Annals of Mathematics **179** (2014), 1121–1174
http://dx.doi.org/10.4007/annals.2014.179.3.7

Bounded gaps between primes

By Yitang Zhang

Abstract

It is proved that

$$\liminf_{n\to\infty}(p_{n+1}-p_n) < 7\times 10^7,$$

where p_n is the n-th prime.

Our method is a refinement of the recent work of Goldston, Pintz and Yıldırım on the small gaps between consecutive primes. A major ingredient of the proof is a stronger version of the Bombieri-Vinogradov theorem that is applicable when the moduli are free from large prime divisors only, but it is adequate for our purpose.

Contents

1. **Introduction**

Let p_n denote the n-th prime. It is conjectured that

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

While a proof of this conjecture seems to be out of reach by present methods, recently Goldston, Pintz and Yildirim [7] have made significant progress toward the weaker conjecture

$$\liminf_{n\to\infty}(p_{n+1}-p_n)<\infty. \tag{1.1}$$

In particular, they prove that if the primes have level of distribution $\vartheta = 1/2+\varpi$ for an (arbitrarily small) $\varpi>0$, then (1.1) will be valid (see [7, Th. 1]). Since the result $\vartheta=1/2$ is known (the Bombieri-Vinogradov theorem), the gap between their result and (1.1) would appear to be, as said in [7], within a hair's breadth. Until very recently, the best result on the small gaps between consecutive primes was due to Goldston, Pintz and Yildirim [8]. This result gives that

$$\liminf_{n\to\infty}\frac{p_{n+1}-p_n}{\sqrt{\log p_n}(\log\log p_n)^2}<\infty. \tag{1.2}$$

One may ask whether the methods in [7], combined with the ideas in Fouvry and Iwaniec [5] and in Bombieri, Friedlander and Iwaniec [1], [2], [3] which are employed to derive some stronger versions of the Bombieri- Vinogradov theorem, would be good enough for proving (1.1) (see [7, Question 1, p. 822]).

In this paper we give an affirmative answer to the above question. We adopt the following notation of [7]. Let

$$\mathcal{H}=\{h_1,\, h_2,\dots,h_{k_0}\} \tag{1.3}$$

be a set composed of distinct nonnegative integers. We say that $\mathcal{H}$ is admissible if $\nu_p(\mathcal{H})<p$ for every prime p, where $\nu_p(\mathcal{H})$ denotes the number of distinct residue classes modulo p occupied by the h_i.

THEOREM 1. *Suppose that $\mathcal{H}$ is admissible with $k_0\ge 3.5\times 10^6$. Then there are infinitely many positive integers n such that the k_0-tuple*

$$\{n+h_1,\, n+h_2,\dots,n+h_{k_0}\} \tag{1.4}$$

contains at least two primes. Consequently, we have

$$\liminf_{n\to\infty}(p_{n+1}-p_n)<7\times 10^7. \tag{1.5}$$

The bound (1.5) results from the fact that the set $\mathcal{H}$ is admissible if it is composed of k_0 distinct primes, each of which is greater than k_0, and the inequality

$$\pi(7\times10^7)-\pi(3.5\times10^6)>3.5\times10^6.$$

This result is, of course, not optimal. The condition $k_0 \geq 3.5 \times 10^6$ is also crude, and there are certain ways to relax it. To replace the right side of (1.5) by a value as small as possible is an open problem that will not be discussed in this paper.

