ANNALS OF MATHEMATICS

Primes in tuples I

By Daniel A. Goldston, János Pintz, and Cem Y. Yildirim

SECOND SERIES, VOL. 170, NO. 2

September, 2009

ANMAAH

Annals of Mathematics, **170** (2009), 819–862

Primes in tuples I

By Daniel A. Goldston, János Pintz, and Cem Y. Yildirim

Abstract

We introduce a method for showing that there exist prime numbers which are very close together. The method depends on the level of distribution of primes in arithmetic progressions. Assuming the Elliott-Halberstam conjecture, we prove that there are infinitely often primes differing by 16 or less. Even a much weaker conjecture implies that there are infinitely often primes a bounded distance apart. Unconditionally, we prove that there exist consecutive primes which are closer than any arbitrarily small multiple of the average spacing, that is,

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

We will quantify this result further in a later paper.

1. Introduction

One of the most important unsolved problems in number theory is to establish the existence of infinitely many prime tuples. Not only is this problem believed to be difficult, but it has also earned the reputation among most mathematicians in the field as hopeless in the sense that there is no known unconditional approach for tackling the problem. The purpose of this paper, the first in a series, is to provide what we believe is a method which could lead to a partial solution for this problem. At present, our results on primes in tuples are conditional on information about the distribution of primes in arithmetic progressions. However, the information needed to prove that there are infinitely often two primes in a given k-tuple for sufficiently large k does not seem to be too far beyond the currently known results. Moreover, we can gain enough in the argument by averaging over many tuples to obtain unconditional results concerning small gaps between primes which go

Goldston was supported by NSF grant DMS-0300563, the NSF Focused Research Group grant 0244660, and the American Institute of Mathematics; Pintz by OTKA grants No. T38396, T43623, T49693 and the Balaton program; Yıldırım by TÜBİTAK.

far beyond anything that has been proved before. Thus, we are able to prove the existence of very small gaps between primes which, however, go slowly to infinity with the size of the primes.

The information on primes we utilize in our method is often referred to as the level of distribution of primes in arithmetic progressions. Let

$$\theta(n) = \begin{cases} \log n & \text{if } n \text{ is prime,} \\ 0 & \text{otherwise,} \end{cases} \tag{1.1}$$

and consider the counting function

$$\theta(N;q,a) = \sum_{\substack{n\le N \\ n\equiv a \pmod q}} \theta(n). \tag{1.2}$$

The Bombieri-Vinogradov theorem states that for any $A > 0$ there is a $B = B(A)$ such that, for $Q = N^{\frac{1}{2}}(\log N)^{-B}$,

$$\sum_{q\le Q} \max_{\substack{a \\ (a,q)=1}} \left|\theta(N;q,a) - \frac{N}{\phi(q)}\right| \ll \frac{N}{(\log N)^A}. \tag{1.3}$$

We say that the primes have *level of distribution* ϑ if (1.3) holds for any $A > 0$ and any $\varepsilon > 0$ with

$$Q = N^{\vartheta-\varepsilon}. \tag{1.4}$$

Elliott and Halberstam [5] conjectured that the primes have level of distribution 1. According to the Bombieri-Vinogradov theorem, the primes are known to have level of distribution $1/2$.

Let n be a natural number and consider the k-tuple

$$(n+h_1, n+h_2, \ldots, n+h_k), \tag{1.5}$$

where $\mathcal{H} = \{h_1, h_2, \ldots, h_k\}$ is a set composed of distinct non-negative integers. If every component of the tuple is a prime we call this a *prime tuple*. Letting n range over the natural numbers, we wish to see how often (1.5) is a prime tuple. For instance, consider $\mathcal{H} = \{0, 1\}$ and the tuple $(n, n+1)$. If $n = 2$, we have the prime tuple $(2, 3)$. Notice that this is the only prime tuple of this form because, for $n > 2$, one of the numbers n or $n+1$ is an even number bigger than 2. On the other hand, if $\mathcal{H} = \{0, 2\}$, then we expect that there are infinitely many prime tuples of the form $(n, n+2)$. This is the twin prime conjecture. In general, the tuple (1.5) can be a prime tuple for more than one n only if for every prime p the h_i's never occupy all of the residue classes modulo p. This is immediately true for all primes $p > k$; so to test this condition we need only to examine small primes. If we denote by $\nu_p(\mathcal{H})$ the number of distinct residue classes modulo p occupied by the integers h_i, then we can avoid p dividing some component of (1.5) for every n by requiring

$$\nu_p(\mathcal{H}) < p \quad \text{for all primes } p. \tag{1.6}$$

If this condition holds we say that $\mathcal{H}$ is *admissible* and we call the tuple (1.5) corresponding to this $\mathcal{H}$ an *admissible tuple*. It is a long-standing conjecture that admissible tuples will infinitely often be prime tuples. Our first result is a step towards confirming this conjecture.

THEOREM 1. *Suppose the primes have level of distribution $\vartheta > 1/2$. Then there exists an explicitly calculable constant $C(\vartheta)$ depending only on ϑ such that any admissible k-tuple with $k \geq C(\vartheta)$ contains at least two primes infinitely often. Specifically, if $\vartheta \geq 0.971$, then this is true for $k \geq 6$.*

Since the 6-tuple $(n, n+4, n+6, n+10, n+12, n+16)$ is admissible, the Elliott-Halberstam conjecture implies that

$$\liminf_{n\to\infty}(p_{n+1} - p_n) \leq 16, \tag{1.7}$$

where the notation p_n is used to denote the n-th prime. This means that $p_{n+1} - p_n \leq 16$ for infinitely many n. Unconditionally, we prove a long-standing conjecture concerning gaps between consecutive primes.

THEOREM 2. *We have*

$$\Delta_1 := \liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\log p_n} = 0. \tag{1.8}$$

There is a long history of results on this topic which we will briefly mention. The inequality $\Delta_1 \leq 1$ is a trivial consequence of the prime number theorem. The first result of type $\Delta_1 < 1$ was proved in 1926 by Hardy and Littlewood [18], who on assuming the Generalized Riemann Hypothesis (GRH) obtained $\Delta_1 \leq 2/3$. This result was improved by Rankin [26] to $\Delta_1 \leq 3/5$, also assuming the GRH. The first unconditional estimate was proved by Erdős [7] in 1940. Using Brun's sieve, he showed that $\Delta_1 < 1 - c$ with an unspecified positive explicitly calculable constant c. His estimate was improved by Ricci [27] in 1954 to $\Delta_1 \leq 15/16$. In 1965, Bombieri and Davenport [2] refined and made unconditional the method of Hardy and Littlewood by substituting the Bombieri-Vinogradov theorem for the GRH, and obtained $\Delta_1 \leq 1/2$. They also combined their method with the method of Erdős and obtained $\Delta_1 \leq 0.4665\ldots$. Their result was further refined by Pilt'ai [25] to $\Delta_1 \leq 0.4571\ldots$, Uchiyama [33] to $\Delta_1 \leq 0.4542\ldots$ and in several steps by Huxley [20], [21] to yield $\Delta_1 \leq 0.4425\ldots$, and finally in 1984 to $\Delta_1 \leq .4393\ldots$ [22]. This was further improved by Fouvry and Grupp [9] to $\Delta_1 \leq .4342\ldots$. In 1988 Maier [23] used his matrix-method to improve Huxley's result to $\Delta_1 \leq e^{-\gamma} \cdot 0.4425\ldots = 0.2484\ldots$, where γ is Euler's constant. Maier's method by itself gives $\Delta_1 \leq e^{-\gamma} = 0.5614\ldots$. The recent version of the method of Goldston and Yıldırım [13] led, without combination with other methods, to $\Delta_1 \le 1/4$.

In a later paper in this series we will prove the quantitative result that

$$\liminf_{n\to\infty} \frac{p_{n+1}-p_n}{(\log p_n)^{\frac{1}{2}}(\log\log p_n)^2} < \infty. \tag{1.9}$$

While Theorem 1 is a striking new result, it also reflects the limitations of our current method. Whether these limitations are real or can be overcome is a critical issue for further investigation. We highlight the following four questions.

Question 1. Can it be proved unconditionally by the current method that there are, infinitely often, bounded gaps between primes? Theorem 1 would appear to be within a hair's breadth of obtaining this result. However, any improvement in the level of distribution ϑ beyond $1/2$ probably lies very deep, and even the GRH does not help. Still, there are stronger versions of the Bombieri-Vinogradov theorem, as found in [3], and the circle of ideas used to prove these results, which may help to obtain this result.

Question 2. Is $\vartheta = 1/2$ a true barrier for obtaining primes in tuples? Soundararajan [31] has demonstrated this is the case for the current argument, but perhaps more efficient arguments may be devised.

Question 3. Assuming the Elliott-Halberstam conjecture, can it be proved that there are three or more primes in admissible k-tuples with large enough k? Even under the strongest assumptions, our method fails to prove anything about more than two primes in a given tuple.

Question 4. Assuming the Elliott-Halberstam conjecture, can the twin prime conjecture be proved with a refinement of our method?

The limitation of our method, identified in Question 3, is the reason we are less successful in finding more than two primes close together. However, we are able to improve on earlier results, in particular the recent results in [13]. For $\nu \ge 1$, let

$$\Delta_\nu = \liminf_{n\to\infty} \frac{p_{n+\nu}-p_n}{\log p_n}. \tag{1.10}$$

Bombieri and Davenport [2] showed $\Delta_\nu \le \nu - 1/2$. This bound was later improved by Huxley [20], [21] to $\Delta_\nu \le \nu - 5/8 + O(1/\nu)$, by Goldston and Yıldırım [13] to $\Delta_\nu \le (\sqrt{\nu} - 1/2)^2$, and by Maier [23] to $\Delta_\nu \le e^{-\gamma}\,(\nu - 5/8 + O\,(1/\nu))$. In proving Theorem 2 we will also show, assuming the primes have level of distribution ϑ,

$$\Delta_\nu \le \max(\nu - 2\vartheta, 0), \tag{1.11}$$

and hence unconditionally $\Delta_\nu \le \nu - 1$. However, by a more complicated argument, we will prove the following result.

THEOREM 3. *Suppose the primes have level of distribution ϑ. Then for $\nu \ge 2$,*

$$\Delta_\nu \le (\sqrt{\nu} - \sqrt{2\vartheta})^2. \tag{1.12}$$

In particular, we have unconditionally, for $\nu \ge 1$,

$$\Delta_\nu \le (\sqrt{\nu} - 1)^2. \tag{1.13}$$

From (1.11) or (1.12) we see that the Elliott-Halberstam conjecture implies that

$$\Delta_2 = \liminf_{n\to\infty} \frac{p_{n+2} - p_n}{\log p_n} = 0. \tag{1.14}$$

We can improve on (1.13) by combining our method with Maier's matrix method [23] to obtain

$$\Delta_\nu \le e^{-\gamma}(\sqrt{\nu} - 1)^2. \tag{1.15}$$

Huxley [20] generalized the results of Bombieri and Davenport [2] for Δ_ν to primes in arithmetic progressions with a fixed modulus. We are able to prove the analogue of (1.15) for primes in arithmetic progressions where the modulus can tend slowly to infinity with the size of the primes considered.

Another extension of our work is that we can find primes in other sets besides intervals. Thus we can prove that there are two primes among the numbers $n + a_i$, $1 \le i \le h$, for $N < n \le 2N$ and the a_i's are given arbitrary integers in the interval $[1, N]$ if $h < C\sqrt{\log N}(\log\log N)^2$ and N is restricted to some sequence N_ν tending to infinity, which avoids Siegel zeros for moduli near to N. It is interesting to note that such a general result can be proved regardless of the distribution of the a_i values, in contrast to our present case where Gallagher's theorem (3.7) requires the a_i's to lie in an interval. The proofs of these results will appear in later papers in this series.

While this paper is our first paper on this subject, we have two other papers that overlap some of the results here. The first paper [15], written jointly with Motohashi, gives a short and simplified proof of Theorems 1 and 2. The second paper [14], written jointly with Graham, uses sieve methods to prove Theorems 1 and 2 and provides applications for tuples of almost-primes (products of a bounded number of primes.)

The present paper is organized as follows. In Section 2, we describe our method and its relation to earlier work. We also state Propositions 1 and 2 which incorporate the key new ideas in this paper. These are developed in a more general form than in [14] or [15] so as to be employable in many applications. In Section 3, we prove Theorems 1 and 2 using these propositions. The method of proof is due to Granville and Soundararajan. In Section 4 we make some further comments on the method used in Section 3. In Section 5 we prove two lemmas needed later. In Section 6, we prove a special case of Proposition 1 which illustrates the key points in the general case. In Section 7 we begin the proof of Proposition 1 which is reduced to evaluating a certain contour integral. In Section 8 we evaluate a more general contour integral that occurs in the proof of both propositions. In Section 9, we prove Proposition 2. In this paper we do not obtain results that are uniform in k, and therefore we assume here that our tuples have a fixed length. However, uniform results are needed for (1.9), and they will be the topic of the next paper in this series. Finally, we prove Theorem 3 in Section 10.

Notation. In the following, c and C will denote (sufficiently) small and (sufficiently) large absolute positive constants, respectively, which have been chosen appropriately. This is also true for constants formed from c or C with subscripts or accents. We unconventionally will allow these constants to be different at different occurrences. Constants implied by pure o, O, $\ll$ symbols will be absolute, unless otherwise stated. $[S]$ is 1 if the statement S is true and is 0 if S is false. The symbol $\sum^{\flat}$ indicates the summation is over squarefree integers, and $\sum'$ indicates the summation variables are pairwise relatively prime.

The ideas used in this paper have developed over many years. We are indebted to many people, not all of whom we can mention. In particular, we would like to thank A. Balog, E. Bombieri, T. H. Chan, J. B. Conrey, P. Deift, D. Farmer, K. Ford, J. Friedlander, S. W. Graham, A. Granville, C. Hughes, D. R. Heath-Brown, A. Ledoan, H. L. Montgomery, Y. Motohashi, Sz. Gy. Revesz, P. Sarnak, J. Sivak, and K. Soundararajan.

2. Approximating prime tuples

Let

$$\mathcal{H} = \{h_1, h_2, \ldots, h_k\} \text{ with } 1 \le h_1, h_2, \ldots, h_k \le h \text{ distinct integers,} \tag{2.1}$$

and let $\nu_p(\mathcal{H})$ denote the number of distinct residue classes modulo p occupied by the elements of $\mathcal{H}$.[1] For squarefree integers d, we extend this definition to $\nu_d(\mathcal{H})$ by multiplicativity. We denote by

$$\mathfrak{S}(\mathcal{H}) := \prod_p \left(1 - \frac{1}{p}\right)^{-k} \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right) \tag{2.2}$$

the singular series associated with $\mathcal{H}$. Since $\nu_p(\mathcal{H}) = k$ for $p > h$, we see that the product is convergent and therefore $\mathcal{H}$ is admissible as defined in (1.6) if and only

[1] The restriction of the set $\mathcal{H}$ to positive integers is only for simplicity, and, if desired, can easily be removed later from all of our results.

if $\mathfrak{S}(\mathcal{H}) \neq 0$. Hardy and Littlewood conjectured an asymptotic formula for the number of prime tuples $(n+h_1, n+h_2, \dots, n+h_k)$, with $1 \le n \le N$, as $N \to \infty$. Let $\Lambda(n)$ denote the von Mangoldt function which equals $\log p$ if $n = p^m$, $m \ge 1$, and zero otherwise. We define

$$\Lambda(n;\mathcal{H}) := \Lambda(n+h_1)\Lambda(n+h_2)\cdots\Lambda(n+h_k) \tag{2.3}$$

and use this function to detect prime tuples and tuples with prime powers in components, the latter of which can be removed in applications. The Hardy-Littlewood prime-tuple conjecture [17] can be stated in the form

$$\sum_{n\le N} \Lambda(n;\mathcal{H}) = N(\mathfrak{S}(\mathcal{H}) + o(1)), \quad \text{as } N \to \infty. \tag{2.4}$$

(This conjecture is trivially true if $\mathcal{H}$ is not admissible.) Except for the prime number theorem (1-tuples), this conjecture is unproved.[2]

The program the first and third authors have been working on since 1999 is to compute approximations for (2.3) with $k \ge 3$ using short divisor sums and to apply the results to problems on primes. The simplest approximation of $\Lambda(n)$ is based on the elementary formula

$$\Lambda(n) = \sum_{d\mid n} \mu(d) \log \frac{n}{d}, \tag{2.5}$$

which can be approximated with the smoothly truncated divisor sum

$$\Lambda_R(n) = \sum_{\substack{d\mid n\\ d\le R}} \mu(d) \log \frac{R}{d}. \tag{2.6}$$

Thus, an approximation for $\Lambda(n;\mathcal{H})$ is given by

$$\Lambda_R(n+h_1)\Lambda_R(n+h_2)\cdots\Lambda_R(n+h_k). \tag{2.7}$$

In [13], Goldston and Yıldırım applied (2.7) to detect small gaps between primes and proved

$$\Delta_1 = \liminf_{n\to\infty} \left(\frac{p_{n+1}-p_n}{\log p_n}\right) \le \frac{1}{4}.$$

In this paper we introduce a new approximation, the idea for which came partly from a paper of Heath-Brown [19] on almost prime tuples. His result is itself a generalization of Selberg's proof from 1951 (see [29, pp. 233–245]) that the polynomial $n(n+2)$ will infinitely often have at most five distinct prime factors, so that the same is true for the tuple $(n, n+2)$. Not only does our approximation have its origin in these papers, but in hindsight the argument of Granville and

[2] Asymptotic results for the number of primes in tuples, unlike the existence result in Theorem 1, are beyond the reach of our method.

