau_{3k}(r) \tag{46}$$

Rewrite the sum

$$= y_{max}^2 (LogR)^{2k} \sum_{r<R^2W} \left(\tau_{3k}(r)\mu(r)E(N,r)^{1/2}\right)\left(\mu(r)E(N,r)^{1/2}\right) \tag{47}$$

Then you will get by using Cauchy Schwartz on both parts between brackets and $E(N,r) \ll N/\varphi(r)$

$$\ll y_{max}^2 (LogR)^{2k} \left(\sum_{r<R^2W} \mu(r)^2 \tau_{3k}^2(r)\frac{N}{\varphi(r)}\right)^{1/2} \left(\sum_{r<R^2W} \mu(r)^2 E(N,r)\right)^{1/2} \tag{48}$$

Remember the primes have level of distribution θ thus the last sum $\ll (N/LogN^{A'})^{1/2}$ for any $A' > 0$. In the first sum they are smaller than $N^{1/2}(LogN)^B$ for a B. Hence this can be rewritten to

$$\ll \frac{y_{max}^2 N}{(LogN)^A} \tag{49}$$

As in the treatment of the first sum the conditions of all $(d_i, e_j) = 1$ can be rewritten by multiplying the expression by $\sum_{s_{i,j}|d_i,e_j} \mu(s_{i,j})$. Again the same requirements may restrict $s_{i,j}$ and will be denoted by $\sum^*$. In the same way the $\varphi([d_i, e_i])$ term can be split, because they are square-free, by the equation

$$\frac{1}{\varphi([d_i, e_i])} = \frac{1}{\varphi(d_i)\varphi(e_i)} \sum_{u_i|d_i,e_i} g(u_i)$$

Where g is the totally multiplicative function defined on the primes by $g(p) = p - 2$. This transforms the main term to

$$\frac{X_N}{\varphi(W)} \sum_{u_1,\dots,u_k} (\prod_{i=1}^{k} g(u_i)) \sum_{s_{1,2},\dots,s_{k,k-1}}^{*} (\prod_{1\leq i.j\leq k} \mu(s_{i,j})) \sum_{\substack{d_1,\dots,d_k \\ e_1,\dots,e_k \\ u_i|d_i,e_i \forall i \\ s_{i,j}|d_i,e_j \forall i\neq j \\ d_m=e_m=1}} \frac{\lambda_{d_1,\dots,d_k} \lambda_{e_1,\dots,e_k}}{\prod_{i=1}^{k} \varphi(d_i) \prod_{i=1}^{k} \varphi(e_i)} \tag{50}$$

A similar substition may be used in this situation we have only one extra demand $r_m = 1$. When $r_m \neq 0$ then $y^{(m)}_{r_1,\dots,r_k}$ is 0. Thus let

$$y^{(m)}_{r_1,\dots,r_k} = \prod_{i=1}^{k} \mu(r_i) g(r_i) \sum_{\substack{d_1,..,d_k \\ r_i|d_i \forall i \\ d_m=1}} \frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^{k} \varphi(d_i)} \tag{51}$$

Hence the main term becomes in a similar way as equations (**??**) to (32)

$$\frac{X_N}{\varphi(W)} \sum_{u_1,\dots,u_k} (\prod_{i=1}^{k} (\frac{\mu(u_i)^2}{g(u_i)}) \sum_{s_{1,2},\dots,s_{k,k-1}} (\prod_{\substack{1\leq i.j\leq k \\ i\neq j}} (\frac{\mu(s_{i,j})}{g(s_{i,j})}) y^{(m)}_{a_1,\dots,a_k} y^{(m)}_{b_1,\dots,b_k} \tag{52}$$

Again there are two different cases for $s_{i,j}$ is 1 or $> D_0$. When $s_{i,j}$ isn't 1 the contribution is by the same calculation as S_1

$$\ll \frac{(y^{(m)}_{max})^2 N}{\varphi(W) LogN} (\sum_{\substack{u<R \\ (u,W)=1}} \frac{\mu(u)^2}{g(u)})^{k-1} (\sum_{s} \frac{\mu(s)^2}{g(s)^2})^{k(k-1)-1} \sum_{s_{i,j}>D_0} \frac{\mu(s_{i,j})^2}{g(s_{i,j})^2}$$

$$\ll \frac{(y^{(m)}_{max})^2\varphi(W)^{k-2}N(logR)^{k-1}}{W^{k-1}D_0 LogN} \tag{53}$$

And by the prime number theorem $X_N = N/LogN + O(N/(LogN)^2)$ which contributes to the error term by

$$\ll \frac{(y^{(m)}_{max})^2 N}{\varphi(W)(LogN)^2}(\sum_{\substack{u<R\\(u,W)=1}} \frac{\mu(u)^2}{g(u)})^{k-1} \ll \frac{(y^{(m)}_{max})^2\varphi(W)^{k-2}N(logR)^{k-3}}{W^{k-1}} \tag{54}$$

Which will be absorbed by the first error term. With this we can rewrite the sum

$$S^{(m)}_2 = \frac{N}{\varphi(W)LogN}\sum_{r_1,\ldots r_k}\frac{(y^{(m)}_{r_1,\ldots,r_k})^2}{\sum_{i=1}^k g(r_i)} + O(\frac{(y^{(m)}_{max})^2\varphi(W)N(LogR)^{k-2}}{W^{k-1}D_0}) + O(\frac{y^{(m)}_{max}N}{(LogN)^A}) \tag{55}$$

Which ends the proof. □

Relating the Variables

Next the new variables of S_1 will be related to S_2^m by the following lemma.

