Chapter 25
Primes in Short Intervals

April 22, 2014

25.1 Preliminaries to the modern theory

A famous unsolved problem concerning prime numbers is the twin prime conjecture, that there are infinitely many pairs of primes which differ by 2. This suggests that there are considerable local oscillations in the primes, and 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. 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.

**Definition 1.** Let $h = h_1, \ldots, h_k$ be a $k$–tuple of distinct non–negative integers and let $\nu_p(h)$ denote the number of different residue classes modulo $p$ among the $h_1, \ldots, h_k$. If $\nu_p(h) < p$ for every $p$, then $h$ is called admissible.

It is clear that if $h$ is inadmissible, then there can only be a finite number of $n$ for which the $n + h_1, \ldots, n + h_k$ are simultaneously prime.
Conjecture 1 (The prime $k$–tuple conjecture). It is conjectured that if $h$ is admissible, then there are infinitely many $n$ such that $n + h_1, \ldots, n + h_k$ are simultaneously prime.

It is useful to establish that there are admissible sets with fairly small largest element. In Theorem 7.16 it is shown that there is a positive constant $C$ such that

$$P(N) = \max_M \sum_{n=M+1}^{M+N} p \mid n \Rightarrow p > N \geq \pi(N) + CN(\log N)^{-2}.$$ 

It is readily seen that if $M$ is chosen to give the maximum, then the $n$ contributing a positive amount to the sum cannot lie in the zero residue to any modulus not exceeding $N$ and that for any $p > N$ the number of $n$ counted is smaller than the number of residue classes modulo $p$. Thus these $n$ form an admissible set. Moreover taking $n - n_0$ where $n_0$ is the smallest such $n$ gives an admissible set which lies in $[0, N]$. One also has the following theorem.

**Theorem 1.** 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 $p$ forms an admissible set. In particular $h = \{0, p_2 - p_1, \ldots, p_k - p_1\}$ 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 $p_j = q \leq k$. \hfill \Box

One can consider applying the Hardy–Littlewood method to this question. Suppose that

$$h_1 < h_2 < \ldots < h_k.$$

Let

$$R(x; h) = \sum_{n \leq x-h_k}^* (\log(n + h_1)) \ldots (\log(n + h_k))$$

where $\sum^*$ indicates that the sum is restricted to $n$ for which $n+h_1, \ldots, n+h_k$ are simultaneously prime, and let

$$S(\alpha) = \sum_{p \leq N} (\log p) e(\alpha p)$$ \hfill (1)
where \( N = \lfloor x - h_k \rfloor \). Then

\[
R(x, h) = \int_{U^k-1} S(\alpha_2 + \cdots + \alpha_k) S(-\alpha_2) \cdots S(-\alpha_k) e(\alpha \cdot h - (\alpha_2 + \cdots + \alpha_k)h_1) d\alpha.
\]

Suppose that we can replace each \( S(\alpha) \) by its expected approximation when \( \alpha \) is “close” to a rational number with a “small” denominator and the contribution from the remaining \( \alpha \) is relatively “small”. We are deliberately rather imprecise as this is purely speculative. Thus if

\[
P = N^\delta
\]

for some small fixed positive number \( \delta \) we would hope to obtain something of the form

\[
R(x, h) \sim J \sum_{q \leq P} \sum_{a}^* \frac{c_q(b)}{\phi(q)^k} c_q(-a_2) \cdots c_q(-a_k) e\left(\frac{a \cdot h - (a_2 + \cdots + a_k)h_1}{q}\right)
\]

where \( \sum^* \) is over \( a \) modulo \( q \) with \((a_2, \ldots, a_k.q) = 1\), \( b = a_2 + \cdots + a_k \), and

\[
J = \int_{U^k-1} T(\beta_2 + \cdots + \beta_k) T(-\beta_2) \cdots T(-\beta_k) e(\beta \cdot h - (\beta_2 + \cdots + \beta_k)h_1) d\beta
\]

and

\[
T(\beta) = \sum_{n=1}^{N} e(\beta n).
\]

It is believed generally that this should hold.

The number \( J \) is the number of \( m_1, \ldots, m_k \) with \( 0 \leq m_j \leq x \) and \( m_j - h_j = n_1 - h_1 \). Hence \( J = x + O(1) \). Thus it is expected that

\[
R(x; h) \sim x \mathcal{S}(h; P)
\]

where

\[
\mathcal{S}(h; P) = \sum_{q \leq P} f(q; h)
\]
and
\[ f(q; h) = \sum_{a}^* \frac{c_q(b)c_q(-a_2) \cdots c_q(-a_k)e\left(\frac{a_2h_2 + \cdots + a_kh_k - bh_1}{q}\right)}{\phi(q)^k}. \] (2)

It is readily verified that \( f \) is a multiplicative function of \( q \). Moreover when \( q = p^t \) with \( t \geq 2 \), since \( (a_2, \ldots, a_k, q) = 1 \) that 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. Now consider the case \( q = p \). Then \( (a_2, \ldots, a_k, p) = 1 \) holds for all \( a \) with \( 1 \leq a \leq p \) except \( a = \ldots = a_k = p \). If we sum over all \( a \) with \( 1 \leq a_j \leq p - 1 \), then \( \nu_p(h) \) is the number of solutions of \( r_1, r_j \equiv h_j - h_1 \) (mod \( p \)) with \( 1 \leq r_j \leq p - 1 \). Thus \( N = p - \nu_p(h) \). The term with \( a = \ldots = a_k = p \) contributes \((p - 1)^k\) and so
\[ f(p; h) = \frac{(p - \nu_p(h))p^{k-1} - (p - 1)^k}{(p - 1)^k} = (1 - \nu_p(h)/p)(1 - 1/p)^{-k} - 1. \] (3)

When \( p \nmid D = \prod_{1 \leq i < j \leq k} |h_j - h_i| \) we have \( \nu_p(h) = k \). Thus \( f(p; h) \ll p^{-2} \).

Hence \( \mathcal{G}(h; P) \) converges absolutely to \( \mathcal{G}(h) \) as \( P \to \infty \) where
\[ \mathcal{G}(h) = \sum_{q=1}^{\infty} f(q; h) \prod_{p}(1 + f(p; h)) = \prod_{p} \left(1 - \frac{\nu_p(h)}{p}\right) \left(1 - \frac{1}{p}\right)^{-k} \] (4)

and
\[ \mathcal{G}(h) \ll_k (\log \log(3D))^k \ll_k (\log \log(3 \max_j |h_j|))^k. \] (5)

It is clear from this that when the \( h_j \) are distinct, if \( h \) is inadmissible, then \( \mathcal{G}(h) = 0 \). If \( h \) is admissible, then we have \( \nu_p(h) \leq \min(k, p - 1) \) and so \( 1 - \nu_p(h)/p \geq 1/p \) when \( p \leq k \) and is \( \geq 1 - k/p \) when \( p > k \). Thus there is a positive number \( C(k) \) such that, when the \( h_j \) are distinct, \( h \) is admissible if and only if
\[ C(k) < \mathcal{G}(h). \] (6)

This suggests a conjecture.

**Conjecture 2.** Suppose that \( h \) is admissible. Then, as \( x \to \infty \),
\[ R(x; h) \sim x \mathcal{G}(h). \]
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 $\mathcal{S}(\mathbf{h})$ and we can expect that this will arise naturally in the analysis.

We can also deduce from (5) and the next theorem that there is a plentiful supply of admissible $k$–tuples.

**Theorem 2** (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

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

In view of the observation above that if $\mathbf{h} \in \mathcal{H}$ is inadmissible, then $\mathcal{S}(\mathbf{h}) = 0$, 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

$$|f(p; \mathbf{h})| \leq \frac{C_k}{p^2}$$

and otherwise

$$|f(p; \mathbf{h})| \leq \frac{C_k}{p}$$

where $C_k$ is a suitable positive number. Let $D = \prod_{1 \leq i < j \leq k} |h_j - h_i|$, so that $D \leq H^{k(k-1)/2}$. Then

$$|f(q; \mathbf{h})| \leq q^{-2}C_k^{(q)}(D, q) \ll_{\varepsilon} q^{\varepsilon-2}(D, q).$$

For convenience we introduce the parameter $Q \geq 1$ which is at our disposal. Then

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

Hence

$$\sum_{q > Q} |f(q; \mathbf{h})| \ll Q^{\varepsilon-1}H^{\varepsilon}. \quad (7)$$
For convenience we write
\[ g(q; h) = \phi(q)^k f(q; h). \tag{8} \]

The case \( k = 2 \) is somewhat special so we treat that first. By (8) and (3),
\[ g(q; h) = \mu(q)^2 \sum_{a=1 \atop (a,q)=1}^q e(a(h_1 - h_2)/q) \]
and so
\[ \sum_{h \in \mathcal{H}} g(q; h) = \mu(q)^2 \sum_{h_2 \leq H} \sum_{a=1 \atop (a,q)=1}^q \sum_{h_1 \leq H} e(a(h_1 - h_2)/q). \]
The innermost sum is \( \ll ||a/q||^{-1} \) and we have \( \sum_{a=1}^{q^{-1}} ||a/q||^{-1} \ll q \log q \). Thus
\[ \sum_{h \in \mathcal{H}} \sum_{1 < q \leq Q} f(q; h) \ll HQ^\varepsilon. \]
The case \( k = 2 \) of the theorem now follows from (4) and (7) with \( Q = H \).