Soundararajan (employed in the proof of Theorems 1 and 2) is essentially the same as the method used in these papers.

In connection with the tuple (1.5), we consider the polynomial

$$P_{\mathcal{H}}(n) = (n + h_1)(n + h_2) \cdots (n + h_k). \tag{2.8}$$

If the tuple (1.5) is a prime tuple then $P_{\mathcal{H}}(n)$ has exactly k prime factors. We detect this condition by using the k-th generalized von Mangoldt function

$$\Lambda_k(n) = \sum_{d|n} \mu(d)\left(\log \frac{n}{d}\right)^k, \tag{2.9}$$

which vanishes if n has more than k distinct prime factors.[3] With this, our prime tuple detecting function becomes

$$\Lambda_k(n; \mathcal{H}) := \frac{1}{k!}\Lambda_k(P_{\mathcal{H}}(n)). \tag{2.10}$$

The normalization factor $1/k!$ simplifies the statement of our results. As we will see in Section 5, this approximation suggests the Hardy-Littlewood type conjecture

$$\sum_{n \le N} \Lambda_k(n; \mathcal{H}) = N\left(\mathfrak{S}(\mathcal{H}) + o(1)\right). \tag{2.11}$$

This is a special case of the general conjecture of Bateman–Horn [1] which is the quantitative form of Schinzel's conjecture [28].

In analogy with (2.6) (when $k = 1$), we approximate $\Lambda_k(n)$ by the smoothed and truncated divisor sum

$$\sum_{\substack{d|n \\ d \le R}} \mu(d)\left(\log \frac{R}{d}\right)^k$$

and define

$$\Lambda_R(n; \mathcal{H}) = \frac{1}{k!} \sum_{\substack{d|P_{\mathcal{H}}(n) \\ d \le R}} \mu(d)\left(\log \frac{R}{d}\right)^k. \tag{2.12}$$

However, as we will see in the next section, this approximation is not adequate to prove Theorems 1 and 2.

A second simple but crucial idea is needed: rather than only approximate prime tuples, one should approximate tuples with primes in many components. Thus, we consider when $P_{\mathcal{H}}(n)$ has $k + \ell$ or fewer distinct prime factors, where

[3] As with $\Lambda(n)$, we overcount the prime tuples by including factors which are proper prime powers, but these can be removed in applications with a negligible error. The slightly misleading notational conflict between the generalized von Mangoldt function Λ_k and Λ_R will only occur in this section.

$0 \le \ell \le k$, and define

$$\Lambda_R(n;\mathcal{H},\ell) = \frac{1}{(k+\ell)!}\sum_{\substack{d\mid P_{\mathcal{H}}(n)\\ d\le R}} \mu(d)\left(\log\frac{R}{d}\right)^{k+\ell}, \tag{2.13}$$

where $|\mathcal{H}| = k$. If $\mathcal{H} = \varnothing$, then $k = \ell = 0$ and we define $\Lambda_R(n;\varnothing,0) = 1$.

The advantage of (2.13) over (2.7) can be seen as follows. If in (2.13) we restrict ourselves to d's with all prime factors larger than h, then the condition $d\mid P_{\mathcal{H}}(n)$ implies that we can write $d = d_1 d_2 \cdots d_k$ uniquely with $d_i \mid n+h_i$, $1 \le i \le k$, the d_i's pairwise relatively prime, and $d_1 d_2 \cdots d_k \le R$. In our application to prime gaps we require that $R \le N^{\frac{1}{4}-\varepsilon}$. On the other hand, on expanding, (2.7) becomes a sum over $d_i \mid n+h_i$, $1 \le i \le k$, with $d_1 \le R$, $d_2 \le R$, $\ldots$, $d_k \le R$. The application to prime gaps here requires that $R^k \le N^{\frac{1}{4}-\varepsilon}$, and so $R \le N^{\frac{1}{4k}-\frac{\varepsilon}{k}}$. Thus (2.7) has a more severe restriction on the range of the divisors. An additional technical advantage is that having one truncation rather than k truncations simplifies our calculations.

Our main results on $\Lambda_R(n;\mathcal{H},\ell)$ are summarized in the following two propositions. Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are, respectively, sets of k_1 and k_2 distinct non-negative integers $\le h$. We always assume that at least one of these sets is nonempty. Let $M = k_1 + k_2 + \ell_1 + \ell_2$.

PROPOSITION 1. *Let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$, $|\mathcal{H}_i| = k_i$, and $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$. If $R \ll N^{\frac{1}{2}}(\log N)^{-4M}$ and $h \le R^C$ for any given constant $C > 0$, then as $R, N \to \infty$,*

$$\begin{aligned}\sum_{n\le N} &\Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\\ &= \binom{\ell_1+\ell_2}{\ell_1}\frac{(\log R)^{r+\ell_1+\ell_2}}{(r+\ell_1+\ell_2)!}(\mathfrak{S}(\mathcal{H})+o_M(1))N.\end{aligned} \tag{2.14}$$

PROPOSITION 2. *Let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$, $|\mathcal{H}_i| = k_i$, $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$, $1 \le h_0 \le h$, and $\mathcal{H}^0 = \mathcal{H} \cup \{h_0\}$. If $R \ll_M N^{\frac{1}{4}}(\log N)^{-B(M)}$ for a sufficiently large positive constant $B(M)$, and $h \le R$, then as $R, N \to \infty$,*

$$\begin{aligned}\sum_{n\le N} &\Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\theta(n+h_0)\\ &= \begin{cases} \binom{\ell_1+\ell_2}{\ell_1}\frac{(\log R)^{r+\ell_1+\ell_2}}{(r+\ell_1+\ell_2)!}(\mathfrak{S}(\mathcal{H}^0)+o_M(1))N & \text{if } h_0 \notin \mathcal{H},\\ \binom{\ell_1+\ell_2+1}{\ell_1+1}\frac{(\log R)^{r+\ell_1+\ell_2+1}}{(r+\ell_1+\ell_2+1)!}(\mathfrak{S}(\mathcal{H})+o_M(1))N & \text{if } h_0 \in \mathcal{H}_1\setminus\mathcal{H}_2,\\ \binom{\ell_1+\ell_2+2}{\ell_1+1}\frac{(\log R)^{r+\ell_1+\ell_2+1}}{(r+\ell_1+\ell_2+1)!}(\mathfrak{S}(\mathcal{H})+o_M(1))N & \text{if } h_0 \in \mathcal{H}_1\cap\mathcal{H}_2.\end{cases}\end{aligned} \tag{2.15}$$

With the assumption that the primes have level of distribution $\vartheta > 1/2$, *i.e.* (1.3) *with* (1.4) *holds, the asymptotics in* (2.15) *hold with* $R \ll N^{\frac{\vartheta}{2}-\varepsilon}$ *and* $h \le R^{\varepsilon}$, *for any fixed* $\varepsilon > 0$.

By relabeling the variables, we obtain the corresponding form if $h_0 \in \mathcal{H}_2$, $h_0 \notin \mathcal{H}_1$.

Propositions 1 and 2 can be strengthened in several ways. We will show that the error terms $o_M(1)$ can be replaced by a series of lower order terms and a prime number theorem type of error term. Moreover, we can make the result uniform for $M \to \infty$ as an explicit function of N and R. This will be proved in a later paper and used in the proof of (1.9).

3. **Proofs of Theorems 1 and 2**

In this section we employ Propositions 1 and 2 and a simple argument due to Granville and Soundararajan to prove Theorems 1 and 2.

For $\ell \ge 0$, $\mathcal{H}_k = \{h_1, h_2, \ldots, h_k\}$, $1 \le h_1, h_2, \ldots, h_k \le h \le R$, we deduce from Proposition 1, for $R \ll N^{\frac{1}{2}}(\log N)^{-B(M)}$ and $R, N \to \infty$, that

$$\sum_{n\le N} \Lambda_R(n;\mathcal{H}_k,\ell)^2 \sim \frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell}. \tag{3.1}$$

For any $h_i \in \mathcal{H}_k$, we have from Proposition 2, for $R \ll N^{\frac{\vartheta}{2}-\varepsilon}$, and $R, N \to \infty$,

$$\sum_{n\le N} \Lambda_R(n;\mathcal{H}_k,\ell)^2\theta(n+h_i) \sim \frac{1}{(k+2\ell+1)!}\binom{2\ell+2}{\ell+1}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell+1}. \tag{3.2}$$

Taking $R = N^{\frac{\vartheta}{2}-\varepsilon}$, we obtain[4]

$$\begin{aligned}\mathcal{S} &:= \sum_{n=N+1}^{2N}\left(\sum_{i=1}^{k}\theta(n+h_i) - \log 3N\right)\Lambda_R(n;\mathcal{H}_k,\ell)^2 \\ &\sim \frac{k}{(k+2\ell+1)!}\binom{2\ell+2}{\ell+1}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell+1} \\ &\quad - \log 3N\,\frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell} \\ &\sim \left(\frac{2k}{k+2\ell+1}\,\frac{2\ell+1}{\ell+1}\log R - \log 3N\right)\frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell}.\end{aligned} \tag{3.3}$$

[4]In (3.3), as well as later in (3.8), the asymptotic sign replaces an error term of size $o(\log N)$ in the parenthesis term after $\log 3N$. We thus make the convention that the asymptotic relationship holds only up to the size of the apparent main term.

Here we note that if $\mathcal{S} > 0$ then there exists an $n \in [N+1, 2N]$ such that at least two of the numbers $n+h_1, n+h_2, \ldots, n+h_k$ will be prime. This occurs when

$$\frac{k}{k+2\ell+1}\frac{2\ell+1}{\ell+1}\vartheta > 1. \tag{3.4}$$

If $k, \ell \to \infty$ with $\ell = o(k)$, then the left-hand side has the limit 2ϑ, and thus (3.4) holds for any $\vartheta > 1/2$ if we choose k and ℓ appropriately depending on ϑ. This proves the first part of Theorem 1. Next, assuming $\vartheta > 20/21$, we see that (3.4) holds with $\ell = 1$ and $k = 7$. This proves the second part of Theorem 1 but with $k = 7$. The case $k = 6$ requires a slightly more complicated argument and is treated later in this section.

The table below gives the values of $C(\vartheta)$, defined in Theorem 1, obtained from (3.4). For a certain ϑ, it gives the smallest k and corresponding smallest ℓ for which (3.4) is true. Here $h(k)$ is the shortest length of any admissible k-tuple, which has been computed by Engelsma [6] by exhaustive search for $1 \le k \le 305$ and covers every value in this table and the next except $h(421)$, where we have taken the upper bound value from [6].

ϑ	k	ℓ	$h(k)$
1	7	1	20
0.95	8	1	26
0.90	9	1	30
0.85	11	1	36
0.80	16	1	60
0.75	21	2	84
0.70	31	2	140
0.65	51	3	252
0.60	111	5	634
0.55	421	10	2956*

* indicates that this value could be an upper bound of the true value.

To prove Theorem 2, we modify the previous proof by considering

$$\tilde{\mathcal{S}} := \sum_{n=N+1}^{2N}\Bigg(\sum_{1\le h_0\le h}\theta(n+h_0) - \nu\log 3N\Bigg)\sum_{\substack{1\le h_1,h_2,\ldots,h_k\le h\\ \text{distinct}}}\Lambda_R(n;\mathcal{H}_k,\ell)^2, \tag{3.5}$$

where ν is a positive integer. To evaluate $\tilde{\mathcal{S}}$, we need the case of Proposition 2 where $h_0 \notin \mathcal{H}_k$:

$$\sum_{n\le N}\Lambda_R(n;\mathcal{H}_k,\ell)^2\,\theta(n+h_0) \sim \frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k\cup\{h_0\})N(\log R)^{k+2\ell}. \tag{3.6}$$

We also need a result of Gallagher [10]: as $h \to \infty$,

$$\sum_{\substack{1 \le h_1, h_2, \dots h_k \le h \\ \text{distinct}}} \mathfrak{S}(\mathcal{H}_k) \sim h^k. \tag{3.7}$$

Taking $R = N^{\frac{\vartheta}{2}-\varepsilon}$, and applying (3.1), (3.2), (3.6), and (3.7), we find that

(3.8)

$$\begin{aligned}
\tilde{\mathcal{S}} \sim \sum_{\substack{1 \le h_1, h_2, \dots, h_k \le h \\ \text{distinct}}} \Bigg(& \frac{k}{(k+2\ell+1)!}\binom{2\ell+2}{\ell+1}\mathfrak{S}(\mathcal{H}_k) N (\log R)^{k+2\ell+1} \\
& + \sum_{\substack{1 \le h_0 \le h \\ h_0 \ne h_i, 1 \le i \le k}} \frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k \cup \{h_0\}) N (\log R)^{k+2\ell} \\
& - \nu \log 3N \frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k) N (\log R)^{k+2\ell} \Bigg) \\
\sim \left(\frac{2k}{k+2\ell+1} \frac{2\ell+1}{\ell+1} \log R + h - \nu \log 3N \right) & \frac{1}{(k+2\ell)!}\binom{2\ell}{\ell} N h^k (\log R)^{k+2\ell}.
\end{aligned}$$

Thus, there are at least $\nu + 1$ primes in some interval $(n, n+h]$, $N < n \le 2N$, provided that

$$h > \left(\nu - \frac{2k}{k+2\ell+1} \frac{2\ell+1}{\ell+1} \left(\frac{\vartheta}{2} - \varepsilon \right) \right) \log N, \tag{3.9}$$

which, on letting $\ell = [\sqrt{k}/2]$ and taking k sufficiently large, gives

$$h > \left(\nu - 2\vartheta + 4\varepsilon + O\left(\frac{1}{\sqrt{k}} \right) \right) \log N. \tag{3.10}$$

This proves (1.11). Theorem 2 is the special case $\nu = 1$ and $\vartheta = 1/2$.