2. **Notation and sketch of the proof**

Notation.

p: a prime number.
a, b, c, h, k, l, m: integers.
d, n, q, r: positive integers.
$\Lambda(q)$: the von Mangoldt function.
$\tau_j(q)$: the divisor function, $\tau_2(q) = \tau(q)$.
$\varphi(q)$: the Euler function.
$\mu(q)$: the Möbius function.
x: a large number.
$\mathcal{L} = \log x$.
y, z: real variables.
$e(y) = \exp\{2\pi i y\}$.
$e_q(y) = e(y/q)$.
$||y||$: the distance from y to the nearest integer.
$m \equiv a(q)$: means $m \equiv a(\operatorname{mod} q)$.
$\bar{c}/d$ means $a/d(\operatorname{mod} 1)$ where $ac \equiv 1(\operatorname{mod} d)$.
$q \sim Q$ means $Q \leq q < 2Q$.
ε: any sufficiently small, positive constant, not necessarily the same in each occurrence.
B: some positive constant, not necessarily the same in each occurrence.
A: any sufficiently large, positive constant, not necessarily the same in each occurrence.
$\eta = 1 + \mathcal{L}^{-2A}$.
$\varkappa_N$: the characteristic function of $[N, \eta N) \cap \mathbf{Z}$.
$\sum^*_{l(\bmod q)}$: a summation over reduced residue classes $l(\operatorname{mod} q)$.
$C_q(a)$: the Ramanujan sum $\sum^*_{l(\bmod q)} e_q(la)$.

We adopt the following conventions throughout our presentation. The set $\mathcal{H}$ given by (1.3) is assumed to be admissible and fixed. We write ν_p for $\nu_p(\mathcal{H})$; similar abbreviations will be used in the sequel. Every quantity depending on $\mathcal{H}$ alone is regarded as a constant. For example, the absolutely convergent product

$$\mathfrak{S} = \prod_p \left(1 - \frac{\nu_p}{p}\right)\left(1 - \frac{1}{p}\right)^{-k_0}$$

is a constant. A statement is valid for any sufficiently small ε and for any sufficiently large A whenever they are involved. The meanings of "sufficiently small" and "sufficiently large" may vary from one line to the next. Constants implied in O or $\ll$, unless specified, will depend on $\mathcal{H}$, ε and A at most.

We first recall the underlying idea in the proof of [7, Th. 1], which consists in evaluating and comparing the sums

$$S_1 = \sum_{n\sim x} \lambda(n)^2 \tag{2.1}$$

and

$$S_2 = \sum_{n\sim x} \bigg(\sum_{i=1}^{k_0} \theta(n+h_i) \bigg) \lambda(n)^2, \tag{2.2}$$

where $\lambda(n)$ is a real function depending on $\mathcal{H}$ and x, and

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

The key point is to prove, with an appropriate choice of λ, that

$$S_2 - (\log 3x) S_1 > 0. \tag{2.3}$$

This implies, for sufficiently large x, that there is a $n \sim x$ such that the tuple (1.4) contains at least two primes.

In [7] the function $\lambda(n)$ mainly takes the form

$$\lambda(n) = \frac{1}{(k_0+l_0)!} \sum_{\substack{d|P(n) \\ d\le D}} \mu(d) \bigg(\log \frac{D}{d} \bigg)^{k_0+l_0}, \qquad l_0 > 0, \tag{2.4}$$

where D is a power of x and

$$P(n) = \prod_{j=1}^{k_0} (n+h_j).$$

Let

$$\Delta(\gamma; d, c) = \sum_{\substack{n\sim x \\ n\equiv c(d)}} \gamma(n) - \frac{1}{\varphi(d)} \sum_{\substack{n\sim x \\ (n,d)=1}} \gamma(n) \quad \text{for} \quad (d,c)=1$$

and

$$\mathcal{C}_i(d) = \{c :\ 1 \le c \le d,\ (c,d)=1,\ P(c-h_i) \equiv 0 (\bmod d)\} \quad \text{for} \quad 1 \le i \le k_0.$$

The evaluations of S_1 and S_2 lead to a relation of the form

$$S_2 - (\log 3x) S_1 = (k_0 \mathcal{T}_2^* - \mathcal{L}\mathcal{T}_1^*) x + O(x\mathcal{L}^{k_0+2l_0}) + O(\mathcal{E})$$

for $D < x^{1/2-\varepsilon}$, where $\mathcal{T}_1^*$ and $\mathcal{T}_2^*$ are certain arithmetic sums (see Lemma 1 below), and

$$\mathcal{E} = \sum_{1\le i\le k_0} \sum_{d<D^2} |\mu(d)|\tau_3(d)\tau_{k_0-1}(d) \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta; d, c)|.$$

