Math 571 Chapter 8 Bounded Gaps in the Primes

Robert C. Vaughan

April 6, 2023

- A famous unsolved problem concerning prime numbers is the twin prime conjecture, namely that there are infinitely many pairs of primes which differ by 2.

- A famous unsolved problem concerning prime numbers is the twin prime conjecture, namely that there are infinitely many pairs of primes which differ by 2.
- Since the average spacing of primes $p \le x$ is $\log x$, this suggests that there are considerable local oscillations in the primes.

- A famous unsolved problem concerning prime numbers is the twin prime conjecture, namely that there are infinitely many pairs of primes which differ by 2.
- Since the average spacing of primes $p \le x$ is $\log x$, this suggests that there are considerable local oscillations in the primes.
- This has motivated a large body of work concerned with investigating the possibility of gaps between primes which are significantly smaller than the average gap.

- A famous unsolved problem concerning prime numbers is the twin prime conjecture, namely that there are infinitely many pairs of primes which differ by 2.
- Since the average spacing of primes $p \leq x$ is $\log x$, this suggests that there are considerable local oscillations in the primes.
- This has motivated a large body of work concerned with investigating the possibility of gaps between primes which are significantly smaller than the average gap.
- Since 2004 a very powerful theory has been developed. This modern theory is motivated by the following observations.

- Consider a k–tuple $h_1, h_2, \ldots, h_k$ of distinct non–negative integers for which it is believed that for infinitely many integers n the $n + h_1, \ldots, n + h_k$ are simultaneously prime.

- Consider a k–tuple $h_1, h_2, \ldots, h_k$ of distinct non–negative integers for which it is believed that for infinitely many integers n the $n + h_1, \ldots, n + h_k$ are simultaneously prime.
- Suppose we use a sieving technique to remove most n for which $n + h_1, \ldots, n + h_k$ are not all prime. Whilst it may not be possible to establish that, for each of the remaining n, the members of the k–tuple $n + h_1, \ldots, n + h_k$ are all prime there is a better chance of finding several primes in many of the k–tuples.

- Consider a k–tuple $h_1, h_2, \ldots, h_k$ of distinct non–negative integers for which it is believed that for infinitely many integers n the $n + h_1, \ldots, n + h_k$ are simultaneously prime.
- Suppose we use a sieving technique to remove most n for which $n + h_1, \ldots, n + h_k$ are not all prime. Whilst it may not be possible to establish that, for each of the remaining n, the members of the k–tuple $n + h_1, \ldots, n + h_k$ are all prime there is a better chance of finding several primes in many of the k–tuples.
- In its simplest form, suppose we are looking for primes in, say $[x, x + y]$. Since the expected number of primes is about $y/\log x$, if we pick an integer at random from the interval it is almost surely composite.

- Consider a k–tuple $h_1, h_2, \ldots, h_k$ of distinct non–negative integers for which it is believed that for infinitely many integers n the $n+h_1, \ldots, n+h_k$ are simultaneously prime.
- Suppose we use a sieving technique to remove most n for which $n+h_1, \ldots, n+h_k$ are not all prime. Whilst it may not be possible to establish that, for each of the remaining n, the members of the k–tuple $n+h_1, \ldots, n+h_k$ are all prime there is a better chance of finding several primes in many of the k–tuples.
- In its simplest form, suppose we are looking for primes in, say $[x, x+y]$. Since the expected number of primes is about $y/\log x$, if we pick an integer at random from the interval it is almost surely composite.
- But suppose we use a sieve to remove multiples of small primes to the extent that the number of remaining elements is about $2y/\log x$.

- Consider a k–tuple $h_1, h_2, \ldots, h_k$ of distinct non–negative integers for which it is believed that for infinitely many integers n the $n + h_1, \ldots, n + h_k$ are simultaneously prime.
- Suppose we use a sieving technique to remove most n for which $n + h_1, \ldots, n + h_k$ are not all prime. Whilst it may not be possible to establish that, for each of the remaining n, the members of the k–tuple $n + h_1, \ldots, n + h_k$ are all prime there is a better chance of finding several primes in many of the k–tuples.
- In its simplest form, suppose we are looking for primes in, say $[x, x + y]$. Since the expected number of primes is about $y/\log x$, if we pick an integer at random from the interval it is almost surely composite.
- But suppose we use a sieve to remove multiples of small primes to the extent that the number of remaining elements is about $2y/\log x$.
- Now if we pick an element at random from this sifted set, then we can expect that it is prime about half the time.

- As it stands just averaging over intervals does not work very well. But it turns out that averaging over suitable k-tuples of integers does.

Definition 1

Let $\mathbf{h} = h_1, \ldots, h_k$ be a k–tuple of distinct non–negative integers and let $\nu_p(\mathbf{h})$ denote the number of different residue classes modulo p among the $h_1, \ldots, h_k$. If $\nu_p(\mathbf{h}) < p$ for every p, then $\mathbf{h}$ is called admissible.

- As it stands just averaging over intervals does not work very well. But it turns out that averaging over suitable k-tuples of integers does.

Definition 1

Let $\mathbf{h} = h_1, \dots, h_k$ be a k–tuple of distinct non–negative integers and let $\nu_p(\mathbf{h})$ denote the number of different residue classes modulo p among the $h_1, \dots, h_k$. If $\nu_p(\mathbf{h}) < p$ for every p, then $\mathbf{h}$ is called admissible.

- It is clear that if $\mathbf{h}$ is inadmissible, then there can only be a finite number of n for which the $n + h_1, \dots, n + h_k$ are simultaneously prime.

Conjecture 2 (The prime k–tuple conjecture)

It is conjectured that if $\mathbf{h}$ is admissible, then there are infinitely many n such that $n + h_1, \dots, n + h_k$ are simultaneously prime.

- It is useful to establish that there are admissible sets with fairly small largest element.

Theorem 3

Suppose that $k \ge 2$ and the primes $p_1, \ldots, p_k$ satisfy $k < p_1 < \ldots < p_k$. Then any translate of the k–tuple $\mathbf{p}$ forms an admissible set. In particular $\mathbf{h} = \{0, p_2 - p_1, \ldots, p_k - p_1\}$ is an admissible set and p_k can be chosen so that $p_k < k \log k + k \log\log k + O(k)$.

- It is useful to establish that there are admissible sets with fairly small largest element.

Theorem 3

Suppose that $k \ge 2$ and the primes $p_1, \ldots, p_k$ satisfy $k < p_1 < \ldots < p_k$. Then any translate of the k–tuple $\mathbf{p}$ forms an admissible set. In particular $\mathbf{h} = \{0, p_2 - p_1, \ldots, p_k - p_1\}$ is an admissible set and p_k can be chosen so that $p_k < k \log k + k \log\log k + O(k)$.

- We remark for future reference that $\pi(105) = 27$ and $\pi(743) = 132$ so that one can take $k = 105$ and there is an admissible 105–tuple with largest element $743 - 107 = 636$.

- *Suppose that* $k \geq 2$ *and the primes* $p_1, \ldots, p_k$ *satisfy* $k < p_1 < \ldots < p_k$. *Then any translate of the* k*–tuple* $\mathbf{p}$ *forms an admissible set. In particular* $\mathbf{h} = \{0, p_2 - p_1, \ldots, p_k - p_1\}$ *is an admissible set and* p_k *can be chosen so that* $p_k < k \log k + k \log\log k + O(k)$.

- *Suppose that* $k \geq 2$ *and the primes* $p_1, \ldots, p_k$ *satisfy* $k < p_1 < \ldots < p_k$. *Then any translate of the* k*–tuple* $\mathbf{p}$ *forms an admissible set. In particular* $\mathbf{h} = \{0, p_2 - p_1, \ldots, p_k - p_1\}$ *is an admissible set and* p_k *can be chosen so that* $p_k < k \log k + k \log\log k + O(k)$.
- *Proof* The last part of the theorem follows from the prime number theorem. To prove the first part, suppose on the contrary that there is a $q > 1$ such that every residue class modulo q contains a p_j. Then $q \leq k < p_1$. On the other hand there is a j such that $p_j \equiv 0 \pmod{q}$ and so $p_j = q \leq k$.

- One can consider applying the Hardy–Littlewood method to this question. Suppose that n is such that

$$h_1 < h_2 < \cdots < h_k, \quad n + h_j = p_j, \quad n \le x.$$

- One can consider applying the Hardy–Littlewood method to this question. Suppose that n is such that

$$h_1 < h_2 < \cdots < h_k, \quad n + h_j = p_j, \quad n \le x.$$

- Then with logarithmic weights we consider

$$R(x;\mathbf{h}) = \sum_{\substack{p_1<p_2<\ldots<p_k\le x+h_k\\ p_k-p_j=h_k-h_j}} (\log p_1)\ldots(\log p_k)$$

and

$$S(\alpha) = \sum_{p\le N} (\log p) e(\alpha p) \tag{1}$$

where $N = \lfloor x + h_k \rfloor$.

- One can consider applying the Hardy–Littlewood method to this question. Suppose that n is such that

$$h_1 < h_2 < \cdots < h_k, \quad n + h_j = p_j, \quad n \le x.$$

- Then with logarithmic weights we consider

$$R(x; \mathbf{h}) = \sum_{\substack{p_1 < p_2 < \ldots < p_k \le x + h_k \\ p_k - p_j = h_k - h_j}} (\log p_1) \ldots (\log p_k)$$

and

$$S(\alpha) = \sum_{p \le N} (\log p) e(\alpha p) \tag{1}$$

where $N = \lfloor x + h_k \rfloor$.

- Then

$$R(x, \mathbf{h}) =$$

$$\int_{\mathfrak{U}^{k-1}} S(-\alpha_1 - \cdots - \alpha_{k-1}) \prod_{j=1}^{k-1} \big(S(\alpha_j) e(\alpha_j (h_k - h_j)) \big) d\boldsymbol{\alpha}.$$

- By the way, it is often more convenient to rearrange the equations $p_j = n + h_j$ connecting the p_j into the form

$$p_j - p_1 = h_j - h_1 \quad (2 \le j \le k).$$

- By the way, it is often more convenient to rearrange the equations $p_j = n + h_j$ connecting the p_j into the form

$$p_j - p_1 = h_j - h_1 \quad (2 \le j \le k).$$

- Also, there is no real loss in generality in supposing that $h_1 = 0$.

- By the way, it is often more convenient to rearrange the equations $p_j = n + h_j$ connecting the p_j into the form

$$p_j - p_1 = h_j - h_1 \quad (2 \le j \le k).$$

- Also, there is no real loss in generality in supposing that $h_1 = 0$.
- Suppose that we can replace each $S(\alpha)$ by its expected approximation when α is "close" to a rational number with a "small" denominator and the contribution from the remaining $\boldsymbol{\alpha}$ is relatively "small". We are deliberately rather imprecise as this is purely speculative.

- Thus if $P = N^\delta$ for some small $\delta > 0$ we would hope to obtain something of the form $R(x, \mathbf{h}) \sim J\times$

$$\sum_{q\le P} \sum_{\mathbf{a}}^{*} \frac{c_q(a_1+\cdots+a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\left(\frac{a_j(h_k-h_j)}{q}\right)$$

where $\sum^*$ is over $\mathbf{a} \pmod q$ with $(a_1,\ldots,a_{k-1},q)=1$ and $J =$

$$\int_{\mathfrak{U}^{k-1}} T(-\beta_1 - .. - \beta_{k-1}) \prod_{j=1}^{k-1} T(\beta_j) e\big(\beta_j(h_k-h_j)\big)\, d\boldsymbol{\beta}$$

and

$$T(\beta) = \sum_{m=1}^{N} e(\beta m).$$

- Thus if $P = N^\delta$ for some small $\delta > 0$ we would hope to obtain something of the form $R(x, \mathbf{h}) \sim J\times$

$$\sum_{q\le P} \sum_{\mathbf{a}}{}^{*} \frac{c_q(a_1+\cdots+a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\left(\frac{a_j(h_k-h_j)}{q}\right)$$

where $\sum^*$ is over $\mathbf{a} \pmod q$ with $(a_1,\ldots,a_{k-1},q)=1$ and $J =$

$$\int_{\mathfrak{U}^{k-1}} T(-\beta_1 - .. - \beta_{k-1}) \prod_{j=1}^{k-1} T(\beta_j) e\big(\beta_j(h_k-h_j)\big)\, d\boldsymbol{\beta}$$

and

$$T(\beta) = \sum_{m=1}^{N} e(\beta m).$$

- It is believed generally that this should hold.

- $J =$

$$\int_{\mathfrak{U}^{k-1}} T(-\beta_1 - .. - \beta_{k-1}) \prod_{j=1}^{k-1} T(\beta_j) e\big(\beta_j(h_k - h_j)\big) d\boldsymbol{\beta}$$

and

$$T(\beta) = \sum_{m=1}^{N} e(\beta m).$$

- $J =$
$$\int_{\mathfrak{U}^{k-1}} T(-\beta_1 - .. - \beta_{k-1}) \prod_{j=1}^{k-1} T(\beta_j) e\big(\beta_j(h_k - h_j)\big) d\boldsymbol{\beta}$$
and
$$T(\beta) = \sum_{m=1}^{N} e(\beta m).$$
- The number J is the number of $m_1, \ldots, m_k$ with $1 \le m_j \le N$ and $m_j = m_k + h_j - h_k$, so that m_j is determined by m_k and so J is the number of m_k with $h_k - h_1 < m_k \le N + h_k = x + O(1)$.

- $J =$

$$\int_{\mathfrak{U}^{k-1}} T(-\beta_1 - .. - \beta_{k-1}) \prod_{j=1}^{k-1} T(\beta_j) e\big(\beta_j(h_k - h_j)\big) d\boldsymbol{\beta}$$

and

$$T(\beta) = \sum_{m=1}^{N} e(\beta m).$$

- The number J is the number of $m_1, \ldots, m_k$ with $1 \le m_j \le N$ and $m_j = m_k + h_j - h_k$, so that m_j is determined by m_k and so J is the number of m_k with $h_k - h_1 < m_k \le N + h_k = x + O(1)$.
- Hence $J = x + O(h_k)$.

- $J =$

$$\int_{\mathfrak{U}^{k-1}} T(-\beta_1 - .. - \beta_{k-1}) \prod_{j=1}^{k-1} T(\beta_j) e\big(\beta_j(h_k - h_j)\big) d\boldsymbol{\beta}$$

and

$$T(\beta) = \sum_{m=1}^{N} e(\beta m).$$

- The number J is the number of $m_1, \ldots, m_k$ with $1 \le m_j \le N$ and $m_j = m_k + h_j - h_k$, so that m_j is determined by m_k and so J is the number of m_k with $h_k - h_1 < m_k \le N + h_k = x + O(1)$.
- Hence $J = x + O(h_k)$.
- Thus it is expected that $R(x; \mathbf{h}) \sim x\mathfrak{S}(\mathbf{h}; P)$ where $\mathfrak{S}(\mathbf{h}; P) = \sum_{q \le P} f(q; \mathbf{h})$ and $f(q; \mathbf{h}) =$

$$\sum_{\mathbf{a}}{}^{*} \frac{c_q(-a_1 - \cdots - a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big).$$

- $\mathfrak{S}(\mathbf{h};P) = \sum_{q\le P} f(q;\mathbf{h})$ and $f(q;\mathbf{h}) =$

$$\sum_{\mathbf{a}}^{*} \frac{c_q(-a_1-\cdots-a_{k-1})}{\phi(q)^k}\prod_{j=1}^{k-1} c_q(a_j)e\Big(\frac{a_j(h_k-h_j)}{q}\Big).$$

- $\mathfrak{S}(\mathbf{h}; P) = \sum_{q \le P} f(q; \mathbf{h})$ and $f(q; \mathbf{h}) =$

$$\sum_{\mathbf{a}}{}^{*} \frac{c_q(-a_1 - \cdots - a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big).$$

- It is readily verified that f is a multiplicative function of q.

- $\mathfrak{S}(\mathbf{h}; P) = \sum_{q \le P} f(q; \mathbf{h})$ and $f(q; \mathbf{h}) =$

$$\sum_{\mathbf{a}}^{*} \frac{c_q(-a_1 - \cdots - a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big).$$

- It is readily verified that f is a multiplicative function of q.
- Moreover when $q = p^t$ with $t \ge 2$, since $(a_1, \dots, a_{k-1}, q) = 1$, for at least one j we have $p \nmid a_j$, and so $c_{p^t}(a_j) = 0$.

- $\mathfrak{S}(\mathbf{h}; P) = \sum_{q \le P} f(q; \mathbf{h})$ and $f(q; \mathbf{h}) =$

$$\sum_{\mathbf{a}}^{*} \frac{c_q(-a_1 - \cdots - a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big).$$

- It is readily verified that f is a multiplicative function of q.
- Moreover when $q = p^t$ with $t \ge 2$, since $(a_1, \ldots, a_{k-1}, q) = 1$, for at least one j we have $p \nmid a_j$, and so $c_{p^t}(a_j) = 0$.
- Thus f has its support on the squarefree numbers.

- $\mathfrak{S}(\mathbf{h}; P) = \sum_{q \le P} f(q; \mathbf{h})$ and $f(q; \mathbf{h}) =$

$$\sum_{\mathbf{a}}{}^{*} \frac{c_q(-a_1 - \cdots - a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big)$$

and f has its support on the squarefree numbers.

- $\mathfrak{S}(\mathbf{h};P) = \sum_{q\le P} f(q;\mathbf{h})$ and $f(q;\mathbf{h}) =$

$$\sum_{\mathbf{a}}{}^{*} \frac{c_q(-a_1-\cdots-a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k-h_j)}{q}\Big)$$

  and f has its support on the squarefree numbers.
- Now consider the case $q=p$.

- $\mathfrak{S}(\mathbf{h}; P) = \sum_{q \le P} f(q; \mathbf{h})$ and $f(q; \mathbf{h}) =$

$$\sum_{\mathbf{a}}{}^{*} \frac{c_q(-a_1 - \cdots - a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big)$$

  and f has its support on the squarefree numbers.
- Now consider the case $q = p$.
- Then $(a_1, \ldots, a_{k-1}, p) = 1$ holds for all $\mathbf{a}$ with $1 \le a_j \le p$ except $a_1 = \cdots = a_{k-1} = p$.

- $\mathfrak{S}(\mathbf{h}; P) = \sum_{q \le P} f(q; \mathbf{h})$ and $f(q; \mathbf{h}) =$

$$\sum_{\mathbf{a}}{}^{*} \frac{c_q(-a_1 - \cdots - a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big)$$

 and f has its support on the squarefree numbers.
- Now consider the case $q = p$.
- Then $(a_1, \ldots, a_{k-1}, p) = 1$ holds for all $\mathbf{a}$ with $1 \le a_j \le p$ except $a_1 = \cdots = a_{k-1} = p$.
- If we sum over all $\mathbf{a}$ with $1 \le a_j \le p$ we obtain $p^{k-1}N$ where N is the number of solutions of $r_j \equiv r_k + h_j - h_k \pmod{p}$ with $1 \le r_j \le p - 1$. Thus r_j is determined by r_k, and $r_k \not\equiv 0$ or $h_k - h_j$ for any j. Thus $N = p - \nu_p(\mathbf{h})$. The term with $a_1 = \ldots = a_{k-1} = p$ contributes $(p-1)^k$ and so $f(p; \mathbf{h}) =$

$$\frac{(p - \nu_p(\mathbf{h}))p^{k-1} - (p-1)^k}{(p-1)^k} = \frac{(1 - \nu_p(\mathbf{h})/p)}{(1 - 1/p)^k} - 1.$$

- $\mathfrak{S}(\mathbf{h};P)=\sum_{q\le P} f(q;\mathbf{h})$ and $f(q;\mathbf{h})=$

$$\sum_{\mathbf{a}}{}^{*}\,\frac{c_q(-a_1-\cdots-a_{k-1})}{\phi(q)^k}\prod_{j=1}^{k-1}c_q(a_j)e\Big(\frac{a_j(h_k-h_j)}{q}\Big).$$

  f is multiplicative, has its support on the squarefree numbers and $f(p;\mathbf{h})=$

$$\frac{(p-\nu_p(\mathbf{h}))p^{k-1}-(p-1)^k}{(p-1)^k}=\frac{(1-\nu_p(\mathbf{h})/p)}{(1-1/p)^k}-1.$$

- $\mathfrak{S}(\mathbf{h}; P) = \sum_{q \le P} f(q; \mathbf{h})$ and $f(q; \mathbf{h}) =$

$$\sum_{\mathbf{a}}{}^{*} \frac{c_q(-a_1 - \cdots - a_{k-1})}{\phi(q)^k} \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big).$$

f is multiplicative, has its support on the squarefree numbers and $f(p; \mathbf{h}) =$

$$\frac{(p - \nu_p(\mathbf{h}))p^{k-1} - (p-1)^k}{(p-1)^k} = \frac{(1 - \nu_p(\mathbf{h})/p)}{(1 - 1/p)^k} - 1.$$

- When $p \nmid D = \prod_{1 \le i < j \le k} |h_j - h_i|$ we have $\nu_p(\mathbf{h}) = k$. Thus $f(p; \mathbf{h}) \ll p^{-2}$. Hence $\mathfrak{S}(\mathbf{h}; P)$ converges absolutely to $\mathfrak{S}(\mathbf{h})$ as $P \to \infty$ where $\mathfrak{S}(\mathbf{h}) = \sum_{q=1}^{\infty} f(q; \mathbf{h})$

$$= \prod_p (1 + f(p; \mathbf{h})) = \prod_p \left(1 - \frac{\nu_p(\mathbf{h})}{p}\right) \left(1 - \frac{1}{p}\right)^{-k}$$

and $\mathfrak{S}(\mathbf{h}) \ll_k (\log\log(3D))^k \ll_k (\log\log(3 \max_j |h_j|))^k$.

- We have

$$\mathfrak{S}(\mathbf{h}) = \prod_p \left(1 - \frac{\nu_p(\mathbf{h})}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}$$

- We have

$$\mathfrak{S}(\mathbf{h}) = \prod_p \left(1 - \frac{\nu_p(\mathbf{h})}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}$$

- Suppose the h_j are distinct.

- We have

$$\mathfrak{S}(\mathbf{h}) = \prod_p \left(1 - \frac{\nu_p(\mathbf{h})}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}$$

- Suppose the h_j are distinct.
- If $\mathbf{h}$ is inadmissible, then $\mathfrak{S}(\mathbf{h}) = 0$.

- We have

$$\mathfrak{S}(\mathbf{h}) = \prod_p \left(1 - \frac{\nu_p(\mathbf{h})}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}$$

- Suppose the h_j are distinct.
- If $\mathbf{h}$ is inadmissible, then $\mathfrak{S}(\mathbf{h}) = 0$.
- If $\mathbf{h}$ is admissible, then we have $\nu_p(\mathbf{h}) \le \min(k, p-1)$ and so $1 - \nu_p(\mathbf{h})/p \ge 1/p$ when $p \le k$ and is $\ge 1 - k/p$ when $p > k$.

- We have

$$\mathfrak{S}(\mathbf{h}) = \prod_p \left(1 - \frac{\nu_p(\mathbf{h})}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}$$

- Suppose the h_j are distinct.
- If $\mathbf{h}$ is inadmissible, then $\mathfrak{S}(\mathbf{h}) = 0$.
- If $\mathbf{h}$ is admissible, then we have $\nu_p(\mathbf{h}) \le \min(k, p-1)$ and so $1 - \nu_p(\mathbf{h})/p \ge 1/p$ when $p \le k$ and is $\ge 1 - k/p$ when $p > k$.
- Thus there is a positive number $C(k)$ such that, when the h_j are distinct, $\mathbf{h}$ is admissible if and only if

$$C(k) < \mathfrak{S}(\mathbf{h}).$$

- We have

$$\mathfrak{S}(\mathbf{h}) = \prod_p \left(1 - \frac{\nu_p(\mathbf{h})}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}$$

- Suppose the h_j are distinct.
- If $\mathbf{h}$ is inadmissible, then $\mathfrak{S}(\mathbf{h}) = 0$.
- If $\mathbf{h}$ is admissible, then we have $\nu_p(\mathbf{h}) \le \min(k, p-1)$ and so $1 - \nu_p(\mathbf{h})/p \ge 1/p$ when $p \le k$ and is $\ge 1 - k/p$ when $p > k$.
- Thus there is a positive number $C(k)$ such that, when the h_j are distinct, $\mathbf{h}$ is admissible if and only if

$$C(k) < \mathfrak{S}(\mathbf{h}).$$

- This suggests a conjecture.

- This suggests a conjecture.

Conjecture 4

Suppose that $\mathbf{h}$ *is admissible. Then, as* $x \to \infty$,

$$R(x; \mathbf{h}) \sim x\mathfrak{S}(\mathbf{h}).$$

- This suggests a conjecture.

Conjecture 4

Suppose that $\mathbf{h}$ *is admissible. Then, as* $x \to \infty$,

$$R(x;\mathbf{h}) \sim x\mathfrak{S}(\mathbf{h}).$$

- This is highly speculative, of course, and establishing this is well beyond what can be done in the current state of knowledge.

- This suggests a conjecture.

Conjecture 4

Suppose that $\mathbf{h}$ *is admissible. Then, as* $x \to \infty$,

$$R(x;\mathbf{h}) \sim x\mathfrak{S}(\mathbf{h}).$$

- This is highly speculative, of course, and establishing this is well beyond what can be done in the current state of knowledge.
- The likelihood of discovering primes in the k–tuple $n+h_1,\ldots,n+h_k$ depends on the avoidance of the zero residue class modulo p for all primes p, so in other words $\mathbf{h}$ needs to be admissible.

- This suggests a conjecture.

Conjecture 4

Suppose that $\mathbf{h}$ *is admissible. Then, as* $x \to \infty$,

$$R(x;\mathbf{h}) \sim x\mathfrak{S}(\mathbf{h}).$$

- This is highly speculative, of course, and establishing this is well beyond what can be done in the current state of knowledge.
- The likelihood of discovering primes in the k–tuple $n+h_1,\ldots,n+h_k$ depends on the avoidance of the zero residue class modulo p for all primes p, so in other words $\mathbf{h}$ needs to be admissible.
- A measure of this is the singular series $\mathfrak{S}(\mathbf{h})$ and we can expect that this will arise naturally in the analysis.

- This suggests a conjecture.

Conjecture 4

Suppose that $\mathbf{h}$ *is admissible. Then, as* $x \to \infty$,

$$R(x; \mathbf{h}) \sim x\mathfrak{S}(\mathbf{h}).$$

- This is highly speculative, of course, and establishing this is well beyond what can be done in the current state of knowledge.
- The likelihood of discovering primes in the k–tuple $n + h_1, \ldots, n + h_k$ depends on the avoidance of the zero residue class modulo p for all primes p, so in other words $\mathbf{h}$ needs to be admissible.
- A measure of this is the singular series $\mathfrak{S}(\mathbf{h})$ and we can expect that this will arise naturally in the analysis.
- We can also deduce from our discussion above and the next theorem that there is a plentiful supply of admissible k–tuples.