Lemma 5. *if* $r_m = 1$ *then*

$$y^{(m)}_{r_1,\ldots,r_k} = \sum_{a_m}\frac{y_{r_1,\ldots,r_{m-1},a_m,r_{m+1},\ldots,r_k}}{\varphi(a_m)} + O(\frac{y_{max}\varphi(W)LogR}{WD_0})$$

Proof. Substitue the expression from (**??**) to (27) in definition (51) we get

$$y^{(m)}_{r_1,\ldots,r_k} = (\prod_{i=1}^k \mu(r_i)g(r_i))\sum_{\substack{d_1,\ldots,d_k\\ r_i|d_i\forall i\\ d_m=1}}(\prod_{i=1}^k\frac{\mu(d_i)d_i}{\varphi(d_i)})\sum_{\substack{a_1,\ldots,a_k\\ d_i|a_i\forall i}}\frac{y_{a_1,\ldots,a_k}}{\prod_{i=1}^k\varphi(a_i)} \tag{56}$$

By switching the summation of d and a variables the last sum is over d. This sum we can calculate explicitly. Because all the d_i, a_i and r_i are square-free and the functions are multiplicative

$$\sum_{\substack{d_1,\dots,d_k \\ d_i|a_i, r_i|d_i \forall i \\ d_m=1}} \prod_{i=1}^{k} \frac{\mu(d_i)d_i}{\varphi(d_i)} = \prod_{i\neq m} \sum_{\substack{d_i \\ d_i|a_i/r_i}} \frac{\mu(d_i)d_i}{\varphi(d_i)} \frac{\mu(r_i)r_i}{\varphi(r_i)} = \prod_{i\neq m} \frac{\mu(a_i/r_i)}{\varphi(a_i/r_i)} \frac{\mu(r_i)r_i}{\varphi(r_i)} \tag{57}$$

Hence we get

$$y^{(m)}_{r_1,\dots,r_k} = (\prod_{i=1}^{k} \mu(r_i)g(r_i)) \sum_{\substack{a_1,\dots,a_k \\ r_i|a_i\forall i}} \frac{y_{a_1,\dots,a_k}}{\prod_{i=1}^{k}\varphi(a_i)} \prod_{i\neq m} \frac{\mu(a_i)r_i}{\varphi(a_i)} \tag{58}$$

Looking at the support of $y^{(m)}_{r_1,\dots,r_k}$ we can restrict the summation over a_i to $(a_i, W) = 1$. Thus either $a_i = r_i$ or $a_i > D_0 r_i$. First look at the contribution of $a_i \neq r_i$ for $j \neq m$. Split the last sum in three parts. The part of $a_j > D_0 r_j$. The part of a_m and the rest. The first and last sums converge thus restricted to a constant. The middle sum is estimated by (14).

$$\ll y_{max}(\prod_{i=1}^{k} g(r_i)r_i)(\sum_{a_j>D_0r_j} \frac{\mu(a_j)^2}{\varphi(a_j)^2})(\sum_{\substack{a_m<R \\ (a_m,W)=1}} \frac{\mu(a_m)^2}{\varphi(a_m)}) \prod_{\substack{1\leq i\leq k \\ i\neq j,m}} (\sum_{r_i|a_i} \frac{\mu(a_i)^2}{\varphi(a_i)^2})$$

$$\ll (\prod_{i=1}^{k} \frac{g(r_i)r_i}{\varphi(r_i)^2}) \frac{y_{max}\varphi(W)LogR}{WD_0} \ll \frac{y_{max}\varphi(W)LogR}{WD_0} \tag{59}$$

Thus estimate the main contribution when $a_j = r_j$ for all $j \neq m$

$$y^{(m)}_{r_1,\dots,r_k} = (\prod_{i=1}^{k} \frac{r_i g(r_i)}{\varphi(r_i)^2}) \sum_{a_m} \frac{y_{r_1,\dots,r_{m-1},a_m,r_{m+1},\dots,r_k}}{\varphi(a_m)} + O(\frac{y_{max}\varphi(W)LogR}{WD_0}) \tag{60}$$

Note that $g(p)p/\varphi(p)^2 = 1 + O(p^{-2})$. Since the contribution is zero unless the product of the r_i is coprime to W every $r_i > D_0$ thus the product in the expression may be replaced by $1 + O(D^{-1})$. This will be dominated by the error term we already have. Thus this gives the result

$$y^{(m)}_{r_1,\dots,r_k} = \sum_{a_m} \frac{y_{r_1,\dots,r_{m-1},a_m,r_{m+1},\dots,r_k}}{\varphi(a_m)} + O(\frac{y_{max}\varphi(W)LogR}{WD_0}) \tag{61}$$

This will end the proof and gives a way te relate both new variables. □

Choosing Suitable y

To complete the proof of proposition 1 we need a suitable choice of y.
The choice of y is such that the ratio between the main terms of S_1 and S_2 is maximized. A second demand is that the y are smooth. Such that it has no dependence on the prime factorisation of the r_i. Remember $r = \prod_{i=1}^{k} r_i$ satisfies $(r, W) = 1$ and $\mu(r)^2 = 1$. Choose

$$y_{r_1,...,r_k} = F(\frac{\log r_i}{\log R}, ..., \frac{\log r_k}{\log R}) \tag{62}$$

for some piecewise differentiable function $F : \mathbb{R}^k \to \mathbb{R}$ supported on $\mathcal{R}_k = \{(x_1, ..., x_k) \in [0,1]^k : \sum_{i=1}^{k} x_i \leq 1\}$. When r is nog coprime to W or not square free set $y_{r_1,...,r_k} = 1$. With this choice suitable asymptotic estimates for S_1 and S_2 can be made.

