Acta Math., 204 (2010), 1–47
DOI: 10.1007/s11511-010-0044-9

Primes in tuples II

by

Daniel A. Goldston

San José State University
San José, CA, U.S.A.

János Pintz

Hungarian Academy of Sciences
Budapest, Hungary

Cem Yalçin Yildirim

Boğaziçi University
Istanbul, Turkey

1. Introduction

In the first paper in this series [7] we proved that, letting p_n denote the nth prime,

$$\Delta = \liminf_{n\to\infty} \frac{p_{n+1}-p_n}{\log p_n} = 0, \tag{1.1}$$

culminating 80 years of work on this problem. Since the average spacing $p_{n+1}-p_n$ in the sequence of primes is asymptotically $\log p_n$, this result showed for the first time that the prime numbers do not eventually become isolated from each other, in the sense that there will always be pairs of primes closer than any fraction of the average spacing. For the history of this problem, we refer the reader to [7] and [18].

The information about primes used to obtain (1.1) is contained in the Bombieri–Vinogradov theorem. Let

$$\theta(N;q,a) = \sum_{\substack{n\leqslant N \\ n\equiv a \,(\mathrm{mod}\, q)}} \theta(n), \quad \text{where } \theta(n) = \begin{cases} \log n, & \text{if } n \text{ is prime,} \\ 0, & \text{otherwise.} \end{cases} \tag{1.2}$$

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

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

The first author was supported in part by an NSF Grant, the second author by OTKA grants No. K 67676, T 43623, T 49693 and the Balaton program, the third author by TÜBITAK.

Thus the primes tend to be equally distributed among the arithmetic progressions modulo q that allow primes, and this holds for the primes up to N and at least for almost all the progressions with modulus q up to nearly $N^{1/2}$. The principle can be quantified by saying that the primes have an *admissible level of distribution* ϑ (or *satisfy a 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 [3] conjectured that the primes have the maximal admissible level of distribution 1, while by the Bombieri–Vinogradov theorem we have immediately that $\frac{1}{2}$ is an admissible level of distribution for the primes. In [7] we proved that *if* the primes satisfy a level of distribution $\vartheta>\frac{1}{2}$ then there is an absolute constant $M(\vartheta)$ for which

$$p_{n+1}-p_n \leqslant M(\vartheta) \quad \text{for infinitely many } n. \tag{1.5}$$

In particular, assuming the Elliott–Halberstam conjecture (or just $\vartheta \geqslant 0.98$) gives

$$p_{n+1}-p_n \leqslant 16 \quad \text{for infinitely many } n. \tag{1.6}$$

These are surprising results because they show that going beyond an admissible level of distribution $\frac{1}{2}$ implies that there are infinitely often bounded gaps between primes, and therefore questions as hard as the twin prime conjecture can nearly be dealt with using this type of information.

Since we obtained our results in 2005 there has been no further progress toward (1.5), and it appears now that an extension of the Bombieri–Vinogradov theorem of sufficient strength to obtain bounded gaps between primes will require some basic new ideas. One can also pursue improving the approximations we used or improving the method used to detect primes, and again this now appears to require some essentially new idea. Our goal in this paper is to extend the current method as much as possible in order to obtain strong quantitative results. In particular, we obtain the following quantitative version of (1.1).

THEOREM 1. *The differences of consecutive primes satisfy*

$$\liminf_{n\to\infty} \frac{p_{n+1}-p_n}{\sqrt{\log p_n}(\log\log p_n)^2} < \infty. \tag{1.7}$$

This result is remarkable in that it shows that there exist pairs of primes nearly within the square root of the average spacing. By comparison, the best result for large gaps between primes [17] is that

$$\limsup_{n\to\infty} \frac{p_{n+1}-p_n}{(\log p_n)(\log\log p_n)(\log\log\log p_n)^{-2}(\log\log\log\log p_n)} \geqslant 2e^{\gamma}, \tag{1.8}$$

where γ is Euler's constant. Thus the best large gap result produces gaps larger than the average by a factor a bit smaller than $\log\log p_n$, while now the small gaps are smaller than the average by a factor a bit bigger than $(\log p_n)^{-1/2}$. In this sense the small gap result now greatly surpasses the large gap result. (There are conflicting conjectures on how large the gap between consecutive primes can get, but all of these conjectures suggest that there can be gaps at least as large as $c(\log p_n)^2$ for some constant c.)

This paper is organized as follows. In §2 we present a generalization of Theorem 1 which applies to many interesting situations and which is the result that we will prove in this paper. Unlike in [7], to obtain our results we need to take into account the possibility of exceptional characters associated with Landau–Siegel zeros. In Theorem 2 we assume that there are no Landau–Siegel zeros in a certain range and are able to obtain our main results including the gaps between primes in Theorem 1 in intervals $[N, 2N]$ for all sufficiently large N. Next, using the Landau–Page theorem, we can find a sequence of ranges which avoid possible Landau–Siegel zeros. Thus we obtain Theorem 3 which unconditionally gives the same results as Theorem 2 but without being able to localize them to a dyadic interval. The proof of these theorems requires substantial refinements of the methods of [7], and in §3 we will discuss some of these refinements and how they arise. The main technical tools needed in our proof, Theorems 4 and 5, are stated in §4. The proofs of Theorems 4 and 5 take up §§5–13. In proving Theorem 5 in our general setting, we need a modified Bombieri–Vinogradov theorem which is the topic of §12. Our method, as in [7], requires a result on the average of the singular series. In [7] the well-known result of Gallagher [5] was used, but in our current setting this result is not applicable, and therefore in §14 we prove a new result well adapted for our needs. With Theorems 4 and 5 in hand together with the new singular series average result, the proofs of Theorems 2 and 3 are completed in §15.

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 will allow these constants to be different at different occurences. Constants implied by pure o, O and $\ll$ symbols will be absolute, unless otherwise stated. The ν times iterated logarithm will be denoted by $\log_\nu N$. $\mathcal{P}$ denotes the set of primes.

Acknowledgment

We thank the referee who was helpful with many corrections and improvements in this paper.

2. A generalization of Theorem 1

Our method will allow us to prove a generalization of Theorem 1 where instead of seeking two neighboring primes of the form $n+j$ and $n+j'$ with

$$1 \leqslant j < j' \leqslant h, \quad h = h(n) = C\sqrt{\log n}\,(\log\log n)^2, \tag{2.1}$$

we look for two primes of the form $n+a_j$ and $n+a_{j'}$, where $n \in [N+1, 2N]$ and

$$\mathcal{A} = \{a_j\}_{j=1}^{h} \subset [1, N] \tag{2.2}$$

is an arbitrarily given set of integers, such that (2.1) is satisfied.

We remark that an extension of this type is a trivial consequence of the prime number theorem if

$$h' = h'(n) > (1+c)\log n, \quad c > 0 \text{ fixed}, \tag{2.3}$$

but that none of the earlier methods of Erdős [4], Bombieri–Davenport [1] and Maier [13] which produce small gaps between primes seem capable of proving a result of this type for any function satisfying

$$h'' = h''(n) < (1-c)\log n, \quad c > 0 \text{ fixed}. \tag{2.4}$$

According to a conjecture of de Polignac [19] from 1849, every even number may be written as the difference of two primes. Although we know that this is true for almost all even numbers, there is no known way to specify these values. Our generalization makes a first step in this direction by proving that we can explicitly find sparse sequences $\mathcal{A}$ such that infinitely many of the elements in $\mathcal{A}-\mathcal{A}$ are differences of two primes, i.e.

$$|(\mathcal{P}-\mathcal{P}) \cap (\mathcal{A}-\mathcal{A})| = \infty. \tag{2.5}$$

(Here we make use of the usual notation $\mathcal{A}-\mathcal{B} = \{a-b : a \in \mathcal{A} \text{ and } b \in \mathcal{B}\}$ for sets $\mathcal{A}$ and $\mathcal{B}$.) Some sequences for which our method applies are:

$$\mathcal{A} = \{k^m\}_{m=1}^{\infty}, \qquad k \geqslant 2 \text{ fixed } (k \in \mathbb{N}), \tag{2.6}$$

$$\mathcal{A} = \{k^{x^2+y^2}\}_{x,y=1}^{\infty}, \quad k \geqslant 2 \text{ fixed } (k \in \mathbb{N}), \tag{2.7}$$

$$\mathcal{A} = \{k^{f(x,y)}\}_{r=1}^{\infty}, \quad k \geqslant 2 \text{ fixed } (k \in \mathbb{N}), \tag{2.8}$$

where the value set $\mathcal{R}=\{r\in\mathbb{N}: f(x,y)=r \text{ for some } x \text{ and } y\}$ satisfies

$$\mathcal{R}(X)=|\{m\in\mathcal{R}: m\leqslant X\}|>C'\sqrt{X}\log^2 X \tag{2.9}$$

(this would happen, for example, for $f(x,y)=x^2+y^m$ with arbitrary $m\geqslant 2$, and for $f(x,y)=x^3+y^3$), or in general any set of type

$$\mathcal{A}=\{k^{r_j}\}_{j=1}^{\infty}, \quad k\geqslant 2 \text{ fixed } (k\in\mathbb{N}), \tag{2.10}$$

if $\mathcal{R}=\{r_j\}_{j=1}^{\infty}\subseteq\mathbb{N}$ satisfies the density condition (2.9). Among these sets the only trivial one is 2^m, that is (2.6) for $k=2$ which has $\lfloor\log X/\log 2\rfloor$ elements below X, thereby corresponding to the case (2.3).

Unfortunately, the possible existence of Landau–Siegel zeros (cf. §12) makes it impossible to formulate a localized version of our result for all $n\in\mathbb{N}$ satisfying the conditions (2.1) and (2.2). Even in the special case when $\mathcal{A}$ is an interval, we cannot guarantee the existence of gaps of size $O\big(\sqrt{\log N}\log_2^2 N\big)$ between primes in any interval of type $[N,2N]$ for any large N.

The first formulation of our main result assumes that there are no exceptional characters in a certain range. (This is hypothesis $S(Y)$ from §12, with $Y=Y(N)=e^{3\sqrt{\log N}}$.)

Theorem 2. *Let us suppose that an $N>N_0$ is given such that for any real primitive character χ mod q, $q\leqslant e^{3\sqrt{\log N}}$, we have*

$$L(s,\chi)\neq 0 \quad \text{for } s\in\left(1-\frac{1}{9\sqrt{\log N}},1\right]. \tag{2.11}$$

Let $\mathcal{A}=\mathcal{A}_N=\{a_j\}_{j=1}^{h}\subseteq[1,N]\cap\mathbb{N}$ be arbitrary ($a_j\neq a_{j'}$ if $j\neq j'$) with

$$h\geqslant C\sqrt{\log N}\log_2^2 N, \tag{2.12}$$

where C is an appropriate absolute constant. Then there exists $n\in[N,2N]$ such that at least two numbers of the form

$$n+a_j \text{ and } n+a_{j'}, \quad 1\leqslant j<j'\leqslant h, \tag{2.13}$$

are primes.

In §12 we show that (2.11), i.e. conjecture $S(Y(N))$, is true for an infinite sequence $N=N_\nu\to\infty$, and thus Theorem 2 implies Theorem 1 by choosing $N=N_\nu$ and

$$\mathcal{A}=\mathcal{A}_N=\{1,2,...,h\} \tag{2.14}$$

with $h=\big\lceil C\sqrt{\log N}\log_2^2 N\big\rceil$. In a similar way, Theorem 2 implies the following result too.

Theorem 3. *Let $\mathcal{A}\subseteq\mathbb{N}$ be an arbitrary sequence satisfying*

$$\mathcal{A}(N)=|\{n\in\mathcal{A}: n\leqslant N\}|>C\sqrt{\log N}\log_2^2 N \quad \text{for } N>N_0. \tag{2.15}$$

Then infinitely many elements of $\mathcal{A}-\mathcal{A}$ can be written as the difference of two primes, that is,

$$|(\mathcal{P}-\mathcal{P})\cap(\mathcal{A}-\mathcal{A})|=\infty. \tag{2.16}$$

3. Some initial considerations

The main tool of our method is an approximation for prime tuples and almost prime tuples. Consider the tuple $(n+h_1, n+h_2, ..., n+h_K)$ as n runs over the integers. If these K values are all primes for some n then we call this a prime tuple, and we wish to examine the existence of prime tuples. A first consideration is that the set of shifts

$$\mathcal{H}=\{h_1, h_2, ..., h_K\}, \quad \text{with } h_j \neq h_{j'} \text{ if } j \neq j', \tag{3.1}$$

imposes divisibility conditions on the components of the tuple which can effect the likelihood of obtaining prime tuples or even preclude the possibility of more than a single prime tuple. Specifically, let $\nu_p(\mathcal{H})$ denote the number of distinct residue classes modulo p occupied by the elements of $\mathcal{H}$, and for squarefree integers d extend this definition to $\nu_d(\mathcal{H})$ multiplicatively. The *singular series* for the set $\mathcal{H}$ is defined to be

$$\mathfrak{S}(\mathcal{H})=\prod_p \left(1-\frac{1}{p}\right)^{-K}\left(1-\frac{\nu_p(\mathcal{H})}{p}\right). \tag{3.2}$$

If $\mathfrak{S}(\mathcal{H})\neq 0$ then $\mathcal{H}$ is called *admissible*. Thus $\mathcal{H}$ is admissible if and only if $\nu_p(\mathcal{H})<p$ for all p, while if $\nu_p(\mathcal{H})=p$ then one component of the tuple is always divisible by p and there can be at most one prime tuple of this form. Hardy and Littlewood [9] conjectured an asymptotic formula for the number of prime tuples $(n+h_1, n+h_2, ..., n+h_K)$, with $1\leqslant n\leqslant N$, as $N\to\infty$. Letting

$$\theta(n)=\begin{cases} \log n, & \text{if } n \text{ is prime,} \\ 0, & \text{otherwise,} \end{cases} \tag{3.3}$$

we define

$$\Lambda(n;\mathcal{H}) := \theta(n+h_1)\theta(n+h_2)\dots\theta(n+h_K) \tag{3.4}$$

and use this function to detect prime tuples. The Hardy–Littlewood prime-tuple conjecture is the asymptotic formula

$$\sum_{n\leqslant N}\Lambda(n;\mathcal{H}) = N(\mathfrak{S}(\mathcal{H})+o(1)), \quad \text{as } N\to\infty, \tag{3.5}$$

which is trivial if $\mathcal{H}$ is not admissible, but is otherwise only known to be true in the case $K=1$, which is the prime number theorem.

The starting point for our method in [7] is to find approximations of $\Lambda(n;\mathcal{H})$ for which we can obtain asymptotic formulas similar to (3.5). A further essential idea is that rather than approximating just prime tuples, we should approximate almost prime

K-tuples with a total of $\leqslant K+\ell$ prime factors in all the components, which, in case $0\leqslant\ell\leqslant K-2$, guarantees at least two of the components being prime. The almost prime tuple approximation used in [7] and which we also use here is

$$\Lambda_R(n;\mathcal{H},\ell) := \frac{1}{(K+\ell)!}\sum_{\substack{d\mid P_{\mathcal{H}}(n)\\ d\leqslant R}}\mu(d)\left(\log\frac{R}{d}\right)^{K+\ell}, \tag{3.6}$$

where $|\mathcal{H}|=K$ and

$$\mathcal{P}_{\mathcal{H}}(n) := (n+h_1)(n+h_2)\ldots(n+h_K). \tag{3.7}$$

Our method for proving (1.1) in [7] is based on a comparison of the two sums

$$\sum_{n\leqslant N}\Lambda_R(n;\mathcal{H},\ell)^2 \quad\text{and}\quad \sum_{n\leqslant N}\theta(n+h_0)\Lambda_R(n;\mathcal{H},\ell)^2. \tag{3.8}$$

An asymptotic formula for the first sum can be obtained if $R\leqslant N^{1/2-\varepsilon}$, while for the second sum we can use an admissible level of distribution of primes ϑ to obtain an asymptotic formula when $R\leqslant N^{\vartheta/2-\varepsilon}$. In [7] it was assumed that K and ℓ are fixed, i.e. independent of N. Using these asymptotic formulas we can now evaluate

$$\mathcal{S}_R := \sum_{n=N+1}^{2N}\left(\sum_{1\leqslant h_0\leqslant h}\theta(n+h_0)-\log 3N\right)\sum_{\substack{1\leqslant h_1,h_2,\ldots,h_K\leqslant h\\ \text{distinct}}}\Lambda_R(n;\mathcal{H},\ell)^2. \tag{3.9}$$

If $\mathcal{S}_R>0$ then the sum over h_0 must have at least two non-zero terms and thus there must be some n and $h_j\neq h_{j'}$ such that $n+h_j$ and $n+h_{j'}$ are both prime. We find with $\vartheta=\frac{1}{2}$ and $h=\lambda\log N$ for any fixed $\lambda>0$ that we can choose K and ℓ for which $\mathcal{S}_R>0$, which proves (1.1). In order to obtain this for any arbitrarily small $\lambda>0$, the fixed K and ℓ are chosen sufficiently large in an appropriate way.

To obtain quantitative bounds to replace (1.1), the first step is to obtain asymptotic formulas which are uniform in K and ℓ so that these can be chosen as functions of N that go to infinity with N. One also needs explicit error terms, and these error terms arise not only from lower order terms and prime number theorem type error terms, but also in (3.8) from the Bombieri–Vinogradov theorem error terms.

We now establish the relations between our parameters that will be used throughout the paper. Recalling the set $\mathcal{A}$ from (2.2), we will always take $\mathcal{H}\subset\mathcal{A}$. Next, R and ℓ will be chosen as

$$K\leqslant h, \quad \ell\asymp\sqrt{K} \quad\text{and}\quad R=(3N)^{1/4-o(1)}. \tag{3.10}$$

We will make use of two important parameters U and V defined by

$$V := \sqrt{\log N} \quad\text{and}\quad U=e^V, \tag{3.11}$$

and will choose K later to be slightly smaller than V. We next denote the product of primes not exceeding V by

$$P := \prod_{p \leqslant V} p, \tag{3.12}$$

where p will always denote primes.

