TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 361, Number 10, October 2009, Pages 5285–5330
S 0002-9947(09)04788-6
Article electronically published on May 27, 2009

SMALL GAPS BETWEEN PRIMES OR ALMOST PRIMES

D. A. GOLDSTON, S. W. GRAHAM, J. PINTZ, AND C. Y. YILDIRIM

ABSTRACT. Let p_n denote the n^{th} prime. Goldston, Pintz, and Yıldırım recently proved that

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

We give an alternative proof of this result. We also prove some corresponding results for numbers with two prime factors. Let q_n denote the n^{th} number that is a product of exactly two distinct primes. We prove that

$$\liminf_{n\to\infty} (q_{n+1}-q_n) \le 26.$$

If an appropriate generalization of the Elliott-Halberstam Conjecture is true, then the above bound can be improved to 6.

1. Introduction

In 1849, A. de Polignac ([23]; see also [5], p. 424) conjectured that every even number is the difference of two primes in infinitely many ways. More generally, we can let $\mathcal{H} = \{h_1, h_2, \ldots, h_k\}$ be a set of k distinct integers. A major open question in number theory is to show that there are infinitely many positive integers n such that $n+h_1, n+h_2, \ldots, n+h_k$ are all prime, provided that $\mathcal{H}$ meets an obvious necessary condition that we call *admissibility*. For each prime p, let $\nu_p(\mathcal{H})$ be the number of distinct residue classes mod p in $\mathcal{H}$. We say that the set $\mathcal{H}$ is *admissible* if $\nu_p(\mathcal{H}) < p$ for all p.

Using heuristics from the circle method, Hardy and Littlewood [14] realized the significance of the singular series $\mathfrak{S}(\mathcal{H})$, defined as

$$\mathfrak{S}(\mathcal{H}) = \prod_p \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right)\left(1 - \frac{1}{p}\right)^{-k} \tag{1.1}$$

for this problem. They made a conjecture about the asymptotic distribution of the numbers n for which $n+h_1, \ldots, n+h_k$ are all prime, which we state here in the following form.

Received by the editors September 17, 2007.
2000 *Mathematics Subject Classification.* Primary 11N25; Secondary 11N36.
Key words and phrases. Primes, almost primes, gaps, Selberg's sieve, applications of sieve methods.
The first author was supported by NSF grant DMS-0300563, the NSF Focused Research Group grant 0244660, and the American Institute of Mathematics.
The second author was supported by a sabbatical leave from Central Michigan University and by NSF grant DMS-070193.
The third author was supported by OTKA grants No. 43623, 49693, 67676 and the Balaton program.
The fourth author was supported by TÜBİTAK.

Conjecture 1. *Let $\varpi(n)$ denote the function*

$$\varpi(n) = \begin{cases} \log n & \text{if } n \text{ is prime,} \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

As N tends to infinity,

$$\sum_{n \le N} \varpi(n+h_1)\varpi(n+h_2)\dots\varpi(n+h_k) = N(\mathfrak{S}(\mathcal{H}) + o(1)). \tag{1.3}$$

From the definition of $\mathfrak{S}(\mathcal{H})$, we see that $\mathfrak{S}(\mathcal{H}) \ne 0$ if and only if $\nu_p(\mathcal{H}) < p$ for all primes p; i.e., if and only if $\mathcal{H}$ is admissible.

The set $\mathcal{H} = \{0, 2\}$ is admissible, so the Hardy-Littlewood conjecture implies that

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

where p_n denotes the n^{th} prime. In an unpublished paper in the *Partitio Numerorum* series, Hardy and Littlewood [15] proved that if the Generalized Riemann Hypothesis is true, then

$$\liminf_{n\to\infty} \left(\frac{p_{n+1} - p_n}{\log p_n} \right) \le \frac{2}{3}.$$

In 1940, Erdős [7] used Brun's sieve to give the first unconditional proof of the inequality

$$\liminf_{n\to\infty} \left(\frac{p_{n+1} - p_n}{\log p_n} \right) < 1.$$

In 1965, Bombieri and Davenport [2] proved unconditionally that

$$\liminf_{n\to\infty} \left(\frac{p_{n+1} - p_n}{\log p_n} \right) \le 0.4665\dots. \tag{1.4}$$

This result was one of the first applications of what is now known as the "Bombieri-Vinogradov Theorem," which we state as follows.

Theorem (Bombieri-Vinogradov)**.** *When $(a, q) = 1$, let $E(x; q, a)$ be defined by the relation*

$$\sum_{\substack{x < n \le 2x \\ n \equiv a \pmod q}} \varpi(n) = \frac{x}{\phi(q)} + E(x; q, a). \tag{1.5}$$

Furthermore, let

$$E(x, q) = \max_{a;(a,q)=1} |E(x, q, a)|, \quad E^*(N, q) = \max_{x \le N} E(x, q). \tag{1.6}$$

If $A > 0$, then there exists $B > 0$ such that if $Q \le N^{1/2} \log^{-B} N$, then

$$\sum_{q \le Q} E^*(N, q) \ll_A N(\log N)^{-A}. \tag{1.7}$$

This result was proved by Bombieri in 1965 [1]. At about the same time, A. I. Vinogradov [28] gave an independent proof of a slightly weaker result. There are numerous proofs of this result available in the literature; see, for example, [4] and [27]. We remark that in the usual definition of $E(x; q, a)$, one takes the sum in (1.5) to be over $n \le x$. However, the above definition is more convenient for our purposes.

The bound (1.4) was improved in several steps by Huxley [19] to 0.4394.... In 1988, Maier [20] used his matrix method to improve the bound to 0.2484.... Maier's method had the shortcoming that it produced a sparse set of gaps; prior authors had shown that small gaps occur in a positive proportion of all cases. Goldston and Yıldırım [9] proved the upper bound of 0.25 for a positive proportion of gaps. Recently, the first, third and fourth authors proved a best possible result in this direction.

Theorem 1 (Goldston, Pintz, and Yıldırım, [10])**.**

$$\liminf_{n\to\infty}\left(\frac{p_{n+1}-p_n}{\log p_n}\right)=0.$$

The proof of Theorem 1 uses, among other things, the Bombieri-Vinogradov Theorem. There are good reasons to believe that the bound in (1.7) holds for larger values of Q. More formally we have the following conjecture.

Hypothesis $BV(\theta)$**.** *Suppose* $1/2<\theta\le 1$*. For all* $A>0$ *and all* $\epsilon>0$*, we have*

$$\sum_{q\le N^{\theta-\epsilon}}|E^*(N;q,a)|\ll_{A,\epsilon} N(\log N)^{-A}. \tag{1.8}$$

If Hypothesis $BV(\theta)$ is true, then we say that the sequence ϖ has a *level of distribution* θ. Thus the Bombieri-Vinogradov Theorem shows that ϖ has a level of distribution 1/2. The statement that ϖ has a level of distribution 1 is known as the "Elliott-Halberstam Conjecture" [6]. Any level of distribution larger than 1/2 will give the following strengthening of Theorem 1.

Theorem 2 (Goldston, Pintz, and Yıldırım [10])**.** *If Hypothesis* $BV(\theta)$ *is true for some fixed* $\theta>1/2$*, then*

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

If Hypothesis $BV(\theta)$ *is true for some* θ *with* $4(8-\sqrt{19})/15=0.97096\ldots<\theta\le 1$, *then*

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

Our first objective here is to give alternative proofs of Theorems 1 and 2. The primary difference in the proofs here and the proofs in [10] comes from the use of Selberg diagonalization and a different choice of sieve coefficients; this will be discussed in more detail below. Our choice of coefficients allows us to give an elementary treatment of the main terms; we will discuss this further after the statement of Theorem 6 below.

Our second objective is to obtain for numbers with a fixed number of prime factors stronger forms of results of the type proved in [10] for primes. Let E_k denote a number with *exactly* k distinct prime factors. This contrasts with the usual definition of "almost-prime", where P_k is used to denote a number with at most k distinct prime factors. Chen [3] proved that there are infinitely many primes p such that $p+2$ is a P_2. While one expects that there are infinitely many primes p such that $p+2$ is an E_2, this appears to be as difficult as the twin prime conjecture. However, we can prove that the limit infimum of gaps between E_2's is bounded.

Theorem 3. *Let q_n denote the n^{th} number that is a product of exactly two primes. Then*

$$\liminf_{n\to\infty}(q_{n+1}-q_n)\le 26.$$

We mention that the bound in the above theorem can be improved to 6 with a more elaborate proof that uses a different weighting function. This will be a topic of a future paper, and we will make some further comments on this after (1.28).

The above theorem uses an analogue of the Bombieri-Vinogradov Theorem for the function $\varpi * \varpi$, which is defined as

$$\varpi * \varpi(n) = \sum_{d|n} \varpi(d)\varpi(n/d).$$

Note that $\varpi * \varpi(n) = 0$ unless n is a product of two primes or n is a square of a prime.

When $(a,r)=1$, we have

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod r}} \varpi * \varpi(n) = \frac{1}{\phi(r)} \sum_{\chi \pmod r} \bar{\chi}(a) \sum_{N<n\le 2N} \varpi * \varpi(n)\chi(n),$$

and we expect the contributions from non-principal characters to show large cancellation, leaving a main term of

$$\frac{1}{\phi(r)} \sum_{N<n\le 2N} \varpi * \varpi(n)\chi_0(n), \tag{1.9}$$

where χ_0 is the principal character mod r. A computation (see Lemma 14) shows that this quantity is asymptotically equal to

$$\frac{N}{\phi(r)}\left(\log N + C_0 - 2\sum_{p|r}\frac{\log p}{p}\right), \tag{1.10}$$

where C_0 is the absolute constant defined in (2.8).

Let $E_2(N;r,a)$ be defined by

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod r}} \varpi * \varpi(n) = \frac{N}{\phi(r)}\left(\log N + C_0 - 2\sum_{p|r}\frac{\log p}{p}\right) + E_2(N;q,a).$$

Parallel to the definitions of $E(N,q)$ and $E^*(N,q)$, we define

$$E_2(N,r) = \max_{a,(a,r)=1} |E_2(N;r,a)|, \quad E_2^*(N,r) = \max_{x\le N} E_2(x,r).$$

Theorem (Bombieri-Vinogradov for $\varpi * \varpi$)**.** *For every $A>0$, there exists $B>0$ such that if $Q \le N^{1/2}\log^{-B} N$,*

$$\sum_{r\le Q} |E_2^*(N,r)| \ll_A N(\log N)^{-A}.$$

This is a special case of a result of Motohashi [22]. Alternatively, one can easily modify Vaughan's Identity for the von Mangoldt function Λ to an identity for $\Lambda * \Lambda$, and then use Vaughan's approach (see [27] or Chapter 28 of [4]) to the Bombieri-Vinogradov Theorem to prove the analogue for $\Lambda * \Lambda$. It is then easy to modify this to a result for $\varpi * \varpi$.

We also propose a natural analogue of Hypothesis $BV(\theta)$.

Hypothesis $BV_2(\theta)$**.** *Suppose* $1/2 < \theta \le 1$*. For all* $A > 0, \epsilon > 0$*, we have*

$$\sum_{q \le N^{\theta-\epsilon}} |E_2^*(N;q)| \ll_{A,\epsilon} N(\log N)^{-A}. \tag{1.11}$$

From this, we obtain the following conditional result.

Theorem 4. *If Hypotheses* $BV(\theta)$ *and* $BV_2(\theta)$ *are both true for some* θ *with* $(75-\sqrt{473})/56 = 0.950918\ldots < \theta \le 1$*, then*

$$\liminf_{n\to\infty} (q_{n+1} - q_n) \le 6.$$

The basic construction for the proofs of Theorems 1 and 2 was inspired by work of Heath-Brown [17] on almost prime-tuples of linear forms. Heath-Brown's work was itself a generalization of Selberg's proof [26] that the polynomial $n(n+2)$ will infinitely often have at most five prime factors, and in such a way that one of n and $n+2$ has at most two prime factors, while the other has at most three prime factors.

Define

$$P(n;\mathcal{H}) = \prod_{h\in\mathcal{H}} (n+h). \tag{1.12}$$

The central idea is to relate the problem to sums of the form

$$\sum_{N<n\le 2N} \Bigg(\sum_{d|P(n;\mathcal{H})} \lambda_d\Bigg)^2 \tag{1.13}$$

and of the form

$$\sum_{N<n\le 2N} \varpi(n) \Bigg(\sum_{d|P(n;\mathcal{H})} \lambda_d\Bigg)^2, \tag{1.14}$$

where one assumes that $\lambda_d = 0$ for $d > R$, and R is a parameter that is chosen to control the size of the error term. One also assumes that $\lambda_d = 0$ when d is not squarefree. By taking squares, we ensure that both sums (1.13) and (1.14) are positive.

To illustrate the relevance of the sums (1.13) and (1.14), we discuss one simple application that is related to the second part of Theorem 2. Let $\mathcal{H}$ be an admissible k-tuple, and consider the sum

$$\mathcal{S} := \sum_{N<n\le 2N} \Bigg\{\sum_{h\in\mathcal{H}} \varpi(n+h) \,-\, (\log 3N)\Bigg\} \Bigg(\sum_{d|P(n;\mathcal{H})} \lambda_d\Bigg)^2. \tag{1.15}$$

For a given n, the inner sum is negative unless there are at least two values $h_i, h_j \in \mathcal{H}$ such that $n+h_i, n+h_j$ are primes. From Theorems 5 and 6 below, one can deduce that if $BV(\theta)$ is true, if $R = N^{\theta/2-\epsilon}$ for $\epsilon > 0$, and if $0 \le \ell \le k$, then

$$\mathcal{S} \gtrsim N\mathfrak{S}(\mathcal{H})(\log R)^{k+2\ell}(\log N) m(k,\ell,\theta),$$

where

$$m(k,\ell,\theta) = \binom{2\ell}{\ell}\frac{1}{(k+2\ell)!}\left\{\frac{k(2\ell+1)(\theta-\epsilon)}{(k+2\ell+1)(\ell+1)} - 1\right\}.$$

This last expression is positive, if for example, $k = 7$, $\ell = 1$, ϵ is sufficiently small, and $20/21 < \theta \le 1$. Consequently, if $BV(1)$ is true, then for any admissible 7-tuple

$\mathcal{H}$, there are infinitely many n and some $h_i, h_j \in \mathcal{H}$ such that $n+h_i, n+h_j$ are both prime. Now

$$\mathcal{H} = \{11, 13, 17, 19, 23, 29, 31\}$$

is an admissible 7-tuple. $\mathcal{H}$ is admissible because if $p \le 7$, then none of the elements in $\mathcal{H}$ are divisible by p, and if $p > 7$, then there are not enough elements in $\mathcal{H}$ to cover all of the residue classes mod p. Now any two elements of $\mathcal{H}$ differ by at most 20, so we conclude that if $BV(1)$ is true, then

$$\liminf_{n\to\infty}(p_{n+1}-p_n) \le 20.$$

To get the stronger bound of 16 given in Theorem 2 we need an extra idea; this will be discussed in Section 7.

The success of the method depends upon making an appropriate choice for the λ_d, and this takes us into the realm of the Selberg upper bound sieve. It is a familiar fact [13, Chap IV, eqn. (1.9)] from the theory of this sieve that

$$\sum_{\substack{N<n\le 2N \\ d|P(n;\mathcal{H})}} 1 = \frac{N}{f(d)} + r_d,$$

where f is a multiplicative function and r_d is a remainder term. (See the first part of Section 3 for the formal definition of f.) Accordingly, an appropriate transformation of the sum in (1.13) leads to consideration of the bilinear form

$$\sum_{d,e} \frac{\lambda_d \lambda_e}{f([d,e])}. \tag{1.16}$$

The typical approach in the Selberg sieve is to choose the λ_d to minimize the form in (1.16). To make this problem feasible, one needs to diagonalize this bilinear form. This can be done by making a change of variables

$$y_r = \mu(r) f_1(r) \sum_d \frac{\lambda_{dr}}{f(dr)}, \tag{1.17}$$

where f_1 is the multiplicative function defined by $f_1 = f * \mu$. (Note that the sum in (1.17) is finite because $\lambda_d = 0$ for $d > R$.) The sum in (1.16) is then transformed into

$$\sum_r \frac{y_r^2}{f_1(r)},$$

and the bilinear form is minimized by taking

$$y_r = \mu^2(r)\frac{\lambda_1}{V}, \tag{1.18}$$

where

$$V = \sum_{r<R} \frac{\mu^2(r)}{f_1(r)}.$$

The minimum of the form in (1.16) is then seen to be

$$\frac{\lambda_1^2}{V}.$$

One usually assumes that $\lambda_1 = 1$, but this is not an essential element of the Selberg sieve, and it is sometimes useful to assign some other non-zero value to λ_1.

The sum in (1.14) can be treated in a similar way. However, the corresponding function f must be replaced by a slightly different function f^*, which will be defined in Section 4. Therefore, the optimal choice of λ_d is different from the optimal choice for the sum in (1.13). However, the basic structure of our approach requires that the same choice of λ_d be used for both sums. We therefore face the problem of making a choice of λ_d that works reasonably well for both problems. A similar choice was faced by Selberg and Heath-Brown in their earlier mentioned work, and they made this choice in different ways. Selberg [26] made a choice of λ_d that was optimal for one problem, and he was able to successfully analyze the effect of this choice for the other problem. Heath-Brown [17] chose

$$\lambda_d = \begin{cases} \mu(d)\left(\dfrac{\log R/d}{\log R}\right)^{k+1} & \text{if } d < R, \\ 0 & \text{otherwise,} \end{cases}$$

k being the number of linear forms under consideration. While this choice is not optimal for either problem, it is asymptotically optimal for the second problem (1.14).

Inspired by Heath-Brown's choice, Goldston, Pintz, and Yıldırım [10] chose

$$\lambda_{d,\ell} = \begin{cases} \mu(d)\dfrac{(\log R/d)^{k+\ell}}{(k+\ell)!} & \text{if } d < R, \\ 0 & \text{otherwise.} \end{cases} \tag{1.19}$$

Here, ℓ is a non-negative integer to be chosen in due course, with $\ell \le k$. With the exponent $k+\ell$, one is effectively using a $k+\ell$-dimensional sieve on a k-dimensional sieve problem. In an upper bound sieve, it is optimal to take the dimension of the sieve to be the same as the dimension of the problem. In the problems considered here, however, it is not the upper bound but the *ratio* of the quantities in (1.13) and (1.14) that is relevant. The presence of the parameter ℓ is essential for the success of their method.