We are now ready to prove the last part of Theorem 1. Consider

$$\begin{aligned}
\mathcal{S}' := & \sum_{n=N+1}^{2N} \left(\sum_{i=1}^{k} \theta(n+h_i) - \log 3N \right) \left(\sum_{\ell=0}^{L} a_\ell \Lambda_R(n; \mathcal{H}_k, \ell) \right)^2 \\
= & \sum_{n=N+1}^{2N} \left(\sum_{i=1}^{k} \theta(n+h_i) - \log 3N \right) \\
& \times \sum_{0 \le \ell_1, \ell_2 \le L} a_{\ell_1} a_{\ell_2} \Lambda_R(n; \mathcal{H}_k, \ell_1) \Lambda_R(n; \mathcal{H}_k, \ell_2) \\
= & \sum_{0 \le \ell_1, \ell_2 \le L} a_{\ell_1} a_{\ell_2} \mathcal{M}_{\ell_1, \ell_2},
\end{aligned} \tag{3.11}$$

where

$$\mathcal{M}_{\ell_1,\ell_2} = \widetilde{M}_{\ell_1,\ell_2} - (\log 3N) M_{\ell_1,\ell_2}, \tag{3.12}$$

say. Applying Propositions 1 and 2 with $R = N^{\frac{\vartheta}{2}-\varepsilon}$, we deduce that

$$M_{\ell_1,\ell_2} \sim \binom{\ell_1+\ell_2}{\ell_1} \frac{(\log R)^{k+\ell_1+\ell_2}}{(k+\ell_1+\ell_2)!} \mathfrak{S}(\mathcal{H}_k) N$$

and

$$\widetilde{M}_{\ell_1,\ell_2} \sim k \binom{\ell_1+\ell_2+2}{\ell_1+1} \frac{(\log R)^{k+\ell_1+\ell_2+1}}{(k+\ell_1+\ell_2+1)!} \mathfrak{S}(\mathcal{H}_k) N.$$

Therefore,

$$\begin{aligned}\mathcal{M}_{\ell_1,\ell_2} \sim \binom{\ell_1+\ell_2}{\ell_1} \mathfrak{S}(\mathcal{H}_k) N \frac{(\log R)^{k+\ell_1+\ell_2}}{(k+\ell_1+\ell_2)!} \\ \times \left(\frac{k(\ell_1+\ell_2+2)(\ell_1+\ell_2+1)}{(\ell_1+1)(\ell_2+1)(k+\ell_1+\ell_2+1)} \log R - \log 3N \right).\end{aligned}$$

Defining $b_\ell = (\log R)^\ell a_\ell$ and $\boldsymbol{b}$ to be the column matrix corresponding to the vector $(b_0, b_1, \ldots, b_L)$, we obtain

$$\begin{aligned} S^*(N, \mathcal{H}_k, \vartheta, \boldsymbol{b}) &:= \frac{1}{\mathfrak{S}(\mathcal{H}_k) N (\log R)^{k+1}} S' \\ &\sim \sum_{0 \le \ell_1, \ell_2 \le L} b_{\ell_1} b_{\ell_2} \binom{\ell_1+\ell_2}{\ell_1} \frac{1}{(k+\ell_1+\ell_2)!} \\ &\qquad \times \left(\frac{k(\ell_1+\ell_2+2)(\ell_1+\ell_2+1)}{(\ell_1+1)(\ell_2+1)(k+\ell_1+\ell_2+1)} - \frac{2}{\vartheta} \right) \\ &\sim \boldsymbol{b}^T \boldsymbol{M} \boldsymbol{b}, \end{aligned} \tag{3.13}$$

where

$$\boldsymbol{M} = \left[\binom{i+j}{i} \frac{1}{(k+i+j)!} \left(\frac{k(i+j+2)(i+j+1)}{(i+1)(j+1)(k+i+j+1)} - \frac{2}{\vartheta} \right) \right]_{0 \le i,j \le L}. \tag{3.14}$$

We need to choose $\boldsymbol{b}$ so that $S^* > 0$ for a given ϑ and minimal k. On taking $\boldsymbol{b}$ to be an eigenvector of the matrix $\boldsymbol{M}$ with eigenvalue λ, we see that

$$S^* \sim \boldsymbol{b}^T \lambda \boldsymbol{b} = \lambda \sum_{i=0}^{k} |b_i|^2 \tag{3.15}$$

will be > 0 provided that λ is positive. Therefore $S^* > 0$ if $\boldsymbol{M}$ has a positive eigenvalue and $\boldsymbol{b}$ is chosen to be the corresponding eigenvector. Using *Mathematica* we computed the values of $C(\vartheta)$ indicated in the following table, which may be compared to the earlier table obtained from (3.4).

ϑ	k	L	$h(k)$
1	6	1	16
0.95	7	1	20
0.90	8	2	26
0.85	10	2	32
0.80	12	2	42
0.75	16	2	60
0.70	22	4	90
0.65	35	4	158
0.60	65	6	336
0.55	193	9	1204

In particular, taking $k=6$, $L=1$, $b_0=1$, and $b_1=b$ in (3.13), we get

$$\begin{aligned}S^* &\sim \frac{1}{8!}\left(96-\frac{112}{\vartheta}+2b\left(18-\frac{16}{\vartheta}\right)+b^2\left(4-\frac{4}{\vartheta}\right)\right)\\ &\sim -\frac{4(1-\vartheta)}{8!\,\vartheta}\left(b^2-2b\frac{18\vartheta-16}{4(1-\vartheta)}-\frac{96\vartheta-112}{4(1-\vartheta)}\right)\\ &\sim -\frac{4(1-\vartheta)}{8!\,\vartheta}\left(\left(b-\frac{18\vartheta-16}{4(1-\vartheta)}\right)^2+\frac{15\vartheta^2-64\vartheta+48}{4(1-\vartheta)^2}\right).\end{aligned}$$

Choosing $b=\frac{18\vartheta-16}{4(1-\vartheta)}$, we then have

$$S^* \sim -\frac{15\vartheta^2-64\vartheta+48}{8!\,\vartheta(1-\vartheta)},$$

of which the right-hand side is >0 if $\vartheta\le 1$ lies between the two roots of the quadratic; this occurs when $4(8-\sqrt{19})/15<\vartheta\le 1$. Thus, there are at least two primes in any admissible tuple $\mathcal{H}_k$ for $k=6$, if

$$\vartheta > \frac{4(8-\sqrt{19})}{15} = 0.97096\ldots\ . \tag{3.16}$$

This completes the proof of Theorem 1. ∎

4. Further remarks on Section 3

We can formulate the method of Section 3 as follows. For a given tuple $\mathcal{H}=\{h_1,h_2,\ldots,h_k\}$ we define

$$Q_1 := \sum_{n=N+1}^{2N} f_R(n;\mathcal{H})^2, \quad Q_2 := \sum_{n=N+1}^{2N}\left(\sum_{i=1}^{k}\theta(n+h_i)\right) f_R(n;\mathcal{H})^2, \tag{4.1}$$

where f should be chosen to make Q_2 large compared with Q_1, and $R = R(N)$ will be chosen later. It is reasonable to assume

$$f_R(n;\mathcal{H}) = \sum_{\substack{d\,|\,P_{\mathcal{H}}(n)\\ d\le R}} \lambda_{d,R}. \tag{4.2}$$

Our goal is to select the $\lambda_{d,R}$ which maximizes

$$\rho = \rho(N;\mathcal{H}, f) := \frac{1}{\log 3N}\left(\frac{Q_2}{Q_1}\right) \tag{4.3}$$

for the purpose of obtaining a good lower bound for ρ. If $\rho > \nu$ for some N and positive integer ν, then there exists an n, $N < n \le 2N$, such that the tuple (1.5) has at least $\nu + 1$ prime components.

This method has much in common with the method introduced for twin primes by Selberg and for general tuples by Heath-Brown. However, they used the divisor function $d(n + h_i)$ in Q_2 in place of $\theta(n + h_i)$ and sought to minimize (4.3) to obtain a good upper bound for ρ. Heath-Brown even chose $f = \Lambda_R(n;\mathcal{H}, 1)$.

As a first example, suppose we choose f as in (2.6) and (2.7), so that

$$f_R(n;\mathcal{H}) = \prod_{i=1}^{k} \Lambda_R(n + h_i). \tag{4.4}$$

By [13], we have, as $R, N \to \infty$,[5]

$$\begin{aligned} Q_1 &\sim N\mathfrak{S}(\mathcal{H})(\log R)^k &&\text{if } R \le N^{\frac{1}{2k}(1-\varepsilon)}, \\ Q_2 &\sim kN\mathfrak{S}(\mathcal{H})(\log R)^{k+1} &&\text{if } R \le N^{\frac{\vartheta}{2k}(1-\varepsilon)}. \end{aligned} \tag{4.5}$$

On taking $R = N^{\frac{\vartheta_0}{2k}}$, $0 < \vartheta_0 < \vartheta$, we see that, as $N \to \infty$,

$$\rho \sim k\frac{\log R}{\log N} \sim \frac{\vartheta_0}{2}. \tag{4.6}$$

Notice that $\rho < 1$, so that we fail to detect primes in tuples. In Section 3, we proved that on choosing $f = \Lambda_R(n;\mathcal{H}, \ell)$, by (3.1) and (3.2), as $N \to \infty$,

$$\rho \sim \frac{k}{k + 2\ell + 1}\,\frac{2\ell + 1}{\ell + 1}\vartheta_0. \tag{4.7}$$

If $\ell = 0$ this gives $\rho \sim \frac{k}{k+1}\vartheta_0$, which, for large k, is twice as large as (4.6), while (4.7) gains another factor of 2 when $\ell \to \infty$ slowly as $k \to \infty$. This finally shows $\rho > 1$ if $\vartheta > 1/2$, but just fails if $\vartheta = 1/2$.

[5]For special reasons, the validity of the formula for Q_2 actually holds here for $R \le N^{\frac{\vartheta}{2(k-1)}(1-\varepsilon)}$ if $k \ge 2$, but this is insignificant for the present discussion.

In (3.11) we chose

$$f_R(n;\mathcal{H}) = \sum_{\ell=0}^{L} \frac{b_\ell}{(\log R)^\ell} \Lambda_R(n;\mathcal{H}_k,\ell) = \sum_{\substack{d \mid P_{\mathcal{H}}(n) \\ d \le R}} \mu(d) P\left(\frac{\log(R/d)}{\log R}\right) \tag{4.8}$$

where P is a polynomial with a k-th order zero at 0. The matrix procedure does not provide a method for analyzing ρ unless L is taken fixed, but the general problem has been solved by Soundararajan [31]. In particular, he showed that $\rho < 1$ if $\vartheta = 1/2$, so that one can not prove there are bounded gaps between primes using (4.8). The exact solution from Soundrarajan's analysis was obtained by a calculus-of-variations argument by Conrey, which gives, as $N \to \infty$,

$$\rho = \frac{k(k-1)}{2\beta}\vartheta_0, \tag{4.9}$$

where β is determined as the solution of the equation

$$\frac{1}{\beta} = \frac{\int_0^1 y^{k-2} q(y)^2\, dy}{\int_0^1 y^{k-1} q'(y)^2\, dy} \quad \text{with } q(y) = J_{k-2}(2\sqrt{\beta}) - y^{1-\frac{k}{2}} J_{k-2}(2\sqrt{\beta y}), \tag{4.10}$$

where J_k is the Bessel function of the first type. Using *Mathematica*, one can check that this gives exactly the values of k in the previous table, which is in agreement with our earlier calculations; but it provides somewhat smaller values of ϑ for which a given k-tuple will contain two primes. Thus, for example, we can replace (3.16) by the result that every admissible 6-tuple will contain at least two primes if

$$\vartheta > .95971\ldots. \tag{4.11}$$

5. **Two lemmas**

In this section we will prove two lemmas needed for the proof of Propositions 1 and 2. The conditions on these lemmas have been constructed in order for them to hold uniformly in the given variables.

The Riemann zeta-function has the Euler product representation, with $s = \sigma + it$,

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}, \quad \sigma > 1. \tag{5.1}$$

The zeta-function is analytic except for a simple pole at $s = 1$, where as $s \to 1$

$$\zeta(s) = \frac{1}{s-1} + \gamma + O(|s-1|). \tag{5.2}$$

(Here γ is Euler's constant.) We need standard information concerning the classical zero-free region of the Riemann zeta-function. By Theorem 3.11 and (3.11.8) in

[32], there exists a small constant $\overline{c} > 0$, for which we assume $\overline{c} \le 10^{-2}$, such that $\zeta(\sigma+it) \ne 0$ in the region

$$\sigma \ge 1 - \frac{4\overline{c}}{\log(|t|+3)} \tag{5.3}$$

for all t. Furthermore, we have

$$\zeta(\sigma+it) - \frac{1}{\sigma-1+it} \ll \log(|t|+3), \quad \frac{1}{\zeta(\sigma+it)} \ll \log(|t|+3), \tag{5.4}$$
$$\frac{\zeta'}{\zeta}(\sigma+it) + \frac{1}{\sigma-1+it} \ll \log(|t|+3),$$

in this region. We will fix this $\overline{c}$ for the rest of the paper (we could take, for instance, $\overline{c} = 10^{-2}$, see [8]). Let $\mathscr{L}$ denote the contour given by

$$s = -\frac{\overline{c}}{\log(|t|+3)} + it. \tag{5.5}$$

LEMMA 1. *For $R \ge C$, $k \ge 2$, $B \le Ck$,*

$$\int_{\mathscr{L}} (\log(|s|+3))^B \left| \frac{R^s}{s^k}\, ds \right| \ll C_1^k R^{-c_2} + e^{-\sqrt{\overline{c}\log R}/2}, \tag{5.6}$$

where C_1, c_2 and the implied constant in $\ll$ depends only on the constant C in the formulation of the lemma. In addition, if $k \le c_3 \log R$ with a sufficiently small c_3 depending only on C, then

$$\int_{\mathscr{L}} (\log(|s|+3))^B \left| \frac{R^s}{s^k}\, ds \right| \ll e^{-\sqrt{\overline{c}\log R}/2}. \tag{5.7}$$

Proof. The left-hand side of (5.6) is, with C_4 depending on C,

$$\begin{aligned} &\ll \int_0^\infty R^{\sigma(t)} \frac{(\log(|t|+4))^B}{(|t|+\overline{c}/2)^k}\, dt \\ &\ll \int_0^{C_4} C_1^k R^{-c_2} dt + \int_{C_4}^{\omega-3} \frac{R^{-\frac{\overline{c}}{\log(|t|+3)}}}{t^{3/2}} dt + \int_{\omega-3}^\infty t^{-3/2} dt \\ &\ll C_1^k R^{-c_2} + e^{-\frac{\overline{c}\log R}{\log\omega}} + \omega^{-\frac{1}{2}}, \end{aligned} \tag{5.8}$$

where now C_1 is a constant depending on C. On choosing $\log\omega = \sqrt{\overline{c}\log R}$, the first part of the lemma follows. The second part is an immediate consequence of the first part. □

The next lemma provides some explicit estimates for sums of the generalized divisor function. Let $\omega(q)$ denote the number of prime factors of a squarefree integer q. For any real number m, we define

$$d_m(q) = m^{\omega(q)}. \tag{5.9}$$

This agrees with the usual definition of the divisor functions when m is a positive integer. Clearly, $d_m(q)$ is a monotonically increasing function of m (for a fixed q), and for real m_1, m_2, and y, we see that

$$d_{m_1}(q)d_{m_2}(q) = d_{m_1 m_2}(q), \quad (d_m(q))^y = d_{m^y}(q). \tag{5.10}$$

Recall that $\sum^\flat$ indicates a sum over squarefree integers. We use the ceiling function $\lceil y \rceil := \min\{n \in \mathbb{Z}; y \le n\}$.

LEMMA 2. *For any positive real m and $x \ge 1$ we have*

$$D'(x,m) := \sum_{q \le x}^{\flat} \frac{d_m(q)}{q} \le (\lceil m \rceil + \log x)^{\lceil m \rceil} \le (m + 1 + \log x)^{m+1}, \tag{5.11}$$

$$D^*(x,m) := \sum_{q \le x}^{\flat} d_m(q) \le x(\lceil m \rceil + \log x)^{\lceil m \rceil} \le x(m + 1 + \log x)^{m+1}. \tag{5.12}$$

Proof. First, we treat the case when m is a positive integer. We prove (5.11) by induction. Observe that the assertion is true for $m = 1$, that is, when $d_1(q) = 1$ by definition. Suppose (5.11) is proved for $m - 1$. Let us denote the smallest term in a given product representation of q by $j = j(q) \le x^{1/m}$. Then this factor can stand at m places, and, therefore, with $q = q' j(q) = q' j$,

$$\begin{aligned} \sum_{q \le x}^{\flat} \frac{d_m(q)}{q} &\le m \sum_{j=1}^{x^{1/m}}{}^{\flat} \frac{1}{j} \sum_{q' \le x/j}^{\flat} \frac{d_{m-1}(q')}{q'} \le m(1 + \log x^{\frac{1}{m}})\,(m - 1 + \log x)^{m-1} \\ &\le (m + \log x)(m + \log x)^{m-1} = (m + \log x)^m. \end{aligned}$$

This completes the induction. For real m, the result holds since $D'(x,m) \le D'(x, \lceil m \rceil)$. We note that (5.12) follows from (5.11) because $D^*(x,m) \le x D'(x,m)$. □

6. A special case of Proposition 1

In this section we prove a special case of Proposition 1 which illustrates the method without involving the technical complications that appear in the general case. This allows us to set up some notation and obtain estimates for use in the general case. We also obtain the result uniformly in k.

Assume $\mathcal{H}$ is nonempty (so that $k \ge 1$), $\ell = 0$, and $\Lambda_R(n; \mathcal{H}, 0) = \Lambda_R(n; \mathcal{H})$.