As just mentioned above, our present treatment requires a much more delicate analysis of the error terms than in [7], and therefore we make an initial simplification to facilitate this analysis. In [7] the irregular behavior of $\nu_p(\mathcal{H})$ for small primes greatly complicated the estimate of the function $G(s_1, s_2)$ and its partial derivatives. We can avoid these difficulties, at least for primes dividing P, by proceeding somewhat similarly to Heath-Brown in [11]. We call a residue class $a\,(\mathrm{mod}\,P)$ *regular with respect to* $\mathcal{H}$ *and* P if

$$(P, P_{\mathcal{H}}(a)) = 1 \tag{3.13}$$

and denote by $A(\mathcal{H}) = A_P(\mathcal{H})$ the set of all regular residue classes mod P. Thus

$$A(\mathcal{H}) := \{a : 1 \leqslant a \leqslant P \text{ and } (P, P_{\mathcal{H}}(a)) = 1\}. \tag{3.14}$$

The number of regular residue classes mod P is clearly

$$|A(\mathcal{H})| = \prod_{p|P} (p - \nu_p(\mathcal{H})) \tag{3.15}$$

and their proportion of all the residue classes mod P is

$$\frac{|A(\mathcal{H})|}{P} = \prod_{p|P} \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right), \tag{3.16}$$

which is positive if $\mathcal{H}$ is admissible. Thus in particular for a given $\mathcal{H}$ and all P there exists at least one regular residue class mod P, if and only if $\mathcal{H}$ is admissible.

With this notation, we now consider the sums

$$\sum_{\substack{n=N+1 \\ n \in A(\mathcal{H}_1) \cap A(\mathcal{H}_2)}}^{2N} \Lambda_R(n; \mathcal{H}_1, \ell) \Lambda_R(n; \mathcal{H}_2, \ell) \tag{3.17}$$

and

$$\sum_{\substack{n=N+1 \\ n \in A(\mathcal{H}_1) \cap A(\mathcal{H}_2)}}^{2N} \Lambda_R(n; \mathcal{H}_1, \ell) \Lambda_R(n; \mathcal{H}_2, \ell) \theta(n + h_0), \tag{3.18}$$

with $h_0 \in [1,h]$, which are asymptotically evaluated in Theorems 4 and 5, respectively. A new feature in the proof of these theorems which does not occur in [7] is that (3.17) and (3.18) are first evaluated for each residue class $a \,(\mathrm{mod}\, P)$ with

$$a \in A(\mathcal{H}_1) \cap A(\mathcal{H}_2) = A(\mathcal{H}_1 \cup \mathcal{H}_2) \tag{3.19}$$

separately, and then the results are added over all regular residue classes mod P. It turns out that the asymptotic main term (and even secondary terms) are independent of the particular choice of the regular residue class a, so this summing presents no difficulty. However, the restriction of the values of n to a single residue class $a \,(\mathrm{mod}\, P)$ in (3.18) requires a stronger form of the Bombieri–Vinogradov theorem (cf. §12).

To detect primes, in place of (3.9) we consider

$$S'_R(N,K,\ell,P) := \frac{1}{Nh^{2K+1}} \sum_{n=N+1}^{2N} \Bigg(\sum_{p:p-n\in\mathcal{A}} \log p - \log 3N \Bigg) \Psi'_R(K,\ell,n,h)^2, \tag{3.20}$$

where

$$\Psi'_R(K,\ell,n,h) := \sum_{\substack{\mathcal{H}:|\mathcal{H}|=K \\ n\in A(\mathcal{H})}} \Lambda_R(n;\mathcal{H},\ell). \tag{3.21}$$

On applying Theorems 4 and 5 we can asymptotically evaluate S'_R, which we carry out in §15. One condition that arises from the main terms is that, in order to prove the existence of prime pairs in intervals of length h, we need

$$h > \frac{C \log N}{K}. \tag{3.22}$$

Since $K \leqslant h$, this immediately implies that

$$h > C\sqrt{\log N}. \tag{3.23}$$

Our goal is to take h as small as possible, and therefore we cannot obtain anything better than (3.23) when using the approximation in (3.6) together with (3.20). Apart from powers of $\log_2 N$, we are able to prove our results for h of this size.

Our actual choices for K and h are

$$K \leqslant c_1 \frac{\sqrt{\log N}}{\log_2^2 N} \quad \text{and} \quad h = \frac{25 \log N}{K} \geqslant \frac{25}{c_1} \sqrt{\log N} \log_2^2 N, \tag{3.24}$$

with a sufficiently small explicitly calculable absolute constant c_1 (to be chosen later). We will need the error terms in Theorems 4 and 5 to be uniform in K with a relative error of size η_1 satisfying

$$\eta_1 < \frac{c}{\sqrt{K}}. \tag{3.25}$$

However, we do not achieve this for all admissible pairs $\mathcal{H}_1$ and $\mathcal{H}_2$ of size K. Instead, for all admissible pairs $\mathcal{H}_1, \mathcal{H}_2$ we obtain a weaker error term, but if

$$K - |\mathcal{H}_1 \cap \mathcal{H}_2| \ll \sqrt{K} \tag{3.26}$$

then we do obtain the error estimate in (3.25). This turns out to be sufficient for our proof, since such pairs $\mathcal{H}_1, \mathcal{H}_2$ will be dominant in (3.20).

4. Two basic theorems

In the following, let N be a sufficiently large integer, c_1 a sufficiently small positive constant,

$$K \leqslant c_1 \frac{\sqrt{\log N}}{\log_2^2 N}, \tag{4.1}$$

$$K \ll k_1, k_2 \leqslant K \quad \text{and} \quad \sqrt{K} \ll \ell_1, \ell_2 \ll \sqrt{K}. \tag{4.2}$$

We will consider sets $\mathcal{H} := \mathcal{H}_1 \cup \mathcal{H}_2$, $\mathcal{H}_1, \mathcal{H}_2 \subseteq [1, N]$, of sizes

$$|\mathcal{H}_1| = k_1, \quad |\mathcal{H}_2| = k_2 \quad \text{and} \quad |\mathcal{H}_1 \cap \mathcal{H}_2| = r. \tag{4.3}$$

Let

$$\overline{m} := K - m \text{ for } m \in [0, K], \quad n^* := \max\{\sqrt{K}, n\} \quad \text{and} \quad \bar{n}^* := (\bar{n})^*. \tag{4.4}$$

Our first main result is the following theorem.

THEOREM 4. *For $N^c < R \leqslant N^{1/2} e^{-c\sqrt{\log N}}$ we have, as $N \to \infty$,*

$$\begin{aligned}\sum_{\substack{n \leqslant N \\ n \in A(\mathcal{H}_1) \cap A(\mathcal{H}_2)}} & \Lambda_R(n; \mathcal{H}_1, \ell_1) \Lambda_R(n; \mathcal{H}_2, \ell_2) \\ &= N \binom{\ell_1 + \ell_2}{\ell_1} \frac{(\log R)^{r + \ell_1 + \ell_2}}{(r + \ell_1 + \ell_2)!} \frac{\mathfrak{S}(\mathcal{H}) P}{|A(\mathcal{H})|} \left(1 + O\left(\frac{K \bar{r}^* \log_2 N}{\log R}\right)\right) \\ &\quad + O\big(N e^{-c\sqrt{\log N}}\big).\end{aligned} \tag{4.5}$$

For the next theorem we suppose that the following instance of the Bombieri–Vinogradov theorem holds (see §12). For a given, sufficiently large N, and recalling the parameter P defined in (3.12), we have

$$\sum_{\substack{q \leqslant Q^* \\ (q, P) = 1}} \max_{(a, q) = 1} \left| \sum_{\substack{N < p \leqslant 2N \\ p \equiv a \,(\mathrm{mod}\, Pq)}} \log p - \frac{N}{\varphi(Pq)} \right| \ll \frac{N}{P} e^{-c\sqrt{\log N}}, \tag{4.6}$$

where

$$Q^* = N^{1/2} P^{-3} e^{-c^* \sqrt{\log N}}, \tag{4.7}$$

with an arbitrary positive constant c^*.

Letting $\mathcal{H}^0 = \mathcal{H} \cup \{h_0\}$, our second main result is as follows.

THEOREM 5. *Suppose that N satisfies* (4.6) *and* (4.7), *and let $N^c \leqslant R \leqslant \sqrt{Q^*}$. Then*

$$\sum_{\substack{N<n\leqslant 2N\\ n\in A(\mathcal{H}_1)\cap A(\mathcal{H}_2)}} \Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\theta(n+h_0)$$
$$= N\frac{C_R(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0)}{(r+\ell_1+\ell_2)!}\binom{\ell_1+\ell_2}{\ell_1}\mathfrak{S}(\mathcal{H}^0)(\log R)^{r+\ell_1+\ell_2} \tag{4.8}$$
$$\times\left(1+O\left(\frac{K\bar{r}^*\log_2 N}{\log R}\right)\right)+O\big(Ne^{-c\sqrt{\log N}/\log_2 N}\big),$$

where

$$C_R(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0)=\begin{cases}1, & \text{if } h_0\notin\mathcal{H},\\ \dfrac{(\ell_1+\ell_2+1)\log R}{(\ell_1+1)(r+\ell_1+\ell_2+1)}, & \text{if } h_0\in\mathcal{H}_1\setminus\mathcal{H}_2,\\ \dfrac{(\ell_1+\ell_2+2)(\ell_1+\ell_2+1)\log R}{(\ell_1+1)(\ell_2+1)(r+\ell_1+\ell_2+1)}, & \text{if } h_0\in\mathcal{H}_1\cap\mathcal{H}_2.\end{cases} \tag{4.9}$$

For the applications to Theorems 1–3, the simpler case $\ell_1=\ell_2=\ell$ will be sufficient.

5. Lemmas

We will use standard properties of the Riemann zeta function $\zeta(s)$. Proceeding slightly differently from [7], we use the zero-free region, with $s=\sigma+it$,

$$\zeta(1+s)\neq 0 \quad \text{for } s\in\mathcal{R}_N:=\left\{s:\sigma\geqslant -\frac{1}{\log_2 N+6\log(|t|+3)}\right\}. \tag{5.1}$$

Further, by Titchmarsh [20, Chapter 3], for $s\in\mathcal{R}_N$ we have

$$\max\left\{\left|\zeta(1+s)-\frac{1}{s}\right|,\left|\frac{1}{\zeta(1+s)}\right|,\left|\frac{\zeta'}{\zeta}(1+s)+\frac{1}{s}\right|\right\}\ll\log(|t|+3). \tag{5.2}$$

In the course of the proof the following contours which lie in the zero-free region $\mathcal{R}_N$ will be used (with U and V given in (3.11)):

$$\begin{aligned} &\mathcal{L}_1:=\left\{s:\sigma=\frac{1}{28V}\text{ and }|t|\leqslant U\right\}, && \mathcal{L}_2:=\left\{s:\sigma=\frac{1}{14V}\text{ and }|t|\leqslant 2U\right\},\\ &\mathcal{L}_3:=\left\{s:\sigma=-\frac{1}{28V}\text{ and }|t|\leqslant U\right\}, && \mathcal{L}_4:=\left\{s:\sigma=-\frac{1}{14V}\text{ and }|t|\leqslant 2U\right\},\\ &\mathcal{L}_5:=\left\{s:|\sigma|\leqslant\frac{1}{28V}\text{ and }|t|=U\right\}, && \mathcal{L}_6:=\left\{s:|\sigma|\leqslant\frac{1}{14V}\text{ and }|t|=2U\right\},\end{aligned} \tag{5.3}$$

and

$$\mathcal{L}'=\mathcal{L}'_0\cup\mathcal{L}'_1, \tag{5.4}$$

where

$$\mathcal{L}'_0=\big\{s=\delta_0e^{i\varphi}:\tfrac12\pi\leqslant\varphi\leqslant\tfrac32\pi\big\}\quad\text{and}\quad \mathcal{L}'_1=\{s=it:\delta_0\leqslant|t|\leqslant U\},$$

with $\delta_0=\big(\sqrt{K}\log_2 N\big)^{-1}$.

Similarly to Lemma 1 of [7], we have the following.

Lemma 1. *Let* $k\log_2^2 N \leqslant \sqrt{\log R}$, $N^c \leqslant R \leqslant N$, $N \geqslant C$, $k \geqslant 2$ *and* $B \leqslant Ck$. *Then*

$$\int_{\mathcal{L}_j} (\log(|t|+3))^B \left| \frac{R^s ds}{s^k} \right| \ll e^{-c\sqrt{\log N}}, \quad 3 \leqslant j \leqslant 6, \tag{5.5}$$

where the constant implied by the $\ll$ *symbol depends only on* C.

Proof. The integral I in (5.5) satisfies, for all j,

$$I \ll \int_0^{2U} R^{-1/28V} \frac{(\log(|t|+3))^B}{\max\{|t|, 1/28V\}^k}\, dt + \int_{|\sigma| \leqslant 1/14V} R^\sigma \frac{d\sigma}{U^{3/2}}$$
$$\ll e^{-c\sqrt{\log N}} \left(\int_0^C (28V)^k\, dt + \int_C^\infty \frac{dt}{t^{3/2}} \right) + e^{\sqrt{\log N}(1/14-1/2)} \ll e^{-c\sqrt{\log N}}. \qquad \square$$

We will prove a generalization of the combinatorial identity (8.16) of [7] in order to evaluate the terms $I_{1,1}$ of §8. Let us define for triplets of integers d, u, y with $d \geqslant 0$, $u \geqslant 0$ and $y+u \geqslant 0$ (to be called *suitable triplets*) the quantity

$$Z(d,u,y) := \frac{1}{u!} \sum_{\substack{m=0 \\ m \geqslant -y}}^{u} \binom{u}{m} (-1)^m \frac{d(d+1) \dots (d+m-1)}{(y+m)!}\, . \tag{5.6}$$

Lemma 2. *We have for any suitable triplet* d, u, y *the relation*

$$Z(d,u,y) = \frac{(y-d+1) \dots (y-d+u)}{u!(y+u)!}. \tag{5.7}$$

Proof. We will prove this by induction on u. For $u=0$ we have trivially, for any non-negative d and y, $Z(d,0,y)=1/y!$ (the empty product in the numerator of Z is 1 by definition). We can suppose that $u \geqslant 1$ and that our statement is true for all suitable triplets $d, u-1, y$. We set, for any real number x and $n<0$,

$$\frac{x}{n!} = 0 \tag{5.8}$$

(in other words, we just neglect in a sum all terms with an $n!$ in the denominator with $n<0$) and fix the notation

$$[S] = \begin{cases} 1, & \text{if the statement } S \text{ is true,} \\ 0, & \text{if the statement } S \text{ is false.} \end{cases}$$

Then we obtain by $\binom{u}{j}=\binom{u-1}{j}+\binom{u-1}{j-1}$ (where we define $\binom{u-1}{u}=\binom{u-1}{-1}=0$),

$$\begin{aligned}
Z(d,u,y) &= \frac{1}{u!}\Bigg(\sum_{\substack{j=0\\ j\geqslant -y}}^{u-1} \binom{u-1}{j}(-1)^j \frac{d(d+1)\dots(d+j-1)}{(y+j)!} \\
&\qquad - \sum_{\substack{j=0\\ j\geqslant -y-1}}^{u-1} \binom{u-1}{j}(-1)^j \frac{d(d+1)\dots(d+j)}{(y+j+1)!}\Bigg) \\
&= \frac{1}{u!}\Bigg(\sum_{\substack{j=0\\ j\geqslant -y}}^{u-1} \binom{u-1}{j}(-1)^j \frac{d(d+1)\dots(d+j-1)}{(y+j)!}\left(1-\frac{d+j}{y+j+1}\right) \\
&\qquad -[-y-1\geqslant 0]\binom{u-1}{-y-1}(-1)^{-y-1}d(d+1)\dots(d-y-2)(d-y-1)\Bigg) \\
&= \frac{1}{u}\frac{1}{(u-1)!} \sum_{\substack{j=0\\ j\geqslant -y-1}}^{u-1} \binom{u-1}{j}(-1)^j \frac{d(d+1)\dots(d+j-1)(y+1-d)}{(y+j+1)!} \\
&= \frac{y+1-d}{u} Z(d,u-1,y+1) \\
&= \frac{(y+1-d)(y+2-d)\dots(y+u-d)}{u!(y+u)!}. \qquad \square
\end{aligned}$$

Finally we mention a simple lemma for the mean value of the generalized divisor function

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

where $\omega(q)$ denotes the number of prime factors of q for a squarefree q.

LEMMA 3. *If $m>0$ and $\nu\geqslant\max\{c'\log(K+1),1\}$, then there exists a constant C' depending on c' such that, for $K\geqslant 1$ and $x\geqslant 1$ we have*

$$\sum_{q\leqslant x}^{\flat} d_m(q) \leqslant x(1+\log x)^{\lceil m\rceil} \tag{5.10}$$

and

$$\sum_{q\leqslant x}^{\flat} \frac{d_{3K}(q)^{1+1/\nu}}{q} \leqslant (1+\log x)^{C'K}, \tag{5.11}$$

where the $\flat$ means that the summation runs over squarefree values of the variable.