In the current exposition, we make a choice that is a hybrid of the above and of Selberg's original approach. Our choice is most easily described in terms of y_r. We choose

$$y_{r,\ell} = y_{r,\ell}(\mathcal{H}) = \begin{cases} \dfrac{\mu^2(r)\mathfrak{S}(\mathcal{H})(\log R/r)^{\ell}}{\ell!} & \text{if } r < R, \\ 0 & \text{otherwise.} \end{cases} \tag{1.20}$$

As motivation for this choice, we note that $y_{r,0}$ is the optimal choice given in (1.18) with $\lambda_1 = V\mathfrak{S}(\mathcal{H})$. Moreover, one can show that

$$\mu(r)f_1(r)\sum_{d<R/r}\frac{\mu(dr)}{f(dr)}\frac{\log^{k+\ell}(R/rd)}{(k+\ell)!} \sim \frac{\mathfrak{S}(\mathcal{H})(\log R/r)^{\ell}}{\ell!}$$

when r is not too close to R. In other words, the choice of $\lambda_{d,\ell}$ in (1.19) gives a value of y_r that is asymptotic to the expression in (1.20).

One can use (1.17) and Möbius inversion to deduce that

$$\frac{\lambda_{d,\ell}}{f(d)} = \mu(d)\sum_r \frac{y_{dr,\ell}}{f_1(rd)}, \tag{1.21}$$

and so, when the choice of $y_{r,\ell}$ of (1.20) is specified, one obtains

$$\lambda_{d,\ell} = \mu(d)\frac{f(d)}{f_1(d)}\frac{\mathfrak{S}(\mathcal{H})}{\ell!}\sum_{\substack{r<R/d\\(r,d)=1}}\frac{\mu^2(r)}{f_1(r)}(\log R/rd)^{\ell} \tag{1.22}$$

when $d < R$. With this choice of $\lambda_{d,\ell}$, we set

$$\Lambda_R(n;\mathcal{H},\ell) = \sum_{d|P(n;\mathcal{H})}\lambda_{d,\ell}. \tag{1.23}$$

As we shall see, this choice $\lambda_{d,\ell}$ allows us to give elementary estimates for the main terms in (1.13) and (1.14).

We also define

$$\beta(\mathcal{H}) = \sum_p \frac{(k-\nu_p(\mathcal{H}))\log p}{p}. \tag{1.24}$$

This sum is finite because $\nu_p = k$ for sufficiently large p.

Theorems 1 through 4 will be derived fairly easily from the following results.

Theorem 5. *Suppose that $\mathcal{H} = \{h_1, \ldots, h_k\}$ is an admissible set, and that $0 \le \ell_1, \ell_2 \le k$. If $R \le N^{1/2-\epsilon}$, then*

$$\begin{aligned}\sum_{N<n\le 2N}&\Lambda_R(n;\mathcal{H},\ell_1)\Lambda_R(n;\mathcal{H},\ell_2)\\ &= \binom{\ell_1+\ell_2}{\ell_1}\mathfrak{S}(\mathcal{H})N\frac{(\log R)^{k+\ell_1+\ell_2}}{(k+\ell_1+\ell_2)!}\left\{1+O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})/\log R)\right\}.\end{aligned} \tag{1.25}$$

The implied constant depends at most on k.

Theorem 6. *Suppose that $\mathcal{H} = \{h_1, \ldots, h_k\}$. Suppose further that Hypothesis $BV(\theta)$ is true and $R \le N^{(\theta-\epsilon)/2}$. If $h_0 \in \mathcal{H}$, $\mathcal{H}$ is admissible, and $0 \le \ell_1, \ell_2 \le k$, then*

(1.26)

$$\begin{aligned}\sum_{N<n\le 2N}&\varpi(n+h_0)\Lambda_R(n;\mathcal{H},\ell_1)\Lambda_R(n;\mathcal{H},\ell_2)\\ &= \binom{\ell_1+\ell_2+2}{\ell_1+1}N\mathfrak{S}(\mathcal{H})\frac{(\log R)^{k+\ell_1+\ell_2+1}}{(k+\ell_1+\ell_2+1)!}\left\{1+O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})/\log R)\right\}.\end{aligned}$$

If $h_0 \notin \mathcal{H}$, $\mathcal{H}^0 = \mathcal{H}\cup\{h_0\}$ is admissible, and $1 \le \ell_1, \ell_2 \le k$, then

(1.27)

$$\begin{aligned}\sum_{N<n\le 2N}&\varpi(n+h_0)\Lambda_R(n;\mathcal{H},\ell_1)\Lambda_R(n;\mathcal{H},\ell_2)\\ &= \binom{\ell_1+\ell_2}{\ell_1}N\mathfrak{S}(\mathcal{H}^0)\frac{(\log R)^{k+\ell_1+\ell_2}}{(k+\ell_1+\ell_2)!}\left\{1+O(\beta(\mathcal{H}^0)\mathfrak{S}(\mathcal{H}^0)/\log R)\right\}.\end{aligned}$$

The implied constants depend at most on k.

With a bit more work, we could allow ℓ_1 or ℓ_2 to be 0 in (1.27). However, we omit this because the only place we use this result is in the proof of Theorem 1, where we will have $\ell_1 = \ell_2 > 0$.

Analogues of Theorems 5 and 6 are given in [10] for $\lambda_{d,\ell}$ given by (1.19). The corresponding main terms in [10] are evaluated with the help of contour integrals in two variables and zero-free regions for the Riemann-zeta function. On the other hand, with the choice of $\lambda_{d,\ell}$ given in (1.22), we are able to give an elementary treatment of the main terms in Theorems 5 and 6.

Theorem 7. *Suppose that $\mathcal{H} = \{h_1, \ldots, h_k\}$ is an admissible set, and that $0 \le \ell_1, \ell_2 \le k$. Suppose that Hypotheses $BV(\theta)$ and $BV_2(\theta)$ are both satisfied, and $R \le N^{(\theta-\epsilon)/2}$. If $h_0 \in \mathcal{H}$, then*

$$\sum_{N<n\le 2N} \varpi * \varpi(n+h_0)\Lambda_R(n;\mathcal{H},\ell_1)\Lambda_R(n;\mathcal{H},\ell_2)$$
$$= \left\{ \binom{\ell_1+\ell_2+2}{\ell_1+1} (N\log N)\mathfrak{S}(\mathcal{H}) \frac{(\log R)^{k+\ell_1+\ell_2+1}}{(k+\ell_1+\ell_2+1)!} \right.$$
$$\left. +2T(k,\ell_1,\ell_2)N\mathfrak{S}(\mathcal{H})\frac{(\log R)^{k+\ell_1+\ell_2+2}}{(k+\ell_1+\ell_2+2)!} \right\} \{1+O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})/\log R)\},$$

where

$$T(k,\ell_1,\ell_2) = -\binom{\ell_1+\ell_2+3}{\ell_2+1} - \binom{\ell_1+\ell_2+3}{\ell_1+1} + \binom{\ell_1+\ell_2+2}{\ell_1+1}.$$

The implied constant depends at most on k.

The reader will note that the sums considered here are more general than the sums in (1.13) and (1.14)–the latter correspond to the case $\ell_1 = \ell_2 = \ell$. We will see in Section 7 that this extra flexibility is useful in applications.

We also remark that the proof of Theorem 1 requires averaging over a set of $\mathcal{H}$, where the elements of $\mathcal{H}$ can be as large as $\log R$. Accordingly, we shall take some extra effort to make our estimates uniform in h, where $h := \max_{h_i\in\mathcal{H}} |h_i|$, and we assume that $h \le \log N$. For our results, it is not necessary to make the estimates in Theorems 5 through 7 uniform in k.

In a forthcoming paper [11], Goldston, Pintz, and Yıldırım will improve Theorem 1 to

$$\liminf_{n\to\infty} \frac{(p_{n+1}-p_n)}{(\log p_n)^{1/2}(\log\log p_n)^2} < \infty. \tag{1.28}$$

This result requires estimates of the kind given in Theorems 5 and 6 that hold uniformly for a wider range of k; this is in contrast to the situation here or in [10], where one needs only arbitrarily large but fixed k.

The implied constants in the error terms of Theorems 6 and 7 are ineffective due to the use of the Bombieri-Vinogradov Theorem, which uses the Siegel-Walfisz Theorem. However, the constants can be made effective by using the procedure of Section 12 of [11]. This procedure deletes the greatest prime factor of the eventually existing exceptional modulus from the sieve process.

A natural question is why we can get bounded gaps for E_2-numbers but not for primes when both sequences have the same level of distribution 1/2. The primary reason is that E_2 numbers are more prevalent than primes. Note, for example, that

$$\sum_{N<n\le 2N} \varpi * \varpi(n) \sim N\log N \sim \sum_{N<n\le 2N} \varpi(n)\log n,$$

but $\varpi * \varpi(n) \lesssim \frac{1}{4}(\log N)^2$ in the range $N < n \leq 2N$. The function $\varpi * \varpi$ used in Theorem 7 is convenient for calculations, but it is not optimal for applications. In a future paper, we will show that by using other functions supported on E_2's, the bound in Theorem 3 can be improved to 6. We will also show that there is a constant C such that for any positive integer r,

$$\liminf_{n\to\infty}(q_{n+r} - q_n) \leq Cre^r.$$

Moreover, these results on E_2's can be used to prove results on consecutive values of the divisor function. For example, we can show that there are infinitely many n such that the equations

$$d(n) = d(n+1) = 24,\ \Omega(n) = \Omega(n+1) = 5,\ \omega(n) = \omega(n+1) = 4$$

hold simultaneously. These results sharpen earlier theorems of Heath-Brown [16] and Schlage-Puchta [25].

Notation. The letters R, N denote real variables tending to infinity. The letter p is always used to denote a prime. The letters d, e, r are usually squarefree numbers; the letters m, n are usually positive integers. $[d, e]$ denotes the least common multiple of d and e. The notation $\omega(n)$ is used to denote the number of distinct prime factors of n. We use ρ to denote the function

$$\rho(r) = 1 + \sum_{p|r} \frac{\log p}{p}.$$

The letters $S, \mathcal{L}, U$, and V, with or without subscripts, are often used to denote sums. The meanings of these symbols are local to sections; e.g., the meaning of S_1 in Section 4 is different from the meaning of S_1 in Section 6.

We use $\sum^\flat$ to denote a summation over squarefree integers. In general, the constants implied by "O" and "$\ll$" will depend on k. Any other dependencies will be explicitly noted. As noted before, k is the size of $\mathcal{H}$; we always assume that $k \geq 2$. The parameter ℓ, with or without subscript, is an integer with $0 \leq \ell \leq k$.

2. Preliminary lemmas

The following two lemmas are classical estimates that have proved useful for handling remainder terms that arise in the Selberg sieve. The results can be found in Halberstam and Richert's book ([12], Lemmas 3.4 and 3.5). We reproduce the proofs here since they are quite short.

Lemma 8. *For any natural number h and for $x \geq 1$,*

$$\sum_{d\leq x}^{\flat} \frac{h^{\omega(d)}}{d} \leq (\log x + 1)^h,$$

$$\sum_{d\leq x}^{\flat} h^{\omega(d)} \leq x(\log x + 1)^h.$$

Proof. For the first inequality, we note that the sum on the left is

$$\sum_{d_1 \dots d_h \leq x} \frac{\mu^2(d_1 \dots d_h)}{d_1 \dots d_h} \leq \left(\sum_{n\leq x} \frac{1}{n}\right)^h \leq (\log x + 1)^h.$$

For the second inequality, we note that the left-hand side is at most

$$x\sum_{d\le x}^{\flat}\frac{h^{\omega(d)}}{d},$$

and we appeal to the first inequality. □

Lemma 9. *Assume Hypothesis $BV(\theta)$, and let h be a positive integer. Given any positive constant U and any $\epsilon>0$, then*

$$\sum_{d<N^{\theta-\epsilon}}^{\flat} h^{\omega(d)}E^*(N,d)\ll_{U,h,\epsilon} N(\log N)^{-U}.$$

Similarly, if Hypothesis $BV_2(\theta)$ is assumed, then

$$\sum_{d<N^{\theta-\epsilon}}^{\flat} h^{\omega(d)}E_2^*(N,d)\ll_{U,h,\epsilon} N(\log N)^{-U}.$$

Proof. We begin by noting the trivial estimate $E^*(N,d)\ll N(\log N)/d$. By Cauchy's inequality

$$\sum_{d<N^{\theta-\epsilon}}^{\flat} h^{\omega(d)}E^*(N,d)\le\left(N\log N\sum_{d<N^{\theta-\epsilon}}^{\flat}\frac{h^{2\omega(d)}}{d}\right)^{1/2}\left(\sum_{d<N^{\theta-\epsilon}}^{\flat}E^*(N,d)\right)^{1/2}$$
$$\ll_{h,\epsilon,A} N(\log N)^{(h^2-A+1)/2}.$$

We have used Lemma 8 and Hypothesis $BV(\theta)$ in the last line. The first result follows by taking $A=h^2+1+2U$. The second result is proved similarly; one uses the trivial bound $E_2^*(N,d)\ll N(\log N)^2/d$. □

Lemma 10. *Suppose $x\ge 1$. If a,b are positive real numbers, both at least 1, then*

$$\int_1^x(\log x/u)^{a-1}(\log u)^{b-1}\frac{du}{u}=(\log x)^{a+b-1}\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Proof. Upon making the change of variables $u=x^v$, the left-hand side becomes

$$(\log x)^{a+b-1}\int_0^1(1-v)^{a-1}v^{b-1}dv.$$

The result follows by the standard formula for the beta-integral. □

Our next lemma is another standard result in the theory of sieves.

Lemma 11. *Suppose that γ is a multiplicative function, and suppose that there are positive real numbers κ, A_1, A_2, L such that*

$$0\le\frac{\gamma(p)}{p}\le 1-\frac{1}{A_1}, \tag{2.1}$$

and

$$-L\le\sum_{w\le p<z}\frac{\gamma(p)\log p}{p}-\kappa\log\frac{z}{w}\le A_2 \tag{2.2}$$

if $2\le w\le z$. Let g be the multiplicative function defined by

$$g(d)=\prod_{p\mid d}\frac{\gamma(p)}{p-\gamma(p)}. \tag{2.3}$$

Then

$$\sum_{d<z}^{\flat} g(d) = c_\gamma \frac{(\log z)^\kappa}{\Gamma(\kappa+1)} \left\{ 1 + O_{A_1,A_2,\kappa}\left(\frac{L}{\log z} \right) \right\},$$

where

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

This is a combination of Lemmas 5.3 and 5.4 of Halberstam and Richert's book [12]. In [12], the hypothesis (2.1) is denoted (Ω_1), and hypothesis (2.2) is denoted $(\Omega_2(\kappa, L))$. As indicated above, the constant implied by "O" may depend on A_1, A_2, κ, but it is independent of L. This will be important in our applications.

Lemma 12. *Suppose that γ and g satisfy the same hypotheses as in the previous lemma. If a is a non-negative integer, then*

$$\sum_{r<R}^{\flat} g(r)(\log R/r)^a = c_\gamma \frac{\Gamma(a+1)}{\Gamma(\kappa+a+1)}(\log R)^{\kappa+a} + O_{A_1,A_2,\kappa,a}\left(L(\log R)^{\kappa+a-1}\right).$$

Proof. When $a = 0$, this is Lemma 11. If $a > 0$, then

$$\begin{aligned}\sum_{r<R}^{\flat} g(r)(\log R/r)^a &= a \sum_{r<R}^{\flat} g(r) \int_r^R (\log R/z)^{a-1} \frac{dz}{z} \\ &= \int_1^R \frac{a(\log R/z)^{a-1}}{z} \sum_{r<z}^{\flat} g(r) dz.\end{aligned}$$

Using Lemma 11, we see that the above is

$$\begin{aligned}&\int_1^R \frac{a(\log R/z)^{a-1}}{z} \left\{ \frac{c_\gamma (\log z)^\kappa}{\Gamma(\kappa+1)} + O(L(\log z)^{\kappa-1}) \right\} dz \\ &= \frac{a c_\gamma}{\Gamma(\kappa+1)} \int_1^R (\log R/z)^{a-1} (\log z)^\kappa \frac{dz}{z} + O\left(aL \int_1^R (\log R/z)^{a-1} (\log z)^{\kappa-1} \frac{dz}{z} \right).\end{aligned}$$

The desired result follows by using Lemma 10. □

Lemma 13. *If $\mathcal{H}$ is admissible and $|h_i| \le h$ for all $h_i \in \mathcal{H}$, then*

$$1 \ll \beta(\mathcal{H}) \ll \log\log 10h, \tag{2.4}$$

and there is a constant w_k (depending only on k) such that

$$1 \ll_k \mathfrak{S}(\mathcal{H}) \ll (\log\log 10h)^{w_k}. \tag{2.5}$$

Proof. Without loss of generality, we may assume that $h \ge 100$; this will simplify the writing of logarithms. We note that $\nu_p < k$ if and only if $p | \Delta(\mathcal{H})$, where

$$\Delta = \Delta(\mathcal{H}) := \prod_{1 \le i < j \le k} |h_i - h_j|. \tag{2.6}$$

Therefore

$$\beta(\mathcal{H}) = \sum_{p|\Delta} (k - \nu_p) \frac{\log p}{p},$$

where we have written ν_p as an abbreviation for $\nu_p(\mathcal{H})$. We may assume without loss of generality that $\Delta \ge 100$.

Now $\nu_2 = 1$ whenever $\mathcal{H}$ is admissible, so we see that $\beta(\mathcal{H}) \geq \log 2/2$. In the opposite direction, we have

$$\begin{aligned}\beta(\mathcal{H}) &\ll \sum_{p \leq \log\Delta} \frac{\log p}{p} + \sum_{\substack{p|\Delta \\ p > \log\Delta}} \frac{\log\log\Delta}{\log\Delta} \\ &\ll \log\log\Delta + \frac{\log\log\Delta}{\log\Delta}\frac{\log\Delta}{\log\log\Delta} \\ &\ll \log\log\Delta + 1.\end{aligned}$$

Finally, note that $\Delta \leq (2h)^{k^2}$, so that $\log\Delta \ll \log h$. This completes the proof of (2.4).

Now consider $\mathfrak{S}(\mathcal{H})$. From the definition of $\mathfrak{S}(\mathcal{H})$, we see that

$$\log\mathfrak{S}(\mathcal{H}) = \sum_p \left\{ \left(\frac{k-\nu_p}{p} \right) + O\left(\frac{1}{p^2} \right) \right\} \ll 1 + \sum_{p|\Delta} \frac{1}{p}. \tag{2.7}$$

The last sum may be bounded in a manner similar to that used for $\beta(\mathcal{H})$. We have

$$\begin{aligned}\sum_{p|\Delta}\frac{1}{p} &\leq \sum_{p\leq\log\Delta}\frac{1}{p} + \sum_{\substack{p|\Delta \\ p>\log\Delta}}\frac{1}{\log\Delta} \\ &\ll \log\log\log\Delta + \frac{1}{\log\Delta}\frac{\log\Delta}{\log\log\Delta} \\ &\ll \log\log\log\Delta.\end{aligned}$$

As noted before, $\log\Delta \ll \log h$. Therefore, there is some constant w_k such that $\log\mathfrak{S}(\mathcal{H}) \leq w_k \log\log\log h$, and the inequality on the right-hand side of (2.5) follows. For the inequality on the left-hand side of (2.5), we also use (2.7) and note that $\nu_p \leq k$ for all p, so $\log\mathfrak{S}(\mathcal{H}) \gg_k \sum_p 1/p^2 \gg_k 1$. □

In our final lemma of this section, we give a computation that was used in (1.10).