For the rest of the proof we suppose that \( k \geq 3 \). Crudely, by (3) for any \( h \)
\[ |g(q; h)| \leq g^*(q) \]
where
\[ g^*(q) = \sum_{a \atop (a,q)=1} |c_q(a_2) \ldots c_q(a_k)c_q(-a_2 - \cdots - a_k)| \]
and this is also a multiplicative function of \( q \) (with its support on the square free numbers). Consider the \( k \) numbers \( a_2, \ldots, a_k, -a_2 - \cdots - a_k \). When \( (a,p) = 1 \) at least two of these numbers are not multiples of \( p \). Moreover in \( g^*(p) \) the terms with exactly \( j \) of the \( a_2, \ldots, a_k, a_2 + \cdots + a_k \) divisible by \( p \) contribute \( (p - 1)^j \) and since the \( a_2, \ldots, a_k, a_2 + \cdots + a_k \) are linearly dependent the number of such terms is at most \( \binom{k}{j}(p - 1)^{k-1-j} \). Hence
\[ g^*(p) \leq 2^k(p - 1)^{k-1} \text{ and } g^*(q)\phi(q)^{-k} \ll q^{\varepsilon-1}. \]
Hence
\[ \sum_{h \in \mathcal{H}} \sum_{1 < q \leq Q} f(q; h) - \sum_{h \in [1,H]^k} \sum_{1 < q \leq Q} f(q; h) \ll H^{k-1} \sum_{q \leq Q} q^{\varepsilon-1} \]
and so
\[ \sum_{h \in \mathcal{H}} \sum_{1 < q \leq Q} f(q; h) - \sum_{h \in [1,H]^k} \sum_{1 < q \leq Q} f(q; h) \ll H^{k-1}Q^\varepsilon. \tag{9} \]
Returning to (2) when $q > 1$ at least two of $a_2, \ldots, a_k, -a_2 - \cdots - a_k$ are non-zero modulo $q$. If there are at least two such of the $a_i$, then we pick two and call them $b_2, b_3$. The remaining $a_i$ can be listed in the form $b_1, \ldots, b_k$ so that $-a_2 - \cdots - a_k = -b_2 - b_3 - \cdots - b_k$. If only one of the $a_i$ is non-zero modulo $q$, then call it $b_2$ and take $b_3 = -a_2 - \cdots - a_k$. In that case any one of the other $a_i$ can be rewritten in the form $-b_2 - b_3 - s \pmod{q}$ where $s$ is the sum of the remaining $a_i$.

Thus

$$\sum_{b \in [1, H]^k} g(q; h) \ll H^{k-2} \sum_{b_2=1}^{q-1} \frac{|c_q(b_1)|}{\|b_3/q\|} \sum_{b_2=1}^{q-1} \frac{|c_q(b_2)|}{\|b_2/q\|} \sum_{b \in [1, q]^{k-3}} |c_q(b_4) \cdots c_q(b_k)c_q(b_2 + \cdots + b_k)|$$

where $b = b_1, \ldots, b_k$ and where the summand over $b$ is taken to be $|c_q(b_2+b_3)|$ when $k = 3$. In general this multiple sum does not exceed

$$\phi(q) \left( \sum_{b=1}^{q} |c_q(b)| \right)^{k-3}$$

Since $|c_q(b)| \leq (q, b)$ the sum here is at most

$$\sum_{r|q} r\phi(q/r) \leq d(q)q.$$

Similarly

$$\sum_{b=1}^{q-1} \frac{|c_q(b)|}{\|b/q\|} \leq \sum_{r|q} \sum_{a=1}^{q/r-1} \frac{\|a/(q/r)\|^{-1}}{r} \ll d(q)q \log q.$$

Therefore

$$\sum_{h \in [1, H]^k} \sum_{1 < q \leq Q} f(q, h) \ll H^{k-2}Q^{1+\varepsilon}.$$
2. As usual let \( \nu_p(h) \) denote the number of different residue classes modulo \( p \) represented by the \( h \). Let \( h_j = (2j - 1)^2 \) \( (j = 1, \ldots, k) \). Prove that \( h \) is an admissible set.

3. Call a set \( h \) of distinct non-negative integers \( h_1, \ldots, h_k \) \( sf\)-admissible when there is no prime \( p \) such that every residue class modulo \( p^2 \) contains at least one of them. Let \( S(x; h) \) denote the number of \( n \leq x \) such that \( n + h_1, \ldots, n + h_k \) are simultaneously squarefree.

(i) Let \( f(n) \) denote the characteristic function of the squarefree numbers. Prove that 
\[
S(x; h) = \sum_{n \leq x} f(n + h_1) \ldots f(n + h_k)
\]
and 
\[
f(n) = \sum_{d^2|n} \mu(d).
\]

(ii) Suppose that \( 0 < \delta < 1/(3k) \) and let \( y = x^\delta \) and 
\[
f(n; y) = \sum_{\substack{d \leq y \\ d^2|n}} \mu(d).
\]
Prove that for \( j = 1, \ldots, k \)
\[
S(x; h) = T_j(x; y) + O(x^{1+\varepsilon} y^{-1})
\]
where 
\[
T_j(x; y) = \sum_{n \leq x} f(n + h_1; y) \ldots f(n + h_j; y) f(n + h_{j+1}) \ldots f(n + h_k).
\]

(iii) Given a \( k \)-tuple of positive integers \( d = d_1, \ldots, d_k \) let \( d = d_1 \ldots d_k \) and given another one \( r \) we use \( d|r \) to mean \( d_j|r_j \) \( (j = 1, \ldots, k) \) and \( d^2 \) to mean \( d_1^2, \ldots, d_k^2 \). Write \( n + h \) for the \( k \)-tuple \( n + h_1, \ldots, n + h_k \). Let \( \rho(d) \) denote the number of solutions of \( d^2|n + h \) in \( n \) modulo \( d^2 \). Prove that \( \rho(d) \leq d^2 \) and 
\[
T_k(x; y) = x \sum_{d_1 \leq y, \ldots, d_k \leq y} \frac{\mu(d_1) \ldots \mu(d_k)}{d^2} \rho(d) + O(y^{3k}).
\]

(iv) Let \( \nu_p(h) \) denote the number of different residue classes modulo \( p^2 \) amongst the \( h_1, \ldots, h_k \). Suppose that \( k = 2 \). Prove that 
\[
S(x; h) = x \prod_p \left( 1 - \frac{\nu_p(h)}{p^2} \right) + O(x^{1-\delta}).
\]
4. Given a \( k \)-tuple of positive integers \( \mathbf{d} = d_1, \ldots, d_k \) let \( d = d_1 \ldots d_k \) and given another one \( \mathbf{r} \) we use \( \mathbf{d}|\mathbf{r} \) to mean \( d_j | r_j \) \((j = 1, \ldots, k)\) and \( d^2 \) to mean \( d_1^2, \ldots, d_k^2 \). Write \( n + \mathbf{h} \) for the \( k \)-tuple \( n + h_1, \ldots, n + h_k \). Let \( \rho(\mathbf{d}) \) denote the number of solutions of \( \mathbf{d}|n+\mathbf{h} \) in \( n \) modulo \( d^2 \) and let \( \rho^*(\mathbf{d}) \) denote the number of solutions of \( \mathbf{d}|n+\mathbf{h} \) in \( n \) modulo \( \text{lcm}[d_1, \ldots, d_k]^2 \). Let \( \nu_p(\mathbf{h}) \) denote the number of different residue classes modulo \( p^2 \) amongst the \( h_1, \ldots, h_k \).

(i) Prove that \( \rho(\mathbf{d}) = d^2 \text{lcm}[d_1, \ldots, d_k]^{-2} \rho^*(\mathbf{d}) \) and \( \rho^*(\mathbf{d}) \leq 1 \).

(ii) Prove that
\[
\sum_{\max(d_j) > y} \frac{\mu(d_1) \ldots \mu(d_k)}{d^2} \rho(\mathbf{d}) \ll \sum_{\max(d_j) > y} \frac{\mu(d_1)^2 \ldots \mu(d_k)^2}{[d_1, \ldots, d_k]^2} \ll \sum_{m > y} g_{\omega(m)} \frac{m^2}{m^2} \ll y^{\varepsilon-1}
\]

and deduce that
\[
T_k(x, y) = x \sum_{m=1}^{\infty} \frac{g(m)}{m^2} + O\left(xy^{\varepsilon-1}\right)
\]
where
\[
g(m) = \sum_{[d_1, \ldots, d_k] = m} \mu(d_1) \ldots \mu(d_k) \rho^*(\mathbf{d}).
\]

(iii) Prove that \( \rho(\mathbf{d}) \) is multiplicative, i.e. given \( \mathbf{d}, \mathbf{e} \), define
\[
\mathbf{d}\mathbf{e} = d_1e_1, \ldots, d_ke_k
\]
and deduce that if \( (d, e) = 1 \), then \( \rho(\mathbf{d}\mathbf{e}) = \rho(\mathbf{d})\rho(\mathbf{e}) \).