Proof. Equation (5.10) follows from

$$\sum_{q\leqslant x}^{\flat} d_m(q) \leqslant x\left(\sum_{q\leqslant x}^{\flat} \frac{d_{\lceil m\rceil}(q)}{q}\right) \leqslant x\left(\sum_{j\leqslant x}\frac{1}{j}\right)^{\lceil m\rceil} \leqslant x(1+\log x)^{\lceil m\rceil}. \tag{5.12}$$

Further, by (5.9), we have

$$d_{3K}(q)^{1+1/\nu} = d_j(q) \tag{5.13}$$

with

$$j = (3K)^{1+1/\nu} \leqslant 9e^{1/c'} K, \tag{5.14}$$

and Lemma 3 follows with $C' = 9e^{1/c'} + 1$. □

6. Preparation for the proof of Theorem 4

Since the preparation for the proof of Theorem 4 is nearly the same as in §6 and §7 of [7] for the analogous Proposition 1 (or 3 or 4), we will briefly summarize it and the reader is referred for the details to [7]. Let

$$\mathcal{H}(p) = \{h'_1, ..., h'_{\nu_p(\mathcal{H})} : h'_j \equiv h_g \pmod p,\ h_g \in \mathcal{H} \text{ for some } g, 1 \leqslant h'_j \leqslant p\}, \tag{6.1}$$

and

$$\bar{\nu}_p(\mathcal{H}_1 \bar{\cap} \mathcal{H}_2) := \nu_p(\mathcal{H}_1(p) \cap \mathcal{H}_2(p)) = \nu_p(\mathcal{H}_1) + \nu_p(\mathcal{H}_2) - \nu_p(\mathcal{H}). \tag{6.2}$$

For any $a \in A(\mathcal{H}_1) \cap A(\mathcal{H}_2)$ (cf. (3.14)), we have, similarly to §7 of [7],

$$\begin{aligned} S_R(N; \mathcal{H}_1, \mathcal{H}_2, \ell_1, \ell_2, P, a) &:= \sum_{\substack{n=N+1 \\ n \equiv a \,(\mathrm{mod}\, P)}}^{2N} \Lambda_R(n; \mathcal{H}_1, \ell_1) \Lambda_R(n; \mathcal{H}_2, \ell_2) \\ &= \frac{N}{P} \mathcal{T}_R(\ell_1, \ell_2, \mathcal{H}_1, \mathcal{H}_2) + O(R^2 (3 \log R)^{7K}), \end{aligned} \tag{6.3}$$

where, denoting by $\int_{(1)}$ the integration over the vertical line $\{s : \mathrm{Re}\, s = 1\}$,

$$\mathcal{T}_R(\ell_1, \ell_2, \mathcal{H}_1, \mathcal{H}_2) := \frac{1}{(2\pi i)^2} \int_{(1)} \int_{(1)} F(s_1, s_2) \frac{R^{s_1}}{s_1^{K+\ell_1+1}} \frac{R^{s_2}}{s_2^{K+\ell_2+1}}\, ds_1\, ds_2, \tag{6.4}$$

$$F(s_1, s_2) := \prod_{p > V} \left(1 - \frac{\nu_p(\mathcal{H}_1)}{p^{1+s_1}} - \frac{\nu_p(\mathcal{H}_2)}{p^{1+s_2}} + \frac{\bar{\nu}_p(\mathcal{H}_1 \bar{\cap} \mathcal{H}_2)}{p^{1+s_1+s_2}}\right); \tag{6.5}$$

here, differently from [7], primes not exceeding V do not appear in $F(s_1, s_2)$ since, by the regularity of a, $(P_{\mathcal{H}_1}(n), P) = (P_{\mathcal{H}_2}(n), P) = 1$.

Let

$$\Delta := \left| \prod_{1 \leqslant j < j' \leqslant K} (h_j - h_{j'}) \right| \leqslant N^{K(K-1)/2}. \tag{6.6}$$

Then, if $p \nmid \Delta$ (consequently, for all sufficiently large primes p),

$$\nu_p(\mathcal{H}_1) = |\mathcal{H}_1| = K, \quad \nu_p(\mathcal{H}_2) = |\mathcal{H}_2| = K \quad \text{and} \quad \bar{\nu}_p(\mathcal{H}_1 \bar{\cap} \mathcal{H}_2) = |\mathcal{H}_1 \cap \mathcal{H}_2| = r. \tag{6.7}$$

We therefore factor out the dominant zeta-factors and write

$$F(s_1, s_2) = G_{\mathcal{H}_1, \mathcal{H}_2}(s_1, s_2) \frac{\zeta(1+s_1+s_2)^d}{\zeta(1+s_1)^a \zeta(1+s_2)^b} \tag{6.8}$$

with a function $G(s_1,s_2)$, regular for $\sigma_j>-\frac{1}{5}$, say, which we write slightly more generally for future application in Theorem 5 as

$$G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)=G(s_1,s_2)=G=G_1G_2G_3=G_1G_4, \tag{6.9}$$

where now $a=b=K$, $d=r$, $\nu_1(p)=\nu_p(\mathcal{H}_1)$, $\nu_2(p)=\nu_p(\mathcal{H}_2)$, $\nu_3(p)=\bar{\nu}_p(\mathcal{H}_1\bar{\cap}\mathcal{H}_2)$,

$$G_1(s_1,s_2)=\prod_{p\leqslant V}\left(1-\frac{1}{p^{1+s_1}}\right)^{-a}\prod_{p\leqslant V}\left(1-\frac{1}{p^{1+s_2}}\right)^{-b}\prod_{p\leqslant V}\left(1-\frac{1}{p^{1+s_1+s_2}}\right)^{d}, \tag{6.10}$$

and

$$\begin{aligned} G_4(s_1,s_2)&=\prod_{p>V}\left(1-\frac{\nu_1(p)}{p^{1+s_1}}-\frac{\nu_2(p)}{p^{1+s_2}}+\frac{\nu_3(p)}{p^{1+s_1+s_2}}\right)\\ &\qquad\times\left(1-\frac{1}{p^{1+s_1}}\right)^{-a}\left(1-\frac{1}{p^{1+s_2}}\right)^{-b}\left(1-\frac{1}{p^{1+s_1+s_2}}\right)^{d}\\ &=\prod_{\substack{p\mid\Delta\\ p>V}}\cdot\prod_{\substack{p\nmid\Delta\\ p>V}}=:G_2(s_1,s_2)G_3(s_1,s_2). \end{aligned} \tag{6.11}$$

Let us use the notation

$$\delta_j:=\max\{0,-\sigma_j\},\ \ j=1,2,\quad \delta:=\delta_1+\delta_2,\quad s_3:=s_1+s_2,\quad s_3^*:=s_1^*+s_2^* \tag{6.12}$$

and

$$\mathcal{R}_N':=\left\{s:\sigma\geqslant-\frac{\frac{1}{2}}{\log_2 N+6\log(|t|+3)}\right\}. \tag{6.13}$$

We will estimate the order of $G(s_1,s_2)$ in the region $s_1,s_2\in\mathcal{R}_N'$ under the more general conditions

$$a,b,d\leqslant K,\quad \nu_j(p)\leqslant K,\ \ j=1,2, \tag{6.14}$$

$$\nu_1(p)=a,\ \ \nu_2(p)=b\ \text{ and }\ \nu_3(p)=d\quad \text{for } p\nmid\Delta. \tag{6.15}$$

(Later we will examine more delicate properties of $G(s_1,s_2)$ with further conditions on a, b, d, $\nu_1(p)$ and $\nu_2(p)$.) We have

$$|G_1(s_1,s_2)|\leqslant\exp\left(C\sum_{p\leqslant V}\frac{K}{p^{1-\delta}}\right)\leqslant e^{CK\log_3 N}, \tag{6.16}$$

$$|G_2(s_1,s_2)|\leqslant\exp\left(C\sum_{p\mid\Delta}\frac{K}{p^{1-\delta}}\right)\leqslant\exp\left(CK\sum_{p\leqslant\log\Delta(1+o(1))}\frac{1}{p^{1-\delta}}\right)\leqslant e^{CK\log_3 N}, \tag{6.17}$$

$$|G_3(s_1,s_2)|\leqslant\exp\left(C\sum_{p>V}\frac{K^2}{p^{2-2\delta}}\right)\leqslant e^{CK^2/V}\leqslant e^{CK}, \tag{6.18}$$

where we made use of the estimates

$$\max\{V^\delta, (\log\Delta)^\delta\} \leqslant (\log^2 N)^{1/2\log_2 N} = e. \tag{6.19}$$

Further, in (6.17), the sum which was originally over $p|\Delta$ has been majorized by using the set of the smallest possible primes which could divide Δ.

Summarizing (6.16)–(6.18), we obtain

$$|G(s_1, s_2)| \leqslant e^{CK\log_3 N} \quad \text{for } s_1, s_2 \in \mathcal{R}'_N, \tag{6.20}$$

and further

$$|F(s_1, s_2)| \leqslant e^{CK\log_3 N}(\log(|t_1|+3)\log(|t_2|+3))^{2K} \quad \text{for } s_1, s_2, s_3 \in \mathcal{R}'_N. \tag{6.21}$$

The above estimate shows that the integrand in (6.4) vanishes as either $|t_1|\to\infty$ or $|t_2|\to\infty$, $s_1, s_2, s_3 \in \mathcal{R}'_N$. We will examine the integral (analogously to (7.15) in [7])

$$I := \mathcal{T}^*_R(d, a, b, u, v, \mathcal{H}_1, \mathcal{H}_2) := \frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)} \frac{D(s_1, s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\,ds_1\,ds_2, \tag{6.22}$$

where we introduce the function $D(s_1, s_2)$, regular for $s_1, s_2, s_3 \in \mathcal{R}'_N$:

$$\begin{aligned} D(s_1, s_2) &:= D_0(s_1, s_2)G(s_1, s_2), \\ D_0(s_1, s_2) &:= \frac{W^d(s_1+s_2)}{W^a(s_1)W^b(s_2)}, \qquad W(s) := s\zeta(1+s), \end{aligned} \tag{6.23}$$

where

$$0 \leqslant d \leqslant a, \quad b \leqslant K, \quad \min\{a, b\} \geqslant cK, \quad \tfrac{1}{8}\sqrt{K} \leqslant u, \quad v \leqslant \sqrt{K} \tag{6.24}$$

and by symmetry we may assume that $u \leqslant v$. (In the applications we will have $|a-b| \leqslant 1$ and $|u-v| \leqslant 1$.)

First step. Move the contour (1) for the integral over s_1 to $\mathcal{L}_1$, and the one over s_2 to $\mathcal{L}_2$. The vertical parts $|t| \geqslant U$ and $|t| \geqslant 2U$, respectively, can be neglected similarly to Lemma 1. After this, move the integral over s_1 from $\mathcal{L}_1$ to $\mathcal{L}_3 \cup \mathcal{L}_5$. The horizontal segments $\mathcal{L}_5$ can again be neglected. We pass a pole of order $u+1$ at $s_1 = 0$, and obtain

$$I = I_1 + I_2 + O\big(e^{-c\sqrt{\log N}}\big), \tag{6.25}$$

where

$$\begin{aligned} I_1 &:= \frac{1}{2\pi i}\int_{\mathcal{L}_2} \operatorname{Res}_{s_1=0}\left(\frac{D(s_1, s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) ds_2 \\ &= \frac{1}{2\pi i}\int_{\mathcal{L}_2} \frac{1}{u!}\left(\sum_{j=0}^{u}\binom{u}{j}(\log R)^{u-j}\frac{\partial^j}{\partial s_1^j}\left(\frac{D(s_1, s_2)}{(s_1+s_2)^d}\right)\bigg|_{s_1=0}\right)\frac{R^{s_2}}{s_2^{v+1}}\,ds_2 \end{aligned} \tag{6.26}$$

and

$$I_2 := \frac{1}{(2\pi i)^2} \int_{\mathcal{L}_2} \int_{\mathcal{L}_3} \frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1} s_2^{v+1} (s_1+s_2)^d} \, ds_1 \, ds_2. \tag{6.27}$$

We denote the complete integrand above by $Z(s_2)$ and express

$$\begin{aligned} \frac{\partial^j}{\partial s_1^j} \left(\frac{D(s_1,s_2)}{(s_1+s_2)^d} \right) \bigg|_{s_1=0} &= (-1)^j \frac{D(0,s_2) d(d+1) \dots (d+j-1)}{s_2^{d+j}} \\ &+ \sum_{j'=1}^{j} \binom{j}{j'} \frac{\partial^{j'}}{\partial s_1^{j'}} D(s_1,s_2) \bigg|_{s_1=0} (-1)^{j-j'} \frac{d(d+1) \dots (d+j-j'-1)}{s_2^{d+j-j'}}, \end{aligned} \tag{6.28}$$

where, for $j=j'$ (including also the case when $j=j'=0$ and $d \geqslant 0$ arbitrary) the empty product in the numerator is 1.

Second step. Let us denote the contribution of the first term in (6.28) to (6.26) by $I_1(j,0)$ and the others by $I_1(j,j')$, $1 \leqslant j' \leqslant j$. Then $I_1(j,0)$ belongs to the main term, while all of the $I_1(j,j')$ with $j' \geqslant 1$ just contribute to the secondary terms. Let us move now the contour $\mathcal{L}_2$ for the integral over s_2 to $\mathcal{L}_4 \cup \mathcal{L}_6$ in (6.26). The horizontal segments $\mathcal{L}_6$ can be neglected again. We pass a pole of order $v+1+d+j-j'$ in case of $I_1(j,j')$ and we obtain in this way

$$\begin{aligned} I_1 &= \frac{1}{u!} \sum_{j=0}^{u} \binom{u}{j} (\log R)^{u-j} \sum_{j'=0}^{j} (-1)^{j-j'} \binom{j}{j'} \frac{d(d+1) \dots (d+j-j'-1)}{(v+d+j-j')!} \\ &\quad \times \sum_{\nu=0}^{v+d+j-j'} \binom{v+d+j-j'}{\nu} (\log R)^{v+d+j-j'-\nu} \frac{\partial^\nu}{\partial s_2^\nu} \frac{\partial^{j'}}{\partial s_1^{j'}} D(s_1,s_2) \bigg|_{s_1=s_2=0} \\ &\quad + \frac{1}{2\pi i} \int_{\mathcal{L}_4} Z(s_2) \, ds_2 + O\big(e^{-c\sqrt{\log N}}\big) \\ &=: I_{1,1} + I_{1,2} + O\big(e^{-c\sqrt{\log N}}\big). \end{aligned} \tag{6.29}$$

7. Estimates of the partial derivatives of $D(s_1, s_2)$

In this section we will estimate the partial derivatives

$$\frac{\partial^j}{\partial s_1^j} \frac{\partial^{j'}}{\partial s_2^{j'}} D(s_1,s_2)$$

of $D(s_1,s_2)$ for $j+j' \leqslant CK$ with $s_j = s_j^*$ in $\mathcal{R}'_N$ for $1 \leqslant j \leqslant 3$. We will often use Cauchy's estimate for functions regular in $|z-z_0| \leqslant \eta$:

$$\frac{1}{j!} |f^{(j)}(z_0)| \leqslant \eta^{-j} \max_{|z-z_0|=\eta} |f(z)|. \tag{7.1}$$

Applying this to $D(s_1,s_2)$, we obtain

$$\frac{1}{j!j'!}\left|\frac{\partial^j}{\partial s_1^j}\frac{\partial^{j'}}{\partial s_2^{j'}}D(s_1^*,s_2^*)\right| \ll \eta^{-(j+j')}\max_{\substack{|s_1'-s_1^*|\leqslant\eta\\ |s_2'-s_2^*|\leqslant\eta}}|D(s_1',s_2')|. \tag{7.2}$$

In order to substitute the above maximum for $D(s_1^*,s_2^*)$, we have to estimate

$$L(s_1,s_2):=\max\left\{\left|\frac{\partial}{\partial s_1}\log D(s_1,s_2)\right|,\left|\frac{\partial}{\partial s_2}\log D(s_1,s_2)\right|\right\} \tag{7.3}$$

for $s_1,s_2,s_3\in\mathcal{R}_N'$; since by the regularity of $\log D(s_1,s_2)$ for $s_j\in\mathcal{R}_N'$, $1\leqslant j\leqslant 3$ (cf. (6.10), (6.11) and (6.23)), we have for $\eta\leqslant\frac{1}{100}(\log_2 N+\log(|t_1|+3)+\log(|t_2|+3))^{-1}$,

$$\left|\frac{D(s_1',s_2')}{D(s_1^*,s_2^*)}\right|\leqslant\exp\biggl(2\eta\max_{\substack{|s_1-s_1^*|\leqslant\eta\\ |s_2-s_2^*|\leqslant\eta}}L(s_1,s_2)\biggr). \tag{7.4}$$

By symmetry, it is enough to deal with

$$L_1(s_1,s_2):=\left|\frac{\partial}{\partial s_1}\log D(s_1,s_2)\right|. \tag{7.5}$$

Since the logarithm is an additive function, using the representation (6.9)–(6.11) and (6.23) of $D(s_1,s_2)$, it is sufficient to examine the factors D_0, G_1, G_2 and G_3 separately.