Lemma 14. *Suppose that q is an integer with all of its prime divisors less than $\sqrt{N}$. Then there is some absolute constant c such that*

$$\sum_{\substack{N<n\leq 2N \\ (n,q)=1}} \varpi * \varpi(n) = N\left(\log N + C_0 - 2\sum_{p|q}\frac{\log p}{p}\right) + O(N\exp(-c\sqrt{\log N})),$$

where

$$C_0 = 2\log 2 - 2\gamma - 1 - 2\sum_p \frac{\log p}{p(p-1)}. \tag{2.8}$$

Proof. We first use the hyperbola method to write

$$\begin{aligned}\sum_{n\leq x}\varpi*\varpi(n) =& 2\sum_{m\leq\sqrt{x}}\varpi(m)\sum_{n\leq x/m}\varpi(n) - \left(\sum_{m\leq\sqrt{x}}\varpi(m)\right)^2 \\ =& 2x\sum_{p\leq\sqrt{x}}\frac{\log p}{p} - x + O\left(x\exp(-c\sqrt{\log x})\right).\end{aligned}$$

Next, we use the estimate

$$\sum_{p\le x}\frac{\log p}{p}=\log x-\gamma-\sum_{p}\frac{\log p}{p(p-1)}+O(\exp(-c\sqrt{\log x})),$$

which can be easily derived by classical methods used in the proof of the Prime Number Theorem; see [21, Section 6.2.1, Exc. 4], for example. We get

$$\sum_{n\le x}\varpi*\varpi(n)=x\log x+C_1x+O(x\exp(-c\sqrt{\log x})),\tag{2.9}$$

where

$$C_1=-2\gamma-2\sum_{p}\frac{\log p}{p(p-1)}-1.$$

We use (2.9) with $x=N$, $x=2N$, and take differences to get

$$\sum_{N<n\le 2N}\varpi*\varpi(n)=N\log N+NC_0+O(N\exp(-c\sqrt{\log N})).\tag{2.10}$$

Finally, we note that for a given integer $q<\sqrt{N}$,

$$\begin{aligned}\sum_{p|q}\sum_{\substack{N<n\le 2N\\(n,q)=p}}\varpi*\varpi(n)=&2\sum_{p|q}\log p\sum_{N/p<n\le 2N/p}\varpi(n)\\=&2N\sum_{p|q}\frac{\log p}{p}+O(N\exp(-c\sqrt{\log N})).\end{aligned}\tag{2.11}$$

The lemma follows by combining (2.10) and (2.11). □

3. Proof of Theorem 5

As we noted in the Introduction, we take $\nu_p(\mathcal{H})$ to be the number of distinct residue classes mod p in $\mathcal{H}$. We extend this definition to arbitrary squarefree moduli d as follows. Let $\mathbb{Z}_d$ be the ring of integers mod d and define

$$\Omega_d(\mathcal{H})=\{a\in\mathbb{Z}_d : P(a;\mathcal{H})\equiv 0 \pmod d\}.\tag{3.1}$$

We define $\nu_d(\mathcal{H})$ to be the cardinality of $\Omega_d(\mathcal{H})$.

Assume that d_1,d_2 are squarefree numbers with $(d_1,d_2)=1$. The Chinese Remainder Theorem gives an isomorphism

$$\xi:\mathbb{Z}_{d_1}\times\mathbb{Z}_{d_2}\to\mathbb{Z}_{d_1d_2}.\tag{3.2}$$

The set $\Omega_{d_1d_2}(\mathcal{H})$ is the image of $\Omega_{d_1}(\mathcal{H})\times\Omega_{d_2}(\mathcal{H})$ under the isomorphism ξ, so $\nu_d(\mathcal{H})$ is multiplicative.

Throughout this section, we will take $\mathcal{H}$ to be a fixed admissible set, and we will usually write ν_d in place of $\nu_d(\mathcal{H})$.

The left-hand side of (1.25) is

$$\begin{aligned}(3.3)\qquad &\sum_{N<n\le 2N}\left(\sum_{d|P(n;\mathcal{H})}\lambda_{d,\ell_1}\right)\left(\sum_{e|P(n;\mathcal{H})}\lambda_{e,\ell_2}\right)\\ &=\sum_{d,e}\lambda_{d,\ell_1}\lambda_{e,\ell_2}\sum_{\substack{N<n\le 2N\\ [d,e]|P(n;\mathcal{H})}}1\\ &=N\sum_{d,e}\frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}}{f([d,e])}+O\left(\sum_{d,e}|\lambda_{d,\ell_1}\lambda_{e,\ell_2}r_{[d,e]}|\right)\\ &=NS_1+O(S_2),\end{aligned}$$

say, where

$$f(d)=\frac{d}{\nu_d} \tag{3.4}$$

and

$$r_d=\sum_{\substack{N<n\le 2N\\ d|P(n;\mathcal{H})}}1-\frac{N}{f(d)}.$$

The estimates of S_1 and S_2 require the following two lemmas. Recall that $f_1 = f*\mu$.

Lemma 15. *We have*

$$\sum_{r<R}\frac{\mu^2(r)}{f_1(r)}(\log R/r)^\ell=\frac{\ell!(\log R)^{k+\ell}}{\mathfrak{S}(\mathcal{H})(k+\ell)!}\left\{1+O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})/\log R)\right\}.$$

Proof. We apply Lemma 12 with

$$\gamma(p)=\nu_p,\quad g(p)=\frac{\nu_p}{p-\nu_p}=\frac{1}{f_1(p)}.$$

Now $\nu_p\le\min(k,p-1)$, so (2.1) holds with $A_1=k+1$. Moreover,

$$-\beta(\mathcal{H})\le\sum_{w\le p<z}\frac{(\nu_p-k)\log p}{p}\le 0$$

and

$$\sum_{w\le p<z}\frac{\log p}{p}=\log(z/w)+O(1).$$

Therefore (2.2) holds with $\kappa=k$, A_2 some constant depending only on k, and

$$L\ll 1+\beta(\mathcal{H})\ll\beta(\mathcal{H}).$$

Finally, we note that

$$c_\gamma=\prod_p\left(1-\frac{\nu_p}{p}\right)^{-1}\left(1-\frac{1}{p}\right)^k=\frac{1}{\mathfrak{S}(\mathcal{H})}.$$

□

Lemma 16. *Let $\lambda_{d,\ell}$ be as defined in* (1.22). *If $d<R$ and d is squarefree, then*

$$|\lambda_{d,\ell}|\ll(\log R)^{k+\ell}.$$

Proof. From (1.22), we see that if d satisfies the hypotheses of the lemma, then

$$\begin{aligned}|\lambda_{d,\ell}| &= \frac{\mathfrak{S}(\mathcal{H})}{\ell!}\frac{f(d)}{f_1(d)}\sum_{\substack{r<R/d\\(r,d)=1}}\frac{\mu^2(r)}{f_1(r)}(\log R/rd)^\ell\\ &= \frac{\mathfrak{S}(\mathcal{H})}{\ell!}\sum_{t|d}\frac{1}{f_1(t)}\sum_{\substack{r<R/d\\(r,d)=1}}\frac{\mu^2(r)}{f_1(r)}(\log R/rd)^\ell.\end{aligned}$$

We move the factor $1/f_1(t)$ inside the sum and write $s = rt$ to get

$$\begin{aligned}|\lambda_{d,\ell}| &= \frac{\mathfrak{S}(\mathcal{H})}{\ell!}\sum_{t|d}\sum_{\substack{r<R/d\\(r,d)=1}}\frac{\mu^2(r)}{f_1(rt)}(\log R/rd)^\ell\\ &= \frac{\mathfrak{S}(\mathcal{H})}{\ell!}\sum_{t|d}\sum_{\substack{s<Rt/d\\(s,d)=t}}\frac{\mu^2(s)}{f_1(s)}(\log Rt/sd)^\ell.\end{aligned}$$

For any $t|d$, we have $Rt/d \le R$, so

$$|\lambda_{d,\ell}| \le \frac{\mathfrak{S}(\mathcal{H})}{\ell!}(\log R)^\ell\sum_{t|d}\sum_{\substack{s<R\\(s,d)=t}}\frac{\mu^2(s)}{f_1(s)}.$$

Now for any $s < R$, there is a unique $t|d$ such that $(s, d) = t$. Therefore

$$|\lambda_{d,\ell}| \le \frac{\mathfrak{S}(\mathcal{H})}{\ell!}(\log R)^\ell\sum_{s<R}\frac{\mu^2(s)}{f_1(s)}.$$

To complete the proof, we use Mertens' Theorem and observe that

$$\begin{aligned}\sum_{s<R}\frac{\mu^2(s)}{f_1(s)} &\le \prod_{p<R}\left(1+\frac{1}{f_1(p)}\right)\\ &= \prod_{p<R}\left(1-\frac{\nu_p}{p}\right)^{-1}\left(1-\frac{1}{p}\right)^k\prod_{p<R}\left(1-\frac{1}{p}\right)^{-k}\\ &\ll \frac{(\log R)^k}{\mathfrak{S}(\mathcal{H})}.\end{aligned}$$

□

We now treat S_1 and S_2. For S_1, we begin by writing

$$\begin{aligned}
S_1 &= \sum_{d,e} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}}{f(d)f(e)} \sum_{\substack{r|d \\ r|e}} f_1(r) \\
&= \sum_r^\flat f_1(r) \left(\sum_d \frac{\lambda_{dr,\ell_1}}{f(dr)} \right) \left(\sum_e \frac{\lambda_{er,\ell_2}}{f(er)} \right) \\
&= \sum_r^\flat \frac{y_{r,\ell_1} y_{r,\ell_2}}{f_1(r)} \\
&= \frac{\mathfrak{S}(\mathcal{H})^2}{\ell_1! \ell_2!} \sum_{r<R} \frac{\mu^2(r)(\log R/r)^{\ell_1+\ell_2}}{f_1(r)},
\end{aligned}$$

where we have used (1.17) and (1.20) in the last two lines. Lemma 15 now yields the estimate

$$S_1 = \binom{\ell_1+\ell_2}{\ell_1} \mathfrak{S}(\mathcal{H}) \frac{(\log R)^{k+\ell_1+\ell_2}}{(k+\ell_1+\ell_2)!} \left\{ 1 + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})/\log R) \right\}.$$

For S_2, we first note that

$$|r_d| \le \nu_d \le k^{\omega(d)}.$$

We also have the bound for $\lambda_{d,\ell}$ given in Lemma 16. Therefore

$$\begin{aligned}
S_2 &= \sum_{d,e<R} |\lambda_{d,\ell_1}\lambda_{e,\ell_2} r_{[d,e]}| \\
&\ll (\log R)^{2k+\ell_1+\ell_2} \sum_{d,e\le R}^\flat k^{\omega([d,e])} \\
&\ll (\log R)^{4k} \sum_{r<R^2}^\flat (3k)^{\omega(r)}.
\end{aligned}$$

Using Lemma 8, we get

$$S_2 \ll R^2 (\log R)^{7k} \ll N^{1-\epsilon}, \tag{3.5}$$

provided $R < N^{1/2-\epsilon}$.

Theorem 5 follows by combining the above estimates for S_1 and S_2.

4. Proof of Theorem 6, Part 1

In this section, we consider Theorem 6 under the assumption that $h_0 \in \mathcal{H}$. Our problem is translation invariant in $\mathcal{H}$, so we may, without loss of generality, assume that $h_0 = 0$ and $0 \in \mathcal{H}$.

Let $\mathcal{L}$ denote the sum on the left-hand side of (1.26). Then

(4.1)
$$\mathcal{L} = \sum_{d,e} \lambda_{d,\ell_1}\lambda_{e,\ell_2} \sum_{\substack{N<n\le 2N \\ [d,e]|P(n;\mathcal{H})}} \varpi(n) = \sum_{d,e} \lambda_{d,\ell_1}\lambda_{e,\ell_2} \sum_{a\in\Omega_{[d,e]}(\mathcal{H})} \sum_{\substack{N<p\le 2N \\ p\equiv a \pmod{[d,e]}}} \log p.$$

Now all prime divisors of $[d,e]$ are $< R$, and $R < N$. Therefore, the innermost sum in (4.1) is 0 if $(a,[d,e]) \ne 1$. Accordingly, we need an analogue of $\Omega_d(\mathcal{H})$ for reduced residue classes. For squarefree d, we define

$$\Omega_d^*(\mathcal{H}) = \{a \in \mathbb{Z}_d : (a,d) = 1 \text{ and } P(a;\mathcal{H}) \equiv 0 \pmod d\}. \tag{4.2}$$

Let $\nu_d^*(\mathcal{H})$ be the cardinality of $\Omega_d^*(\mathcal{H})$. For brevity, we will usually write ν_d^* in place of $\nu_d^*(\mathcal{H})$.

When d_1, d_2 are squarefree and $(d_1, d_2) = 1$, the set $\Omega_{d_1 d_2}^*(\mathcal{H})$ is the image of $\Omega_{d_1}^* \times \Omega_{d_2}^*$ under the isomorphism ξ of (3.2). Therefore, the function ν^* is multiplicative. Moreover, when p is prime,

$$\nu_p^* = \nu_p - 1,$$

because we are assuming that $0 \in \mathcal{H}$.

In this context, the most natural analogue of $\mathfrak{S}(\mathcal{H})$ is the product

$$\mathfrak{S}^*(\mathcal{H}) = \prod_p \left(1 - \frac{\nu_p^*}{p-1}\right)\left(1 - \frac{1}{p}\right)^{-k+1}. \tag{4.3}$$

Note, however, that

$$\begin{aligned}\mathfrak{S}^*(\mathcal{H}) &= \prod_p \left(1 - \frac{\nu_p - 1}{p-1}\right)\left(1 - \frac{1}{p}\right)^{-k+1} \\ &= \prod_p \left(\frac{p-\nu_p}{p-1}\right)\left(\frac{p-1}{p}\right)\left(1 - \frac{1}{p}\right)^{-k} \\ &= \mathfrak{S}(\mathcal{H}).\end{aligned} \tag{4.4}$$

Returning to $\mathcal{L}$, we write this sum as

$$\mathcal{L} = \sum_{d,e} \lambda_{d,\ell_1}\lambda_{e,\ell_2} \sum_{a \in \Omega_{[d,e]}^*(\mathcal{H})} \sum_{\substack{N < p \le 2N \\ p \equiv a \pmod{[d,e]}}} \log p = NS + O(T), \tag{4.5}$$

where

$$S = \sum_{d,e} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}\nu_{[d,e]}^*}{\phi([d,e])} \tag{4.6}$$

and

$$T = \sum_{d,e} |\lambda_{d,\ell_1}\lambda_{e,\ell_2}| \nu_{[d,e]}^* E^*(N, [d,e]).$$

By Lemma 16 and Lemma 9,

$$T \ll (\log R)^{2k+\ell_1+\ell_2} \sum_{r<R^2}^{\flat} (3k-3)^{\omega(r)} E^*(N,r) \ll N/\log N. \tag{4.7}$$

We now consider the sum S. We shall define

$$f^*(r) = \frac{\phi(r)}{\nu_r^*}. \tag{4.8}$$

However, we need to take some care with this definition because there may be terms with $\nu_r^* = 0$. However, $\nu_p^* = k-1$ for all but finitely many primes p, so there are at most finitely many primes p such that $\nu_p^* = 0$. We define

$$A = A(\mathcal{H}) = \prod_{\substack{p \\ \nu_p^*(\mathcal{H})=0}} p, \tag{4.9}$$

and we use the definition in (4.8) for any r with $(r, A) = 1$. We define f_1^*, a function analogous to f_1, by taking

$$f_1^*(r) = (f^* * \mu)(r)$$

for r with $(r,A)=1$. For future reference, we note that if p is a prime and $p\nmid A$, then

$$f^*(p)=\frac{p-1}{\nu_p-1},\quad f_1^*(p)=\frac{p-\nu_p}{\nu_p-1}.$$

With this definition of f^*, we now have

$$S=\sideset{}{'}\sum_{d,e}\frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}}{f^*(d)f^*(e)}\sum_{\substack{r|d\\ r|e}}f_1^*(r).$$

Here, and in the sequel, we use $\sum'$ to denote that the sum is over values of the indices that are relatively prime to A. Interchanging the order of summation, we get

$$\begin{aligned}S&=\sideset{}{'}\sum_{r}f_1^*(r)\left(\sideset{}{'}\sum_{d}\frac{\lambda_{dr,\ell_1}}{f^*(dr)}\right)\left(\sideset{}{'}\sum_{e}\frac{\lambda_{er,\ell_2}}{f^*(er)}\right)\\&=\sideset{}{'}\sum_{r}\frac{y^*_{r,\ell_1}y^*_{r,\ell_2}}{f_1^*(r)},\end{aligned}\tag{4.10}$$

where the quantity $y^*_{r,\ell}$ is analogous to $y_{r,\ell}$ and is defined as

$$y^*_{r,\ell}=\begin{cases}\mu(r)f_1^*(r)\sideset{}{'}\sum_{d}\dfrac{\lambda_{dr,\ell}}{f^*(dr)} & \text{if } (r,A)=1 \text{ and } r<R,\\ 0 & \text{otherwise.}\end{cases}\tag{4.11}$$

Upon using (1.21), the original definition of $\lambda_{d,\ell}$, we see that

$$\begin{aligned}\frac{\mu(r)y^*_{r,\ell}}{f_1^*(r)}&=\sideset{}{'}\sum_{d}\frac{\lambda_{dr,\ell}}{f^*(dr)}=\sideset{}{'}\sum_{d}\frac{\mu(dr)}{f^*(dr)}f(dr)\sum_{t}\frac{y_{rdt,\ell}}{f_1(rdt)}\\&=\frac{\mu(r)f(r)}{f^*(r)f_1(r)}\sideset{}{'}\sum_{\substack{d\\(d,r)=1}}\frac{\mu(d)f(d)}{f^*(d)}\sum_{t}\frac{y_{rdt,\ell}}{f_1(dt)}\\&=\frac{\mu(r)f(r)}{f^*(r)f_1(r)}\sum_{\substack{m\\(m,r)=1}}\frac{y_{rm,\ell}}{f_1(m)}\sideset{}{'}\sum_{d|m}\frac{\mu(d)f(d)}{f^*(d)}.\end{aligned}$$

Note that m can be any squarefree integer; we need not have $(m,A)=1$. Now

$$\begin{aligned}\sideset{}{'}\sum_{d|m}\frac{\mu(d)f(d)}{f^*(d)}&=\prod_{p|m,p\nmid A}\left(1-\frac{f(p)}{f^*(p)}\right)\\&=\prod_{p|m,p\nmid A}\left(\frac{p-\nu_p}{\nu_p(p-1)}\right)\\&=\prod_{p|m}\left(\frac{p-\nu_p}{\nu_p(p-1)}\right).\end{aligned}$$

We may drop the condition that $p\nmid A$ in the last line because when $p|A$, $\nu_p=1$, and $(p-\nu_p)/(\nu_p(p-1))=1$. Therefore

$$\frac{1}{f_1(m)}\sideset{}{'}\sum_{d|m}\frac{\mu(d)f(d)}{f^*(d)}=\prod_{p|m}\frac{p-\nu_p}{\nu_p(p-1)f_1(p)}=\frac{1}{\phi(m)}.\tag{4.12}$$

Moreover,

$$\frac{f_1^*(r)f(r)}{f^*(r)f_1(r)} = \frac{r}{\phi(r)} \tag{4.13}$$

when $(r, A) = 1$, and so

$$y_{r,\ell}^* = \mu^2(r)\frac{\mathfrak{S}(\mathcal{H})}{\ell!}\frac{r}{\phi(r)}\sum_{\substack{m<R/r\\(m,r)=1}}\frac{\mu^2(m)}{\phi(m)}(\log R/rm)^\ell \tag{4.14}$$

when $(r, A) = 1$.