Lemma 6. *Let $\kappa, A_1, A_2, L > 0$, Let γ be a multiplicative function satisfying*

$$0 \leq \frac{\gamma(p)}{p} \leq 1 - A_1$$

and

$$-L \leq \sum_{w \leq p \leq z} \frac{\gamma(p)\log p}{p} - \kappa \log z/w \leq A_2$$

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

$$\sum_{d<z} \mu(d)^2 g(d) G(\frac{\log d}{\log z}) = \mathscr{G}\frac{(\log z)^\kappa}{\Gamma(\kappa)} \int_0^1 G(x)x^{\kappa-1}dx + O_{A_1,A_2,\kappa}(\mathscr{G} L G_{max}(\log z)^{\kappa-1})$$

where

$$\mathscr{G} = \prod_p (q - \frac{\gamma(p)}{p})^{-1}(q - \frac{1}{p})^\kappa$$

The constant implied by O is independent of L and G

Proof. This proof is divided in two lemmas in [3]. Lemmas two and three prove this result with a slight notation difference. Which again is divided in multiple lemmas and based on explicit estimates of selbergs upper bounds. Lemmas 5.3 and 5.4 [5]
The next two lemmas will finish the estimation of S_1 and S_2^m this will conclude the proof of Proposition 1. The next lemma will estimate S_1

Lemma 7. *Let $y_{r_1,\dots,r_k}$ be given in terms of a piecewise differentiable function F supported on $\mathcal{R}_k = \{(x_1,..,x_k) \in [0,1]^k : |sum_{i=1}^k x_i \leq 1\}$ by (62). Let*

$$F_{max} = \sup_{(t_1,\dots,t_k)\in[0,1]^k} |F(t_1,...,t_k)| + \sum_{i=1}^{k} |\frac{\delta F}{\delta t_i}(t_1,...,t_k)| \tag{63}$$

Then

$$S_1 = \frac{\varphi(W)^k N(\log R)^k}{W^{k+1}} I_k(F) + O(\frac{F_{max}^2 \varphi(W)^k N(\log N)^k}{W^{k+1} D_0}) \tag{64}$$

where

$$I_k(F) = \int_0^1 ... \int_0^1 F(t_1,...,t_k)^2 dt_1...dt_k \tag{65}$$

Proof. We substitute the choice of y (62) into the expression of S_1 which was rewritten in lemma (2). There are no restrictions given for the *k-tuples* which are supported on y so we will add the restrictions to the sum. This gives

$$S_1 = \frac{N}{W} \sum_{\substack{u_1,\dots u_k \\ (u_i,u_j)=1 \forall i\neq j \\ (u_i,W)=1\forall i}} (\prod_{i=1}^{k} \frac{\mu(u_i)^2}{\varphi(u_i)}) F(\frac{\log u_i}{\log R}, ..., \frac{\log u_k}{\log R})^2 + O(\frac{y_{max}^2 N(LogR)^k \varphi(W)^k}{W^k D_0}) \tag{66}$$

Note that the requirement of $(u_i, u_j) = 1$ can be dropped at the cost of an reasonably sized error. This is because every prime divider they have in common has to be bigger than D_0 for they both are coprime with W. This error is of size

$$\ll \frac{F_{max}^2 N}{W} \sum_{p>D_0} \sum_{\substack{u_1,\dots u_k \\ p|u_i,u_j \\ (u_i,W)=1\forall i}} \prod_{i=1}^{k} \frac{\mu(r_i)^2}{\varphi(u_i)}$$

$$\ll \frac{F_{max}^2 N}{W} \sum_{p>D_0} \frac{1}{(p-1)^2} (\sum_{\substack{u<R \\ (u,W)=1}} \frac{\mu(u)^2}{\varphi(u)})^k \ll \frac{F_{max}^2 N (\log R)^k \varphi(W)^k}{W^{k+1} D_0} \tag{67}$$

The second sum is rewritten in the same way as in the estimations of 2 to $\varphi^k (\log R)^k / W^k$ while in the first sum "hier moet nog een afschatting" and rewrite it to $1/D_0$. Which will give the desired result.
Thus the sum that needs to be evaluated is

$$\sum_{\substack{u_i,...,u_k \\ (u_i,W)=1 \forall i}} (\prod_{i=1}^{k} \frac{\mu(u_1)^2}{\varphi(u_i)}) F(\frac{\log u_i}{\log R}, ..., \frac{\log u_k}{\log R})^2 \tag{68}$$

By k-fold application of Lemma.6 the sum can be estimated. Take for each application $\kappa = 1$ and

$$\gamma(p) = \begin{cases} 1, & p \not| W \\ 0, & otherwise \end{cases} \tag{69}$$

$$L \ll 1 + \sum_{p|W} \frac{\log p}{p} \ll \log D_0 \tag{70}$$

and A_1 and A_2 constants of suitable size.

Lemma 8. *If $\gamma(p)$, κ stated as above then the lower limit L stated in Lemma 6 will be* $\log D_0$

Proof. One chooses $\kappa = 1$ and no terms are counted when $p < D_0$ and use the maximum of $\log z/w$. The prime counting function $\pi(x)$ will be used. But not $x/\log x$ a stricter approximation is needed. A stricter approxomation is $\pi(x) = \int_2^x dt/\log t + O(x/(\log x)^A$ [1]

$$\begin{aligned} \pi(x) = \left[\frac{t}{\log t}\right]_2^x - \int_2^x t \frac{d}{dt}\left(\frac{1}{\log t}\right) dt + O(x/(\log x)^A) \\ \frac{x}{\log x} + \frac{x}{(\log x)^2} + O(x/(\log x)^3) \end{aligned} \tag{71}$$

Look at $\sum_{D_0<p<R} \log p/p$ over primes.

$$\sum_{D_0<p<R} \log p/p = \int_{D_0}^{R} \frac{\log x}{x} d\pi(x) = \left[\pi(x)\frac{\log x}{x}\right]_{D_0}^{R} - \int_{D_0}^{R} \pi(x) \frac{d}{dx}\left(\frac{\log x}{x}\right)$$