We will choose a positive η such that

$$\eta\leqslant\frac{\frac{1}{100}}{\log_2 N+\log T},\quad T=T_1+T_2,\quad T_j=|t_j|+3,\quad j=1,2, \tag{7.6}$$

where, by $\delta=\delta_1+\delta_2\leqslant 2/\log_2 N$, $V=(\log N)^{1/2}$, $\log\Delta\leqslant K^2\log N\leqslant\log^2 N$, we have

$$\max\{V^\delta,V^\eta,(\log\Delta)^\delta,(\log\Delta)^\eta\}\ll 1. \tag{7.7}$$

We have, by (5.2) and (6.23),

$$\frac{\partial}{\partial s_1}(\log D_0(s_1,s_2))=d\biggl(\frac{\zeta'}{\zeta}(1+s_1+s_2)+\frac{1}{s_1+s_2}\biggr)-a\biggl(\frac{\zeta'}{\zeta}(1+s_1)+\frac{1}{s_1}\biggr)\ll K\log T. \tag{7.8}$$

Further we have

$$\frac{\partial}{\partial s_1}(\log G_1(s_1,s_2))=\sum_{p\leqslant V}\frac{\log p}{p^{1+s_1}}\biggl(\frac{dp^{-s_2}}{1-p^{-(1+s_1+s_2)}}-\frac{a}{1-p^{-(1+s_1)}}\biggr)\ll K\log_2 N. \tag{7.9}$$

Similarly to (6.17), we obtain by (7.7),

$$\begin{aligned}\frac{\partial}{\partial s_1}&(\log G_2(s_1,s_2))\\ &=\sum_{\substack{p\mid\Delta\\ p>V}}\frac{\log p}{p^{1+s_1}}\biggl(\frac{\nu_1(p)-\nu_3(p)p^{-s_2}}{1-\nu_1(p)p^{-(1+s_1)}-\nu_2(p)p^{-(1+s_2)}+\nu_3(p)p^{-(1+s_1+s_2)}}\\ &\qquad\qquad -\frac{a}{1-p^{-(1+s_1)}}+\frac{dp^{-s_2}}{1-p^{-(1+s_1+s_2)}}\biggr)\\ &\ll K\sum_{p\mid\Delta}\frac{\log p}{p^{1-\delta}}\\ &\ll K\sum_{p\leqslant\log\Delta(1+o(1))}\frac{\log p}{p^{1-\delta}}\\ &\ll K\log_2\Delta\\ &\ll K\log_2 N.\end{aligned}\tag{7.10}$$

Finally, analogously to (6.18), we have by (6.15),

$$\begin{aligned}\frac{\partial}{\partial s_1}(\log G_3(s_1,s_2))&=\sum_{\substack{p\nmid\Delta\\ p>V}}\frac{\log p}{p^{1+s_1}}\biggl(\frac{a-dp^{-s_2}}{1-ap^{-(1+s_1)}-bp^{-(1+s_2)}+dp^{-(1+s_1+s_2)}}\\ &\qquad\qquad -\frac{a}{1-p^{-(1+s_1)}}+\frac{dp^{-s_2}}{1-p^{-(1+s_1+s_2)}}\biggr)\\ &\ll\sum_{p>V}\frac{\log p}{p^{1-\delta_1}}\biggl(a\frac{K}{p^{1-\delta}}+dp^{\delta_2}\frac{K}{p^{1-\delta}}\biggr)\\ &\ll K^2\sum_{p>V}\frac{\log p}{p^{2-2\delta}}\\ &\ll\frac{K^2}{V^{1-2\delta}}\\ &\ll\frac{K^2}{V}\\ &\ll K.\end{aligned}\tag{7.11}$$

Summarizing (7.3)–(7.11), we have

$$\max_{\substack{|s_1'-s_1^*|\leqslant\eta\\ |s_2'-s_2^*|\leqslant\eta}}|D(s_1',s_2')|\leqslant e^{C\eta K(\log_2 N+\log T)}|D(s_1^*,s_2^*)|,\quad\text{if } s_1^*,s_2^*,s_3^*\in\mathcal{R}_N'.\tag{7.12}$$

Hence, (7.2) and (7.12) imply by the choice $1/\eta=100K(\log_2 N+\log T)$ the following estimate.

Lemma 4. *We have for* $s_1, s_2, s_3 \in \mathcal{R}'_N$,

$$\frac{1}{j!j'!}\left|\frac{\partial^j}{\partial s_1^j}\frac{\partial^{j'}}{\partial s_2^{j'}}D(s_1,s_2)\right| \ll (CK(\log_2 N + \log T))^{j+j'}|D(s_1,s_2)|. \tag{7.13}$$

The above estimate is sufficient for our purposes at every point (s_1, s_2) apart from $(0,0)$, which will appear in the main term. We will show an analogous result for the point $(0,0)$ where η in (7.6) will be replaced by the larger value

$$\eta_0 = \frac{1}{\bar{d}^* \log_2 N}, \tag{7.14}$$

where we use the notation $\bar{d}$ and $\bar{d}^*$ of (4.4). Next, the following result holds.

Lemma 5. *We have*

$$\frac{1}{j!j'!}\left|\frac{\partial}{\partial s_1^j}\frac{\partial^{j'}}{\partial s_2^{j'}}D(s_1,s_2)\right|_{s_1=s_2=0} \ll (\bar{d}^* \log_2 N)^{j+j'}D(0,0).$$

Proof. Let $d_1 = a-d$. Analogously to (7.8)–(7.9), we have, for $|s_1|, |s_2| \leqslant \eta_0$,

$$\begin{aligned}\frac{\partial}{\partial s_1}\log D_0(s_1,s_2) &= d\frac{W'}{W}(s_1+s_2) - a\frac{W'}{W}(s_1)\\ &= d\left(\frac{W'}{W}(s_1+s_2) - \frac{W'}{W}(s_1)\right) - (a-d)\frac{W'}{W}(s_1)\\ &\ll K\eta_0 + d_1\\ &\ll \frac{K}{\bar{d}^*} + d_1\\ &\ll \sqrt{K} + d_1\\ &\ll \bar{d}^*,\end{aligned} \tag{7.15}$$

and

$$\begin{aligned}\frac{\partial}{\partial s_1}(\log G_1(s_1,s_2)) &= \sum_{p\leqslant V}\frac{\log p}{p^{1+s_1}}\left(\frac{d}{p^{s_2}-p^{-(1+s_1)}} - \frac{d}{1-p^{-(1+s_1)}} + \frac{d-a}{1-p^{-(1+s_1)}}\right)\\ &\ll \sum_{p\leqslant V}\frac{\log p}{p^{1-\delta_1}}\left(\frac{K|p^{s_2}-1|}{p^{-\delta_2}} + d_1\right)\\ &\ll K\sum_{p\leqslant V}\frac{\eta_0 \log^2 p}{p^{1-\delta}} + d_1\sum_{p\leqslant V}\frac{\log p}{p^{1-\delta}}\\ &\ll V^\delta(K\eta_0 \log V + d_1)\log V\\ &\ll K\eta_0 \log_2^2 N + d_1 \log_2 N\\ &\ll \bar{d}^* \log_2 N.\end{aligned} \tag{7.16}$$

The treatment of G_2 will be similar to this and (7.10). By

$$|\nu_1(p) - \nu_3(p)| = |\nu_p(\mathcal{H}_1) - \nu_p(\mathcal{H}_1(p)\cap\mathcal{H}_2(p))| \leqslant |\mathcal{H}_1 \setminus \mathcal{H}_2| = a-d,$$

we have

$$
\begin{aligned}
&\frac{\partial}{\partial s_1}(\log G_2(s_1,s_2))\\
&\quad=\sum_{\substack{p\mid\Delta\\ p>V}}\frac{\log p}{p^{1+s_1}}\Bigg(\frac{\nu_1(p)-\nu_3(p)+\nu_3(p)(1-p^{-s_2})}{1-\nu_1(p)p^{-(1+s_1)}-\nu_2(p)p^{-(1+s_2)}+\nu_3(p)p^{-(1+s_1+s_2)}}\\
&\qquad\qquad-\frac{d_1}{1-p^{-(1+s_1)}}-d\bigg(\frac{1}{1-p^{-(1+s_1)}}-\frac{1}{p^{s_2}-p^{-(1+s_1)}}\bigg)\Bigg)\\
&\quad\ll\sum_{\substack{p\mid\Delta\\ p>V}}\frac{\log p}{p^{1-\delta}}(d_1+K(p^{\eta_0}-1))\\
&\quad\ll d_1\sum_{p\mid\Delta}\frac{\log p}{p^{1-\delta}}+K\eta_0\sum_{\substack{p\mid\Delta\\ p\leqslant e^{2/\eta_0}}}\frac{\log^2 p}{p^{1-\delta}}+K\sum_{p\mid\Delta}\frac{\log p}{e^{1/\eta_0}} \qquad (7.17)\\
&\quad\ll d_1\sum_{p\leqslant(1+o(1))\log\Delta}\frac{\log p}{p^{1-\delta}}+K\eta_0\sum_{p\leqslant(1+o(1))\log\Delta}\frac{\log^2 p}{p^{1-\delta}}+\frac{K\log\Delta}{(\log N)^{\bar d^*}}\\
&\quad\ll d_1\log_2\Delta+K\eta_0\log_2^2\Delta+\frac{1}{(\log N)^{\sqrt{K}-3}}\\
&\quad\ll (d_1+K\eta_0\log_2 N)\log_2 N\\
&\quad\ll \bar d^*\log_2 N.
\end{aligned}
$$

Finally we have, similarly to above and (7.11), using $a-dp^{-s_2}=a-d+d\big(1-p^{-s_2}\big)$,

$$
\begin{aligned}
&\frac{\partial}{\partial s_1}\log G_3(s_1,s_2)\\
&\quad=\sum_{\substack{p\nmid\Delta\\ p>V}}\frac{\log p}{p^{1+s_1}}\Bigg(a\bigg(\frac{1}{1-p^{-(1+s_1+s_2)}}-\frac{1}{1-p^{-(1+s_1)}}\bigg)\\
&\qquad+(a-dp^{-s_2})\bigg(\frac{1}{1-ap^{-(1+s_1)}-bp^{-(1+s_2)}+dp^{-(1+s_1+s_2)}}-\frac{1}{1-p^{-(1+s_1+s_2)}}\bigg)\Bigg)\\
&\quad\ll\sum_{p>V}\frac{\log p}{p^{1-\delta_1}}\bigg(\frac{K}{p^{1-\delta}}+\frac{K}{p^{1-\delta}}(d_1+K(p^{\eta_0}-1))\bigg)\\
&\quad\ll Kd_1\sum_{p>V}\frac{\log p}{p^{2-2\delta}}+K^2\eta_0\sum_{V<p\leqslant e^{1/\eta_0}}\frac{\log^2 p}{p^{2-2\delta}}+K^2\sum_{p>e^{1/\eta_0}}\frac{\log p}{p^{2-2\delta-\eta_0}} \qquad (7.18)\\
&\quad\ll\frac{Kd_1}{V^{1-2\delta}}+\frac{K^2\eta_0\log V}{V^{1-2\delta}}+\frac{K^2}{(\log N)^{(1-2\delta-\eta_0)\bar d^*}}\\
&\quad\ll\frac{K}{V}(d_1+K\eta_0\log_2 N)+o(1)\\
&\quad\ll d_1+K\eta_0\log_2 N\\
&\quad\ll \bar d^*.
\end{aligned}
$$

Now, (7.15)–(7.18) imply, by symmetry for $j=1,2$, that

$$\left|\frac{\partial}{\partial s_j}\log D(s_1,s_2)\right| \ll \frac{1}{\eta_0} \quad \text{for } |s_1|,|s_2|\leqslant \eta_0, \tag{7.19}$$

and therefore, similarly to (7.12), we have

$$\max_{\substack{|s_1'|\leqslant\eta_0\\ |s_2'|\leqslant\eta_0}} |D(s_1',s_2')| \ll D(0,0), \tag{7.20}$$

which by (7.2) proves Lemma 5. □

8. Evaluation of the integral I_1

This section will be devoted to the examination of $I_{1,1}$, the sum of the residues in (6.29).

The rather complicated formula (6.29) yields the main term and all secondary terms of the form $(\log R)^m$ exclusively for $m\in[d,d+u+v-1]$ and will additionally contribute to other secondary terms for $m\in[0,d-1]$. However, from the terms $I_{1,1}(j,j',\nu)$ belonging to the triplet (j,j',ν) in the triple summation, only those with $\nu=0$ and $j'=0$ contribute to the main term of order $(\log R)^{d+u+v}$, since in all other terms the exponent of $\log R$ is $d+u+v-j'-\nu$.

We now have to work more carefully than in [7]. For example, by the aid of Lemma 2 (a generalization of (8.16) in [7]) we will exactly evaluate the coefficients $A_{j',\nu}$ of

$$\frac{1}{j'!\nu!}\frac{\partial^\nu}{\partial s_2^\nu}\frac{\partial^{j'}}{\partial s_1^{j'}}D(s_1,s_2)(\log R)^{v+d+u-j'-\nu}$$

in (6.29) as follows. Let $j',\nu\geqslant 0$,

$$m:=j-j'\geqslant 0 \quad\text{and}\quad y:=v+d-\nu, \tag{8.1}$$

where we may assume, by (6.29), that

$$\nu\leqslant v+d+m, \quad\text{that is}\quad m\geqslant \nu-v-d=-y. \tag{8.2}$$

Then we have from (6.29), by notation (5.6), (8.1) and Lemma 2,

$$\begin{aligned} A_{j',\nu} &= \frac{j'!\nu!}{u!}\sum_{\substack{m=0\\ m\geqslant -y}}^{u-j'}\binom{u}{m+j'}(-1)^m\binom{m+j'}{j'}\frac{d(d+1)\dots(d+m-1)}{(v+d+m-\nu)!\nu!}\\ &= \sum_{\substack{m=0\\ m\geqslant -y}}^{u-j'}\frac{(-1)^m}{(u-j'-m)!m!}\frac{d(d+1)\dots(d+m-1)}{(v+d+m-\nu)!}\\ &= Z(d,u-j,v+d-\nu)\\ &= \frac{(v-\nu+1)\dots(v-\nu+u-j')}{(u-j')!(d+v-\nu+u-j')!}. \end{aligned} \tag{8.3}$$

We have to compare $A_{j',\nu}$ with $A_{0,0}$. This will be furnished by the following lemma.

LEMMA 6. *We have*

$$|A'_{j',\nu}| := \left|\frac{A_{j',\nu}}{A_{0,0}}\right| \leqslant (CK)^{j'+\nu}.$$

Proof. Recalling (6.24), we have

$$|A'_{j',\nu}| = \frac{(d+v+u)!}{(d+v+u-\nu-j')!}\,\frac{(u-j'+1)\dots u}{(v+u-j'+1)\dots(v+u)}|A''_{j',\nu}| \leqslant (CK)^{j'+\nu}|A''_{j',\nu}|,$$

where

$$|A''_{j',\nu}| = \frac{|(v-\nu+1)\dots(v-\nu+u-j')|}{(v+1)\dots(v+u-j')}. \tag{8.4}$$

If $\nu \leqslant 2(v+1)$, then clearly $|A''_{j',\nu}| \leqslant 1$, so we may suppose that

$$\nu = B(v+1), \quad B > 2. \tag{8.5}$$

In this case we have, by $u-j' \leqslant u \leqslant v < v+1$,

$$|A''_{j',\nu}| \leqslant \left(\frac{\nu}{v+1}\right)^{u-j'} \leqslant B^{v+1} = B^{\nu/B} < 2^{\nu}, \tag{8.6}$$

since the maximum of $x^{1/x}$ in $[1,\infty)$ is attained at $x=e$ and $e^{1/e} < 2$. □

Now we are ready to evaluate the crucial term $I_{1,1}$ by the aid of Lemmas 2, 5 and 6. Namely, by (4.1) and (4.4), $R \gg N^c$, (6.29), (8.3) and (6.24), we have

$$\begin{aligned} I_{1,1} &= A_{0,0}(\log R)^{d+u+v}\left(D(0,0) + \sum_{j'=0}^{u}\sum_{\substack{\nu=0\\ j'+\nu\geqslant 1}}^{v+d+u-j'} \frac{A'_{j',\nu}}{(\log R)^{j'+\nu}}\frac{1}{j'!\nu!}\frac{\partial^{j'}}{\partial s_1^{j'}}\frac{\partial^{\nu}}{\partial s_2^{\nu}}D(0,0)\right)\\ &= Z(d,u,v+d)(\log R)^{d+u+v}D(0,0)\left(1+O\left(\sum_{j'=0}^{\infty}\sum_{\substack{\nu=0\\ j'+\nu\geqslant 1}}^{\infty}\left(\frac{K\bar{d}^*\log_2 N}{\log R}\right)^{j'+\nu}\right)\right)\\ &= \frac{1}{(d+v+u)!}\binom{v+u}{u}(\log R)^{d+v+u}D(0,0)\left(1+O\left(\frac{K\bar{d}^*\log_2 N}{\log R}\right)\right). \end{aligned}\tag{8.7}$$

The integral $I_{1,2}$ in (6.29) does not contribute to the main term and can be estimated relatively easily due to the presence of the term R^{s_2} $(s_2 \in \mathcal{L}_4)$. In fact, choosing

$$\frac{1}{\eta} = 100(\log_2 N + \log T)$$

as earlier, we obtain by Lemma 4, (5.2), (6.20) and (6.26)–(6.28) for any $s_2 \in \mathcal{L}_4$,

$$\begin{aligned} Z(s_2) &\ll \sum_{j=0}^{u}\sum_{j'=0}^{j}\frac{(\log R)^{u-j}(CK)^{j-j'}}{(u-j)!(j-j')!}(CK(\log_2 N+\log T))^{j'}\frac{|D(0,s_2)|R^{\sigma_2}}{|s_2|^{d+j-j'+v+1}}\\ &\ll e^{C(\sqrt{K}\log_2 N + K\log_3 N)}\frac{(\log(|t_2|+3))^{3K+O(\sqrt{K})}R^{\sigma_2}}{|s_2|^{b+O(\sqrt{K})}}. \end{aligned}\tag{8.8}$$

Now Lemma 1 yields immediately, by (6.24) and $R \gg N^c$,

$$I_{1,2} = \frac{1}{2\pi i}\int_{\mathcal{L}_4} Z(s_2)\,ds_2 \ll e^{C(\sqrt{K}\log_2 N + K\log_3 N) - c\sqrt{\log N}} \ll e^{-c\sqrt{\log N}}. \tag{8.9}$$

We may summarize (8.7) and (8.9), by $D(0,0) = D_0(0,0)G(0,0) = G(0,0) \neq 0$ (which is true by the admissibility condition), by (6.9)–(6.11) and (6.23), as the following.