For the inner sum, we use Lemma 12 with

$$\gamma(p) = \begin{cases} 1 & \text{if } p \nmid r, \\ 0 & \text{if } p | r. \end{cases}$$

The hypotheses (2.1) and (2.2) are satisfied with $\kappa = 1$, some absolute constants A_1, A_2, and

$$L = \sum_{p|r}\frac{\log p}{p} + O(1).$$

Let

$$\rho(r) = 1 + \sum_{p|r}\frac{\log p}{p}, \tag{4.15}$$

so that $L \ll \rho(r)$. With this choice of γ, we have

$$c_\gamma = \prod_{p|r}\left(1 - \frac{1}{p}\right) = \frac{\phi(r)}{r}.$$

We therefore conclude that

$$\sum_{\substack{m<R/r\\(m,r)=1}}\frac{\mu^2(m)}{\phi(m)}(\log R/rm)^\ell = \frac{\phi(r)}{r}\frac{(\log R/r)^{\ell+1}}{\ell+1} + O\left(\rho(r)(\log 2R/r)^\ell\right). \tag{4.16}$$

We remark parenthetically that Hildebrand [18] gave a more precise formula for this sum in the case $\ell = 0$. It is possible to use his result to derive a more accurate version of (4.16), but the above version is sufficient for our purposes.

From (4.16) and (4.14), we deduce that when $(r, A) = 1$ and $r < R$,

$$y_{r,\ell}^* = \mu^2(r)\frac{\mathfrak{S}(\mathcal{H})}{(\ell+1)!}(\log R/r)^{\ell+1} + O\left(\frac{\mu^2(r)\rho(r)r}{\phi(r)}\mathfrak{S}(\mathcal{H})(\log 2R/r)^\ell\right). \tag{4.17}$$

We plug this back into our formula for S in (4.10) and use the simple observation that $\rho(r)r/\phi(r) \ll \log r \ll \log R$ to get

$$S = \sum_{r<R}{}' \frac{y_{r,\ell_1}^* y_{r,\ell_2}^*}{f_1^*(r)} = V + O\left(\mathfrak{S}(\mathcal{H})^2(\log R)^{\ell_1+\ell_2+1}W\right), \tag{4.18}$$

where

$$V = \frac{\mathfrak{S}(\mathcal{H})^2}{(\ell_1+1)!(\ell_2+1)!}\sum_{r<R}{}' \frac{\mu^2(r)}{f_1^*(r)}(\log R/r)^{\ell_1+\ell_2+2} \tag{4.19}$$

and

$$W = \sum_{r<R}{}' \frac{\mu^2(r)}{f_1^*(r)} \frac{\rho(r)r}{\phi(r)}. \tag{4.20}$$

We will use Lemma 12 for V. We will need to estimate a similar sum in Section 6, so it is convenient to have the following lemma that is general enough to cover both situations.

Lemma 17. *If d is squarefree, $d < R$, and a is a non-negative integer, then*

$$\begin{aligned}\sum_{\substack{r<R/d \\ (r,d)=1}}{}' \frac{\mu^2(r)}{f_1^*(r)} (\log R/dr)^a =& \frac{1}{\mathfrak{S}(\mathcal{H})} \frac{a!}{(k+a-1)!} (\log R/d)^{k+a-1} \prod_{p|d} \left(\frac{p-\nu_p}{p-1}\right) \\ &+ O\left((\beta(\mathcal{H}) + \rho(d))(\log 2R/d)^{k+a-2}\right).\end{aligned}$$

Proof. We apply Lemma 12 with

$$\gamma(p) = \begin{cases} \dfrac{p\nu_p^*}{p-1} & \text{if } (p,d) = 1, \\ 0 & \text{if } p|d. \end{cases}$$

With this definition for γ, we have

$$g(p) = \frac{\gamma(p)}{p-\gamma(p)} = \frac{1}{f_1^*(p)}$$

when $(p, Ad) = 1$. Moreover,

$$\nu_p^* = \nu_p - 1 \le \min(k-1, p-2),$$

so (2.1) is true with $A_1 = k$. For (2.2), we first note that

$$\begin{aligned}\sum_{w\le p<z} \frac{\gamma(p)\log p}{p} &= \sum_{\substack{w\le p<z \\ (p,d)=1}} \frac{(\nu_p - 1)\log p}{p-1} \\ &= (k-1) \sum_{w\le p<z} \frac{\log p}{p-1} - \sum_{\substack{w\le p<z \\ (p,d)=1}} \frac{(k-\nu_p)\log p}{p-1} - \sum_{\substack{w\le p<z \\ p|d}} \frac{(k-1)\log p}{p-1}.\end{aligned}$$

Now

$$\begin{aligned}\sum_{w\le p<z} \frac{\log p}{p-1} &= \log(z/w) + O(1), \\ \sum_{\substack{w\le p<z \\ (p,d)=1}} \frac{(k-\nu_p)\log p}{p-1} &\le \beta(\mathcal{H}) + O(1), \\ \sum_{\substack{w\le p<z \\ p|d}} \frac{(k-1)\log p}{p-1} &\le (k-1)\rho(d),\end{aligned}$$

so (2.2) is satisfied with $\kappa = k-1$, A_2 some constant depending only on k, and $L = \beta(\mathcal{H}) + (k-1)\rho(d) + O(1) \ll \beta(\mathcal{H}) + \rho(d)$. Finally, we note that in this situation,

$$c_\gamma = \prod_p \left(1 - \frac{\nu_p^*}{p-1}\right)^{-1} \left(1-\frac{1}{p}\right)^{k-1} \prod_{p|d} \left(1 - \frac{\nu_p^*}{p-1}\right)$$
$$= \frac{1}{\mathfrak{S}(\mathcal{H})} \prod_{p|d} \left(\frac{p-\nu_p}{p-1}\right)$$

by (4.4). □

From the previous lemma, with $d = 1$, we see that

(4.21)
$$V = \frac{\mathfrak{S}(\mathcal{H})}{(\ell_1+1)!(\ell_2+1)!} \frac{(\ell_1+\ell_2+2)!}{(k+\ell_1+\ell_2+1)!} (\log R)^{k+\ell_1+\ell_2+1} + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2 (\log R)^{k+\ell_1+\ell_2})$$
$$= \binom{\ell_1+\ell_2+2}{\ell_1+1} \mathfrak{S}(\mathcal{H}) \frac{(\log R)^{k+\ell_1+\ell_2+1}}{(k+\ell_1+\ell_2+1)!} + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2}).$$

The sum W may be estimated by relatively trivial means. Now

(4.22)
$$W = \sum_{r<R}{}' \frac{\mu^2(r)r}{f_1^*(r)\phi(r)} \left(1 + \sum_{p|r} \frac{\log p}{p}\right)$$
$$= \sum_{r<R}{}' \frac{\mu^2(r)r}{f_1^*(r)\phi(r)} + \sum_{p<R}{}' \frac{\log p}{p} \sum_{\substack{r<R \\ p|r}}{}' \frac{\mu^2(r)r}{f_1^*(r)\phi(r)}$$
$$= \sum_{r<R}{}' \frac{\mu^2(r)r}{f_1^*(r)\phi(r)} + \sum_{p<R}{}' \frac{\log p}{f_1^*(p)\phi(p)} \sum_{\substack{r<R/p \\ (r,p)=1}}{}' \frac{\mu^2(r)r}{f_1^*(r)\phi(r)}$$
$$\ll \left(1 + \sum_{p<R}{}' \frac{\log p}{f_1^*(p)\phi(p)}\right) W^* \ll W^*,$$

where

(4.23)
$$W^* = \sum_{r<R}{}' \frac{\mu^2(r)r}{f_1^*(r)\phi(r)} = \sum_{r<R}{}^\flat \frac{\nu_r^* h(r)}{r}$$

and

(4.24)
$$h(r) = \prod_{p|r} \frac{p^2}{(p-\nu_p)(p-1)}.$$

Let $h_1 = h * \mu$, so that

$$h_1(d) = \prod_{p|d} \frac{p(\nu_p+1)-\nu_p}{(p-1)(p-\nu_p)}.$$

Then

(4.25)
$$W^* = \sum_{r<R}{}^\flat \frac{\nu_r^*}{r} \sum_{d|r} h_1(d) = \sum_{d<R}{}^\flat \frac{h_1(d)\nu_d^*}{d} \sum_{\substack{r<R/d \\ (r,d)=1}}{}^\flat \frac{\nu_r^*}{r} \le \prod_{p<R}\left(1 + \frac{h_1(p)\nu_p^*}{p}\right) \sum_{r<R}{}^\flat \frac{\nu_r^*}{r}.$$

The sum on the right-hand side of (4.25) is $\ll (\log R)^{k-1}$ by Lemma 8. The product is $\ll 1$ because

$$\sum_{p<R} \log\left(1 + \frac{h_1(p)\nu_p^*}{p}\right) \ll \sum_{p<R} \frac{\nu_p^2}{(p-1)(p-\nu_p)} \ll 1.$$

We conclude that $W^* \ll (\log R)^{k-1}$, and so

$$W \ll (\log R)^{k-1}. \tag{4.26}$$

Combining the above with the estimate in (4.21) gives

(4.27)

$$S = \binom{\ell_1+\ell_2+2}{\ell_1+1}\mathfrak{S}(\mathcal{H})\frac{(\log R)^{k+\ell_1+\ell_2+1}}{(k+\ell_1+\ell_2+1)!} + O\left(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2}\right).$$

The first part of Theorem 6 (statement (1.26)) now follows by combining (4.5), (4.7), and (4.27).

5. Proof of Theorem 6, Part 2

In this section, we consider Theorem 6 in the case $h_0 \notin \mathcal{H}$. As in the previous section, our problem is translation invariant, so we we may assume that $0 \notin \mathcal{H}$ and $\mathcal{H}^0 = \mathcal{H} \cup \{0\}$. Consequently, $P(n;\mathcal{H}^0) = nP(n;\mathcal{H})$.

Now let $\mathcal{L}$ be the left-hand side of (1.27). If n is a prime with $N < n \le 2N$, then 1 is the only divisor of n less than N. When $d < R < N$, we have $d|P(n;\mathcal{H})$ if and only if $d|P(n;\mathcal{H}^0)$. Consequently,

(5.1)

$$\mathcal{L} = \sum_{d,e} \lambda_{d,\ell_1}\lambda_{e,\ell_2} \sum_{\substack{N<n\le 2N \\ [d,e]|P(n;\mathcal{H}^0)}} \varpi(n) = \sum_{d,e} \lambda_{d,\ell_1}\lambda_{e,\ell_2} \sum_{a\in\Omega^*_{[d,e]}(\mathcal{H}^0)} \sum_{\substack{N<p\le 2N \\ p\equiv a \pmod{[d,e]}}} \log p.$$

Parallel to the argument in (4.5) through (4.7), we find that

$$\mathcal{L} = NS + O(T),$$

where

$$S = \sum_{d,e} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}\nu^*_{[d,e]}(\mathcal{H}^0)}{\phi([d,e])} \tag{5.2}$$

and

$$T = \sum_{d,e} |\lambda_{d,\ell_1}\lambda_{e,\ell_2}|\nu^*_{[d,e]}(\mathcal{H}^0)E^*(N,[d,e]) \ll N/\log N.$$

Therefore

$$\mathcal{L} = NS + O(N/\log N). \tag{5.3}$$

The rest of this section is devoted to evaluating the sum S.

For brevity, we write $\nu_r^\dagger$ for $\nu_r^*(\mathcal{H}^0)$. Let

$$A_0 = A(\mathcal{H}^0) = \prod_{\substack{p \\ \nu_p^\dagger = 0}} p.$$

For squarefree r with $(r, A_0) = 1$, we define

$$f^\dagger(r) = \frac{\phi(r)}{\nu_r^\dagger} = \prod_{p|r}\left(\frac{p-1}{\nu_p^\dagger}\right) \tag{5.4}$$

and

$$f_1^\dagger(r) = f^\dagger * \mu(r) = \prod_{p|r} \left(\frac{p-1-\nu_p^\dagger}{\nu_p^\dagger} \right). \tag{5.5}$$

Note that

$$\nu_p^\dagger = \begin{cases} \nu_p & \text{if } 0 \notin \Omega_p(\mathcal{H}), \\ \nu_p - 1 & \text{if } 0 \in \Omega_p(\mathcal{H}). \end{cases}$$

We are assuming that $0 \notin \mathcal{H}$, so there are only finitely many primes p with $0 \in \Omega_p(\mathcal{H})$. Let

$$B_0 = B_0(\mathcal{H}) = \prod_{\substack{p \\ \nu_p^\dagger = \nu_p - 1}} p = \prod_{\substack{p \\ 0 \in \Omega_p(\mathcal{H})}} p. \tag{5.6}$$

In fact, $0 \in \Omega_p(\mathcal{H})$ if and only if p divides h for some $h \in \mathcal{H}$. Therefore B_0 is the squarefree kernel of the product of all elements of $\mathcal{H}$.

For future reference, we note that when $(r, A_0) = 1$,

$$f^\dagger(r) = \prod_{\substack{p|r \\ p \nmid B_0}} \left(\frac{p-1}{\nu_p} \right) \prod_{\substack{p|r \\ p|B_0}} \left(\frac{p-1}{\nu_p - 1} \right)$$

and

$$f_1^\dagger(r) = \prod_{\substack{p|r \\ p \nmid B_0}} \left(\frac{p-1-\nu_p}{\nu_p} \right) \prod_{\substack{p|r \\ p|B_0}} \left(\frac{p-\nu_p}{\nu_p - 1} \right).$$

With the above definitions of $f^\dagger$ and $f_1^\dagger$, we may write

$$\begin{aligned} S &= \sum_{d,e}{}' \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}}{f^\dagger([d,e])} \\ &= \sum_{d,e}{}' \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}}{f^\dagger(d) f^\dagger(e)} \sum_{\substack{r|d \\ r|e}} f_1^\dagger(r) \\ &= \sum_{r}{}' f_1^\dagger(r) \left(\sum_{d}{}' \frac{\lambda_{dr,\ell_1}}{f^\dagger(dr)} \right) \left(\sum_{e}{}' \frac{\lambda_{er,\ell_2}}{f^\dagger(er)} \right), \end{aligned}$$

where $\sum'$ denotes that the sum is over values of the indices that are relatively prime to A_0. We get

$$S = \sum_{r}{}' \frac{y_{r,\ell_1}^\dagger y_{r,\ell_2}^\dagger}{f_1^\dagger(r)}, \tag{5.7}$$

where we define

$$y_{r,\ell}^\dagger = \begin{cases} \mu(r) f_1^\dagger(r) \sum_{d}{}' \dfrac{\lambda_{dr,\ell}}{f^\dagger(dr)} & \text{if } (r, A_0) = 1 \text{ and } r < R, \\ 0 & \text{otherwise.} \end{cases} \tag{5.8}$$

Upon using (1.21), our original definition of $\lambda_{d,\ell}$, we see that

$$\begin{aligned}\frac{\mu(r)y^\dagger_{r,\ell}}{f_1^\dagger(r)} &= \sum_d{}' \frac{\lambda_{dr,\ell}}{f^\dagger(dr)} = \sum_d{}' \frac{\mu(dr)}{f^\dagger(dr)} f(dr) \sum_t \frac{y_{rdt,\ell}}{f_1(rdt)}\\ &= \frac{\mu(r)f(r)}{f^\dagger(r)f_1(r)} \sum_{\substack{d\\(d,r)=1}}{}' \frac{\mu(d)f(d)}{f^\dagger(d)} \sum_t \frac{y_{rdt,\ell}}{f_1(dt)}\\ &= \frac{\mu(r)f(r)}{f^\dagger(r)f_1(r)} \sum_{\substack{m\\(m,r)=1}} \frac{y_{rm,\ell}}{f_1(m)} \sum_{d|m}{}' \frac{\mu(d)f(d)}{f^\dagger(d)}.\end{aligned}$$

Now

$$\sum_{d|m}{}' \frac{\mu(d)f(d)}{f^\dagger(d)} = \prod_{\substack{p|m\\p\nmid A_0}} \left(1 - \frac{f(p)}{f^\dagger(p)}\right) = \prod_{p|m} \left(1 - \frac{p\nu_p^\dagger}{(p-1)\nu_p}\right).$$

The condition $p \nmid A_0$ can be dropped because $\nu_p^\dagger = 0$ when $p|A_0$. Therefore

$$\begin{aligned}\sum_{d|m}{}' \frac{\mu(d)f(d)}{f^\dagger(d)} &= \prod_{\substack{p|m\\p\nmid B_0}} \left(1 - \frac{p}{p-1}\right) \prod_{\substack{p|m\\p|B_0}} \left(1 - \frac{p(\nu_p - 1)}{\nu_p(p-1)}\right)\\ &= \frac{\mu(m)}{\phi(m)} f_2(m),\end{aligned}$$

where f_2 is the multiplicative function defined by

$$f_2(p) = \begin{cases} 1 & \text{if } p \nmid B_0,\\ -f_1(p) & \text{if } p|B_0.\end{cases} \tag{5.9}$$

In other words,

$$f_2(m) = \mu((m, B_0)) f_1((m, B_0)).$$

Therefore

$$\begin{aligned}y^\dagger_{r,\ell} &= \mu^2(r) \frac{f_1^\dagger(r)f(r)}{f^\dagger(r)f_1(r)} \sum_{\substack{m<R/r\\(m,r)=1}} \frac{y_{rm,\ell}}{f_1(m)} \frac{\mu(m)}{\phi(m)} f_2(m)\\ &= \mu^2(r) \frac{\mathfrak{S}(\mathcal{H})}{\ell!} \frac{f_1^\dagger(r)f(r)}{f^\dagger(r)f_1(r)} \sum_{\substack{m<R/r\\(m,r)=1}} \frac{\mu(m)f_2(m)(\log R/rm)^\ell}{f_1(m)\phi(m)}.\end{aligned} \tag{5.10}$$

The sum

$$\sum_{\substack{m=1\\(m,r)=1}}^\infty \frac{\mu(m)f_2(m)}{f_1(m)\phi(m)}$$

converges, and so one would expect that

$$y^\dagger_{r,\ell} \sim \mu^2(r) \frac{\mathfrak{S}(\mathcal{H})}{\ell!} (\log R/r)^\ell \frac{f_1^\dagger(r)f(r)}{f^\dagger(r)f_1(r)} \sum_{\substack{m=1\\(m,r)=1}}^\infty \frac{\mu(m)f_2(m)}{f_1(m)\phi(m)}$$

when $r < R$ and $(r, A_0) = 1$. From Lemma 18 below, we would then obtain

$$y^\dagger_{r,\ell} \sim \mu^2(r) \frac{\mathfrak{S}(\mathcal{H}^0)}{\ell!} (\log R/r)^\ell,$$

and we will ultimately prove this. This asymptotic relation should be compared to (1.20) and (4.17).