(iv) Prove that \( g(m) \) is multiplicative and has its support on the squarefree numbers.

(v) Deduce that
\[
\sum_{m=1}^{\infty} g(m) \frac{1}{m^2} = \prod_p \left( 1 + g(p)p^{-2} \right).
\]

(vi) Prove that \( 1 + g(p)p^{-2} = 1 - \nu_p(\mathbf{h})p^{-2} \).

(vii) (Pillai [1936].) Prove that
\[
S(x; h) = x \prod_p \left( 1 - \frac{\nu_p(\mathbf{h})}{p^2} \right) + O(x^{1-\delta})
\]
and hence that if $h$ is $sf$-admissible, then there are infinitely many $n$ such that $n + h_j$ are simultaneously square free for $j = 1, \ldots, k$.

5. Find the minimal diameter of 20-tuples which are $sf$-admissible, i.e. $\max h_j - h_i$ is minimal.

25.2 Maynard’s Theorem

The principal idea is to use artefacts from sieve theory, especially the Selberg sieve, not directly in the form of a sieve but as a means to enhance terms of special interest, particularly those terms with relatively few prime factors. As a preliminary observation consider the starting point for the Selberg upper bound sieve (Chapter 3) in the form

$$\sum_{a \in A} \left( \sum_{q \leq R \atop q \mid a} \lambda_q \right)^2$$

and recall that one is planning to minimise this under the assumptions that $\lambda_1 = 1$ and that

$$A_d = \sum_{a \in A \atop d \mid a} 1$$

can be approximated by an expression of the form

$$\frac{Xg(d)}{d}$$

where $X$ is a good approximation to $A_1$ and $g$ is multiplicative. The minimising choice of $\lambda_q$ is given by

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

where

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

Typically one applies this when the sieve is of dimension $k$, e.g.

$$\sum_{p \leq y} g(p) \frac{\log p}{p} = k \log y + O(1).$$
Under this kind of condition one might expect that

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

and so \( \lambda_q \) could be replaced by

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

Indeed 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 \( \lambda_q \) can be applied quite effectively to any sieving question where the dimension is \( k \).

Let \( 1_{\mathbb{P}} \) denote the characteristic function of the set of primes \( \mathbb{P} \). Then the basic idea is to construct an expression of the form

\[ \sum_{N \leq n \leq 2N} \left( \sum_{j=1}^{k} 1_{\mathbb{P}}(n + h_j) - \rho \right) \left( \sum_{q \leq R} \lambda_q \right)^2 \]

where \( Z = \prod_{i=1}^{k} (n + h_i) \). A wrinkle introduced by Goldston, Pintz and Yildirim is to use a more general \( \lambda_q \) of the form

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

where \( f \) is at our disposal. If one can show that this is positive, then it follows that there are \( n \) such that the number of primes amongst the \( n + h_j \) is at least \( \lfloor \rho \rfloor + 1 \).

Following Maynard [to appear] we will use a more sophisticated version of this. Let \( n + h \) denote the \( k \)-tuple \( n + h_1, \ldots, n + h_k \) and let \( d \) denote the \( k \)-tuple \( d_1, \ldots, d_k \). We generally use the notation that given a \( k \)-tuple \( d \) of positive integers \( d \) denotes \( d_1 \ldots d_k \) and given another one \( r \), then \( d | r \) means that \( d_j | r_j \) for each \( j \). We also use \([d, e]\) to denote the \( k \)-tuple \( \text{lcm}[d_1, e_1], \ldots, \text{lcm}[d_k, e_k] \). First of all we do some initial sieving for small primes so as to simplify some later expressions and a simple way to do this is to restrict our attention to a given residue class \( a \) modulo \( q \) where

\[ q = \prod_{p \leq Q} p, \quad Q = \log \log \log N \] (10)
and $N$ is a large integer parameter. When $h$ is admissible we can suppose that there is an $a$ modulo $q$ such that for $1 \leq j \leq j$ 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. The immediate effect of this can be seen via the heuristic argument of §25.1. If one supposes in addition that $n \equiv a$ modulo $q$, then the singular series takes the shape

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

for large $N$.

Now we consider

$$\sum_{N < n \leq 2N \atop n \equiv a \pmod{q}} \left( \sum_{j=1}^{k} \mathbb{1}_P(n + h_j) - \rho \right) \left( \sum_{d \leq R \atop (d, q) = 1} \lambda(d) \right)^2. \quad (11)$$

In the first instance we ought to consider

$$\lambda(d) = \mu(d) g(d)$$

for some suitable $g$. However we shall be carrying out diagonalisation of quadratic forms in the $\lambda$ and this leads to a natural representation of the $\lambda(d)$, when $d$ is squarefree with $(d, q) = 1$, in the form

$$\lambda(d) = \mu(d) \sum_{d \mid r \atop (r, q) = 1} \frac{\mu(r)^2}{\phi(r)} f \left( \frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R} \right). \quad (12)$$

We further suppose that

$$\text{supp} f = \mathcal{R} = \{ x \in [0, 1]^k : x_1 + \cdots + x_k \leq 1 \}. \quad (13)$$

There are two major tasks to be undertaken. The first is to obtain a good approximation to $(11)$ with $(12)$ for a wide class $\mathcal{F}$ of $f$. In practice this means good approximations $S^*(f)$ and $T^*(f)$ to $S(f)$ and $T(f)$ where

$$S(f) = \sum_{j=1}^{k} S_j(f)$$
with
\[ S_j(f) = S_j = \sum_{N < n \leq 2N} \mathbb{1}_\mathcal{P}(n + h_j) \left( \sum_{d \leq R \atop (d, q) = 1} \lambda(d) \right)^2. \] (14)

\[ T(f) = T = \sum_{N \leq n \leq 2N} \left( \sum_{d \leq R \atop (d, q) = 1} \lambda(d) \right)^2. \] (15)

The second is then to maximise the ratio
\[ \frac{S^*(f)}{T^*(f)} \]
over the class \( \mathcal{F} \). The optimal solution to this latter task is not known, although the former can be carried out for a very wide class, for example for \( f \) for which the partial derivatives are continuous on \( \mathcal{R} \), and even this requirement can be relaxed somewhat.

A major input into the approximation for \( S_j(f) \) will be the Bombieri–Vinogradov theorem \((20.??)\) or a variant thereof. We define the level \( \theta \) of distribution for the prime numbers to be the assumption that for every sufficiently small positive \( \delta \) and every \( A > 0 \) we have
\[ \sum_{q \leq x^\theta} \max \sup \left| \pi(y; q, a) - \frac{\text{li}(y)}{\phi(q)} \right| \ll_{\delta, A} x(\log x)^{-A}. \]

The Bombieri–Vinogradov theorem tells us that \( \theta = \frac{1}{2} \) is permissible. However it is useful to be able to see at once the consequence of the Elliott–Halberstam conjecture \((\theta = 1)\) or some intermediate improvement in the Bombieri–Vinogradov theorem.

Let \( \mathcal{R}_j \) denote the set of \( j-1 \)-tuples \( t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_k \) with \( t \in \mathcal{R} \) 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 \( t^* = t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_k \) with \( t_i \geq 0 \) and \( t_1 + \cdots + t_{j-1} + t_{j+1} + \cdots + t_k \leq 1 \) the function \( f_j^*(t_j) = f(t) \) 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
\[ F = \sup_{t \in \mathcal{R}} |f(t)| + \sum_{j=1}^k \sup_{t \in \mathcal{R}_j} \int_0^1 \left| \frac{\partial f}{\partial t_j}(t) \right| dt_j. \]
Theorem 3 (Maynard). Let \( k \geq 2 \). Suppose that the primes have level of distribution \( \theta \) where \( 0 < \theta \leq 1 \), let \( \delta \) be a sufficiently small positive number and let \( N \) be a large positive integer. Put \( R = N^{\frac{\theta}{2} - \delta} \), define \( Q \) and \( q \) as in (10) and \( f \) and \( R \) as in (13), and assume \( f \in \mathcal{F} \). Let \( h \) be an admissible set and choose \( a \) modulo \( q \) so that for each \( j \) we have \((a + h_j, q) = 1\). Let

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

and

\[
J = \int_{[0,1]^k} f(t)^2 dt
\]

and let \( S(f) \) and \( T(f) \) be as in (14) and (15). Then, as \( N \to \infty \),

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

and

\[
T(f) = \frac{(1 + o(1)) \phi(q)^k N (\log R)^k}{q^{k+1} J} J.
\]

In particular

\[
\frac{S(f)}{T(f)} = (1 + o(1)) \left( \frac{\theta}{2} - \delta \right) \frac{\sum_{j=1}^k I_j}{J}.
\]

The proof of Theorem 3 is divided into several stages. Fortunately the treatments of \( S(f) \) and \( T(f) \) are similar. Initially we do not assume (12) but suppose only that the \( \lambda(d) \) are general real valued functions with support satisfying \( d_1 \ldots d_k = d \leq R \), \((d, q) = 1\) and \( d \) squarefree. Thus it can be supposed that \((d_i, d_j) = 1\) when \( i \neq j \). We begin with the normal diagonalisation process. To this end it is useful to define the multiplicative function \( \phi_2(n) \) by \( \phi_2(p) = p - 2 \) and \( \phi_2(p^t) = 0 \) when \( t \geq 2 \).