LEMMA 7. *The integral I_1 in* (6.26) *satisfies the asymptotic*

$$I_1 = \frac{1}{(d+v+u)!}\binom{v+u}{u}(\log R)^{d+v+u}G(0,0)\left(1+O\left(\frac{K\bar{d}^*\log_2 N}{\log R}\right)\right) + O\left(e^{-c\sqrt{\log N}}\right). \tag{8.10}$$

9. Estimate of the integral I_2

For I_2 in (6.27), after interchange of the two integrations, we move the contour $\mathcal{L}_2$ for the inner integral over s_2 to the left to $\mathcal{L}_4$ passing a pole of order d at $s_2 = -s_1$ if $|t_2| \leqslant U$ and a pole of order $v+1$ at $s_2 = 0$, and obtain

$$\begin{aligned} I_2 &= \frac{1}{2\pi i}\int_{\mathcal{L}_3} \operatorname*{Res}_{s_2=-s_1}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) ds_1 \\ &\quad + \frac{1}{2\pi i}\int_{\mathcal{L}_3} \operatorname*{Res}_{s_2=0}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) ds_1 \\ &\quad + \frac{1}{(2\pi i)^2}\int_{\mathcal{L}_4}\int_{\mathcal{L}_3} F(s_1,s_2)\frac{R^{s_1}}{{s_1}^{a+u+1}}\frac{R^{s_2}}{{s_2}^{b+v+1}}\,ds_1\,ds_2 + O\left(e^{-c\sqrt{\log N}}\right) \\ &=: I_{2,1} + I_{2,2} + I_{2,3} + O\left(e^{-c\sqrt{\log N}}\right). \end{aligned} \tag{9.1}$$

By the argument of Lemma 1 and (6.21), the third integral $I_{2,3}$ is $\ll e^{-c\sqrt{\log N}}$. The second integral $I_{2,2}$ is completely analogous to $I_{1,2}$ in (6.29), which was estimated by $e^{-c\sqrt{\log N}}$ in (8.9), the only change being that the roles of s_1 and s_2 are interchanged.

The residue in $I_{2,1}$ is zero if $d=0$, while for $d \geqslant 1$ we have

$$\begin{aligned} \operatorname*{Res}_{s_2=-s_1}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right) &= \lim_{s_2\to -s_1}\frac{1}{(d-1)!}\frac{\partial^{d-1}}{\partial s_2^{d-1}}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}}\right) \\ &= \frac{1}{(d-1)!}\sum_{j=0}^{d-1}\mathcal{B}_j(s_1,\mathcal{H}_1,\mathcal{H}_2)(\log R)^{d-1-j'}, \end{aligned} \tag{9.2}$$

where

$$\mathcal{B}_{j'}(s_1,\mathcal{H}_1,\mathcal{H}_2) = \binom{d-1}{j'}\sum_{\nu=0}^{j'}\binom{j'}{\nu}\frac{\partial^{j'-\nu}}{\partial s_2^{j'-\nu}}D(s_1,s_2)\bigg|_{s_2=-s_1}\frac{(-1)^\nu(v+1)\dots(v+\nu)}{(-1)^{\nu+v+1}s_1^{u+v+\nu+2}}. \tag{9.3}$$

We thus obtain

$$I_{2,1} = \frac{1}{(d-1)!} \sum_{j'=0}^{d-1} \mathcal{C}_{j'}(\mathcal{H}_1, \mathcal{H}_2)(\log R)^{d-1-j'}, \tag{9.4}$$

where

$$\mathcal{C}_{j'}(\mathcal{H}_1, \mathcal{H}_2) = \frac{1}{2\pi i} \int_{\mathcal{L}_3} \mathcal{B}_{j'}(s_1, \mathcal{H}_1, \mathcal{H}_2)\, ds_1, \quad j' = 0, 1, 2, ..., d-1. \tag{9.5}$$

It remains to estimate these quantities, which are independent of R.

We are allowed to transform the contour $\mathcal{L}_3$ in (9.5) to the contour $\mathcal{L}'$, defined in (5.4). Our task is now the estimation of the integral $\mathcal{C}_{j'}$ on the new contour $\mathcal{L}' = \mathcal{L}'_0 \cup \mathcal{L}'_1$, since the integral on the horizontal segments $|t| = U$ is $O\big(e^{-c\sqrt{\log N}}\big)$.

10. Comparison of $D(s, -s)$ and $D(0, 0)$

We have seen in §7 that by Lemma 4 we can estimate

$$\frac{\partial^j}{\partial s_1^j} \frac{\partial^{j'}}{\partial s_2^{j'}} D(s_1, s_2)$$

with the aid of $D(s_1, s_2)$. We will show now how to estimate $|D(s, -s)|/D(0, 0)$ from above when s is on the contour $\mathcal{L}'$, which, together with Lemma 10, will play a crucial role in the estimation of $I_{2,1}$, which is the main part of I_2.

First we note that if $s \in \mathcal{L}'$ is on the semicircle $\mathcal{L}'_0$, then by (7.12) we obtain

$$|D(s, -s)| \leqslant e^{C\sqrt{K}} D(0, 0), \quad s \in \mathcal{L}'_0. \tag{10.1}$$

Thus, in the following, we may suppose that

$$s = it, \quad t > 0, \tag{10.2}$$

since $|D(-it, it)| = |D(it, -it)|$. Our main results in this section are the following estimates.

LEMMA 8. *We have, for any real t,*

$$|D(it, -it)| \leqslant \max\{1, |t|\}^{-(a+b)/2} D(0, 0).$$

Together with (10.1), this implies the following result.

LEMMA 9. *We have, for $s \in \mathcal{L}'$,*

$$|D(s, -s)| \leqslant e^{C\sqrt{K}} \max\{1, |t|\}^{-(a+b)/2} D(0, 0),$$

where $\mathcal{L}'$ was defined in (5.4).

From (6.9) and (6.23) we have $D=D_0G_1G_4$. We first examine the behavior of the functions $D_0(s,-s)$ and $G_1(s,-s)$ on the imaginary axis, which requires a lemma concerning $W(s)$ from (6.23).

LEMMA 10. *There exist positive absolute constants t_0 and t_1, with $t_1>1$, such that*

$$|W(it)|\geqslant e^{t^2/6}\geqslant 1=W(0) \quad \textit{for } |t|\leqslant t_0, \tag{10.3}$$

$$|W(it)|\geqslant t^{2/3} \quad \textit{for } |t|\geqslant t_1. \tag{10.4}$$

Proof. We will use that in a neighborhood of $s=0$ we have, for the entire function $W(s)$, the representation

$$W(s)=1+\gamma_0 s+\sum_{\nu=1}^{\infty}\gamma_\nu s^{\nu+1}, \tag{10.5}$$

where $\gamma_0=\gamma$ is Euler's constant and (see [12, notes on p. 49])

$$\gamma_0=\gamma=0.5772157\ldots, \quad \gamma_1=0.07281\ldots. \tag{10.6}$$

This implies that

$$\begin{aligned}|W(it)|^2&=W(it)W(-it)\\&=(1+i\gamma t-\gamma_1t^2+O(t^3))(1-i\gamma t-\gamma_1t^2+O(t^3))\\&=1+t^2(\gamma^2-2\gamma_1)+O(t^3),\end{aligned} \tag{10.7}$$

as $t\to 0$. Now (10.6)–(10.7) prove (10.3) for $|t|\leqslant t_0$, while (10.4) clearly holds by (5.2). □

Remark. If $|W(it)|\geqslant 1$ for any t (which could be checked by computers, since t_0 and t_1 are explicitly calculable), then the following simple lemma is not necessary.

LEMMA 11. *Given any positive constants B_0, B_1 and ε, we have, for any $t\in[B_0,B_1]$ and any $X>C(B_0,B_1,\varepsilon)$,*

$$J(t,X):=\prod_{p\leqslant X}\frac{|1-p^{-1-it}|}{1-p^{-1}}\geqslant c(B_0,B_1)(\log X)^{1/2-\varepsilon}. \tag{10.8}$$

Proof. Fix t. Since every factor is at least 1, we can neglect those with $\cos(t\log p)>0$. On the other hand, if $\cos(t\log p)\leqslant 0$, then we have

$$\log\left(\frac{|1-p^{-1-it}|}{1-p^{-1}}\right)>\log\frac{1}{1-p^{-1}}>\frac{1}{p}. \tag{10.9}$$

The primes satisfying $\cos(t\log p)\leqslant 0$ are in intervals of the type

$$I_j=[e^{\pi(2j+1/2)/t},e^{\pi(2j+3/2)/t}]=:[e^{m_j},e^{m_j+\pi/t}] \tag{10.10}$$

and $I_j \subset [1, X]$ if $2\pi\bigl(j+\frac{3}{4}\bigr)/t \leqslant \log X$, that is, if

$$j \leqslant \frac{t\log X}{2\pi} - \frac{3}{4} =: j^*. \tag{10.11}$$

Using the prime number theorem, we obtain by partial summation

$$\sum_{p\in I_j} \frac{1}{p} \sim \int_{e^{m_j}}^{e^{m_j+\pi/t}} \frac{dx}{x\log x} = \log \frac{2j+\frac{3}{2}}{2j+\frac{1}{2}} = \frac{1}{2j} + O\biggl(\frac{1}{j^2}\biggr). \tag{10.12}$$

Hence, by (10.9), we have

$$\log J(t, X) > \sum_{1\leqslant j\leqslant j^*} \frac{1-\varepsilon}{2j} + O(1) > \frac{1-\varepsilon}{2} \log_2 X - c'(B_0, B_1), \tag{10.13}$$

which completes the proof. □

Remark. Working more carefully, we could prove Lemma 11 with $(\log X)^{1/2-\varepsilon}$ replaced by $\log X$, but the result we just obtained suffices for our needs.

Taking into account the trivial relation

$$\frac{1}{|1-p^{-1-it}|} \leqslant \frac{1}{|1-p^{-1}|}, \tag{10.14}$$

we obtain, from Lemmas 10 and 11, the following result.

LEMMA 12. *We have, with a sufficiently small constant* $t_0 < 1$,

$$E_0(t) := \left| \frac{D_0(it, -it)G_1(it, -it)}{D_0(0,0)G_1(0,0)} \right| \leqslant e^{-(a+b)t^2/6}, \quad \textit{if } |t| \leqslant t_0, \tag{10.15}$$

and for any $|t| > t_0$ *and* $N > N_0$,

$$E_0(t) \leqslant e^{-(a+b)} \max\{1, |t|\}^{-(a+b)/2}. \tag{10.16}$$

Proof. By (10.14) and the definition of D_0 in (6.23), we clearly have

$$|G_1(it, -it)| \leqslant G_1(0,0), \tag{10.17}$$

and

$$|D_0(it, -it)| = |W(it)|^{-(a+b)}, \quad D_0(0,0) = W(0) = 1, \tag{10.18}$$

which immediately imply

$$E_0(t) \leqslant |W(it)|^{-(a+b)}. \tag{10.19}$$

Hence, by Lemma 10, we have (10.15), and (10.16) for $|t|\geqslant t_1$, where t_1 is a sufficiently large constant. (Actually $t_1>e^6$ is sufficient.) Letting

$$C(t_0,t_1):=\max_{t_0\leqslant t\leqslant t_1}\frac{1}{|W(it)|}, \tag{10.20}$$

we have, by Lemma 11,

$$\begin{aligned}E_0(t)=(J(|t|,V)|W(it)|)^{-(a+b)}&\leqslant\left(\frac{c(t_0,t_1)(\log V)^{1/3}}{C(t_0,t_1)}\right)^{-(a+b)}\\&\leqslant(c\log_2 N)^{-(a+b)/3}\leqslant(et_1)^{-(a+b)},\end{aligned} \tag{10.21}$$

which proves (10.16). □

We will continue our study of $D(it,-it)$ with that of

$$L_4(t):=\log\frac{|G_4(it,-it)|}{G_4(0,0)}=\operatorname{Re}\log\frac{G_4(it,-it)}{G_4(0,0)}. \tag{10.22}$$

We first divide each term by $(1+\nu_3(p)/p)(1-1/p)^d$ in the product representation (6.11) of both $G_4(it,-it)$ and $G_4(0,0)$. After this, we take the logarithm of each term and use the formula

$$\log(1-z)=-\left(z+\sum_{m=2}^{\infty}\frac{z^m}{m}\right),\quad \text{if } |z|<1. \tag{10.23}$$

We now separate the effect of the linear terms and those of order $m\geqslant 2$ and write accordingly

$$L_4(t)=L_{4,1}(t)+L_{4,2}(t). \tag{10.24}$$

We have, by the trivial relations $\nu_1(p)\leqslant a$ and $\nu_2(p)\leqslant b$,

$$L_{4,1}(t)=\sum_{p>V}\left(\frac{a+b}{p}-\frac{\nu_1(p)+\nu_2(p)}{p(1+\nu_3(p)/p)}\right)(\cos(t\log p)-1)\leqslant 0. \tag{10.25}$$

(We remark that the sum is convergent, since $\nu_1(p)=a$ and $\nu_2(p)=b$ for $p\nmid\Delta$.)

The contribution of the higher-order terms of G_4 to $L_{4,2}$ which do not involve the functions $\nu_j(p)$ are majorized for any t by

$$(a+b)\sum_{p>V}\sum_{m=2}^{\infty}\frac{2}{mp^m}\leqslant\frac{C(a+b)}{V\log V}\leqslant\frac{C}{\log_2^3 N}, \tag{10.26}$$

while for small t they are majorized by

$$C(a+b)t^2\sum_{p>V}\frac{\log^2 p}{p^2}\leqslant\frac{Ct^2}{\log_2 N}. \tag{10.27}$$

Similarly, using (3.11), (4.1) and (6.24), the contribution of the higher-order terms of G_4 to $L_{4,2}$ which involve the functions $\nu_j(p)$ are majorized for any t by

$$\sum_{p>V}\sum_{m=2}^{\infty}\frac{2(a+b)^m}{mp^m}\leqslant\frac{C(a+b)^2}{V\log V}\leqslant\frac{C(a+b)}{\log_2^3 N},\tag{10.28}$$

while for small values of t they are majorized by

$$\begin{aligned}\sum_{p>V}\sum_{m=2}^{\infty}\sum_{j'=0}^{m}\binom{m}{j'}&\frac{\nu_1(p)^{j'}\nu_2(p)^{m-j'}}{m(p+\nu_3(p))^m}(1-\cos((m-2j')t\log p))\\&\leqslant C\sum_{p>V}\sum_{m=2}^{\infty}\frac{(a+b)^m}{mp^m}t^2m^2\log^2 p\\&\leqslant C(a+b)^2t^2\sum_{p>V}\frac{\log^2 p}{p^2}\\&\leqslant\frac{C(a+b)^2t^2\log V}{V}\\&\leqslant\frac{C(a+b)t^2}{\log_2 N}.\end{aligned}\tag{10.29}$$

Summarizing (10.22)–(10.29), we have proved the following result.

LEMMA 13. *We have*

$$\frac{|G_4(it,-it)|}{G_4(0,0)}\leqslant e^{C(a+b)\min\{1,t^2\}/\log_2 N}.$$

Comparing the above with (10.15)–(10.16), we see that (10.15)–(10.16) remain valid if we multiply them by $G_4(it,-it)/G(0,0)$. This proves Lemma 8.

11. Estimate of $I_{2,1}$ and evaluation of I

In this section we will estimate the integral $I_{2,1}$ based on formulas (9.3)–(9.5), using Lemmas 4 and 9.

We first obtain from the above lemmas for $s\in\mathcal{L}'$, as $j'\leqslant d\leqslant\min\{a,b\}\leqslant K$, $v\leqslant\sqrt{K}$ and $a+b\gg K$,

$$\begin{aligned}\mathcal{B}_{j'}(s,\mathcal{H}_1,\mathcal{H}_2)&\ll\frac{d^{j'}}{|s|^{u+v+2}}|D(s,-s)|\sum_{\nu=0}^{j'}(CK(\log_2 N+\log T))^{j'-\nu}\prod_{j=1}^{\nu}\frac{1+v/j}{|s|}\\&\ll\frac{e^{C\sqrt{K}}D(0,0)d^{j'}(\log(|t|+3))^{j'}}{\max\{1,\sqrt{|t|}\}^{a+b}}\frac{\delta_0^{-(u+v)}}{|s|^2}\sum_{\nu=0}^{j'}(CK\log_2 N)^{j'-\nu}(K\log_2 N)^{\nu}\\&\ll e^{C\sqrt{K}}(CK^2\log_2 N)^{j'}\frac{\delta_0^{-(u+v)}}{|s|^2}D(0,0).\end{aligned}\tag{11.1}$$

Integrating the above upper bound along $\mathcal{L}'$, we obtain

$$\mathcal{C}_{j'}(\mathcal{H}_1,\mathcal{H}_2) \ll e^{C\sqrt{K}}(CK^2\log_2 N)^{j'}\delta_0^{-(u+v+1)}D(0,0). \tag{11.2}$$

Finally, summation over $j'\leqslant d-1$ yields in (9.4), as $R\gg N^c$ and $u+v\gg\sqrt{K}$,

$$\begin{aligned} I_{2,1} &\ll \frac{e^{C\sqrt{K}}D(0,0)\delta_0^{-(u+v+1)}(\log R)^{d-1}}{(d-1)!}\sum_{j'=0}^{d-1}\left(\frac{CK^2\log_2 N}{\log R}\right)^{j'} \\ &\ll \frac{e^{C(u+v)}D(0,0)(\log R)^{d-1}\big(\sqrt{K}\log_2 N\big)^{u+v+1}}{(d-1)!} \\ &\ll \frac{D(0,0)(\log R)^{d+u+v}}{(d+u+v)!}\left(\frac{CK^{3/2}\log_2 N}{\log R}\right)^{u+v+1} \\ &\ll \frac{D(0,0)(\log R)^{d+u+v}}{(d+u+v)!}(\log N)^{-\sqrt{K}/50}. \end{aligned} \tag{11.3}$$

This implies, by (9.1) and (9.4), that

$$I_2 \ll \frac{D(0,0)(\log R)^{d+u+v}}{(d+u+v)!}(\log N)^{-\sqrt{K}/50}+e^{-c\sqrt{\log N}}. \tag{11.4}$$

This yields, by Lemma 7, the final asymptotic evaluation of I in (6.22), since $D(0,0)=G(0,0)$, as

$$I=\frac{1}{(d+v+u)!}\binom{v+u}{u}(\log R)^{d+v+u}G(0,0)\left(1+O\left(\frac{K\bar{d}^*\log_2 N}{\log R}\right)\right)+O\big(e^{-c\sqrt{\log N}}\big), \tag{11.5}$$

where

$$G(0,0)=\prod_{p\nmid P}\left(1-\frac{\nu_p(\mathcal{H})}{p}\right)\prod_{p}\left(1-\frac{1}{p}\right)^{-|\mathcal{H}|}=\frac{\mathfrak{S}(\mathcal{H})P}{|A(\mathcal{H})|}, \tag{11.6}$$

thereby proving Theorem 4.