- Counting admissible k-tuples in a box.

Theorem 5 (Gallagher)

Suppose that $k \ge 2$ and $\mathcal{H}$ is the set of k–tuples $\mathbf{h}$ of distinct integers $h_1, \ldots, h_k$ with $1 \le h_j \le H$, and let $\mathcal{A}$ be the subset of those $\mathbf{h}$ which are also admissible. Then

$$\sum_{\mathbf{h}\in\mathcal{A}} \mathfrak{S}(\mathbf{h}) = H^k + O(H^{k-1+\varepsilon}).$$

- Counting admissible k-tuples in a box.

Theorem 5 (Gallagher)

Suppose that $k \ge 2$ and $\mathcal{H}$ is the set of k–tuples $\mathbf{h}$ of distinct integers $h_1, \dots, h_k$ with $1 \le h_j \le H$, and let $\mathcal{A}$ be the subset of those $\mathbf{h}$ which are also admissible. Then

$$\sum_{\mathbf{h}\in\mathcal{A}} \mathfrak{S}(\mathbf{h}) = H^k + O(H^{k-1+\varepsilon}).$$

- In view of the observation above that if $\mathbf{h} \in \mathcal{H}$ is inadmissible, then $\mathfrak{S}(\mathbf{h}) = 0$, it suffices to prove the conclusion with $\mathcal{A}$ replaced by $\mathcal{H}$.

- Counting admissible k-tuples in a box.

Theorem 5 (Gallagher)

Suppose that $k \geq 2$ and $\mathcal{H}$ is the set of k–tuples $\mathbf{h}$ of distinct integers $h_1, \ldots, h_k$ with $1 \leq h_j \leq H$, and let $\mathcal{A}$ be the subset of those $\mathbf{h}$ which are also admissible. Then

$$\sum_{\mathbf{h} \in \mathcal{A}} \mathfrak{S}(\mathbf{h}) = H^k + O(H^{k-1+\varepsilon}).$$

- In view of the observation above that if $\mathbf{h} \in \mathcal{H}$ is inadmissible, then $\mathfrak{S}(\mathbf{h}) = 0$, it suffices to prove the conclusion with $\mathcal{A}$ replaced by $\mathcal{H}$.
- When $\nu_p(\mathbf{h}) = k$, $f(q) = f(q; \mathbf{h})$ satisfies $|f(p; \mathbf{h})| \leq \dfrac{C_k}{p^2}$ and otherwise $|f(p; \mathbf{h})| \leq \dfrac{C_k}{p}$, for some C_k.

- Counting admissible k-tuples in a box.

Theorem 5 (Gallagher)

Suppose that $k \ge 2$ and $\mathcal{H}$ is the set of k–tuples $\mathbf{h}$ of distinct integers $h_1, \ldots, h_k$ with $1 \le h_j \le H$, and let $\mathcal{A}$ be the subset of those $\mathbf{h}$ which are also admissible. Then

$$\sum_{\mathbf{h}\in\mathcal{A}} \mathfrak{S}(\mathbf{h}) = H^k + O(H^{k-1+\varepsilon}).$$

- In view of the observation above that if $\mathbf{h} \in \mathcal{H}$ is inadmissible, then $\mathfrak{S}(\mathbf{h}) = 0$, it suffices to prove the conclusion with $\mathcal{A}$ replaced by $\mathcal{H}$.
- When $\nu_p(\mathbf{h}) = k$, $f(q) = f(q;\mathbf{h})$ satisfies $|f(p;\mathbf{h})| \le \dfrac{C_k}{p^2}$ and otherwise $|f(p;\mathbf{h})| \le \dfrac{C_k}{p}$, for some C_k.
- Then $|f(q;\mathbf{h})| \le q^{-2} C_k^{\omega(q)}(D,q) \ll_\varepsilon q^{\varepsilon-2}(D,q)$.

- Then $|f(q;\mathbf{h})| \le q^{-2} C_k^{\omega(q)}(D,q) \ll_\varepsilon q^{\varepsilon-2}(D,q)$ where As above, $D = \prod_{1\le i<j\le k} |h_j - h_i|$, so that $D \le H^{k(k-1)/2}$.

- Then $|f(q;\mathbf{h})| \le q^{-2}C_k^{\omega(q)}(D,q) \ll_\varepsilon q^{\varepsilon-2}(D,q)$ where As above, $D=\prod_{1\le i<j\le k}|h_j-h_i|$, so that $D\le H^{k(k-1)/2}$.
- For convenience we introduce the parameter $Q\ge 1$ which is at our disposal. Then

$$\sum_{q>Q}|f(q;\mathbf{h})| \ll \sum_{r|D} r \sum_{\substack{q>Q\\(D,q)=r}} q^{\varepsilon-2}$$

$$\ll \sum_{r|D} r^{\varepsilon-1}\sum_{t>Q/r} t^{\varepsilon-2} \ll Q^{\varepsilon-1}d(D).$$

- Then $|f(q;\mathbf{h})| \le q^{-2}C_k^{\omega(q)}(D,q) \ll_\varepsilon q^{\varepsilon-2}(D,q)$ where As above, $D=\prod_{1\le i<j\le k}|h_j-h_i|$, so that $D\le H^{k(k-1)/2}$.
- For convenience we introduce the parameter $Q\ge 1$ which is at our disposal. Then

$$\sum_{q>Q}|f(q;\mathbf{h})| \ll \sum_{r|D} r \sum_{\substack{q>Q\\(D,q)=r}} q^{\varepsilon-2}$$
$$\ll \sum_{r|D} r^{\varepsilon-1} \sum_{t>Q/r} t^{\varepsilon-2} \ll Q^{\varepsilon-1}d(D).$$

- Hence

$$\sum_{q>Q}|f(q;\mathbf{h})| \ll Q^{\varepsilon-1}H^{\varepsilon}. \tag{2}$$

- Then $|f(q;\mathbf{h})| \le q^{-2}C_k^{\omega(q)}(D,q) \ll_\varepsilon q^{\varepsilon-2}(D,q)$ where As above, $D=\prod_{1\le i<j\le k}|h_j-h_i|$, so that $D\le H^{k(k-1)/2}$.
- For convenience we introduce the parameter $Q\ge 1$ which is at our disposal. Then

$$\sum_{q>Q}|f(q;\mathbf{h})| \ll \sum_{r|D} r \sum_{\substack{q>Q\\(D,q)=r}} q^{\varepsilon-2}$$
$$\ll \sum_{r|D} r^{\varepsilon-1}\sum_{t>Q/r} t^{\varepsilon-2} \ll Q^{\varepsilon-1}d(D).$$

- Hence

$$\sum_{q>Q}|f(q;\mathbf{h})| \ll Q^{\varepsilon-1}H^{\varepsilon}. \qquad (2)$$

- We take $Q=H$ and sum over the elements of $\mathcal{H}$ to obtain the bound $\ll H^{k-1+2\varepsilon}$.

- The case $k = 2$ is special so we treat that first. Then $f(q;\mathbf{h}) = \frac{\mu(q)^2}{\phi(q)^2} \sum_{\substack{a=1\\(a,q)=1}}^{q} e(a(h_1-h_2)/q)$ and so

$$\sum_{\mathbf{h}\in\mathcal{H}} f(q;\mathbf{h}) = \frac{\mu(q)^2}{\phi(q)^2} \sum_{h_2\le H} \sum_{\substack{a=1\\(a,q)=1}}^{q} \sum_{\substack{h_1\le H\\h_1\ne h_2}} e(a(h_1-h_2)/q).$$

- The case $k = 2$ is special so we treat that first. Then $$f(q;\mathbf{h}) = \frac{\mu(q)^2}{\phi(q)^2} \sum_{\substack{a=1\\(a,q)=1}}^{q} e(a(h_1-h_2)/q)$$ and so

$$\sum_{\mathbf{h}\in\mathcal{H}} f(q;\mathbf{h}) = \frac{\mu(q)^2}{\phi(q)^2} \sum_{h_2\le H} \sum_{\substack{a=1\\(a,q)=1}}^{q} \sum_{\substack{h_1\le H\\h_1\ne h_2}} e(a(h_1-h_2)/q).$$

- The innermost sum is $\ll \|a/q\|^{-1}$ and we have $\sum_{a=1}^{q-1} \|a/q\|^{-1} \ll q\log q$.

- The case $k=2$ is special so we treat that first. Then $f(q;\mathbf{h}) = \dfrac{\mu(q)^2}{\phi(q)^2} \displaystyle\sum_{\substack{a=1\\(a,q)=1}}^{q} e(a(h_1-h_2)/q)$ and so

$$\sum_{\mathbf{h}\in\mathcal{H}} f(q;\mathbf{h}) = \frac{\mu(q)^2}{\phi(q)^2} \sum_{h_2\le H} \sum_{\substack{a=1\\(a,q)=1}}^{q} \sum_{\substack{h_1\le H\\h_1\ne h_2}} e(a(h_1-h_2)/q).$$

- The innermost sum is $\ll \|a/q\|^{-1}$ and we have $\sum_{a=1}^{q-1}\|a/q\|^{-1} \ll q\log q$.
- Thus $\displaystyle\sum_{\mathbf{h}\in\mathcal{H}} f(1;\mathbf{h}) = H^2+O(H)$, since $f(1;\mathbf{h})=1$ and $\operatorname{card}\mathcal{H} = H^2+O(H)$, and we have $\displaystyle\sum_{\mathbf{h}\in\mathcal{H}}\sum_{1<q\le Q} f(q;\mathbf{h}) \ll HQ^{\varepsilon}$, so $Q=H$ gives case $k=2$

- Now suppose $k \geq 3$, and write $g(q;\mathbf{h}) = \phi(q)^k f(q;\mathbf{h})$

$$= \sum_{\mathbf{a}}{}^{*} c_q(-a_1 - \cdots - a_{k-1}) \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big).$$

- Now suppose $k \geq 3$, and write $g(q;\mathbf{h}) = \phi(q)^k f(q;\mathbf{h})$

$$= \sum_{\mathbf{a}}^{*} c_q(-a_1 - \cdots - a_{k-1}) \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big).$$

- Then $|g(q\,;\mathbf{h})| \leq g^*(q)$ where

$$g^*(q) = \sum_{\substack{\mathbf{a} \\ (\mathbf{a},q)=1}} |c_q(a_1)\dots c_q(a_{k-1}) c_q(-a_1 - \cdots - a_{k-1})|$$

and this is also a multiplicative function of q with its support on the square free numbers.

- Now suppose $k \ge 3$, and write $g(q;\mathbf{h}) = \phi(q)^k f(q;\mathbf{h})$

$$= \sum_{\mathbf{a}}{}^{*} c_q(-a_1 - \cdots - a_{k-1}) \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k - h_j)}{q}\Big).$$

- Then $|g(q\,;\mathbf{h})| \le g^*(q)$ where

$$g^*(q) = \sum_{\substack{\mathbf{a} \\ (\mathbf{a},q)=1}} |c_q(a_1)\dots c_q(a_{k-1})c_q(-a_1 - \cdots - a_{k-1})|$$

and this is also a multiplicative function of q with its support on the square free numbers.
- Thus

$$\sum_{\mathbf{h}\in[1,H]^k\setminus\mathcal{H}} \sum_{1\le q\le Q} f(q;\mathbf{h}) \ll H^{k-1} \prod_{p\le Q} \Big(1 + \frac{g^*(p)}{(p-1)^k.}\Big)$$

- Thus

$$\sum_{\mathbf{h}\in[1,H]^k\setminus\mathcal{H}} \sum_{1\le q\le Q} f(q;\mathbf{h}) \ll H^{k-1} \prod_{p\le Q} \left(1+\frac{g^*(p)}{(p-1)^k}\right).$$

where

$$g^*(p) = \sum_{\substack{\mathbf{a}\\(\mathbf{a},p)=1}} |c_p(a_1)\dots c_p(a_{k-1})c_p(-a_1-\dots-a_{k-1})|.$$

- Thus
$$\sum_{\mathbf{h}\in[1,H]^k\setminus\mathcal{H}}\ \sum_{1\le q\le Q} f(q;\mathbf{h}) \ll H^{k-1}\prod_{p\le Q}\left(1+\frac{g^*(p)}{(p-1)^k}\right).$$
where
$$g^*(p) = \sum_{\substack{\mathbf{a}\\(\mathbf{a},p)=1}} |c_p(a_1)\dots c_p(a_{k-1})c_p(-a_1-\dots-a_{k-1})|.$$
- Consider the k numbers $a_1,\dots,a_{k-1},-a_1-\dots-a_{k-1}$. When $(\mathbf{a},p)=1$ at least two of these numbers are not multiples of p. Moreover in $g^*(p)$ the terms with exactly j of the $a_1,\dots,a_{k-1},a_1+\dots+a_{k-1}$ divisible by p contribute $(p-1)^j$ and since the $a_1,\dots,a_{k-1},a_1+\dots+a_{k-1}$ are linearly dependent the number of such terms is at most $\binom{k}{j}(p-1)^{k-1-j}$.

- Thus
$$\sum_{\mathbf{h}\in[1,H]^k\setminus\mathcal{H}} \sum_{1\le q\le Q} f(q;\mathbf{h}) \ll H^{k-1} \prod_{p\le Q}\left(1+\frac{g^*(p)}{(p-1)^k}\right).$$
where
$$g^*(p) = \sum_{\substack{\mathbf{a}\\(\mathbf{a},p)=1}} |c_p(a_1)\dots c_p(a_{k-1})c_p(-a_1-\dots-a_{k-1})|.$$
- Consider the k numbers $a_1,\dots,a_{k-1},-a_1-\dots-a_{k-1}$. When $(\mathbf{a},p)=1$ at least two of these numbers are not multiples of p. Moreover in $g^*(p)$ the terms with exactly j of the $a_1,\dots,a_{k-1},a_1+\dots+a_{k-1}$ divisible by p contribute $(p-1)^j$ and since the $a_1,\dots,a_{k-1},a_1+\dots+a_{k-1}$ are linearly dependent the number of such terms is at most $\binom{k}{j}(p-1)^{k-1-j}$.
- Hence $g^*(p)\le 2^k(p-1)^{k-1}$ and
$$\sum_{\mathbf{h}\in[1,H]^k\setminus\mathcal{H}} \sum_{1\le q\le Q} f(q;\mathbf{h}) \ll H^{k-1}Q^{\varepsilon}.$$

- Consider $\displaystyle\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h})$ where $q>1$ and $g(q;\mathbf{h})$

$$= \sum_{\mathbf{a}}{}^{*} c_q(-a_1-\cdots-a_{k-1}) \prod_{j=1}^{k-1} c_q(a_j) e\Big(\frac{a_j(h_k-h_j)}{q}\Big).$$

- Consider $\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h})$ where $q>1$ and $g(q;\mathbf{h})$

$$= \sum_{\mathbf{a}}{}^{*} c_q(-a_1-\cdots-a_{k-1})\prod_{j=1}^{k-1} c_q(a_j)e\Big(\frac{a_j(h_k-h_j)}{q}\Big).$$

- At least two of $a_1,\ldots,a_{k-1},-a_1-\cdots-a_{k-1}$ are $\not\equiv 0$ (mod q). If there are at least two $a_i\not\equiv 0$, then pick two and call them b_1,b_2. List the rest as $b_3,\ldots,b_{k-1}$. Note $-a_1-\cdots-a_{k-1}=-b_1-\cdots-b_{k-1}$. If only one of the $a_i\not\equiv 0$, then call it b_1, and put $b_2=-a_1-\cdots-a_{k-1}$. Then any of the other a_i can be rewritten $-b_1-b_2-s$ (mod q) where s is the sum of the remaining a_t. Hence

$$\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h}) \ll H^{k-2}\sum_{b_1=1}^{q-1}\frac{|c_q(b_1)|}{\|b_1/q\|}\sum_{b_2=1}^{q-1}\frac{|c_q(b_2)|}{\|b_2/q\|}\times.$$

$$\sum_{b_3,\ldots,b_{k-1}\in[1,q]^{k-3}} |c_q(b_1+\cdots+b_{k-1})|\prod_{j=3}^{k-1}|c_q(b_j)|.$$

- Hence

$$\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h}) \ll H^{k-2} \sum_{b_1=1}^{q-1} \frac{|c_q(b_1)|}{\|b_1/q\|} \sum_{b_2=1}^{q-1} \frac{|c_q(b_2)|}{\|b_2/q\|} \times.$$

$$\sum_{\mathbf{b}\in[1,q]^{k-3}} |c_q(b_1+\cdots+b_{k-1})| \prod_{j=3}^{k-1} |c_q(b_j)|$$

where $\mathbf{b} = b_3, \dots, b_{k-1}$.

- Hence
$$\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h}) \ll H^{k-2} \sum_{b_1=1}^{q-1} \frac{|c_q(b_1)|}{\|b_1/q\|} \sum_{b_2=1}^{q-1} \frac{|c_q(b_2)|}{\|b_2/q\|} \times.$$
$$\sum_{\mathbf{b}\in[1,q]^{k-3}} |c_q(b_1+\cdots+b_{k-1})| \prod_{j=3}^{k-1} |c_q(b_j)|$$
where $\mathbf{b} = b_3, \ldots, b_{k-1}$.
- The inner sum does not exceed $\phi(q)\Big(\sum_{b=1}^{q}|c_q(b)|\Big)^{k-3}$.

- Hence
$$\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h}) \ll H^{k-2} \sum_{b_1=1}^{q-1} \frac{|c_q(b_1)|}{\|b_1/q\|} \sum_{b_2=1}^{q-1} \frac{|c_q(b_2)|}{\|b_2/q\|} \times.$$
$$\sum_{\mathbf{b}\in[1,q]^{k-3}} |c_q(b_1+\cdots+b_{k-1})| \prod_{j=3}^{k-1} |c_q(b_j)|$$
where $\mathbf{b} = b_3, \ldots, b_{k-1}$.
- The inner sum does not exceed $\phi(q)\Big(\sum_{b=1}^{q} |c_q(b)|\Big)^{k-3}$.
- As $|c_q(b)| \le (q,b)$ the sum here is $\le \sum_{r|q} r\phi(q/r) \le d(q)q$.

- Hence
$$\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h}) \ll H^{k-2} \sum_{b_1=1}^{q-1} \frac{|c_q(b_1)|}{\|b_1/q\|} \sum_{b_2=1}^{q-1} \frac{|c_q(b_2)|}{\|b_2/q\|} \times .$$
$$\sum_{\mathbf{b}\in[1,q]^{k-3}} |c_q(b_1+\cdots+b_{k-1})| \prod_{j=3}^{k-1} |c_q(b_j)|$$
where $\mathbf{b} = b_3, \ldots, b_{k-1}$.
- The inner sum does not exceed $\phi(q)\Big(\sum_{b=1}^{q} |c_q(b)|\Big)^{k-3}$.
- As $|c_q(b)| \le (q,b)$ the sum here is
$\le \sum_{r|q} r\phi(q/r) \le d(q)q$.
- Similarly
$$\sum_{b=1}^{q-1} \frac{|c_q(b)|}{\|b/q\|} \le \sum_{r|q} r \sum_{a=1}^{q/r-1} \|a/(q/r)\|^{-1} \ll d(q)q\log q.$$

- Hence

$$\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h}) \ll H^{k-2} \sum_{b_1=1}^{q-1} \frac{|c_q(b_1)|}{\|b_1/q\|} \sum_{b_2=1}^{q-1} \frac{|c_q(b_2)|}{\|b_2/q\|} \times .$$

$$\sum_{\mathbf{b}\in[1,q]^{k-3}} |c_q(b_1+\cdots+b_{k-1})| \prod_{j=3}^{k-1} |c_q(b_j)|.$$

$$\ll H^{k-2} d(q)^2 q^2 (\log q)^2 \phi(q) d(q)^{k-3} q^{k-3}.$$

- Hence

$$\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h}) \ll H^{k-2}\sum_{b_1=1}^{q-1}\frac{|c_q(b_1)|}{\|b_1/q\|}\sum_{b_2=1}^{q-1}\frac{|c_q(b_2)|}{\|b_2/q\|}\times.$$

$$\sum_{\mathbf{b}\in[1,q]^{k-3}} |c_q(b_1+\cdots+b_{k-1})|\prod_{j=3}^{k-1}|c_q(b_j)|.$$

$$\ll H^{k-2}d(q)^2q^2(\log q)^2\phi(q)d(q)^{k-3}q^{k-3}.$$

- Therefore

$$\sum_{\mathbf{h}\in[1,H]^k}\sum_{1<q\le Q} f(q,\ \mathbf{h}) \ll H^{k-2}Q^{1+\varepsilon}.$$

- Hence

$$\sum_{\mathbf{h}\in[1,H]^k} g(q;\mathbf{h}) \ll H^{k-2}\sum_{b_1=1}^{q-1}\frac{|c_q(b_1)|}{\|b_1/q\|}\sum_{b_2=1}^{q-1}\frac{|c_q(b_2)|}{\|b_2/q\|}\times.$$

$$\sum_{\mathbf{b}\in[1,q]^{k-3}} |c_q(b_1+\cdots+b_{k-1})|\prod_{j=3}^{k-1}|c_q(b_j)|.$$

$$\ll H^{k-2}d(q)^2q^2(\log q)^2\phi(q)d(q)^{k-3}q^{k-3}.$$

- Therefore

$$\sum_{\mathbf{h}\in[1,H]^k}\sum_{1<q\le Q} f(q,\ \mathbf{h}) \ll H^{k-2}Q^{1+\varepsilon}.$$

- The term $q=1$ contributes H^k and so $Q=H$ gives the theorem

- The principal idea is to use the Selberg sieve to enhance the chances of finding primes.

- The principal idea is to use the Selberg sieve to enhance the chances of finding primes.
- The starting point for the Selberg upper bound sieve is

$$\sum_{a\in\mathcal{A}}\Big(\sum_{\substack{q\le R\\ q|a}}\lambda_q\Big)^2.$$

- The principal idea is to use the Selberg sieve to enhance the chances of finding primes.
- The starting point for the Selberg upper bound sieve is
$$\sum_{a\in\mathcal{A}}\Bigl(\sum_{\substack{q\le R\\ q|a}}\lambda_q\Bigr)^2.$$
- One is planning to minimise this under the assumptions 1. $\lambda_1 = 1$ and 2. that
$$A_d = \sum_{\substack{a\in\mathcal{A}\\ d|a}} 1$$
can be approximated by $\dfrac{Xg(d)}{d}$ where g is multiplicative.

- The principal idea is to use the Selberg sieve to enhance the chances of finding primes.
- The starting point for the Selberg upper bound sieve is

$$\sum_{a\in\mathcal{A}}\Bigg(\sum_{\substack{q\le R\\ q|a}}\lambda_q\Bigg)^2.$$

- One is planning to minimise this under the assumptions 1. $\lambda_1 = 1$ and 2. that

$$A_d = \sum_{\substack{a\in\mathcal{A}\\ d|a}} 1$$

can be approximated by $\dfrac{Xg(d)}{d}$ where g is multiplicative.
- The minimising choice of λ_q is given by

$$\lambda_q = \mu(q)\frac{S(R,q)}{S(R,1)}\prod_{p|q}\left(\frac{p}{p-g(p)}\right)$$

where $S(R,q) = \displaystyle\sum_{r\le R/q,(r,q)=1}\mu(r)^2\prod_{p|r}\frac{g(p)}{p-g(p)}$.

- The minimising choice of λ_q is given by

$$\lambda_q = \mu(q)\frac{S(R,q)}{S(R,1)}\prod_{p|q}\left(\frac{p}{p-g(p)}\right)$$

where $S(R,q) = \displaystyle\sum_{r\le R/q,(r,q)=1}\mu(r)^2\prod_{p|r}\frac{g(p)}{p-g(p)}.$

- The minimising choice of λ_q is given by

$$\lambda_q = \mu(q)\frac{S(R,q)}{S(R,1)}\prod_{p|q}\left(\frac{p}{p-g(p)}\right)$$

where $S(R,q) = \displaystyle\sum_{r\le R/q,(r,q)=1} \mu(r)^2\prod_{p|r}\frac{g(p)}{p-g(p)}.$

- Typically this is applied when the sieve has dimension k, e.g.

$$\sum_{p\le y} g(p)\frac{\log p}{p} = k\log y + O(1).$$

- The minimising choice of λ_q is given by

$$\lambda_q = \mu(q)\frac{S(R,q)}{S(R,1)}\prod_{p|q}\left(\frac{p}{p-g(p)}\right)$$

where $S(R,q) = \displaystyle\sum_{r\le R/q,(r,q)=1}\mu(r)^2\prod_{p|r}\frac{g(p)}{p-g(p)}.$

- Typically this is applied when the sieve has dimension k, e.g.

$$\sum_{p\le y} g(p)\frac{\log p}{p} = k\log y + O(1).$$

- Under this kind of condition one might expect that

$$S(R,q)\sim C(\log R/q)^k\prod_{p|q}\frac{p-g(p)}{p}$$

and so λ_q could be replaced by

$$\lambda_q = \mu(q)\frac{\log^k(R/q)}{\log^k R} = \mu(q)\left(1-\frac{\log q}{\log R}\right)^k$$

- We expect that $S(R,q) \sim C(\log R/q)^k \prod_{p|q} \frac{p-g(p)}{p}$ and so λ_q could be replaced by

$$\lambda_q = \mu(q)\frac{\log^k(R/q)}{\log^k R} = \mu(q)\left(1 - \frac{\log q}{\log R}\right)^k$$

- We expect that $S(R,q) \sim C(\log R/q)^k \prod_{p|q} \frac{p-g(p)}{p}$ and so λ_q could be replaced by

$$\lambda_q = \mu(q)\frac{\log^k(R/q)}{\log^k R} = \mu(q)\left(1 - \frac{\log q}{\log R}\right)^k$$

- This is correct, and whilst there is some loss in precision in the final conclusion there is one significant advantage, namely that this choice of λ_q can be applied effectively to any sieving question where the dimension is k.