Lemma 4. For \( j = 1, \ldots, k \) let

\[
\kappa_j(r) = \mu(r) \phi_2(r) \sum_{d \mid r} \frac{\lambda(d)}{\phi(d)}.
\]
where $\sum^j$ indicates that the summation variable is a $k$-tuple, say $d$, which is restricted by $d_j = 1$, and let

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

Then

$$\lambda(d) = \mu(d)\phi(d) \sum_{d \mid r} j \frac{\kappa_j(r)}{\phi_2(r)}$$  \hspace{1cm} (16)$$

and

$$\lambda(d) = \mu(d) \sum_{d \mid r} \frac{\kappa(r)}{\phi(r)}.$$  \hspace{1cm} (17)$$

Proof. Consider

$$\sum_{d \mid r} j \frac{\kappa_j(r)}{\phi_2(r)}$$

On substituting the definition of $\kappa_j$ this becomes

$$\sum_{d \mid r} j \frac{\mu(r) \sum_{s \mid r} \lambda(s)}{\phi(s)} = \sum_{s \mid r} \frac{\lambda(s)}{\phi(s)} \sum_{d \mid r} \mu(r).$$

The innermost sum is a sum over $r_1, \ldots, r_{j-1}, r_{j+1}, \ldots, r_k$ with $d_1|s_i|d_i$, and the general term is $\mu(r) = \mu(r_1) \cdots \mu(r_{j-1}) \mu(r_{k-1}) \cdots \mu(r_k)$. Thus the sum over $r_i$ is $\mu(d_i) \sum_{t_i|s_i/d_i} \mu(t_i) = 0$ unless $s_i = d_i$ in which case it is $\mu(d_i)$. Thus $d_j = 1$ and

$$\sum_{d \mid r} j \frac{\kappa_j(r)}{\phi_2(r)} = \mu(d) \frac{\lambda(d)}{\phi(d)},$$

which is equivalent to (16).

The inversion formula (17) follows in the same way.

Lemma 5. Let

$$K_j = \max_r |\kappa_j(r)|, \quad K = \max_r |\kappa(r)|.$$
Then for any fixed \( A > 0 \)
\[
S_j = \frac{N}{\phi(q) \log N} \sum_r j \kappa_j(r)^2 + O \left( \frac{K_j^2 \phi(q)^{k-2} N (\log R)^{k-2}}{q^{k-1}Q} \right)
\]
and
\[
T = \frac{N}{q} \sum_r \kappa(r)^2 + O \left( \frac{K^2 N (\log R)^k}{qQ} \right).
\]

**Proof.** We set the pattern with \( S_j \). Not only do we need to substitute \( \kappa_j \) for \( \lambda \) in the main term but we need suitable bounds for the \( \lambda(d) \) in any error terms which arise. Moreover, we need to do so in terms of \( \kappa \) and \( \kappa_j \) rather than \( \lambda \).

We square out and invert the order of summation. Thus
\[
S_j = \sum_{d,e \in \mathbb{N}} \frac{\lambda(d) \lambda(e)}{\gcd(d, e)} \sum_{N < n \leq 2N} 1 \gamma(n + h_j).
\]

We recall that for \( \lambda(d) \neq 0 \) we have \( d \) squarefree and \( (d, q) = 1 \). Therefore \( (d_h, d_v) = 1 \) when \( u \neq v \). Likewise for \( e \). 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 \( n \equiv a \pmod{q} \). Then the innermost sum can be rewritten as
\[
\sum_{N + h_j < p \leq 2N + h_j} \frac{1}{p \equiv h_j - h_i \pmod{d_i, e_i} \pmod{d_j, e_j}}.
\]

By construction \( (a + h_j, q) = 1 \) and \( (h_j - h_i, de) = 1 \) when \( i \neq j \). Let
\[
m = q^{k-1} \prod_{i=1}^{k} \gcd(d_i, e_i),
\]
\[
X = \int_{N + h_j}^{2N + h_j} \frac{dt}{\log t}.
\]
and

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

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

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

By (16), on taking the maximum over \(d\) with \(d_j = 1\) we have

$$\max_{d,d_j=1} |\lambda(d)| \leq \max_{d,d_j=1} \phi(d) \sum_{r \mid (d,q)=1}^j \frac{K_j \mu(r)^2}{\phi_2(r)} = K_j \max_{d} \phi(d) \sum_{s \mid (d,q)=1}^j \frac{\mu(s)^2}{\phi_2(s)}.$$

Thus,

$$\max_{d,d_j=1} |\lambda(d)| \leq K_j \max_{d} \phi(d) \prod_{p \mid d} \left(1 + \frac{1}{p - 2}\right)^{k-1} \ll K_j (\log R)^{k-1}.$$

A concomitant argument shows that

$$\max_{d} |\lambda(d)| \ll K (\log R)^k. \quad (18)$$

Now consider the number of ways that the modulus \(m/q\) can arise in \(E\). This is squarefree. Let \(p\) be a prime factor of \(m/q\). Then \(p\) divides exactly one of the \([d_i, e_i]\). There are \(k - 1\) choices of \(i\) and for any one choice there are 3 possibilities, \(p \mid (d_i, e_i)\); \(p \mid d_i\) and \(p \nmid e_i\); and \(p \nmid d_i\) and \(p \mid e_i\). Thus there are \((3(k - 1))^{\omega(m/q)} \leq (3k)^{\omega(m)}\) possible \(d, e\) which give rise to \(m\). Thus

$$E \ll K_j^2 (\log R)^{2k} \sum_{m \leq qR^2} \mu(m)^2 (3k)^{2\omega(m)} \max_{(b,m)=1} \sup_{x \leq 2N} \left| \pi(x; m, b) - \frac{\text{li}(x)}{\phi(m)} \right|.$$

We have

$$\sum_{m \leq qR^2} \mu(m)^2 (3k)^{2\omega(m)} \max_{(b,m)=1} \sup_{x \leq 2N} \left| \pi(x; m, b) - \frac{\text{li}(x)}{\phi(m)} \right| \ll \sum_{m \leq qR^2} \mu(m)^2 (3k)^{2\omega(m)} Nm^{-1} \ll N(\log N)^{(3k)^2}.$$
Hence, by Bombieri’s theorem

\[ E \ll K_j^2 N (\log N)^{-A}. \]

It remains to deal with the main term for \( S_j \) and 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 \( \phi(m) = \phi(q) \prod_{i \neq j} \phi([d_i, e_i]), \frac{1}{\phi([d_i, e_i])} = \frac{\phi((d_i, e_i))}{\phi(d_i)\phi(e_i)} \) and \( \phi((d_i, e_i)) = \sum_{n_i | d_i, n_i | e_i} \phi_2(n_i) \). Hence

\[ \frac{1}{\phi(m)} = \frac{1}{\phi(q)\phi(d)\phi(e)} \sum_{n|d,n|e} \phi_2(n). \]

We substitute this in the main term and invert the order of summation to obtain

\[ \frac{X}{\phi(q)} \sum_{n} \phi_2(n) \sum_{d,e} \frac{\lambda(d)\lambda(e)}{\phi(d)\phi(e)}. \]

We now take the first step in dealing with the condition \((d_u e_v) = 1\) for \( u \neq v \). We replace it by

\[ \sum_{s_{uv} | d_u, s_{uv} | e_v} \mu(s_{uv}). \]

There are various observations we can make with regard to the \( s_{uv} \). We have \( n_u | d_u \). Thus \((d_v, n_u) = 1\). Hence \((s_{uv}, n_u) = 1\). Likewise \((s_{uv}, n_v) = 1\). Also, when \( w \neq v \), \( s_{uw} | e_w \) and \((e_v, e_w) = 1\). Hence \((s_{uv}, s_{uw}) = 1\). Likewise, when \( w \neq u \), \((s_{uw}, s_{uv}) = 1\) and so in summary

\[ (s_{uv}, n_u) = 1, \quad (s_{uv}, n_v) = 1, \quad (s_{uv}, s_{uw}) = 1, \quad (s_{uw}, s_{uv}) = 1. \quad (19) \]