Let $\varpi > 0$ be a small constant. If

$$D = x^{1/4+\varpi} \tag{2.5}$$

and k_0 is sufficiently large in terms of ϖ, then, with an appropriate choice of l_0, one can prove that

$$k_0\mathcal{T}_2^* - \mathcal{L}\mathcal{T}_1^* \gg \mathcal{L}^{k_0+2l_0+1}. \tag{2.6}$$

In this situation the error $\mathcal{E}$ can be efficiently bounded if the primes have level of distribution $\vartheta > 1/2+2\varpi$, but one is unable to prove it by present methods. On the other hand, for $D = x^{1/4-\varepsilon}$, the Bombieri-Vinogradov theorem is good enough for bounding $\mathcal{E}$, but the relation (2.6) cannot be valid, even if a more general form of $\lambda(n)$ is considered (see Soundararajan [13]).

Our first observation is that, in the sums $\mathcal{T}_1^*$ and $\mathcal{T}_2^*$, the contributions from the terms with d having a large prime divisor are relatively small. Thus, if we impose the constraint $d|\mathcal{P}$ in (2.4), where $\mathcal{P}$ is the product of the primes less than a small power of x, the resulting main term is still $\gg \mathcal{L}^{k_0+2l_0+1}$ with D given by (2.5).

Our second observation, which is the most novel part of the proof, is that with D given by (2.5) and with the constraint $d|\mathcal{P}$ imposed in (2.4), the resulting error

$$\sum_{1\le i\le k_0} \sum_{\substack{d<D^2\\ d|\mathcal{P}}} \tau_3(d)\tau_{k_0-1}(d) \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta; d, c)| \tag{2.7}$$

can be efficiently bounded. This is originally due to the simple fact that if $d|\mathcal{P}$ and d is not too small, say $d > x^{1/2-\varepsilon}$, then d can be factored as

$$d = rq \tag{2.8}$$

with the range for r flexibly chosen (see Lemma 4 below). Thus, roughly speaking, the characteristic function of the set $\{d:\ x^{1/2-\varepsilon} < d < D^2,\ d|\mathcal{P}\}$ may be treated as a well-factorable function (see Iwaniec [11]). The factorization (2.8) is crucial for bounding the error terms.

It suffices to prove Theorem 1 with

$$k_0 = 3.5 \times 10^6,$$

which is henceforth assumed. Let D be as in (2.5) with

$$\varpi = \frac{1}{1168}.$$

Let $g(y)$ be given by

$$g(y) = \frac{1}{(k_0+l_0)!}\left(\log\frac{D}{y}\right)^{k_0+l_0} \quad \text{if} \quad y < D$$

and

$$g(y) = 0 \quad \text{if} \quad y \geq D,$$

where

$$l_0 = 180.$$

Write

$$D_1 = x^{\varpi}, \qquad \mathcal{P} = \prod_{p<D_1} p, \tag{2.9}$$

$$D_0 = \exp\{\mathcal{L}^{1/k_0}\}, \qquad \mathcal{P}_0 = \prod_{p\leq D_0} p. \tag{2.10}$$

In the case $d|\mathcal{P}$ and d is not too small, the factor q in (2.8) may be chosen such that $(q, \mathcal{P}_0) = 1$. This will considerably simplify the argument.

We choose

$$\lambda(n) = \sum_{d|(P(n),\mathcal{P})} \mu(d)g(d). \tag{2.11}$$

In the proof of Theorem 1, the main terms are not difficult to handle, since we deal with a fixed $\mathcal{H}$. This is quite different from [7] and [8], in which various sets $\mathcal{H}$ are involved in the argument to derive results like (1.2).