- We expect that $S(R,q) \sim C(\log R/q)^k \prod_{p|q} \frac{p-g(p)}{p}$ and so λ_q could be replaced by

$$\lambda_q = \mu(q)\frac{\log^k(R/q)}{\log^k R} = \mu(q)\left(1 - \frac{\log q}{\log R}\right)^k$$

- This is correct, and whilst there is some loss in precision in the final conclusion there is one significant advantage, namely that this choice of λ_q can be applied effectively to any sieving question where the dimension is k.
- Let $\mathbf{1}_{\mathbb{P}}$ denote the characteristic function of the set of primes $\mathbb{P}$ and write $Z = \prod_{i=1}^{k}(n+h_i)$. Then the idea of Goldston, Pintz and Yıldırım is to construct the expression

$$\sum_{N\le n\le 2N}\left(\sum_{j=1}^{k}\mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{q\le R\\ q|Z(n;\mathbf{h})}}\lambda_q\Bigg)^2$$

- We expect that $S(R,q) \sim C(\log R/q)^k \prod_{p|q} \frac{p-g(p)}{p}$ and so λ_q could be replaced by

$$\lambda_q = \mu(q)\frac{\log^k(R/q)}{\log^k R} = \mu(q)\left(1-\frac{\log q}{\log R}\right)^k$$

- This is correct, and whilst there is some loss in precision in the final conclusion there is one significant advantage, namely that this choice of λ_q can be applied effectively to any sieving question where the dimension is k.
- Let $\mathbf{1}_{\mathbb{P}}$ denote the characteristic function of the set of primes $\mathbb{P}$ and write $Z = \prod_{i=1}^k (n+h_i)$. Then the idea of Goldston, Pintz and Yıldırım is to construct the expression

$$\sum_{N\le n\le 2N}\left(\sum_{j=1}^k \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{q\le R\\ q|Z(n;\mathbf{h})}}\lambda_q\Bigg)^2$$

- If this is positive, then it follows that there are n such that there are at least $\lfloor\rho\rfloor+1$ primes amongst the $n+h_j$.

- $$\sum_{N\le n\le 2N}\Big(\sum_{j=1}^{k}\mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\Big)\Big(\sum_{\substack{q\le R\\ q|Z(n;\mathbf{h})}}\lambda_q\Big)^2$$

- $$\sum_{N \le n \le 2N} \Big(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j) - \rho \Big) \Big(\sum_{\substack{q \le R \\ q \mid Z(n;\mathbf{h})}} \lambda_q \Big)^2$$

- A wrinkle introduced by Goldston, Pintz and Yıldırım is to use a more general λ_q of the form

$$\lambda_q = \mu(q) f\left(\frac{\log q}{\log R} \right)$$

where f is at our disposal.

- $$\sum_{N\le n\le 2N}\left(\sum_{j=1}^{k}\mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{q\le R\\ q|Z(n;\mathbf{h})}}\lambda_q\Bigg)^2$$

- A wrinkle introduced by Goldston, Pintz and Yıldırım is to use a more general λ_q of the form

$$\lambda_q=\mu(q)f\left(\frac{\log q}{\log R}\right)$$

  where f is at our disposal.

- Following Maynard we will use a more sophisticated version of this.

- Let $n + \mathbf{h}$ denote the k–tuple $n + h_1, \ldots, n + h_k$ and let $\mathbf{d}$ denote the k–tuple $d_1, \ldots, d_k$. We generally use the notation that given a k–tuple $\mathbf{d}$ of positive integers d denotes $d_1 \ldots d_k$ and given another one $\mathbf{r}$, then $\mathbf{d}|\mathbf{r}$ means that $d_j|r_j$ for each j. We also use $[\mathbf{d}, \mathbf{e}]$ to denote the k–tuple $\mathrm{lcm}[d_1, e_1], \ldots, \mathrm{lcm}[d_k, e_k]$.

- Let $n + \mathbf{h}$ denote the k–tuple $n + h_1, \ldots, n + h_k$ and let $\mathbf{d}$ denote the k–tuple $d_1, \ldots, d_k$. We generally use the notation that given a k–tuple $\mathbf{d}$ of positive integers d denotes $d_1 \ldots d_k$ and given another one $\mathbf{r}$, then $\mathbf{d}|\mathbf{r}$ means that $d_j|r_j$ for each j. We also use $[\mathbf{d}, \mathbf{e}]$ to denote the k–tuple $\mathrm{lcm}[d_1, e_1], \ldots, \mathrm{lcm}[d_k, e_k]$.
- One wrinkle is to do some initial sieving for small primes so as to simplify some later expressions and s simple way to do this is to restrict our attention to a given residue class a modulo q where

$$q = \prod_{p \le Q} p, \quad Q = \log\log\log N \tag{3}$$

and N is a large integer parameter

- Let $n+\mathbf{h}$ denote the k–tuple $n+h_1,\ldots,n+h_k$ and let $\mathbf{d}$ denote the k–tuple $d_1,\ldots,d_k$. We generally use the notation that given a k–tuple $\mathbf{d}$ of positive integers d denotes $d_1\ldots d_k$ and given another one $\mathbf{r}$, then $\mathbf{d}|\mathbf{r}$ means that $d_j|r_j$ for each j. We also use $[\mathbf{d},\mathbf{e}]$ to denote the k–tuple $\operatorname{lcm}[d_1,e_1],\ldots,\operatorname{lcm}[d_k,e_k]$.
- One wrinkle is to do some initial sieving for small primes so as to simplify some later expressions and s simple way to do this is to restrict our attention to a given residue class a modulo q where

$$q=\prod_{p\le Q}p,\quad Q=\log\log\log N \tag{3}$$

  and N is a large integer parameter
- When $\mathbf{h}$ is admissible we can suppose that there is an a modulo q such that for $1\le j\le k$ we have $(a+h_j,q)=1$.

- Let $n+\mathbf{h}$ denote the k–tuple $n+h_1,\ldots,n+h_k$ and let $\mathbf{d}$ denote the k–tuple $d_1,\ldots,d_k$. We generally use the notation that given a k–tuple $\mathbf{d}$ of positive integers d denotes $d_1\ldots d_k$ and given another one $\mathbf{r}$, then $\mathbf{d}|\mathbf{r}$ means that $d_j|r_j$ for each j. We also use $[\mathbf{d},\mathbf{e}]$ to denote the k–tuple $\operatorname{lcm}[d_1,e_1],\ldots,\operatorname{lcm}[d_k,e_k]$.
- One wrinkle is to do some initial sieving for small primes so as to simplify some later expressions and s simple way to do this is to restrict our attention to a given residue class a modulo q where

$$q=\prod_{p\le Q}p,\quad Q=\log\log\log N \tag{3}$$

  and N is a large integer parameter
- When $\mathbf{h}$ is admissible we can suppose that there is an a modulo q such that for $1\le j\le k$ we have $(a+h_j,q)=1$.
- To see that this holds observe that it holds for each prime divisor of q and then apply the Chinese Remainder Theorem.

- $q = \prod_{p \le Q} p, \quad Q = \log\log\log N$ and N is a large integer parameter, and when $\mathbf{h}$ is admissible there is an a modulo q such that for $1 \le j \le k$ we have $(a + h_j, q) = 1$.

- $q = \prod_{p \le Q} p, \quad Q = \log\log\log N$ and N is a large integer parameter, and when $\mathbf{h}$ is admissible there is an a modulo q such that for $1 \le j \le k$ we have $(a + h_j, q) = 1$.
- The immediate effect of this can be seen *via* the heuristic argument based on the Hardy-Littlewood method which we saw earlier. If one supposes in addition that $n \equiv a$ modulo q, then the singular series takes the shape

$$\mathfrak{S}(\mathbf{h}) = \prod_{p > Q} \left(1 - \frac{k}{p}\right)\left(1 - \frac{1}{p}\right)^{-k} \sim 1$$

for large N.

- Thus Maynard was lead to consider

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \left(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2.$$

- Thus Maynard was lead to consider

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \left(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2.$$

- In the first instance we might presume to take

$$\lambda(\mathbf{d})=\mu(d)g(\mathbf{d})$$

for some suitable g.

- Thus Maynard was lead to consider

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \left(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{\mathbf{d}\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2.$$

- In the first instance we might presume to take

$$\lambda(\mathbf{d})=\mu(d)g(\mathbf{d})$$

for some suitable g.

- However when diagonalising the quadratic forms in the λ and trying to keep control of the support for the $\mathbf{d}$ it transpires that it is natural to suppose that if d is squarefree and $(d,q)=1$, then

$$\lambda(\mathbf{d})=\mu(d)d\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}\\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R},\dots,\frac{\log r_k}{\log R}\right). \quad (4)$$

- If d is squarefree and $(d, q) = 1$, then

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r} \\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right).$$

- If d is squarefree and $(d,q)=1$, then

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r} \\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R}\right).$$

- It is further supposed that

$$\mathrm{supp} f = \mathcal{R} = \{\mathbf{x} \in [0,1]^k : x_1 + \cdots x_k \le 1\}.$$

- If d is squarefree and $(d,q)=1$, then

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r} \\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right).$$

- It is further supposed that

$$\mathrm{supp} f = \mathcal{R} = \{\mathbf{x} \in [0,1]^k : x_1 + \cdots x_k \le 1\}.$$

- This is equivalent to $r_1 \dots r_k \le R$, which gives natural control of the variables.

- To repeat, we consider

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \left(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2. \qquad (5)$$

and if d is squarefree and $(d,q)=1$, then take

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}\\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R},\dots,\frac{\log r_k}{\log R}\right). \qquad (6)$$

We further suppose that $\mathcal{F}$ is a class of "smooth" f satisfying

$$\mathrm{supp} f = \mathcal{R} = \{\mathbf{x}\in[0,1]^k : x_1+\cdots x_k \le 1\}. \qquad (7)$$

- To repeat, we consider

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \left(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2. \quad (5)$$

and if d is squarefree and $(d,q)=1$, then take

$$\lambda(\mathbf{d})=\mu(d)d\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}\\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R},\dots,\frac{\log r_k}{\log R}\right). \quad (6)$$

We further suppose that $\mathcal{F}$ is a class of "smooth" f satisfying

$$\mathrm{supp} f=\mathcal{R}=\{\mathbf{x}\in[0,1]^k : x_1+\cdots x_k\le 1\}. \quad (7)$$

- There are two major tasks to be undertaken. The first is to obtain a good approximation to (5) with (6) for a wide class of f in $\mathcal{F}$.

- $$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \left(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2.$$

- $$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \Big(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\Big)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$
- This means good approximations $S^*(f)$ and $T^*(f)$ to

$$S(f)=\sum_{j=1}^{k} S_j(f)$$

where

$$S_j(f)=S_j=\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$

$$T(f)=T=\sum_{\substack{N\le n\le 2N\\ n\equiv a \pmod q}} \Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$

- Thus we obtain

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \left(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\Bigg(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2 \sim S^*(f)-\rho T^*(f).$$

- Thus we obtain

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \left(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\right)\left(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\right)^2 \sim S^*(f)-\rho T^*(f).$$

- We want this to be positive, but with ρ as large as possible.

- Thus we obtain

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \Big(\sum_{j=1}^{k} \mathbf{1}_{\mathbb{P}}(n+h_j)-\rho\Big)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2 \sim S^*(f)-\rho T^*(f).$$

- We want this to be positive, but with ρ as large as possible.
- This means that the second task is to choose f to maximise the ratio
$$\frac{S^*(f)}{T^*(f)}.$$

- To approximate

$$S_j(f) = S_j = \sum_{\substack{N<n\le 2N \\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Bigg(\sum_{\substack{d\le R \\ \mathbf{d}|n+\mathbf{h} \\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2.$$

it is natural to use the Bombieri–Vinogradov theorem.

- To approximate

$$S_j(f) = S_j = \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$

  it is natural to use the Bombieri–Vinogradov theorem.
- We define the *level* θ of distribution for the prime numbers to be the assumption that for every sufficiently small positive δ and every $A>0$ we have

$$\sum_{m\le x^{\theta-\delta}} \max_{(a,m)=1} \sup_{y\le x} \left|\pi(y;m,a) - \frac{\mathrm{li}(y)}{\phi(m)}\right| \ll_{\delta,A} x(\log x)^{-A}.$$

- To approximate

$$S_j(f) = S_j = \sum_{\substack{N<n\le 2N \\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R \\ \mathbf{d}|n+\mathbf{h} \\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$

  it is natural to use the Bombieri–Vinogradov theorem.
- We define the *level* θ of distribution for the prime numbers to be the assumption that for every sufficiently small positive δ and every $A>0$ we have

$$\sum_{m\le x^{\theta-\delta}} \max_{(a,m)=1} \sup_{y\le x} \left|\pi(y;m,a) - \frac{\mathrm{li}(y)}{\phi(m)}\right| \ll_{\delta,A} x(\log x)^{-A}.$$

- The Bombieri–Vinogradov theorem tells us that $\theta = \frac{1}{2}$ is permissible.

- To approximate
$$S_j(f) = S_j = \sum_{\substack{N<n\le 2N \\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R \\ \mathbf{d}|n+\mathbf{h} \\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$
it is natural to use the Bombieri–Vinogradov theorem.
- We define the *level* θ of distribution for the prime numbers to be the assumption that for every sufficiently small positive δ and every $A>0$ we have
$$\sum_{m\le x^{\theta-\delta}} \max_{(a,m)=1} \sup_{y\le x} \left|\pi(y;m,a) - \frac{\mathrm{li}(y)}{\phi(m)}\right| \ll_{\delta,A} x(\log x)^{-A}.$$
- The Bombieri–Vinogradov theorem tells us that $\theta = \frac{1}{2}$ is permissible.
- However it is useful to be able to see any consequence of any $\theta > 1/2$, especially the Elliott–Halberstam conjecture $(\theta = 1)$.

- To approximate

$$S_j(f) = S_j = \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$

it is natural to use the Bombieri–Vinogradov theorem.

- We define the *level* θ of distribution for the prime numbers to be the assumption that for every sufficiently small positive δ and every $A>0$ we have

$$\sum_{m\le x^{\theta-\delta}} \max_{(a,m)=1} \sup_{y\le x} \left| \pi(y;m,a) - \frac{\mathrm{li}(y)}{\phi(m)} \right| \ll_{\delta,A} x(\log x)^{-A}.$$

- The Bombieri–Vinogradov theorem tells us that $\theta=\frac{1}{2}$ is permissible.
- However it is useful to be able to see any consequence of any $\theta>1/2$, especially the Elliott–Halberstam conjecture $(\theta=1)$.
- Moreover we will see that any $\theta>0$ is good enough for bounded gaps.

- Let me remind you of the way in which the Selberg sieve worked.

- Let me remind you of the way in which the Selberg sieve worked.
- For some squarefree P and non-negative $a(m)$ we are interested in

$$\sum_{(m,P)=1} a(m) \le \sum_{m} a(m) \left(\sum_{\substack{d\in\mathcal{D}\\ d|m}} \lambda(d) \right)^2 .$$

- Let me remind you of the way in which the Selberg sieve worked.
- For some squarefree P and non-negative $a(m)$ we are interested in

$$\sum_{(m,P)=1} a(m) \le \sum_m a(m) \Bigg(\sum_{\substack{d\in\mathcal{D} \\ d|m}} \lambda(d) \Bigg)^2 .$$

- For a divisor closed subset of the divisors of P we rewrote this as

$$\sum_{d\in\mathcal{D}} \sum_{e\in\mathcal{D}} \lambda(d)\lambda(e) \sum_{\substack{m \\ [d,e]|m}} a(m).$$

- Let me remind you of the way in which the Selberg sieve worked.
- For some squarefree P and non-negative $a(m)$ we are interested in

$$\sum_{(m,P)=1} a(m) \leq \sum_{m} a(m) \left(\sum_{\substack{d\in\mathcal{D} \\ d|m}} \lambda(d) \right)^2 .$$

- For a divisor closed subset of the divisors of P we rewrote this as

$$\sum_{d\in\mathcal{D}} \sum_{e\in\mathcal{D}} \lambda(d)\lambda(e) \sum_{\substack{m \\ [d,e]|m}} a(m).$$

- We also supposed that for $d \in \mathcal{D}$ and some $\rho \in \mathcal{M}$ we have

$$\sum_{\substack{m \\ [d,e]|m}} a(m) = X\rho(d) + R_d.$$

- Let me remind you of the way in which the Selberg sieve worked.
- For some squarefree P and non-negative $a(m)$ we are interested in

$$\sum_{(m,P)=1} a(m) \le \sum_m a(m) \left(\sum_{\substack{d\in\mathcal{D}\\ d|m}} \lambda(d) \right)^2 .$$

- For a divisor closed subset of the divisors of P we rewrote this as

$$\sum_{d\in\mathcal{D}} \sum_{e\in\mathcal{D}} \lambda(d)\lambda(e) \sum_{\substack{m\\ [d,e]|m}} a(m).$$

- We also supposed that for $d \in \mathcal{D}$ and some $\rho \in \mathcal{M}$ we have

$$\sum_{\substack{m\\ [d,e]|m}} a(m) = X\rho(d) + R_d.$$

- I changed from f to ρ here for notational convenience.

- Then the main term becomes

$$X \sum_{d \in \mathcal{D}} \sum_{e \in \mathcal{D}} \rho([d, e]) \lambda(d) \lambda(e).$$

- Then the main term becomes

$$X\sum_{d\in\mathcal{D}}\sum_{e\in\mathcal{D}}\rho([d,e])\lambda(d)\lambda(e).$$

- We were able to diagonalise this as

$$X\sum_{l}\left(\prod_{p|l}\frac{1-\rho(p)}{\rho(p)}\right)\left(\sum_{\substack{r\\ l|r}}\rho(r)\lambda(r)\right)^2.$$

- Then the main term becomes

$$X \sum_{d\in\mathcal{D}} \sum_{e\in\mathcal{D}} \rho([d,e])\lambda(d)\lambda(e).$$

- We were able to diagonalise this as

$$X \sum_{l} \left(\prod_{p|l} \frac{1-\rho(p)}{\rho(p)} \right) \left(\sum_{\substack{r \\ l|r}} \rho(r)\lambda(r) \right)^2 .$$

- We then applied the invertible mapping

$$\omega(l) = \sum_{\substack{r \\ l|r}} \rho(r)\lambda(r).$$

- Then the main term becomes

$$X\sum_{d\in\mathcal{D}}\sum_{e\in\mathcal{D}}\rho([d,e])\lambda(d)\lambda(e).$$

- We were able to diagonalise this as

$$X\sum_{l}\left(\prod_{p|l}\frac{1-\rho(p)}{\rho(p)}\right)\left(\sum_{\substack{r\\ l|r}}\rho(r)\lambda(r)\right)^2.$$

- We then applied the invertible mapping

$$\omega(l)=\sum_{\substack{r\\ l|r}}\rho(r)\lambda(r).$$

- Note that at this stage λ can be pretty arbitrary, and certainly does not have to be optimal.

- Then the main term becomes

$$X \sum_{d \in \mathcal{D}} \sum_{e \in \mathcal{D}} \rho([d,e]) \lambda(d) \lambda(e).$$

- We were able to diagonalise this as

$$X \sum_{l} \left(\prod_{p|l} \frac{1-\rho(p)}{\rho(p)} \right) \left(\sum_{\substack{r \\ l|r}} \rho(r) \lambda(r) \right)^2 .$$

- We then applied the invertible mapping

$$\omega(l) = \sum_{\substack{r \\ l|r}} \rho(r) \lambda(r).$$

- Note that at this stage λ can be pretty arbitrary, and certainly does not have to be optimal.
- We want to carry this out for $S_j(f)$ and $T(f)$. There are some differences of detail, but not of principle.

- $S_j(f) = \sum_{\substack{N<n\le 2N \\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Bigg(\sum_{\substack{d\le R \\ \mathbf{d}|n+\mathbf{h} \\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2.$

- $$S_j(f) = \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$
- When we looked at k dimensional sieves previously we would have considered $d|(n+h_1)\dots(n+h_k)$. Now we are being more prescriptive in that we assume some control over $(d, n+h_k) = d_j$. Thus we suppose that $\mathbf{d}|n+\mathbf{h}$.

- $$S_j(f) = \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$
- When we looked at k dimensional sieves previously we would have considered $d|(n+h_1)\dots(n+h_k)$. Now we are being more prescriptive in that we assume some control over $(d, n+h_k) = d_j$. Thus we suppose that $\mathbf{d}|n+\mathbf{h}$.
- I believe this was done to give better control over the d_i in the later analysis, but I do not think it loses anything of consequence.

- $$S_j(f) = \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$
- When we looked at k dimensional sieves previously we would have considered $d|(n+h_1)\dots(n+h_k)$. Now we are being more prescriptive in that we assume some control over $(d,n+h_k)=d_j$. Thus we suppose that $\mathbf{d}|n+\mathbf{h}$.
- I believe this was done to give better control over the d_i in the later analysis, but I do not think it loses anything of consequence.
- Since we have to deal with $T(f)$ as well, we are pretty much forced to choose $\lambda(\mathbf{d})$ corresponding to a k-dimensional sieve, although in $S_j(f)$ since one of the variables is prescribed to be prime we would only need a $k-1$-dimensional sieve.

- We have

$$S_j(f) = \sum_{\substack{N<n\le 2N \\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R \\ \mathbf{d}|n+\mathbf{h} \\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2$$

$$= \sum_{\substack{N<n\le 2N \\ n\equiv a \pmod q \\ n+h_j\in\mathbb{P}}} \Bigg(\sum_{\substack{d\le R \\ \mathbf{d}|n+\mathbf{h} \\ d_j=1,\,(d,q)=1}} \lambda(\mathbf{d})\Bigg)^2.$$

- We have

$$S_j(f) = \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2$$

$$= \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q\\ n+h_j\in\mathbb{P}}} \Bigg(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ d_j=1,\,(d,q)=1}} \lambda(\mathbf{d})\Bigg)^2.$$

- Thus although we gain a $(\log N)^{-1}$ by using Bombieri-Vinogradov, we do not get anything small for the sum over d_1 so we lose back something like a $\log R$.

- We have

$$S_j(f) = \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2$$

$$= \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q\\ n+h_j\in\mathbb{P}}} \Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ d_j=1,\,(d,q)=1}} \lambda(\mathbf{d})\Big)^2.$$

- Thus although we gain a $(\log N)^{-1}$ by using Bombieri-Vinogradov, we do not get anything small for the sum over d_1 so we lose back something like a $\log R$.
- On the other hand, since the prime factors p of the d satisfy $p > Q = \log\log\log N$, any factors like

$$\prod_{p|d} \frac{p^k - kp^{k-1}}{(p-1)^k}$$

are going to be close to 1, at least on average and so won't differ in any important way from the $k-1$ version.

- Recall that we plan to take

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r} \\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R}\right)$$

with

$$\operatorname{supp} f = \mathcal{R} = \{x \in [0,1]^k : x_1 + \cdots + x_k \le 1\}.$$

- Recall that we plan to take

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r} \\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right)$$

with

$$\operatorname{supp} f = \mathcal{R} = \{x \in [0,1]^k : x_1 + \cdots + x_k \le 1\}.$$

- In the 1-dimensional sieve we had $f = 1$, and showing that

$$\mu(d)d \sum_{\substack{r \le R \\ d|r}} \frac{\mu(r)^2}{\phi(r)} = \frac{\mu(d)d}{\phi(d)} \sum_{\substack{s \le R/d \\ (s,d)=1}} \frac{\mu(s)^2}{\phi(s)} \sim \mu(d) \log \frac{R}{d}$$

was relatively easy.

- Recall that we plan to take

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r} \\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right)$$

with

$$\operatorname{supp} f = \mathcal{R} = \{x \in [0,1]^k : x_1 + \cdots + x_k \le 1\}.$$

- In the 1-dimensional sieve we had $f = 1$, and showing that

$$\mu(d)d \sum_{\substack{r \le R \\ d|r}} \frac{\mu(r)^2}{\phi(r)} = \frac{\mu(d)d}{\phi(d)} \sum_{\substack{s \le R/d \\ (s,d)=1}} \frac{\mu(s)^2}{\phi(s)} \sim \mu(d) \log \frac{R}{d}$$

was relatively easy.
- Now we need to entertain the possibility that f will be more complicated and we need to apply partial summation, maybe in more than one dimension.

- Recall that we plan to take

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r} \\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R}\right)$$

with

$$\operatorname{supp} f = \mathcal{R} = \{x \in [0,1]^k : x_1 + \cdots + x_k \le 1\}.$$

- In the 1-dimensional sieve we had $f = 1$, and showing that

$$\mu(d)d \sum_{\substack{r \le R \\ d|r}} \frac{\mu(r)^2}{\phi(r)} = \frac{\mu(d)d}{\phi(d)} \sum_{\substack{s \le R/d \\ (s,d)=1}} \frac{\mu(s)^2}{\phi(s)} \sim \mu(d) \log \frac{R}{d}$$

was relatively easy.
- Now we need to entertain the possibility that f will be more complicated and we need to apply partial summation, maybe in more than one dimension.
- We need to set up some notation.

- Let $\mathcal{R}_j$ denote the set of k–tuples $t_1, \ldots . t_{j-1}, t_{j+1}, \ldots, t_k$ with $\mathbf{t} \in \mathcal{R}$ for some t_j.

- Let $\mathcal{R}_j$ denote the set of k–tuples $t_1, \dots .t_{j-1}, t_{j+1}, \dots, t_k$ with $\mathbf{t} \in \mathcal{R}$ for some t_j.
- We define $\mathcal{F}$ to be the class of functions f, not identically 0, defined on $\mathcal{R}$ such that for each j, given $\mathbf{t}^* = t_1, \dots, t_{j-1}, t_{j+1}, \dots, t_k$ with $t_i \geq 0$ and $t_1 + \dots + t_{j-1} + t_{j+1} + \dots + t_k \leq 1$ the function $f_j(t_j) = f(\mathbf{t})$ is absolutely continuous on $[0, 1 - t_1 - \dots - t_{j-1} - t_{j+1} - \dots - t_k]$.

- Let $\mathcal{R}_j$ denote the set of k–tuples $t_1, \ldots. t_{j-1}, t_{j+1}, \ldots, t_k$ with $\mathbf{t} \in \mathcal{R}$ for some t_j.
- We define $\mathcal{F}$ to be the class of functions f, not identically 0, defined on $\mathcal{R}$ such that for each j, given $\mathbf{t}^* = t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_k$ with $t_i \geq 0$ and $t_1 + \cdots + t_{j-1} + t_{j+1} + \cdots + t_k \leq 1$ the function $f_j(t_j) = f(\mathbf{t})$ is absolutely continuous on $[0, 1 - t_1 - \cdots - t_{j-1} - t_{j+1} - \cdots - t_k]$.
- Given an $f \in \mathcal{F}$ it is useful first to extend its definition to $[0,1]^k$ by taking it to be 0 outside $\mathcal{R}$ and then to define a suitable metric.

$$F = \sup_{\mathbf{t} \in \mathcal{R}} |f(\mathbf{t})| + \sum_{j=1}^{k} \sup_{\mathbf{t}^* \in \mathcal{R}_j} \int_0^1 \left| \frac{\partial f}{\partial t_j}(\mathbf{t}) \right| dt_j.$$

Theorem 6 (Maynard)

Let $k \ge 2$. Suppose the primes have level of distribution θ and $N > N_0(\delta)$. Let $R = N^{\frac{\theta}{2}-\delta}$, and Q, q, $\mathcal{R}$ and $f \in \mathcal{F}$ be as above. Assume $\mathbf{h}$ is admissible and that for each j, $(a+h_j, q) = 1$. Let $J = \int_{[0,1]^k} f(\mathbf{t})^2 d\mathbf{t}$,

$$I_j = \int_{[0,1]^{k-1}} \left(\int_0^1 f(\mathbf{t}) dt_j \right)^2 dt_1 \dots dt_{j-1} dt_{j+1} \dots dt_k,$$

$$S(f) = \frac{(1+o(1))\phi(q)^k N (\log R)^{k+1}}{q^{k+1} \log N} \sum_{j=1}^{k} I_j$$

and $T(f) = \dfrac{(1+o(1))\phi(q)^k N (\log R)^k}{q^{k+1}} J$ *as* $N \to \infty$. *In particular* $\dfrac{S(f)}{T(f)} = (1+o(1)) \left(\dfrac{\theta}{2} - \delta \right) \dfrac{\sum_{j=1}^{k} I_j}{J}$.