Thus

\[ \sum_{n} \phi_2(n) \sum_{d,e} \frac{\lambda(d)\lambda(e)}{\phi(d)\phi(e)} \]

\[ = \sum_{n} \phi_2(n) \prod_{s_{uv} | d_u, s_{uv} | e_v} \mu(s_{uv}) \left( \sum_{d} \frac{\lambda(d)}{\phi(d)} \right) \left( \sum_{e} \frac{\lambda(e)}{\phi(e)} \right). \]
where $\sum$ indicates that (19) holds. We now substitute the $\kappa_j$ for the $\lambda$. Thus the above becomes

$$\sum_n \frac{1}{\phi_2(n)} \sum_{s_{uv}} \left( \prod_{u \neq v} \frac{\mu(s_{uv})}{\phi_2(s_{uv})^2} \right) \kappa_j(a)\kappa_j(b)$$

where $a = a_1, \ldots, a_k$, $b = b_1, \ldots, b_k$ and

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

In particular $a = b = ns$ where $s = \prod_{u \neq v} s_{uv}$. Thus the main term is

$$\frac{X}{\phi(q)} \sum_n \frac{1}{\phi_2(n)} \sum_{s_{uv}} \mu(s) s_{uv} \mu(s)\kappa_j(a)\kappa_j(b).$$

Since $n_j = 1$ the terms with $s > 1$ contribute

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

The inner sum is

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

and the sum over $n$ is

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

Thus the total contribution from the terms with $s > 1$ is

$$\frac{K_j^2 \phi(q)^{k-2} N (\log R)^{k-2}}{q^{k-1}Q}.$$ 

For the remaining terms we have $a = b = n$. Thus they give

$$\frac{X}{\phi(q)} \sum_n \frac{j}{\phi_2(n)} \kappa_j(n)^2.$$
We recall that $X = \int_{N + h_j}^{2N + h_j} \frac{dt}{\log t} = \frac{N}{\log N} + O\left(\frac{N}{(\log N)^2}\right)$. Moreover

$$
\frac{1}{\phi(q)} \sum_n \frac{\kappa_j(n)^2}{\phi_2(n)} \ll \frac{K_j^2}{\phi(q)} \prod_{Q < p \leq R} \left(1 + \frac{1}{p - 2}\right)^{k-1} \ll \frac{K_j^2(\phi(q))^{k-2}(\log R)^{k-1}}{q^{k-1}}.
$$

This completes the proof of the approximation for $S_j$.

The proof of the approximation for $T$ is essentially the same, except that we do not use Bombieri’s theorem and we do no 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 $O(1)$. By (18) we see that the total contribution arising from this error is

$$
\ll K^2 R^2 (\log R)^{4k-2}
$$

which is acceptable. Then just as the function $\phi$ now plays the rôle that $\phi_2$ played earlier, so the $\kappa_j$ is replaced by its understudy $\kappa$. Then the process of replacing $\lambda$ by $\kappa$ is identical, as is the elimination of the restriction $(d_u, e_v) = 1$.

The functions $\kappa_j$ and $\kappa$ introduced in Lemma 4 are clearly related to each other, as can be seen explicitly by (16) and (17). Thus when we insert the (16) into the definition of $\kappa_j$ and invert the order of summation we obtain (when $r_j = 1$)

$$
\kappa_j(r) = \mu(r)\phi_2(r) \sum_{s|r s|r s} \frac{\kappa(s)}{\phi(s)} \sum_{d|r|s} \frac{\mu(d)d}{\phi(d)}.
$$

Write $e_i = d_i/r_i$ and $t_i = s_i/r_i$. Then the inner sum is

$$
\frac{\mu(r)r}{\phi(r)} \sum_{e_i|t} \frac{\mu(e)e}{\phi(e)} = \frac{\mu(r)r\mu(t/t_j)}{\phi(r)\phi(t/t_j)} = \frac{r\mu(s/s_j)}{\phi(s/s_j)}.
$$

On using the notation $rt$ for $r_1t_1, \ldots, r_kt_k$ we find that

$$
\kappa_j(r) = \frac{r\phi_2(r)}{\phi(r)^2} \sum_t \frac{\kappa(rt)\mu(t)\phi(t_j)\mu(t_j)}{\phi(t)^2}.
$$
The terms with \( t > t_j \) contribute

\[ \ll K \sum_{t_j \leq R} \frac{\mu(t_j)^2}{\phi(t_j)} \sum_{n > 1} \frac{(k - 1)^{\omega(n)} \mu(n)^2}{\phi(n)^2} \]

and we have

\[ \sum_{t_j \leq R} \frac{\mu(t_j)^2}{\phi(t_j)} \ll \left( \prod_{Q < p \leq R} \frac{p}{p - 1} \right) \ll \frac{\phi(q)}{q} \log R \]

and

\[ \left( -1 + \prod_{Q > p} \left( 1 + \frac{k - 1}{(p - 1)^2} \right) \right) \ll Q^{-1}. \]

Since also

\[ \frac{r \phi_2(r)^2}{\phi(r)} = 1 + O(1/Q) \]

it follows that, when \( r_j = 1 \),

\[ \kappa_j(r) = \sum_{t_j} \frac{\kappa(r')}{\phi(t_j)} + O \left( \frac{K \phi(q) \log R}{qQ} \right) \]  (20)

where \( r' = r_1, \ldots, r_j - 1, t_j, r_{j+1}, \ldots, r_k \).

Having come this far, we should take stock. The ultimate aim is to maximise the ratio

\[ \frac{\sum_{r} \kappa_j(r)^2}{\phi_2(r)} \]

We now make the reasonable assumption that

\[ \kappa(r) = f \left( \frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R} \right) \]  (21)

where \( f \in \mathcal{F} \) which through (17) gives (12).

The final step of the proof of Theorem 3 is to obtain smooth approximations to the main terms in Lemma 4. We have standard methods of carrying this out when \( k = 1 \), i.e. \( r = r_1 \). We adopt the simple expedient of establishing a suitable one-dimensional approximation and then applying it \( k \)-times.
Suppose that \( g : [0, 1] \to \mathbb{R} \). Then we call \( g \) \( l \)-piecewise absolutely continuous on \([0, 1]\) when associated with \( g \) there is a partition \( a_0 = 0 < a_1 < \ldots < a_l = 1 \) of \([0, 1]\) such that for \( 1 \leq j \leq l \)

1. \( g_+(a_{j-1}) = \lim_{x \to a_{j-1}^+} g(x) \) and \( g_-(a_j) = \lim_{x \to a_j^-} g(x) \) both exist, and

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

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

We observe in passing that in practice it suffices for our application that \( g' \) is continuous except for at most one \( x \) in \([0, 1]\) where \( g \) and \( g' \) have jump discontinuities.

**Lemma 6.** Suppose that \( \eta : \mathbb{N} \to \mathbb{R} \) is multiplicative with its support on the squarefree numbers, and that \( 0 \leq \eta(p) \leq 2 \) and there is a constant \( C \) such that whenever \( p > C \) we have

\[
\left| \frac{\eta(p)}{p} - \frac{1}{p} \right| \leq \frac{C}{p^2}.
\]

Suppose also that \( g \in \mathcal{G}(l, G) \) and \( m \in \mathbb{N} \). Then

\[
\sum_{\substack{n \leq x \\
(n,m)=1}} \eta(n) g \left( \frac{\log n}{\log x} \right)
= A_m \int_0^1 g(v) \, dv \log x + O \left( lG \left( 1 + \sum_{p \mid m} \frac{\log p}{p} \right) \prod_{p \mid m} \left( 1 + \frac{1}{p} \right) \right)
\]

where

\[
A_m = \frac{\phi(m)}{m} \prod_{p \mid m} \left( 1 + \frac{\eta(p)}{p} \right) \left( 1 - \frac{1}{p} \right).
\]

We also have

\[
A_m \ll \frac{\phi(m)}{m}.
\]
In order to make a comparison with the main term, which is of size
\[ \sim \frac{\phi(m)}{m} \log x \int_0^1 g(v)dv, \]
it is useful to observe that the error term is
\[ \ll G \frac{\phi(m)}{m} (\log \log 3m)^3. \]

**Proof.** We begin with the case when \( g \) is identically 1. Also we may suppose that \( \eta(p) = 0 \) when \( p|m \). Let \( \rho \) be the multiplicative function with \( \rho(p) = \eta(p) - 1/p, \rho(p^2) = -\eta(p)/p, \eta(p^t) = 0 \) \((t \geq 3)\) and let \( \nu = 0 \) or \( 1 \). Then
\[ \sum_{u|n} \frac{\rho(n/u)}{u} = \eta(n) \]
and
\[ \sum_{l \leq y} (\log 2l)\nu|\rho(l)| \ll \sum_{rsst \leq y \atop r|m, (st, m) = 1} (\log 2rst)^\nu \frac{\mu(rst)^2}{rsst} \sum_{u|st} \frac{C_{\nu(u)}}{u} \]
\[ \ll \left( 1 + \sum_{p|m} \frac{\log p}{p} \right) \prod_{p|m} \left( 1 + \frac{1}{p} \right). \]
Also
\[ \sum_{x<l \leq y} |\rho(l)| \ll \frac{1}{\log x} \sum_{l} (\log l)|\rho(l)|. \]

Therefore
\[ \sum_{n \leq x} \eta(n) = \sum_{u,v \leq x} \frac{\rho(v)}{u} = \sum_{v \leq x} \rho(v) \left( \log \frac{x}{v} + O(1) \right) \]
\[ = A_m \log x + O \left( \left( 1 + \sum_{p|m} \frac{\log p}{p} \right) \prod_{p|m} \left( 1 + \frac{1}{p} \right) \right). \]