PROPOSITION 3. *Suppose*

$$k \ll_{\eta_0} (\log R)^{\frac{1}{2} - \eta_0} \text{ with an arbitrarily small fixed } \eta_0 > 0, \tag{6.1}$$

and $h \le R^C$, *with* C *any fixed positive number*; *then*

$$\sum_{n=1}^{N} \Lambda_R(n;\mathcal{H}) = \mathfrak{S}(\mathcal{H})N + O(Ne^{-c\sqrt{\log R}}) + O\big(R(2\log R)^{2k}\big). \tag{6.2}$$

This result motivates the conjecture (2.11).

Proof. We have

$$\mathcal{S}_R(N;\mathcal{H}) := \sum_{n=1}^{N} \Lambda_R(n;\mathcal{H}) = \frac{1}{k!}\sum_{d\le R}\mu(d)\left(\log\frac{R}{d}\right)^k \sum_{\substack{1\le n\le N\\ d\,|\,P_{\mathcal{H}}(n)}} 1. \tag{6.3}$$

If for a prime p we have $p|P_{\mathcal{H}}(n)$, then among the solutions $n \equiv -h_i \pmod p$, $1 \le i \le k$, there will be $\nu_p(\mathcal{H})$ distinct solutions modulo p. For d squarefree we then have by multiplicativity $\nu_d(\mathcal{H})$ distinct solutions for n modulo d which satisfy $d|P_{\mathcal{H}}(n)$, and for each solution, n runs through a residue class modulo d. Hence we see that

$$\sum_{\substack{1\le n\le N\\ d\,|\,P_{\mathcal{H}}(n)}} 1 = \nu_d(\mathcal{H})\left(\frac{N}{d} + O(1)\right). \tag{6.4}$$

Trivially, $\nu_q(\mathcal{H}) \le k^{\omega(q)} = d_k(q)$ for squarefree q. Therefore, we conclude that

(6.5)

$$\begin{aligned}\mathcal{S}_R(N;\mathcal{H}) &= N\left(\frac{1}{k!}\sum_{d\le R}\frac{\mu(d)\nu_d(\mathcal{H})}{d}\left(\log\frac{R}{d}\right)^k\right) + O\left(\frac{(\log R)^k}{k!}\sideset{}{^\flat}\sum_{d\le R}\nu_d(\mathcal{H})\right)\\ &= N\mathcal{T}_R(\mathcal{H}) + O\big(R(k+\log R)^{2k}\big),\end{aligned}$$

by Lemma 2.

Let (a) denote the contour $s = a + it$, $-\infty < t < \infty$. We apply the formula

$$\frac{1}{2\pi i}\int_{(c)} \frac{x^s}{s^{k+1}}\,ds = \begin{cases} 0 & \text{if } 0 < x \le 1,\\ \frac{1}{k!}(\log x)^k & \text{if } x \ge 1,\end{cases} \tag{6.6}$$

for $c > 0$, and have that

$$\mathcal{T}_R(\mathcal{H}) = \frac{1}{2\pi i}\int_{(1)} F(s)\frac{R^s}{s^{k+1}}ds, \tag{6.7}$$

where, letting $s = \sigma + it$ and assuming $\sigma > 0$,

$$F(s) = \sum_{d=1}^{\infty}\frac{\mu(d)\nu_d(\mathcal{H})}{d^{1+s}} = \prod_p\left(1 - \frac{\nu_p(\mathcal{H})}{p^{1+s}}\right). \tag{6.8}$$

Since $\nu_p(\mathcal{H}) = k$ for all $p > h$,

$$F(s) = \frac{G_{\mathcal{H}}(s)}{\zeta(1+s)^k}, \tag{6.9}$$

where by (5.1)

$$\begin{aligned} G_{\mathcal{H}}(s) &= \prod_p \left(1 - \frac{\nu_p(\mathcal{H})}{p^{1+s}}\right)\left(1 - \frac{1}{p^{1+s}}\right)^{-k} \\ &= \prod_p \left(1 + \frac{k - \nu_p(\mathcal{H})}{p^{1+s}} + O_h\left(\frac{k^2}{p^{2+2\sigma}}\right)\right), \end{aligned} \tag{6.10}$$

which is analytic and uniformly bounded for $\sigma > -1/2 + \delta$ for any $\delta > 0$. Also, by (2.2) we see that

$$G_{\mathcal{H}}(0) = \mathfrak{S}(\mathcal{H}). \tag{6.11}$$

From (5.4) and (6.9), the function $F(s)$ satisfies the bound

$$F(s) \ll |G_{\mathcal{H}}(s)|(C \log(|t| + 3))^k \tag{6.12}$$

in the region on and to the right of $\mathcal{L}$. Here $G_{\mathcal{H}}(s)$ is analytic and bounded in this region, and has a dependence on both k and the size h of the components of $\mathcal{H}$. We note that $\nu_p(\mathcal{H}) = k$ not only when $p > h$, but whenever $p \nmid \Delta$, where

$$\Delta := \prod_{1 \le i < j \le k} |h_j - h_i|, \tag{6.13}$$

since then all k of the h_i's are distinct modulo p. We now introduce an important parameter U that is used throughout the rest of the paper. We want U to be an upper bound for $\log \Delta$, and since trivially $\Delta \le h^{k^2}$ we choose

$$U := Ck^2 \log(2h) \tag{6.14}$$

and have

$$\log \Delta \le U. \tag{6.15}$$

We now prove, for $-1/4 < \sigma \le 1$,

$$|G_{\mathcal{H}}(s)| \ll \exp(5kU^{\delta} \log\log U), \quad \text{where } \delta = \max(-\sigma, 0). \tag{6.16}$$

We treat separately the different pieces of the product defining $G_{\mathcal{H}}$. First, by use of the inequality $\log(1 + x) \le x$ for $x \ge 0$, we have

$$\left|\prod_{p\le U}\left(1-\frac{\nu_p(\mathcal{H})}{p^{1+s}}\right)\right|\le\prod_{p\le U}\left(1+\frac{k}{p^{1-\delta}}\right)$$
$$=\exp\left(\sum_{p\le U}\log\left(1+\frac{k}{p^{1-\delta}}\right)\right)\le\exp\left(\sum_{p\le U}\frac{k}{p^{1-\delta}}\right)$$
$$\le\exp\left(kU^{\delta}\sum_{p\le U}\frac{1}{p}\right)\ll\exp\left(kU^{\delta}\log\log U\right).$$

Second, by the same estimates and the inequality $(1-x)^{-1}\le 1+3x$ for $0\le x\le 2/3$, we see that

$$\left|\prod_{p\le U}\left(1-\frac{1}{p^{1+s}}\right)^{-k}\right|\le\left(\prod_{p\le U}\left(1-\frac{1}{p^{1-\delta}}\right)^{-1}\right)^{k}$$
$$\le\left(\prod_{p\le U}\left(1+\frac{3}{p^{1-\delta}}\right)\right)^{k}\quad\left(\text{since }\frac{1}{p^{1-\delta}}\le\frac{1}{2^{3/4}}<\frac{2}{3}\right),$$
$$\ll\exp\left(3kU^{\delta}\log\log U\right).$$

Hence, the terms in the product for $G_{\mathcal{H}}(s)$ with $p\le U$ are

$$\ll\exp\big(4kU^{\delta}\log\log U\big).$$

For the terms $p>U$, we first consider those for which $p|\Delta$. In absolute value, they are

$$\le\prod_{\substack{p|\Delta\\ p>U}}\left(1+\frac{k}{p^{1-\delta}}\right)\left(1+\frac{3}{p^{1-\delta}}\right)^{k}\le\exp\Big(\sum_{\substack{p|\Delta\\ p>U}}\frac{4k}{p^{1-\delta}}\Big).$$

Since there are fewer than $(1+o(1))\log\Delta<U$ primes with $p|\Delta$, the sum above is increased if we replace these terms with the integers between U and $2U$. Therefore the right-hand side above is

$$\le\exp\left(4k\sum_{U<n\le 2U}\frac{1}{n^{1-\delta}}\right)\le\exp\left(4k(2U)^{\delta}\sum_{U<n\le 2U}\frac{1}{n}\right)\le\exp(4kU^{\delta}).$$

Finally, if $p>U$

$$\left|\frac{k}{p^{1+s}}\right|\le\frac{k}{U^{1-\delta}}\le\frac{1}{2},$$

so that in absolute value the terms with $p > U$ and $p \nmid \Delta$ are

$$
\begin{aligned}
&= \left| \prod_{\substack{p \nmid \Delta \\ p > U}} \left(1 - \frac{k}{p^{1+s}}\right)\left(1 - \frac{1}{p^{1+s}}\right)^{-k} \right| \\
&= \left| \exp\left(\sum_{\substack{p \nmid \Delta \\ p > U}} \left(-\sum_{\nu=1}^{\infty} \frac{1}{\nu}\Big(\frac{k}{p^{1+s}}\Big)^{\nu} + k \sum_{\nu=1}^{\infty} \frac{1}{\nu}\Big(\frac{1}{p^{1+s}}\Big)^{\nu} \right)\right)\right| \\
&\le \exp\left(\sum_{p > U} \sum_{\nu=2}^{\infty} \frac{2}{\nu}\Big(\frac{k}{p^{1-\delta}}\Big)^{\nu}\right) \le \exp\left(\sum_{p > U} \sum_{\nu=2}^{\infty} \Big(\frac{k}{p^{1-\delta}}\Big)^{\nu}\right) \\
&\le \exp\left(2k^2 \sum_{n > U} \frac{1}{n^{2-2\delta}}\right) \le \exp\left(\frac{4k^2 U^{\delta}}{U^{1-\delta}}\right) \le \exp\left(2kU^{\delta}\right).
\end{aligned}
$$

Thus, the terms with $p > U$ contribute $\le \exp\left(6kU^{\delta}\right)$, from which we obtain (6.16).

In conclusion, for $h \ll R^C$ (where $C > 0$ is fixed and as large as we wish) and for s on or to the right of $\mathcal{L}$, we have

$$F(s) \ll (C \log(|t| + 3))^k \exp(5kU^{\delta} \log\log U). \tag{6.17}$$

Returning to the integral in (6.7), we see that the integrand vanishes as $|t| \to \infty$, $-1/4 < \sigma \le 1$. By (6.9) we see that in moving the contour from (1) to the left to $\mathcal{L}$ we either pass through a simple pole at $s = 0$ when $\mathcal{H}$ is admissible (so that $\mathfrak{S}(\mathcal{H}) \ne 0$), or we pass through a regular point at $s = 0$ when $\mathcal{H}$ is not admissible. In either case, we have by virtue of (5.2), (6.11), (6.14), (6.17), and Lemma 1, for any k satisfying (6.1),

$$\mathcal{T}_R(\mathcal{H}) = G_{\mathcal{H}}(0) + \frac{1}{2\pi i} \int_{\mathcal{L}} F(s) \frac{R^s}{s^{k+1}} ds = \mathfrak{S}(\mathcal{H}) + O(e^{-c\sqrt{\log R}}). \tag{6.18}$$

Equation (6.2) now follows from this and (6.5). □

Remark. The exponent $1/2$ in the restriction $k \ll (\log R)^{1/2-\eta_0}$ is not significant. Using Vinogradov's zero-free region for $\zeta(s)$ we could replace $1/2$ by $3/5$.

7. First part of the proof of Proposition 1

Let

$$\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2,\ |\mathcal{H}_1| = k_1,\ |\mathcal{H}_2| = k_2,\ k = k_1 + k_2, \quad r = |\mathcal{H}_1 \cap \mathcal{H}_2|,\ M = k_1 + k_2 + \ell_1 + \ell_2. \tag{7.1}$$

Thus $|\mathcal{H}| = k - r$. We prove Proposition 1 in the following sharper form.

PROPOSITION 4. *Let $h \ll R^C$, where C is any positive fixed constant. As $R, N \to \infty$, we have*

$$\sum_{n\le N} \Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2) = \binom{\ell_1+\ell_2}{\ell_1}\frac{(\log R)^{r+\ell_1+\ell_2}}{(r+\ell_1+\ell_2)!}\mathfrak{S}(\mathcal{H})N + N\sum_{j=1}^{r+\ell_1+\ell_2}\mathcal{D}_j(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2)(\log R)^{r+\ell_1+\ell_2-j} + O_M(Ne^{-c\sqrt{\log R}}) + O(R^2(3\log R)^{3k+M}), \tag{7.2}$$

where the $\mathcal{D}_j(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2)$'s are functions independent of R and N which satisfy the bound

$$\mathcal{D}_j(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2) \ll_M (\log U)^{C_j} \ll_M (\log\log 10h)^{C_j'} \tag{7.3}$$

where U is as defined in (6.14) and C_j and C_j' are two positive constants depending on M.

Proof. We can assume that both $\mathcal{H}_1$ and $\mathcal{H}_2$ are nonempty since the case where one of these sets is empty can be covered in the same way we did in the case of $\ell = 0$ in Section 6. Thus $k \ge 2$ and we have

(7.4)

$$\mathcal{S}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2) := \sum_{n=1}^{N}\Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)$$

$$= \frac{1}{(k_1+\ell_1)!(k_2+\ell_2)!}\sum_{d,e\le R}\mu(d)\mu(e)\left(\log\frac{R}{d}\right)^{k_1+\ell_1}\left(\log\frac{R}{e}\right)^{k_2+\ell_2}\sum_{\substack{1\le n\le N\\ d|P_{\mathcal{H}_1}(n)\\ e|P_{\mathcal{H}_2}(n)}} 1.$$

For the inner sum, we let $d = a_1a_{12}$, $e = a_2a_{12}$ where $(d,e) = a_{12}$. Thus a_1, a_2, and a_{12} are pairwise relatively prime, and the divisibility conditions $d|P_{\mathcal{H}_1}(n)$ and $e|P_{\mathcal{H}_2}(n)$ become $a_1|P_{\mathcal{H}_1}(n)$, $a_2|P_{\mathcal{H}_2}(n)$, $a_{12}|P_{\mathcal{H}_1}(n)$, and $a_{12}|P_{\mathcal{H}_2}(n)$. As in Section 6, we get $\nu_{a_1}(\mathcal{H}_1)$ solutions for n modulo a_1, and $\nu_{a_2}(\mathcal{H}_2)$ solutions for n modulo a_2. If $p|a_{12}$, then from the two divisibility conditions we have $\nu_p(\mathcal{H}_1(p)\cap\mathcal{H}_2(p))$ solutions for n modulo p, where

$$\mathcal{H}(p) = \{h'_1,\dots,h'_{\nu_p(\mathcal{H})} : h'_j \equiv h_i \in \mathcal{H} \text{ for some } i, 1\le h'_j\le p\}.$$

Notice that $\mathcal{H}(p) = \mathcal{H}$ if $p > h$. Alternatively, we can avoid this definition which is necessary only for small primes by defining

$$\overline{\nu}_p(\mathcal{H}_1\overline{\cap}\mathcal{H}_2) := \nu_p(\mathcal{H}_1(p)\cap\mathcal{H}_2(p)) := \nu_p(\mathcal{H}_1)+\nu_p(\mathcal{H}_2)-\nu_p(\mathcal{H}) \tag{7.5}$$

and then extending this definition to squarefree numbers by multiplicativity.[6] Thus we see that

$$\sum_{\substack{1\le n\le N\\ d\mid P_{\mathcal{H}_1}(n)\\ e\mid P_{\mathcal{H}_2}(n)}} 1 = \nu_{a_1}(\mathcal{H}_1)\nu_{a_2}(\mathcal{H}_2)\overline{\nu}_{a_{12}}(\mathcal{H}_1\overline{\cap}\,\mathcal{H}_2)\left(\frac{N}{a_1a_2a_{12}} + O(1)\right),$$

and have

(7.6)

$$\begin{aligned}
&\mathcal{S}_R(N;\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2) = \frac{N}{(k_1+\ell_1)!(k_2+\ell_2)!}\\
&\quad\times \sum_{\substack{a_1a_{12}\le R\\ a_2a_{12}\le R}}{}^{'} \frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2\nu_{a_1}(\mathcal{H}_1)\nu_{a_2}(\mathcal{H}_2)\overline{\nu}_{a_{12}}(\mathcal{H}_1\overline{\cap}\,\mathcal{H}_2)}{a_1a_2a_{12}}\\
&\quad\times\left(\log\frac{R}{a_1a_{12}}\right)^{k_1+\ell_1}\left(\log\frac{R}{a_2a_{12}}\right)^{k_2+\ell_2}\\
&\quad+O\left((\log R)^M \sum_{\substack{a_1a_{12}\le R\\ a_2a_{12}\le R}}{}^{'} \mu(a_1)^2\mu(a_2)^2\mu(a_{12})^2\nu_{a_1}(\mathcal{H}_1)\nu_{a_2}(\mathcal{H}_2)\overline{\nu}_{a_{12}}(\mathcal{H}_1\overline{\cap}\,\mathcal{H}_2)\right)\\
&\quad= N\mathcal{T}_R(\ell_1,\ell_2;\mathcal{H}_1,\mathcal{H}_2) + O(R^2(3\log R)^{3k+M}),
\end{aligned}$$

where $\sum'$ indicates the summands are pairwise relatively prime. Notice that by Lemma 2, the error term was bounded by

$$\begin{aligned}
&\ll (\log R)^M \sum_{q\le R^2}^{\flat}\sum_{q=a_1a_2a_{12}} d_k(q) = (\log R)^M \sum_{q\le R^2}^{\flat} d_3(q)d_k(q)\\
&= (\log R)^M \sum_{q\le R^2}^{\flat} d_{3k}(q) \ll R^2(3\log R)^{3k+M}.
\end{aligned}$$

By (6.6), we have

(7.7)

$$\mathcal{T}_R(\ell_1,\ell_2;\mathcal{H}_1,\mathcal{H}_2) = \frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)} F(s_1,s_2)\frac{R^{s_1}}{{s_1}^{k_1+\ell_1+1}}\frac{R^{s_2}}{{s_2}^{k_2+\ell_2+1}}\,ds_1\,ds_2,$$

where, by letting $s_j = \sigma_j + it_j$ and assuming $\sigma_1,\sigma_2 > 0$,

[6]We are establishing a convention here that for $\overline{\nu}_p$ we take intersections modulo p.