- The proof is divided into several stages. Fortunately the treatments of $S(f)$ and $T(f)$ are similar.

- The proof is divided into several stages. Fortunately the treatments of $S(f)$ and $T(f)$ are similar.
- Initially we do not assume anything about the $\lambda(\mathbf{d})$ apart from supposing that the $\lambda(\mathbf{d})$ are general real valued functions with support satisfying $d_1 \dots d_k = d \le R$, $(d, q) = 1$ and d squarefree.

- The proof is divided into several stages. Fortunately the treatments of $S(f)$ and $T(f)$ are similar.
- Initially we do not assume anything about the $\lambda(\mathbf{d})$ apart from supposing that the $\lambda(\mathbf{d})$ are general real valued functions with support satisfying $d_1 \ldots d_k = d \le R$, $(d, q) = 1$ and d squarefree.
- Of course, then $(d_i, d_j) = 1$ when $i \ne j$.

- The proof is divided into several stages. Fortunately the treatments of $S(f)$ and $T(f)$ are similar.
- Initially we do not assume anything about the $\lambda(\mathbf{d})$ apart from supposing that the $\lambda(\mathbf{d})$ are general real valued functions with support satisfying $d_1 \dots d_k = d \le R$, $(d, q) = 1$ and d squarefree.
- Of course, then $(d_i, d_j) = 1$ when $i \ne j$.
- We begin with the diagonalisation process, and it is useful to define the multiplicative function $\phi_2(n)$ by $\phi_2(p) = p - 2$ and $\phi_2(p^t) = 0$ when $t \ge 2$.

- The proof is divided into several stages. Fortunately the treatments of $S(f)$ and $T(f)$ are similar.
- Initially we do not assume anything about the $\lambda(\mathbf{d})$ apart from supposing that the $\lambda(\mathbf{d})$ are general real valued functions with support satisfying $d_1 \dots d_k = d \le R$, $(d, q) = 1$ and d squarefree.
- Of course, then $(d_i, d_j) = 1$ when $i \ne j$.
- We begin with the diagonalisation process, and it is useful to define the multiplicative function $\phi_2(n)$ by $\phi_2(p) = p - 2$ and $\phi_2(p^t) = 0$ when $t \ge 2$.
- Then the diagonalisation process can be summarised by the following lemma

Lemma 7

For $j = 1, \ldots, k$ let

$$\kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)},$$

where $\sum^j$ indicates that the summation variable is a k–tuple, say $\mathbf{d}$, which is restricted by $d_j = 1$, and let

$$\kappa(\mathbf{r}) = \mu(r)\phi(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}} \frac{\lambda(\mathbf{d})}{d}.$$

Then $\displaystyle \frac{\mu(d)}{\phi(d)}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$ *and* $\displaystyle \frac{\mu(d)}{d}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}} \frac{\kappa(\mathbf{r})}{\phi(r)}.$

- To summarize

$$\kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)},$$

$$\lambda(\mathbf{d}) = \mu(d)\phi(d) \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)},$$

$$\kappa(\mathbf{r}) = \mu(r)\phi(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}} \frac{\lambda(\mathbf{d})}{d} \text{ and } \lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}} \frac{\kappa(\mathbf{r})}{\phi(r)}.$$

- To summarize

$$\kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)},$$

$$\lambda(\mathbf{d}) = \mu(d)\phi(d) \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)},$$

$$\kappa(\mathbf{r}) = \mu(r)\phi(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}} \frac{\lambda(\mathbf{d})}{d} \text{ and } \lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}} \frac{\kappa(\mathbf{r})}{\phi(r)}.$$

- In the k dimensional case this looks familiar and the $k-1$ dimensional case does not look too bad. However the use of k-tuples $\mathbf{d}$, etc., makes for some complications.

- This is Möbius inversion. Consider $\displaystyle\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$

- This is Möbius inversion. Consider $\displaystyle\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$
- and substitute in the definition of κ_j to obtain

$$\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \mu(r) \sum_{\substack{\mathbf{s}\\ \mathbf{r}|\mathbf{s}}}^{j} \frac{\lambda(\mathbf{s})}{\phi(s)} = \sum_{\mathbf{d}|\mathbf{s}}^{j} \frac{\lambda(\mathbf{s})}{\phi(s)} \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}|\mathbf{s}}} \mu(r)$$

$$= \sum_{\mathbf{t}}^{j} \frac{\lambda(\mathbf{dt})}{\phi(dt)} \mu(d) \sum_{\substack{\mathbf{u}\\ \mathbf{u}|\mathbf{t}}} \mu(u).$$

Note that the $s = dt$ are square free, the t_i are pairwise coprime, and hence the u_i are pairwise coprime and so are the d_i. Also $(t, d) = 1$. Thus the u_i are free to range over a complete set of divisors of t_i.

- This is Möbius inversion. Consider $\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$
- and substitute in the definition of κ_j to obtain

$$\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \mu(r) \sum_{\substack{\mathbf{s}\\ \mathbf{r}|\mathbf{s}}}^{j} \frac{\lambda(\mathbf{s})}{\phi(s)} = \sum_{\mathbf{d}|\mathbf{s}}^{j} \frac{\lambda(\mathbf{s})}{\phi(s)} \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}|\mathbf{s}}} \mu(r)$$
$$= \sum_{\mathbf{t}}^{j} \frac{\lambda(\mathbf{dt})}{\phi(dt)} \mu(d) \sum_{\substack{\mathbf{u}\\ \mathbf{u}|\mathbf{t}}} \mu(u).$$

  Note that the $s = dt$ are square free, the t_i are pairwise coprime, and hence the u_i are pairwise coprime and so are the d_i. Also $(t, d) = 1$. Thus the u_i are free to range over a complete set of divisors of t_i.
- Also $s_j = 1$, so $d_j = t_j = 1$.

- This is Möbius inversion. Consider $\displaystyle\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$
- and substitute in the definition of κ_j to obtain

$$\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \mu(r) \sum_{\substack{\mathbf{s}\\ \mathbf{r}|\mathbf{s}}}^{j} \frac{\lambda(\mathbf{s})}{\phi(s)} = \sum_{\mathbf{d}|\mathbf{s}}^{j} \frac{\lambda(\mathbf{s})}{\phi(s)} \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}|\mathbf{s}}} \mu(r)$$

$$= \sum_{\mathbf{t}}^{j} \frac{\lambda(\mathbf{dt})}{\phi(dt)} \mu(d) \sum_{\substack{\mathbf{u}\\ \mathbf{u}|\mathbf{t}}} \mu(u).$$

Note that the $s = dt$ are square free, the t_i are pairwise coprime, and hence the u_i are pairwise coprime and so are the d_i. Also $(t, d) = 1$. Thus the u_i are free to range over a complete set of divisors of t_i.

- Also $s_j = 1$, so $d_j = t_j = 1$.
- The sum over u_i is 0 unless $t_i = 1$. Thus it all collapses down to $\dfrac{\mu(d)}{\phi(d)}\lambda(\mathbf{d})$.

- This is Möbius inversion. Consider $\displaystyle\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$
- and substitute in the definition of κ_j to obtain

$$\sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \mu(r) \sum_{\substack{\mathbf{s}\\ \mathbf{r}|\mathbf{s}}}^{j} \frac{\lambda(\mathbf{s})}{\phi(s)} = \sum_{\mathbf{d}|\mathbf{s}}^{j} \frac{\lambda(\mathbf{s})}{\phi(s)} \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}|\mathbf{s}}} \mu(r)$$

$$= \sum_{\mathbf{t}}^{j} \frac{\lambda(\mathbf{dt})}{\phi(dt)} \mu(d) \sum_{\substack{\mathbf{u}\\ \mathbf{u}|\mathbf{t}}} \mu(u).$$

  Note that the $s = dt$ are square free, the t_i are pairwise coprime, and hence the u_i are pairwise coprime and so are the d_i. Also $(t, d) = 1$. Thus the u_i are free to range over a complete set of divisors of t_i.
- Also $s_j = 1$, so $d_j = t_j = 1$.
- The sum over u_i is 0 unless $t_i = 1$. Thus it all collapses down to $\dfrac{\mu(d)}{\phi(d)}\lambda(\mathbf{d})$.
- The other inversion formula follows in the same way.

- The core of the proof is the following lemma.

Lemma 8

Let

$$K_j = \max_{\mathbf{r}} |\kappa_j(\mathbf{r})|, \quad K = \max_{\mathbf{r}} |\kappa(\mathbf{r})|.$$

Then

$$S_j(f) = \frac{N}{\phi(q)\log N} \sum_{\mathbf{r}}^{j} \frac{\kappa_j(\mathbf{r})^2}{\phi_2(r)} + O\left(\frac{K_j^2 \phi(q)^{k-2} N(\log R)^{k-2}}{q^{k-1}Q} \right)$$

and

$$T(f) = \frac{N}{q} \sum_{\mathbf{r}} \frac{\kappa(\mathbf{r})^2}{\phi(r)} + O\left(\frac{K^2 N(\log R)^k}{qQ} \right).$$

- Let $K_j = \max_{\mathbf{r}} |\kappa_j(\mathbf{r})|, \quad K = \max_{\mathbf{r}} |\kappa(\mathbf{r})|$. Then

$$S_j(f) = \frac{N}{\phi(q)\log N} \sum_{\mathbf{r}}^{j} \frac{\kappa_j(\mathbf{r})^2}{\phi_2(r)} + O\left(\frac{K_j^2 \phi(q)^{k-2} N (\log R)^{k-2}}{q^{k-1} Q} \right)$$

and $T(f) = \frac{N}{q} \sum_{\mathbf{r}} \frac{\kappa(\mathbf{r})^2}{\phi(r)} + O\left(\frac{K^2 N (\log R)^k}{qQ} \right)$.

- Let $K_j = \max_{\mathbf{r}} |\kappa_j(\mathbf{r})|, \quad K = \max_{\mathbf{r}} |\kappa(\mathbf{r})|$. Then

$$S_j(f) = \frac{N}{\phi(q)\log N} \sum_{\mathbf{r}}^{j} \frac{\kappa_j(\mathbf{r})^2}{\phi_2(r)} + O\left(\frac{K_j^2 \phi(q)^{k-2} N (\log R)^{k-2}}{q^{k-1} Q} \right)$$

and $T(f) = \frac{N}{q} \sum_{\mathbf{r}} \frac{\kappa(\mathbf{r})^2}{\phi(r)} + O\left(\frac{K^2 N (\log R)^k}{qQ} \right)$.

- If $\kappa(r)$ were normalised so that $\kappa(r) \approx (\log R)^{-k}$, then we would have

$$\sum_{\mathbf{r}} \frac{\kappa(\mathbf{r})^2}{\phi(r)} \approx (\log R)^{-2k} \sum_{r_1 \dots r_k \le R} \frac{\mu(r_1 \dots r_k)^2}{\phi(r_1 \dots r_k)} \approx (\log R)^{-k}.$$

- Let $K_j = \max_{\mathbf{r}} |\kappa_j(\mathbf{r})|, \quad K = \max_{\mathbf{r}} |\kappa(\mathbf{r})|$. Then

$$S_j(f) = \frac{N}{\phi(q)\log N}\sum_{\mathbf{r}}^{j} \frac{\kappa_j(\mathbf{r})^2}{\phi_2(r)} + O\left(\frac{K_j^2\phi(q)^{k-2}N(\log R)^{k-2}}{q^{k-1}Q}\right)$$

and $T(f) = \frac{N}{q}\sum_{\mathbf{r}}\frac{\kappa(\mathbf{r})^2}{\phi(r)} + O\left(\frac{K^2N(\log R)^k}{qQ}\right)$.

- If $\kappa(r)$ were normalised so that $\kappa(r) \approx (\log R)^{-k}$, then we would have

$$\sum_{\mathbf{r}}\frac{\kappa(\mathbf{r})^2}{\phi(r)} \approx (\log R)^{-2k}\sum_{r_1\dots r_k\le R}\frac{\mu(r_1\dots r_k)^2}{\phi(r_1\dots r_k)} \approx (\log R)^{-k}.$$

- Likewise $\sum_{\mathbf{r}}^{j}\frac{\kappa_j(\mathbf{r})^2}{\phi_2(r)} \approx (\log R)^{1-k}$. So we are in the right ballpark!

- Consider first

$$S_j(f) = \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Big(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Big)^2$$

. We need to insert the information about distribution into residue classes and in the main term replace $\lambda(\mathbf{d})$ by $\kappa_j(\mathbf{d})$.

- Consider first

$$S_j(f) = \sum_{\substack{N<n\le 2N\\ n\equiv a \pmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j)\Bigg(\sum_{\substack{d\le R\\ \mathbf{d}|n+\mathbf{h}\\ (d,q)=1}} \lambda(\mathbf{d})\Bigg)^2$$

. We need to insert the information about distribution into residue classes and in the main term replace $\lambda(\mathbf{d})$ by $\kappa_j(\mathbf{d})$.

- Squaring out we obtain

$$S_j(f) = \sum_{\substack{\mathbf{d},\mathbf{e}\\ d_j=e_j=1}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N<n\le 2N\\ [\mathbf{d},\mathbf{e}]|n+\mathbf{h}\\ n\equiv a \bmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j).$$

- Squaring out we obtain

$$S_j(f) = \sum_{\substack{\mathbf{d},\mathbf{e} \\ d_j=e_j=1}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N<n\le 2N \\ [\mathbf{d},\mathbf{e}]|n+\mathbf{h} \\ n\equiv a \bmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j).$$

- Squaring out we obtain

$$S_j(f) = \sum_{\substack{\mathbf{d},\mathbf{e} \\ d_j=e_j=1}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N<n\le 2N \\ [\mathbf{d},\mathbf{e}]|n+\mathbf{h} \\ n\equiv a \bmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j).$$

- We recall that for $\lambda(\mathbf{d}) \neq 0$ we have d squarefree and $(d,q)=1$. Therefore $(d_u,d_v)=1$ when $u\neq v$. Likewise for $\mathbf{e}$. Also if $p|n+h_u$ and $p|n+h_v$, then $p|h_v-h_u$ and this is impossible since $p>\log\log\log N > \max|h_v-h_u|$. Thus, when $u\neq v$, $([d_u,e_u],[d_v,e_v])=1$, whence $(d_u,e_v)=1$. Since $d_j=e_j=1$ we have $[d_j,e_j]=1$. Hence in the inner sum we are left with the system of congruences $n\equiv -h_i \pmod{[d_i,e_i]}$ $i\neq j$ and $n\equiv a \pmod{q}$.

- Squaring out we obtain

$$S_j(f) = \sum_{\substack{\mathbf{d},\mathbf{e} \\ d_j=e_j=1}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N<n\le 2N \\ [\mathbf{d},\mathbf{e}]|n+\mathbf{h} \\ n\equiv a \bmod q}} \mathbf{1}_{\mathbb{P}}(n+h_j).$$

- We recall that for $\lambda(\mathbf{d}) \ne 0$ we have d squarefree and $(d,q)=1$. Therefore $(d_u,d_v)=1$ when $u\ne v$. Likewise for $\mathbf{e}$. Also if $p|n+h_u$ and $p|n+h_v$, then $p|h_v-h_u$ and this is impossible since $p>\log\log\log N>\max|h_v-h_u|$. Thus, when $u\ne v$, $([d_u,e_u],[d_v,e_v])=1$, whence $(d_u,e_v)=1$. Since $d_j=e_j=1$ we have $[d_j,e_j]=1$. Hence in the inner sum we are left with the system of congruences $n\equiv -h_i \pmod{[d_i,e_i]}$ $i\ne j$ and $n\equiv a \pmod{q}$.
- Then the innermost sum can be rewritten as

$$\sum_{\substack{N+h_j<p\le 2N+h_j \\ p\equiv h_j-h_i \bmod [d_i,e_i]\,(i\ne j) \\ p\equiv a+h_j \bmod q}} 1.$$

- Thus

$$S_j(f) = \sum_{\substack{\mathbf{d},\mathbf{e} \\ d_j=e_j=1}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N+h_j<p\le 2N+h_j \\ p\equiv h_j-h_i \bmod [d_i,e_i]\,(i\ne j) \\ p\equiv a+h_j \bmod q}} 1.$$

- Thus

$$S_j(f) = \sum_{\substack{\mathbf{d},\mathbf{e} \\ d_j=e_j=1}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N+h_j<p\le 2N+h_j \\ p\equiv h_j-h_i \bmod [d_i,e_i]\,(i\ne j) \\ p\equiv a+h_j \bmod q}} 1.$$

- We have $(a+h_j,q)=1$ and $(h_j-h_i,de)=1\,(i\neq j)$.

- Thus

$$S_j(f) = \sum_{\substack{\mathbf{d},\mathbf{e} \\ d_j=e_j=1}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N+h_j<p\le 2N+h_j \\ p\equiv h_j-h_i \bmod [d_i,e_i]\,(i\ne j) \\ p\equiv a+h_j \bmod q}} 1.$$

- We have $(a+h_j,q)=1$ and $(h_j-h_i,de)=1\,(i\ne j)$.
- Let $m = q\prod_{i=1}^{k}[d_i,e_i]$, $X_j = \int_{N+h_j}^{2N+h_j} \frac{dt}{\log t}$ and

$$E = \sum_{\mathbf{d},\mathbf{e}}^{*} |\lambda(\mathbf{d})\lambda(\mathbf{e})| \max_{(b,m)=1} \sup_{x\le 2N+H} \left| \pi(x;m,b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|$$

where $\sum^*$ indicates the restrictions $d_j=e_j=1$ and $(d_u,e_v)=1$ when $u\ne v$, and $H=\max_j h_j$.

- Thus

$$S_j(f) = \sum_{\substack{\mathbf{d},\mathbf{e} \\ d_j=e_j=1}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N+h_j<p\le 2N+h_j \\ p\equiv h_j-h_i \bmod [d_i,e_i]\,(i\ne j) \\ p\equiv a+h_j \bmod q}} 1.$$

- We have $(a+h_j,q)=1$ and $(h_j-h_i,de)=1\,(i\ne j)$.
- Let $m=q\prod_{i=1}^{k}[d_i,e_i]$, $X_j=\int_{N+h_j}^{2N+h_j}\frac{dt}{\log t}$ and

$$E=\sum_{\mathbf{d},\mathbf{e}}^{*}|\lambda(\mathbf{d})\lambda(\mathbf{e})|\max_{(b,m)=1}\sup_{x\le 2N+H}\left|\pi(x;m,b)-\frac{\mathrm{li}(x)}{\phi(m)}\right|$$

where $\sum^*$ indicates the restrictions $d_j=e_j=1$ and $(d_u,e_v)=1$ when $u\ne v$, and $H=\max_j h_j$.
- Then

$$S_j(f)=X_j\sum_{\mathbf{d},\mathbf{e}}^{*}\frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}+O(E).$$

- We need to bound the $\lambda(\mathbf{d})$. Recall that by Lemma 7

$$\frac{\mu(d)}{\phi(d)}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$$

- We need to bound the $\lambda(\mathbf{d})$. Recall that by Lemma 7

$$\frac{\mu(d)}{\phi(d)}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$$

- Hence

$$\begin{aligned}\max_{\mathbf{d}, d_j=1} |\lambda(\mathbf{d})| &\le \max_{\mathbf{d}, d_j=1} \phi(d) \sum_{\substack{\mathbf{r}\in\mathcal{D} \\ \mathbf{d}|\mathbf{r} \\ (d,q)=1}}^{j} \frac{K_j \mu(r)^2}{\phi_2(r)} \\ &= K_j \max_{\mathbf{d}} \frac{\phi(d)}{\phi_2(d)} \sum_{\substack{\mathbf{sd}\in\mathcal{D} \\ (s,dq)=1}}^{j} \frac{\mu(s)^2}{\phi_2(s)} \\ &\le K_j \max_{\mathbf{d}} \frac{\phi(d)}{\phi_2(d)} \prod_{\substack{Q<p\le R \\ p\nmid d}} \left(1+\frac{1}{p-2}\right)^{k-1} \\ &\ll K_j (\log R)^{k-1}.\end{aligned}$$

- We need to bound the $\lambda(\mathbf{d})$. Recall that by Lemma 7

$$\frac{\mu(d)}{\phi(d)}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$$

- Hence

$$\max_{\mathbf{d},d_j=1} |\lambda(\mathbf{d})| \le \max_{\mathbf{d},d_j=1} \phi(d) \sum_{\substack{\mathbf{r}\in\mathcal{D}\\ \mathbf{d}|\mathbf{r}\\ (d,q)=1}}^{j} \frac{K_j\mu(r)^2}{\phi_2(r)}$$

$$= K_j \max_{\mathbf{d}} \frac{\phi(d)}{\phi_2(d)} \sum_{\substack{\mathbf{sd}\in\mathcal{D}\\ (s,dq)=1}}^{j} \frac{\mu(s)^2}{\phi_2(s)}$$

$$\le K_j \max_{\mathbf{d}} \frac{\phi(d)}{\phi_2(d)} \prod_{\substack{Q<p\le R\\ p\nmid d}} \left(1+\frac{1}{p-2}\right)^{k-1}$$

$$\ll K_j(\log R)^{k-1}.$$

- Similarly $\max_{\mathbf{d}} |\lambda(\mathbf{d})| \ll K(\log R)^k$.

- Recall the error

$$E = \sum_{\mathbf{d},\mathbf{e}}{}^{*} |\lambda(\mathbf{d})\lambda(\mathbf{e})| \max_{(b,m)=1} \sup_{x \le 2N+H} \left| \pi(x; m, b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|$$

- Recall the error

$$E = \sum_{\mathbf{d},\mathbf{e}}^{*} |\lambda(\mathbf{d})\lambda(\mathbf{e})| \max_{(b,m)=1} \sup_{x\le 2N+H} \left| \pi(x;m,b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|$$

- Here m/q depends on the $\mathbf{d}$, $\mathbf{e}$. We need to know how many times the same m can arise.

- Now consider the number of ways that the modulus m/q can arise in E.

- Now consider the number of ways that the modulus m/q can arise in E.
- In other words, how many choices of $d_1, \ldots, d_k, e_1, \ldots, e_k$ give rise to m?

- Now consider the number of ways that the modulus m/q can arise in E.
- In other words, how many choices of $d_1, \ldots, d_k, e_1, \ldots, e_k$ give rise to m?
- Since $(d_u, e_v) = 1$ for all $u \neq v$, it follows that m/q, and so m, is squarefree.

- Now consider the number of ways that the modulus m/q can arise in E.
- In other words, how many choices of $d_1, \ldots, d_k, e_1, \ldots, e_k$ give rise to m?
- Since $(d_u, e_v) = 1$ for all $u \neq v$, it follows that m/q, and so m, is squarefree.
- Also as $d_1 \ldots d_k | \prod_{i=1}^{k}[d_i, e_i] = m/q$ and $d_j = 1$ the number of possibilities for $\mathbf{d}$ is at most $d_k(m/q)$, and likewise for $\mathbf{e}$.

- Now consider the number of ways that the modulus m/q can arise in E.
- In other words, how many choices of $d_1, \dots, d_k, e_1, \dots, e_k$ give rise to m?
- Since $(d_u, e_v) = 1$ for all $u \neq v$, it follows that m/q, and so m, is squarefree.
- Also as $d_1 \dots d_k | \prod_{i=1}^k [d_i, e_i] = m/q$ and $d_j = 1$ the number of possibilities for $\mathbf{d}$ is at most $d_k(m/q)$, and likewise for $\mathbf{e}$.
- Thus $E \ll K_j^2 (\log R)^{2k} E'$ where $E' =$

$$\sum_{m \leq qR^2} \mu(m)^2 d_k(m)^2 \max_{(b,m)=1} \sup_{x \leq 2N+H} \left| \pi(x; m, b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|.$$

- We have $E \ll K_j^2(\log R)^{2k}E'$ where $E' =$

$$\sum_{m\le qR^2} \mu(m)^2 d_k(m)^2 \max_{(b,m)=1} \sup_{x\le 2N+H} \left|\pi(x;m,b) - \frac{\mathrm{li}(x)}{\phi(m)}\right|.$$

- We have $E \ll K_j^2 (\log R)^{2k} E'$ where $E' =$

$$\sum_{m \le qR^2} \mu(m)^2 d_k(m)^2 \max_{(b,m)=1} \sup_{x \le 2N+H} \left| \pi(x; m, b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|.$$

- The extra factor $d_k(m)^2$ is a minor irritant in the application of the Bombieri-Vinogradov theorem or equivalents and we deal with it by applying the Cauchy-Schwarz inequality.