By the prime number theorem [4, p. 352] the first part is $O(1)$ for the second part the above estimation of $\pi(x)$ is used. This gives.

$$= O(1) + \int_{D_0}^{R} \left(\frac{x}{\log x} + \frac{x}{(\log x)^2}\right)\left(\frac{\log x}{x^2} - \frac{1}{x^2}\right) dx$$

$$= O(1) + \int_{D_0}^{R} \frac{1}{x} - \frac{1}{x(\log x)^2} dx = C + \log R/D_0$$

This constant $C > 0$ thus our estimation becomes

$$-\log D_0 \leq C + \log R/D_0 - \log R \leq C \tag{72}$$

Thus $L = \log D_0$ □

It will be shown how this works when $k = 2$. When $\kappa = 1$ and because W is square-free

$$\mathscr{G} = \prod_{p|W}\left(1 - \frac{1}{p}\right) = \frac{\varphi(W)}{W} \tag{73}$$

Thus we get

$$\sum_{u_1<R} \frac{\mu(u_1)^2}{\varphi(u_1)} \sum_{u_2<R} \frac{\mu(u_2)^2}{\varphi(u_2)} F(\frac{\log u_1}{\log R}, \frac{\log u_2}{\log R})^2$$

$$= \sum_{u_1<R} \frac{\mu(u_1)^2}{\varphi(u_1)} \left(\frac{\varphi(W)(\log R)}{W} \int\limits_0^1 F(x, \frac{\log u_2}{\log R})^2 dx + O(\frac{\varphi(W)F_{max}^2(\log D_0)}{W}) \right)$$

$$= \frac{\varphi(W)(\log R)}{W} \left(\int\limits_0^1 \frac{\varphi(W)(\log R)}{W} \int\limits_0^1 F(x, y)^2 dx + O(\frac{\varphi(W)F_{max}^2(\log D_0)}{W}) dy \right) + O(\frac{\varphi(W)F_{max}^2(\log D_0)}{W})$$

$$= \frac{\varphi(W)^2(\log R)^2}{W^2} I_2(F) + \int_0^1 O(\frac{\varphi(W)^2 F_{max}^2(\log D_0)\log R}{W^2}) dy + O(\frac{\varphi(W)F_{max}^2(\log D_0)}{W})$$

$$= \frac{\varphi(W)^2(\log R)^2}{W^2} I_2(F) + O(\frac{\varphi(W)^2 F_{max}^2(\log D_0)\log R}{W^2}) \tag{74}$$

If this is applied k times to the sum in equation (68) we get

$$\sum_{\substack{u_1,\ldots,u_k \\ (u_i,W)=1 \forall i}} (\prod_{i=1}^{k} \frac{\mu(u_1)^2}{\varphi(u_i)}) F(\frac{\log r_i}{\log R}, \ldots, \frac{\log r_k}{\log R})^2$$

$$= \frac{\varphi(W)^k (\log R)^k}{W^k} I_k(F) + O(\frac{\varphi(W)^k F_{max}^k (\log D_0)(\log R)^{k-1}}{W^k}) \qquad (75)$$

And by combining (75), (67) and (66) this ends the proof of lemma (7) $\square$

To end the proof of the proposition the following lemma needs to be proved.

Lemma 9. *Let $y_{r_1,...,r_k}$, F and F_{max} be as defined in Lemma 7. Then we have*

$$S_2^{(m)} = \frac{\varphi(W)^k N (\log R)^{k+1}}{W^{k+1} \log N} J_k^{(m)}(F) + O(\frac{F_{max}^2 \varphi(W)^k N (\log N)^k}{W^{k+1} D_0}) \qquad (76)$$

where

$$J_k^{(m)}(F) = \int_0^1 ... \int_0^1 (\int_0^1 F(t_1, ..., t_k) dt_m)^2 dt_1 ... dt_{m-1} dt_{m+1} ... dt_k \qquad (77)$$

Proof. The proof is similar to the proof of Lemma 7. First estimate $y^{(m)}_{r_1,...,r_k}$. Recall $y^{(m)}_{r_1,...,r_k} = 0$ unless $r_m = 1$ and $r = \prod_{i=1}^k r_i$ satisfies $(r, W) = 1$ and $\mu(r)^2 = 1$. Then $y^{(m)}_{r_1,...,r_k} = 0$ is given by Lemma 5. First the case when $y^{(m)}_{r_1,...,r_k} \neq 0$ is checked. Substitute (62) in the expression of Lemma 5 and this gives

$$y^{(m)}_{r_1,...,r_k} = \sum_{(u, W\prod_{i=1}^k) r_i) = 1} \frac{\mu(u)^2}{\varphi(u)} F(\frac{\log r_i}{\log R}, ..., \frac{\log r_{m-1}}{\log R}, \frac{\log u}{\log R}, \frac{\log r_{m+1}}{\log R}, ..., \frac{\log r_k}{\log R})$$

$$+ O(\frac{F_{max} \varphi(W) \log R}{W D_0}) \qquad (78)$$

Which makes $y^{(m)}_{max} \ll \varphi(W) F_{max} (\log R)/W$. Now estimate the sum over u. Lemma 6 is used with $\kappa = 1$,

$$\gamma(p) = \begin{cases} 1, & p \not| W \prod_{i=1}^k r_i \\ 0, & otherwise \end{cases} \qquad (79)$$

$$L \ll 1 + \sum_{p | W \prod_{i=1}^k r_i} \frac{\log p}{p} \ll \sum_{p < \log R} \frac{\log p}{p} + \sum_{\substack{p | W \prod_{i=1}^k r_i \\ p > \log R}} \frac{\log \log R}{\log R} \ll \log \log N \qquad (80)$$

and with A_1, A_2 suitable fixed constants. In the same way as in lemma 7 it is easy to see $\mathscr{G} = \varphi(W) \prod_{i=1}^k \varphi(r_i) / W \prod_{i=1}^k r_i$ which gives

$$y^{(m)}_{r_1,\dots,r_k} = (\log R)\frac{\varphi(W)}{W}(\prod_{i=1}^{k}\frac{\varphi(r_i)}{r_i})F^{(m)}_{r_1,\dots,r_k} + O(\frac{F_{max}\varphi(W)\log R}{WD_0}) \qquad (81)$$

where

$$F^{(m)}_{r_1,\dots,r_k} = \int_0^1 F\left(\frac{\log r_1}{\log R},\dots,\frac{\log r_{m-1}}{\log R},t_m,\frac{\log_{r_{m+1}}}{\log R},\dots,\frac{\log r_k}{\log R}\right)dt_m \qquad (82)$$