(7.8)

$$F(s_1,s_2)=\sum_{1\le a_1,a_2,a_{12}<\infty}\frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2\nu_{a_1}(\mathcal{H}_1)\nu_{a_2}(\mathcal{H}_2)\overline{\nu}_{a_{12}}(\mathcal{H}_1\overline{\cap}\mathcal{H}_2)}{{a_1}^{1+s_1}{a_2}^{1+s_2}{a_{12}}^{1+s_1+s_2}}$$
$$=\prod_p\left(1-\frac{\nu_p(\mathcal{H}_1)}{p^{1+s_1}}-\frac{\nu_p(\mathcal{H}_2)}{p^{1+s_2}}+\frac{\overline{\nu}_p(\mathcal{H}_1\overline{\cap}\mathcal{H}_2)}{p^{1+s_1+s_2}}\right).$$

Since for all $p>h$ we have $\nu_p(\mathcal{H}_1)=k_1$, $\nu_p(\mathcal{H}_2)=k_2$, and $\nu_p(\mathcal{H}_1\cap\mathcal{H}_2)=r$, we factor out the dominant zeta-factors and write

$$F(s_1,s_2)=G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)\frac{\zeta(1+s_1+s_2)^r}{\zeta(1+s_1)^{k_1}\zeta(1+s_2)^{k_2}}, \tag{7.9}$$

where by (5.1)

(7.10)

$$G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)=\prod_p\left(\frac{\left(1-\frac{\nu_p(\mathcal{H}_1)}{p^{1+s_1}}-\frac{\nu_p(\mathcal{H}_2)}{p^{1+s_2}}+\frac{\overline{\nu}_p(\mathcal{H}_1\overline{\cap}\mathcal{H}_2)}{p^{1+s_1+s_2}}\right)\left(1-\frac{1}{p^{1+s_1+s_2}}\right)^r}{\left(1-\frac{1}{p^{1+s_1}}\right)^{k_1}\left(1-\frac{1}{p^{1+s_2}}\right)^{k_2}}\right)$$

is analytic and uniformly bounded for $\sigma_1,\sigma_2>-1/4+\delta$, for any fixed $\delta>0$. Also, from (2.2), (7.1), and (7.5) we see immediately that

$$G_{\mathcal{H}_1,\mathcal{H}_2}(0,0)=\mathfrak{S}(\mathcal{H}). \tag{7.11}$$

Furthermore, the same argument leading to (6.16) shows that for s_1, s_2 on $\mathcal{L}$ or to the right of $\mathcal{L}$

$$G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)\ll\exp(CkU^{\delta_1+\delta_2}\log\log U), \tag{7.12}$$

with $\delta_i=-\min(\sigma_i,0)$ and U as defined in (6.14). We define

$$W(s):=s\zeta(1+s) \tag{7.13}$$

and

$$D(s_1,s_2)=G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)\frac{W(s_1+s_2)^r}{W(s_1)^{k_1}W(s_2)^{k_2}}, \tag{7.14}$$

so that

(7.15)

$$\mathcal{T}_R(\ell_1,\ell_2;\mathcal{H}_1,\mathcal{H}_2)=\frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)}D(s_1,s_2)\frac{R^{s_1+s_2}}{{s_1}^{\ell_1+1}{s_2}^{\ell_2+1}(s_1+s_2)^r}\,ds_1ds_2.$$

To complete the proof of Proposition 1, we need to evaluate this integral. We will also need to evaluate a similar integral in the proof of Proposition 2, where the parameters k_1, k_2, and r have several slightly different relationships with $\mathcal{H}_1$ and $\mathcal{H}_2$, and G is slightly altered. Therefore we change notation to handle these situations simultaneously.

8. Completion of the proof of Proposition 1: Evaluating an integral

Let

$$\mathcal{T}_R^*(a,b,d,u,v,h) := \frac{1}{(2\pi i)^2} \int_{(1)}\int_{(1)} \frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\, ds_1\, ds_2, \tag{8.1}$$

where

$$D(s_1,s_2) = \frac{G(s_1,s_2)W^d(s_1+s_2)}{W^a(s_1)W^b(s_2)} \tag{8.2}$$

and W is from (7.13). We assume $G(s_1,s_2)$ is regular on $\mathscr{L}$ and to the right of $\mathscr{L}$ and satisfies the bound, with $\delta_i = -\min(\sigma_i, 0)$,

$$G(s_1,s_2) \ll_M \exp(CMU^{\delta_1+\delta_2}\log\log U), \quad \text{where } U = CM^2\log(2h). \tag{8.3}$$

LEMMA 3. *Suppose that*

$$0 \le a,b,d,u,v \le M, \quad a+u \ge 1, \quad b+v \ge 1, \quad d \le \min(a,b), \tag{8.4}$$

where M is a large constant and our estimates may depend on M. Let $h \ll R^C$, with C any positive fixed constant. Then we have, as $R \to \infty$,

$$\mathcal{T}_R^*(a,b,d,u,v,h) = \binom{u+v}{u}\frac{(\log R)^{u+v+d}}{(u+v+d)!}G(0,0) + \sum_{j=1}^{u+v+d} \mathfrak{D}_j(a,b,d,u,v,h)(\log R)^{u+v+d-j} + O_M(e^{-c\sqrt{\log R}}), \tag{8.5}$$

where the $\mathfrak{D}_j(a,b,d,u,v,h)$'s are functions independent of R which satisfy the bound

$$\mathfrak{D}_j(a,b,d,u,v,h) \ll_M (\log U)^{C_j} \ll_M (\log\log 10h)^{C_j'} \tag{8.6}$$

for some positive constants C_j, C_j' depending on M.

Proof. One would expect to proceed exactly as in Section 6 by moving both contours to the left to $\mathscr{L}$. There is, however, a complication because the integrand now contains the function $\zeta(1+s_1+s_2)$ which necessitates also that s_1+s_2 be restricted to the region to the right of $\mathscr{L}$ if we wish to use the bounds in (5.4).[7] By

[7]This was pointed out to us by J. Sivak and also Y. Motohashi and was handled in similar ways in [30] and in [15]; we have also adopted this approach here.

the conditions of Lemma 3, (5.4), and (8.3), we have

$$(8.7)\qquad \frac{D(s_1,s_2)}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}$$
$$\ll_M \frac{\exp(CMU^{\delta_1+\delta_2}\log\log U)\big(\log(|t_1|+3)\log(|t_2|+3)\big)^{2M}\max(1,|s_1+s_2|^{-d})}{|s_1|^{a+u+1}|s_2|^{b+v+1}}$$

provided s_1, s_2, and s_1+s_2 are on or to the right of $\mathscr{L}$. We next let

$$(8.8)\qquad V = e^{\sqrt{\log R}}$$

and define the contours, for $j=1$ or 2,

$$(8.9)\qquad \begin{aligned} L_j' &= \Big(\frac{4^{-j}c}{\log V}\Big) = \Big\{\frac{4^{-j}c}{\log V}+it:\ -\infty<t<\infty\Big\},\\ L_j &= \Big\{\frac{4^{-j}c}{\log V}+it:\ |t|\le 4^{-j}V\Big\},\\ \mathscr{L}_j &= \Big\{-\frac{4^{-j}c}{\log V}+it:\ |t|\le 4^{-j}V\Big\},\\ H_j &= \Big\{\sigma_j \pm i4^{-j}V:\ |\sigma_j|\le \frac{4^{-j}c}{\log V}\Big\}. \end{aligned}$$

By (8.7) the integrand in (8.1) vanishes as $|t_1|\to\infty$ or $|t_2|\to\infty$ provided s_1 and s_2 are to the right of L_2'. We first shift the contours (1) for the integrals over s_1 and s_2 to L_1' and L_2', respectively. Next, we truncate these contours so that they may be replaced with L_1 and L_2. In doing this there are two error terms which are estimated by (8.7). For example the error term coming from L_1' and the truncated piece of L_2' is

$$\ll_M (\log U)^{CM}(\log V)^M V^{\frac{5c}{16}}\left(\int_{-\infty}^{\infty}\frac{(\log|t|+3)^{2M}}{|\frac{c}{4\log V}+it|^{a+u+1}}\,dt\right) \times\left(\int_{V/16}^{\infty}\frac{(\log t)^{2M}}{t^2}\,dt\right)$$
$$\ll_M \frac{(\log V)^{6M}}{V^{1-\frac{5c}{16}}} \ll_M e^{-c\sqrt{\log R}}.$$

Hence

$$(8.10)\qquad \mathscr{T}_R^* = \frac{1}{(2\pi i)^2}\int_{L_2}\int_{L_1}\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\,ds_1\,ds_2 + O_M(e^{-c\sqrt{\log R}}).$$

To replace the s_1-contour along L_1 with the contour along $\mathscr{L}_1$ we consider the rectangle formed by L_1, H_1, and $\mathscr{L}_1$ which contains poles of the integrand as a function of s_1 at $s_1=0$ and $s_1=-s_2$. Hence we see that

$$
\begin{aligned}
(8.11)\qquad \mathcal{T}_R^* &= \frac{1}{2\pi i}\int_{L_2} \operatorname*{Res}_{s_1=0}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) ds_2 \\
&+ \frac{1}{2\pi i}\int_{L_2} \operatorname*{Res}_{s_1=-s_2}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) ds_2 \\
&+ \frac{1}{(2\pi i)^2}\int_{L_2}\int_{\mathcal{L}_1\cup H_1} \frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\, ds_1\, ds_2 + O_M(e^{-c\sqrt{\log R}}).
\end{aligned}
$$

Here the contours along $\mathcal{L}_1$ and H_1 are oriented clockwise. In the first and third integrals we move the contour over L_2 to $\mathcal{L}_2$ in the same fashion, but now we only pass a pole at $s_2 = 0$. Thus we obtain

(8.12)

$$
\begin{aligned}
T_R^* &= \operatorname*{Res}_{s_2=0}\operatorname*{Res}_{s_1=0} \frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d} \\
&\quad + \frac{1}{2\pi i}\int_{\mathcal{L}_2\cup H_2} \operatorname*{Res}_{s_1=0}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) ds_2 \\
&\quad + \frac{1}{2\pi i}\int_{\mathcal{L}_1\cup H_1} \operatorname*{Res}_{s_2=0}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) ds_1 \\
&\quad + \frac{1}{2\pi i}\int_{L_2} \operatorname*{Res}_{s_1=-s_2}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) ds_2 \\
&\quad + \frac{1}{(2\pi i)^2}\int_{\mathcal{L}_2\cup H_2}\int_{\mathcal{L}_1\cup H_1} \frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\, ds_1\, ds_2 + O_M(e^{-c\sqrt{\log N}}) \\
&:= I_0 + I_1 + I_2 + I_3 + I_4 + O_M(e^{-c\sqrt{\log R}}).
\end{aligned}
$$

We will see that the residue I_0 provides the main term and some of the lower order terms, the integral I_3 provides the remaining lower order terms, and the integrals I_1, I_2, and I_4 are error terms.

We consider first I_0. At $s_1 = 0$ there is a pole of order $\le u+1$, and therefore[8] by Leibniz's rule we have

$$
\operatorname*{Res}_{s_1=0} \frac{D(s_1,s_2)R^{s_1}}{s_1^{u+1}(s_1+s_2)^d} = \frac{1}{u!}\sum_{i=0}^{u}\binom{u}{i}(\log R)^{u-i}\frac{\partial^i}{\partial s_1^i}\left(\frac{D(s_1,s_2)}{(s_1+s_2)^d}\right)\Bigg|_{s_1=0}
$$

[8] If $G(0,0) = 0$ then the order of the pole is u or less, but the formula we use to compute the residue is still valid. In this situation one or more of the initial terms will have the value zero.

and

$$\frac{\partial^i}{\partial s_1^i}\left(\frac{D(s_1,s_2)}{(s_1+s_2)^d}\right)\Bigg|_{s_1=0} = (-1)^i\,\frac{D(0,s_2)d(d+1)\cdots(d+i-1)}{s_2^{d+i}}$$
$$+\sum_{j=1}^{i}\binom{i}{j}\frac{\partial^j}{\partial s_1^j}D(s_1,s_2)\Bigg|_{s_1=0}(-1)^{i-j}\,\frac{d(d+1)\cdots(d+i-j-1)}{s_2^{d+i-j}},$$

where in case of $i=j$ (including the case when $i=j=0$ and $d\ge 0$ arbitrary) the empty product in the numerator is 1. We conclude that

$$\operatorname*{Res}_{s_1=0}\frac{D(s_1,s_2)R^{s_1}}{s_1^{u+1}(s_1+s_2)^d} = \sum_{i=0}^{u}\sum_{j=0}^{i}\frac{a(i,j)(\log R)^{u-i}}{s_2^{d+i-j}}\frac{\partial^j}{\partial s_1^j}D(s_1,s_2)\Bigg|_{s_1=0} \tag{8.13}$$

with $a(i,j)$ as given explicitly in the previous equations. To complete the evaluation of I_0, we see that the (i,j)th term contributes to I_0 a pole at $s_2=0$ of order $v+1+d+i-j$ (or less), and therefore by Leibniz's formula

$$\operatorname*{Res}_{s_2=0}\frac{R^{s_2}}{s_2^{v+1+d+i-j}}\frac{\partial^j}{\partial s_1^j}D(s_1,s_2)\Bigg|_{s_1=0}$$
$$=\frac{1}{(v+d+i-j)!}\sum_{m=0}^{v+d+i-j}\binom{v+d+i-j}{m}(\log R)^{v+d+i-j-m}\frac{\partial^m}{\partial s_2^m}\frac{\partial^j}{\partial s_1^j}D(s_1,s_2)\Bigg|_{\substack{s_1=0\\ s_2=0}}.$$

This completes the evaluation of I_0, and we conclude

$$I_0=\sum_{i=0}^{u}\sum_{j=0}^{i}\sum_{m=0}^{v+d+i-j} b(i,j,m)\left(\frac{\partial^m}{\partial s_2^m}\frac{\partial^j}{\partial s_1^j}D(s_1,s_2)\Bigg|_{\substack{s_1=0\\ s_2=0}}\right)(\log R)^{u+v+d-j-m}, \tag{8.14}$$

where

$$b(i,j,m)=(-1)^{i-j}\binom{u}{i}\binom{i}{j}\binom{v+d+i-j}{m}\frac{d(d+1)\cdots(d+i-j-1)}{u!(v+d+i-j)!}. \tag{8.15}$$

The main term is of order $(\log R)^{u+v+d}$ and occurs when $j=m=0$. Therefore, it is given by

$$G(0,0)(\log R)^{u+v+d}\left(\frac{1}{u!}\sum_{i=0}^{u}\binom{u}{i}(-1)^i\,\frac{d(d+1)\cdots(d+i-1)}{(v+d+i)!}\right).$$

It is not hard to prove that

$$\frac{1}{u!}\sum_{i=0}^{u}\binom{u}{i}(-1)^i\,\frac{d(d+1)\cdots(d+i-1)}{(v+d+i)!}=\binom{u+v}{u}\frac{1}{(u+v+d)!}, \tag{8.16}$$

from which we conclude that the main term is

$$G(0,0)\binom{u+v}{u}\frac{1}{(u+v+d)!}(\log R)^{d+u+v}. \tag{8.17}$$

Motohashi found the following approach which avoids proving (8.16) directly and which can be used to simplify some of the previous analysis. Granville also made a similar observation. The residue we are computing is equal to

$$\frac{1}{(2\pi i)^2}\int\limits_{\Gamma_2}\int\limits_{\Gamma_1}\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\,ds_1\,ds_2,$$

where Γ_1 and Γ_2 are the circles $|s_1|=\rho$ and $|s_2|=2\rho$, respectively, with a small $\rho>0$. When $s_1=s$ and $s_2=sw$, this is equal to

$$\frac{1}{(2\pi i)^2}\int\limits_{\Gamma_3}\int\limits_{\Gamma_1}\frac{D(s,sw)R^{s(w+1)}}{s^{u+v+d+1}w^{v+1}(w+1)^d}\,ds\,dw,$$

with Γ_3 the circle $|w|=2$. The main term is obtained from the constant term $G(0,0)$ in the Taylor expansion of $D(s,sw)$ and, therefore, equals

$$G(0,0)\frac{(\log R)^{u+v+d}}{(u+v+d)!}\,\frac{1}{2\pi i}\int\limits_{\Gamma_3}\frac{(w+1)^{u+v}}{w^{v+1}}\,dw=G(0,0)\frac{(\log R)^{u+v+d}}{(u+v+d)!}\binom{u+v}{v},$$

by the binomial expansion.