- We have $E \ll K_j^2 (\log R)^{2k} E'$ where $E' =$

$$\sum_{m \le qR^2} \mu(m)^2 d_k(m)^2 \max_{(b,m)=1} \sup_{x \le 2N+H} \left| \pi(x;m,b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|.$$

- The extra factor $d_k(m)^2$ is a minor irritant in the application of the Bombieri-Vinogradov theorem or equivalents and we deal with it by applying the Cauchy-Schwarz inequality.
- Thus $E' \le (E_1 E_2)^{1/2}$ where

$$E_1 = \sum_{m \le qR^2} \max_{(b,m)=1} \sup_{x \le 2N+H} \left| \pi(x;m,b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|$$

and $E_2 =$

$$\sum_{m \le qR^2} \mu(m)^2 d_k(m)^4 \max_{(b,m)=1} \sup_{x \le 2N+H} \left| \pi(x;m,b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|.$$

- $E' \le (E_1E_2)^{1/2}$ and $E_2 =$

$$\sum_{m\le qR^2} \mu(m)^2 d_k(m)^4 \max_{(b,m)=1} \sup_{x\le 2N+H} \left| \pi(x;m,b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|.$$

- $E' \le (E_1 E_2)^{1/2}$ and $E_2 =$

$$\sum_{m \le qR^2} \mu(m)^2 d_k(m)^4 \max_{(b,m)=1} \sup_{x \le 2N+H} \left| \pi(x; m, b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|.$$

- Crudely we have

$$\max_{(b,m)=1} \sup_{x \le 2N+H} \left| \pi(x; m, b) - \frac{\mathrm{li}(x)}{\phi(m)} \right| \ll \frac{N}{m}.$$

- $E' \le (E_1 E_2)^{1/2}$ and $E_2 =$

$$\sum_{m\le qR^2} \mu(m)^2 d_k(m)^4 \max_{(b,m)=1} \sup_{x\le 2N+H} \left| \pi(x;m,b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|.$$

- Crudely we have

$$\max_{(b,m)=1} \sup_{x\le 2N+H} \left| \pi(x;m,b) - \frac{\mathrm{li}(x)}{\phi(m)} \right| \ll \frac{N}{m}.$$

- Thus $E_2 \ll N \sum_{m\le qR^2} \mu(m)^2 d_k(m)^4 m^{-1}$

$$\ll N \prod_{p\le qR^2} \left(1 + \frac{k^4}{p}\right) \ll N(\log N)^{k^4}$$

- $E' \le (E_1 E_2)^{1/2}$ and $E_2 =$

$$\sum_{m \le qR^2} \mu(m)^2 d_k(m)^4 \max_{(b,m)=1} \sup_{x \le 2N+H} \left| \pi(x; m, b) - \frac{\mathrm{li}(x)}{\phi(m)} \right|.$$

- Crudely we have

$$\max_{(b,m)=1} \sup_{x \le 2N+H} \left| \pi(x; m, b) - \frac{\mathrm{li}(x)}{\phi(m)} \right| \ll \frac{N}{m}.$$

- Thus $E_2 \ll N \sum_{m \le qR^2} \mu(m)^2 d_k(m)^4 m^{-1}$

$$\ll N \prod_{p \le qR^2} \left(1 + \frac{k^4}{p} \right) \ll N(\log N)^{k^4}$$

- Hence, by our assumption that the level of distribution is θ and the choice of $R = N^{\frac{\theta}{2} - \delta}$ we have $E \ll K_j^2 N(\log N)^{-A}$.

- Thus we have established that

$$S_j(f) = X_j \sum_{\mathbf{d},\mathbf{e}}{}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} + O(K_j^2 N(\log N)^{-A})$$

where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $X_j = \int_{N+h_j}^{2N+h_j} \frac{dt}{\log t}$, $Q = \log\log\log N$, $q = \prod_{p\le Q} p$, $H = \max_j h_j$ and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \neq v$.

- Thus we have established that

$$S_j(f) = X_j \sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} + O(K_j^2 N (\log N)^{-A})$$

where $m = q \prod_{i=1}^{k} [d_i, e_i]$, $X_j = \displaystyle\int_{N+h_j}^{2N+h_j} \frac{dt}{\log t}$,
$Q = \log\log\log N$, $q = \prod_{p \le Q} p$, $H = \max_j h_j$ and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$.
- It remains to deal with the main term $S_j(f)$.

- Thus we have established that
$$S_j(f) = X_j \sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} + O(K_j^2 N(\log N)^{-A})$$
where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $X_j = \displaystyle\int_{N+h_j}^{2N+h_j} \frac{dt}{\log t}$,
$Q = \log\log\log N$, $q = \prod_{p\le Q} p$, $H = \max_j h_j$ and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$.
- It remains to deal with the main term $S_j(f)$.
- We want to diagonalise it.

- Thus we have established that

$$S_j(f) = X_j \sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} + O(K_j^2 N(\log N)^{-A})$$

where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $X_j = \int_{N+h_j}^{2N+h_j} \frac{dt}{\log t}$,
$Q = \log\log\log N$, $q = \prod_{p\le Q} p$, $H = \max_j h_j$ and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$.

- It remains to deal with the main term $S_j(f)$.
- We want to diagonalise it.
- If we had $k = j = 2$, then the sum would just be

$$\sum_{d_1,e_1} \frac{\lambda(d_1)\lambda(e_1)}{\phi([d_1, e_1])}$$

and we could imitate the Selberg sieve.

- Thus we have established that

$$S_j(f) = X_j \sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} + O(K_j^2 N(\log N)^{-A})$$

where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $X_j = \int_{N+h_j}^{2N+h_j} \frac{dt}{\log t}$, $Q = \log\log\log N$, $q = \prod_{p\le Q} p$, $H = \max_j h_j$ and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$.

- It remains to deal with the main term $S_j(f)$.
- We want to diagonalise it.
- If we had $k = j = 2$, then the sum would just be

$$\sum_{d_1,e_1} \frac{\lambda(d_1)\lambda(e_1)}{\phi([d_1, e_1])}$$

and we could imitate the Selberg sieve.

- With this in mind, it is desirable to rid ourselves of the condition that $(d_u, e_v) = 1$ when $u \ne v$.

- We have to deal with $\displaystyle\sum_{\mathbf{d},\mathbf{e}}{}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$ where $\displaystyle m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$.

- We have to deal with $\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^{*}$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$.
- It is desirable to rid ourselves of the condition that $(d_u, e_v) = 1$ when $u \ne v$.

- We have to deal with $\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \neq v$.
- It is desirable to rid ourselves of the condition that $(d_u, e_v) = 1$ when $u \neq v$.
- That this is possible without undue effect on the main term is due to the prior sieving resulting from the choice of the residue class a modulo q. Thus any primes p which can potentially divide (d_u, e_v) satisfy $p > Q$.

- We have to deal with $\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \neq v$.
- It is desirable to rid ourselves of the condition that $(d_u, e_v) = 1$ when $u \neq v$.
- That this is possible without undue effect on the main term is due to the prior sieving resulting from the choice of the residue class a modulo q. Thus any primes p which can potentially divide (d_u, e_v) satisfy $p > Q$.
- Now $\frac{\phi(m)}{\phi(q)} = \prod_{i\neq j}\phi([d_i, e_i])$ & $\frac{1}{\phi([d_i, e_i])} = \frac{\phi((d_i, e_i))}{\phi(d_i)\phi(e_i)}$.

- We have to deal with $\sum^*_{\mathbf{d},\mathbf{e}} \dfrac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$.
- It is desirable to rid ourselves of the condition that $(d_u, e_v) = 1$ when $u \ne v$.
- That this is possible without undue effect on the main term is due to the prior sieving resulting from the choice of the residue class a modulo q. Thus any primes p which can potentially divide (d_u, e_v) satisfy $p > Q$.
- Now $\dfrac{\phi(m)}{\phi(q)} = \prod_{i\ne j}\phi([d_i, e_i])$ & $\dfrac{1}{\phi([d_i, e_i])} = \dfrac{\phi((d_i, e_i))}{\phi(d_i)\phi(e_i)}$.
- Hence $\dfrac{1}{\phi(m)} = \dfrac{1}{\phi(q)\phi(d)\phi(e)}\prod_{i\ne j}\phi((d_i, e_i))$.

- $\displaystyle\sum_{\mathbf{d},\mathbf{e}}{}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$ where $m = q\displaystyle\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$, and

$$\frac{1}{\phi(m)} = \frac{1}{\phi(q)\phi(d)\phi(e)} \prod_{i\ne j} \phi\big((d_i, e_i)\big)$$

- $\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p \le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$, and
$$\frac{1}{\phi(m)} = \frac{1}{\phi(q)\phi(d)\phi(e)} \prod_{i \ne j} \phi\big((d_i, e_i)\big)$$
- Also $p - 1 = 1 + (p - 2)$, so for squarefree l we have $\phi(l) = \sum_{t|l} \phi_2(t)$.

- $\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^{*}$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$, and

$$\frac{1}{\phi(m)} = \frac{1}{\phi(q)\phi(d)\phi(e)} \prod_{i\ne j} \phi((d_i, e_i))$$

- Also $p - 1 = 1 + (p - 2)$, so for squarefree l we have $\phi(l) = \sum_{t|l} \phi_2(t)$.
- Hence $\phi((d_i, e_i)) = \sum_{n_i|d_i, n_i|e_i} \phi_2(n_i)$.

- $\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^{*}$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$, and

$$\frac{1}{\phi(m)} = \frac{1}{\phi(q)\phi(d)\phi(e)}\prod_{i\ne j}\phi((d_i, e_i))$$

- Also $p - 1 = 1 + (p - 2)$, so for squarefree l we have $\phi(l) = \sum_{t|l}\phi_2(t)$.
- Hence $\phi((d_i, e_i)) = \sum_{n_i|d_i, n_i|e_i}\phi_2(n_i)$.
- We substitute this in the main term and invert the order of summation to obtain

$$\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} = \sum_{\mathbf{n}}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{\mathbf{d},\mathbf{e}\\ \mathbf{n}|\mathbf{d},\mathbf{n}|\mathbf{e}}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(d)\phi(e)}.$$

- $\displaystyle\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} = \sum_{\mathbf{n}}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{\mathbf{d},\mathbf{e}\\ \mathbf{n}|\mathbf{d},\mathbf{n}|\mathbf{e}}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(d)\phi(e)}$ where $\displaystyle m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$

- $$\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} = \sum_{\mathbf{n}}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{\mathbf{d},\mathbf{e} \\ \mathbf{n}|\mathbf{d},\mathbf{n}|\mathbf{e}}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(d)\phi(e)}$$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$
- We now begin to deal with $(d_u, e_v) = 1$ for $u \ne v$. We replace it by $\displaystyle\sum_{s_{uv}|d_u, s_{uv}|e_v} \mu(s_{uv})$.

- $\displaystyle\sum_{\mathbf{d},\mathbf{e}}{}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} = \sum_{\mathbf{n}}{}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{\mathbf{d},\mathbf{e} \\ \mathbf{n}|\mathbf{d},\mathbf{n}|\mathbf{e}}}{}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(d)\phi(e)}$ where $\displaystyle m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$
- We now begin to deal with $(d_u, e_v) = 1$ for $u \ne v$. We replace it by $\displaystyle\sum_{s_{uv}|d_u, s_{uv}|e_v} \mu(s_{uv})$.
- There are various observations with regard to the s_{uv}.

- $$\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} = \sum_{\mathbf{n}}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{\mathbf{d},\mathbf{e}\\ \mathbf{n}|\mathbf{d},\mathbf{n}|\mathbf{e}}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(d)\phi(e)}$$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$
- We now begin to deal with $(d_u, e_v) = 1$ for $u \ne v$. We replace it by $\displaystyle\sum_{s_{uv}|d_u, s_{uv}|e_v} \mu(s_{uv})$.
- There are various observations with regard to the s_{uv}.
- We have $n_u|d_u$, so $(e_v, n_u) = 1$. Hence $(s_{uv}, n_u) = 1$.

- $$\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} = \sum_{\mathbf{n}}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{\mathbf{d},\mathbf{e} \\ \mathbf{n}|\mathbf{d},\mathbf{n}|\mathbf{e}}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(d)\phi(e)}$$ where $m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$
- We now begin to deal with $(d_u, e_v) = 1$ for $u \ne v$. We replace it by $\displaystyle\sum_{s_{uv}|d_u, s_{uv}|e_v} \mu(s_{uv})$.
- There are various observations with regard to the s_{uv}.
- We have $n_u|d_u$, so $(e_v, n_u) = 1$. Hence $(s_{uv}, n_u) = 1$.
- Likewise $(s_{uv}, n_v) = 1$.

- $\displaystyle\sum_{\mathbf{d},\mathbf{e}}{}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} = \sum_{\mathbf{n}}{}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{\mathbf{d},\mathbf{e}\\ \mathbf{n}|\mathbf{d},\mathbf{n}|\mathbf{e}}}{}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(d)\phi(e)}$ where $\displaystyle m = q\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p\le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$
- We now begin to deal with $(d_u, e_v) = 1$ for $u \ne v$. We replace it by $\displaystyle\sum_{s_{uv}|d_u, s_{uv}|e_v} \mu(s_{uv})$.
- There are various observations with regard to the s_{uv}.
- We have $n_u|d_u$, so $(e_v, n_u) = 1$. Hence $(s_{uv}, n_u) = 1$.
- Likewise $(s_{uv}, n_v) = 1$.
- Also, when $w \ne v$, we have $s_{uw}|e_w$ and $(e_v, e_w) = 1$. Hence$(s_{uv}, s_{uw}) = 1$.

- $\displaystyle\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} = \sum_{\mathbf{n}}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{\mathbf{d},\mathbf{e} \\ \mathbf{n}|\mathbf{d},\mathbf{n}|\mathbf{e}}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(d)\phi(e)}$ where

  $m = q\displaystyle\prod_{i=1}^{k}[d_i, e_i]$, $q = \prod_{p \le Q} p$, and $\sum^*$ indicates the restrictions $d_j = e_j = 1$ and $(d_u, e_v) = 1$ when $u \ne v$
- We now begin to deal with $(d_u, e_v) = 1$ for $u \ne v$. We replace it by $\displaystyle\sum_{s_{uv}|d_u, s_{uv}|e_v} \mu(s_{uv})$.
- There are various observations with regard to the s_{uv}.
- We have $n_u|d_u$, so $(e_v, n_u) = 1$. Hence $(s_{uv}, n_u) = 1$.
- Likewise $(s_{uv}, n_v) = 1$.
- Also, when $w \ne v$, we have $s_{uw}|e_w$ and $(e_v, e_w) = 1$. Hence$(s_{uv}, s_{uw}) = 1$.
- Likewise, when $w \ne u$, $(s_{uv}, s_{wv}) = 1$ and so in summary

$$(s_{uv}, n_u) = (s_{uv}, n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1.$$

- Thus *via* $\displaystyle\sum_{s_{uv}|d_u, s_{uv}|e_v} \mu(s_{uv})$, $\displaystyle\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} =$

$$\sum_{\mathbf{n}}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{s_{uv} \\ u\neq v}}^{\dagger} \Big(\prod_{u\neq v} \mu(s_{uv}) \Big) \Bigg(\sum_{\substack{\mathbf{d} \\ \mathbf{n}|\mathbf{d} \\ s_{uv}|d_u}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)} \Bigg) \Bigg(\sum_{\substack{\mathbf{e} \\ \mathbf{n}|\mathbf{e} \\ s_{uv}|e_v}}^{j} \frac{\lambda(\mathbf{e})}{\phi(e)} \Bigg)$$

with $\sum^{\dagger}$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1$.

- Thus *via* $\sum\limits_{s_{uv}|d_u, s_{uv}|e_v} \mu(s_{uv})$, $\sum\limits_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} =$

$$\sum_{\mathbf{n}}^{j} \frac{\phi_2(n)}{\phi(q)} \sum_{\substack{s_{uv} \\ u\neq v}}^{\dagger} \Big(\prod_{u\neq v} \mu(s_{uv}) \Big) \Bigg(\sum_{\substack{\mathbf{d} \\ \mathbf{n}|\mathbf{d} \\ s_{uv}|d_u}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)} \Bigg) \Bigg(\sum_{\substack{\mathbf{e} \\ \mathbf{n}|\mathbf{e} \\ s_{uv}|e_v}}^{j} \frac{\lambda(\mathbf{e})}{\phi(e)} \Bigg)$$

with $\sum^{\dagger}$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1.$

- This is not yet a diagonal form, but it is progress.

- Thus *via* $\displaystyle\sum_{s_{uv}|d_u,s_{uv}|e_v}\mu(s_{uv})$, $\displaystyle\sum_{\mathbf{d},\mathbf{e}}^{*}\frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)} =$

$$\sum_{\mathbf{n}}^{j}\frac{\phi_2(n)}{\phi(q)}\sum_{\substack{s_{uv}\\u\neq v}}^{\dagger}\Big(\prod_{u\neq v}\mu(s_{uv})\Big)\Bigg(\sum_{\substack{\mathbf{d}\\ \mathbf{n}|\mathbf{d}\\ s_{uv}|d_u}}^{j}\frac{\lambda(\mathbf{d})}{\phi(d)}\Bigg)\Bigg(\sum_{\substack{\mathbf{e}\\ \mathbf{n}|\mathbf{e}\\ s_{uv}|e_v}}^{j}\frac{\lambda(\mathbf{e})}{\phi(e)}\Bigg)$$

with $\sum^{\dagger}$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1$.
- This is not yet a diagonal form, but it is progress.
- We sub $\displaystyle\frac{\kappa_j(\mathbf{r})}{\mu(r)\phi_2(r)} = \sum_{\substack{\mathbf{d}\\ \mathbf{r}|\mathbf{d}}}^{j}\frac{\lambda(\mathbf{d})}{\phi(d)}$ for λ, so $\displaystyle\sum_{\mathbf{d},\mathbf{e}}^{*}\frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$

$$= \sum_{\mathbf{n}}^{j}\frac{1}{\phi(q)\phi_2(n)}\sum_{\substack{s_{uv}\\u\neq v}}^{\dagger}\Big(\prod_{u\neq v}\frac{\mu(s_{uv})}{\phi_2(s_{uv})^2}\Big)\kappa_j(\mathbf{a})\kappa_j(\mathbf{b})$$

where $\mathbf{a} = a_1,\ldots,a_k$, $\mathbf{b} = b_1,\ldots,b_k$ are factors of $\mathbf{d}$, $\mathbf{e}$,

$$a_u = n_u\prod_{\substack{v\\v\neq u}}s_{uv},\quad b_v = n_v\prod_{\substack{u\\u\neq v}}s_{uv}.$$

- Thus $\displaystyle\sum_{\mathbf{d},\mathbf{e}}^{*} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$

$$= \sum_{\mathbf{n}}^{j} \frac{1}{\phi(q)\phi_2(n)} \sum_{\substack{s_{uv}\\ u\neq v}}^{\dagger} \left(\prod_{u\neq v} \frac{\mu(s_{uv})}{\phi_2(s_{uv})^2} \right) \kappa_j(\mathbf{a})\kappa_j(\mathbf{b})$$

where $\mathbf{a} = a_1, \dots, a_k$, $\mathbf{b} = b_1, \dots, b_k$,

$$a_u = n_u \prod_{\substack{v\\ v\neq u}} s_{uv}, \quad b_v = n_v \prod_{\substack{u\\ u\neq v}} s_{uv}$$

and $\sum^{\dagger}$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1$.

- Thus $\sum^*_{\mathbf{d},\mathbf{e}} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$

$$= \sum_{\mathbf{n}}^{j} \frac{1}{\phi(q)\phi_2(n)} \sum^{\dagger}_{\substack{s_{uv}\\ u\neq v}} \left(\prod_{u\neq v} \frac{\mu(s_{uv})}{\phi_2(s_{uv})^2} \right) \kappa_j(\mathbf{a})\kappa_j(\mathbf{b})$$

where $\mathbf{a} = a_1, \ldots, a_k$, $\mathbf{b} = b_1, \ldots, b_k$,

$$a_u = n_u \prod_{\substack{v\\ v\neq u}} s_{uv}, \quad b_v = n_v \prod_{\substack{u\\ u\neq v}} s_{uv}$$

and $\sum^\dagger$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1.$
- In particular $a = b = ns$ where $s = \prod_{u\neq v} s_{uv}$.

- Thus $\sum^*_{\mathbf{d},\mathbf{e}} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi(m)}$

$$= \sum^j_{\mathbf{n}} \frac{1}{\phi(q)\phi_2(n)} \sum^\dagger_{\substack{s_{uv}\\ u\neq v}} \left(\prod_{u\neq v} \frac{\mu(s_{uv})}{\phi_2(s_{uv})^2} \right) \kappa_j(\mathbf{a})\kappa_j(\mathbf{b})$$

where $\mathbf{a} = a_1, \ldots, a_k$, $\mathbf{b} = b_1, \ldots, b_k$,

$$a_u = n_u \prod_{\substack{v\\ v\neq u}} s_{uv}, \quad b_v = n_v \prod_{\substack{u\\ u\neq v}} s_{uv}$$

and $\sum^\dagger$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1$.
- In particular $a = b = ns$ where $s = \prod_{u\neq v} s_{uv}$.
- Thus the main term is

$$\sum^j_{\mathbf{n}} \frac{1}{\phi(q)\phi_2(n)} \sum^\dagger_{\substack{s_{uv}\\ u\neq v}} \frac{\mu(s)}{\phi_2(s)^2} \kappa_j(\mathbf{a})\kappa_j(\mathbf{b}).$$

- The main term is $\sum_{\mathbf{n}}^{j} \frac{1}{\phi(q)\phi_2(n)} \sum_{\substack{s_{uv} \\ u \neq v}}^{\dagger} \frac{\mu(s)}{\phi_2(s)^2} \kappa_j(\mathbf{a})\kappa_j(\mathbf{b})$

  and $\sum^{\dagger}$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1$.

- The main term is $\displaystyle\sum_{\mathbf{n}}^{j} \frac{1}{\phi(q)\phi_2(n)} \sum_{\substack{s_{uv} \\ u\neq v}}^{\dagger} \frac{\mu(s)}{\phi_2(s)^2} \kappa_j(\mathbf{a})\kappa_j(\mathbf{b})$

  and $\sum^\dagger$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1$.
- Since $n_j = 1$ the terms with $s > 1$ contribute

$$\ll \frac{K_j^2}{\phi(q)} \sum_{\substack{n\le R \\ (n,q)=1}} \frac{d_{k-1}(n)\mu(n)^2}{\phi_2(n)} \sum_{\substack{s>1 \\ (s,q)=1}} \frac{d_{k(k-1)}(s)\mu(s)^2}{\phi_2(s)^2}.$$

- The main term is $\sum_{\mathbf{n}}^{j} \frac{1}{\phi(q)\phi_2(n)} \sum_{\substack{s_{uv} \\ u\neq v}}^{\dagger} \frac{\mu(s)}{\phi_2(s)^2} \kappa_j(\mathbf{a})\kappa_j(\mathbf{b})$
  and $\sum^{\dagger}$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1$.
- Since $n_j = 1$ the terms with $s > 1$ contribute

$$\ll \frac{K_j^2}{\phi(q)} \sum_{\substack{n\le R \\ (n,q)=1}} \frac{d_{k-1}(n)\mu(n)^2}{\phi_2(n)} \sum_{\substack{s>1 \\ (s,q)=1}} \frac{d_{k(k-1)}(s)\mu(s)^2}{\phi_2(s)^2}.$$

- The inner sum is
  $\ll -1 + \prod_{p>Q} \left(1 + \frac{k(k-1)}{(p-2)^2}\right) \ll \frac{1}{Q\log Q}$

- The main term is $\displaystyle\sum_{\mathbf{n}}^{j} \frac{1}{\phi(q)\phi_2(n)} \sum_{\substack{s_{uv} \\ u\neq v}}^{\dagger} \frac{\mu(s)}{\phi_2(s)^2}\kappa_j(\mathbf{a})\kappa_j(\mathbf{b})$

  and $\sum^{\dagger}$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1$.
- Since $n_j = 1$ the terms with $s > 1$ contribute

$$\ll \frac{K_j^2}{\phi(q)} \sum_{\substack{n\le R\\(n,q)=1}} \frac{d_{k-1}(n)\mu(n)^2}{\phi_2(n)} \sum_{\substack{s>1\\(s,q)=1}} \frac{d_{k(k-1)}(s)\mu(s)^2}{\phi_2(s)^2}.$$

- The inner sum is
  $\displaystyle\ll -1 + \prod_{p>Q}\left(1+\frac{k(k-1)}{(p-2)^2}\right) \ll \frac{1}{Q\log Q}$
- and the sum over n is

$$\ll \prod_{Q<p\le R}(1+(k-1)/(p-2)) \ll (\phi(q)(\log R)/q)^{k-1}.$$

- The main term is $\displaystyle\sum_{\mathbf{n}}^{j} \frac{1}{\phi(q)\phi_2(n)} \sum_{\substack{s_{uv} \\ u\neq v}}^{\dagger} \frac{\mu(s)}{\phi_2(s)^2} \kappa_j(\mathbf{a})\kappa_j(\mathbf{b})$

  and $\sum^{\dagger}$: $(s_{uv}, n_u n_v) = (s_{uv}, s_{uw}) = (s_{uv}, s_{wv}) = 1$.
- Since $n_j = 1$ the terms with $s > 1$ contribute

$$\ll \frac{K_j^2}{\phi(q)} \sum_{\substack{n\le R \\ (n,q)=1}} \frac{d_{k-1}(n)\mu(n)^2}{\phi_2(n)} \sum_{\substack{s>1 \\ (s,q)=1}} \frac{d_{k(k-1)}(s)\mu(s)^2}{\phi_2(s)^2}.$$

- The inner sum is
  $\displaystyle \ll -1 + \prod_{p>Q}\left(1 + \frac{k(k-1)}{(p-2)^2}\right) \ll \frac{1}{Q\log Q}$
- and the sum over n is

$$\ll \prod_{Q<p\le R} (1 + (k-1)/(p-2)) \ll (\phi(q)(\log R)/q)^{k-1}.$$

- Thus the total contribution from the terms with
  $s = \prod_{u\neq v} s_{uv} > 1$ is $\displaystyle \frac{K_j^2 \phi(q)^{k-2}(\log R)^{k-1}}{q^{k-1}Q}$.

- For the terms with $s = 1$ we have $\mathbf{a} = \mathbf{b} = \mathbf{n}$. Thus the main term becomes

$$\sum_{\mathbf{n}}^{j} \frac{\kappa_j(\mathbf{n})^2}{\phi(q)\phi_2(n)} + O\left(\frac{K_j^2\phi(q)^{k-2}(\log R)^{k-1}}{q^{k-1}Q}\right)$$

- For the terms with $s = 1$ we have $\mathbf{a} = \mathbf{b} = \mathbf{n}$. Thus the main term becomes

$$\sum_{\mathbf{n}}^{j} \frac{\kappa_j(\mathbf{n})^2}{\phi(q)\phi_2(n)} + O\left(\frac{K_j^2 \phi(q)^{k-2}(\log R)^{k-1}}{q^{k-1}Q}\right)$$

- Recall that this is multiplied by

$$\int_{N+h_j}^{2N+h_j} \frac{d\alpha}{\log \alpha} = \frac{N}{\log N} + O\left(\frac{N}{(\log N)^2}.\right)$$

- For the terms with $s = 1$ we have $\mathbf{a} = \mathbf{b} = \mathbf{n}$. Thus the main term becomes

$$\sum_{\mathbf{n}}^{j} \frac{\kappa_j(\mathbf{n})^2}{\phi(q)\phi_2(n)} + O\left(\frac{K_j^2\phi(q)^{k-2}(\log R)^{k-1}}{q^{k-1}Q}\right)$$

- Recall that this is multiplied by

$$\int_{N+h_j}^{2N+h_j} \frac{d\alpha}{\log\alpha} = \frac{N}{\log N} + O\left(\frac{N}{(\log N)^2}.\right)$$

- Since

$$\sum_{\mathbf{n}}^{j} \frac{\kappa_j(\mathbf{n})^2}{\phi(q)\phi_2(n)} \ll \frac{K_j^2}{\phi(q)} \sum_{n\le R} \frac{d_{k-1}(n)}{phi_2(n)} \ll \frac{K_j^2}{\phi(q)}(\log R)^{k-1}$$

the complete main term is seen to be

$$\frac{N}{\log N}\sum_{\mathbf{n}}^{j} \frac{\kappa_j(\mathbf{n})^2}{\phi(q)\phi_2(n)} + O\left(\frac{NK_j^2\phi(q)^{k-2}(\log R)^{k-1}}{q^{k-1}Q(\log N)}\right).$$

- This completes the proof of the approximation for S_j.