If this is substituted in the expression (40). The term of y_{max} is to the power of two. Thus the error term becomes $(F^2_{max}\varphi(W)^kN(\log R)^k)/(W^{k+1}D_0)$. The sum obtained is

$$S^{(m)}_2 = \frac{\varphi(W)N(\log R)^2F^2_{max}}{W^2\log N}\sum_{\substack{r_1,\dots,r_k\\(r_i,W)=1\forall i\\(r_i,r_j)=1\forall i\neq j\\r_m=1}}\left(\prod_{i=1}^{k}\frac{\mu(r_i)^2\varphi(r_i)}{g(r_i)r_i}\right)(F^{(m)}_{r_1,\dots,r_k})^2+$$

$$O\left(\frac{F^2_{max}\varphi(W)^kN(\log R)^k}{W^{k+1}D_0}\right) \qquad (83)$$

In the same way as in the first sum the condition $(r_i,r_j)=1$ can be removed. This will introduce an error of size

$$\ll \frac{\varphi(W)N(\log R)^2F^2_{max}}{W^2\log N}\left(\sum_{p>D_0}\frac{\varphi(p)^2}{g(p)^2p^2}\right)\left(\sum_{\substack{r<R\\(r,W)=1}}\frac{\mu(r)^2\varphi(r)}{g(r)r}\right)^{k-1}$$

The first sum will converge to a constant of $\ll 1/D_0$. The second sum can be estimated by $\varphi(W)^{k-1}(\log R)^{k-1}/W^{k-1} \leq \varphi(W)^{k-1}(\log N)^{k-1}/W^{k-1}$ which give an error term of

$$\ll \frac{F^2_{max}\varphi(W)^kN(\log N)^k}{W^{k+1}D_0} \qquad (84)$$

Now the last part of the sum needs to be evaluated. In the same way as the first sum we will use Lemma 6. Thus the sum we evaluate is

$$\sum_{\substack{r_1,\dots,rm-1,r_{m+1},\dots,r_k\\(r_i,W)=1\forall i}}\left(\prod_{1\leq i\leq ki\neq j}\frac{\mu(r_i)^2\varphi(r_i)}{g(r_i)r_i}\right)(F^{(m)}_{r_1,\dots,r_k})^2 \qquad (85)$$

By applying Lemma 6 $k-1$ times with $\kappa = 1$ and

$$\gamma(p) = \begin{cases} 1 + \frac{1}{p^2-p-1}, & p \not| W \\ 0, & otherwise \end{cases} \tag{86}$$

$$L \ll 1 + \sum_{p|W} \frac{\log p}{p} \ll \log D_0 \tag{87}$$

and A_1, A_2 suitable fixed constants. It is easy to see that error term (84) dominates the dominant error term we get when using this lemma $k-1$ times. Thus this gives

$$S_2^{(m)} = \frac{\varphi(W)^k N(\log R)^{k+1}}{W^{k+1}\log N} J_k^{(m)}(F) + O(\frac{F_{max}^2 \varphi(W)^k N(\log N)^k}{W^{k+1} D_0}) \tag{88}$$

where

$$J_k^{(m)}(F) = \int_0^1 ... \int_0^1 (\int_0^1 F(t_1, ..., t_k) dt_m)^2 dt_1 ... dt_{m-1} dt_{m+1} ... dt_k \tag{89}$$

as required. □

Comparing the sums

In the previous chapter we estimated the sums. Now it needs to be proven that by taking these estimations and a suitable k-tuple that there are infinitely many integers n such that several, but at least two, of the $n + h_i$ are prime. In the next proposition we will prove this and we will introduce a way to calculate the bound.

Proposition 4. *Let the primes have level of distribution $\theta > 0$. Let $\delta > 0$ and $\mathcal{H} = h_1, ..., h_k$ an admissible set. Let $I_k(F)$ and $J_k^{(m)}(F)$ be given as in Proposition 1, let $\mathcal{S}_k$ denote the set of piecewise differentiable functions $F : [0,1]^k \to \mathbb{R}$ supported on $\mathcal{R}_k = \{(x_1, ..., x_k) \in [0,1]^k : \sum_{i=1}^k x_i \leq 1\}$ with $I_k(F) \neq 0$ and $J_k^{(m)} \neq 0$ for each m. Let*

$$M_k = \sup_{F \in \mathcal{S}_k} \frac{\sum_{m=1}^k J_k^{(m)}}{I_k(F)}, \quad r_k = \lceil \tfrac{\theta M_k}{2} \rceil \tag{90}$$

Then there are infinitely many integers n such that at least r_k of the $n + h_i (1 \leq i \leq k)$ are prime. In particular, $\liminf_n (p_{n+r_k-1} - p_n) \leq \max_{1 \leq i,j \leq j}(h_i - h_j)$.

This proposition will tell when we have an upper bound $\theta M_k > 2$ for a k then we will have infinitely many n where $n + h_i$ with h_i in a admissible set of order k. Where at least two will be prime.