Lemma 18. *If r is squarefree and $(r, A_0) = 1$, then*

$$\frac{f_1^\dagger(r)f(r)}{f^\dagger(r)f_1(r)} \sum_{\substack{m=1\\(m,r)=1}}^{\infty} \frac{\mu(m)f_2(m)}{f_1(m)\phi(m)} = \frac{\mathfrak{S}(\mathcal{H}^0)}{\mathfrak{S}(\mathcal{H})}.$$

Proof. For r satisfying our hypotheses, it is convenient to define

$$F(r) = \frac{f_1^\dagger(r)f(r)}{f^\dagger(r)f_1(r)} \text{ and } G(r) = \sum_{\substack{m=1\\(m,r)=1}}^{\infty} \frac{\mu(m)f_2(m)}{f_1(m)\phi(m)}, \tag{5.11}$$

so that the left-hand side of the proposed result is $F(r)G(r)$. We begin by noting that

$$F(r) = \prod_{p|r} F(p) = \prod_{p|r} \frac{p(p-1-\nu_p^\dagger)}{(p-1)(p-\nu_p)}.$$

Moreover,

$$\begin{aligned} G(r) &= \prod_{p\nmid r} \left(1 - \frac{f_2(p)}{\phi(p)f_1(p)}\right) \\ &= \prod_{\substack{p\nmid B_0\\p\nmid r}} \frac{p(p-1-\nu_p)}{(p-1)(p-\nu_p)} \prod_{\substack{p|B_0\\p\nmid r}} \frac{p}{p-1} \\ &= \prod_{p\nmid r} \frac{p(p-1-\nu_p^\dagger)}{(p-1)(p-\nu_p)} = \prod_{p\nmid r} F(p). \end{aligned}$$

In the last line, we used the fact that $\nu_p^\dagger = \nu_p$ if $p \nmid B_0$ and $\nu_p^\dagger = \nu_p - 1$ if $p \mid B_0$. Combining the last two results yields

$$F(r)G(r) = \prod_p \frac{p(p-1-\nu_p^\dagger)}{(p-1)(p-\nu_p)} = \prod_p F(p). \tag{5.12}$$

On the other hand, if we replace $\mathcal{H}$ by $\mathcal{H}^0$ and k by $k+1$ in (4.4), then we obtain

$$\mathfrak{S}(\mathcal{H}^0) = \mathfrak{S}^*(\mathcal{H}^0) = \prod_p \left(1 - \frac{\nu_p^\dagger}{p-1}\right)\left(1 - \frac{1}{p}\right)^{-k}.$$

We combine this with the definition of $\mathfrak{S}(\mathcal{H})$ given in (1.1) to get

$$\frac{\mathfrak{S}(\mathcal{H}^0)}{\mathfrak{S}(\mathcal{H})} = \prod_p \frac{p(p-1-\nu_p^\dagger)}{(p-1)(p-\nu_p)} = \prod_p F(p). \tag{5.13}$$

The lemma follows by comparing this with (5.12). □

Lemma 19. *Suppose $\ell \geq 1$. If $r < R$ and $(r, A_0) = 1$, then*

$$y_{r,\ell}^\dagger = \mu^2(r)\frac{\mathfrak{S}(\mathcal{H}^0)}{\ell!}(\log R/r)^\ell + O\left(\mu^2(r)\beta(\mathcal{H}^0)\mathfrak{S}(\mathcal{H}^0)(\log 2R/r)^{\ell-1}\right). \tag{5.14}$$

Proof. From the definition of $y^{\dagger}_{r,\ell}$ in (5.8), the lemma is trivial if r is not squarefree. For the remainder of the proof, we assume that r is squarefree, $(r, A_0) = 1$, and $r < R$.

We start from the expression for $y^{\dagger}_{r,\ell}$ given in (5.10). For a given m in the inner sum, write $m = \delta n$, where $\delta | B_0$ and $(n, B_0) = 1$. Then $f_2(m) = \mu(\delta) f_1(\delta)$ and

$$\frac{\mu(m) f_2(m)}{f_1(m)\phi(m)} = \frac{\mu^2(\delta)\mu(n)}{\phi(\delta)\phi(n) f_1(n)}.$$

Therefore (5.10) may be transformed into

$$y^{\dagger}_{r,\ell} = \frac{\mathfrak{S}(\mathcal{H})F(r)}{\ell!} \sum_{\substack{\delta | B_0 \\ (\delta, r)=1}} \frac{\mu^2(\delta)}{\phi(\delta)} \sum_{\substack{n < R/r\delta \\ (n, rB_0)=1}} \frac{\mu(n)}{f_1(n)\phi(n)} (\log R/r\delta n)^{\ell}.$$

If we set

$$B_1 = \prod_{\substack{p | B_0 \\ p \nmid r}} p = \frac{B_0}{(B_0, r)}, \tag{5.15}$$

then the above equation for $y^{\dagger}_{r,\ell}$ may be written as

$$y^{\dagger}_{r,\ell} = \frac{\mathfrak{S}(\mathcal{H})F(r)}{\ell!} \sum_{\delta | B_1} \frac{\mu^2(\delta)}{\phi(\delta)} \sum_{\substack{n < R/r\delta \\ (n, rB_1)=1}} \frac{\mu(n)}{f_1(n)\phi(n)} (\log R/r\delta n)^{\ell}. \tag{5.16}$$

For future reference, note that $B_0 | rB_1$.

Now let

$$Y(x; d, \ell) = \sum_{\substack{n < x \\ (n,d)=1}} \frac{\mu(n)}{f_1(n)\phi(n)} (\log x/n)^{\ell}, \tag{5.17}$$

so that the innermost sum in (5.16) is $Y(R/r\delta; rB_1, \ell)$.

Now assume that $\ell \geq 1$. We begin our analysis of Y by writing

$$\begin{aligned} Y(x; d, \ell) &= \sum_{\substack{n < x \\ (n,d)=1}} \frac{\mu(n)}{f_1(n)\phi(n)} \int_n^x \ell(\log x/u)^{\ell-1} \frac{du}{u} \\ &= \int_1^x \frac{\ell(\log x/u)^{\ell-1}}{u} \sum_{\substack{n < u \\ (n,d)=1}} \frac{\mu(n)}{f_1(n)\phi(n)} du \\ &= Y_1(x; d, \ell) - Y_2(x; d, \ell), \end{aligned} \tag{5.18}$$

where

$$Y_1(x; d, \ell) = \int_1^x \frac{\ell(\log x/u)^{\ell-1}}{u} \sum_{\substack{n=1 \\ (n,d)=1}}^{\infty} \frac{\mu(n)}{f_1(n)\phi(n)} du \tag{5.19}$$

and

$$Y_2(x; d, \ell) = \int_1^x \frac{\ell(\log x/u)^{\ell-1}}{u} \sum_{\substack{n \geq u \\ (n,d)=1}} \frac{\mu(n)}{f_1(n)\phi(n)} du. \tag{5.20}$$

We see immediately that

$$Y_1(x;d,\ell) = (\log x)^\ell \prod_{p\nmid d}\left(1 - \frac{1}{f_1(p)\phi(p)}\right).$$

If we assume that $B_0|d$, then we may write

$$Y_1(x;d,\ell) = (\log x)^\ell \prod_{p\nmid d} F(p). \tag{5.21}$$

For $Y_2(x;d,\ell)$ we bound the sum inside the integrand as

$$\left|\sum_{\substack{n\ge u\\(n,d)=1}} \frac{\mu(n)}{f_1(n)\phi(n)}\right| \le \sum_{n\ge u}\frac{\mu^2(n)}{f_1(n)\phi(n)} = \int_u^\infty \left(\sum_{u\le n<v}\frac{\mu^2(n)n}{f_1(n)\phi(n)}\right)\frac{dv}{v^2}. \tag{5.22}$$

Now let

$$W^\dagger(v) = \sum_{n<v}\frac{\mu^2(n)n}{f_1(n)\phi(n)}. \tag{5.23}$$

This sum is very similar to the sum W^* defined in (4.23); in fact,

$$W^\dagger(v) = \sideset{}{^\flat}\sum_{n<v}\frac{\nu_n h(n)}{n},$$

where h was defined in (4.24). We have, similarly to (4.25),

$$W^\dagger(v) = \sideset{}{^\flat}\sum_{n<v}\frac{\nu_n}{n}\sum_{d|n}h_1(d) = \sideset{}{^\flat}\sum_{d<v}\frac{h_1(d)\nu_d}{d}\sideset{}{^\flat}\sum_{\substack{n<v/d\\(n,d)=1}}\frac{\nu_n}{n} \le \prod_{p<n}\left(1+\frac{h_1(p)\nu_p}{p}\right)\sideset{}{^\flat}\sum_{n<v}\frac{\nu_n}{n}.$$

The sum on the right-hand side is $\ll (\log 2v)^k$ by Lemma 8. The product on the right-hand side is $\ll 1$ because

$$\sum_{p<v}\log\left(1+\frac{h_1(p)\nu_p}{p}\right) \ll \sum_{p<v}\frac{\nu_p^2}{(p-1)(p-\nu_p)} \ll 1.$$

Therefore

$$W^\dagger(v) \ll (\log 2v)^k. \tag{5.24}$$

Now we use (5.24) in (5.22) to get

$$\left|\sum_{\substack{n\ge u\\(n,d)=1}} \frac{\mu(n)}{f_1(n)\phi(n)}\right| \ll \int_u^\infty \frac{(\log 2v)^k}{v^2}dv \ll \frac{(\log 2u)^k}{u}.$$

We use this in (5.20) to get

$$Y_2(x;d,\ell) \ll (\log 2x)^{\ell-1}\int_1^x (\log 2v)^k\frac{dv}{v^2} \ll (\log 2x)^{\ell-1}. \tag{5.25}$$

Combining this with (5.21) gives

$$Y(x;d,\ell) = (\log x)^\ell\prod_{p\nmid d}F(p) + O((\log 2x)^{\ell-1}) \tag{5.26}$$

when $B_0|d$.

Now we use (5.26) with $d = rB_1$ in (5.16) to obtain

$$y^{\dagger}_{r,\ell} = \frac{\mathfrak{S}(\mathcal{H})}{\ell!}\left(\prod_{p\nmid B_1} F(p)\right)\sum_{\substack{\delta|B_1\\ \delta<R/r}} \frac{\mu^2(\delta)}{\phi(\delta)}(\log R/r\delta)^{\ell} + O\left(\mathfrak{S}(\mathcal{H})F(r)\sum_{\substack{\delta|B_1\\ \delta<R/r}} \frac{\mu^2(\delta)}{\phi(\delta)}(\log 2R/r\delta)^{\ell-1}\right). \tag{5.27}$$

The error term in (5.27) is

$$\begin{aligned}
&\ll \mathfrak{S}(\mathcal{H})F(r)(\log 2R/r)^{\ell-1}\sum_{\delta|B_1}\frac{\mu^2(\delta)}{\phi(\delta)}\\
&\ll \mathfrak{S}(\mathcal{H})\left(\prod_{p|rB_1}F(p)\right)(\log 2R/r)^{\ell-1}\\
&\ll \mathfrak{S}(\mathcal{H}^0)(\log 2R/r)^{\ell-1}\left(\prod_{p\nmid rB_1}F(p)\right)^{-1}.
\end{aligned}$$

We have used (5.13) in the last line. Now when $p \nmid B_0$,

$$F(p)^{-1} = \left(1 - \frac{\nu_p}{(p-1)(p-\nu_p)}\right)^{-1} = 1 + O(1/p^2),$$

so

$$\left(\prod_{p\nmid rB_1} F(p)\right)^{-1} \ll 1.$$

Therefore the error term in (5.27) is

$$\ll \mathfrak{S}(\mathcal{H}^0)(\log 2R/r)^{\ell-1}. \tag{5.28}$$

Now we consider the main term in (5.27), which we write as

$$\frac{\mathfrak{S}(\mathcal{H})}{\ell!}\left(\prod_{p\nmid B_1}F(p)\right)\{M_1 - M_2 - M_3\}, \tag{5.29}$$

where

$$\begin{aligned}
M_1 &= (\log R/r)^{\ell}\sum_{\delta|B_1}\frac{\mu^2(\delta)}{\phi(\delta)},\\
M_2 &= (\log R/r)^{\ell}\sum_{\substack{\delta|B_1\\ \delta\geq R/r}}\frac{\mu^2(\delta)}{\phi(\delta)},\\
M_3 &= \sum_{\substack{\delta|B_1\\ \delta<R/r}}\frac{\mu^2(\delta)}{\phi(\delta)}\left\{(\log R/r)^{\ell} - (\log R/r\delta)^{\ell}\right\}.
\end{aligned}$$

For M_1, we note that

$$\sum_{\delta|B_1} \frac{\mu^2(\delta)}{\phi(\delta)} = \prod_{p|B_1} \frac{p}{p-1} = \prod_{p|B_1} F(p).$$

Therefore

$$\frac{\mathfrak{S}(\mathcal{H})}{\ell!} M_1 \prod_{p\nmid B_1} F(p) = \frac{\mathfrak{S}(\mathcal{H})}{\ell!} (\log R/r)^\ell \prod_p F(p) = \frac{\mathfrak{S}(\mathcal{H}^0)}{\ell!} (\log R/r)^\ell, \tag{5.30}$$

by (5.13).

For M_2, we note that

$$\sum_{\substack{\delta|B_1 \\ \delta\geq R/r}} \frac{\mu^2(\delta)}{\phi(\delta)} \ll \sum_{\delta|B_1} \frac{\mu^2(\delta)}{\phi(\delta)} \frac{\log\delta}{(\log 2R/r)}$$

and

$$\begin{aligned}\sum_{\delta|B_1} \frac{\mu^2(\delta)}{\phi(\delta)} \log\delta &= \sum_{\delta|B_1} \frac{\mu^2(\delta)}{\phi(\delta)} \sum_{p|\delta} \log p = \sum_{p|B_1} \frac{\log p}{p-1} \sum_{\delta|B_1/p} \frac{\mu^2(\delta)}{\phi(\delta)} \\ &= \sum_{p|B_1} \frac{\log p}{p-1} \frac{B_1/p}{\phi(B_1/p)} = \frac{B_1}{\phi(B_1)} \sum_{p|B_1} \frac{\log p}{p} \\ &= F(B_1) \sum_{p|B_1} \frac{\log p}{p}.\end{aligned}$$

Now if $p|B_1$, then $p|B_0$ and $\nu_p(\mathcal{H}^0) \leq k$. Therefore

$$\sum_{p|B_1} \frac{\log p}{p} \leq \sum_p \frac{(k+1-\nu_p(\mathcal{H}^0))\log p}{p} = \beta(\mathcal{H}^0).$$

Consequently,

$$\sum_{\delta|B_1} \frac{\mu^2(\delta)}{\phi(\delta)} \log\delta \ll F(B_1)\beta(\mathcal{H}^0), \tag{5.31}$$

and so

$$\begin{aligned}\frac{\mathfrak{S}(\mathcal{H})}{\ell!} M_2 \prod_{p\nmid B_1} F(p) &\ll (\log 2R/r)^{\ell-1} \mathfrak{S}(\mathcal{H})\beta(\mathcal{H}^0)F(B_1) \prod_{p\nmid B_1} F(p) \\ &\ll \mathfrak{S}(\mathcal{H}^0)\beta(\mathcal{H}^0)(\log 2R/r)^{\ell-1}.\end{aligned} \tag{5.32}$$

For M_3, we note that when $\delta \leq R/r$,

$$\begin{aligned}(\log R/r)^\ell &- (\log R/r\delta)^\ell \\ &= (\log\delta)\left\{(\log R/r)^{\ell-1} + (\log R/r\delta)(\log R/r)^{\ell-2} + \ldots + (\log R/r\delta)^{\ell-1}\right\} \\ &\ll (\log\delta)(\log R/r)^{\ell-1}.\end{aligned}$$

Thus

$$M_3 \ll (\log 2R/r)^{\ell-1} \sum_{\delta|B_1} \frac{\mu^2(\delta)}{\phi(\delta)} \log\delta \ll (\log 2R/r)^{\ell-1} F(B_1)\beta(\mathcal{H}^0)$$

by (5.31), and so

$$\frac{\mathfrak{S}(\mathcal{H})}{\ell!} M_3 \prod_{p\nmid B_1} F(p) \ll (\log 2R/r)^{\ell-1} \mathfrak{S}(\mathcal{H}^0)\beta(\mathcal{H}^0). \tag{5.33}$$

Combining the estimates (5.28), (5.30), (5.32), and (5.33) gives the proof of Lemma 19. □

In reference to the above lemma, we remark that with a bit more work we could give an estimate valid for $y_{r,0}$ with a somewhat weaker error term. However, we omit this because it is not necessary for the proof of Theorem 1.