By Cauchy's inequality, the error (2.7) is efficiently bounded via the following

Theorem 2. *For $1 \leq i \leq k_0$ we have*

$$\sum_{\substack{d<D^2\\ d|\mathcal{P}}} \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta; d, c)| \ll x\mathcal{L}^{-A}. \tag{2.12}$$

The proof of Theorem 2 is described as follows. First, applying combinatorial arguments (see Lemma 6 below), we reduce the proof to estimating the sum of $|\Delta(\gamma; d, c)|$ with certain Dirichlet convolutions γ. There are three types of the convolutions involved in the argument. Write

$$x_1 = x^{3/8+8\varpi}, \qquad x_2 = x^{1/2-4\varpi}. \tag{2.13}$$

In the first two types the function γ is of the form $\gamma = \alpha * \beta$ such that the following hold:

(A_1) $\alpha = (\alpha(m))$ is supported on $[M, \eta^{j_1}M)$, $j_1 \leq 19$, $\alpha(m) \ll \tau_{j_1}(m)\mathcal{L}$.

(A_2) $\beta = (\beta(n))$ is supported on $[N, \eta^{j_2}N)$, $j_2 \leq 19$, $\beta(n) \ll \tau_{j_2}(n)\mathcal{L}$, $x_1 < N < 2x^{1/2}$.

For any q, r and a satisfying $(a,r)=1$, the following "Siegel-Walfisz" assumption is satisfied:

$$\sum_{\substack{n\equiv a(r)\\(n,q)=1}}\beta(n)-\frac{1}{\varphi(r)}\sum_{(n,qr)=1}\beta(n)\ll\tau_{20}(q)N\mathcal{L}^{-200A}.$$

(A$_3$) $j_1+j_2\le 20, \qquad [MN,\eta^{20}MN)\subset[x,2x)$.

We say that γ is of Type I if $x_1<N\le x_2$; we say that γ is of Type II if $x_2<N<2x^{1/2}$.

In the Type I and II estimates we combine the dispersion method in [5] and [1] with the factorization (2.8). (Here r is close to N in the logarithmic scale.) Due to the fact that the modulo d is at most slightly greater than $x^{1/2}$ in the logarithmic scale, after reducing the problem to estimating certain incomplete Kloosterman sums, we need only to save a small power of x from the trivial estimates; a variant of Weil's bound for Kloosterman sums (see Lemma 11) will fulfill it. Here the condition $N>x_1$, which may be slightly relaxed, is essential.

We say that γ is of Type III if it is of the form $\gamma=\alpha*\varkappa_{N_1}*\varkappa_{N_2}*\varkappa_{N_3}$ such that α satisfies (A$_1$) with $j_1\le 17$ and such that the following hold:

(A$_4$) $N_3\le N_2\le N_1$, $MN_1\le x_1$.

(A$_5$) $[MN_1N_2N_3,\eta^{20}MN_1N_2N_3)\subset[x,2x)$.

The Type III estimate essentially relies on the Birch-Bombieri result in the appendix to [6] (see Lemma 12), which is employed by Friedlander and Iwaniec [6] and by Heath-Brown [10] to study the distribution of $\tau_3(n)$ in arithmetic progressions. This result in turn relies on Deligne's proof of the Riemann Hypothesis for varieties over finite fields (the Weil Conjecture) [4]. We estimate each $\Delta(\gamma;d,c)$ directly. However, if one applies the method in [6] alone, efficient estimates will be valid only for $MN_1\ll x^{3/8-5\varpi/2-\varepsilon}$. Our argument is carried out by combining the method in [6] with the factorization (2.8) (here r is relatively small); the latter will allow us to save a factor $r^{1/2}$.

In our presentation, all the $\alpha(m)$ and $\beta(n)$ are real numbers.