*Proof.*Let $S = S_2 - \rho S_1$. If $S > 0$ for all large N and $\rho > 1$ then there are infinitely many integers n such that at least two of the $n + h_i$ are prime.

Put $R = N^{\theta/2 - \epsilon}$ for a small $\epsilon > 0$. By the definition of M_k, a $F_0 \in \mathcal{S}_k$ can be chosen such that $\sum_{m=1}^k J_k^{(m)}(F_0) > (M_k - \epsilon) I_k(F_0) = \sum_{m=1}^k J_k^{(m)}(F_0) - \epsilon I_k(F_0)$. By using Proposition 1, we can rewrite both sums with given $\lambda_{d_1,...,d_k}$ such that

$$S = \frac{\varphi(W)^k N (\log R)^k}{W^{k+1}} \left(\frac{\log R}{\log N} \sum_{m=1}^k J_k^{(m)}(F_0) - \rho I_k(F_0) + o(1) \right)$$

$$\leq \frac{\varphi(W)^k N(\log R)^k I_k(F_0)}{W^{k+1}} \left((\frac{\theta}{2} - \epsilon)(M_k - \epsilon) - \rho + o(1) \right) \tag{91}$$

When $\rho = \theta M_k/2 - \delta$ by choosing ϵ sufficiently small then $S > 0$ for all large N. Thus there are infinitely many integers n for which at least $\lfloor \rho + 1 \rfloor$ of the $n + h_i$ are prime. Since $\lfloor \rho + 1 \rfloor = \lceil \theta M_k/2 \rceil$ when δ is sufficiently small the result is proven. □

To get the result of bounded gaps between primes an esimation of M_k and θ is needed. We need θ to be as large as possible. Remember if the primes have 'level of distribution' θ if for any $A > 0$, we have

$$\sum_{q \leq x^\theta} \max_{(a,q)=1} \left| \pi(x; q, a) - \frac{\pi(x)}{\varphi(q)} \right| \ll_A \frac{x}{(\log x)^A} \tag{92}$$

By the Bombieri-Vinogradov theorem it is proven when $\theta < 1/2$. If we find $M_k > 4$ for a k and an admissible set $\mathcal{H}_k$ with k elements there are infinitely many $n \in \mathbb{N}$ such that two of the $n + h_i$ are prime.

Obtaining a lower bound of M_k for small k

Let $\mathcal{S}_k$ denote the set of piecewise differentiable function $F : [0,1]^k \to \mathbb{R}$ supported on $\mathcal{R}_k = \{(x_1, ..., x_k) : \sum_{i=1}^k x_i \leq 1\}$ such that $I_k(F) \neq 0$ and $J_K^{(m)}(F) \neq 0$ for each m. If a lower bound is obtained, with these requirement, this will be

$$M_k = \sup_{F \in \mathcal{S}_k} \frac{\sum_{m=1}^k J_k^{(m)}(F)}{I_k(F)} \tag{93}$$

To obtain this lower bound we will consider approximations to the optimal function F of the form

$$F(t_1, ..., t_k) = \begin{cases} P(t_1, ..., t_k), & if (t_1, ..., t_k) \in \mathcal{R}_k \\ 0, & otherwise \end{cases} \tag{94}$$

for polynomials P.

Lemma 10. *Let $P_j = sum_{i=1}^k t_i^j$ denote the j^{th} symmetric power sum polynomial. Then we have*

$$\int\limits_{\mathcal{R}_k} (1 - P_1)^a P_j^b dt_1 ... dt_k = \frac{a!}{(k + jb + a)!} G_{b,j}(k) \tag{95}$$

where

$$G_{b,j}(x) = b! \sum_{r=1}^{b} \binom{k}{r} \sum_{\substack{b_1,\ldots,b_r \geq 1 \\ \sum_{i=1}^{r} b_i = b}} \prod_{i=1}^{r} \frac{(jb_1)!}{b_i!} \tag{96}$$

Proof. First by induction on k it follows

$$\int\limits_{\mathcal{R}_k} (1 - \sum_{i=1}^{k} t_i)^a \prod_{i=1}^{k} t_i^{a_i} dt_1, ..., dt_k = \frac{a! \prod_{i=1}^{k} a_i!}{(k + a + \sum_{i=1}^{k} a_i)!} \tag{97}$$

First consider the integration with respect to t_1. The limits of the integration are 0 and $1 - \sum_{i=2}^{k} t_i$ for $(t_2, ..., t_k) \in \mathcal{R}_{k-1}$. By substituing $v = t_1/(1 - \sum_{i=2}^{k} t_i)$ we find

$$\int_0^{1-\sum_{i=2}^{k} t_i} (1-\sum_{i=1}^{k} t_i)^a (\prod_{i=1}^{k} t_i^{a_i}) dt_1 = (\prod_{i=2}^{k} t_i^{a_i}))(1-\sum_{i=2}^{k} t_i)^{a+a_i+1} \int_0^1 (1-v)^a v^{a_1} dv$$

$$= \frac{a! a_1!}{(a + a_1 + 1)!} (\prod_{i=2}^{k} t_i^{a_i})(1 - \sum_{i=2}^{k} t_i)^{a+a_i+1} \tag{98}$$

Where the beta function identity is used $\int_0^1 t^a (1-t)^b dt = a!b!/(a+b+1)!$. The next step will give a fraction before the main term of $(a+a_1+1)!a_2!/(a+a_1+a_2+2)!$ thus the identity follows by induction.
By the multinomial theorem,

$$P_j^b = (\sum_{i=1}^{k} t_i^j)^b = \sum_{\substack{b_1,\ldots,b_k \\ \sum_{i=1}^{k} b_i = b}} \frac{b!}{\prod_{i=1}^{k} b_i!} \prod_{i=1}^{k} t_i^{jb_i} \tag{99}$$