12. A Bombieri–Vinogradov type theorem

In the present section we will prove a modified Bombieri–Vinogradov theorem, where the examined moduli are all multiples of a single modulus M. It would facilitate our task if we were entitled to use the following hypothesis.

Hypothesis $S(Y)$. If $L(1-\delta,\chi)=0$ for $\delta>0$ and a real primitive character $\chi\,(\mathrm{mod}\,q)$, $q\leqslant Y$, then

$$\delta \geqslant \frac{1}{3\log Y} \tag{12.1}$$

for an explicitly calculable absolute constant $Y>C_0$.

We note that we have the effective unconditional estimate ([6], [15]), valid for $q>q_0$:

$$\delta \geqslant \frac{1}{\sqrt{q}}. \tag{12.2}$$

A further observation (similar to that of Maier [13]) is that by the Landau–Page theorem (cf. Davenport [2, §14], with some constant c in place of $\frac{1}{3}$, or Pintz [16] with $\left(\frac{1}{2}+o(1)\right)$) for any given Y there is at most one real primitive character χ_1 which does not fulfill (12.1). This makes it possible to turn hypothesis $S(Y)$ into a theorem, valid for a sequence $Y=Y_n\to\infty$ (for $n>n_0$, an explicitly calculable absolute constant) with

$$Y_n \leqslant e^{\sqrt{Y_{n-1}}}. \tag{12.3}$$

In order to show this, suppose that (12.1) is false for a sufficiently large Y', i.e. by (12.2) there exists a character $\chi_1 \bmod q_1 \leqslant Y'$ such that $L(1-\delta_1,\chi_1)=0$ with

$$\frac{1}{\sqrt{Y'}} \leqslant \min\left\{\frac{1}{\sqrt{q_1}}, c_0\right\} \leqslant \delta_1 < \frac{1}{3\log Y'}. \tag{12.4}$$

Let us choose $\widetilde{Y}>Y'$ in such a way that

$$\widetilde{Y} = e^{1/3\delta_1}, \quad \text{that is} \quad \delta_1 = \frac{1}{3\log \widetilde{Y}}. \tag{12.5}$$

Then, for any other zero $1-\delta_2$ belonging to a real primitive character $\chi_2 \bmod q_2$, $q_2\leqslant\widetilde{Y}$, we have by the Landau–Page theorem in the version of Pintz [16],

$$\max\{\delta_1,\delta_2\} > \frac{1}{3\log\widetilde{Y}}, \quad \text{that is} \quad \delta_2 > \frac{1}{3\log\widetilde{Y}}. \tag{12.6}$$

Now, (12.4)–(12.6) show that (12.1) is true for a value $Y=\widetilde{Y}$ satisfying

$$Y' < \widetilde{Y} < e^{\sqrt{Y'}/3}. \tag{12.7}$$

We can formulate this in the following way.

Lemma 14. *Hypothesis $S(Y)$ holds for a sequence $Y_n\to\infty$ with*

$$Y_n \leqslant e^{\sqrt{Y_{n-1}}}, \tag{12.8}$$

where Y_0 can be chosen with $Y_0<C_0$, an explicitly calculable absolute constant.

An alternative to this lemma and this approach would be to use Heath-Brown's theorem [10] (but only in case of Theorem 1) according to which either

(i) $S(Y)$ holds for every $Y>C$, with some absolute constant C, or

(ii) there are infinitely many twin primes.

The significance of the real zeros in hypothesis $S(Y)$ is that a similar inequality holds with $\operatorname{Re}\varrho$ in place of $1-\delta$ and with a constant c_0 in place of $\frac{1}{3}$ if $\operatorname{Im}\varrho$ is not too large; this is the standard zero-free region of L-functions (cf. Davenport [2, §14]).

Lemma 15. *There exists an explicitly calculable absolute constant* $c_0<\frac{1}{3}$ *such that* $L(s,\chi)\neq 0$ *in the region*

$$\sigma>1-\frac{c_0}{\log q(|t|+3)}, \tag{12.9}$$

apart from possible real exceptional zeros of real L-functions.

Now we are in a position to formulate and prove the following theorem (which is similar to but stronger than Lemma 6 of Maier [13]).

Theorem 6. *Let c^* be an arbitrary, fixed positive constant. Let $Y=Y(X)$ be a strictly monotonically increasing continuous function of X with*

$$e^{2\sqrt{\log X}}\leqslant Y(X)\leqslant X. \tag{12.10}$$

Then there exists a sequence $X_n\to\infty$ satisfying $X_1<C_0'$, an explicitly calculable constant, with the following property. Let $X=X_n$, $\mathcal{L}=\log X$, $M\leqslant\min\{\frac{1}{4}\sqrt{Y(X)},X^{1/8}\}$ be a natural number,

$$Q^*=X^{1/2}M^{-3}e^{-c^*\sqrt{\log X}}, \tag{12.11}$$

and

$$E^*(X,q):=\max_{x\leqslant X}\max_{(a,q)=1}|E(X,q,a)|:=\max_{x\leqslant X}\max_{(a,q)=1}\Bigg|\sum_{\substack{p\leqslant x\\ p\equiv a\ (\mathrm{mod}\ q)}}\log p-\frac{x}{\varphi(q)}\Bigg|. \tag{12.12}$$

Then

$$\sum_{\substack{q\leqslant Q^*\\ (q,M)=1}}E^*(X,Mq)\ll\frac{X}{M}\mathcal{L}^{15}e^{-c_2\log X/\log Y(X)}, \tag{12.13}$$

where $c_2=\min\{\frac{1}{6}c^*,\frac{1}{4}c_0\}$.

Remark. The above theorem holds with any X, for which $S(Y(X))=S(Y)$ is true, i.e.

$$L(s,\chi)\neq 0\quad\text{for } s\in\left(1-\frac{1}{3\log Y},1\right] \tag{12.14}$$

holds without exception for all real primitive characters $\chi \bmod q$, where

$$q\leqslant Y=Y(X). \tag{12.15}$$

Proof. We will choose our sequence $X_n = Y^{-1}(Y_n)$, where Y_n is the sequence supplied by Lemma 14 (for which $S(Y)$ is true) and Y^{-1} is the inverse function of $Y(X)$. Alternatively, if (12.14)–(12.15) hold, then we can choose X as an arbitrary sufficiently large number. In both cases $S(Y)$, i.e. (12.14)–(12.15), hold. Using the explicit formula for primes in arithmetic progressions with $T^* = \sqrt{X}\log^2 X$ ($\varrho=\beta+i\gamma=1-\delta+i\gamma$ denotes a generic zero of an L-function) we obtain (cf. Davenport [2, §19]) for any a with $(a,q)=1$, $q \leqslant Q^*$ and $y \leqslant X$ the relation

$$E(y,q,a) = -\frac{1}{\varphi(q)} \sum_{\chi(q)} \bar{\chi}(a) \sum_{\substack{\varrho=\varrho_\chi \\ \beta\geqslant 1/2 \\ |\gamma|\leqslant T^*}} \frac{y^\varrho}{\varrho} + O(\mathcal{L}^2\sqrt{y}). \tag{12.16}$$

The effect of the last error term is clearly suitable, $O\big(Q^*\mathcal{L}^2\sqrt{X}\,\big)$ in total. We can classify zeros of all primitive L-functions mod $q\widetilde{M} \leqslant Q^*M$ ($\widetilde{M} \mid M$) up to height T^* into $O(\mathcal{L}^4)$ classes $B(\varkappa,\lambda,\mu,\nu)$ by Lemma 15, as

$$\widetilde{M} \in \left[\frac{M_\lambda}{2}, M_\lambda\right), \quad q \in \left[\frac{Q_\nu}{2}, Q_\nu\right), \quad \gamma \in \left[\frac{T_\mu}{2}, T_\mu\right), \quad \delta \in \left[\frac{\varkappa c_0}{\mathcal{L}}, \frac{(\varkappa+1)c_0}{\mathcal{L}}\right), \tag{12.17}$$

where

$$M_\lambda = 2^\lambda \leqslant 2M, \quad Q_\nu = 2^\nu \leqslant 2Q^*, \quad T_\mu = 2^\mu \leqslant 2T^* \quad \text{and} \quad \frac{\varkappa c_0}{\mathcal{L}} \leqslant \frac{1}{2}, \tag{12.18}$$

with the additional class of index 0: $\gamma \in [0,1) = [0,T_0)$. The set of quadruples $(\varkappa,\lambda,\mu,\nu)$ satisfying (12.18) with $\nu \geqslant 1$, $\mu \geqslant 0$, $\lambda \geqslant 1$ and $\varkappa \geqslant 0$ will be denoted by $\mathcal{B}$.

In this case we have clearly by (12.16), similarly to Davenport [2, §28],

$$\sum_{\substack{q\leqslant Q^* \\ (q,M)=1}} E^*(X,qM) \ll \frac{X}{M}\mathcal{L}^6 \max_{(\varkappa,\lambda,\mu,\nu)\in\mathcal{B}} \frac{N^*(1-(\varkappa+1)c_0/\mathcal{L}, M_\lambda Q_\nu, T_\mu)}{Q_\nu T_\mu} X^{-c_0\varkappa/\mathcal{L}}, \tag{12.19}$$

where

$$N^*(\sigma,Q,T) = \sum_{Q/2<q\leqslant Q} \sum_{\substack{\chi(q) \\ \chi \text{ primitive}}} \sum_{\substack{\varrho=\varrho_\chi \\ \beta\geqslant\sigma \\ |\gamma|\leqslant T}} 1. \tag{12.20}$$

We will see that, in order to prove our theorem, it will be enough to prove, for any quadruple (δ, M', Q, T) with the property (cf. (12.9) and (12.14)–(12.15)) that

$$\begin{aligned} \frac{c_0}{\log QM'T} \leqslant \delta \leqslant \frac{1}{2}, \quad & 2\leqslant M' \leqslant M, \quad 2 \leqslant Q \leqslant Q^*, \quad 1 \leqslant T \leqslant T^*, \quad \text{or} \\ \frac{c_0}{\log Y} \leqslant \delta \leqslant \frac{1}{2}, \quad & 2\leqslant M' \leqslant M, \quad Q \leqslant \sqrt{Y}, \quad T = T_0 = 1, \\ 0 \leqslant \delta \leqslant \frac{1}{2}, \quad & 2\leqslant M' \leqslant M, \quad Q > \sqrt{Y}, \quad T = T_0 = 1, \end{aligned} \tag{12.21}$$

the crucial inequality

$$N^*(1-\delta, M'Q, T) \ll \mathcal{L}^9 QTX^\delta e^{-c\log X/\log Y} \tag{12.22}$$

with some positive absolute constant c. The first line in (12.21) is meant to cover all non-real zeros, the second and third lines are meant to cover the real zeros.

We will use Theorem 12.2 of Montgomery [14]:

$$N^*(1-\delta, Q, T) \ll (Q^2T)^{3\delta/(1+\delta)}(\log QT)^9. \tag{12.23}$$

(We do not need, for the range $\delta \leqslant \frac{1}{5}$, the stronger inequality of Theorem 12.2 of [14] with the exponent $3\delta/(1+\delta)$ replaced by the smaller $2\delta/(1-\delta)$.) Since $3\delta/(1+\delta) \leqslant 1$, (12.22) will follow if we can show that

$$M^{6\delta}(Q)^{6\delta/(1+\delta)-1} \ll X^\delta e^{-c_2\sqrt{\log X}} \quad \text{with } c_2 = \tfrac{1}{6}c^*. \tag{12.24}$$

As in the range $0 \leqslant \delta \leqslant \frac{1}{2}$ we have $6\delta/(1+\delta)-1 \leqslant 2\delta$, this is true by the definition

$$Q^* = X^{1/2}M^{-3}e^{-c^*\sqrt{\log X}},$$

if $\delta \geqslant \frac{1}{12}$.

In case $\delta \leqslant \frac{1}{12}$ we have, by (12.23),

$$N^*(1-\delta, M'Q, T) \ll (QT)^{1/2}M^{6\delta}. \tag{12.25}$$

If we have here $QT \geqslant e^{\sqrt{\log X}}$, then (12.25) directly implies (12.22), since

$$\frac{N^*(1-\delta, M'Q, T)}{QT} \ll (QT)^{-1/2}M^{6\delta} \ll X^\delta e^{-\sqrt{\log X}/2}. \tag{12.26}$$

If $\delta \leqslant \frac{1}{12}$ and $QT \leqslant e^{\sqrt{\log X}} \leqslant \sqrt{Y}$, then we may assume $\delta \geqslant c_0/\log MQT$ by (12.9) or $\delta \geqslant 1/3\log Y > c_0/\log Y$ by (12.14)–(12.15), since the modulus of the corresponding primitive character is $q\widetilde{M} \leqslant 4QM \leqslant Y$. Hence,

$$\left(\frac{M^6}{X}\right)^\delta \leqslant X^{-\delta/4} \leqslant \exp\left(-\frac{c_0}{8}\min\left\{\frac{\log X}{\log QT}, \frac{\log X}{\log M}, \frac{\log X}{\log\sqrt{Y}}\right\}\right) = e^{-c_0\log X/8\log\sqrt{Y}}. \quad \square$$

Remark. The condition for Q^* could be weakened to $Q^* < X^{1/2}M^{-1}e^{-c^*\sqrt{\log X}}$, but this has no significance in our application.

13. Proof of Theorem 5

The method of proof of Theorem 5 is quite similar to that of Theorem 4. The basic difference is that instead of the trivial problem of the distribution of integers in arithmetic progressions we have to use properties of the distribution of primes in arithmetic progressions. Since we have to consider the (weighted) sum of the error terms in the formula for the number of primes in arithmetic progressions, the Bombieri–Vinogradov theorem can help us. However, due to the relatively weak estimate of the original Bombieri–Vinogradov theorem, it does not lead to better results than

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

That is partly why we need to use Theorem 6 instead. Our situation is even more complicated here, since we need the moduli of the progressions to be multiples of a number V. Fortunately our present Theorem 6 solves this problem in a completely satisfactory way, even without loss if $P=M\leqslant e^{(1+o(1))\sqrt{\log N}}$ which is now the case by $V=\sqrt{\log N}$.

We will suppose that $N=\frac{1}{3}X_n$, n is sufficiently large, $M=P$ and $Y(X)=e^{3\sqrt{\log X}}$ in Theorem 6. (If we use Heath-Brown's theorem [10] we may assume hypothesis $S(Y)$ for any N, and then N can be an arbitrary, sufficiently large integer.)

In the course of the proof we will follow closely the analogous proofs of Propositions 4 and 5 in §§7–9 of [7], so we will sometimes omit details. Let

$$\mathbf{\Theta}(x;q,a) := \sum_{\substack{p\leqslant x\\ p\equiv a \,(\mathrm{mod}\, q)}} \log p = [(a,q)=1]\frac{x}{\phi(q)} + E(x;q,a), \tag{13.1}$$

where $[S]$ is 1 if the statement S is true and 0 if S is false. We have for a regular residue class $\tilde{a}$ with respect to $\mathcal{H}$ and P,

$$\begin{aligned}\widetilde{S}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,P,\tilde{a},h_0) &:= \sum_{n=N+1}^{2N} \Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\theta(n+h_0)\\ &= \frac{1}{(K+\ell_1)!(K+\ell_2)!}\sum_{d,e\leqslant R}\mu(d)\mu(e)\\ &\quad\times\left(\log\frac{R}{d}\right)^{K+\ell_1}\left(\log\frac{R}{e}\right)^{K+\ell_2}\sum_{\substack{1\leqslant n\leqslant N\\ n\equiv\tilde{a}\,(\mathrm{mod}\,P)\\ d\mid P_{\mathcal{H}_1}(n)\\ e\mid P_{\mathcal{H}_2}(n)}}\theta(n+h_0).\end{aligned} \tag{13.2}$$

For the inner sum, we let $d=a_1a_{12}$ and $e=a_2a_{12}$, where $(d,e)=a_{12}$, and thus a_1, a_2 and a_{12} are pairwise relatively prime. In the following, we may suppose that

$$(d,P)=(e,P)=1,$$

otherwise the last sum would be zero, since by the regularity of $\tilde{a}$ we have

$$(P_{\mathcal{H}_1}(n), P) = (P_{\mathcal{H}_2}(n), P) = 1.$$

The n for which $d|P_{\mathcal{H}_1}(n)$ and $e|P_{\mathcal{H}_2}(n)$ cover certain residue classes modulo $[d,e]$. If $n \equiv b' \pmod{a_1a_2a_{12}}$ is such a residue class, then letting $m = n + h_0 \equiv b' + h_0 \pmod{a_1a_2a_{12}}$, $b \equiv b' \pmod{a_1a_2a_{12}}$ and $b \equiv \tilde{a} \pmod{P}$, we see that this residue class contributes to the inner sum

$$\begin{aligned}
&\sum_{\substack{N+1+h_0 \leqslant m \leqslant 2N+h_0 \\ m \equiv b+h_0 \,(\mathrm{mod}\, a_1a_2a_{12}P)}} \theta(m) \\
&\quad = \theta(2N+h_0; a_1a_2a_{12}P, b+h_0) - \theta(N+h_0; a_1a_2a_{12}P, b+h_0) \\
&\quad = [(b+h_0, a_1a_2a_{12}P) = 1] \frac{N}{\phi(a_1a_2a_{12}P)} + O(E^*(3N, a_1a_2a_{12}P)).
\end{aligned} \tag{13.3}$$

We need to determine the number of these residue classes where $(b+h_0, a_1a_2a_{12}P) = 1$ so that the main term is non-zero. The condition $(\tilde{a}+h_0, P) = (b+h_0, P) = 1$ is equivalent to $\tilde{a}$ being regular with respect to $\mathcal{H}^0$, since $\tilde{a}$ is regular with respect to $\mathcal{H}$. Thus we will assume from now on that $\tilde{a}$ is regular with respect to $\mathcal{H}^0$. 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 $\nu_p^*(\mathcal{H}_1)$ for this number in either case, where we define for a set $\mathcal{G}$,

$$\nu_p^*(\mathcal{G}) = \nu_p(\mathcal{G}^0) - 1, \tag{13.4}$$

with

$$\mathcal{G}^0 = \mathcal{G} \cup \{h_0\}. \tag{13.5}$$

We extend this definition to $\nu_d^*(\mathcal{H}_1)$ for squarefree numbers d, by multiplicativity. (The function ν_d^* is familiar in sieve theory, see [8].) The same applies for $\nu_d^*(\mathcal{H}_2)$ and $\bar{\nu}_d^*(\mathcal{H}_1 \bar{\cap} \mathcal{H}_2)$, as in (6.2).