We can now complete the estimate of S. From (5.7) and Lemma 19, we see that

$$S = V^\dagger + O\left(\mathfrak{S}(\mathcal{H}^0)^2 \beta(\mathcal{H}^0)(\log R)^{\ell_1+\ell_2-1} W^\dagger\right), \tag{5.34}$$

where

$$V^\dagger = \frac{\mathfrak{S}(\mathcal{H}^0)^2}{\ell_1!\ell_2!} \sideset{}{'}\sum_{r<R} \frac{\mu^2(r)}{f_1^\dagger(r)} (\log R/r)^{\ell_1+\ell_2}$$

and

$$W^\dagger = \sideset{}{'}\sum_{r<R} \frac{\mu^2(r)}{f_1^\dagger(r)}.$$

Now $V^\dagger$ is the same as the sum V in (4.19) except that $\mathcal{H}$ has been replaced by $\mathcal{H}^0$, k has been replaced by $k+1$, and ℓ_1, ℓ_2 have been replaced by $\ell_1 - 1, \ell_2 - 1$ respectively. From (4.21), we see that

$$V^\dagger = \binom{\ell_1+\ell_2}{\ell_1} \mathfrak{S}(\mathcal{H}^0) \frac{(\log R)^{k+\ell_1+\ell_2}}{(k+\ell_1+\ell_2)!} + O(\beta(\mathcal{H}^0)\mathfrak{S}(\mathcal{H}^0)^2 (\log R)^{k+\ell_1+\ell_2-1}). \tag{5.35}$$

For $W^\dagger$, we use Lemma 17 with $a=0$, $d=1$, f^* replaced by $f^\dagger$, and k replaced by $k+1$ to get

$$W^\dagger \ll (\log R)^k. \tag{5.36}$$

Now we combine (5.34), (5.35), and (5.36) to get

$$S = \binom{\ell_1+\ell_2}{\ell_1} \mathfrak{S}(\mathcal{H}^0) \frac{(\log R)^{k+\ell_1+\ell_2}}{(k+\ell_1+\ell_2)!} + O(\beta(\mathcal{H}^0)\mathfrak{S}(\mathcal{H}^0)^2 (\log R)^{k+\ell_1+\ell_2-1}). \tag{5.37}$$

Equation (1.27) now follows by combining this with (5.3).

6. Proof of Theorem 7

We may again assume, without loss of generality, that $h_0 = 0$. Accordingly, we assume throughout this section that $0 \in \mathcal{H}$.

Let $\mathcal{L}$ denote the sum on the left-hand side in the statement of Theorem 7. Then

$$\mathcal{L} = \sum_{d,e} \lambda_{d,\ell_1} \lambda_{e,\ell_2} \sum_{a \in \Omega_{[d,e]}(\mathcal{H})} \sum_{\substack{N<n\le 2N \\ n \equiv a \pmod{[d,e]}}} \varpi * \varpi(n). \tag{6.1}$$

In this sum, we have $d, e < R < \sqrt{N}$, so $[d,e]$ has no prime divisors exceeding $\sqrt{N}$. On the other hand, if $N < n \le 2N$ and $\varpi * \varpi(n) > 0$, then n is a product of two primes, at least one of which must exceed $\sqrt{N}$. Therefore, the inner sum in (6.1) will be 0 unless $(a,[d,e]) = 1$ or $(a,[d,e]) = p$ for some prime $p < R$.

We write

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2,$$

where $\mathcal{L}_1$ is the sum in (6.1) with the extra condition that $(a,[d,e])=1$, and $\mathcal{L}_2$ is the sum in (6.1) with the extra condition that $(a,[d,e])=p$ for some prime p.

Before analyzing $\mathcal{L}_2$, it is useful to note that when r is squarefree and $(a,r)=p$,

$$\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod r}} \varpi*\varpi(n) = 2\log p \sum_{\substack{\frac{N}{p}<m\le\frac{2N}{p}\\ m\equiv\frac{a}{p} \pmod{\frac{r}{p}}}} \varpi(m)$$

$$= \frac{2N}{\phi(r)}\frac{(\log p)\phi(p)}{p} + O(E^*(N/p,r/p)\log p).$$

When r is squarefree and p is a prime dividing r, we define

$$\Omega^*_{r,p}(\mathcal{H}) = \{a\in\mathbb{Z}_r : (a,r)=p \text{ and } P(a;\mathcal{H})\equiv 0 \pmod r\}. \tag{6.2}$$

Let $\nu^*_{r,p}=\nu^*_{r,p}(\mathcal{H})$ be the cardinality of $\Omega^*_{r,p}(\mathcal{H})$.

We take $d_1=p, d_2=r/p$ in (3.2), and we see that $\Omega^*_{r,p}(\mathcal{H})$ is the image of the set $\{0\}\times\Omega^*_{r/p}$ under the isomorphism ξ of (3.2). Therefore

$$\nu^*_{r,p}=\nu^*_{r/p}.$$

Using the above information, we find that

$$\mathcal{L}_2 = \sum_{d,e}\lambda_{d,\ell_1}\lambda_{e,\ell_2}\sum_{p|[d,e]}\sum_{a\in\Omega^*_{[d,e],p}}\sum_{\substack{N<n\le 2N\\ n\equiv a \pmod{[d,e]}}}\varpi*\varpi(n) \tag{6.3}$$

$$= 2N\sum_{d,e}\frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}}{\phi([d,e])}\sum_{p|[d,e]}\nu^*_{[d,e]/p}\frac{(\log p)\phi(p)}{p} + O\left((\log N)^{4k}\mathcal{E}_2\right),$$

where

$$\mathcal{E}_2 = \sum_{r<R^2}^{\flat} 3^{\omega(r)}\sum_{\substack{p|r\\ p<R}}\nu^*_{r/p}E^*(N/p,r/p)(\log p).$$

Upon writing $r=pm$ and changing the order of summation, we find that

$$\mathcal{E}_2 \le 3\sum_{p<R}\log p\sum_{m<R^2/p}^{\flat}3^{\omega(m)}\nu^*_m E^*(N/p,m).$$

By Lemma 9, the inner sum is $\ll (N/p)(\log N/p)^{-4k-2} \ll (N/p)(\log N)^{-4k-2}$. Summing over p, we get

$$\mathcal{E}_2 \ll N(\log N)^{-4k-1}.$$

Therefore

$$\mathcal{L}_2 = 2N\sum_{d,e}\frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}}{\phi([d,e])}\sum_{p|[d,e]}\frac{\nu^*_{[d,e]/p}(\log p)\phi(p)}{p} + O(N/\log N). \tag{6.4}$$

Now we turn our attention to $\mathcal{L}_1$. From our definitions and (1.10), we have

$$\mathcal{L}_1 = \sum_{d,e}\lambda_{d,\ell_1}\lambda_{e,\ell_2}\sum_{a\in\Omega^*_{[d,e]}}\sum_{\substack{N\le n<2N\\ n\equiv a \pmod{[d,e]}}}\varpi*\varpi(n) \tag{6.5}$$

$$= N\sum_{d,e}\frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}\nu^*_{[d,e]}}{\phi([d,e])}\left(\log N + C_0 - 2\sum_{p|[d,e]}\frac{\log p}{p}\right) + O(\mathcal{E}_1),$$

where

$$\mathcal{E}_1 = (\log R)^{4k} \sum_{r<R^2}^{\flat} 3^{\omega(r)} \nu_r^* E_2^*(N,r).$$

By Lemma 9, $\mathcal{E}_1 \ll N/\log N$.

Combining our estimates for $\mathcal{L}_1$ and $\mathcal{L}_2$, we find that

$$\mathcal{L} = N(\log N + C_0)S_1 - 2NS_2 + 2NS_3 + O(N/\log N), \tag{6.6}$$

where

$$S_1 = \sum_{d,e} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}\nu^*_{[d,e]}}{\phi([d,e])},$$

$$S_2 = \sum_{d,e} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}\nu^*_{[d,e]}}{\phi([d,e])} \sum_{p\mid[d,e]} \frac{\log p}{p},$$

$$S_3 = \sum_{d,e} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}}{\phi([d,e])} \sum_{p\mid[d,e]} \nu^*_{[d,e]/p} \frac{(\log p)\phi(p)}{p}.$$

We have already encountered the sum S_1; it is the same as the sum S defined in (4.6). From (4.27), we see that

$$S_1 = \binom{\ell_1+\ell_2+2}{\ell_1+1} \mathfrak{S}(\mathcal{H}) \frac{(\log R)^{k+\ell_1+\ell_2+1}}{(k+\ell_1+\ell_2+1)!} + O\left(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2 (\log R)^{k+\ell_1+\ell_2}\right). \tag{6.7}$$

Of the remaining two sums, S_3 is more important, so we concentrate on it first. We begin by interchanging the order of summation in S_3; this yields

$$S_3 = \sum_p \frac{\log p}{p} U(p), \tag{6.8}$$

where

$$U(p) = \sum_{\substack{d,e \\ p\mid[d,e]}} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}\nu^*_{[d,e]/p}}{\phi([d,e]/p)}. \tag{6.9}$$

We decompose $U(p)$ as

$$U(p) = U_1(p) + U_2(p) + U_3(p), \tag{6.10}$$

where

$$U_1(p,\ell_1,\ell_2) = \sum_{\substack{d,e \\ p\mid d,\, p\nmid e}} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}\nu^*_{[d,e]/p}}{\phi([d,e]/p)},$$

$$U_2(p,\ell_1,\ell_2) = \sum_{\substack{d,e \\ p\nmid d,\, p\mid e}} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}\nu^*_{[d,e]/p}}{\phi([d,e]/p)},$$

$$U_3(p,\ell_1,\ell_2) = \sum_{\substack{d,e \\ p\mid d,\, p\mid e}} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}\nu^*_{[d,e]/p}}{\phi([d,e]/p)}.$$

Going back to (6.8), we will write

$$S_3 = S_{3,1} + S_{3,2} + S_{3,3}, \tag{6.11}$$

where

$$S_{3,i} = \sum_p \frac{\log p}{p} U_i(p, \ell_1, \ell_2).$$

We will ultimately see that each $S_{3,i}$ corresponds to one of the terms in the quantity $T(k, \ell_1, \ell_2)$ defined in the statement of Theorem 7. More precisely, we will show that when $1 \le i \le 3$,

$$S_{3,i} = T_i \mathfrak{S}(\mathcal{H}) \frac{(\log R)^{k+\ell_1+\ell_2+2}}{(k+\ell_1+\ell_2+2)!} \left\{1 + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})/\log R)\right\},$$

where

$$T_1 = -\binom{\ell_1+\ell_2+3}{\ell_2+1}, \quad T_2 = -\binom{\ell_1+\ell_2+3}{\ell_1+1}, \quad T_3 = \binom{\ell_1+\ell_2+2}{\ell_1+1}.$$

We note that $U_2(p, \ell_1, \ell_2)$ is the same as $U_1(p, \ell_1, \ell_2)$ except that the roles of ℓ_1, ℓ_2 have been reversed; i.e., $U_2(p, \ell_1, \ell_2) = U_1(p, \ell_2, \ell_1)$. Accordingly, we will concentrate on evaluating $U_1(p, \ell_1, \ell_2)$ and $U_3(p, \ell_1, \ell_2)$. For brevity, we will usually write these as $U_1(p)$ and $U_3(p)$.

The evaluations of $U_1(p)$ and $U_3(p)$ will require use of the quantity $y^*_{r,\ell}$ defined in (4.11), as well as a new quantity $z^*_{r,p,\ell}$. The latter is defined as

$$z^*_{r,p,\ell} = \begin{cases} \mu(pr) f_1^*(r) \sum_d{}' \dfrac{\lambda_{drp,\ell}}{f^*(dr)} & \text{if } r < R/p \text{ and } (r, A) = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6.12}$$

As in Section 4, we use $\sum'$ to denote that the sum is over values of the indices that are relatively prime to A. Note that $z^*_{r,p,\ell} = 0$ if $(p, r) \neq 1$. On the other hand, the condition $p|A$ (i.e., $\nu_p^* = 0$) does not imply that $z^*_{r,p,\ell} = 0$. However, one can easily show that if $p \nmid A$, then

$$z^*_{r,p,\ell} = \left(\frac{p-1}{p-\nu_p}\right) y^*_{rp,\ell}. \tag{6.13}$$

We now give three lemmas that we will use for the evaluation of S_1 and S_3.

Lemma 20. *If $p < R$, then*

$$U_1(p) = -\sum_{\substack{r \\ (r,p)=1}}{}' \frac{z^*_{r,p,\ell_1} y^*_{r,\ell_2}}{f_1^*(r)} - \frac{\nu_p^*}{p-1} \sum_{\substack{r \\ (r,p)=1}}{}' \frac{z^*_{r,p,\ell_1} z^*_{r,p,\ell_2}}{f_1^*(r)} \quad \textit{and} \tag{6.14}$$

$$U_3(p) = \sum_{\substack{r \\ (r,p)=1}}{}' \frac{z^*_{r,p,\ell_1} z^*_{r,p,\ell_2}}{f_1^*(r)}. \tag{6.15}$$

Proof. The sum $U_1(p)$ may be written as

$$U_1(p) = \sideset{}{'}\sum_{\substack{d,e\\ p\nmid e}} \frac{\lambda_{dp,\ell_1}\lambda_{e,\ell_2}}{\phi([d,e])}\nu^*_{[d,e]} = \sideset{}{'}\sum_{\substack{d,e\\ p\nmid e}} \frac{\lambda_{dp,\ell_1}\lambda_{e,\ell_2}}{f^*([d,e])} = \sideset{}{'}\sum_{\substack{d,e\\ p\nmid e}} \frac{\lambda_{dp,\ell_1}\lambda_{e,\ell_2}}{f^*(d)f^*(e)} \sum_{\substack{r|d\\ r|e}} f_1^*(r) \tag{6.16}$$

$$= \sideset{}{'}\sum_{\substack{r\\ (r,p)=1}} f_1^*(r) \left(\sideset{}{'}\sum_{d} \frac{\lambda_{drp,\ell_1}}{f^*(dr)} \right) \left(\sideset{}{'}\sum_{\substack{e\\ p\nmid e}} \frac{\lambda_{er,\ell_2}}{f^*(er)} \right).$$

In the last expression, the first sum in parentheses is $\mu(pr)z^*_{r,p,\ell_1}/f_1^*(r)$. The innermost sum is

$$\sideset{}{'}\sum_{e} \frac{\lambda_{er,\ell_2}}{f^*(er)} - \sideset{}{'}\sum_{\substack{e\\ p|e}} \frac{\lambda_{er,\ell_2}}{f^*(er)} = \frac{\mu(r)y^*_{r,\ell_2}}{f_1^*(r)} - \sideset{}{'}\sum_{\substack{e\\ p|e}} \frac{\lambda_{er,\ell_2}}{f^*(er)}.$$

We claim that

$$\sideset{}{'}\sum_{\substack{e\\ p|e}} \frac{\lambda_{er,\ell_2}}{f^*(er)} = \frac{\nu_p^*\mu(pr)z^*_{r,p,\ell_2}}{(p-1)f_1^*(r)}. \tag{6.17}$$

If $\nu_p^* = 0$, then both sides of (6.17) are 0. If $\nu_p^* \neq 0$, then

$$\sideset{}{'}\sum_{\substack{e\\ p|e}} \frac{\lambda_{er,\ell_2}}{f^*(er)} = \sideset{}{'}\sum_{e} \frac{\lambda_{epr,\ell_2}}{f^*(epr)} = \frac{\mu(pr)z^*_{r,p,\ell_2}}{f_1^*(r)f^*(p)},$$

and (6.17) follows again.

Going back to (6.16), we find that

$$U_1(p) = \sideset{}{'}\sum_{\substack{r\\ (r,p)=1}} f_1^*(r) \left(\frac{\mu(rp)z^*_{r,p,\ell_1}}{f_1^*(r)} \right) \left(\frac{\mu(r)y^*_{r,\ell_2}}{f_1^*(r)} - \frac{\mu(rp)z^*_{r,p,\ell_2}\nu_p^*}{f_1^*(r)(p-1)} \right),$$

and (6.14) follows.

For $U_3(p)$, observe that

$$U_3(p) = \sideset{}{'}\sum_{d,e} \frac{\lambda_{dp,\ell_1}\lambda_{ep,\ell_2}}{f^*([d,e])} = \sideset{}{'}\sum_{d,e} \frac{\lambda_{dp,\ell_1}\lambda_{ep,\ell_2}}{f^*(d)f^*(e)} \sum_{\substack{r|d\\ r|e}} f_1^*(r)$$

$$= \sideset{}{'}\sum_{\substack{r\\ (r,p)=1}} f_1^*(r) \left(\sideset{}{'}\sum_{d} \frac{\lambda_{drp,\ell_1}}{f^*(dr)} \right) \left(\sideset{}{'}\sum_{e} \frac{\lambda_{erp,\ell_1}}{f^*(er)} \right)$$

$$= \sideset{}{'}\sum_{\substack{r\\ (r,p)=1}} \frac{z^*_{r,p,\ell_1}z^*_{r,p,\ell_2}}{f_1^*(r)},$$

and this yields (6.15). □

Lemma 21. *If $r < R/p$ and $(r, A) = 1$, then*

$$z^*_{r,p,\ell} = \mu^2(rp)\frac{\mathfrak{S}(\mathcal{H})}{(\ell+1)!}\left(\frac{p-1}{p-\nu_p}\right)(\log R/rp)^{\ell+1} + O\left(\mu^2(rp)\frac{\rho(rp)rp}{\phi(rp)}\mathfrak{S}(\mathcal{H})(\log 2R/rp)^{\ell}\right). \tag{6.18}$$

We remark that the error term could be simplified; it is obvious that

$$\frac{\rho(rp)rp}{\phi(rp)} \ll \frac{\rho(r)r}{\phi(r)}.$$

However, we prefer to write it as above to emphasize the connection between $y^*_{r,\ell}$ and $z^*_{r,p,\ell}$. In fact, this lemma follows immediately from (6.13) and (4.17) when $\nu^*_p \neq 0$. However, the following argument works whether or not $\nu^*_p = 0$.

Proof. The result is trivial if rp is not squarefree, because both sides of (6.18) are 0 in this case. For the rest of this proof, we assume that rp is squarefree. Note that this assumption implies that $(r,p)=1$.