- This completes the proof of the approximation for S_j.
- The proof of the approximation for $T(f)$ is essentially the same, except that we do not use Bombieri's theorem and we do not have the restriction that $d_j = 1$ to contend with.

- This completes the proof of the approximation for S_j.
- The proof of the approximation for $T(f)$ is essentially the same, except that we do not use Bombieri's theorem and we do not have the restriction that $d_j = 1$ to contend with.
- Thus on the initial application of the Chinese Remainder Theorem the main term is

$$\frac{N}{m}$$

and the error term is $O(1)$.

- This completes the proof of the approximation for S_j.
- The proof of the approximation for $T(f)$ is essentially the same, except that we do not use Bombieri's theorem and we do not have the restriction that $d_j = 1$ to contend with.
- Thus on the initial application of the Chinese Remainder Theorem the main term is
$$\frac{N}{m}$$
and the error term is $O(1)$.
- Since
$$\max_{\mathbf{d}} |\lambda(\mathbf{d})| \ll K(\log R)^k$$
we see that the total contribution arising from this error is
$$\ll K^2 R^2 (\log R)^{4k-2}$$
which is acceptable since $R = N^{\theta/2-\delta}$.

- This completes the proof of the approximation for S_j.
- The proof of the approximation for $T(f)$ is essentially the same, except that we do not use Bombieri's theorem and we do not have the restriction that $d_j = 1$ to contend with.
- Thus on the initial application of the Chinese Remainder Theorem the main term is
$$\frac{N}{m}$$
and the error term is $O(1)$.
- Since
$$\max_{\mathbf{d}} |\lambda(\mathbf{d})| \ll K(\log R)^k$$
we see that the total contribution arising from this error is
$$\ll K^2R^2(\log R)^{4k-2}$$
which is acceptable since $R = N^{\theta/2-\delta}$.
- Then just as the function ϕ now plays the rôle that ϕ_2 played earlier, so the κ_j is replaced by its understudy κ. The process of replacing λ by κ is identical, as is the elimination of the restriction $(d_u, e_v) = 1$.

- To summarize, we have established Lemma 8.
  Let

$$K_j = \max_{\mathbf{r}} |\kappa_j(\mathbf{r})|, \quad K = \max_{\mathbf{r}} |\kappa(\mathbf{r})|.$$

Then

$$S_j(f) = \frac{N}{\phi(q)\log N}\sum_{\mathbf{r}}^{j} \frac{\kappa_j(\mathbf{r})^2}{\phi_2(r)} + O\left(\frac{K_j^2\phi(q)^{k-2}N(\log R)^{k-2}}{q^{k-1}Q}\right)$$

and

$$T(f) = \frac{N}{q}\sum_{\mathbf{r}} \frac{\kappa(\mathbf{r})^2}{\phi(r)} + O\left(\frac{K^2N(\log R)^k}{qQ}\right).$$

- We have initially defined κ and κ_j in terms of λ.

$$\kappa(\mathbf{r}) = \mu(r)\phi(r)\sum_{\substack{\mathbf{d}\\ \mathbf{r}|\mathbf{d}}} \frac{\lambda(\mathbf{d})}{d}.$$

$$\kappa_j(\mathbf{r}) = \mu(r)\phi_2(r)\sum_{\substack{\mathbf{d}\\ \mathbf{r}|\mathbf{d}}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)} \quad (j = 1, \ldots, k),$$

where $\sum^j$ indicates that the summation variable is a k–tuple, say $\mathbf{d}$, which is restricted by $d_j = 1$

- We have initially defined κ and κ_j in terms of λ.

$$\kappa(\mathbf{r}) = \mu(r)\phi(r)\sum_{\substack{\mathbf{d}\\ \mathbf{r}|\mathbf{d}}} \frac{\lambda(\mathbf{d})}{d}.$$

$$\kappa_j(\mathbf{r}) = \mu(r)\phi_2(r)\sum_{\substack{\mathbf{d}\\ \mathbf{r}|\mathbf{d}}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)} \quad (j = 1, \ldots, k),$$

where $\sum^j$ indicates that the summation variable is a k–tuple, say $\mathbf{d}$, which is restricted by $d_j = 1$

- In Lemma 7 we showed they are invertible. $\frac{\mu(d)}{\phi(d)}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$ and $\frac{\mu(d)}{d}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}} \frac{\kappa(\mathbf{r})}{\phi(r)}$.

- We have initially defined κ and κ_j in terms of λ.

$$\kappa(\mathbf{r}) = \mu(r)\phi(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}} \frac{\lambda(\mathbf{d})}{d}.$$

$$\kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)} \quad (j = 1, \ldots, k),$$

where $\sum^j$ indicates that the summation variable is a k–tuple, say $\mathbf{d}$, which is restricted by $d_j = 1$

- In Lemma 7 we showed they are invertible.
$\frac{\mu(d)}{\phi(d)}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$ and $\frac{\mu(d)}{d}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r} \\ \mathbf{d}|\mathbf{r}}} \frac{\kappa(\mathbf{r})}{\phi(r)}$.
- Thus as in the Selberg sieve, rather than choosing first λ, we can instead choose κ, and then the values of λ, and so κ_j, will follow.

- $\dfrac{\mu(d)}{d}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}} \dfrac{\kappa(\mathbf{r})}{\phi(r)}$ and $\dfrac{\mu(d)}{\phi(d)}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \dfrac{\kappa_j(\mathbf{r})}{\phi_2(r)}$.

- $$\frac{\mu(d)}{d}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}} \frac{\kappa(\mathbf{r})}{\phi(r)} \text{ and } \frac{\mu(d)}{\phi(d)}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}.$$
- You may recall that it was asserted in (4) that we would choose

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}\\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right).$$

- $\displaystyle \frac{\mu(d)}{d}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}} \frac{\kappa(\mathbf{r})}{\phi(r)}$ and $\displaystyle \frac{\mu(d)}{\phi(d)}\lambda(\mathbf{d}) = \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}}}^{j} \frac{\kappa_j(\mathbf{r})}{\phi_2(r)}$.
- You may recall that it was asserted in (4) that we would choose

$$\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{r}\\ \mathbf{d}|\mathbf{r}\\ (r,q)=1}} \frac{\mu(r)^2}{\phi(r)} f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right).$$

- The motivation for this was the knowledge that this can be achieved by simply taking

$$\kappa(\mathbf{r}) = f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right).$$

- It is useful to have an estimate for κ_j in terms of κ.

- It is useful to have an estimate for κ_j in terms of κ.
- $\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{s} \\ \mathbf{d}|\mathbf{s}}} \frac{\kappa(\mathbf{s})}{\phi(s)}, \quad \kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)}$

- It is useful to have an estimate for κ_j in terms of κ.
- $\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{s} \\ \mathbf{d}|\mathbf{s}}} \frac{\kappa(\mathbf{s})}{\phi(s)}, \quad \kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)}$
- Thus $\kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{s} \\ \mathbf{r}|\mathbf{s}}} \frac{\kappa(\mathbf{s})}{\phi(s)} \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}|\mathbf{s}}}^{j} \frac{\mu(d)d}{\phi(d)}$.

- It is useful to have an estimate for κ_j in terms of κ.
- $\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{s} \\ \mathbf{d}|\mathbf{s}}} \frac{\kappa(\mathbf{s})}{\phi(s)}, \quad \kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum^{j}_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}} \frac{\lambda(\mathbf{d})}{\phi(d)}$
- Thus $\kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{s} \\ \mathbf{r}|\mathbf{s}}} \frac{\kappa(\mathbf{s})}{\phi(s)} \sum^{j}_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}|\mathbf{s}}} \frac{\mu(d)d}{\phi(d)}$.
- Write $e_i = d_i/r_i$ and $t_i = s_i/r_i$. Then the inner sum is

$$\frac{\mu(r)r}{\phi(r)} \sum_{\substack{\mathbf{e} \\ \mathbf{e}|\mathbf{t} \\ e_j=1}} \frac{\mu(e)e}{\phi(e)} = \frac{\mu(r)r\mu(t/t_j)}{\phi(r)\phi(t/t_j)}.$$

- It is useful to have an estimate for κ_j in terms of κ.
- $\lambda(\mathbf{d}) = \mu(d)d \sum_{\substack{\mathbf{s} \\ \mathbf{d}|\mathbf{s}}} \frac{\kappa(\mathbf{s})}{\phi(s)}, \quad \kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}}}^{j} \frac{\lambda(\mathbf{d})}{\phi(d)}$
- Thus $\kappa_j(\mathbf{r}) = \mu(r)\phi_2(r) \sum_{\substack{\mathbf{s} \\ \mathbf{r}|\mathbf{s}}} \frac{\kappa(\mathbf{s})}{\phi(s)} \sum_{\substack{\mathbf{d} \\ \mathbf{r}|\mathbf{d}|\mathbf{s}}}^{j} \frac{\mu(d)d}{\phi(d)}$.
- Write $e_i = d_i/r_i$ and $t_i = s_i/r_i$. Then the inner sum is

$$\frac{\mu(r)r}{\phi(r)} \sum_{\substack{\mathbf{e} \\ \mathbf{e}|\mathbf{t} \\ e_j=1}} \frac{\mu(e)e}{\phi(e)} = \frac{\mu(r)r\mu(t/t_j)}{\phi(r)\phi(t/t_j)}.$$

- Using $\mathbf{rt}(=\mathbf{s})$ for $r_1t_1, \ldots, r_kt_k$,

$$\kappa_j(\mathbf{r}) = \frac{r\phi_2(r)}{\phi(r)^2} \sum_{\mathbf{t}} \kappa(\mathbf{rt}) \frac{\mu(t)\phi(t_j)\mu(t_j)}{\phi(t)^2}.$$

- $\kappa_j(\mathbf{r}) = \frac{r\phi_2(r)}{\phi(r)^2} \sum_{\mathbf{t}} \kappa(\mathbf{rt}) \frac{\mu(t)\phi(t_j)\mu(t_j)}{\phi(t)^2}.$

- $\kappa_j(\mathbf{r}) = \frac{r\phi_2(r)}{\phi(r)^2} \sum_{\mathbf{t}} \kappa(\mathbf{rt}) \frac{\mu(t)\phi(t_j)\mu(t_j)}{\phi(t)^2}.$
- The $t > t_j$ contribute

$$\ll K \sum_{\substack{t_j \le R \\ (t_j,q)=1}} \frac{\mu(t_j)^2}{\phi(t_j)} \sum_{\substack{n>1 \\ (n,q)=1}} \frac{(k-1)^{\omega(n)}\mu(n)^2}{\phi(n)^2}$$

- $\kappa_j(\mathbf{r}) = \dfrac{r\phi_2(r)}{\phi(r)^2} \sum_{\mathbf{t}} \kappa(\mathbf{rt}) \dfrac{\mu(t)\phi(t_j)\mu(t_j)}{\phi(t)^2}.$
- The $t > t_j$ contribute

$$\ll K \sum_{\substack{t_j \le R \\ (t_j,q)=1}} \frac{\mu(t_j)^2}{\phi(t_j)} \sum_{\substack{n>1 \\ (n,q)=1}} \frac{(k-1)^{\omega(n)}\mu(n)^2}{\phi(n)^2}$$

- The inner sum is $-1 + \prod_{p>Q}\left(1 + \dfrac{k-1}{(p-1)^2}\right) \ll Q^{-1}$. and

we have $\displaystyle \sum_{\substack{t_j \le R \\ (t_j,q)=1}} \frac{\mu(t_j)^2}{\phi(t_j)} \ll \prod_{Q<p\le R} \frac{p}{p-1} \ll \frac{\phi(q)}{q} \log R.$

- $\kappa_j(\mathbf{r}) = \dfrac{r\phi_2(r)}{\phi(r)^2}\sum_{\mathbf{t}}\kappa(\mathbf{rt})\dfrac{\mu(t)\phi(t_j)\mu(t_j)}{\phi(t)^2}$.
- The $t > t_j$ contribute

$$\ll K \sum_{\substack{t_j\le R\\(t_j,q)=1}} \frac{\mu(t_j)^2}{\phi(t_j)} \sum_{\substack{n>1\\(n,q)=1}} \frac{(k-1)^{\omega(n)}\mu(n)^2}{\phi(n)^2}$$

- The inner sum is $-1+\prod_{p>Q}\left(1+\dfrac{k-1}{(p-1)^2}\right) \ll Q^{-1}$. and

we have $\displaystyle\sum_{\substack{t_j\le R\\(t_j,q)=1}} \frac{\mu(t_j)^2}{\phi(t_j)} \ll \prod_{Q<p\le R}\frac{p}{p-1} \ll \frac{\phi(q)}{q}\log R.$

- Since also $\left(\dfrac{r\phi_2(r)}{\phi(r)}\right)^2 = 1 + O(1/Q)$ it follows when $r_j = 1$,

$$\kappa_j(\mathbf{r}) = \sum_{t_j}\frac{\kappa(\mathbf{r}')}{\phi(t_j)} + O\left(\frac{K\phi(q)\log R}{qQ}\right)$$

where $\mathbf{r}' = r_1,\ldots,r_{j-1},t_j,r_{j+1},\ldots,r_k$.

- The final step of the proof of Maynard's theorem is to obtain smooth approximations to the main terms.

- The final step of the proof of Maynard's theorem is to obtain smooth approximations to the main terms.
- We already did this for the Selberg sieve, i.e. $k = 1$.

- The final step of the proof of Maynard's theorem is to obtain smooth approximations to the main terms.
- We already did this for the Selberg sieve, i.e. $k = 1$.
- We adopt the expedient of establishing a one–dimensional approximation and applying it k–times.

- The final step of the proof of Maynard's theorem is to obtain smooth approximations to the main terms.
- We already did this for the Selberg sieve, i.e. $k = 1$.
- We adopt the expedient of establishing a one–dimensional approximation and applying it k–times.
- Suppose that $g : [0, 1] \to \mathbb{R}$. Then we call g *l–piecewise absolutely continuous on* $[0, 1]$ when there is a partition $a_0 = 0 < a_1 < \ldots < a_l = 1$ of $[0, 1]$ so that for $1 \le j \le l$
  1. $g_+(a_{j-1}) = \lim_{x\to a_{j-1}+} g(x)$ & $g_-(a_j) = \lim_{x\to a_j-} g(x)$ exist,
  2. g is absolutely continuous on $[a_{j-1}, a_j]$ when we replace $g(a_{j-1})$ and $g(a_j)$ by $g_+(a_{j-1})$ and $g_-(a_j)$ respectively.

- The final step of the proof of Maynard's theorem is to obtain smooth approximations to the main terms.
- We already did this for the Selberg sieve, i.e. $k = 1$.
- We adopt the expedient of establishing a one–dimensional approximation and applying it k–times.
- Suppose that $g : [0,1] \to \mathbb{R}$. Then we call g *l–piecewise absolutely continuous on* $[0,1]$ when there is a partition $a_0 = 0 < a_1 < \ldots < a_l = 1$ of $[0,1]$ so that for $1 \le j \le l$
  1. $g_+(a_{j-1}) = \lim_{x\to a_{j-1}+} g(x)$ & $g_-(a_j) = \lim_{x\to a_j-} g(x)$ exist,
  2. g is absolutely continuous on $[a_{j-1}, a_j]$ when we replace $g(a_{j-1})$ and $g(a_j)$ by $g_+(a_{j-1})$ and $g_-(a_j)$ respectively.
- We define $\mathcal{G}(l, G)$ to be the class of l–piecewise absolutely continuous functions g on $[0,1]$ such that

$$\sup_{v\in[0,1]} |g(v)| + \int_0^1 |g'(v)|dv \le G.$$

- The final step of the proof of Maynard's theorem is to obtain smooth approximations to the main terms.
- We already did this for the Selberg sieve, i.e. $k = 1$.
- We adopt the expedient of establishing a one–dimensional approximation and applying it k–times.
- Suppose that $g : [0,1] \to \mathbb{R}$. Then we call g *l–piecewise absolutely continuous on* $[0,1]$ when there is a partition $a_0 = 0 < a_1 < \ldots < a_l = 1$ of $[0,1]$ so that for $1 \le j \le l$
  1. $g_+(a_{j-1}) = \lim_{x \to a_{j-1}+} g(x)$ & $g_-(a_j) = \lim_{x \to a_j-} g(x)$ exist,
  2. g is absolutely continuous on $[a_{j-1}, a_j]$ when we replace $g(a_{j-1})$ and $g(a_j)$ by $g_+(a_{j-1})$ and $g_-(a_j)$ respectively.
- We define $\mathcal{G}(l, G)$ to be the class of l–piecewise absolutely continuous functions g on $[0,1]$ such that

$$\sup_{v \in [0,1]} |g(v)| + \int_0^1 |g'(v)| dv \le G.$$

- In practice it suffices that g' is continuous except for at most one x in $[0,1]$ where g and g' have jump discontinuities.

- We establish

Lemma 9

Suppose $\eta:\mathbb{N}\to\mathbb{R}$ *is multiplicative, supported on the squarefree numbers, that* $0\le\eta(p)\le 2$. $\eta(2)<2$ *and there is a* $C>0$ *such that, whenever* $p>C$, $\left|\eta(p)-\frac{1}{p}\right|\le\frac{C}{p^2}$. *Suppose also* $g\in\mathcal{G}(I,G)$ *and* $m\in\mathbb{N}$. *Then* $\displaystyle\sum_{\substack{n\le x\\(n,m)=1}}\eta(n)g\left(\frac{\log n}{\log x}\right)=$

$$A_m\int_0^1 g(v)dv\log x+O\left(IG\left(1+\sum_{p|m}\frac{\log p}{p}\right)\prod_{p|m}\left(1+\frac{1}{p}\right)\right)$$

where $\displaystyle A_m=\frac{\phi(m)}{m}\prod_{p\nmid m}(1+\eta(p))\left(1-\frac{1}{p}\right)$. *We also have*

$$A_m\ll\phi(m)/m.$$

- Then $\displaystyle\sum_{\substack{n\le x\\(n,m)=1}} \eta(n) g\left(\frac{\log n}{\log x}\right) =$

$$A_m \int_0^1 g(v)dv \log x + O\left(IG\left(1+\sum_{p|m}\frac{\log p}{p}\right)\prod_{p|m}\left(1+\frac{1}{p}\right)\right)$$

- Then $\sum_{\substack{n\le x\\(n,m)=1}} \eta(n)g\left(\frac{\log n}{\log x}\right) =$

$$A_m\int_0^1 g(v)dv\log x + O\left(lG\left(1+\sum_{p|m}\frac{\log p}{p}\right)\prod_{p|m}\left(1+\frac{1}{p}\right)\right)$$

- Although there is nothing very deep in this, the generality creates a lot of detail.

- Then $\displaystyle\sum_{\substack{n\le x\\(n,m)=1}} \eta(n)g\left(\frac{\log n}{\log x}\right) =$

$$A_m\int_0^1 g(v)dv\log x + O\left(lG\left(1+\sum_{p|m}\frac{\log p}{p}\right)\prod_{p|m}\left(1+\frac{1}{p}\right)\right)$$

- Although there is nothing very deep in this, the generality creates a lot of detail.
- We proceed first to look at the special case when g is identically 1. Of course $\eta(n)$ is itself fairly general, but it is close to $1/n$, and we use this. The fact that the support is just the squarefree numbers is a further complication.

- We extend η to a totally multiplicative function $\eta^*(n)$ by

$$\eta^*(p^k) = \eta(p)^k.$$

- We extend η to a totally multiplicative function $\eta^*(n)$ by

$$\eta^*(p^k) = \eta(p)^k.$$

- Now we compare $\eta^*(n)$ with the function $1/n$.

- We extend η to a totally multiplicative function $\eta^*(n)$ by
$$\eta^*(p^k) = \eta(p)^k.$$
- Now we compare $\eta^*(n)$ with the function $1/n$.
- To this end let ρ be the multiplicative function with
$$\rho(p^k) = \eta(p)^{k-1}\big(\eta(p) - 1/p\big) \quad (k > 0).$$

- We extend η to a totally multiplicative function $\eta^*(n)$ by
$$\eta^*(p^k) = \eta(p)^k.$$
- Now we compare $\eta^*(n)$ with the function $1/n$.
- To this end let ρ be the multiplicative function with
$$\rho(p^k) = \eta(p)^{k-1}\big(\eta(p) - 1/p\big) \quad (k > 0).$$
- Then, for some positive constant C_1, $|\rho(p^k)| \le \dfrac{C_1^k}{p^{k+1}}$, and

- We extend η to a totally multiplicative function $\eta^*(n)$ by
$$\eta^*(p^k) = \eta(p)^k.$$
- Now we compare $\eta^*(n)$ with the function $1/n$.
- To this end let ρ be the multiplicative function with
$$\rho(p^k) = \eta(p)^{k-1}\big(\eta(p) - 1/p\big) \quad (k > 0).$$
- Then, for some positive constant C_1, $|\rho(p^k)| \le \dfrac{C_1^k}{p^{k+1}}$, and
- $\displaystyle\sum_{u=0}^{k} \rho(p^u)p^{u-k} = \sum_{u=0}^{k} \eta(p)^u p^{u-k} - \sum_{u=1}^{k} \eta(p)^{u-1} p^{u-1-k} =$ $\eta^*(p^k)$. Thus $\displaystyle\eta^*(n) = \sum_{v|n} v^{-1}\rho(n/v)$.

- We extend η to a totally multiplicative function $\eta^*(n)$ by
$$\eta^*(p^k) = \eta(p)^k.$$
- Now we compare $\eta^*(n)$ with the function $1/n$.
- To this end let ρ be the multiplicative function with
$$\rho(p^k) = \eta(p)^{k-1}\big(\eta(p) - 1/p\big) \quad (k > 0).$$
- Then, for some positive constant C_1, $|\rho(p^k)| \le \dfrac{C_1^k}{p^{k+1}}$, and
- $\displaystyle\sum_{u=0}^{k} \rho(p^u)p^{u-k} = \sum_{u=0}^{k} \eta(p)^u p^{u-k} - \sum_{u=1}^{k} \eta(p)^{u-1} p^{u-1-k} = \eta^*(p^k)$. Thus $\displaystyle\eta^*(n) = \sum_{v|n} v^{-1}\rho(n/v)$.
- We now use the "Rankin trick" to estimate $\displaystyle\sum_{w>y} |\rho(w)|$.

- We extend η to a totally multiplicative function $\eta^*(n)$ by

$$\eta^*(p^k) = \eta(p)^k.$$

- Now we compare $\eta^*(n)$ with the function $1/n$.
- To this end let ρ be the multiplicative function with

$$\rho(p^k) = \eta(p)^{k-1}\big(\eta(p) - 1/p\big) \quad (k > 0).$$

- Then, for some positive constant C_1, $|\rho(p^k)| \le \dfrac{C_1^k}{p^{k+1}}$, and
- $\displaystyle\sum_{u=0}^{k} \rho(p^u)p^{u-k} = \sum_{u=0}^{k} \eta(p)^u p^{u-k} - \sum_{u=1}^{k} \eta(p)^{u-1} p^{u-1-k} =$ $\eta^*(p^k)$. Thus $\displaystyle\eta^*(n) = \sum_{v|n} v^{-1}\rho(n/v)$.
- We now use the "Rankin trick" to estimate $\displaystyle\sum_{w>y} |\rho(w)|$.
- Let $0 < \tau < 1$. Then $\displaystyle\sum_{w>y} |\rho(w)| \le y^{-\tau} \sum_{w=1}^{\infty} w^{\tau}|\rho(w)|$.