Now we apply this to general \( g \in \mathcal{G}(l, G) \). Let
\[ E(x) = \sum_{n \leq x} \eta(n) - A_m \log x \]
and choose \( a_j \) as provided by the definition of \( \mathcal{G}(l, G) \). When \( x^{a_j-1} < n \leq x^{a_j} \) we have
\[ g \left( \frac{\log n}{\log x} \right) = g_-(a_j) - \int_{\log n \atop \log x}^{a_j} g'(v)dv \]
except possibly when \( n = x^{a_j} \) in which case the two sides differ by \( \ll G \). Now we multiply by \( \eta(n) \), sum over the \( n \in (x^{a_{j-1}}, x^{a_j}] \), interchange the order of summation and integration and apply the formula for \( E \) to obtain

\[
(A_m(\log x)(a_j - a_{j-1}) + E(x^{a_j}) - E(x^{a_{j-1}}))g_-(a_j)
- \int_{a_{j-1}}^{a_j} (A_m(\log x)(v - a_{j-1}) + E(x^v) - E(x^{a_{j-1}}))g'(v)dv + O(G).
\]

We integrate the main term by parts to obtain

\[
\int_{a_{j-1}}^{a_j} A_m(\log x)g(v)dv
\]

which on summing over \( j \) gives the desired main term. We insert the bound for \( E \) given by the first part of the proof and sum over \( j \). This completes the proof of the lemma.

We are now in a position to complete the proof of Theorem 3. We now make the choice (21) for some \( f \) in \( F \). To simplify some of the formulae we then extend the definition of \( f \) to \([0, 1]^k\) by taking \( f \) to be 0 outside \( R \). Again we concentrate on \( S_j \) rather than \( T \). We recall that \( \kappa_j(r) = 0 \) unless \( r_j = 1, (r, q) = 1 \) and \( r \) is squarefree, in which case, by (20) and (21), we have

\[
\kappa_j(r) = \sum_{t_j} \frac{\mu(t_j)^2}{\phi(t_j)} f \left( \frac{\log r_1, \ldots, \log r_{j-1}, \log t_j, \log r_{j-1}, \ldots, \log r_k}{\log R, \log R, \log R, \ldots, \log R} \right)
+ O \left( \frac{F \phi(q) \log R}{qQ} \right)
\]

where \( r' = r_1, \ldots, r_{j-1}t_j, r_{j+1}, \ldots, r_k \). Thus

\[
K_j \ll F \frac{\phi(q)}{q} \log R.
\]

Moreover, by Lemma 6, with \( \eta(p) = 1/p \) and \( m = qr \) we have

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

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

Recall that this holds when \( r_j = 1, (r, q) = 1 \) and \( r \) is squarefree and that otherwise \( \kappa_j(r) = 0 \). Thus, by Lemma 5,

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

The general arithmetical factor in the main term in the sum can be rewritten as

\[ \prod_{i=1}^k \frac{\mu(r_i) \phi(r_i)}{\phi_2(r_i)r_i} \]

provided that the sum over \( r \) is restricted to \( r \) with \((r_u, r_v) = 1\) when \( u \neq v \). However if \((r_u, r_v) > 1\), then there is a prime \( p > Q \) such that \( p|r_u \) and \( p|r_v \). Therefore when we remove the condition \((r_u, r_v) = 1\) the total error in so doing is

\[ \ll \frac{F^2 \phi(q)N(\log R)^2}{q^2 \log N} \sum_{p>Q} \phi(p)^2 (\sum_{(n, q) = 1}^n \frac{\mu(n)^2 \phi(n)}{\phi_2(n)n})^{k-1} \]

\[ \ll \frac{F^2 \phi(q)^k N(\log R)^k}{q^{k+1}Q}. \]

Thus the sum in the main term can be replaced by

\[ \sum_{(r, q) = 1}^r f_j(r)^2 \prod_{i=1}^k \frac{\mu(r_i)^2 \phi(r_i)}{\phi_2(r_i)r_i}. \]

Now we apply Lemma 6 to each variable \( r_i \) in turn, i.e \( k - 1 \) times, with

\[ \eta(p) = \frac{p - 1}{(p - 2)p} - \frac{1}{p} = \frac{1}{p(p - 2)} \]

and \( m = q \). Thus

\[ S_j = \frac{\phi(q)^k N(\log R)^{k+1}}{q^{k+1} \log N} I_j + O \left( \frac{F^2 \phi(q)^k N(\log R)^k}{q^{k+1}Q} \right). \]
where $I_j$ is as in Theorem 3. This gives the first part of that theorem. The second part follows in the same way.

**Theorem 7** (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 3. 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 8** (Zhang). There are bounded gaps in the sequence of primes.

This is immediate from Theorems 3, 7 and the fact that there are admissible sets with $k$ elements as provided, for example, by Theorem 1.

**Corollary 9** (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 10** (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 $h$ contained in $\mathcal{G}$ such that there are infinitely many $n$ for which each member of the $m$-tuple $n + h$ is prime. Then for $l > l_0(m)$

$$l^m \geq M(m, l, \mathcal{G}) \gg_m l^m$$

and

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

dе 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 11. 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 $p_2 - p_1 = d$.

Proof of Theorem 7. Let

$$x = \frac{k/ \log k}{\log(k/ \log k)}$$

and let $\xi$ be the positive solution to

$$1 + \xi x = e^\xi.$$ 

Then

$$\frac{e^\xi}{\xi} > x$$

and so, for $k$ sufficiently large

$$\log k - \log \log k < \xi < \log k.$$ 

Let $g : [0, \infty) \to \mathbb{R}$ be defined by

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

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

$$\alpha = \int_0^\infty g(y) dy = 1,$$

$$\beta = \int_0^\infty g(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi e^\xi},$$

$$\gamma = \int_0^\infty yg(y)^2 dy = \frac{1}{\xi} - \frac{1}{\xi^2} + \frac{1}{\xi^2 e^\xi},$$

$$\tau = \int_0^\infty y^2 g(y)^2 dy = \frac{x}{\xi^2} - \frac{2}{\xi^2} + \frac{1}{\xi^3} - \frac{1}{\xi^3 e^\xi}.$$

We now take

$$f(t) = \begin{cases} 
\prod_{i=1}^k g(kt) & t \in \mathcal{R}, \\
0 & t \notin \mathcal{R}.
\end{cases}$$
Since $f$ is symmetric we have $I_j(f) = I_k(f)$ for every $j \leq k$. Thus 

$$
\rho \geq \frac{k I_k(f)}{J(f)}
$$

and we now proceed to estimate $I_k(f)$ and $J(f)$. It is clear that with this choice most of the mass of $f$ is close to the axes and so the boundary condition $t_1 + \cdots + t_k \leq 1$ on $R$ is relatively unimportant. Since we are concerned with only a lower bound for $\rho$, lower and upper bounds for $I_k(f)$ and $J$ respectively will suffice. An upper bound for $J(f)$ is easy. We have 

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

Thus we can concentrate on $I_k(f)$. Let $S$ denote the set of $k-1$–tuples $y_1, \ldots, y_{k-1}$ with $y_i \geq 0$ and $y_1 + \cdots + y_{k-1} \leq k - x$. Then we have 

\[
\begin{align*}
    kI_k(f) &= k \int_{\mathbb{R}^{k-1}} \left( \prod_{i=1}^{k-1} g(kt_i)^2 \right) \left( \int_0^{1-t_1-\cdots-t_{k-1}} g(kt_k)dt_k \right)^2 dt_1 \cdots dt_{k-1} \\
    &\geq k^{-k} \alpha^2 \int_{S} \prod_{i=1}^{k-1} g(y_i)^2 dy \\
    &= k^{-k} \alpha^2 \beta^{k-1} - E
\end{align*}
\]

where 

$$
E = k^{-k} \alpha^2 \int_{S^*} \prod_{i=1}^{k-1} g(y_i)^2 dy
$$

and

$$
S^* = [0, \infty)^{k-1} \setminus S.
$$

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 $y \in S^*$ is equivalent to $y_1 + \cdots + y_{k-1} \geq k - x$ and this in turn is equivalent to 