3. **Lemmas**

In this section we introduce a number of prerequisite results, some of which are quoted from the literature directly. Results given here may not be in the strongest forms, but they are adequate for the proofs of Theorems 1 and 2.

LEMMA 1. *Let $\varrho_1(d)$ and $\varrho_2(d)$ be the multiplicative functions supported on square-free integers such that*

$$\varrho_1(p)=\nu_p, \qquad \varrho_2(p)=\nu_p-1.$$

Let

$$\mathcal{T}_1^* = \sum_{d_0} \sum_{d_1} \sum_{d_2} \frac{\mu(d_1 d_2)\varrho_1(d_0 d_1 d_2)}{d_0 d_1 d_2} g(d_0 d_1) g(d_0 d_2)$$

and

$$\mathcal{T}_2^* = \sum_{d_0} \sum_{d_1} \sum_{d_2} \frac{\mu(d_1 d_2)\varrho_2(d_0 d_1 d_2)}{\varphi(d_0 d_1 d_2)} g(d_0 d_1) g(d_0 d_2).$$

We have

$$\mathcal{T}_1^* = \frac{1}{(k_0 + 2l_0)!} \binom{2l_0}{l_0} \mathfrak{S} (\log D)^{k_0 + 2l_0} + o(\mathcal{L}^{k_0 + 2l_0}) \tag{3.1}$$

and

$$\mathcal{T}_2^* = \frac{1}{(k_0 + 2l_0 + 1)!} \binom{2l_0 + 2}{l_0 + 1} \mathfrak{S} (\log D)^{k_0 + 2l_0 + 1} + o(\mathcal{L}^{k_0 + 2l_0 + 1}). \tag{3.2}$$

Proof. The sum $\mathcal{T}_1^*$ is the same as the sum $\mathcal{T}_R(l_1, l_2; \mathcal{H}_1, \mathcal{H}_2)$ in [7, eq. (7.6)] with

$$\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H} \quad (k_1 = k_2 = k_0), \quad l_1 = l_2 = l_0, \quad R = D,$$

so (3.1) follows from [7, Lemma 3]; the sum $\mathcal{T}_2^*$ is the same as the sum $\tilde{\mathcal{T}}_R(l_1, l_2; \mathcal{H}_1, \mathcal{H}_2, h_0)$ in [7, eq. (9.12)] with

$$\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}, \quad l_1 = l_2 = l_0, \quad h_0 \in \mathcal{H}, \quad R = D,$$

so (3.2) also follows from [7, Lemma 3]. □

Remark. A generalization of this lemma can be found in [13].

LEMMA 2. *Let*

$$\mathcal{A}_1(d) = \sum_{(r,d)=1} \frac{\mu(r)\varrho_1(r)}{r} g(dr)$$

and

$$\mathcal{A}_2(d) = \sum_{(r,d)=1} \frac{\mu(r)\varrho_2(r)}{\varphi(r)} g(dr).$$

Suppose that $d < D$ and $|\mu(d)| = 1$. Then we have

$$\mathcal{A}_1(d) = \frac{\vartheta_1(d)}{l_0!} \mathfrak{S} \left(\log \frac{D}{d} \right)^{l_0} + O(\mathcal{L}^{l_0 - 1 + \varepsilon}) \tag{3.3}$$

and

$$\mathcal{A}_2(d) = \frac{\vartheta_2(d)}{(l_0 + 1)!} \mathfrak{S} \left(\log \frac{D}{d} \right)^{l_0 + 1} + O(\mathcal{L}^{l_0 + \varepsilon}), \tag{3.4}$$

where $\vartheta_1(d)$ and $\vartheta_2(d)$ are the multiplicative functions supported on square-free integers such that

$$\vartheta_1(p) = \left(1 - \frac{\nu_p}{p}\right)^{-1}, \qquad \vartheta_2(p) = \left(1 - \frac{\nu_p - 1}{p - 1}\right)^{-1}.$$