- $|\rho(p^k)| \le \dfrac{C_1^k}{p^{k+1}}, \ \displaystyle\sum_{w>y} |\rho(w)| \le y^{-\tau} \sum_{w=1}^{\infty} w^{\tau} |\rho(w)|$

- $|\rho(p^k)| \le \dfrac{C_1^k}{p^{k+1}}$, $\displaystyle\sum_{w>y} |\rho(w)| \le y^{-\tau} \sum_{w=1}^{\infty} w^{\tau}|\rho(w)|$
- The sum here converges because

$$\prod_p \left(1 + \sum_{k=1}^{\infty} p^{k\tau}|\rho(p^k)|\right) \ll \prod_p \left(1 + \sum_{k=1}^{\infty} p^{k\tau-k-1} C_1^k\right).$$

- $|\rho(p^k)| \le \frac{C_1^k}{p^{k+1}}$, $\sum_{w>y} |\rho(w)| \le y^{-\tau} \sum_{w=1}^{\infty} w^\tau |\rho(w)|$
- The sum here converges because

$$\prod_p \Big(1 + \sum_{k=1}^{\infty} p^{k\tau} |\rho(p^k)|\Big) \ll \prod_p \Big(1 + \sum_{k=1}^{\infty} p^{k\tau - k - 1} C_1^k\Big).$$

- Hence $\sum_{\substack{z \le y \\ (z,m)=1}} \rho(z) = D(m) + O(y^{-\tau})$ where $D(m) =$

$$\prod_{p \nmid m} \Big(1 + \sum_{k=1}^{\infty} \eta(p)^{k-1} \big(\eta(p) - 1/p\big)\Big) = \prod_{p \nmid m} \frac{1 - 1/p}{1 - \eta(p)}.$$

- $|\rho(p^k)| \le \dfrac{C_1^k}{p^{k+1}}, \ \sum_{w>y} |\rho(w)| \le y^{-\tau} \sum_{w=1}^{\infty} w^\tau |\rho(w)|$
- The sum here converges because

$$\prod_p \Big(1 + \sum_{k=1}^{\infty} p^{k\tau} |\rho(p^k)|\Big) \ll \prod_p \Big(1 + \sum_{k=1}^{\infty} p^{k\tau-k-1} C_1^k\Big).$$

- Hence $\sum_{\substack{z\le y\\(z,m)=1}} \rho(z) = D(m) + O(y^{-\tau})$ where $D(m) =$

$$\prod_{p\nmid m} \Big(1 + \sum_{k=1}^{\infty} \eta(p)^{k-1} \big(\eta(p) - 1/p\big)\Big) = \prod_{p\nmid m} \frac{1-1/p}{1-\eta(p)}.$$

- Therefore $\sum_{\substack{v\le x\\(v,m)=1}} \eta^*(v) = \sum_{\substack{w\le x\\(w,m)=1}} \dfrac{1}{w} \sum_{\substack{z\le x/w\\(z,m)=1}} \rho(z) =$

$$\sum_{\substack{w\le x\\(w,m)=1}} \frac{1}{w} \big(D(m) + O(w^\tau x^{-\tau})\big) = \sum_{\substack{w\le x\\(w,m)=1}} \frac{D(m)}{w} + O(1).$$

- Therefore $\displaystyle\sum_{\substack{v\le x\\(v,m)=1}} \eta^*(v) = \sum_{\substack{w\le x\\(w,m)=1}} \frac{D(m)}{w} + O(1).$

- Therefore $\displaystyle\sum_{\substack{v\le x\\(v,m)=1}} \eta^*(v) = \sum_{\substack{w\le x\\(w,m)=1}} \frac{D(m)}{w} + O(1).$
- The sum here is
$$\sum_{v|m}\frac{\mu(v)}{v}\sum_{u\le x/v}\frac{1}{u} = \sum_{v|m}\frac{\mu(v)}{v}\big(\log(x/v) + C_0 + O(v/x)\big)$$
$$= \frac{\phi(m)}{m}(\log x + C_0) - \sum_{v|m}\frac{\mu(v)\log v}{v} + O\big(d(m)/x\big).$$

- Therefore $\displaystyle\sum_{\substack{v\le x\\(v,m)=1}} \eta^*(v) = \sum_{\substack{w\le x\\(w,m)=1}} \frac{D(m)}{w} + O(1).$
- The sum here is

$$\sum_{v|m}\frac{\mu(v)}{v}\sum_{u\le x/v}\frac{1}{u} = \sum_{v|m}\frac{\mu(v)}{v}\big(\log(x/v)+C_0+O(v/x)\big)$$

$$=\frac{\phi(m)}{m}(\log x + C_0) - \sum_{v|m}\frac{\mu(v)\log v}{v} + O\big(d(m)/x\big).$$

- Hence $\displaystyle\sum_{\substack{v\le x\\(v,m)=1}} \eta^*(v) = \frac{\phi(m)}{m}D(m)\log x$

$$+O\left(\frac{d(m)}{x}+\left(1+\sum_{p|m}\frac{\log p}{p}\right)\prod_{p|m}\left(1+\frac{1}{p}\right)\right)$$

- Hence $$\sum_{\substack{v\le x\\(v,m)=1}} \eta^*(v) = \frac{\phi(m)}{m} D(m)\log x$$
$$+O\left(\frac{d(m)}{x} + \left(1+\sum_{p|m}\frac{\log p}{p}\right)\prod_{p|m}\left(1+\frac{1}{p}\right)\right)$$

- Hence
$$\sum_{\substack{v\le x\\(v,m)=1}} \eta^*(v) = \frac{\phi(m)}{m}D(m)\log x$$
$$+O\left(\frac{d(m)}{x} + \left(1+\sum_{p|m}\frac{\log p}{p}\right)\prod_{p|m}\left(1+\frac{1}{p}\right)\right)$$
- Thus
$$\sum_{\substack{n\le x\\(n,m)=1}} \eta(n) = \sum_{\substack{n\le x\\(n,m)=1}} \mu(n)^2\eta^*(n) =$$
$$\sum_{\substack{u\le\sqrt{x}\\(u,m)=1}} \mu(u)\eta^*(u)^2 \sum_{\substack{v\le x/u^2\\(v,m)=1}} \eta^*(v) = D_1(m)\log x$$
$$+O\left(\frac{d(m)}{x} + \left(1+\sum_{p|m}\frac{\log p}{p}\right)\prod_{p|m}\left(1+\frac{1}{p}\right)\right)$$

where $D_1(m) = \frac{\phi(m)}{m}D(m)\sum_{\substack{u=1\\(u,m)=1}}^{\infty} \mu(u)\eta^*(u)^2$

- Thus $$\sum_{\substack{n\le x\\(n,m)=1}} \eta(n) = D_1(m)\log x$$
$$+O\left(\frac{d(m)}{x} + \Big(1+\sum_{p|m}\frac{\log p}{p}\Big)\prod_{p|m}\Big(1+\frac{1}{p}\Big)\right)$$
where $$D_1(m) = \frac{\phi(m)}{m}D(m)\sum_{\substack{u=1\\(u,m)=1}}^{\infty} \mu(u)\eta^*(u)^2 =$$
$$\frac{\phi(m)}{m}\prod_{p\nmid m}\left(1-\eta(p)^2\right)(1-1/p)(1-\eta(p))^{-1} = A_m$$

- Thus $\displaystyle\sum_{\substack{n\le x\\(n,m)=1}} \eta(n) = D_1(m)\log x$

$$+O\left(\frac{d(m)}{x} + \left(1+\sum_{p|m}\frac{\log p}{p}\right)\prod_{p|m}\left(1+\frac{1}{p}\right)\right)$$

where $\displaystyle D_1(m) = \frac{\phi(m)}{m}D(m)\sum_{\substack{u=1\\(u,m)=1}}^{\infty} \mu(u)\eta^*(u)^2 =$

$$\frac{\phi(m)}{m}\prod_{p\nmid m}\left(1-\eta(p)^2\right)(1-1/p)(1-\eta(p))^{-1} = A_m$$

- Now we apply this to general g by partial summation.

- Now we apply this to general $g \in \mathcal{G}(l, G)$.

- Now we apply this to general $g \in \mathcal{G}(I, G)$.
- Let $E(x) = \sum_{\substack{n \le x \\ (n,m)=1}} \eta(n) - A_m \log x$ and choose a_j as in the definition of $\mathcal{G}(I, G)$. When $x^{a_{j-1}} < n \le x^{a_j}$, $g\left(\frac{\log n}{\log x}\right) = g_-(a_j) - \int_{\frac{\log n}{\log x}}^{a_j} g'(v)dv$ except when $n = x^{a_j}$ when the two sides differ by $\ll G$.

- Now we apply this to general $g \in \mathcal{G}(I, G)$.
- Let $E(x) = \sum_{\substack{n\le x\\(n,m)=1}} \eta(n) - A_m \log x$ and choose a_j as in the definition of $\mathcal{G}(I, G)$. When $x^{a_{j-1}} < n \le x^{a_j}$,
$$g\left(\frac{\log n}{\log x}\right) = g_-(a_j) - \int_{\frac{\log n}{\log x}}^{a_j} g'(v)dv \text{ except when } n = x^{a_j}$$
when the two sides differ by $\ll G$.
- Multiply by $\eta(n)$, sum over $n \in (x^{a_{j-1}}, x^{a_j}]$, interchange the order of summation and integration and apply E to get
$$\big(A_m(\log x)(a_j - a_{j-1}) + E(x^{a_j}) - E(x^{a_{j-1}})\big)g_-(a_j) + O(G)$$
$$-\int_{a_{j-1}}^{a_j} \big(A_m(\log x)(v - a_{j-1}) + E(x^v) - E(x^{a_{j-1}})\big)g'(v)dv.$$

- Now we apply this to general $g \in \mathcal{G}(I, G)$.
- Let $E(x) = \sum_{\substack{n \le x\\ (n,m)=1}} \eta(n) - A_m \log x$ and choose a_j as in the definition of $\mathcal{G}(I, G)$. When $x^{a_{j-1}} < n \le x^{a_j}$,
$$g\left(\frac{\log n}{\log x}\right) = g_-(a_j) - \int_{\frac{\log n}{\log x}}^{a_j} g'(v)dv \text{ except when } n = x^{a_j}$$
when the two sides differ by $\ll G$.
- Multiply by $\eta(n)$, sum over $n \in (x^{a_{j-1}}, x^{a_j}]$, interchange the order of summation and integration and apply E to get
$$\big(A_m(\log x)(a_j - a_{j-1}) + E(x^{a_j}) - E(x^{a_{j-1}})\big)g_-(a_j) + O(G)$$
$$- \int_{a_{j-1}}^{a_j} \big(A_m(\log x)(v - a_{j-1}) + E(x^v) - E(x^{a_{j-1}})\big)g'(v)dv.$$
- Integrate main term by parts to give $\int_{a_{j-1}}^{a_j} A_m(\log x)g(v)dv$ which on summing over j gives the main term.

- Now we apply this to general $g \in \mathcal{G}(I, G)$.
- Let $E(x) = \sum_{\substack{n\le x\\(n,m)=1}} \eta(n) - A_m \log x$ and choose a_j as in the definition of $\mathcal{G}(I, G)$. When $x^{a_{j-1}} < n \le x^{a_j}$,
$$g\left(\frac{\log n}{\log x}\right) = g_-(a_j) - \int_{\frac{\log n}{\log x}}^{a_j} g'(v)dv$$ except when $n = x^{a_j}$ when the two sides differ by $\ll G$.
- Multiply by $\eta(n)$, sum over $n \in (x^{a_{j-1}}, x^{a_j}]$, interchange the order of summation and integration and apply E to get
$$\big(A_m(\log x)(a_j - a_{j-1}) + E(x^{a_j}) - E(x^{a_{j-1}})\big)g_-(a_j) + O(G)$$
$$-\int_{a_{j-1}}^{a_j} \big(A_m(\log x)(v - a_{j-1}) + E(x^v) - E(x^{a_{j-1}})\big)g'(v)dv.$$
- Integrate main term by parts to give $\int_{a_{j-1}}^{a_j} A_m(\log x)g(v)dv$ which on summing over j gives the main term.
- Insert the bound for E from earlier and sum over j.

- We now complete the proof of Maynard's theorem.

- We now complete the proof of Maynard's theorem.
- We finally assume that $\kappa(\mathbf{r}) = f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R}\right)$ for some f in $\mathcal{F}$.

- We now complete the proof of Maynard's theorem.
- We finally assume that $\kappa(\mathbf{r}) = f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R}\right)$ for some f in $\mathcal{F}$.
- To simplify some of the formulæ we then extend the definition of f to $[0,1]^k$ by taking f to be 0 outside $\mathcal{R}$.

- We now complete the proof of Maynard's theorem.
- We finally assume that $\kappa(\mathbf{r}) = f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R}\right)$ for some f in $\mathcal{F}$.
- To simplify some of the formulæ we then extend the definition of f to $[0,1]^k$ by taking f to be 0 outside $\mathcal{R}$.
- Again we concentrate on S_j rather than T.

- We now complete the proof of Maynard's theorem.
- We finally assume that $\kappa(\mathbf{r}) = f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right)$ for some f in $\mathcal{F}$.
- To simplify some of the formulæ we then extend the definition of f to $[0,1]^k$ by taking f to be 0 outside $\mathcal{R}$.
- Again we concentrate on S_j rather than T.
- Recall that $\kappa_j(\mathbf{r}) = 0$ unless $r_j = 1$, $(r,q) = 1$ and r is squarefree, in which case $\kappa_j(\mathbf{r}) = \sum_{t_j} \frac{\mu(t_j)^2}{\phi(t_j)} \times$

$$f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_{j-1}}{\log R}, \frac{\log t_j}{\log R}, \frac{\log r_{j+1}}{\log R}, \dots, \frac{\log r_k}{\log R}\right) + O\left(\frac{F\phi(q)\log R}{qQ}\right)$$

where $\mathbf{r}' = r_1, \dots, r_{j-1}, t_j, r_{j+1}, \dots, r_k$.

- We now complete the proof of Maynard's theorem.
- We finally assume that $\kappa(\mathbf{r}) = f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R}\right)$ for some f in $\mathcal{F}$.
- To simplify some of the formulæ we then extend the definition of f to $[0,1]^k$ by taking f to be 0 outside $\mathcal{R}$.
- Again we concentrate on S_j rather than T.
- Recall that $\kappa_j(\mathbf{r}) = 0$ unless $r_j = 1$, $(r,q) = 1$ and r is squarefree, in which case $\kappa_j(\mathbf{r}) = \sum_{t_j} \frac{\mu(t_j)^2}{\phi(t_j)} \times$

$$f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_{j-1}}{\log R}, \frac{\log t_j}{\log R}, \frac{\log r_{j+1}}{\log R}, \ldots, \frac{\log r_k}{\log R}\right) + O\left(\frac{F\phi(q)\log R}{qQ}\right)$$

where $\mathbf{r}' = r_1, \ldots, r_{j-1}, t_j, r_{j+1}, \ldots, r_k$.
- Thus $K_j \ll F\frac{\phi(q)}{q}\log R$.

- $\kappa_j(\mathbf{r}) = 0$ unless $r_j = 1$, $(r, q) = 1$ and r is squarefree, in which case $\kappa_j(\mathbf{r}) = \sum_{t_j} \frac{\mu(t_j)^2}{\phi(t_j)} \times$

$$f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_{j-1}}{\log R}, \frac{\log t_j}{\log R}, \frac{\log r_{j+1}}{\log R}, \dots, \frac{\log r_k}{\log R}\right) + O\left(\frac{F\phi(q)\log R}{qQ}\right)$$

and $K_j \ll F\dfrac{\phi(q)}{q}\log R$.

- $\kappa_j(\mathbf{r}) = 0$ unless $r_j = 1$, $(r, q) = 1$ and r is squarefree, in which case $\kappa_j(\mathbf{r}) = \sum_{t_j} \frac{\mu(t_j)^2}{\phi(t_j)} \times$

$$f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_{j-1}}{\log R}, \frac{\log t_j}{\log R}, \frac{\log r_{j+1}}{\log R}, \dots, \frac{\log r_k}{\log R}\right) + O\left(\frac{F\phi(q)\log R}{qQ}\right)$$

  and $K_j \ll F\dfrac{\phi(q)}{q}\log R$.

- Thus, by the last lemma, with $\eta(p) = 1/(p-1)$ and $m = qr$, when $r_j = 1$, $(r, q) = 1$ and r is squarefree

$$\kappa_j(\mathbf{r}) = (\log R)\frac{\phi(qr)}{qr}f_j(\mathbf{r}) + O\left(\frac{F\phi(q)\log R}{qQ}\right)$$

  where $f_j(\mathbf{r}) =$

$$\int_0^1 f\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_{j-1}}{\log R}, u_j, \frac{\log r_{j-1}}{\log R}, \dots, \frac{\log r_k}{\log R}\right) du_j.$$

- We have $\kappa_j(\mathbf{r}) = (\log R)\dfrac{\phi(qr)}{qr} f_j(\mathbf{r}) + O\left(\dfrac{F\phi(q)\log R}{qQ}\right)$ where $f_j(\mathbf{r}) =$

$$\int_0^1 f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_{j-1}}{\log R}, u_j, \frac{\log r_{j-1}}{\log R}, \ldots, \frac{\log r_k}{\log R}\right) du_j.$$

- We have $\kappa_j(\mathbf{r}) = (\log R)\dfrac{\phi(qr)}{qr} f_j(\mathbf{r}) + O\left(\dfrac{F\phi(q)\log R}{qQ}\right)$ where $f_j(\mathbf{r}) =$

$$\int_0^1 f\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_{j-1}}{\log R}, u_j, \frac{\log r_{j-1}}{\log R}, \ldots, \frac{\log r_k}{\log R}\right) du_j.$$

- Thus, by Lemma 8, $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2 \log N} \times$

$$\sum_{\substack{\mathbf{r} \\ (r,q)=1}}^{j} \frac{\mu(r)^2\phi(r)^2}{\phi_2(r)r^2} f_j(\mathbf{r})^2 + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$

- We have $\kappa_j(\mathbf{r}) = (\log R)\dfrac{\phi(qr)}{qr} f_j(\mathbf{r}) + O\left(\dfrac{F\phi(q)\log R}{qQ}\right)$ where $f_j(\mathbf{r}) =$

$$\int_0^1 f\left(\frac{\log r_1}{\log R},\dots,\frac{\log r_{j-1}}{\log R},u_j,\frac{\log r_{j-1}}{\log R},\dots,\frac{\log r_k}{\log R}\right)du_j.$$

- Thus, by Lemma 8, $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2\log N}\times$

$$\sum_{\substack{\mathbf{r}\\(r,q)=1}}^{j} \frac{\mu(r)^2\phi(r)^2}{\phi_2(r)r^2} f_j(\mathbf{r})^2 + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$

- We will repeatedly use, without further comment, that if $\tau(p) \ll p^{-2}$, then we have $\displaystyle\prod_{p>Q}(1+\tau(p)) = 1 + O(1/Q)$ and so such products can be replaced by 1 in the analysis. We have $\dfrac{\phi(r)^2}{\phi_2(r)r} = \displaystyle\prod_{p|r}\frac{(p-1)^2}{(p-1)^2-1}$ and each prime factor of r exceeds Q, so this is $1 + O(Q^{-1})$.

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2 \log N} \times$

$$\sum_{\substack{\mathbf{r} \\ (r,q)=1}}^{j} \frac{\mu(r)^2}{r} f_j(\mathbf{r})^2 + O\left(\frac{F^2 \phi(q)^k N(\log R)^k}{q^{k+1} Q} \right).$$

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2\log N}\times$
$$\sum_{\substack{\mathbf{r}\\(r,q)=1}}^{j} \frac{\mu(r)^2}{r} f_j(\mathbf{r})^2 + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$
- As r is squarefree, the general arithmetical factor in the sum can be rewritten as $\prod_{i=1}^k \frac{\mu(r_i)^2}{r_i}$ provided that the sum over $\mathbf{r}$ is restricted to $\mathbf{r}$ with $(r_u, r_v)=1$ when $u \neq v$.

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2 \log N} \times$

$$\sum_{\substack{\mathbf{r} \\ (r,q)=1}}^{j} \frac{\mu(r)^2}{r} f_j(\mathbf{r})^2 + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$

- As r is squarefree, the general arithmetical factor in the sum can be rewritten as $\prod_{i=1}^{k} \frac{\mu(r_i)^2}{r_i}$ provided that the sum over $\mathbf{r}$ is restricted to $\mathbf{r}$ with $(r_u, r_v){=}1$ when $u \neq v$.
- If we add in any $(r_u, r_v) > 1$, the extra $\mathbf{r}$ have a prime $p > Q$ such that $p|r_u$ and $p|r_v$ for some $u \neq v$.

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2 \log N} \times$
$$\sum_{\substack{\mathbf{r} \\ (r,q)=1}}^{j} \frac{\mu(r)^2}{r} f_j(\mathbf{r})^2 + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$
- As r is squarefree, the general arithmetical factor in the sum can be rewritten as $\prod_{i=1}^{k} \frac{\mu(r_i)^2}{r_i}$ provided that the sum over $\mathbf{r}$ is restricted to $\mathbf{r}$ with $(r_u, r_v){=}1$ when $u \neq v$.
- If we add in any $(r_u, r_v) > 1$, the extra $\mathbf{r}$ have a prime $p > Q$ such that $p|r_u$ and $p|r_v$ for some $u \neq v$.
- Therefore the total error introduced is $\ll$
$$\frac{\phi(q)N\log^2 R}{q^2 \log N} \sum_{p>Q} \frac{F^2}{p^2} \left(\sum_{n<R} \frac{1}{n}\right)^{k-1} \ll \frac{F^2\phi(q)^k N \log^k R}{q^{k+1}Q}$$

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2\log N}\times$
$$\sum_{\substack{\mathbf{r}\\(r,q)=1}}^{j} \frac{\mu(r)^2}{r} f_j(\mathbf{r})^2 + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$
- As r is squarefree, the general arithmetical factor in the sum can be rewritten as $\prod_{i=1}^k \frac{\mu(r_i)^2}{r_i}$ provided that the sum over $\mathbf{r}$ is restricted to $\mathbf{r}$ with $(r_u, r_v){=}1$ when $u \neq v$.
- If we add in any $(r_u, r_v) > 1$, the extra $\mathbf{r}$ have a prime $p > Q$ such that $p|r_u$ and $p|r_v$ for some $u \neq v$.
- Therefore the total error introduced is $\ll$
$$\frac{\phi(q)N\log^2 R}{q^2\log N}\sum_{p>Q}\frac{F^2}{p^2}\left(\sum_{n<R}\frac{1}{n}\right)^{k-1} \ll \frac{F^2\phi(q)^k N\log^k R}{q^{k+1}Q}$$
- Thus the sum in the main term can be replaced by
$$\sum_{\substack{\mathbf{r}\\(r,q)=1}}^{j} f_j(\mathbf{r})^2\prod_{i=1}^{k}\frac{\mu(r_i)^2}{r_i}.$$

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2 \log N} \times$

$$\sum_{\substack{\mathbf{r} \\ (r,q)=1}}^{j} f_j(\mathbf{r})^2 \prod_{i=1}^{k} \frac{\mu(r_i)^2}{r_i} + O\left(\frac{F^2 \phi(q)^k N(\log R)^k}{q^{k+1} Q} \right).$$

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2\log N}\times$

$$\sum_{\substack{\mathbf{r}\\(r,q)=1}}^{j} f_j(\mathbf{r})^2 \prod_{i=1}^{k} \frac{\mu(r_i)^2}{r_i} + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$

- Now we apply Lemma 9 to each variable r_i in turn, i.e $k-1$ times, with

$$\eta(p) = \frac{1}{p}$$

and $m = q$.

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2\log N}\times$
$$\sum_{\substack{\mathbf{r}\\(r,q)=1}}^{j} f_j(\mathbf{r})^2 \prod_{i=1}^{k} \frac{\mu(r_i)^2}{r_i} + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$
- Now we apply Lemma 9 to each variable r_i in turn, i.e $k-1$ times, with
$$\eta(p) = \frac{1}{p}$$
and $m = q$.
- Each time we obtain a factor $\prod_{p>Q}(1+\eta(p))(1-1/p) = \prod_{p>Q}(1-p^{-2}) = 1+O(1/Q)$. Thus
$$S_j(f) = \frac{\phi(q)^k N(\log R)^{k+1}}{q^{k+1}\log N} I_j + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right)$$
where I_j is as in Theorem 6.

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2 \log N} \times$

$$\sum_{\substack{\mathbf{r} \\ (r,q)=1}}^{j} f_j(\mathbf{r})^2 \prod_{i=1}^{k} \frac{\mu(r_i)^2}{r_i} + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$

- Now we apply Lemma 9 to each variable r_i in turn, i.e $k-1$ times, with

$$\eta(p) = \frac{1}{p}$$

and $m = q$.

- Each time we obtain a factor $\prod_{p>Q}(1+\eta(p))(1-1/p) = \prod_{p>Q}(1-p^{-2}) = 1+O(1/Q)$. Thus

$$S_j(f) = \frac{\phi(q)^k N(\log R)^{k+1}}{q^{k+1}\log N} I_j + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right)$$

where I_j is as in Theorem 6.

- This gives the first part of that theorem.

- Thus $S_j(f) = \frac{\phi(q)N(\log R)^2}{q^2 \log N} \times$

$$\sum_{\substack{\mathbf{r} \\ (r,q)=1}}^{j} f_j(\mathbf{r})^2 \prod_{i=1}^{k} \frac{\mu(r_i)^2}{r_i} + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right).$$

- Now we apply Lemma 9 to each variable r_i in turn, i.e $k-1$ times, with
$$\eta(p) = \frac{1}{p}$$
and $m = q$.
- Each time we obtain a factor $\prod_{p>Q}(1+\eta(p))(1-1/p) = \prod_{p>Q}(1-p^{-2}) = 1+O(1/Q)$. Thus
$$S_j(f) = \frac{\phi(q)^k N(\log R)^{k+1}}{q^{k+1}\log N} I_j + O\left(\frac{F^2\phi(q)^k N(\log R)^k}{q^{k+1}Q}\right)$$
where I_j is as in Theorem 6.
- This gives the first part of that theorem.
- The second part follows in the same way.

Theorem 10 (Maynard)

Suppose that when $k \ge 2$, we take $f \in \mathcal{F}$ and then $I_j = I_j(f)$ and $J = J(f)$ are as in Theorem 6. Let $\rho = \sup_{f\in\mathcal{F}} \frac{\sum_{j=1}^k I_j(f)}{J(f)}$. *Then, for k sufficiently large,* $\rho > \log k - \log\log k - 1$.

-

Theorem 10 (Maynard)

Suppose that when $k \geq 2$, we take $f \in \mathcal{F}$ and then $I_j = I_j(f)$ and $J = J(f)$ are as in Theorem 6. Let $\rho = \sup_{f \in \mathcal{F}} \frac{\sum_{j=1}^{k} I_j(f)}{J(f)}$. *Then, for k sufficiently large,* $\rho > \log k - \log\log k - 1$.

-

Corollary 11 (Zhang)

There are bounded gaps in the sequence of primes.

- This is immediate from Theorems 6, 10 and the fact that there are admissible sets with k elements as provided, for example, by Theorem 3.

Theorem 10 (Maynard)

Suppose that when $k \geq 2$, we take $f \in \mathcal{F}$ and then $I_j = I_j(f)$ and $J = J(f)$ are as in Theorem 6. Let $\rho = \sup_{f \in \mathcal{F}} \frac{\sum_{j=1}^{k} I_j(f)}{J(f)}$. Then, for k sufficiently large, $\rho > \log k - \log\log k - 1$.

-

Corollary 11 (Zhang)

There are bounded gaps in the sequence of primes.

- This is immediate from Theorems 6, 10 and the fact that there are admissible sets with k elements as provided, for example, by Theorem 3.

Corollary 12 (Maynard, Tao)

For each $m \in \mathbb{N}$ we have $\liminf_{n\to\infty}(p_{n+m} - p_n) \ll m^2 e^{4m}$.

-

Corollary 13 (Maynard)

Let $m \in \mathbb{N}$ and let $\mathcal{G} = \{g_1, \ldots, g_l\}$ be a set of l distinct non–negative integers. Let $M(m, l, \mathcal{G})$ be the number of admissible m–tuples contained in $\mathcal{G}$ and let $N(m, l, \mathcal{G})$ be the number of admissible m–tuples $\mathbf{h}$ contained in $\mathcal{G}$ such that there are infinitely many n for which each member of the m–tuple $n + \mathbf{h}$ is prime. Then, for $l > l_0(m)$,

$$l^m \geq M(m, l, \mathcal{G}) \gg_m l^m \text{ and } \frac{N(m, l, \mathcal{G})}{M(m, l, \mathcal{G})} \gg_m 1.$$

-

- de Polignac's conjecture [1849] asserts that every even integer is the difference of infinitely many pairs of primes. That the conjecture holds for a positive proportion of all even integers follows on taking $m = 2$ and $g_j = 2j - 2$ in the previous corollary, for then number of solutions of $g_{j_2} - g_{j_1} = 2d$ is at most l and so there must be $\gg l^2/l = 1$ different differences $g_{j_2} - g_{j_1}$ arising from the admissible pairs counted by $N(2, l, \mathcal{G})$.

Corollary 14

There is an infinite subset $\mathbb{D}$ of $\mathbb{N}$ with positive lower asymptotic density such that for each $d \in \mathbb{D}$ there are infinitely many pairs of primes p_1, p_2 such that $p_2 - p_1 = d$.

- Let $\varpi = \frac{k/\log k}{\log(k/\log k)}$ and ξ be the positive solution to $1 + \xi\varpi = e^{\xi}$.

- Let $\varpi = \frac{k/\log k}{\log(k/\log k)}$ and ξ be the positive solution to $1 + \xi\varpi = e^{\xi}$.
- Then $e^{\xi}/\xi > \varpi$ and, for k sufficiently large, $\log k - \log\log k < \xi < \log k$.