$$
\frac{y_1 + \cdots + y_{k-1}}{k-1} - \sigma \geq \frac{k - x - \sigma(k-1)}{k-1} = \frac{(1 - \sigma)(k-1) - x + 1}{k-1}.
$$
For $k$ sufficiently large we have
\[(1 - \sigma)(k - 1) - x + 1 > 0\]
and
\[1 - \sigma - \frac{x - 1}{k - 1} = \xi^{-1} + O(\xi^{-2}).\]
In particular if $y \in S^*$, then
\[\left(\frac{y_1 + \cdots + y_{k-1}}{k - 1} - \sigma\right)^2 \zeta^2 \geq 1\]
where
\[\zeta = \left(1 - \sigma - \frac{x - 1}{k - 1}\right)^{-1} = \xi + O(1).\]
Hence
\[E \leq k^{-k} \alpha^2 \zeta^2 \int_{[0, \infty)^{k-1}} \left(\frac{y_1 + \cdots + y_{k-1}}{k - 1} - \sigma\right)^2 \prod_{i=1}^{k-1} g(y_i)^2 dy.
\]
We now square out the expression
\[\left(\frac{y_1 + \cdots + y_{k-1}}{k - 1} - \sigma\right)^2 = \frac{2}{(k - 1)^2} \sum_{1 \leq i < j \leq k-1} y_iy_j + \frac{1}{(k - 1)^2} \sum_{i=1}^{k-1} y_i^2 - \frac{2\sigma}{k - 1} \sum_{i=1}^{k-1} y_i + \sigma^2\]
and evaluate the various integrals with reference to the integrals evaluated above. Thus
\[E \leq k^{-k} \alpha^2 \zeta^2 \left(\frac{k - 2}{k - 1} \gamma^2 \beta^{k-3} + \frac{1}{k - 1} \tau \beta^{k-2} - 2\sigma \gamma \beta^{k-2} + \sigma^2 \beta^{k-1}\right).
\]
By the definition of $\sigma$ this becomes
\[k^{-k} \alpha^2 \zeta^2 \beta^{k-3} \frac{\tau \beta - \gamma^2}{k - 1} < k^{-k} \alpha^2 \zeta^2 \beta^{k-2} \frac{\tau}{k - 1}.
\]
Thus
\[\rho > \beta^{-1} \left(1 - \frac{\zeta^2 \tau}{\beta(k - 1)}\right).\]
We have
\[ \log k - \log \log k < \xi = \log k - \log \log k + O(1), \]
\[ \beta^{-1} = \xi + O(\xi^{-1} \log k), \]
\[ \zeta^2 = \xi^2 + O(\xi), \]
\[ \tau = x\xi^{-2} + O(\xi^{-2}), \]
\[ \frac{1}{k - 1} = \frac{1}{k} + O(k^{-2}). \]
Thus
\[ \frac{\zeta^2 \tau}{\beta(k - 1)} = (\xi + O(1)) xk^{-1} = \frac{\xi + O(1)}{(\log k) \log(k/\log k)} = \frac{1}{\log k} + O((\log k)^{-2}). \]
Hence
\[ \rho > \xi \left(1 + O(k^{-1} \log k)\right) \left(1 - \frac{1}{\log k} + O((\log k)^{-2})\right) > \log k - \log \log k - 1 \] if \( k \) is sufficiently large. \( \square \)

**Proof of Corollary 9.** Let \( C \) be a constant chosen so that for every \( m \in \mathbb{N} \) we have
\[ \frac{Cme^{4m}}{4m + \log m + \log C} > e^{2+4m}. \]
Hence for \( k \geq \max(3, Cme^{4m}) \) we have
\[ \frac{k}{\log k} \geq e^{2+4m} \]
and so
\[ \log k - \log \log k - 1 > 4m + 1. \]
Thus if \( k \) is large enough
\[ \left(\frac{1}{4} - \frac{1}{k}\right) (\log k - \log \log k - 1) > m. \]
Taking the level of distribution \( \theta \) to be \( \frac{1}{2} \) and choosing \( \delta = \frac{2}{k} \) we see that every admissible \( k \)-tuple \( h \) has the property that there are infinitely many \( n \) such that the \( k \)-tuple \( n + h \) contains at least \( m \) primes. By Theorem 1 there is a an admissible \( k \)-tuple of diameter \( \ll k \log k \ll m^2 e^{4m}. \) \( \square \)
Proof of Corollary 10. Let \( k = \lceil \max(3, Cm^{4m}) \rceil \) be as in the proof of Corollary 9 and let \( h \) be an admissible \( k \)-tuple. By considering all possible \( m \)-tuples \( h' = h_1', \ldots, h_m' \) which are subsets of \( h \) we see that at least one has the property that there are infinitely many \( n \) such that \( n + h_1', \ldots, n + h_m' \) are simultaneously prime, i.e. the prime \( m \)-tuple conjecture holds for this \( m \)-tuple.

Starting from \( G \) we construct a subset \( G' \) by successively removing elements from \( G \). Given a prime \( p \) and a finite set \( L \) of integers we can construct a subset as follows. Let \( L(p; h) = \{ n \in L : n \equiv h \ (\text{mod} \ p) \} \) and \( L(p; h) = \text{card}L(p; h) \). Choose an \( h \) for which \( L(p; h) \) is minimal and take \( L' = L \setminus L(p; h) \). Then \( \text{card}L' \geq (1 - 1/p)\text{card}L \). We apply this operation successively to \( G \) for \( p \leq k \) giving a subset \( G' \) which satisfies

\[
\text{card}G' \geq \text{card}G \prod_{p \leq k} \left( 1 - \frac{1}{p} \right) \gg m \ l.
\]

Thus on taking \( l \) to be sufficiently large we have \( s = \text{card}G' > k \). Every subset \( h \) of \( G' \) of cardinality \( k \) is an admissible set since it omits a residue class modulo \( p \) for every \( p \leq k \). There are \( \binom{s}{k} \) such \( h \) and, from above, each one contains at least one \( m \)-tuple \( h' \) for which the prime \( m \)-tuples conjecture holds. Subsets \( b \) of \( G' \) of cardinality \( k \) which contain \( h' \) are exactly those in which the \( k - m \) remaining elements of \( b \) are chosen at random from the \( s - m \) remaining elements of \( G' \). Thus there are precisely \( \binom{s-m}{k-m} \) such \( b \). Hence there are at least \( \binom{s}{k}/\binom{s-m}{k-m} = \frac{(s-m+1)\cdots(s-1)s}{(k-m+1)\cdots(k-1)k} \gg m \ s^m \gg m \ l^m \) admissible subsets of \( G \) of cardinality \( m \) which satisfy the prime \( m \)-tuples conjecture. On the other hand there are \( \binom{l}{m} \leq l^m \) subsets \( h \) of \( G \) of cardinality \( m \), and this completes the proof of Corollary 10. \( \square \)

25.2.1 Exercises

1. Suppose that \( k \geq 2 \). Let \( R_k \subset [0, 1]^k \) be defined by \( R_k = \{ t : t_i \geq 0, t_1 + \cdots + t_k \leq 1 \} \), and let \( m \in \mathbb{N} \) and \( f(t) = (1 - t_1 - \cdots - t_k)^m \). Given \( t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_{k} \in [0, 1]^{k-1} \) with \( t_1 + \cdots + t_{j-1} + t_{j+1} + \cdots + t_{k} \leq 1 \) let \( A_j \) denote the interval \([0, 1 - t_1 - \cdots - t_{j-1} - t_{j+1} - \cdots - t_{k}] \) (and take it to be the empty set otherwise) and define

\[
I_j(f) = \int_0^1 \cdots \int_0^1 \left( \int_{A_j} f(t) \, dt \right)^2 \, dt_1 \cdots dt_{j-1} \, dt_{j+1} \cdots dt_{k}
\]
and

\[ J(f) = \int_{\mathcal{R}_k} f(t)^2 \, dt. \]

(i) Prove that \( \sum_{j=1}^{k} I_j(f) = \frac{k(2m+2)!}{(2m+1+k)!(m+1)^2} \).

(ii) Prove that \( J(f) = \frac{(2m)!}{(2m+k)!} \).

(iii) Prove that \( \sum_{j=1}^{k} I_j(f) \frac{J(f)}{J(f)} = 4 \left( 1 - \frac{1}{2m+2} \right) \left( 1 - \frac{2m+1}{2m+1+k} \right) \).

(iv) (Goldston, Pintz, Yıldırım) Prove that if the level \( \theta \) of distribution of the primes satisfies \( \theta > \frac{1}{2} \), then there are infinitely many bounded gaps in the sequence of primes.

2. Let \( \mathcal{R}_k \) be as in question 1. For \( t \in \mathcal{R}_k \) let \( \alpha_k(t) = t_1 + \cdots + t_k \) and \( \beta_k(t) = t_1^2 + \cdots + t_k^2 \).

(i) Suppose that \( a \) and \( a_j \) are non-negative integers. Prove, by induction on \( k \) or otherwise, that

\[ \int_{\mathcal{R}_k} (1 - \alpha_k(t))^a \prod_{j=1}^{k} t_j^{a_j} \, dt = \frac{a! \prod_{j=1}^{k} a_j!}{(k+a+\sum_{j=1}^{k} a_j)!}. \]

(ii) Suppose that \( a \) and \( b \) are non-negative integers. Prove that

\[ \int_{\mathcal{R}_k} (1 - \alpha_k(t))^a \beta_k(t)^b \, dt = \frac{a! b!}{(k+a+2b)!} \sum_{b_1+\cdots+b_k=b} \prod_{j=1}^{k} \frac{(2b_j)!}{b_j!}. \]

(The multinomial theorem applied to \( \beta_k^b \) is useful here.)