To complete the analysis of I_0, we only need to show that the partial derivatives of $D(s_1,s_2)$ at $(0,0)$ satisfy the bounds given in the lemma. For this, we use Cauchy's estimate for derivatives

$$|f^{(j)}(z_0)|\le\max_{|z-z_0|=\eta}|f(z)|\frac{j!}{\eta^j}, \tag{8.18}$$

if $f(z)$ is analytic for $|z-z_0|\le\eta$. In the application below we will choose z_0 on $\mathscr{L}$ or to the right of $\mathscr{L}$ and

$$\eta=\frac{1}{C\log U\log T},\quad \text{where } T=|s_1|+|s_2|+3. \tag{8.19}$$

Thus we see the whole circle $|z-z_0-1|=\eta$ will remain in the region (5.3) and the estimates (5.4) hold in this circle. (We remind the reader that the generic constants c, C take different values at different appearances.) Thus, we have for s_1, s_2 on $\mathscr{L}$

or to the right of $\mathscr{L}$, and $j \le M$, $m \le 2M$,

$$\begin{aligned}(8.20)\qquad & \frac{\partial^m}{\partial s_2^m}\frac{\partial^j}{\partial s_1^j} D(s_1,s_2) \\ & \le \; j!m!(C\log U\log T)^{j+m} \max_{|s_1^*-s_1|\le\eta, |s_2^*-s_2|\le\eta} |D(s_1^*,s_2^*)| \\ & \ll_M \exp(CMU^{\delta_1+\delta_2}\log\log U)(\log T)^{6M} \frac{\max(1,|s_1+s_2|)^d}{\max(1,|s_1|)^a \max(1,|s_2|)^b},\end{aligned}$$

which, if $\max(|s_1|,|s_2|) \le C$, reduces to

$$(8.21)\qquad \frac{\partial^m}{\partial s_2^m}\frac{\partial^j}{\partial s_1^j} D(s_1,s_2) \ll_M \exp\left(CMU^{\delta_1+\delta_2}\log\log U\right).$$

In particular, we have

$$(8.22)\qquad \frac{\partial^m}{\partial s_2^m}\frac{\partial^j}{\partial s_1^j} D(s_1,s_2)\bigg|_{\substack{s_1=0\\ s_2=0}} \ll_M (\log U)^{CM}.$$

We conclude from (8.14), (8.17), and (8.22) that I_0 provides the main term and some of the secondary terms in Lemma 3 which satisfy the stated bound.

We now consider I_1. By (8.13) and (8.20),

(8.23)

$$I_1 \ll_M (\log R)^u \int_{\mathscr{L}_2\cup H_2} e^{CMU^{\delta_2}\log\log U} \frac{(\log(|t_2|+3))^{3M}\max(1,|s_2|^d)}{|s_2|^{v+1+d}\max(1,|s_2|^b)} |R^{s_2}||ds_2|.$$

By (8.4) we have $b+v \ge 1$, along H_2, $|R^{s_2}| \ll e^{c\sqrt{\log R}}$, and along both $\mathscr{L}_2$ and H_2 we have $U^{\delta_2} \ll 1$. When $|s_2| \ge 1$,

$$\frac{\max(1,|s_2|^d)}{|s_2|^{v+1+d}\max(1,|s_2|^b)} \ll \frac{1}{|s_2|^{v+1+b}} \ll \frac{1}{|s_2|^2},$$

and therefore the contribution from H_2 to I_1 is

$$\ll_M \frac{(\log R)^{7M/2-1/2}}{V^2} e^{c\sqrt{\log R}} \ll_M e^{-c\sqrt{\log R}}.$$

Similarly the integral along $\mathscr{L}_2$ is bounded by

$$\begin{aligned}& \ll_M (\log R)^{2M} R^{-\frac{c}{16\log V}} \int_{-V}^{V} (\log(|t|+3))^{3M} \min\left(\frac{1}{(\log V)^{-3M}}, \frac{1}{t^2}\right) dt \\ & \ll_M (\log R)^{3M} R^{-\frac{c}{16\log V}} \\ & \ll_M e^{-c\sqrt{\log R}},\end{aligned}$$

and therefore I_1 also satisfies this bound. The same bound holds for I_2 since it is with relabeling equal to I_1. Further, I_4 also satisfies this bound by the same argument on applying (8.7) and noting that $|s_1+s_2| \gg \frac{c}{\log V}$ in I_4.

Finally, we examine I_3, which only occurs if $d \geq 1$:

(8.24)
$$\operatorname*{Res}_{s_1=-s_2}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) = \lim_{s_1\to -s_2}\frac{1}{(d-1)!}\frac{\partial^{d-1}}{\partial {s_1}^{d-1}}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{{s_1}^{u+1}{s_2}^{v+1}}\right)$$
$$= \frac{1}{(d-1)!}\sum_{i=0}^{d-1}\mathcal{B}_i(s_2)(\log R)^{d-1-i},$$

where

(8.25)
$$\mathcal{B}_i(s_2) = \binom{d-1}{i}\sum_{j=0}^{i}\binom{i}{j}\frac{\partial^{i-j}}{\partial s_1^{i-j}}D(s_1,s_2)\bigg|_{s_1=-s_2}\frac{(-1)^j(u+1)\cdots(u+j)}{(-1)^{u+j+1}s_2^{u+v+j+2}}.$$

Therefore by (8.12), (8.24), and (8.25),

$$I_3 = \frac{1}{(d-1)!}\sum_{i=0}^{d-1}\mathcal{C}_i(\log R)^{d-1-i}, \tag{8.26}$$

where

$$\mathcal{C}_i = \frac{1}{2\pi i}\int_{L_2}\mathcal{B}_i(s_2)\,ds_2, \quad 0 \leq i \leq d-1. \tag{8.27}$$

By (8.20) and (8.25) we see that for s_2 to the right of $\mathcal{L}$

$$\mathcal{B}_i(s_2) \ll_M \exp\left(CMU^{|\delta_2|}\log\log U\right)\frac{\left(\log(|t_2|+3)\right)^{4M}}{|t_2|^{u+v+a+b+2}\max(1,|t_2|^i)}. \tag{8.28}$$

In (8.27) we may shift the contour L_2 to the imaginary axis with a semicircle of radius $1/\log U$ centered at and to the right of $s_2=0$. Further, we can extend this contour to the complete imaginary axis with an error $O_M(e^{-c\sqrt{\log R}})$ using (8.28) and the same argument used above (8.10). Letting

$$\mathcal{L}' = \left\{s = it : |t| \geq \frac{1}{\log U}\right\} \cup \left\{s = \frac{e^{i\vartheta}}{\log U} : -\frac{\pi}{2} \leq \vartheta \leq \frac{\pi}{2}\right\} \tag{8.29}$$

oriented from $-i\infty$ up to $i\infty$, we conclude

$$\mathcal{C}_i = \frac{1}{2\pi i}\int_{\mathcal{L}'}\mathcal{B}_i(s_2)\,ds_2 + O_M(e^{-c\sqrt{\log R}}), \quad 0 \leq i \leq d-1. \tag{8.30}$$

The integral here is independent of R but depends on h. Therefore this provides in (8.26) some further lower order terms in Lemma 3. The contribution to $\mathscr{C}_i$ from the integral along the imaginary axis is

$$\ll_M (\log U)^{u+v+i+a+b+1} \exp(CM \log\log U) \ll_M (\log U)^{C'M}. \tag{8.31}$$

This expression also bounds the contribution to $\mathscr{C}_i$ from the semicircle contour, completing the evaluation of I_3. Combining our results, we obtain Lemma 3. $\square$

9. Proof of Proposition 2

We introduce some standard notation associated with (1.2) and (1.3). Let

$$\theta(x;q,a) := \sum_{\substack{p\le x\\ p\equiv a \,(\mathrm{mod}\, q)}} \log p = [(a,q)=1]\frac{x}{\phi(q)} + E(x;q,a), \tag{9.1}$$

where $[S]$ is 1 if the statement S is true and is 0 if S is false. Next, we define

$$E'(x,q) := \max_{a,\ (a,q)=1} |E(x;q,a)|, \quad E^*(x,q) = \max_{y\le x} E'(y,q). \tag{9.2}$$

In this paper we only need level of distribution results for E', but usually these results are stated in the stronger form for E^*. Thus, for some $1/2 \le \vartheta \le 1$, we assume, given any $A>0$ and $\varepsilon>0$, that

$$\sum_{q\le x^{\vartheta-\varepsilon}} E^*(x,q) \ll_{A,\varepsilon} \frac{x}{(\log x)^A}. \tag{9.3}$$

This is known to hold with $\vartheta = 1/2$.

We prove the following stronger version of Proposition 2. Let

$$C_R(\ell_1,\ell_2,\mathscr{H}_1,\mathscr{H}_2,h_0) = \begin{cases} 1 & \text{if } h_0 \notin \mathscr{H}, \\ \dfrac{(\ell_1+\ell_2+1)\log R}{(\ell_1+1)(r+\ell_1+\ell_2+1)} & \text{if } h_0 \in \mathscr{H}_1\setminus\mathscr{H}_2, \\ \dfrac{(\ell_1+\ell_2+2)(\ell_1+\ell_2+1)\log R}{(\ell_1+1)(\ell_2+1)(r+\ell_1+\ell_2+1)} & \text{if } h_0 \in \mathscr{H}_1\cap\mathscr{H}_2. \end{cases}$$

By relabeling the variables we obtain the corresponding form if $h_0 \in \mathscr{H}_2\setminus\mathscr{H}_1$. We continue to use the notation (7.1).

PROPOSITION 5. *Suppose $h \ll R$. Given any positive A, there exists $B = B(A,M)$ such that for*

$$R \ll_{M,A} N^{\frac14}/(\log N)^B \quad \textit{and} \quad R,N\to\infty \tag{9.4}$$

we have

$$\sum_{n=1}^{N} \Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\theta(n+h_0) \tag{9.5}$$
$$= \frac{C_R(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0)}{(r+\ell_1+\ell_2)!}\binom{\ell_1+\ell_2}{\ell_1}\mathfrak{S}(\mathcal{H}^0)N(\log R)^{r+\ell_1+\ell_2}$$
$$+ N\sum_{j=1}^{r}\mathcal{D}_j(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0)(\log R)^{r+\ell_1+\ell_2-j} + O_{M,A}\left(\frac{N}{(\log N)^A}\right),$$

where the $\mathcal{D}_j(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0)$*'s are functions independent of R and N which satisfy the bound*

$$\mathcal{D}_j(\mathcal{H}_1,\mathcal{H}_2,h_0) \ll_M (\log U)^{C_j} \ll_M (\log\log 10h)^{C_j'} \tag{9.6}$$

for some positive constants C_j, C_j' *depending on* M. *If conjecture* (9.3) *holds, then* (9.5) *holds for* $R \ll_M N^{\frac{\vartheta}{2}-\varepsilon}$ *and* $h \le R^{\varepsilon}$, *for any given* $\varepsilon > 0$.

Proof. We assume that both $\mathcal{H}_1$ and $\mathcal{H}_2$ are nonempty so that $k_1 \ge 1$ and $k_2 \ge 1$. The proof in the case when one of these sets is empty is much easier and may be obtained by an argument analogous to that of Section 6. We have

$$\widetilde{\mathcal{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0) := \sum_{n=1}^{N}\Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\theta(n+h_0) \tag{9.7}$$
$$= \frac{1}{(k_1+\ell_1)!(k_2+\ell_2)!}\sum_{d,e\le R}\mu(d)\mu(e)\left(\log\frac{R}{d}\right)^{k_1+\ell_1}\left(\log\frac{R}{e}\right)^{k_2+\ell_2}$$
$$\times \sum_{\substack{1\le n\le N\\ d\mid P_{\mathcal{H}_1}(n)\\ e\mid P_{\mathcal{H}_2}(n)}}\theta(n+h_0).$$

To treat the inner sum above, let $d = a_1a_{12}$ and $e = a_2a_{12}$, where $(d,e) = a_{12}$, so that a_1, a_2, and a_{12} are pairwise relatively prime. As in Section 7, the n for which $d\mid P_{\mathcal{H}_1}(n)$ and $e\mid P_{\mathcal{H}_2}(n)$ cover certain residue classes modulo $[d,e]$. If $n \equiv b \pmod{a_1a_2a_{12}}$ is such a residue class, then letting

$$m = n + h_0 \equiv b + h_0 \pmod{a_1a_2a_{12}},$$

we see that this residue class contributes to the inner sum

$$\sum_{\substack{1+h_0\le m\le N+h_0\\ m\equiv b+h_0 \pmod{a_1a_2a_{12}}}}\theta(m) \tag{9.8}$$
$$= \theta(N+h_0;a_1a_2a_{12},b+h_0) - \theta(h_0;a_1a_2a_{12},b+h_0)$$
$$= [(b+h_0,a_1a_2a_{12})=1]\frac{N}{\phi(a_1a_2a_{12})} + E(N;a_1a_2a_{12},b+h_0) + O(h\log N).$$

We must determine the number of these residue classes where $(b+h_0, a_1a_2a_{12}) = 1$ so that the main term is non-zero. If $p|a_1$, then $b \equiv -h_j \pmod p$ for some $h_j \in \mathcal{H}_1$, and therefore $b + h_0 \equiv h_0 - h_j \pmod p$. Thus, if h_0 is distinct modulo p from all the $h_j \in \mathcal{H}_1$, then all $\nu_p(\mathcal{H}_1)$ residue classes satisfy the relatively prime condition, while otherwise $h_0 \equiv h_j \pmod p$ for some $h_j \in \mathcal{H}_1$ leaving $\nu_p(\mathcal{H}_1) - 1$ residue classes with a non-zero main term. We introduce the notation $\nu_p^*({\mathcal{H}_1}^0)$ for this number in either case, where we define for a set $\mathcal{H}$ and integer h_0

$$\nu_p^*(\mathcal{H}^0) = \nu_p(\mathcal{H}^0) - 1, \tag{9.9}$$

where

$$\mathcal{H}^0 = \mathcal{H} \cup \{h_0\}. \tag{9.10}$$

We extend this definition to $\nu_d^*(\mathcal{H}^0)$ for squarefree numbers d by multiplicativity. The function ν_d^* is familiar in sieve theory; see [16]. A more algebraic discussion of ${\nu_d}^*$ may also be found in [14], [15]. We define $\overline{\nu}_d^*\big((\mathcal{H}_1 \overline{\cap} \mathcal{H}_2)^0\big)$ as in (7.5).