Thus by applying this the following result will be obtained

$$\int\limits_{\mathcal{R}_k} (1 - P_1)^a P_j^b dt_1 ... dt_k = \sum_{\substack{b_1,\ldots,b_k \\ \sum_{i=1}^{k} b_i = b}} \frac{b!}{\prod_{i=1}^{k} b_i!} \int\limits_{\mathcal{R}_k} (1 - \sum_{i=1}^{k} t_i)^a \prod_{i=1}^{k} t_i^b dt_1, ..., dt_k$$

$$= \frac{b! a!}{(k + a + jb)!} \sum_{\substack{b_1,\ldots,b_k \\ \sum_{i=1}^{k} b_i = b}} \prod_{i=1}^{k} \frac{(jb_i)!}{b_i!} \tag{100}$$

For computations b will be small, and so it is convenient to split the summation over the b_i are non-zero. Given an integer r, there are $\binom{k}{r}$ ways of choosing r of $b_1, ..., b_k$ to be non-zero. Thus

$$\sum_{\substack{b_1,...,b_k \\ \sum_{i=1}^k b_i=b}} \prod_{i=1}^k \frac{(jb_i)!}{b_i!} = \sum_{r=1}^b \binom{k}{r} \sum_{\substack{b_1,...,b_k\geq 1 \\ \sum_{i=1}^k b_i=b}} \prod_{i=1}^k \frac{(jb_i)!}{b_i!} \tag{101}$$

This gives the result. □

Now a lemma is used to write $I_k(F)$ and $J_k^{(m)}(F)$ in terms of manageable expression with this choice of P.

Lemma 11. *Let F be given in terms of a polynomial P by (94). Let P be given in terms of a polynomial expression in the symmetric power polynomials* $P_1 = \sum_{i=1}^k t_i$ *and* $P_2 = \sum_{i=1}^k t_i^2$ *by* $P = \sum_{i=1}^d a_i(1-P_1)^{b_i}P_2^{c_i}$ *for constants* $a_i \in \mathbb{R}$ *and non-negative integers* b_i, c_i*. Then for each* $1 \leq m \leq k$ *we have*

$$I_k(F) = \sum_{1\leq i,j\leq d} a_i a_j \frac{(b_i+b_j)!G_{c_i+c_j,2}(k)}{(k+b_i+b_j+2c_i+2c_j)!}$$

$$J_k^{(m)}(F) = \sum_{1\leq i,j\leq d} a_i a_j \sum_{c_1'=0}^{c_i} \sum_{c_2'=0}^{c_j} \binom{c_i}{c_1'}\binom{c_i}{c_1'} \frac{\gamma_{b_i,b_j,c_i,c_j,c_1',c_2'} G_{c_1'+c_2'}(k-1)}{(k+b_i+b_j+2c_i+2c_j+1)!}$$

where

$$\gamma_{b_i,b_j,c_i,c_j,c_1',c_2'} = \frac{b_i!b_j!(2c_i-2c_1')!(2c_j'-2c_2')!(b_i+b_j+2c_i+2c_j-2c_1'-2c_2'+2)!}{(b_i+2c_i-2c_1'+1)!(b_j+2c_j-2c_2'+1)!}$$

and G is the polynomial given by Lemma 10

Proof. First consider $I_k(F)$. Using Lemma 10

$$I_k(F) = \int_{\mathcal{R}_k} P^2 dt_1...dt_k = \sum_{1\leq i,j\leq d} a_i a_j \iint_{\mathcal{R}_k} (1-P_1)^{b_i+b_j} P_2^{c_i+c_j} dt_1...dt_k$$

$$= \sum_{1\leq i,j\leq k} a_i a_j \frac{(b_i+b_j)!G_{c_i+c_j,2}(k)}{(k+b_i+b_j+2c_i+2c_j)!} \tag{102}$$

Thus the first part is proven. For the second sum remember that F is symmetric in $t_1, ..., t_k$ we see that $J_k^{(m)}(F)$ is independent of m, thus is suffices to only consider $J_k^{(1)}$. The first integral becomes

$$\int_0^{1-\sum_{i=2}^k}(1-P_1)^b P_2^c dt_1 = \sum_{c'=0}^{c}\binom{c}{c'}(\sum_{i=2}^k t_i^2)^{c'}\int_0^{1-\sum_{i=2}^k t_i}(1-\sum_{i=1}^k t_i)^b t_1^{2c-2c'}dt_1$$

Where P_2^c is split such we have one part with t_1 and the rest which can be handled as a constant.

$$= \sum_{c'=0}^{c}\binom{c}{c'}(P_2')^{c'}(1-P_1)^{b+2c-2c'+1}\int_0^1 (1-u)^b u^{2c-c'}du$$

$$= \sum_{c'=0}^{c}\binom{c}{c'}(P_2')^{c'}(1-P_1)^{b+2c-2c'+1}\frac{b!(2c-2c')!}{(b+2c-2c'+1)!} \tag{103}$$

Here $P_1' = \sum_{i=2}^k$ and $P_2' = \sum_{i=2}^k t_i^2$. Hence

$$(\int_0^1 F dt_1)^2 = (\sum_{i=1}^d a_i \int_0^{1-\sum_{i=2}^k t_i}(1-P_1)^{b_i}P_2^{c_i}dt_1)^2$$

$$= \sum_{1\le i,j\le d} a_i a_j \sum_{c_1'=0}^{c_i}\sum_{c_2'=0}^{c_j}\binom{c_i}{c_1'}\binom{c_j}{c_2'}(P_2')^{c_1'+c_2'}(1-P_1')^{b_i+b_j+2c_i+2c_j-2c_1'-2c_2'+2}$$

$$\times\frac{b_i!b_j!(2c_i-2c_1')!}{(b_i+2c_i-2c_1'+1)!(b_j+2c_j-2c_2'+1)!} \tag{104}$$