We start by observing that

$$\begin{aligned}\frac{\mu(rp)z^*_{r,p,\ell}}{f^*_1(r)} &= \sum_d{}' \frac{\lambda_{drp,\ell}}{f^*(dr)} = \sum_d{}' \frac{\mu(drp)}{f^*(dr)} f(drp) \sum_t \frac{y_{drpt,\ell}}{f_1(drpt)} \\ &= \frac{\mu(rp)f(rp)}{f^*(r)f_1(rp)} \sum_d{}' \frac{\mu(d)f(d)}{f^*(d)} \sum_t \frac{y_{rpdt,\ell}}{f_1(dt)} \\ &= \frac{\mu(rp)f(rp)}{f^*(r)f_1(rp)} \sum_{\substack{m \\ (m,rp)=1}} \frac{y_{rpm,\ell}}{f_1(m)} \sum_{d|m}{}' \frac{\mu(d)f(d)}{f^*(d)} \\ &= \frac{\mu(rp)f(rp)}{f^*(r)f_1(rp)} \sum_{\substack{m \\ (m,rp)=1}} \frac{y_{rpm,\ell}}{\phi(m)}.\end{aligned}$$

In the last line, we have used the relation (4.12). If we also use (4.13), we find that

$$\begin{aligned}z^*_{r,p,\ell} &= \frac{f(r)f^*_1(r)f(p)}{f_1(r)f^*(r)f_1(p)} \sum_{\substack{m<R/rp \\ (m,rp)=1}} \frac{y_{rpm,\ell}}{\phi(m)} \\ &= \frac{\mathfrak{S}(\mathcal{H})}{\ell!} \frac{rp}{\phi(rp)} \frac{(p-1)}{(p-\nu_p)} \sum_{\substack{m<R/rp \\ (m,rp)=1}} \frac{\mu^2(m)}{\phi(m)} (\log R/rpm)^\ell.\end{aligned} \tag{6.19}$$

We then use (4.16) to complete the proof. □

Lemma 22. *If a,b are non-negative integers, then*

$$\sum_{p<R} \frac{(\log p)^{a+1}(\log R/p)^b}{p} = \frac{a!b!}{(a+b+1)!}(\log R)^{a+b+1} + O_{a,b}((\log R)^{a+b}) \text{ and}$$

$$\sum_{p<R} \frac{(\log p)^{a+1}(\log R/p)^b}{p^2} \ll_a (\log R)^b.$$

Proof. Let $E(u)$ be defined by the relation

$$\sum_{p\le u} \frac{\log p}{p} = \log u + E(u).$$

It is well-known that $E(u) \ll 1$. The first sum in the lemma is

$$\sum_{p<R} \frac{(\log p)^{a+1}(\log R/p)^b}{p} = \int_1^R (\log u)^a (\log R/u)^b \frac{du}{u} + \int_1^R (\log u)^a (\log R/u)^b dE(u).$$

By Lemma 10, the first integral is

$$\frac{a!b!}{(a+b+1)!}(\log R)^{a+b+1}.$$

Using integration by parts, we see that the second integral is

$$\int_1^R E(u)\frac{d}{du}\left\{(\log u)^a(\log R/u)^b\right\} du \ll_{a,b} (\log R)^{a+b}.$$

This proves the first statement. The second statement is easier; we simply note that

$$\sum_{p<R} \frac{(\log p)^{a+1}(\log R/p)^b}{p^2} \ll (\log R)^b \sum_p \frac{(\log p)^{a+1}}{p^2} \ll_a (\log R)^b.$$

□

Evaluation of $S_{3,3}$. From Lemmas 20 and 21, we see that

(6.20)
$$U_3(p) = \frac{\mathfrak{S}(\mathcal{H})^2}{(\ell_1+1)!(\ell_2+1)!}\left(\frac{p-1}{p-\nu_p}\right)^2 V_3(p) + O\left(\mathfrak{S}(\mathcal{H})^2(\log R)^{\ell_1+\ell_2+1}W(p)\right),$$

where

(6.21)
$$V_3(p) = \sum_{\substack{r<R/p \\ (r,p)=1}}{}' \frac{\mu^2(r)}{f_1^*(r)}(\log R/rp)^{\ell_1+\ell_2+2} \text{ and}$$

(6.22)
$$W(p) = \sum_{r<R/p}{}' \frac{\mu^2(r)\rho(r)r}{f_1^*(r)\phi(r)}.$$

$W(p)$ is majorized by the sum W defined in (4.20), and, using (4.26), we see that

(6.23)
$$W(p) \ll (\log R)^{k-1}.$$

From Lemma 17, we see that

(6.24)
$$V_3(p) = \frac{(\ell_1+\ell_2+2)!}{\mathfrak{S}(\mathcal{H})}\left(\frac{p-\nu_p}{p-1}\right)\frac{(\log R/p)^{k+\ell_1+\ell_2+1}}{(k+\ell_1+\ell_2+1)!} + O\left(\beta(\mathcal{H})(\log R)^{k+\ell_1+\ell_2}\right).$$

We combine the above estimates for $V_3(p)$ and $W(p)$ with (6.20) to get

(6.25)
$$U_3(p) = \binom{\ell_1+\ell_2+2}{\ell_1+1}\mathfrak{S}(\mathcal{H})\left(\frac{p-1}{p-\nu_p}\right)\frac{(\log R/p)^{k+\ell_1+\ell_2+1}}{(k+\ell_1+\ell_2+1)!} + O\left(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2}\right).$$

We can now finish our estimation of $S_{3,3}$. From our definition and from (6.25), we get

$$\begin{aligned} S_{3,3} &= \sum_{p<R} \frac{\log p}{p} U_3(p) \\ &= \binom{\ell_1+\ell_2+2}{\ell_1+1} \frac{\mathfrak{S}(\mathcal{H})}{(k+\ell_1+\ell_2+1)!} \sum_{p<R} \left(\frac{\log p}{p}\right)\left(\frac{p-1}{p-\nu_p}\right)(\log R/p)^{k+\ell_1+\ell_2+1} \\ &\qquad + O\left(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2} \sum_{p<R} \frac{\log p}{p}\right). \end{aligned}$$

Now $(p-1)/(p-\nu_p) = 1 + O(1/p)$, so we may use Lemma 22 to get

$$\begin{aligned} S_{3,3} = &\binom{\ell_1+\ell_2+2}{\ell_1+1} \mathfrak{S}(\mathcal{H}) \frac{(\log R)^{k+\ell_1+\ell_2+2}}{(k+\ell_1+\ell_2+2)!} \\ &+ O\left(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2+1}\right). \end{aligned} \tag{6.26}$$

Evaluation of $S_{3,1}$. The evaluation of $S_{3,1}$ proceeds similarly to the evaluation of $S_{3,3}$, but it is somewhat more involved. We start by defining

$$U_4(p) = \sideset{}{'}\sum_{\substack{r \\ (r,p)=1}} \frac{y^*_{r,\ell_2} z^*_{r,p,\ell_1}}{f_1^*(r)} \tag{6.27}$$

and

$$S_4 = \sum_{p<R} \frac{\log p}{p} U_4(p). \tag{6.28}$$

Then (6.14) may be rewritten as

$$U_1(p) = -U_4(p) - \frac{\nu_p^*}{p-1} U_3(p), \tag{6.29}$$

and we may also write

$$S_{3,1} = -S_4 - \sum_{p<R} \frac{(\log p)\nu_p^*}{p(p-1)} U_3(p). \tag{6.30}$$

From (6.25), we see that

$$\begin{aligned} \sum_{p<R} \frac{(\log p)\nu_p^*}{p(p-1)} U_3(p) &\ll \beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2+1} \sum_p \frac{(\log p)\nu_p^*}{p^2} \\ &\ll \beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2+1}. \end{aligned} \tag{6.31}$$

Now we concentrate on $U_4(p)$ and S_4. From (4.17) and Lemma 21, we see that

$$U_4(p) = \frac{\mathfrak{S}(\mathcal{H})^2}{(\ell_1+1)!(\ell_2+1)!}\left(\frac{p-1}{p-\nu_p}\right) V_4(p) + O\left(\mathfrak{S}(\mathcal{H})^2(\log R)^{\ell_1+\ell_2+1} W(p)\right), \tag{6.32}$$

where $W(p)$ was defined in (6.22) and

$$V_4(p) = \sideset{}{'}\sum_{\substack{r<R/p \\ (r,p)=1}} \frac{\mu^2(r)}{f_1^*(r)} (\log R/r)^{\ell_2+1} (\log R/rp)^{\ell_1+1}. \tag{6.33}$$

We write $\log R/r = \log p + \log R/rp$ and use the binomial theorem to get

$$V_4(p) = \sum_{j=0}^{\ell_2+1} \binom{\ell_2+1}{j} (\log p)^j \sideset{}{'}\sum_{\substack{r<R/p\\(r,p)=1}} \frac{\mu^2(r)}{f_1^*(r)} (\log R/rp)^{\ell_1+\ell_2+2-j}.$$

We apply Lemma 17 to the inner sum, and we get

(6.34)

$$V_4(p) = \frac{1}{\mathfrak{S}(\mathcal{H})} \left(\frac{p-\nu_p}{p-1}\right) \sum_{j=0}^{\ell_2+1} \binom{\ell_2+1}{j} \frac{(\ell_1+\ell_2+2-j)!}{(k+\ell_1+\ell_2+1-j)!} (\log p)^j (\log R/p)^{k+\ell_1+\ell_2+1-j} + O(\beta(\mathcal{H})(\log 2R)^{k+\ell_1+\ell_2}).$$

Using this together with (6.32) and (6.23) gives

$$U_4(p) = \frac{\mathfrak{S}(\mathcal{H})}{(\ell_1+1)!(\ell_2+1)!} U_5(p) + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2 (\log R)^{k+\ell_1+\ell_2}), \tag{6.35}$$

where

$$U_5(p) = \sum_{j=0}^{\ell_2+1} \binom{\ell_2+1}{j} \frac{(\ell_1+\ell_2+2-j)!}{(k+\ell_1+\ell_2+1-j)!} (\log p)^j (\log R/p)^{k+\ell_1+\ell_2+1-j}. \tag{6.36}$$

For future reference, we note the crude estimate

$$U_1(p) \ll \beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2 (\log R)^{k+\ell_1+\ell_2+1} \tag{6.37}$$

that is implicit in the combination of (6.29), (6.35), (6.36), and (6.25).

Using (6.28) and (6.35), we see that

$$S_4 = \frac{\mathfrak{S}(\mathcal{H})}{(\ell_1+1)!(\ell_2+1)!} \sum_{p<R} \frac{\log p}{p} U_5(p) + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2 (\log R)^{k+\ell_1+\ell_2+1}). \tag{6.38}$$

We apply Lemma 22 to get

$$\sum_{p<R} \left(\frac{\log p}{p}\right) (\log p)^j (\log R/p)^{k+\ell_1+\ell_2+1-j} = \frac{j!(k+\ell_1+\ell_2+1-j)!}{(k+\ell_1+\ell_2+2)!} (\log R)^{k+\ell_1+\ell_2+2} + O((\log R)^{k+\ell_1+\ell_1+1}).$$

Using this in (6.38) gives

$$S_4 = \mathfrak{S}(\mathcal{H}) \frac{(\log R)^{k+\ell_1+\ell_2+2}}{(k+\ell_1+\ell_2+2)!} \sum_{j=0}^{\ell_2+1} \binom{\ell_1+\ell_2+2-j}{\ell_2+1-j} + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2 (\log R)^{k+\ell_1+\ell_2+1}). \tag{6.39}$$

To treat the sum of binomial coefficients in the above, we make a change of variables $j = \ell_2 + 1 - i$. The sum then becomes

$$\sum_{i=0}^{\ell_2+1} \binom{\ell_1+1+i}{i} = \sum_{i=0}^{\ell_2+1} \left\{ \binom{\ell_1+2+i}{i} - \binom{\ell_1+1+i}{i-1} \right\}, \tag{6.40}$$

provided we make the usual convention that

$$\binom{\ell_1+1}{-1} = 0.$$

The sum on the right-hand side of (6.40) is telescoping, so

$$\sum_{i=0}^{\ell_2+1} \binom{\ell_1+1+i}{i} = \binom{\ell_1+\ell_2+3}{\ell_2+1}.$$

Putting this information into (6.39) gives our final estimate for S_4; i.e.,

$$S_4 = \binom{\ell_1+\ell_2+3}{\ell_2+1}\mathfrak{S}(\mathcal{H})\frac{(\log R)^{k+\ell_1+\ell_2+2}}{(k+\ell_1+\ell_2+2)!} + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2+1}). \tag{6.41}$$

From this, together with (6.30) and (6.31), we get

$$S_{3,1} = -\binom{\ell_1+\ell_2+3}{\ell_2+1}\mathfrak{S}(\mathcal{H})\frac{(\log R)^{k+\ell_1+\ell_2+2}}{(k+\ell_1+\ell_2+2)!} + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2+1}). \tag{6.42}$$

As we noted earlier, $S_{3,2}$ is the same as $S_{3,1}$ with the roles of ℓ_1 and ℓ_2 reversed. Therefore

$$S_{3,2} = -\binom{\ell_1+\ell_2+3}{\ell_1+1}\mathfrak{S}(\mathcal{H})\frac{(\log R)^{k+\ell_1+\ell_2+2}}{(k+\ell_1+\ell_2+2)!} + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2+1}). \tag{6.43}$$

Combining (6.42), (6.43), and (6.26) gives

$$S_3 = T(k,\ell_1,\ell_2)\mathfrak{S}(\mathcal{H})\frac{(\log R)^{k+\ell_1+\ell_2+2}}{(k+\ell_1+\ell_2+2)!} + O(\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})^2(\log R)^{k+\ell_1+\ell_2+1}), \tag{6.44}$$

where $T(k,\ell_1,\ell_2)$ is as defined in Theorem 7.

Finally, we will quickly dispatch S_2. We rewrite this sum as

$$S_2 = \sideset{}{'}\sum_{d,e} \frac{\lambda_{d,\ell_1}\lambda_{e,\ell_2}}{f^*([d,e])} \sum_{p|[d,e]} \frac{\log p}{p} = \sideset{}{'}\sum_{p} \frac{\log p}{pf^*(p)} U(p),$$

where $U(p)$ was defined in (6.9). We employ the crude estimate

$$U(p) \ll \mathfrak{S}(\mathcal{H})^2\beta(\mathcal{H})(\log R)^{k+\ell_1+\ell_2+1}.$$

This is easily seen by combining (6.10), (6.37), (6.25), and using the symmetry between $U_1(p)$ and $U_2(p)$. The sum

$$\sideset{}{'}\sum_{p\le R} \frac{\log p}{pf^*(p)}$$

is $\ll 1$. Combining the above gives the bound

$$S_2 \ll \mathfrak{S}(\mathcal{H})^2\beta(\mathcal{H})(\log R)^{k+\ell_1+\ell_2+1}. \tag{6.45}$$

The proof of Theorem 7 is completed by combining (6.6) together with the final estimates for S_1, S_2, S_3, which are (6.7), (6.45), and (6.44) respectively.

7. Proofs of Theorems 1 through 4

Let $\mathcal{H} = \{h_1, h_2, \ldots, h_k\}$ be an arbitrary admissible k-tuple. Without loss of generality, we may specify that

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

It is also useful to assume that

$$h_k \leq \log N. \tag{7.1}$$

With this hypothesis, we see from Lemma 13 that the error terms in Theorems 5, 6, and 7 satisfy

$$\beta(\mathcal{H})\mathfrak{S}(\mathcal{H})/\log N \ll (\log\log\log N)^{w_k+1}/\log N \ll (\log\log N)/\log N.$$

Consider the sum

$$\mathcal{S}_1 := \sum_{N<n\leq 2N} \left\{\sum_{h\in\mathcal{H}} \varpi(n+h) \,-\, (\log 3N)\right\} \left(\sum_{\ell=0}^{L} b_\ell (\log R)^{-\ell}\Lambda_R(n;\mathcal{H},\ell)\right)^2. \tag{7.2}$$

For a given n, the sum inside the brackets is non-positive unless there are at least two distinct values, $h_i, h_j \in \mathcal{H}$ such that $n+h_i, n+h_j$ are primes. Consequently, if we can show that the sum in (7.2) is $\gg N\mathfrak{S}(\mathcal{H})(\log R)^{k+1}$, then we can conclude that $\liminf_{n\to\infty}(p_{n+1}-p_n) \leq h_k - h_1$.

Expanding the square in (7.2), we see that

$$\mathcal{S}_1 = \sum_{0\leq \ell_1,\ell_2\leq L} b_{\ell_1} b_{\ell_2} (\log R)^{-\ell_1-\ell_2}\mathcal{M}_1(\ell_1,\ell_2),$$

where

$$\mathcal{M}_1(\ell_1,\ell_2) = \sum_{N\leq n<2N} \left\{\sum_{h\in\mathcal{H}} \varpi(n+h) - (\log 3N)\right\} \Lambda_R(n;\mathcal{H},\ell_1)\Lambda_R(n;\mathcal{H},\ell_2).$$

We assume Hypothesis $BV(\theta)$, and we use Theorems 5 and 6 with $R = N^{(\theta-\epsilon)/2}$ to get

$$\begin{aligned}\mathcal{M}_1(\ell_1,\ell_2) &\sim \binom{\ell_1+\ell_2+2}{\ell_1+1} N\mathfrak{S}(\mathcal{H})k\frac{(\log R)^{k+\ell_1+\ell_2+1}}{(k+\ell_1+\ell_2+1)!} \\ &\qquad - \binom{\ell_1+\ell_2}{\ell_1} N\mathfrak{S}(\mathcal{H})\frac{(\log R)^{k+\ell_1+\ell_2}\log N}{(k+\ell_1+\ell_2)!} \\ &\sim N\mathfrak{S}(\mathcal{H})(\log R)^{k+\ell_1+\ell_2}(\log N)(m(k,\ell_1,\ell_2,\theta)-\epsilon'),\end{aligned}$$

where

$$m(k,\ell_1,\ell_2,\theta) = \binom{\ell_1+\ell_2}{\ell_1}\frac{1}{(k+\ell_1+\ell_2)!}\left(\frac{k(\ell_1+\ell_2+1)(\ell_1+\ell_2+2)}{(k+\ell_1+\ell_2+1)(\ell_1+1)(\ell_2+1)}\frac{\theta}{2}-1\right), \tag{7.3}$$

and $\epsilon' = \epsilon'(k,\ell_1,\ell_2,\epsilon)$ goes to 0 as ϵ goes to 0.

Define $\mathbf{b} = (b_0, b_1, \ldots, b_L)$. Then (we suppress the ϵ' term)

$$\begin{aligned}\mathcal{S}_1^*(N, \mathcal{H}, \theta, \mathbf{b}) &:= \frac{\mathcal{S}_1}{N\mathfrak{S}(\mathcal{H})(\log R)^k \log N} \\ &\sim \sum_{0\le \ell_1, \ell_2 \le L} b_{\ell_1} b_{\ell_2} m(k, \ell_1, \ell_2, \theta) \\ &= \mathbf{b}^T \mathbf{M} \mathbf{b},\end{aligned} \tag{7.4}$$

where $\mathbf{M} = \mathbf{M}(k, \theta)$ is the matrix

$$\mathbf{M} = [m(k, i, j, \theta)]_{0\le i,j\le L}\,.$$

Our goal is to pick $\mathbf{b}$ to make $\mathcal{S}_1^* > 0$ for a given θ and minimal k. This is easily determined by picking $\mathbf{b}$ to be an eigenvector of the matrix $\mathbf{M}$ with eigenvalue λ, in which case

$$\mathcal{S}_1^* \sim \mathbf{b}^T \lambda \mathbf{b} = \lambda \sum_{i=0}^{L} b_i^2.$$

This will be positive provided λ is positive. We conclude that $\mathcal{S}_1^* > 0$ if $\mathbf{M}$ has a positive eigenvalue and $\mathbf{b}$ is chosen to be the corresponding eigenvector.