Proof. Recall that D_0 is given by (2.10). Since $\varrho_1(r) \le \tau_{k_0}(r)$, we have trivially

$$\mathcal{A}_1(d) \ll 1 + \Big(\log(D/d)\Big)^{2k_0+l_0},$$

so we may assume $D/d > \exp\{(\log D_0)^2\}$ without loss of generality. Write $s = \sigma + it$. For $\sigma > 0$ we have

$$\sum_{(r,d)=1} \frac{\mu(r)\varrho_1(r)}{r^{1+s}} = \vartheta_1(d,s)G_1(s)\zeta(1+s)^{-k_0},$$

where

$$\vartheta_1(d,s) = \prod_{p|d}\left(1 - \frac{\nu_p}{p^{1+s}}\right)^{-1}, \qquad G_1(s) = \prod_p \left(1 - \frac{\nu_p}{p^{1+s}}\right)\left(1 - \frac{1}{p^{1+s}}\right)^{-k_0}.$$

It follows that

$$\mathcal{A}_1(d) = \frac{1}{2\pi i}\int_{(1/\mathcal{L})} \frac{\vartheta_1(d,s)G_1(s)}{\zeta(1+s)^{k_0}}\frac{(D/d)^s\,ds}{s^{k_0+l_0+1}}.$$

Note that $G_1(s)$ is analytic and bounded for $\sigma \ge -1/3$. We split the line of integration into two parts according to $|t| \le D_0$ and $|t| > D_0$. By a well-known result on the zero-free region for $\zeta(s)$, we can move the line segment $\{\sigma = 1/\mathcal{L},\ |t| \le D_0\}$ to

$$\Big\{\sigma = -\kappa(\log D_0)^{-1},\ |t| \le D_0\Big\},$$

where $\kappa > 0$ is a certain constant, and apply some standard estimates to deduce that

$$\mathcal{A}_1(d) = \frac{1}{2\pi i}\int_{|s|=1/\mathcal{L}} \frac{\vartheta_1(d,s)G_1(s)}{\zeta(1+s)^{k_0}}\frac{(D/d)^s\,ds}{s^{k_0+l_0+1}} + O(\mathcal{L}^{-A}).$$

Note that $\vartheta_1(d,0) = \vartheta_1(d)$ and

$$\vartheta_1(d,s) - \vartheta_1(d) = \vartheta_1(d,s)\vartheta_1(d)\sum_{l|d}\frac{\mu(l)\varrho_1(l)}{l}(1 - l^{-s}).$$

If $|s| \le 1/\mathcal{L}$, then $\vartheta_1(d,s) \ll (\log\mathcal{L})^B$ so that, by trivial estimation,

$$\vartheta_1(d,s) - \vartheta_1(d) \ll \mathcal{L}^{\varepsilon-1}.$$

On the other hand, by Cauchy's integral formula, for $|s| \le 1/\mathcal{L}$ we have

$$G_1(s) - \mathfrak{S} \ll 1/\mathcal{L}.$$

It follows that

$$\frac{1}{2\pi i}\int_{|s|=1/\mathcal{L}} \frac{\vartheta_1(d,s)G_1(s)}{\zeta(1+s)^{k_0}}\frac{(D/d)^s\,ds}{s^{k_0+l_0+1}} - \frac{1}{2\pi i}\vartheta_1(d)\mathfrak{S}\int_{|s|=1/\mathcal{L}}\frac{(D/d)^s\,ds}{s^{l_0+1}} \ll \mathcal{L}^{l_0-1+\varepsilon}.$$

This leads to (3.3).