Next, the divisibility conditions $a_2 | P_{\mathcal{H}_2}(n)$, $a_{12}|P_{\mathcal{H}_1}(n)$, and $a_{12}|P_{\mathcal{H}_2}(n)$ are handled as in Section 7 together with the above considerations. Since $E(n;q,a) \ll (\log N)$ if $(a,q) > 1$ and $q \le N$, we conclude that

$$\sum_{\substack{1 \le n \le N \\ d | P_{\mathcal{H}_1}(n) \\ e | P_{\mathcal{H}_2}(n)}} \theta(n+h_0) = \nu_{a_1}^*({\mathcal{H}_1}^0)\nu_{a_2}^*({\mathcal{H}_2}^0)\overline{\nu}_{a_{12}}^*\left((\mathcal{H}_1 \overline{\cap} \mathcal{H}_2)^0\right)\frac{N}{\phi(a_1a_2a_{12})} + O\Big(d_k(a_1a_2a_{12})\Big(\max_{\substack{b \\ (b,a_1a_2a_{12})=1}} \big|E(N;a_1a_2a_{12},b)\big| + h(\log N)\Big)\Big). \tag{9.11}$$

Substituting this into (9.7) we obtain for $\widetilde{\mathcal{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0)$ the value

$$\frac{N}{(k_1+\ell_1)!(k_2+\ell_2)!} \times \sum_{\substack{a_1a_{12}\le R \\ a_2a_{12} \le R}} \frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2 \nu_{a_1}^*({\mathcal{H}_1}^0)\nu_{a_2}^*({\mathcal{H}_2}^0)\overline{\nu}_{a_{12}}^*\left((\mathcal{H}_1 \overline{\cap} \mathcal{H}_2)^0\right)}{\phi(a_1a_2a_{12})} \times \left(\log\frac{R}{a_1a_{12}}\right)^{k_1+\ell_1}\left(\log\frac{R}{a_2a_{12}}\right)^{k_2+\ell_2} + O\Bigg((\log R)^M \sum_{\substack{a_1a_{12}\le R \\ a_2a_{12} \le R}} d_k(a_1a_2a_{12})E'(N,a_1a_2a_{12})\Bigg) + O(hR^2(3\log N)^{M+3k+1}),$$

that is to say,

$$\tag{9.12} \begin{aligned}\widetilde{\mathcal{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0) &= N\widetilde{\mathcal{T}}_R(\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0)\\ &\quad + O\big((\log R)^M \mathcal{E}_k(N)\big) + O\big(hR^2(3\log N)^{M+3k+1}\big),\end{aligned}$$

where the last error term was obtained using Lemma 2. To estimate the first error term we use Lemma 2, (1.3), and the trivial estimate $E'(N,q) \le (2N/q)\log N$ for $q \le N$ to find, uniformly for $k \le \sqrt{(\log N)/18}$, that

$$\tag{9.13} \begin{aligned} |\mathcal{E}_k(N)| &\le \sideset{}{^\flat}\sum_{q\le R^2} d_k(q) \max_{\substack{b\\(b,q)=1}} \big|E(N;q,b)\big| \sum_{q=a_1a_2a_{12}} 1\\ &= \sideset{}{^\flat}\sum_{q\le R^2} d_k(q)d_3(q)E'(N,q) \le \sqrt{\sideset{}{^\flat}\sum_{q\le R^2}\frac{d_{3k}(q)^2}{q}}\sqrt{\sum_{q\le R^2} q(E'(N,q))^2}\\ &\le \sqrt{(\log N)^{9k^2}}\sqrt{2N\log N}\sqrt{\sum_{q\le R^2} E'(N,q)} \ll N(\log N)^{(9k^2+1-A)/2},\end{aligned}$$

provided $R^2 \ll N^{\frac{1}{2}}/(\log N)^B$. On relabeling, we conclude that given any positive integers A and M there is a positive constant $B = B(A,M)$ so that for $R \ll N^{\frac{1}{4}}/(\log N)^B$ and $h \le R$,

$$\tag{9.14} \widetilde{\mathcal{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0) = N\widetilde{\mathcal{T}}_R(\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0) + O_M\left(\frac{N}{(\log N)^A}\right).$$

Using (9.3) with any $\vartheta > 1/2$, we see that (9.14) holds for the longer range $R \ll_M N^{\frac{\vartheta}{2}-\varepsilon}$, $h \ll N^{\varepsilon}$.

Returning to the main term in (9.12), we have by (6.6) that

$$\tag{9.15} \begin{aligned}\widetilde{\mathcal{T}}_R(\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0) &= \frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)} F(s_1,s_2)\frac{R^{s_1}}{{s_1}^{k_1+\ell_1+1}}\frac{R^{s_2}}{{s_2}^{k_2+\ell_2+1}}ds_1ds_2,\end{aligned}$$

where, by letting $s_j = \sigma_j + it_j$ and assuming $\sigma_1,\sigma_2 > 0$,

$$\tag{9.16} \begin{aligned} F(s_1,s_2) &= \sum_{1\le a_1,a_2,a_{12}<\infty} \frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2\nu^*_{a_1}({\mathcal{H}_1}^0)\nu^*_{a_2}({\mathcal{H}_2}^0)\overline{\nu}^*_{a_{12}}\big((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\big)}{\phi(a_1){a_1}^{s_1}\phi(a_2){a_2}^{s_2}\phi(a_{12}){a_{12}}^{s_1+s_2}}\\ &= \prod_p\left(1 - \frac{\nu^*_p({\mathcal{H}_1}^0)}{(p-1)p^{s_1}} - \frac{\nu^*_p({\mathcal{H}_2}^0)}{(p-1)p^{s_2}} + \frac{\overline{\nu}^*_p\big((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\big)}{(p-1)p^{s_1+s_2}}\right).\end{aligned}$$

We now consider three cases.

Case 1. Suppose $h_0 \notin \mathcal{H}$. Then we have, for $p > h$,

$$\nu_p^*({\mathcal{H}_1}^0) = k_1, \quad \nu_p^*({\mathcal{H}_2}^0) = k_2, \quad \overline{\nu}_p^*\left((\mathcal{H}_1 \overline{\cap} \mathcal{H}_2)^0\right) = r.$$

Therefore in this case we define $G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2)$ by

$$F(s_1, s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2)\frac{\zeta(1+s_1+s_2)^r}{\zeta(1+s_1)^{k_1}\zeta(1+s_2)^{k_2}}. \tag{9.17}$$

Case 2. Suppose $h_0 \in \mathcal{H}_1$ but $h_0 \notin \mathcal{H}_2$. (By relabeling this also covers the case where $h_0 \in \mathcal{H}_2$ and $h_0 \notin \mathcal{H}_1$.) Then for $p > h$

$$\nu_p^*({\mathcal{H}_1}^0) = k_1 - 1, \quad \nu_p^*({\mathcal{H}_2}^0) = k_2, \quad \overline{\nu}_p^*\left((\mathcal{H}_1 \overline{\cap} \mathcal{H}_2)^0\right) = r.$$

Therefore, we define $G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2)$ by

$$F(s_1, s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2)\frac{\zeta(1+s_1+s_2)^r}{\zeta(1+s_1)^{k_1-1}\zeta(1+s_2)^{k_2}}. \tag{9.18}$$

Case 3. Suppose $h_0 \in \mathcal{H}_1 \cap \mathcal{H}_2$. Then for $p > h$

$$\nu_p^*({\mathcal{H}_1}^0) = k_1 - 1, \quad \nu_p^*({\mathcal{H}_2}^0) = k_2 - 1, \quad \overline{\nu}_p^*\left((\mathcal{H}_1 \overline{\cap} \mathcal{H}_2)^0\right) = r - 1.$$

Thus, we define $G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2)$ by

$$F(s_1, s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1, s_2)\frac{\zeta(1+s_1+s_2)^{r-1}}{\zeta(1+s_1)^{k_1-1}\zeta(1+s_2)^{k_2-1}}. \tag{9.19}$$

In each case, G is analytic and uniformly bounded for $\sigma_1, \sigma_2 > -c$, with any $c < 1/4$.

We now show that in all three cases

$$G_{\mathcal{H}_1,\mathcal{H}_2}(0,0) = \mathfrak{S}(\mathcal{H}^0). \tag{9.20}$$

Notice that in Cases 2 and 3 we have $\mathcal{H}^0 = \mathcal{H}$. By (5.1), (7.5), (9.9), and (9.16), we find in all three cases

$$\begin{aligned} G_{\mathcal{H}_1,\mathcal{H}_2}(0,0) &= \prod_p \left(1 - \frac{\nu_p({\mathcal{H}_1}^0) + \nu_p({\mathcal{H}_2}^0) - \overline{\nu}_p((\mathcal{H}_1 \overline{\cap} \mathcal{H}_2)^0) - 1}{p-1}\right)\left(1 - \frac{1}{p}\right)^{-a(\mathcal{H}_1,\mathcal{H}_2,h_0)} \\ &= \prod_p \left(1 - \frac{\nu_p(\mathcal{H}^0) - 1}{p-1}\right)\left(1 - \frac{1}{p}\right)^{-a(\mathcal{H}_1,\mathcal{H}_2,h_0)}, \end{aligned} \tag{9.21}$$

where $a(\mathcal{H}_1, \mathcal{H}_2, h_0) = k_1 + k_2 - r = k - r$ in Case 1; $a(\mathcal{H}_1, \mathcal{H}_2, h_0) = (k_1 - 1) + k_2 - r = k - r - 1$ in Case 2; and $a(\mathcal{H}_1, \mathcal{H}_2, h_0) = (k_1 - 1) + (k_2 - 1) - (r - 1) = k - r - 1$

in Case 3. Hence, in Case 1 we have

$$(9.22)\qquad G_{\mathcal{H}_1,\mathcal{H}_2}(0,0)=\prod_p\Big(\frac{p-\nu_p(\mathcal{H}^0)}{p-1}\Big)\Big(1-\frac{1}{p}\Big)^{-(k-r)}$$
$$=\prod_p\Big(1-\frac{\nu_p(\mathcal{H}^0)}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-(k-r+1)}=\mathfrak{S}(\mathcal{H}^0),$$

while in Cases 2 and 3 we have

$$(9.23)\quad G_{\mathcal{H}_1,\mathcal{H}_2}(0,0)=\prod_p\Big(\frac{p-\nu_p(\mathcal{H})}{p-1}\Big)\Big(1-\frac{1}{p}\Big)^{-(k-r-1)}$$
$$=\prod_p\Big(1-\frac{\nu_p(\mathcal{H})}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-(k-r)}=\mathfrak{S}(\mathcal{H})\quad(=\mathfrak{S}(\mathcal{H}^0)).$$

We are now ready to evaluate $\mathcal{T}_R(\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0)$. There are two differences between the functions F and G that appear in (9.16)–(9.19) and the earlier (7.8)–(7.10). The first difference is that a factor of p in the denominator of the Euler product in (7.8) has been replaced by $p-1$, which only affects the value of constants in calculations. The second difference is the relationship between k_1, k_2, and r, which affects the residue calculations of the main terms. However, the analysis of lower order terms and the error analysis are essentially unchanged and, therefore, we only need to examine the main terms. We use Lemma 3 here to cover all of the cases. Taking into account (9.17)–(9.19) we have in Case 1 that $a=k_1, b=k_2, d=r, u=\ell_1, v=\ell_2$; in Case 2 that $a=k_1-1, b=k_2, d=r, u=\ell_1+1, v=\ell_2$; and in Case 3 that $a=k_1-1, b=k_2-1, d=r-1, u=\ell_1+1, v=\ell_2+1$. By (9.22) and (9.23), the proof of Propositions 5 and 2 is thus complete. □

10. **Proof of Theorem 3**

For convenience, we agree in our notation below that we consider every set of size k with a multiplicity $k!$ according to all permutations of the elements $h_i\in\mathcal{H}$, unless mentioned otherwise. While unconventional, this will clarify some of the calculations.

To prove Theorem 3 we consider in place of (3.5)

$$(10.1)\quad \mathcal{S}_R(N,k,\ell,h,\nu)$$
$$:=\frac{1}{Nh^{2k+1}}\sum_{n=N+1}^{2N}\Big(\sum_{1\le h_0\le h}\theta(n+h_0)-\nu\log 3N\Big)\Big(\sum_{\substack{\mathcal{H}\subset\{1,2,\ldots,h\}\\|\mathcal{H}|=k}}\Lambda_R(n;\mathcal{H},\ell)\Big)^2$$
$$=\widetilde{M}_R(N,k,\ell,h)-\nu\frac{\log 3N}{h}M_R(N,k,\ell,h),$$

where

$$M_R(N,k,\ell,h) = \frac{1}{Nh^{2k}} \sum_{n=N+1}^{2N} \Bigg(\sum_{\substack{\mathcal{H}\subset\{1,2,\dots,h\}\\ |\mathcal{H}|=k}} \Lambda_R(n;\mathcal{H},\ell) \Bigg)^2, \tag{10.2}$$

$$\widetilde{M}_R(N,k,\ell,h) = \frac{1}{Nh^{2k+1}} \sum_{n=N+1}^{2N} \Bigg(\sum_{1\le h_0\le h} \theta(n+h_0) \Bigg) \Bigg(\sum_{\substack{\mathcal{H}\subset\{1,2,\dots,h\}\\ |\mathcal{H}|=k}} \Lambda_R(n;\mathcal{H},\ell) \Bigg)^2. \tag{10.3}$$

To evaluate M_R and $\widetilde{M}_R$ we multiply out the sum and apply Propositions 1 and 2. We need to group the pairs of sets $\mathcal{H}_1$ and $\mathcal{H}_2$ according to the size of the intersection $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$, and thus $|\mathcal{H}| = |\mathcal{H}_1 \cup \mathcal{H}_2| = 2k - r$. Let us choose now a set $\mathcal{H}$ and here, exceptionally, we disregard the permutation of the elements in $\mathcal{H}$. (However for $\mathcal{H}_1$ and $\mathcal{H}_2$ we take into account all permutations.) Given the set $\mathcal{H}$ of size $2k-r$, we can choose $\mathcal{H}_1$ in $\binom{2k-r}{k}$ ways. Afterwards, we can choose the intersection set in $\binom{k}{r}$ ways. Finally, we can arrange the elements both in $\mathcal{H}_1$ and $\mathcal{H}_2$ in $k!$ ways. This gives

$$\binom{2k-r}{k}\binom{k}{r}(k!)^2 = (2k-r)!\binom{k}{r}^2 r! \tag{10.4}$$

choices for $\mathcal{H}_1$ and $\mathcal{H}_2$, when we take into account the permutation of the elements in $\mathcal{H}_1$ and $\mathcal{H}_2$. If we consider in the summation every union set $\mathcal{H}$ of size j just once, independently of the arrangement of the elements, then Gallagher's theorem (3.7) may be formulated as

$$\sideset{}{^*}\sum_{\substack{\mathcal{H}\subset\{1,2,\dots,h\}\\ |\mathcal{H}|=j}} \mathfrak{S}(\mathcal{H}) \sim \frac{h^j}{j!}, \tag{10.5}$$

where $\sum^*$ indicates every set is counted just once. Applying this, we obtain on letting

$$x = \frac{\log R}{h}, \tag{10.6}$$

and using Proposition 1, that

(10.7)

$$\begin{aligned} M_R(N,k,\ell,h) &\sim \frac{1}{Nh^{2k}} \sum_{r=0}^{k} (2k-r)!\binom{k}{r}^2 r! \binom{2\ell}{\ell} \frac{(\log R)^{2\ell+r}}{(r+2\ell)!} N \sideset{}{^*}\sum_{|\mathcal{H}|=2k-r} \mathfrak{S}(\mathcal{H}) \\ &\sim \binom{2\ell}{\ell}(\log R)^{2\ell} \sum_{r=0}^{k} \binom{k}{r}^2 \frac{x^r}{(r+1)\cdots(r+2\ell)}. \end{aligned}$$

By Proposition 2 and (10.5),

$$\widetilde{M}_R(N,k,\ell,h) \sim \frac{1}{Nh^{2k+1}} \sum_{r=0}^{k} (2k-r)! \binom{k}{r}^2 r!\, Z_r, \tag{10.8}$$

where, abbreviating $a = \frac{2\ell+1}{\ell+1} = \binom{2\ell+1}{\ell+1}\binom{2\ell}{\ell}^{-1} = \frac{1}{2}\binom{2\ell+2}{\ell+1}\binom{2\ell}{\ell}^{-1}$, we have