With $k = 6$ and $L = 1$, we find that

$$\mathbf{M} = \frac{1}{8!}\begin{bmatrix} 48\theta - 56 & 9\theta - 8 \\ 9\theta - 8 & 2\theta - 2\end{bmatrix}.$$

The determinant of $8!\mathbf{M}$ is $15\theta^2 - 64\theta + 48$, which is negative if $4(8 - \sqrt{19})/15 < \theta \le 1$. Since the determinant is the product of the eigenvalues, we conclude that $\mathbf{M}$ has a positive eigenvalue for θ in this range. Consequently, if $\mathcal{H}$ is an admissible 6-tuple, then there are infinitely many n such that at least two of the numbers $n + h_1, \ldots, n + h_6$ are prime. We complete the proof of the second part of Theorem 2 by taking

$$\mathcal{H} = \{7, 11, 13, 17, 19, 23\}.$$

$\mathcal{H}$ is admissible because for $p \le 5$, none of the elements in $\mathcal{H}$ are divisible by p, and for $p \ge 7$, there are not enough elements to cover all of the residue classes mod p.

To prove the first part of Theorem 2, we again use (7.4); however, we use the trivial choice $b_\ell = 1$ for some specific ℓ, and $b_i = 0$ for all other i. Then

$$\mathcal{S}_1^* \sim m(k, \ell, \ell, \theta) = \binom{2\ell}{\ell}\frac{1}{(k + 2\ell)!}\left(\frac{2k(2\ell + 1)}{(k + 2\ell + 1)(\ell + 1)}\frac{\theta}{2} - 1\right) - \epsilon'.$$

The above is positive if

$$\theta > \left(\frac{1}{2} + \frac{1}{4\ell + 2}\right)\left(1 + \frac{2\ell + 1}{k}\right).$$

The right-hand side approaches $1/2$ if $\ell, k \to \infty$ with $\ell = o(k)$.

The above argument just fails when $\theta = 1/2$. To remedy this, we modify (7.2) by taking h to be a parameter to be chosen later, with $h \le \log N$. We then sum over all admissible size k subsets $\mathcal{H}$ of $\{1, \ldots, h\}$. Specifically, we take

$$\tilde{\mathcal{S}}_1 = \sum_{\substack{\mathcal{H}\subseteq\{1,\ldots,h\} \\ |\mathcal{H}|=k \\ \mathcal{H} \text{ admissible}}} \sum_{N<n\le 2N} \left\{\sum_{1\le h_0\le h} \varpi(n + h_0) \;-\; (\log 3N)\right\} \Lambda_R^2(n; \mathcal{H}, \ell). \tag{7.5}$$

We apply Theorems 5 and 6 to the sum $\tilde{\mathcal{S}}_1$ for those terms when $\mathcal{H}$ and $\mathcal{H}\cup\{h_0\}$ are both admissible. There may be terms with $\mathcal{H}$ admissible but $\mathcal{H}\cup\{h_0\}$ not admissible; for these terms we apply the trivial bound

$$\sum_{N<n\le 2N}\sum_{1\le h_0\le h}\varpi(n+h_0)\Lambda_R(n;\mathcal{H},\ell)^2\ge 0.$$

We find that

$$\begin{aligned}(7.6)\qquad \tilde{\mathcal{S}}_1 \gtrsim &\binom{2\ell+2}{\ell+1}\frac{N(\log R)^{k+2\ell+1}}{(k+2\ell+1)!}\sum_{1\le h_0\le h}\sum_{\substack{\mathcal{H}\subseteq\{1,\ldots,h\}\\ |\mathcal{H}|=k,h_0\in\mathcal{H}}}\mathfrak{S}(\mathcal{H})\\ &+\binom{2\ell}{\ell}\frac{N(\log R)^{k+2\ell}}{(k+2\ell)!}\sum_{1\le h_0\le h}\sum_{\substack{\mathcal{H}\subseteq\{1,\ldots,h\}\\ |\mathcal{H}|=k,h_0\notin\mathcal{H}}}\mathfrak{S}\left(\mathcal{H}\cup\{h_0\}\right)\\ &-\binom{2\ell}{\ell}\frac{N(\log N)(\log R)^{k+2\ell}}{(k+2\ell)!}\sum_{\substack{\mathcal{H}\subseteq\{1,\ldots,h\}\\ |\mathcal{H}|=k}}\mathfrak{S}(\mathcal{H}).\end{aligned}$$

We have dropped the condition that $\mathcal{H}$ is admissible in the above sums; we may do so because $\mathfrak{S}(\mathcal{H})=0$ when $\mathcal{H}$ is not admissible.

Now we observe that

$$\sum_{1\le h_0\le h}\sum_{\substack{\mathcal{H}\subseteq\{1,\ldots,h\}\\ |\mathcal{H}|=k,h_0\in\mathcal{H}}}\mathfrak{S}(\mathcal{H})=k\sum_{\substack{\mathcal{H}\subseteq\{1,\ldots,h\}\\ |\mathcal{H}|=k}}\mathfrak{S}(\mathcal{H})\sim\frac{kh^k}{k!}.$$

In the above, equality occurs from noting that every relevant set $\mathcal{H}$ occurs k times in the initial sum, and the asymptotic relation is a theorem of Gallagher [8]. We also have that

$$\sum_{1\le h_0\le h}\sum_{\substack{\mathcal{H}\subseteq\{1,\ldots,h\}\\ |\mathcal{H}|=k,h_0\notin\mathcal{H}}}\mathfrak{S}(\mathcal{H}\cup\{h_0\})=(k+1)\sum_{\substack{\mathcal{H}\subseteq\{1,\ldots,h\}\\ |\mathcal{H}|=k+1}}\mathfrak{S}(\mathcal{H})\sim\frac{h^{k+1}}{k!}.$$

Returning to the evaluation of $\tilde{\mathcal{S}}_1$, we find that

$$\tilde{\mathcal{S}}_1\gtrsim\binom{2\ell}{\ell}\frac{N(\log N)(\log R)^{k+2\ell}h^k}{k!(k+2\ell)!}\tilde{b}_1(k,\ell,h),$$

where

$$\tilde{b}_1(k,\ell,h)=2\cdot\frac{2\ell+1}{\ell+1}\cdot\frac{k}{k+2\ell+1}\cdot\frac{\log R}{\log N}+\frac{h}{\log N}-1.$$

Unconditionally, we may take $\theta=1/2$, so $\log R/\log N=1/4-\epsilon$. We get two primes in some interval $(n,n+h], N<n\le 2N$, provided $\tilde{b}_1(k,\ell,h)>0$. This is equivalent to

$$\begin{aligned}\frac{h}{\log N}>&1-\frac{2k}{k+2\ell+1}\cdot\frac{2\ell+1}{\ell+1}\cdot\left(\frac{1}{4}-\epsilon\right)\\ =&\frac{k+4\ell^2+6\ell+2+4\epsilon(k+2k\ell)}{2(1+\ell)(1+2\ell+k)}.\end{aligned}$$

On letting $\ell=[\sqrt{k}]$ and taking k sufficiently large, we see that this is valid with $h/\log N$ arbitrarily small. This proves Theorem 1.

For the proofs of Theorem 3 and Theorem 4, we note that if $N < n \le 2N$, then

$$\varpi * \varpi(n) \le \frac{(\log 3N)^2}{2}.$$

Accordingly, we consider

$$\mathcal{S}_2 := \sum_{N<n\le 2N} \left\{ \sum_{h\in\mathcal{H}} \varpi * \varpi(n+h) \;-\; \frac{(\log 3N)^2}{2} \right\} \times \left(\sum_{\ell=0}^{L} b_\ell (\log R)^{-\ell} \Lambda_R(n;\mathcal{H},\ell) \right)^2. \tag{7.7}$$

The term n contributes a negative amount unless there are two values $h_i, h_j \in \mathcal{H}$ such that $n+h_i, n+h_j$ are products of two primes. The values of n for which any $n+h$ is a square of a prime contribute $\ll N^{1/2}(\log N)^{2k+2}$, and this contribution may be absorbed into the error terms of our estimates.

We assume Hypotheses $BV(\theta)$ and $BV_2(\theta)$, and we argue along the same lines as in the proof of Theorem 2. When $R = N^{(\theta-\epsilon)/2}$, we obtain

$$\mathcal{S}_2 = \sum_{0\le \ell_1,\ell_2\le L} b_{\ell_1} b_{\ell_2} (\log R)^{-\ell_1-\ell_2} \mathcal{M}_2(\ell_1,\ell_2),$$

where

$$\mathcal{M}_2 \sim \mathfrak{S}(\mathcal{H}) N (\log N)^2 (\log R)^{k+\ell_1+\ell_2} (m_2(k,\ell_1,\ell_2,\theta) - \epsilon'),$$

$$m_2(k,\ell_1,\ell_2,\theta) = m_{21} + m_{22} - m_{23},$$

$$m_{21} = \binom{\ell_1+\ell_2+2}{\ell_1+1} \frac{k}{(k+\ell_1+\ell_2+1)!} \frac{\theta}{2},$$

$$m_{22} = 2\left\{ \binom{\ell_1+\ell_2+2}{\ell_1+1} - \binom{\ell_1+\ell_2+3}{\ell_1+1} - \binom{\ell_1+\ell_2+3}{\ell_2+1} \right\} \frac{k}{(k+\ell_1+\ell_2+2)!} \frac{\theta^2}{4},$$

$$m_{23} = \frac{1}{2} \binom{\ell_1+\ell_2}{\ell_1} \frac{1}{(k+\ell_1+\ell_2)!},$$

and $\epsilon' = \epsilon'(k,\ell_1,\ell_2,\epsilon) \to 0$ as $\epsilon \to 0$.

Let $\mathbf{b}$ be as defined before. Then (suppressing the ϵ' term)

$$\begin{aligned} \mathcal{S}_2^*(N,\mathcal{H},\theta,\mathbf{b}) &:= \frac{\mathcal{S}_2}{N\mathfrak{S}(\mathcal{H})(\log R)^k (\log N)^2} \sim \sum_{0\le \ell_1,\ell_2\le L} b_{\ell_1} b_{\ell_2} m_2(k,\ell_1,\ell_2,\theta) \\ &= \mathbf{b}^T \mathbf{M}_2 \mathbf{b}, \end{aligned}$$

where $\mathbf{M}_2 = \mathbf{M}_2(k,\theta)$ is the matrix

$$\mathbf{M}_2 = [m_2(k,i,j,\theta)]_{0\le i,j\le L}.$$

We first prove Theorem 4. As in the proof of Theorem 2, we wish to show that there is some $\mathbf{b}$ such that $\mathcal{S}_2^* > 0$ for a given θ and minimal k. Taking $k=3$ and $L=1$, we find that

$$\mathbf{M}_2 = \frac{1}{480} \begin{bmatrix} -24\theta^2 + 60\theta - 40 & -7\theta^2 + 18\theta - 10 \\ -7\theta^2 + 18\theta - 10 & -2\theta^2 + 6\theta - 4 \end{bmatrix}. \tag{7.8}$$

If we take $b_0 = 1, b_1 = 4$, then we find that

$$\mathbf{b}^T \mathbf{M}_2 \mathbf{b} = -\frac{7\theta^2}{30} + \frac{5\theta}{8} - \frac{23}{60}.$$

This is positive whenever

$$\frac{75-\sqrt{473}}{56} < \theta \leq 1.$$

Finally, we note that $\mathcal{H} = \{5, 7, 11\}$ is an admissible triple, so this completes the proof of Theorem 4.

We can also prove Theorem 4 with a slightly wider range of allowable θ by taking the determinant of the matrix in (7.8). A numerical calculation shows that this determinant has a zero at $\theta = 0.943635\ldots$.

For the proof of Theorem 3, we take $k = 8, L = 2, \theta = 1/2 - \epsilon$, and we find that

$$\mathbf{M}_2 = \frac{1}{14!}\begin{bmatrix} -216216 & 8736 & 3458 \\ 8736 & -364 & 14 \\ 3458 & 14 & -36 \end{bmatrix}.$$

With

$$b_0 = 1, b_1 = 16, b_2 = 16,$$

we find that

$$14!\mathbf{b}^T\mathbf{M}\mathbf{b} = 78760 > 0.$$

Now $\mathcal{H} = \{11, 13, 17, 19, 23, 29, 31, 37\}$ is an admissible octuple, so this completes the proof of Theorem 3.

We make one final comment regarding the proofs that make use of bilinear forms in $\mathbf{b}$. By taking

$$\sum_{\ell=0}^{L} b_\ell (\log R)^{-\ell} \Lambda_R(n; \mathcal{H}, \ell)$$

in the definitions of $\mathcal{S}_1$ and $\mathcal{S}_2$, we are in essence using

$$y_r = \mathfrak{S}(\mathcal{H}) \sum_{\ell=0}^{L} \frac{b_\ell}{\ell!}\left(\frac{\log R/r}{\log R}\right)^\ell.$$

In other words, we have essentially replaced $(\log R/r)^\ell$ in (1.20) by a polynomial in $\log R/r$.

Acknowledgements

We thank Tsz-Ho Chan, Yoichi Motohashi, and the anonymous referees for their comments on this paper. Part of the work for this paper was done at the American Institute of Mathematics, where the second author was visiting in Fall 2004. He thanks them for their hospitality and excellent working environment.

References

[1] E. Bombieri, *The large sieve,* Mathematika **12** (1965), 201–225. MR0197425 (33:5590)
[2] E. Bombieri and H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. Ser. A **293** (1966), 1–18. MR0199165 (33:7314)
[3] J.-R. Chen, *On the representation of a large even integer as the sum of a prime and a product of at most two primes*, Scientia Sinica **16** (1973), 157-176. MR0434997 (55:7959)
[4] H. Davenport, *Multiplicative Number Theory*, Second Edition, revised by H.L. Montgomery, Springer, Berlin, Heidelberg, New York, 1980. MR606931 (82m:10001)
[5] L.E. Dickson, *History of the Theory of Numbers*, Vol. I, Chelsea, New York.
[6] P.D.T.A. Elliott and H. Halberstam, *A conjecture in prime number theory,* Symposia Mathematica **4** (INDAM, Rome, 1968/69), 59–72, Academic Press, London. MR0276195 (43:1943)
[7] P. Erdős, *The difference of consecutive primes,* Duke Math. J. **6** (1940), 438–441. MR0001759 (1:292h)

[8] P.X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), 4–9, *Corrigendum*, Mathematika **28** (1981), 86. MR0409385 (53:13140)
[9] D.A. Goldston and C.Y. Yıldırım, *Higher correlations of divisor sums related to primes III: Small gaps between primes*, Proc. London Math. Soc., to appear.
[10] D. A. Goldston, J. Pintz, and C.Y. Yıldırım, *Primes in tuples I*, Annals of Mathematics, to appear.
[11] D. A. Goldston, J. Pintz, and C.Y. Yıldırım, *Primes in tuples II*, preprint.
[12] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, New York, 1974. MR0424730 (54:12689)
[13] H. Halberstam and K.F. Roth, *Sequences*, 2$^{\text{nd}}$ Edition, Springer-Verlag, New York, 1983. MR687978 (83m:10094)
[14] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum': III On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70. MR1555183
[15] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum': VII*, unpublished manuscript; see [24].
[16] D. R. Heath-Brown, The divisor function at consecutive integers, *Mathematika* **31** (1984), 141–149. MR762186 (86c:11071)
[17] D.R. Heath-Brown, *Almost prime k-tuples*, Mathematika **44** (1997), 245-266. MR1600529 (99a:11106)
[18] A. Hildebrand, *Über die punktweise Konvergenz von Ramanujan-Entwicklungen zahlentheoretischer Funktionen*, Acta Arithmetica **44** (1984), 109-140. MR774094 (86d:11078)
[19] M.N. Huxley, *An application of the Fouvry-Iwaniec theorem*, Acta Arithmetica **43** (1984), 441-443. MR756293 (85k:11043)
[20] H. Maier, *Small differences between prime numbers.* Michigan Math. Journal **351** (1988), 323–344. MR978303 (90e:11126)
[21] H.L. Montgomery and R.C. Vaughan, *Multiplicative Number Theory I: Classical Theory (Cambridge Studies in Advanced Mathematics)*, Cambridge University Press, 2007. MR2378655
[22] Y. Motohashi, *An induction principle for the generalization of Bombieri's prime number theorem*, Proc. Japan Acad. **52** (1976), 273–275. MR0422179 (54:10171)
[23] A. de Polignac, *Six propostions arithmologiques déduites de crible d'Ératosthène,* Nouv. Ann. Math. **8** (1849), 423-429.
[24] R. A. Rankin, *The difference between consecutive prime numbers II.* Proc. Cambbridge Philos. Soc. **36** (1940), 255-266. MR0001760 (1:292i)
[25] J.-C. Schlage-Puchta, The equation $\omega(n) = \omega(n+1)$, *Mathematika* **50** (2003), no. 1-2 (2005), 99–101. MR2136354 (2005k:11198)
[26] A. Selberg, *Lectures on Sieves*, Collected Papers, Volume II, Springer, 1992, pp. 65–247. MR1295844 (95g:01032)
[27] R.C. Vaughan, *An elementary method in prime number theory,* Acta Arith. **37** (1980), 111–115. MR598869 (82c:10055)
[28] A.I. Vinogradov, *On the density hypothesis for Dirichlet L-functions,* Izv. Akad. Nauk SSSR Ser. Mat. **29** (1965), 903-934. Corrigendum, *loc. cit.* **30** (1966), 719–720.

Department of Mathematics, San Jose State University, San Jose, California 95192
E-mail address: goldston@math.sjsu.edu

Department of Mathematics, Central Michigan University, Mt. Pleasant, Michigan 48859
E-mail address: graha1sw@cmich.edu

Rényi Mathematical Institute of the Hungarian Academy of Sciences, H-1053 Budapest, Realtanoda u. 13–15, Hungary
E-mail address: pintz@renyi.hu

Department of Mathematics, Boğaziçi University, Istanbul 34342, Turkey – and – Feza Gürsey Enstitüsü, Çengelköy, Istanbul, P.K. 6, 81220, Turkey
E-mail address: yalciny@boun.edu.tr