3. (Maynard) (i) Let \( k = 5 \). In the notation of the previous question, when \( t \in \mathcal{R}_5 \), let

\[ f(t) = (1 - \alpha_5(t)) \beta_5(t) + \frac{7}{10} (1 - \alpha_5(t))^2 + \frac{1}{14} \beta_5(t)^2 - \frac{3}{14} (1 - \alpha_5(t)). \]

Prove that

\[ \sum_{j=1}^{5} I_j(f) = \frac{1417255}{708216}. \]

(ii) Prove that if the level of distribution \( \theta \) is 1, then

\[ \liminf_{n \to \infty} p_{n+1} - p_n \leq 12. \]
25.3 Notes

§1. The study of the distribution of \( k \)-tuples of primes is an area which is, as we write (April 22, 2014), very active and indeed has a considerable air of excitement about it with substantial progress and widespread consequences.

There is a long history of investigation into the existence of \( k \)-tuple of primes, but until the last decade or so, in spite of the use of quite heavy machinery, there has been more speculation than progress. The twin prime conjecture that there are infinitely many prime pairs which differ by 2 was already well known when de Polignac conjectured (1849) the every even number is the difference of infinitely many pair of primes. The general prime \( k \)-tuples conjecture, Conjecture 1, and a quantitative form equivalent to Conjecture 2, were made by Hardy and Littlewood [1922], and their justification was, as here, by a process which ignores any minor arcs and assumes good approximations to the generating function on the major arcs.

Theorem 2 can be found in Gallager [1976].

§2. In an unpublished manuscript, originally intended as part of the seminal series of papers “On some problems of partitio numerorum”, in which they assumed the Generalised Riemann Hypothesis (GRH), Hardy and Littlewood applied their eponymous method to show that

\[
\liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} \leq \frac{2}{3}.
\]

By a different method based on Brun’s sieve Erdős [1940] showed that there is a constant \( c < 1 \) such that unconditionally

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

Rankin [1940] refined the Hardy–Littlewood argument to obtain \( c = 3/5 \) on GRH, and in [1950] obtained \( c = (42/43)(3/5) = 0.5860 \ldots \) on GRH by combining his method with that of Erdős. Ricci [1954] improved the sieve method so as to obtain \( c = 15/16 \) unconditionally.

More substantial progress was made by Bombieri and Davenport. It is of some methodological and historic interest. In a similar manner to Hardy and Littlewood, and Rankin, they consider

\[
I = \int_0^1 |S(\alpha)|^2 K(\alpha) d\alpha
\]
where $S(\alpha)$ is as in (1) and

$$K(\alpha) = \left| \sum_{h=1}^{H} e(\alpha h) \right|^2 = \sum_{h=-H}^{H} (H - |h|) e(\alpha h)$$

is the Fejér kernel. They then evaluate $I$ in two different ways. First of all, by the orthogonality of the additive characters $e(n\alpha)$,

$$I = H \sum_{p \leq x} (\log p)^2 + 2 \sum_{h=1}^{H} (H - h) R(x; h)$$

where

$$R(x; h) = \sum_{\substack{p, p' \leq x \\ p' - p = h}} (\log p)(\log p').$$

The first sum is

$$Hx \log x + O(Hx).$$

The positivity of the integrand means that a lower bound can be obtained by considering major arcs only. Moreover here the Bombieri–Vinogradov theorem can be used to make the major arcs almost as wide and numerous as would pertain if one assumed the Generalised Riemann Hypothesis. This was the first use of the Bombieri–Vinogradov theorem, and indeed it is this work which stimulated Bombieri’s discovery. Thus it can be shown that the integral is asymptotically at least

$$H^2 x + \frac{1}{2} Hx \log x$$

and so the asymptotic lower bound

$$\sum_{h=1}^{H} (H - h) R(x; h) \geq \frac{1}{2} H^2 x - \frac{1}{4} Hx \log x$$

implies that $\liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} \leq \frac{1}{2}$. They also apply the Selberg sieve to obtain an upper bound for the terms on the left with $h$ near $H$, and here the Bombieri–Vinogradov theorem is also applied in the sieve. This gives

$$c = \frac{2 + \sqrt{3}}{8} = 0.4665 \ldots.$$
More refined kernels $K(\alpha)$ were introduced by Pil’tai [1972] and Huxley [1973, 1977] to obtain small improvements,

\[ c = 0.4571 \ldots, \ c = 0.4463 \ldots, \ c = 0.4425 \ldots \]

respectively. A larger improvement was made by Maier [1988] by the adaptation of his matrix method to the argument so as to take advantage of oscillations in the primes over short intervals and this lead to the known bound being reduced by a factor of $e^{-\gamma}$ where $\gamma$ is Euler’s constant.

\[ c = 0.2484 \ldots. \]

In the last decade or so there have been a series of major advances. In a seminal paper Goldston, Pintz, Yıldırım [2009] proved that

\[ \liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} = 0 \]

In fact they were able to prove rather more than this. For example they obtain [2010] an explicit upper bound

\[ p_{n+} - p_n \ll (\log p_n)^{1/2}(\log \log p_n)^2 \]  \hspace{1cm} (22)

which is satisfied for infinitely many $n$. and they showed that if the level of distribution exceeds $\frac{1}{2}$, then there are infinitely many bounded gaps between primes. Indeed, if the level of distribution can be taken to be 1 (the Elliott–Halberstam conjecture [1970]), they were able to show that infinitely often $p_{n+1} - p_n \leq 16$. All subsequent work is based on their method. There have been two sensational developments. Zhang [2014] proved a version of the Bombieri–Vinogradov theorem in which the moduli of the arithmetic progressions are restricted to being numbers with only relatively small prime factors but, crucially, the level of distribution exceeds $\frac{1}{2}$ by a small amount. Then, although the moduli are restricted, nevertheless the modified Bombieri–Vinogradov theorem contains enough information to enable an adaptation of the Goldston, Pintz, Yıldırım machinery to work. Thus Zhang obtains

\[ \liminf_{n \to \infty} p_{n+1} - p_n \leq 70,000,000. \]  \hspace{1cm} (23)

Then Maynard [to appear], returning to a version of the Goldston, Pintz, Yıldırım methods which predates their [2009] paper and which had been
aborted as unsuccessful, was able to adapt their methods to establish infinitely many bounded gaps between the primes even if one only has a positive level of distribution for the primes. In particular, using the Bombieri–Vinogradov theorem Maynard obtains

\[ \liminf_{n \to \infty} p_{n+1} - p_n \leq 600. \tag{24} \]

These most recent methods involve quite heavy computations to obtain the sharpest bounds. For example, in the notation of exercise 25.2.1.2, Maynard considers Theorem 3 with

\[ f(t) = \sum_{i=1}^{d} a_i (1 - \alpha_k(t))^b_i \beta_k(t)^{c_i} \]

and finds that (c.f. Exercise 25.2.1.2)

\[ \frac{\sum_{j=1}^{k} I_j(f)}{J(f)} = \frac{a^T M a}{a^T N a} \]

where the \( d \times d \) positive definite matrices \( M, N \) depend on the exponents \( b_i, c_i \). He shows that this ratio is maximised when \( a \) is an eigenvector of \( MN^{-1} \) corresponding to the largest eigenvalue. He then takes \( k = 105 \) and considers all choices of \( b_i, c_i \) with \( b_i + 2c_i \leq 11 \), so that \( d = 42 \). It transpires that the largest eigenvalue is

\[ 4.0020697 \ldots \]

and so an appeal to Theorem 3 establishes that for any admissible 105–tuple \( h \) there are infinitely many \( n \) such that \( n + h \) contains at least two primes. He then displays a known admissible 105–tuple of diameter 600 discovered by Englesma to establish (24). Maynard’s method also shows that when the level of distribution is 1 we have

\[ \liminf_{n \to \infty} p_{n+1} - p_n \leq 12, \tag{25} \]

for which see exercise 25.2.1.3.

There is an ongoing Polymath project at

lead by Tao to combine all the methods, especially those of Maynard and Zhang, and the record as of 11th April 2014 is

$$\lim \inf_{n \to \infty} p_{n+1} - p_n \leq 252. \quad (26)$$

The methods described here have very great flexibility and have many potential applications. One is to Dickson’s conjecture [1904] which states that if the $g_i$, $h_i$ are integers and $\prod_{i=1}^k (g_i n + h_i)$ has no fixed prime divisor, then there are infinitely many $n$ such that the $g_i n + h_i$ are simultaneously prime. Another undertaken by Pintz [to appear] is to deal with questions involving consecutive primes and arithmetic progressions. In yet another Goldston, Graham, Pintz, Yıldırım [2011] have considered $n$ for which $d(n) = d(n + 1)$, $\omega(n) = \omega(n + 1)$ and $\Omega(n) = \Omega(n + 1)$ simultaneously. There are also applications to cognate problems in algebraic number fields.

25.4 References

Dickson, L. E. (1904), A new extension of Dirichlet’s theorem on prime numbers, Messenger of Math. 33(1904), 155-161.
Erdős, P. [1940], The difference of consecutive primes, Duke Math. J. 6 (1940), 438-441.
Maynard, J. [to appear], Small gaps between primes, arXiv:1311.4600v2 [math.NT].