(10.9)

$$\begin{aligned} Z_r := \binom{2\ell}{\ell} \frac{(\log R)^{2\ell+r}}{(r+2\ell)!} \Bigg\{ & r \sum_{|\mathcal{H}|=2k-r}^{*} 2a \frac{\log R}{r+2\ell+1} \mathfrak{S}(\mathcal{H}) N \\ & + (2k-2r) \sum_{|\mathcal{H}|=2k-r}^{*} a \frac{\log R}{r+2\ell+1} \mathfrak{S}(\mathcal{H}) N + \sum_{|\mathcal{H}|=2k-r}^{*} \sum_{\substack{h_0=1 \\ h_0 \notin \mathcal{H}}}^{h} \mathfrak{S}(\mathcal{H}^0) N \Bigg\} \\ \sim N \binom{2\ell}{\ell} \frac{(\log R)^{2\ell+r}}{(r+2\ell)!} & \left\{ \frac{h^{2k-r}}{(2k-r)!} \frac{2ak \log R}{r+2\ell+1} + \frac{2k-r+1}{(2k-r+1)!} h^{2k-r+1} \right\}. \end{aligned}$$

In the last sum we took into account which element of $\mathcal{H}^0$ is h_0, which can be chosen in $2k-r+1$ ways. Thus we obtain

$$\begin{aligned} &\widetilde{M}_R(N,k,\ell,h) \\ &\quad \sim \binom{2\ell}{\ell} (\log R)^{2\ell} \sum_{r=0}^{k} \binom{k}{r}^2 \frac{x^r}{(r+1)\cdots(r+2\ell)} \left(\frac{2ak}{r+2\ell+1} x + 1 \right). \end{aligned} \tag{10.10}$$

We conclude, on introducing the parameters

$$\varphi = \frac{1}{\ell+1}, \text{ (so that } a = 2-\varphi), \quad \Theta = \frac{\log R}{\log 3N}, \text{ (so that } R = (3N)^{\Theta}), \tag{10.11}$$

that

$$S_R(N,k,\ell,h,\nu) \sim \binom{2\ell}{\ell} (\log R)^{2\ell} P_{k,\ell,\nu}(x), \tag{10.12}$$

where

$$P_{k,\ell,\nu}(x) = \sum_{r=0}^{k} \binom{k}{r}^2 \frac{x^r}{(r+1)\cdots(r+2\ell)} \left(1 + x \left(\frac{4(1-\frac{\varphi}{2})k}{r+2\ell+1} - \frac{\nu}{\Theta} \right) \right). \tag{10.13}$$

Let

$$h = \lambda \log 3N, \quad \text{so that } x = \frac{\Theta}{\lambda}. \tag{10.14}$$

The analysis of when $S > 0$ now depends on the polynomial $P_{k,\ell,\nu}(x)$. We examine this polynomial as $k, \ell \to \infty$ in such a way that $\ell = o(k)$. In the first place, the size of the terms of the polynomial are determined by the factor

$$g(r) = \binom{k}{r}^2 x^r,$$

and since $g(r) > g(r-1)$ is equivalent to

$$r < \frac{k+1}{1+\frac{1}{\sqrt{x}}}$$

we should expect the polynomial is controlled by terms with r close to $k/(z+1)$, where

$$z = \frac{1}{\sqrt{x}}. \tag{10.15}$$

Consider now the sign of each term. For small x, the terms in the polynomial are positive, but they become negative when

$$1 + x\left(\frac{4(1-\frac{\varphi}{2})k}{r+2\ell+1} - \frac{\nu}{\Theta}\right) < 0.$$

When $r = k/(z+1)$ and $k, \ell \to \infty$, $\ell = o(k)$, we have heuristically

$$1 + x\left(\frac{4(1-\frac{\varphi}{2})k}{r+2\ell+1} - \frac{\nu}{\Theta}\right) \approx 1 + \frac{1}{z^2}\left(\frac{4k}{\frac{k}{z+1}} - \frac{\nu}{\Theta}\right) = \frac{1}{z^2}\left((z+2)^2 - \frac{\nu}{\Theta}\right).$$

Therefore, the terms will be positive for r near $k/(z+1)$ if $z > \sqrt{\nu/\Theta} - 2$, which is equivalent to $\lambda > (\sqrt{\nu} - 2\sqrt{\Theta})^2$. Since we can take Θ as close to $\vartheta/2$ as we wish, this implies Theorem 3. To make this argument precise, we choose r_0 slightly smaller than $g(r)$ maximal, and prove that all the negative terms together contribute less than the single term r_0, which will be positive for z and thus λ close to the values above.

For the proof, we may assume $\nu \geq 2$ and $1/2 \leq \vartheta_0 \leq 1$ are fixed, with $\vartheta_0 < 1$ in case $\nu = 2$. (The case $\nu = 1$ is covered by Theorem 2, and the case $\nu = 2$, $\vartheta_0 = 1$, $E_2 = 0$ is covered by (1.11) proved in Section 3.) First, we choose ε_0 as a sufficiently small fixed positive number. We will choose ℓ sufficiently large, depending on ν, ϑ_0, ε_0, and set

$$k = (\ell+1)^2 = \varphi^{-2}, \quad \ell > \ell_0(\nu, \vartheta_0, \varepsilon_0), \quad \text{so that } \varphi < \varphi_0(\nu, \vartheta_0, \varepsilon_0). \tag{10.16}$$

Furthermore, we choose

$$\Theta = \frac{\log R}{\log 3N} = \frac{\vartheta_0(1-\varphi)}{2}, \tag{10.17}$$

and (because of our assumptions on ν) we can define

$$z_0 := \sqrt{2\nu/\vartheta_0} - 2 > 0. \tag{10.18}$$

Thus, we see that

$$1+\frac{1}{{z_0}^2}\left(\frac{4k}{\frac{k}{z_0+1}}-\frac{2\nu}{\vartheta_0}\right)=\frac{1}{{z_0}^2}\left((z_0+2)^2-\frac{2\nu}{\vartheta_0}\right)=0, \tag{10.19}$$

Let us choose now

$$r_0=\left[\frac{k+1}{z_0+1}\right], \quad r_1=r_0+\varphi k=r_0+\ell+1, \tag{10.20}$$

and put

$$z=z_0(1+\varepsilon_0). \tag{10.21}$$

The linear factor in each term of $P_{k,\ell,\nu}(x)$ is, for $r_0\le r\le r_1$,

(10.22)
$$\begin{aligned}1+x\left(\frac{4(1-\frac{\varphi}{2})k}{r+2\ell+1}-\frac{\nu}{\Theta}\right)&=1+\frac{1}{z_0^2(1+\varepsilon_0)^2}\left(\frac{4k(1+O(\varphi))}{\frac{k}{z_0+1}+O(k\varphi)}-\frac{2\nu}{\vartheta_0(1-\varphi)}\right)\\&=1+\frac{-z_0^2+O(\sqrt{\nu}\varphi)+O(\nu\varphi)}{z_0^2(1+\varepsilon_0)^2}\\&>c(\nu,\vartheta_0)\varepsilon_0 \quad \text{if } \ \varphi<\varphi_0(\nu,\vartheta,\varepsilon_0),\end{aligned}$$

where $c(\nu,\vartheta_0)>0$ is a constant. Letting

$$f(r):=\binom{k}{r}^2\frac{x^r}{(r+1)\cdots(r+2\ell)}, \tag{10.23}$$

we have, for any $r_2>r_1$,

$$\begin{aligned}\frac{f(r_2)}{f(r_0)}&<\prod_{r_0<r\le r_2}\left(\frac{k+1-r}{r}\cdot\frac{1}{z}\right)^2<\prod_{r_0<r\le r_1}\left(\frac{k+1-r}{r}\cdot\frac{1}{z}\right)^2\\&<\left(\left(\frac{k+1}{r_0+1}-1\right)\frac{1}{z}\right)^{2\ell}\le\left(\frac{z_0+1-1}{z_0(1+\varepsilon_0)}\right)^{2\ell}<e^{-\varepsilon_0\ell}.\end{aligned} \tag{10.24}$$

Thus, the total contribution in absolute value of the negative terms of $P_{k,\ell,\nu}(x)$ will be, for sufficiently large ℓ, at most

$$k\left(1+\frac{4(k+\nu)}{z^2}\right)e^{-\varepsilon_0\ell}f(r_0)<e^{-\varepsilon_0\ell/2}f(r_0), \tag{10.25}$$

while that of the single term r_0 will be by (10.22) at least

$$c(\nu,\vartheta_0)\varepsilon_0 f(r_0)>e^{-\varepsilon_0\ell/2}f(r_0) \ \text{ if } \ell>\ell_0(\nu,\vartheta_0,\varepsilon_0). \tag{10.26}$$

This shows that $P_{k,\ell,\nu}(x)>0$. Hence, we must have at least $\nu+1$ primes in some interval

$$[n+1,n+h]=[n+1,n+\lambda\log 3N], \quad n\in[N+1,2N], \tag{10.27}$$

where

$$\lambda = \Theta z^2 < \frac{\vartheta_0}{2} z_0^2 (1+\varepsilon_0)^2 = (1+\varepsilon_0)^2 (\sqrt{\nu} - \sqrt{2\vartheta_0})^2. \tag{10.28}$$

Since ε_0 can be chosen arbitrarily small, this proves Theorem 3.

References

[1] P. T. BATEMAN and R. A. HORN, A heuristic asymptotic formula concerning the distribution of prime numbers, *Math. Comp.* **16** (1962), 363–367. MR 26 #6139 Zbl 0105.03302

[2] E. BOMBIERI and H. DAVENPORT, Small differences between prime numbers, *Proc. Roy. Soc. Ser.* A **293** (1966), 1–18. MR 33 #7314 Zbl 0151.04201

[3] E. BOMBIERI, J. B. FRIEDLANDER, and H. IWANIEC, Primes in arithmetic progressions to large moduli. III, *J. Amer. Math. Soc.* **2** (1989), 215–224. MR 89m:11087 Zbl 0674.10036

[4] H. DAVENPORT, *Multiplicative Number Theory*, second ed., *Grad. Texts Math.* **74**, Springer-Verlag, New York, 1980, Revised by Hugh L. Montgomery. MR 82m:10001 Zbl 0453.10002

[5] P. D. T. A. ELLIOTT and H. HALBERSTAM, A conjecture in prime number theory, in *Symposia Mathematica*, *Vol.* IV (INDAM, *Rome*, 1968/69), Academic Press, London, 1970, pp. 59–72. MR 43 #1943 Zbl 0238.10030

[6] T. J. ENGELSMA, k-tuple permissible patterns, 2005. Available at http://www.opertech.com/primes/k-tuples.html

[7] P. ERDÖS, The difference of consecutive primes, *Duke Math. J.* **6** (1940), 438–441. MR 1,292h Zbl 0023.29801

[8] K. FORD, Zero-free regions for the Riemann zeta function, in *Number Theory for the Millennium*, II (Urbana, IL, 2000), A K Peters, Natick, MA, 2002, pp. 25–56. MR 2003k:11136 Zbl 1034.11045

[9] É. FOUVRY and F. GRUPP, On the switching principle in sieve theory, *J. Reine Angew. Math.* **370** (1986), 101–126. MR 87j:11092 Zbl 0588.10051

[10] P. X. GALLAGHER, On the distribution of primes in short intervals, *Mathematika* **23** (1976), 4–9. MR 53 #13140 Zbl 0346.10024

[11] D. A. GOLDSTON, On Bombieri and Davenport's theorem concerning small gaps between primes, *Mathematika* **39** (1992), 10–17. MR 93h:11102 Zbl 0758.11037

[12] D. A. GOLDSTON and C. Y. YILDIRIM, Higher correlations of divisor sums related to primes. I. Triple correlations, *Integers* **3** (2003), A5, 66 pp. MR 2004h:11075 Zbl 1118.11039

[13] ———, Higher correlations of divisor sums related to primes III: Small gaps between primes, *Proc. London Math. Soc.* **95** (2007), 653–686. Zbl 1134.11034

[14] D. A. GOLDSTON, S. W. GRAHAM, J. PINTZ, and C. Y. YILDIRIM, Small gaps between primes and almost primes, *Trans. Amer. Math. Soc.* **361** (2009), 5285–5330.

[15] D. A. GOLDSTON, Y. MOTOHASHI, J. PINTZ, and C. Y. YILDIRIM, Small gaps between primes exist, *Proc. Japan Acad. Ser.* A *Math. Sci.* **82** (2006), 61–65. MR 2007a:11135 Zbl 05123005

[16] H. HALBERSTAM and H.-E. RICHERT, *Sieve Methods*, *London Math. Soc. Monogr.*, No. 4, Academic Press, New York, 1974. MR 54 #12689 Zbl 0298.10026

[17] 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. MR 1555183

[18] G. H. HARDY and J. E. LITTLEWOOD, unpublished manuscript, see [26].

[19] D. R. HEATH-BROWN, Almost-prime k-tuples, *Mathematika* **44** (1997), 245–266. MR 99a:11106 Zbl 0886.11052

[20] M. N. HUXLEY, On the differences of primes in arithmetical progressions, *Acta Arith.* **15** (1968/1969), 367–392. MR 39 #5494 Zbl 0186.36402

[21] ______, Small differences between consecutive primes. II, *Mathematika* **24** (1977), 142–152. MR 57 #5925 Zbl 0367.10038

[22] M. HUXLEY, An application of the Fouvry-Iwaniec theorem, *Acta Arith.* **43** (1984), 441–443. MR 85k:11043 Zbl 0542.10036

[23] H. MAIER, Small differences between prime numbers, *Michigan Math. J.* **35** (1988), 323–344. MR 90e:11126 Zbl 0671.10037

[24] H. L. MONTGOMERY, *Topics in Multiplicative Number Theory*, *Lecture Notes in Math.* **227**, Springer-Verlag, New York, 1971. MR 49 #2616 Zbl 0216.03501

[25] G. Z. PILT'AI, On the size of the difference between consecutive primes, *Issledovania po teorii chisel* **4** (1972), 73–79.

[26] R. A. RANKIN, The difference between consecutive prime numbers. II, *Proc. Cambridge Philos. Soc.* **36** (1940), 255–266. MR 1,292i Zbl 0025.30702

[27] G. RICCI, Sull'andamento della differenza di numeri primi consecutivi, *Riv. Mat. Univ. Parma* **5** (1954), 3–54. MR 16,675e Zbl 0058.27602

[28] A. SCHINZEL and W. SIERPIŃSKI, Sur certaines hypothèses concernant les nombres premiers, *Acta Arith.* 4, 185–208; Erratum **5** (1958), 259. MR 21 #4936 Zbl 0082.25802

[29] A. SELBERG, *Collected Papers. Vol.* II, Springer-Verlag, New York, 1991. MR 95g:01032 Zbl 0729.11001

[30] J. SIVAK, Méthodes de crible appliquées aux sommes de Kloosterman et aux petits écarts entre nombres premiers, 2005, Thèse de Doctorat de l'Université Paris Sud (Paris XI). Available at http://www.math.u-psud.fr/∼sivak/these.pdf

[31] K. SOUNDARARAJAN, Small gaps between prime numbers: the work of Goldston-Pintz-Yıldırım, *Bull. Amer. Math. Soc.* **44** (2007), 1–18. MR 2007k:11150 Zbl 05135876

[32] E. C. TITCHMARSH, *The Theory of the Riemann Zeta-Function*, second ed., The Clarendon Press, Oxford University Press, New York, 1986. MR 88c:11049 Zbl 0601.10026

[33] S. UCHIYAMA, On the difference between consecutive prime numbers, *Acta Arith.* **27** (1975), 153–157. MR 51 #3085 Zbl 0301.10037

(Received September 27, 2005)
(Revised July 22, 2006)

E-mail address: goldston@math.sjsu.edu
DEPARTMENT OF MATHEMATICS, SAN JOSE STATE UNIVERSITY, ONE WASHINGTON SQUARE, SAN JOSE, CA 95192-0130, UNITED STATES

E-mail address: pintz@renyi.hu
RÉNYI MATHEMATICAL INSTITUTE OF MATHEMATICS, P.O. BOX 127, 1364 BUDAPEST, HUNGARY

E-mail address: yalciny@boun.edu.tr
BOĞAZIÇI UNIVERSITY, DEPARTMENT OF MATHEMATICS, BEBEK, 34342 İSTANBUL, TURKEY

and

FEZA GÜRSEY ENSTITÜSÜ, KULELI MAHALLESI, ŞEKIP AYHAN ÖZIŞIK CADDESI 44, 34684 ÇENGELKÖY, İSTANBUL, TURKEY