Since $E(n; q, a) \ll \log N$ if $(a, q) > 1$ and $q \leqslant N$, we conclude that

$$\begin{aligned}
\sum_{\substack{N+1 \leqslant n \leqslant 2N \\ n \equiv \tilde{a}\,(P) \\ d|P_{\mathcal{H}_1}(n) \\ e|P_{\mathcal{H}_2}(n)}} \theta(n+h_0) &= \nu_{a_1}^*(\mathcal{H}_1)\nu_{a_2}^*(\mathcal{H}_2)\bar{\nu}_{a_{12}}^*(\mathcal{H}_1 \bar{\cap} \mathcal{H}_2) \frac{N}{\phi(a_1a_2a_{12}P)} \\
&\quad + O(d_K(a_1a_2a_{12}) E^*(3N; a_1a_2a_{12}P)).
\end{aligned} \tag{13.6}$$

Let $\sum^{(P)}$ denote that the summation variables are relatively prime to P and to each other. Substituting this into (13.2), we conclude, as $\ell_j \leqslant K$, that

$$\begin{aligned}
&\tilde{\mathcal{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,P,\tilde{a},h_0) \\
&\quad = \frac{N}{\varphi(P)(K+\ell_1)!(K+\ell_2)} \sum_{\substack{a_1a_{12}\leqslant R\\ a_2a_{12}\leqslant R}}^{(P)} \frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2\nu_{a_1}^*(\mathcal{H}_1)\nu_{a_2}^*(\mathcal{H}_2)\bar{\nu}_{a_{12}}^*(\mathcal{H}_1\bar{\cap}\mathcal{H}_2)}{\phi(a_1a_2a_{12})} \\
&\qquad\qquad \times\left(\log\frac{R}{a_1a_{12}}\right)^{K+\ell_1}\left(\log\frac{R}{a_2a_{12}}\right)^{K+\ell_2} \\
&\qquad + O\Bigg((\log R)^{4K} \sum_{\substack{a_1a_{12}\leqslant R\\ a_2a_{12}\leqslant R}}^{(P)} d_K(a_1a_2a_{12})E^*(3N;a_1a_2a_{12}P)\Bigg) \\
&\quad = \frac{N}{\varphi(P)}\widetilde{\mathcal{T}}_R(\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0)+O((\log R)^{4K}\mathcal{E}_K(N)).
\end{aligned} \tag{13.7}$$

Since $R^2\leqslant Q^*$, we obtain from Theorem 6, by the trivial estimate

$$|E(X,Pq,a)|\leqslant \frac{2X\log X}{Pq} \quad \text{for } Pq\leqslant X,$$

Lemma 3 and by Hölder's inequality with parameters $\alpha=\nu+1$ and $\beta=(\nu+1)/\nu$, where $\nu\in\mathbb{Z}^+$, $c'\log(K+1)\leqslant\nu\leqslant c''\log(K+1)$, uniformly for $K\leqslant(\log N)/2C$ ($\sum^{\flat *}$ means summation over squarefree integers which are relatively prime to P),

$$\begin{aligned}
\mathcal{E}_K(N) &\leqslant \sum_{q\leqslant Q^*}^{\flat *} d_K(q)E^*(3N,Pq)\sum_{q=a_1a_2a_{12}} 1 \\
&\leqslant \sum_{q\leqslant Q^*}^{\flat *} d_K(q)d_3(q)E^*(3N,Pq) \\
&= \sum_{q\leqslant Q^*}^{\flat *} \frac{d_{3K}(q)}{q^{1/\beta}}q^{1/\beta}E^*(3N,Pq) \\
&\leqslant \left(\sum_{q\leqslant Q^*}^{\flat} \frac{d_{3K}(q)^\beta}{q}\right)^{1/\beta}\left(\sum_{q\leqslant Q^*}^{\flat *} q^{\alpha/\beta}E^*(3N,Pq)^\alpha\right)^{1/\alpha} \\
&\leqslant \left(1+\frac{\log N}{2}\right)^{CK}\left(\frac{6N\log 3N}{P}\right)^{\nu/(\nu+1)}\left(\sum_{q\leqslant Q^*}^{\flat *} E^*(3N,Pq)\right)^{1/(\nu+1)} \\
&\ll (\log N)^{CK+1}\frac{N}{P}e^{-c_2\sqrt{\log N}/(\nu+1)} \\
&\leqslant \frac{N}{P}e^{(CK+1)\log_2 N-c_2\sqrt{\log N}/(\nu+1)} \\
&\leqslant \frac{N}{P}e^{-c\sqrt{\log N}/\log(K+1)},
\end{aligned} \tag{13.8}$$

since, by (3.24), K satisfies the inequality

$$K \log_2 N < c\frac{\sqrt{\log N}}{\log K}. \tag{13.9}$$

From (13.9) we have, finally,

$$(\log R)^{4K} \mathcal{E}_K(N) \leqslant \frac{N}{P} e^{-c\sqrt{\log N}/\log(K+1)}. \tag{13.10}$$

So, our task is reduced to the evaluation of $\widetilde{T}_R$ which is very similar to $\mathcal{T}_R^*$ in (6.22). Due to the more general treatment of $\mathcal{T}_R^*$ in §5 than needed, the crucial part, the error analysis will remain the same. The difference will only be the fact that we have now $\varphi(a_1a_2a_{12})$ in the denominator in (13.7) in place of $a_1a_2a_{12}$. Therefore $\nu_j(p)/p^{1+s_j}$ has to be replaced by $\nu_j(p)/(p-1)p^{s_j}$ in the definition of $F(s_1, s_2)$ and $G(s_1, s_2)$ in (6.5) and (6.11) (where $s_j = s_1$, $s_j = s_2$ or $s_j = s_3 = s_1 + s_2$). However, factors of the type $1 - p^{-(1+s_j)}$ remain unchanged, since they arise from the zeta-factors.

Summarizing our results above, we have

$$\widetilde{S}_R(N; \mathcal{H}_1, \mathcal{H}_2, \ell_1, \ell_2, P, \tilde{a}, h_0) = \frac{N}{\varphi(P)} \mathcal{T}_R(\ell_1, \ell_2; \mathcal{H}_1, \mathcal{H}_2) + O\left(\frac{N}{P} e^{-c\sqrt{\log N}/\log_2 N}\right), \tag{13.11}$$

where

$$\mathcal{T}_R(\ell_1, \ell_2, \mathcal{H}_1, \mathcal{H}_2) := \frac{1}{(2\pi i)^2} \int_{(1)} \int_{(1)} F(s_1, s_2) \frac{R^{s_1}}{s_1^{K+\ell_1+1}} \frac{R^{s_2}}{s_2^{K+\ell_2+1}}\, ds_1\, ds_2, \tag{13.12}$$

$$F(s_1, s_2) := \prod_{p>V} \left(1 - \frac{\nu_1(p)}{(p-1)p^{s_1}} - \frac{\nu_2(p)}{(p-1)p^{s_2}} + \frac{\nu_3(p)}{(p-1)p^{s_3}}\right) \tag{13.13}$$

and for this paragraph we have, with notation (6.2),

$$\nu_j(p) = \nu_p^*(\mathcal{H}_j) = \nu_p(\mathcal{H}_j^0) - 1,\ \ j = 1, 2, \quad \text{and} \quad \nu_3(p) = \bar{\nu}_p((\mathcal{H}_1 \bar{\cap} \mathcal{H}_2)^0) - 1. \tag{13.14}$$

To factor out the dominant zeta-factors we now write, in place of (6.8),

$$F(s_1, s_2) = G_{\mathcal{H}_1, \mathcal{H}_2}(s_1, s_2) \frac{\zeta(1+s_1+s_2)^{|(\mathcal{H}_1 \cap \mathcal{H}_2)^0| - 1}}{\zeta(1+s_1)^{|\mathcal{H}_1^0|-1} \zeta(1+s_2)^{|\mathcal{H}_2^0|-1}} \tag{13.15}$$

and define accordingly G_j, $1 \leqslant j \leqslant 4$, as in (6.9)–(6.11), with

$$a = |\mathcal{H}_1^0| - 1, \quad b = |\mathcal{H}_2^0| - 1 \quad \text{and} \quad d = |(\mathcal{H}_1 \cap \mathcal{H}_2)^0| - 1, \tag{13.16}$$

and with $\nu_j(p)p^{-s_j}/(p-1)$ in place of $\nu_j(p)p^{-1-s_j}$.

Similarly to [7, §9], by symmetry we have to consider three cases:

Case 1. $h_0 \notin \mathcal{H}$, that is $a=K$, $b=K$ and $d=r$.

Case 2. $h_0 \in \mathcal{H}_1 \setminus \mathcal{H}_2$, that is $a=K-1$, $b=K$ and $d=r$.

Case 3. $h_0 \in \mathcal{H}_1 \cap \mathcal{H}_2$, that is $a=K-1$, $b=K-1$ and $d=r-1$.

(Cases 1 and 3 are basically the same.)

Since the results of the previous section are more general, they apply to the error analysis here and we only have to evaluate $G(0,0)$ in Cases 1–3. Similarly to [7, §9], we have, by (13.14),

$$\nu_1(p)+\nu_2(p)-\nu_3(p)=\nu_p(\mathcal{H}_1^0)+\nu_p(\mathcal{H}_2^0)-\bar{\nu}_p(\mathcal{H}_1^0 \bar{\cap} \mathcal{H}_2^0)-1=\nu_p(\mathcal{H}^0)-1, \tag{13.17}$$

$$a+b-d=|\mathcal{H}^0|-1. \tag{13.18}$$

Hence, from the analogies of (6.9)–(6.11), we have now

$$G_1(0,0)=\prod_{p\leqslant V}\left(1-\frac{1}{p}\right)^{-(|\mathcal{H}^0|-1)}=\left(\frac{P}{\varphi(P)}\right)^{|\mathcal{H}^0|-1}, \tag{13.19}$$

$$\begin{aligned} G_4(0,0) &= \prod_{p>V}\left(1-\frac{\nu_p(\mathcal{H}^0)-1}{p-1}\right)\left(\frac{p}{p-1}\right)^{|\mathcal{H}^0|-1} \\ &= \prod_{p>V}\left(\frac{p-\nu_p(\mathcal{H}^0)}{p}\right)\left(1-\frac{1}{p}\right)^{-|\mathcal{H}^0|} =: \overline{\mathfrak{S}}_V(\mathcal{H}^0). \end{aligned} \tag{13.20}$$

Taking into account the term $\varphi(P)$ in the denominator in (13.11), we obtain

$$\frac{G(0,0)}{\varphi(P)}=\frac{1}{P}\prod_{p\leqslant V}\left(1-\frac{1}{p}\right)^{-|\mathcal{H}^0|}\overline{\mathfrak{S}}_V(\mathcal{H}^0). \tag{13.21}$$

Further, by the comparison of (13.12), (13.15) and (6.22), we have

$$u=K+\ell_1-a=K+1-|\mathcal{H}_1^0|+\ell_1, \quad v=K+1-|\mathcal{H}_2^0|+\ell_2 \quad \text{and} \quad d=|\mathcal{H}_1^0 \cap \mathcal{H}_2^0|-1. \tag{13.22}$$

The evaluation (11.5) of the crucial integral I defined in (6.22) yields, using in our case (13.11)–(13.15), the relation

$$\begin{aligned} &\widetilde{S}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,P,\tilde{a},h_0) \\ &\quad = N\frac{G(0,0)}{\varphi(P)}\frac{1}{(d+v+u)!}\binom{v+u}{u}(\log R)^{d+v+u}\left(1+O\left(\frac{K\bar{d}^*\log_2 N}{\log R}\right)\right) \\ &\qquad +O\left(\frac{N}{\varphi(P)}e^{-c\sqrt{\log N}/\log_2 N}\right). \end{aligned} \tag{13.23}$$

Let us observe that on the right-hand side the residue class $\tilde a$ does not appear at all. Therefore we can add this together for all $|A(\mathcal{H}^0)|$ regular residue classes $\tilde a \,(\mathrm{mod}\, P)$ with respect to $\mathcal{H}^0$ and P, since the contribution of those with $(\tilde a+h_0, P)>1$ is zero, as mentioned after (13.3). Taking into account the trivial relations (3.15)–(3.16) for $\mathcal{H}^0$ in place of $\mathcal{H}$, we obtain from (13.21),

$$\begin{aligned}\sum_{\tilde a\in A(\mathcal{H}^0)} \frac{G(0,0)}{\varphi(P)} &= \frac{|A(\mathcal{H}^0)|}{P} \prod_{p\leqslant V}\left(1-\frac{1}{p}\right)^{-|\mathcal{H}^0|} \overline{\mathfrak{S}}_V(\mathcal{H}^0) \\ &= \prod_{p\leqslant V}\left(1-\frac{\nu_p(\mathcal{H}^0)}{p}\right)\left(1-\frac{1}{p}\right)^{-|\mathcal{H}^0|} \overline{\mathfrak{S}}_V(\mathcal{H}^0) = \mathfrak{S}(\mathcal{H}^0).\end{aligned} \tag{13.24}$$

Inserting this into (13.23), we obtain by (13.11),

$$\begin{aligned}&\widetilde{\widetilde{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,P,h_0) \\ &\quad := \sum_{\tilde a\in A(\mathcal{H})} \widetilde{S}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,P,\tilde a,h_0) \\ &\quad = N\frac{1}{(d+v+u)!}\binom{v+u}{u}(\log R)^{d+v+u}\mathfrak{S}(\mathcal{H}^0)\left(1+O\left(\frac{K\bar d^*\log_2 N}{\log R}\right)\right) \\ &\qquad + O(Ne^{-c\sqrt{\log N}/\log_2 N}).\end{aligned} \tag{13.25}$$

Now, from a brief examination of the values of the parameters a, b and d in Cases 1, 2 and 3 (after (13.16)) and (13.22), we see that

$$\begin{aligned}&\frac{1}{(d+v+u)!}\binom{v+u}{u}(\log R)^{d+v+u} \\ &\quad = C_R(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0)\frac{1}{(r+\ell_1+\ell_2)!}\binom{\ell_1+\ell_2}{\ell_1}(\log R)^{r+\ell_1+\ell_2}.\end{aligned} \tag{13.26}$$

The relations (13.25)–(13.26) prove Theorem 5.

14. The sum of the singular series $\mathfrak{S}(\mathcal{H})$

Let

$$B_{\mathcal{A}}(k) = B(k) = \sum_{\substack{\mathcal{H}\subset\mathcal{A}\\ |\mathcal{H}|=k}} \mathfrak{S}(\mathcal{H}), \tag{14.1}$$

where all sets $\mathcal{H}=\{h_1,h_2,\dots,h_k\}\subseteq\mathcal{A}\subseteq[1,N]$ are counted with multiplicity $k!$, according to all possible permutations of the h_j, and $|\mathcal{A}|=h$.

By Gallagher's theorem [5] we have, for fixed k and $\mathcal{A}=[1,h]$, as $h\to\infty$,

$$B_{\mathcal{A}}(k) = h^k(1+O_{k,\varepsilon}(h^{-1/2+\varepsilon})). \tag{14.2}$$

This is not uniform in k, but up to some level $k \leqslant f(h)$ one could still show that $B_{\mathcal{A}}(k) \sim h^k$. However, we will use here a completely different approach. We do not prove (14.2), just (see Lemma 16) the weaker relation that $B_{\mathcal{A}}(k)/h^k$ is, apart from a factor $1+o(1)$, non-decreasing as a function of k, at least as long as $k=o(h/\log h)$. This result is fortunately completely sufficient for our purposes.

Further, our method is much more general and works for any set $\mathcal{A}$ with $\mathcal{A} \subseteq [1,N]$, $|\mathcal{A}|=h$.

We remark that the asymptotic $B_{\mathcal{A}}(k) \sim h^k$ is probably not true if $\mathcal{A}$ is arbitrary, and even for $\mathcal{A}=[1,h]$ it might fail if k is as large as $h/(\log h)^C$.