- Let $\varpi = \frac{k/\log k}{\log(k/\log k)}$ and ξ be the positive solution to $1+\xi\varpi = e^{\xi}$.
- Then $e^{\xi}/\xi > \varpi$ and, for k sufficiently large, $\log k - \log\log k < \xi < \log k$.
- Let $g : [0,\infty) \to \mathbb{R}$ be defined by

$$g(y) = \begin{cases} \frac{1}{1+\xi y} & 0 \le y \le \varpi, \\ 0 & \varpi < y. \end{cases}$$

- Let $\varpi = \frac{k/\log k}{\log(k/\log k)}$ and ξ be the positive solution to $1 + \xi\varpi = e^{\xi}$.
- Then $e^{\xi}/\xi > \varpi$ and, for k sufficiently large, $\log k - \log\log k < \xi < \log k$.
- Let $g : [0,\infty) \to \mathbb{R}$ be defined by

$$g(y) = \begin{cases} \frac{1}{1+\xi y} & 0 \le y \le \varpi, \\ 0 & \varpi < y. \end{cases}$$

- We need to compute various integrals which we denote by $\alpha, \beta, \gamma, \tau$ as follows.

$$\alpha = \int_0^\infty g(y)dy = 1, \quad \beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^{\xi}},$$
$$\gamma = \int_0^\infty y g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi^2} + \frac{1}{\xi^2 e^{\xi}},$$
$$\tau = \int_0^\infty y^2 g(y)^2 dy = \frac{\varpi}{\xi^2} - \frac{2}{\xi^2} + \frac{1}{\xi^3} - \frac{1}{\xi^3 e^{\xi}}.$$

- We now take

$$f(\mathbf{t}) = \begin{cases} \prod_{i=1}^{k} g(kt_i) & \mathbf{t} \in \mathcal{R}, \\ 0 & \mathbf{t} \notin \mathcal{R}. \end{cases}$$

- We now take

$$f(\mathbf{t}) = \begin{cases} \prod_{i=1}^{k} g(kt_i) & \mathbf{t} \in \mathcal{R}, \\ 0 & \mathbf{t} \notin \mathcal{R}. \end{cases}$$

- Since f is symmetric we have $I_j(f) = I_k(f)$ for every $j \le k$. Thus $\rho \ge \dfrac{kI_k(f)}{J(f)}$ and we now proceed to estimate $I_k(f)$ and $J(f)$.

- We now take

$$f(\mathbf{t}) = \begin{cases} \prod_{i=1}^{k} g(kt_i) & \mathbf{t} \in \mathcal{R}, \\ 0 & \mathbf{t} \notin \mathcal{R}. \end{cases}$$

- Since f is symmetric we have $I_j(f) = I_k(f)$ for every $j \le k$. Thus $\rho \ge \dfrac{kI_k(f)}{J(f)}$ and we now proceed to estimate $I_k(f)$ and $J(f)$.
- With this choice most of the mass of f is close to the axes. $g(kt) = \frac{1}{1+kt\xi} \sim \frac{1}{tk\log k}$. Thus for $t \gg 1/\big(k(\log k)^{1/2}\big)$ we have $g(kt) \ll (\log k)^{-1/2}$.

- We now take

$$f(\mathbf{t}) = \begin{cases} \prod_{i=1}^{k} g(kt_i) & \mathbf{t} \in \mathcal{R}, \\ 0 & \mathbf{t} \notin \mathcal{R}. \end{cases}$$

- Since f is symmetric we have $I_j(f) = I_k(f)$ for every $j \le k$. Thus $\rho \ge \dfrac{kI_k(f)}{J(f)}$ and we now proceed to estimate $I_k(f)$ and $J(f)$.
- With this choice most of the mass of f is close to the axes. $g(kt) = \frac{1}{1+kt\xi} \sim \frac{1}{tk\log k}$. Thus for $t \gg 1/\big(k(\log k)^{1/2}\big)$ we have $g(kt) \ll (\log k)^{-1/2}$.
- Thus the boundary condition $t_1 + \cdots + t_k \le 1$ on $\mathcal{R}$ is relatively unimportant.

- We now take

$$f(\mathbf{t}) = \begin{cases} \prod_{i=1}^{k} g(kt_i) & \mathbf{t} \in \mathcal{R}, \\ 0 & \mathbf{t} \notin \mathcal{R}. \end{cases}$$

- Since f is symmetric we have $I_j(f) = I_k(f)$ for every $j \le k$. Thus $\rho \ge \dfrac{kI_k(f)}{J(f)}$ and we now proceed to estimate $I_k(f)$ and $J(f)$.
- With this choice most of the mass of f is close to the axes. $g(kt) = \frac{1}{1+kt\xi} \sim \frac{1}{tk \log k}$. Thus for $t \gg 1/\big(k(\log k)^{1/2}\big)$ we have $g(kt) \ll (\log k)^{-1/2}$.
- Thus the boundary condition $t_1 + \cdots + t_k \le 1$ on $\mathcal{R}$ is relatively unimportant.
- Since we are concerned with only a lower bound for ρ, lower and upper bounds for $I_k(f)$ and J respectively will suffice.

- An upper bound for $J(f)$ is easy. We have

$$J(f) \le \int_{[0,\infty)^k} \prod_{i=1}^{k} g(kt_i)^2 d\mathbf{t} = k^{-k}\beta^k.$$

- An upper bound for $J(f)$ is easy. We have

$$J(f) \le \int_{[0,\infty)^k} \prod_{i=1}^{k} g(kt_i)^2 d\mathbf{t} = k^{-k}\beta^k.$$

- We can concentrate on a lower bound for $I_k(f)$. We want to let kt_k have the full range of its support so restrict the $t_1, \dots, t_{k-1}$ to $kt_1 + \cdots + kt_{k-1} \le k - \varpi$.

- An upper bound for $J(f)$ is easy. We have

$$J(f) \le \int_{[0,\infty)^k} \prod_{i=1}^{k} g(kt_i)^2 d\mathbf{t} = k^{-k}\beta^k.$$

- We can concentrate on a lower bound for $I_k(f)$. We want to let kt_k have the full range of its support so restrict the $t_1, \ldots, t_{k-1}$ to $kt_1 + \cdots + kt_{k-1} \le k - \varpi$.
- Then we define $\mathcal{S}$ to be the set of $k-1$–tuples $y_1, \ldots, y_{k-1}$ with $y_i \ge 0$ and $y_1 + \cdots + y_{k-1} \le k - \varpi$.
- Thus $kI_k(f) =$

$$k \int_{\mathcal{R}_{k-1}} \left(\int_0^{1-t_1-\cdots-t_{k-1}} g(kt_k)dt_k \right)^2 \prod_{i=1}^{k-1} g(kt_i)^2 dt_1..t_{k-1}$$

$$\ge k^{-k}\alpha^2 \int_{\mathcal{S}} \prod_{i=1}^{k-1} g(y_i)^2 d\mathbf{y} = k^{-k}\alpha^2\beta^{k-1} - E$$

where $E = \dfrac{\alpha^2}{k^k} \displaystyle\int_{\mathcal{S}^*} \prod_{i=1}^{k-1} g(y_i)^2 d\mathbf{y}$ and $\mathcal{S}^* = [0,\infty)^{k-1} \setminus \mathcal{S}$.

- Thus $kI_k(f) \geq k^{-k}\alpha^2 \int_{\mathcal{S}} \prod_{i=1}^{k-1} g(y_i)^2 d\mathbf{y} = k^{-k}\alpha^2\beta^{k-1} - E$ where $E = \dfrac{\alpha^2}{k^k} \int_{\mathcal{S}^*} \prod_{i=1}^{k-1} g(y_i)^2 d\mathbf{y}$ and $\mathcal{S}^* = [0,\infty)^{k-1} \setminus \mathcal{S}$.

- Thus $kI_k(f) \ge k^{-k}\alpha^2 \int_{\mathcal{S}} \prod_{i=1}^{k-1} g(y_i)^2 d\mathbf{y} = k^{-k}\alpha^2\beta^{k-1} - E$

  where $E = \frac{\alpha^2}{k^k}\int_{\mathcal{S}^*} \prod_{i=1}^{k-1} g(y_i)^2 d\mathbf{y}$ and $\mathcal{S}^* = [0,\infty)^{k-1} \setminus \mathcal{S}$.
- Let $\sigma = \gamma/\beta = \frac{1-\xi^{-1}+\xi^{-1}e^{-\xi}}{1-e^{-\xi}} = 1 - \frac{1}{\xi} + \frac{1}{e^{\xi}-1}$. The condition $\mathbf{y} \in \mathcal{S}^*$ is equivalent to $y_1 + \cdots + y_{k-1} \ge k - \varpi$ and this in turn is equivalent to $\frac{y_1+\cdots+y_{k-1}}{k-1} - \sigma$

$$\ge \frac{k-\varpi-\sigma(k-1)}{k-1} = 1 - \sigma - \frac{\varpi-1}{k-1}.$$

  For k sufficiently large we have

$$(1-\sigma)(k-1) - \varpi + 1 = \frac{1}{\xi}\left(1-\frac{1}{\varpi}\right)(k-1) - \varpi + 1$$
$$= \frac{k}{\xi} + O\left(\frac{k}{(\log k)^2}\right) = \xi^{-1} + O(\xi^{-2}) > 0$$

  and $1 - \sigma - \frac{\varpi-1}{k-1} = \xi^{-1} + O(\xi^{-2})$.

- $\sigma = \gamma/\beta = 1 - \dfrac{1}{\xi} + \dfrac{1}{e^{\xi} - 1}$.
  $\mathbf{y} \in \mathcal{S}^*$ is equivalent to

$$\frac{y_1 + \cdots + y_{k-1}}{k-1} - \sigma \geq 1 - \sigma - \frac{\varpi - 1}{k-1} = \xi^{-1} + O\big(\xi^{-2}\big).$$

- $\sigma = \gamma/\beta = 1 - \dfrac{1}{\xi} + \dfrac{1}{e^{\xi} - 1}$.
  $\mathbf{y} \in \mathcal{S}^*$ is equivalent to

$$\frac{y_1 + \cdots + y_{k-1}}{k-1} - \sigma \geq 1 - \sigma - \frac{\varpi - 1}{k-1} = \xi^{-1} + O\big(\xi^{-2}\big).$$

- Thus if $\mathbf{y} \in \mathcal{S}^*$, then

$$\left(\frac{y_1 + \cdots + y_{k-1}}{k-1} - \sigma\right)^2 \zeta^2 \geq 1$$

where $\zeta = \left(1 - \sigma - \dfrac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$.

- $\sigma = \gamma/\beta = 1 - \frac{1}{\xi} + \frac{1}{e^{\xi} - 1}$.
  $\mathbf{y} \in \mathcal{S}^*$ is equivalent to

$$\frac{y_1 + \cdots + y_{k-1}}{k-1} - \sigma \geq 1 - \sigma - \frac{\varpi - 1}{k-1} = \xi^{-1} + O(\xi^{-2}).$$

- Thus if $\mathbf{y} \in \mathcal{S}^*$, then

$$\left(\frac{y_1 + \cdots + y_{k-1}}{k-1} - \sigma\right)^2 \zeta^2 \geq 1$$

where $\zeta = \left(1 - \sigma - \frac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$.

- Hence $E \leq$

$$\frac{\alpha^2\zeta^2}{k^k} \int_{[0,\infty)^{k-1}} \left(\frac{y_1 + \cdots + y_{k-1}}{k-1} - \sigma\right)^2 \prod_{i=1}^{k-1} g(y_i)^2 d\mathbf{y}.$$

A variant of the "Rankin trick".

- $\sigma = \gamma/\beta$, $\zeta = \left(1 - \sigma - \frac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$, $E \le$

$$\frac{\alpha^2\zeta^2}{k^k} \int_{[0,\infty)^{k-1}} \left(\frac{y_1 + \cdots + y_{k-1}}{k-1} - \sigma\right)^2 \prod_{i=1}^{k-1} g(y_i)^2 d\mathbf{y}.$$

$$\alpha = \int_0^\infty g(y)dy, \quad \beta = \int_0^\infty g(y)^2 dy,$$

$$\gamma = \int_0^\infty y g(y)^2 dy, \quad \tau = \int_0^\infty y^2 g(y)^2 dy.$$

- $\sigma = \gamma/\beta$, $\zeta = \left(1 - \sigma - \frac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$, $E \le$

$$\frac{\alpha^2\zeta^2}{k^k}\int_{[0,\infty)^{k-1}}\left(\frac{y_1+\cdots+y_{k-1}}{k-1}-\sigma\right)^2\prod_{i=1}^{k-1}g(y_i)^2d\mathbf{y}.$$

$$\alpha = \int_0^\infty g(y)dy, \quad \beta = \int_0^\infty g(y)^2dy,$$
$$\gamma = \int_0^\infty yg(y)^2dy, \quad \tau = \int_0^\infty y^2g(y)^2dy.$$

- We now square out $\left(\frac{y_1+\cdots+y_{k-1}}{k-1}-\sigma\right)^2 =$

$$\sum_{1\le i<j\le k-1}\frac{2y_iy_j}{(k-1)^2}+\sum_{i=1}^{k-1}\frac{y_i^2}{(k-1)^2}-\sum_{i=1}^{k-1}\frac{2\sigma y_i}{k-1}+\sigma^2$$

and evaluate this with reference to α, etc. Thus $E \le$

$$\frac{\alpha^2\zeta^2}{k^k}\left(\frac{k-2}{k-1}\gamma^2\beta^{k-3}+\frac{\tau\beta^{k-2}}{k-1}-2\sigma\gamma\beta^{k-2}+\sigma^2\beta^{k-1}\right).$$

- $\sigma = \gamma/\beta$, $\zeta = \left(1 - \sigma - \dfrac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$,

$$\alpha = \int_0^\infty g(y)dy = 1, \quad \beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^\xi},$$

$$\gamma = \int_0^\infty y g(y)^2 dy, \quad \tau = \int_0^\infty y^2 g(y)^2 dy.$$

$$E \le \frac{\alpha^2 \zeta^2 \beta^{k-3}}{k^k} \left(\frac{k-2}{k-1}\gamma^2 + \frac{\tau\beta}{k-1} - 2\sigma\gamma\beta + \sigma^2\beta^2 \right)$$

- $\sigma = \gamma/\beta$, $\zeta = \left(1 - \sigma - \frac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$,

$$\alpha = \int_0^\infty g(y)dy = 1, \quad \beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^\xi},$$

$$\gamma = \int_0^\infty y g(y)^2 dy, \quad \tau = \int_0^\infty y^2 g(y)^2 dy.$$

$$E \le \frac{\alpha^2 \zeta^2 \beta^{k-3}}{k^k} \left(\frac{k-2}{k-1}\gamma^2 + \frac{\tau\beta}{k-1} - 2\sigma\gamma\beta + \sigma^2\beta^2\right)$$

- By definition of σ,

$$E \le \frac{\alpha^2 \zeta^2 \beta^{k-3}(\tau\beta - \gamma^2)}{k^k(k-1)} < \frac{\alpha^2 \zeta^2 \beta^{k-2}\tau}{k^k(k-1)}.$$

- $\sigma = \gamma/\beta$, $\zeta = \left(1 - \sigma - \frac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$,

$$\alpha = \int_0^\infty g(y)dy = 1, \quad \beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^\xi},$$
$$\gamma = \int_0^\infty y g(y)^2 dy, \quad \tau = \int_0^\infty y^2 g(y)^2 dy.$$

$$E \le \frac{\alpha^2 \zeta^2 \beta^{k-3}}{k^k} \left(\frac{k-2}{k-1}\gamma^2 + \frac{\tau\beta}{k-1} - 2\sigma\gamma\beta + \sigma^2\beta^2 \right)$$

- By definition of σ,
$$E \le \frac{\alpha^2\zeta^2\beta^{k-3}(\tau\beta - \gamma^2)}{k^k(k-1)} < \frac{\alpha^2\zeta^2\beta^{k-2}\tau}{k^k(k-1)}.$$
- We showed above that $J(f) \le k^{-k}\beta^k$ and $kI_k(f) \ge k^{-k}\alpha^2\beta^{k-1} - E$

- $\sigma = \gamma/\beta$, $\zeta = \left(1 - \sigma - \frac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$,

$$\alpha = \int_0^\infty g(y)dy = 1, \quad \beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^\xi},$$
$$\gamma = \int_0^\infty y g(y)^2 dy, \quad \tau = \int_0^\infty y^2 g(y)^2 dy.$$

$$E \le \frac{\alpha^2 \zeta^2 \beta^{k-3}}{k^k} \left(\frac{k-2}{k-1}\gamma^2 + \frac{\tau\beta}{k-1} - 2\sigma\gamma\beta + \sigma^2\beta^2 \right)$$

- By definition of σ,
$$E \le \frac{\alpha^2 \zeta^2 \beta^{k-3}(\tau\beta - \gamma^2)}{k^k(k-1)} < \frac{\alpha^2 \zeta^2 \beta^{k-2}\tau}{k^k(k-1)}.$$
- We showed above that $J(f) \le k^{-k}\beta^k$ and $kI_k(f) \ge k^{-k}\alpha^2\beta^{k-1} - E$
- Thus
$$\rho > \beta^{-1}\left(1 - \frac{\zeta^2\tau}{\beta(k-1)}\right).$$

- $\zeta = \left(1 - \sigma - \dfrac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$, $\rho > \beta^{-1} - \frac{\zeta^2\tau}{\beta^2(k-1)}$,

$$\beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^\xi}, \quad \beta^{-1} = \xi + O(\xi k^{-1} \log k),$$

$$\tau = \int_0^\infty y^2 g(y)^2 dy = \frac{\varpi}{\xi^2} - \frac{2}{\xi^2} + \frac{1}{\xi^3} - \frac{1}{\xi^3 e^\xi},$$

- $\zeta = \left(1 - \sigma - \dfrac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$, $\rho > \beta^{-1} - \frac{\zeta^2\tau}{\beta^2(k-1)}$,

$$\beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^\xi}, \quad \beta^{-1} = \xi + O(\xi k^{-1}\log k),$$

$$\tau = \int_0^\infty y^2 g(y)^2 dy = \frac{\varpi}{\xi^2} - \frac{2}{\xi^2} + \frac{1}{\xi^3} - \frac{1}{\xi^3 e^\xi},$$

- $\varpi = \frac{k/\log k}{\log(k/\log k)}$ and ξ is the positive root of $1 + \xi\varpi = e^\xi$, so $\log k - \log\log k < \xi = \log k - \log\log k + O(1)$,

$$\zeta^2 = \xi^2 + O(\xi),\ \tau = \varpi\xi^{-2} + O(\xi^{-2}),\ \frac{1}{k-1} = \frac{1}{k} + O(k^{-2}).$$

- $\zeta = \left(1 - \sigma - \dfrac{\varpi - 1}{k-1}\right)^{-1} = \xi + O(1)$, $\rho > \beta^{-1} - \frac{\zeta^2\tau}{\beta^2(k-1)}$,

$$\beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^\xi}, \quad \beta^{-1} = \xi + O(\xi k^{-1}\log k),$$

$$\tau = \int_0^\infty y^2 g(y)^2 dy = \frac{\varpi}{\xi^2} - \frac{2}{\xi^2} + \frac{1}{\xi^3} - \frac{1}{\xi^3 e^\xi},$$

- $\varpi = \frac{k/\log k}{\log(k/\log k)}$ and ξ is the positive root of $1 + \xi\varpi = e^\xi$, so $\log k - \log\log k < \xi = \log k - \log\log k + O(1)$,

$$\zeta^2 = \xi^2 + O(\xi),\ \tau = \varpi\xi^{-2} + O(\xi^{-2}),\ \frac{1}{k-1} = \frac{1}{k} + O(k^{-2}).$$

- Thus $\dfrac{\zeta^2\tau}{\beta(k-1)} = \big(\xi + O(1)\big)\dfrac{\varpi}{k} = \dfrac{1}{\log k} + O\big(\log^{-2} k\big)$.

- $\zeta = \left(1 - \sigma - \dfrac{\varpi - 1}{k - 1}\right)^{-1} = \xi + O(1)$, $\rho > \beta^{-1} - \frac{\zeta^2\tau}{\beta^2(k-1)}$,

$$\beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^\xi}, \quad \beta^{-1} = \xi + O(\xi k^{-1}\log k),$$

$$\tau = \int_0^\infty y^2 g(y)^2 dy = \frac{\varpi}{\xi^2} - \frac{2}{\xi^2} + \frac{1}{\xi^3} - \frac{1}{\xi^3 e^\xi},$$

- $\varpi = \frac{k/\log k}{\log(k/\log k)}$ and ξ is the positive root of $1 + \xi\varpi = e^\xi$, so $\log k - \log\log k < \xi = \log k - \log\log k + O(1)$,

$$\zeta^2 = \xi^2 + O(\xi),\ \tau = \varpi\xi^{-2} + O(\xi^{-2}),\ \frac{1}{k-1} = \frac{1}{k} + O(k^{-2}).$$

- Thus $\dfrac{\zeta^2\tau}{\beta(k-1)} = (\xi + O(1))\dfrac{\varpi}{k} = \dfrac{1}{\log k} + O\big(\log^{-2} k\big)$.
- Hence, if $k > k_0$, we have $\rho > \beta^{-1}\left(1 - \frac{\zeta^2\tau}{\beta(k-1)}\right)$

$$> \xi\left(1 + O\big(k^{-1}\log k\big)\right)\left(1 - \frac{1}{\log k} + O\big((\log k)^{-2}\big)\right)$$

$$> \log k - \log\log k - 1.$$

- Let $\rho = \sup_{f\in\mathcal{F}} \frac{\sum_{j=1}^k I_j(f)}{J(f)}$. Then, for k sufficiently large, $\rho > \log k - \log\log k - 1$.

- Let $\rho = \sup_{f\in\mathcal{F}} \frac{\sum_{j=1}^{k} I_j(f)}{J(f)}$. Then, for k sufficiently large, $\rho > \log k - \log\log k - 1$.
- This completes the proof of Maynard's second theorem. Applied to his first theorem this gives

$$\sup_{f\in\mathcal{F}} \frac{S(f)}{T(f)} > \left(\frac{\theta}{2} - \delta\right)(\log k - \log\log k - 1).$$

- Let $\rho = \sup_{f\in\mathcal{F}} \dfrac{\sum_{j=1}^{k} I_j(f)}{J(f)}$. Then, for k sufficiently large, $\rho > \log k - \log\log k - 1$.
- This completes the proof of Maynard's second theorem. Applied to his first theorem this gives

$$\sup_{f\in\mathcal{F}} \frac{S(f)}{T(f)} > \left(\frac{\theta}{2} - \delta\right)(\log k - \log\log k - 1).$$

- Thus if the level of distribution $\theta > 0$, then we can choose any large k and any admissible k-tuple and deduce that infinitely often there are bounded gaps in the primes.

- We now prove Corollary 12 (Maynard, Tao).
  For each $m \in \mathbb{N}$ we have $\liminf_{n\to\infty}(p_{n+m} - p_n) \ll m^2 e^{4m}$.

- We now prove Corollary 12 (Maynard, Tao).
For each $m \in \mathbb{N}$ we have $\liminf_{n\to\infty}(p_{n+m} - p_n) \ll m^2 e^{4m}$.
- Let C be chosen so that for every $m \in \mathbb{N}$ we have

$$\frac{Cme^{4m}}{4m + \log m + \log C} > e^{2+4m}.$$

- We now prove Corollary 12 (Maynard, Tao).
For each $m \in \mathbb{N}$ we have $\liminf_{n\to\infty}(p_{n+m} - p_n) \ll m^2 e^{4m}$.
- Let C be chosen so that for every $m \in \mathbb{N}$ we have

$$\frac{Cme^{4m}}{4m + \log m + \log C} > e^{2+4m}.$$

- Then for $k \geq \max(3, Cme^{4m})$ we have $\dfrac{k}{\log k} \geq e^{2+4m}$ and so $\log k - \log\log k - 1 > 4m + 1$.

- We now prove Corollary 12 (Maynard, Tao).
  For each $m \in \mathbb{N}$ we have $\liminf_{n\to\infty}(p_{n+m} - p_n) \ll m^2 e^{4m}$.
- Let C be chosen so that for every $m \in \mathbb{N}$ we have

$$\frac{Cme^{4m}}{4m + \log m + \log C} > e^{2+4m}.$$

- Then for $k \ge \max(3, Cme^{4m})$ we have $\dfrac{k}{\log k} \ge e^{2+4m}$ and so $\log k - \log\log k - 1 > 4m + 1$.
- Thus if C is large enough (≥ 1 should do actually),

$$\left(\frac{1}{4} - \frac{1}{k}\right)(\log k - \log\log k - 1) > m.$$

- We now prove Corollary 12 (Maynard, Tao).
For each $m \in \mathbb{N}$ we have $\liminf_{n\to\infty}(p_{n+m} - p_n) \ll m^2 e^{4m}$.
- Let C be chosen so that for every $m \in \mathbb{N}$ we have

$$\frac{Cme^{4m}}{4m + \log m + \log C} > e^{2+4m}.$$

- Then for $k \ge \max(3, Cme^{4m})$ we have $\dfrac{k}{\log k} \ge e^{2+4m}$ and so $\log k - \log\log k - 1 > 4m + 1$.
- Thus if C is large enough (≥ 1 should do actually),

$$\left(\frac{1}{4} - \frac{1}{k}\right)(\log k - \log\log k - 1) > m.$$

- With level of distribution θ to be $\frac{1}{2}$ and $\delta = \frac{1}{k}$, as in the deduction of Zhang's theorem we see $\rho > m$ and so any admissible k–tuple $\mathbf{h}$ is such that there are infinitely many n such that the k-tuple $n + \mathbf{h}$ contains at least m primes.

- We now prove Corollary 12 (Maynard, Tao).
For each $m \in \mathbb{N}$ we have $\liminf_{n\to\infty}(p_{n+m} - p_n) \ll m^2 e^{4m}$.
- Let C be chosen so that for every $m \in \mathbb{N}$ we have

$$\frac{Cme^{4m}}{4m + \log m + \log C} > e^{2+4m}.$$

- Then for $k \geq \max(3, Cme^{4m})$ we have $\dfrac{k}{\log k} \geq e^{2+4m}$ and so $\log k - \log\log k - 1 > 4m + 1$.
- Thus if C is large enough (≥ 1 should do actually),

$$\left(\frac{1}{4} - \frac{1}{k}\right)(\log k - \log\log k - 1) > m.$$

- With level of distribution θ to be $\frac{1}{2}$ and $\delta = \frac{1}{k}$, as in the deduction of Zhang's theorem we see $\rho > m$ and so any admissible k–tuple $\mathbf{h}$ is such that there are infinitely many n such that the k-tuple $n + \mathbf{h}$ contains at least m primes.
- By Gallagher's Theorem there is a an admissible k–tuple of diameter $\ll k \log k \ll m^2 e^{4m}$.