All terms are treated as constants in the other integrals except for the terms with P_1' and P_2' by Lemma 10 these terms are

$$\int_{\mathcal{R}_k}(1-P_1')^b(P_2')^c = \frac{b!}{(k-1+b+c)!}G_{c,2}(k-1) \tag{105}$$

with $b = c_1' + c_2'$ and $c = b_i + b_j + 2c_i + 2c_j - 2c_1' - 2c_2' + 2$ combining these, (105) and (104), will give the result. In Lemma 11 it has been proven that $I_k(F)$ and $J_k^{(m)}(F)$ can be written as quadratic forms of the a_i. Thus both terms can be writen as matrices over coefficients $a = (a_1, ..., a_d)$ of P. Moreover these will be positive definite real quadratic forms. Thus in particular M_k can be written as

$$M_k = \sup_{F\in\mathcal{S}_k}\frac{\sum_{m=1}^k J_k^{(m)}(F)}{I_k(F)} = \sup_{F\in\mathcal{S}_k}\frac{\mathbf{a}^T M_2\mathbf{a}}{\mathbf{a}^T M_1\mathbf{a}} \tag{106}$$

for two rational symmetric positive definite matrices M_1, M_2, which can be calculate explicitly in terms of k for any choice of the exponents b_i, c_i. This can be maximized and has a solution

Lemma 12. *Let M_1, M_2 be real, symmetric positive definite matrices. Then*

$$\frac{\boldsymbol{a}^T M_2 \boldsymbol{a}}{\boldsymbol{a}^T M_1 \boldsymbol{a}} \tag{107}$$

is maximized when $\boldsymbol{a}$ is an eigenvector of $M_1^{-1}M_2$ corresponding tot the largest eigenvalue of $M_1^{-1}M_2$. The value of the ratio at its maximum is this largest eigenvalue.

Proof. Multiplying $\mathbf{a}$ by a non zero scalar doesn't change the ratio, so we may assume without loss of generality that $\mathbf{a}^T M_1 \mathbf{a} = 1$. By using Lagrangian multipliers, $\mathbf{a}^T M_2 \mathbf{a}$ is maximized subjest to $\mathbf{a}^T M_1 \mathbf{a}$ when

$$L(\mathbf{a}, \lambda) = \mathbf{a}^T M_2 \mathbf{a} - \lambda(\mathbf{a}^T M_1 \mathbf{a} - 1) \tag{108}$$

is stationary. This occurs when (using the symmetricity of M_1, M_2)

$$0 = \frac{\delta L}{\delta a_i} = ((2M_2 - 2M_1)\mathbf{a})_i \tag{109}$$

for each i. This implies that (recalling M_1 is positive definite so invertible)

$$M_1^{-1} M_2 \mathbf{a} = \lambda \mathbf{a} \tag{110}$$

thus $\mathbf{a}^T M_1 \mathbf{a} = \lambda^{-1} \mathbf{a}^T M_2 \mathbf{a}$ □

Length of Bounded Gaps

By Lemma 12 a lower bound for M_k can be obtained. If F is chosen in the same way as in (94) the eigenvalues can be calculated.

Theorem 2.

$$i) \liminf_n (p_{n+1} - p_n) \leq 600$$

Proof. When $k = 105$ the largest eigenvalue of $M_1^{-1}M_2$ is

$$\lambda \approx 4.0030697 > 4 \tag{111}$$

Thus $M_{105} > 4$. Because $\theta = 1/2 - \epsilon$ for every $\epsilon > 0$ by the Bombieri-Vinogradov Theorem. Proposition 4 can be used. Calculating $r_k \lceil \frac{\theta M_k}{2} \rceil = \lceil (1/2) \times 4.0030697/2 \rceil = 2$. Hence if an admissible set with $k = 105$ elements is given. The longest interval in the admissible set will give a upperbound for the interval length. The admissible set that can be chosen is of length 600 thus by Proposition 4, Theorem 2 is proven. □

Theorem 3. *Assume that the primes have level of distribution θ for all $\theta < 1$. Then*

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

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

Proof. In the same way as in Theorem 2 it can be shown $M_{105} > 4$ thus $r_{105} = 3$ when $\theta < 1$ in the same way $r_5 = 2$. By using the admissible set $\mathcal{H} = \{0, 2, 6, 8, 12\}$ the first part is proven. Because $r_{105} = 3$ with this distribution of primes by Proposition 4 this means there are infinitely many n such that there are 3 primes in the set $n+h_i$ with h_i in the admissible set. This means if the same admissible set is used as in Theorem 2 the second part is proven. □

Bibliografie

[1] Tom M. Apostol. *Introduction to Analytic Number Theory*. Springer-Verlag, 1976.

[2] C. Y. Yildrim D. A. Goldston, J. Pintz. Positive proportioh of small gaps in consecutive primes. *arXiv:1103.3986*, 2011.

[3] J.Pintz C.Y.Yildrim D.A.Goldston, S.W.Graham. Small gaps between products of primes. *Proc. London Math. Soc.*, 3:741–774, 2009.

[4] G.H.Hardy and E.M.Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, 2008.

[5] H.-E.Richert H. Halberstam. *Sieve Methods*. Academic Press, 1974.

[6] James Maynard. Small gaps between primes. *arXiv:1311.4600v2*, 2013.

[7] M.Ram Murty. *Problems in Analytic Number Theory*. Springer-Verslag New York, Inc, 2001.

[8] D.H.J. Polymath. The bounded gaps between primes polymath project - a retrospective. *arXiv:1409.8361*, 2014.

[9] D. Zagier. Newman's short proof of the prime number theorem. *The American Mathematical Monthly*, 104:705–708, 1997.

[10] Y. Zhang. Bounded gaps between primes. *Annals of Mathematics*, 179:1121–1174.