Let h and N be sufficiently large,

$$\begin{gathered} k \leqslant \log N, \quad h^2 \leqslant z = \log^5 N, \\ Z = P(z) = \prod_{p \leqslant z} p, \quad Y = Y_z = \prod_{p \leqslant z} \left(1 - \frac{1}{p}\right)^{-1} \sim e^{\gamma} \log z, \end{gathered} \tag{14.3}$$

$$Q := Q_z := \{n : (n, P(z)) = 1\} \quad \text{and} \quad M := \sum_{\substack{1 \leqslant n \leqslant Z \\ n \in Q}} 1 = \frac{Z}{Y}. \tag{14.4}$$

Then we have, for a fixed set $\mathcal{H}$ consisting of k distinct elements $h_j \in [1,N]$, similarly to §6, the density of z-quasi prime tuples of pattern $\mathcal{H}$, using (6.6):

$$\begin{aligned} R(\mathcal{H}) &:= \frac{1}{Z} \sum_{\substack{j=1 \\ P_{\mathcal{H}}(j) \in Q}}^{Z} 1 \\ &= \prod_{p \leqslant z} \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right) \\ &= Y^{-k} \prod_{p \leqslant z} \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right) \left(1 - \frac{1}{p}\right)^{-k} \\ &= Y^{-k} \mathfrak{S}(\mathcal{H}) \exp\left(O\left(k \sum_{\substack{p > z \\ p \mid \Delta}} \frac{1}{p} + k^2 \sum_{\substack{p > z \\ p \nmid \Delta}} \frac{1}{p^2}\right)\right) \\ &= Y^{-k} \mathfrak{S}(\mathcal{H}) \exp\left(O\left(k \sum_{p \mid \Delta} \frac{\log p}{z \log z} + \frac{k^2}{z \log z}\right)\right) \\ &= Y^{-k} \mathfrak{S}(\mathcal{H}) \exp\left(O\left(\frac{k^3 \log N}{z \log z} + \frac{k^2}{z \log z}\right)\right) \\ &= Y^{-k} \mathfrak{S}(\mathcal{H}) \left(1 + O\left(\frac{1}{\log N}\right)\right), \end{aligned} \tag{14.5}$$

uniformly in k, h, z and N satisfying (14.3). Let further

$$S^*_{\mathcal{A}}(k) := \frac{1}{h^k} \sum_{\substack{\mathcal{H}\subset\mathcal{A}\\|\mathcal{H}|=k}} \mathfrak{S}(\mathcal{H}) = \frac{B_{\mathcal{A}}(k)}{h^k}. \tag{14.6}$$

LEMMA 16. *If* $k<\varepsilon h/\log_2 N$, *then*

$$S^*_{\mathcal{A}}(k+1) \geqslant S^*_{\mathcal{A}}(k)\left(1+O(\varepsilon)+O\left(\frac{1}{\log N}\right)\right). \tag{14.7}$$

Proof. For $j\in[1,Z]$ let

$$f_j = \sum_{\substack{j'\\ j+a_{j'}\in Q}} 1 \quad \text{and} \quad b_j = b_j(k) = f_j(f_j-1)\dots(f_j-k+1). \tag{14.8}$$

Then $b_j(k)$ is the number of all k-tuples of z-quasiprimes of the type $j+a_{j_\nu}$, $a_{j_\nu}\in\mathcal{A}$ ($\nu=1,\dots,k$, $1\leqslant j_\nu\leqslant h$, j_ν distinct), calculated with all $k!$ permutations, while f_j is the number of z-quasiprimes of the form $j+a_{j'}$. For every pair $j,j'\in[1,h]$, we obviously have

$$f_j \geqslant f_{j'} \iff b_j \geqslant b_{j'}, \tag{14.9}$$

and therefore

$$\frac{1}{Z}\sum_{j=1}^{Z} b_j f_j \geqslant \left(\frac{1}{Z}\sum_{j=1}^{Z} f_j\right)\left(\frac{1}{Z}\sum_{j=1}^{Z} b_j\right). \tag{14.10}$$

The above formula follows from

$$2\left(Z\sum_{j=1}^{Z} b_j f_j - \sum_{j=1}^{Z} f_j \sum_{j=1}^{Z} b_j\right) = \sum_{j=1}^{Z}\sum_{j'=1}^{Z} (f_j-f_{j'})(b_j-b_{j'}) \geqslant 0. \tag{14.11}$$

We further have $b_j(k+1)=b_j(k)(f_j-k)=b_jf_j-kb_j$, and by calculating in two different ways how many times all pairs $(j,\mathcal{H})$ ($|\mathcal{H}|=k$) satisfy the relation $P_{\mathcal{H}}(j)\in Q$, we obtain

$$\frac{1}{Z}\sum_{j=1}^{Z} b_j(k) = \frac{1}{Z}\sum_{j=1}^{Z}\sum_{\substack{|\mathcal{H}|=k\\ P_{\mathcal{H}}(j)\in Q}} 1 = \frac{1}{Z}\sum_{|\mathcal{H}|=k}\sum_{\substack{j=1\\ P_{\mathcal{H}}(j)\in Q}}^{Z} 1 = \sum_{|\mathcal{H}|=k} R(\mathcal{H}), \tag{14.12}$$

while

$$\frac{1}{Z}\sum_{j=1}^{Z} f_j = \frac{hM}{Z} = \frac{h}{Y}. \tag{14.13}$$

Thus (14.10) and (14.13) imply, by $b_j f_j = b_j(k+1) + kb_j$, that

$$\frac{1}{Z}\sum_{j=1}^{Z} b_j(k+1) + k\frac{1}{Z}\sum_{j=1}^{Z} b_j(k) \geqslant \frac{h}{Y}\frac{1}{Z}\sum_{j=1}^{Z} b_j(k). \tag{14.14}$$

Hence, using (14.12), we obtain

$$\sum_{|\mathcal{H}|=k+1} R(\mathcal{H}) \geqslant \left(\frac{h}{Y} - k\right) \sum_{|\mathcal{H}|=k} R(\mathcal{H}). \tag{14.15}$$

Multiplying by Y^{k+1} on both sides, we obtain, by (14.5),

$$\sum_{|\mathcal{H}|=k+1} \mathfrak{S}(\mathcal{H}) \geqslant h\left(1 + O\left(\frac{kY}{h}\right) + O\left(\frac{1}{\log N}\right)\right) \sum_{|\mathcal{H}|=k} \mathfrak{S}(\mathcal{H}). \tag{14.16}$$

Now, dividing by h^{k+1} on both sides, we obtain (14.7) by $Y \ll \log_2 N$. □

15. Proof of Theorem 2

Theorems 4 and 5 allow us to express the quantity $S'_R(N, K, \ell, P)$ in (3.20) in terms of

$$S^*_{\mathcal{A}}(k) = S^*(k) := \frac{B_{\mathcal{A}}(k)}{h^k} := \frac{1}{h^k} \sum_{\substack{\mathcal{H}\subset\mathcal{A} \\ |\mathcal{H}|=k}} \mathfrak{S}(\mathcal{H}), \tag{15.1}$$

where we consider two sets $\mathcal{H}$ and $\mathcal{H}'$ different if they contain the same elements in different permutations. The value of the parameter k will be between K and $2K+1$, since in the application the sum on the right-hand side of (15.1) will be applied for sets of the type $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$, $|\mathcal{H}_j| = K$, or to $\mathcal{H}^0$.

The derivation of the proof of Theorem 2 from our present Theorems 4 and 5 will be nearly the same as that of the main result (Theorem 3) of [7] from Propositions 1 and 2 in [7], which appears in §10 of [7], so we will be brief. Although the restrictions for K and h will be quite different here, nearly everything will be valid without any change in the present case. Our analysis refers now to the case $\nu = 1$ of §10 in [7].

Let us choose, somewhat differently from [7],

$$R = (3N)^{\Theta} = (3N)^{1/4-\xi}, \quad \xi = \frac{c}{\sqrt{\log N}}, \quad V = \sqrt{\log N}, \tag{15.2}$$

$$K = 16(\ell+1)^2 = \frac{16}{\varphi^2}, \quad \text{that is} \quad \ell + 1 = \frac{1}{\varphi} = \frac{\sqrt{K}}{4}, \tag{15.3}$$

$$x = \frac{K}{100} = \frac{\log R}{h}, \quad \text{that is} \quad h = \frac{100 \log R}{K}\left(\sim \frac{25 \log N}{K}\right), \tag{15.4}$$

$$r_0 = (1-2\varphi)K, \quad r_1 = (1-\varphi)K, \tag{15.5}$$

$$f(r) = \binom{K}{r}^2 \frac{x^r}{(r+1)\dots(r+2\ell)}, \quad \bar{r}^* = \max\left\{\sqrt{K}, K-r\right\} \quad \text{and} \quad t(r) = \frac{\bar{r}^*}{\varphi K}, \tag{15.6}$$

and suppose that our crucial parameter K satisfies

$$K \leqslant c_0 \frac{\sqrt{\log N}}{\log_2^2 N}, \tag{15.7}$$

with a sufficiently small (explicitly calculable) absolute constant c_0.

In the course of the proof of our present Theorem 2 (similar to [7, §10]), a very important role is played by the fact that, although the sums evaluated in Theorems 4 and 5 depend on the actual choice of $\mathcal{H}_1$ and $\mathcal{H}_2$, the asymptotic formulas for them depends just on the set $\mathcal{H}=\mathcal{H}_1\cup\mathcal{H}_2$ and on the size of $\mathcal{H}_1\cap\mathcal{H}_2$. On the other hand, the size of the error terms may depend on the actual choice of $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}$. This dependence is made explicit in our present refined version, at least in the sense that we show an asymptotic which is more precise if $r=|\mathcal{H}_1\cap\mathcal{H}_2|$ is near $K=|\mathcal{H}_j|$.

We have seen in [7] that taking any given set $\mathcal{H}$ of given size $k=2K-r\in[K,2K]$, we can write it in

$$(2K-r)!\binom{K}{r}^2 r! \tag{15.8}$$

ways as the union of two sets $\mathcal{H}_1$ and $\mathcal{H}_2$ of size K, $|\mathcal{H}_1\cap\mathcal{H}_2|=r$, if we consider the sets $\mathcal{H}_j$ and $\mathcal{H}_j'$ to be different when the permutation of the same elements is different (cf. [7, (10.4)]). We can now apply Theorems 4 and 5 in order to obtain, similarly to [7, §10],

$$S_R'(N,K,\ell,P) = \binom{2\ell}{\ell}(\log R)^{2\ell} P_{K,\ell}^*(x), \tag{15.9}$$

with

$$P_{K,\ell}^*(x) \geqslant \sum_{r=0}^{K} f(r) S^*(2K-r)\left(1+O(\eta_2)+x\left(\frac{4K(1-\varphi/2)}{r+2\ell+1} - \frac{1}{\Theta} + O(\eta_1)\right)\right), \tag{15.10}$$

where the error terms arising from Theorems 4, 5 and Lemma 16 are now

$$\eta_1 = \frac{K\bar{r}^* \log_2 N}{\log N} = \frac{4K^{3/2} t(r) \log_2 N}{\log N} \quad \text{and} \quad \eta_2 = \frac{1}{\log_2^3 N}, \tag{15.11}$$

and by our choice of Θ in (15.2) we have

$$\frac{1}{\Theta} = 4 + O(\eta_3) \quad \text{and} \quad \eta_3 = \frac{1}{\sqrt{\log N}}. \tag{15.12}$$

We will examine the quantity in the parenthesis after x in (15.10) which is clearly monotonic in r (apart from the error terms). By (15.3)–(15.6), for $r\leqslant r_1=K-\varphi K=K-4\sqrt{K}$ (that is, $t(r)\geqslant 1$), we have

$$r+2\ell+1<K-t(r)\varphi K+\tfrac{1}{2}\sqrt{K}=K\bigl(1-t(r)\varphi+\tfrac{1}{8}\varphi\bigr), \tag{15.13}$$

and therefore

$$\begin{aligned}\frac{4(1-\varphi/2)K}{r+2\ell+1}-\frac{1}{\Theta}+O(\eta_1)&>4\left(t(r)-\frac{5}{8}\right)\varphi+O(\eta_1+\eta_3)\\ &>\frac{16}{\sqrt{K}}\frac{3}{8}t(r)-\frac{Ct(r)K^{3/2}\log_2 N}{\log N}-\frac{C}{\sqrt{\log N}}>0,\end{aligned} \tag{15.14}$$

as $t(r)\geqslant 1$ and (15.7).

On the other hand, as in [7, (10.24)–(10.25)], the contribution of all terms $r>r_1$ to $P^*_{K,\ell}(x)$ is bounded by

$$e^{-\sqrt{K}}f(r_0)\max_{r>r_1}S^*(2K-r), \tag{15.15}$$

because $f(r)$ quickly decreases for $r>r_1$. (These are the terms where the quantity in the parenthesis after x in (15.10) may be negative.) We have

$$\frac{f(r_1)}{f(r_0)}=\prod_{r_0<r\leqslant r_1}\left(\frac{K-r+1}{r}\frac{\sqrt{K}}{10}\right)^2\leqslant\left(\frac{2\varphi K}{K-2\varphi K}\frac{\sqrt{K}}{10}\right)^{2\varphi K}\leqslant(0.81\dots)^{8\sqrt{K}}=e^{-1.6\sqrt{K}}. \tag{15.16}$$

However, all terms $r\leqslant r_1$ have a positive contribution and that of $r=r_0$ is at least

$$f(r_0)S^*(2K-r_0)\left(1+O\left(\frac{1}{\log_2^3 N}\right)\right). \tag{15.17}$$

Now the quasi-monotonic property, Lemma 16, implies that

$$\frac{S^*(2K-r_0)}{\max\limits_{r>r_1}S^*(2K-r)}>e^{-(K-r_0)C/\log_2^3 N}=e^{-8C\sqrt{K}/\log_2^3 N}. \tag{15.18}$$

Consequently, the positive term belonging to r_0 dominates all possibly negative terms belonging to $r>r_1$, and therefore we have

$$P^*_{K,\ell}(x)>0,\quad \text{that is}\quad S'_R(N,K,\ell,P)>0. \tag{15.19}$$

This, by (3.20), proves the existence of some $n\in[N+1,2N]$ with

$$\sum_{\substack{p=n+a_\nu\\ a_\nu\in\mathcal{A}}}\log p>\log 3N, \tag{15.20}$$

and thereby the existence of two primes $p',p''\in[N+1,3N]$ with

$$0\neq p''-p'\in\mathcal{A}-\mathcal{A}. \tag{15.21}$$

This proves Theorem 2 if we choose K maximal, satisfying the restriction (15.7).

References

[1] Bombieri, E. & Davenport, H., Small differences between prime numbers. *Proc. Roy. Soc. Ser. A*, 293 (1966), 1–18.
[2] Davenport, H., *Multiplicative Number Theory.* Graduate Texts in Mathematics, 74. Springer, New York, 2000.
[3] Elliott, P. D. T. A. & Halberstam, H., A conjecture in prime number theory, in *Symposia Mathematica*, Vol. IV (INDAM, Rome, 1968/69), pp. 59–72. Academic Press, London, 1970.
[4] Erdős, P., The difference of consecutive primes. *Duke Math. J.*, 6 (1940), 438–441.
[5] Gallagher, P. X., On the distribution of primes in short intervals. *Mathematika*, 23 (1976), 4–9.
[6] Goldfeld, D. M. & Schinzel, A., On Siegel's zero. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 2 (1975), 571–583.
[7] Goldston, D. A., Pintz, J. & Yildirim, C. Y., Primes in tuples I. *Ann. of Math.*, 170 (2009), 819–862.
[8] Halberstam, H. & Richert, H. E., *Sieve Methods.* London Math. Soc. Monogr. Ser., 4. Academic Press, London–New York, 1974.
[9] Hardy, G. H. & Littlewood, J. E., Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes. *Acta Math.*, 44 (1923), 1–70.
[10] Heath-Brown, D. R., Prime twins and Siegel zeros. *Proc. Lond. Math. Soc.*, 47 (1983), 193–224.
[11] — Almost-prime k-tuples. *Mathematika*, 44 (1997), 245–266.
[12] Ivić, A., *The Riemann zeta-function.* John Wiley, New York, 1985.
[13] Maier, H., Small differences between prime numbers. *Michigan Math. J.*, 35 (1988), 323–344.
[14] Montgomery, H. L., *Topics in Multiplicative Number Theory.* Lecture Notes in Mathematics, 227. Springer, Berlin–Heidelberg, 1971.
[15] Pintz, J., Elementary methods in the theory of L-functions. II. On the greatest real zero of a real L-function. *Acta Arith.*, 31 (1976), 273–289.
[16] — Elementary methods in the theory of L-functions. VIII. Real zeros of real L-functions. *Acta Arith.*, 33 (1977), 89–98.
[17] — Very large gaps between consecutive primes. *J. Number Theory*, 63 (1997), 286–301.
[18] — Approximations to the Goldbach and twin prime problem and gaps between consecutive primes, in *Probability and Number Theory* (Kanazawa, 2005), Adv. Stud. Pure Math., 49, pp. 323–365. Math. Soc. Japan, Tokyo, 2007.
[19] de Polignac, A., Six propositions arithmologiques déduites du crible d'Ératosthène. *Nouv. Ann. Math.*, 8 (1849), 423–429.
[20] Titchmarsh, E. C., *The Theory of the Riemann Zeta-Function.* Oxford University Press, New York, 1986.

Daniel A. Goldston
Department of Mathematics
San José State University
One Washington Square
San José, CA 95192
U.S.A.
goldston@math.sjsu.edu

János Pintz
Alfréd Rényi Institute of Mathematics
Hungarian Academy of Sciences
Reáltanoda u. 13–15
HU-1053 Budapest
Hungary
pintz@renyi.hu

Cem Yalçin Yildirim
Department of Mathematics
Boğaziçi University
Bebek
Istanbul 34342
Turkey
yalciny@boun.edu.tr

Received November 15, 2007
Received in revised form October 21, 2008