The proof of (3.4) is analogous. We have only to note that

$$\mathcal{A}_2(d) = \frac{1}{2\pi i}\int_{(1/\mathcal{L})} \frac{\vartheta_2(d,s)G_2(s)}{\zeta(1+s)^{k_0-1}}\frac{(D/d)^s\,ds}{s^{k_0+l_0+1}}$$

with

$$\vartheta_2(d,s) = \prod_{p|d}\left(1-\frac{\nu_p-1}{(p-1)p^s}\right)^{-1},\ \ G_2(s) = \prod_p\left(1-\frac{\nu_p-1}{(p-1)p^s}\right)\left(1-\frac{1}{p^{1+s}}\right)^{1-k_0},$$

and $G_2(0) = \mathfrak{S}$. □

Lemma 3. *We have*

$$\sum_{d<x^{1/4}} \frac{\varrho_1(d)\vartheta_1(d)}{d} = \frac{(1+4\varpi)^{-k_0}}{k_0!}\mathfrak{S}^{-1}(\log D)^{k_0} + O(\mathcal{L}^{k_0-1}) \tag{3.5}$$

and

$$\sum_{d<x^{1/4}} \frac{\varrho_2(d)\vartheta_2(d)}{\varphi(d)} = \frac{(1+4\varpi)^{1-k_0}}{(k_0-1)!}\mathfrak{S}^{-1}(\log D)^{k_0-1} + O(\mathcal{L}^{k_0-2}). \tag{3.6}$$

Proof. Noting that $\vartheta_1(p)/p = 1/(p-\nu_p)$, for $\sigma > 0$ we have

$$\sum_{d=1}^{\infty} \frac{\varrho_1(d)\vartheta_1(d)}{d^{1+s}} = B_1(s)\zeta(1+s)^{k_0},$$

where

$$B_1(s) = \prod_p\left(1+\frac{\nu_p}{(p-\nu_p)p^s}\right)\left(1-\frac{1}{p^{1+s}}\right)^{k_0}.$$

Hence, by Perron's formula,

$$\sum_{d<x^{1/4}} \frac{\varrho_1(d)\vartheta_1(d)}{d} = \frac{1}{2\pi i}\int_{1/\mathcal{L}-iD_0}^{1/\mathcal{L}+iD_0} \frac{B_1(s)\zeta(1+s)^{k_0}x^{s/4}}{s}\,ds + O(D_0^{-1}\mathcal{L}^B).$$

Note that $B_1(s)$ is analytic and bounded for $\sigma \ge -1/3$. Moving the path of integration to $[-1/3-iD_0, -1/3+iD_0]$, we see that the right side above is equal to

$$\frac{1}{2\pi i}\int_{|s|=1/\mathcal{L}} \frac{B_1(s)\zeta(1+s)^{k_0}x^{s/4}}{s}\,ds + O(D_0^{-1}\mathcal{L}^B).$$

Since, by Cauchy's integral formula, $B_1(s) - B_1(0) \ll 1/\mathcal{L}$ for $|s| = 1/\mathcal{L}$, and

$$B_1(0) = \prod_p\left(\frac{p}{p-\nu_p}\right)\left(1-\frac{1}{p}\right)^{k_0} = \mathfrak{S}^{-1},$$

it follows that

$$\sum_{d<x^{1/4}} \frac{\varrho_1(d)\vartheta_1(d)}{d} = \frac{1}{k_0!}\mathfrak{S}^{-1}\left(\frac{\mathcal{L}}{4}\right)^{k_0} + O(\mathcal{L}^{k_0-1}).$$

This leads to (3.5) since $\mathcal{L}/4 = (1+4\varpi)^{-1}\log D$ by (2.5).

The proof of (3.6) is analogous. We have only to note that, for $\sigma > 0$,

$$\sum_{d=1}^{\infty} \frac{\varrho_2(d)\vartheta_2(d)}{\varphi(d)p^s} = B_2(s)\zeta(1+s)^{k_0-1}$$

with

$$B_2(s) = \prod_p \left(1 + \frac{\nu_p - 1}{(p-\nu_p)p^s}\right)\left(1 - \frac{1}{p^{1+s}}\right)^{k_0-1},$$

and $B_2(0) = \mathfrak{S}^{-1}$. □

Recall that D_1 and $\mathcal{P}$ are given by (2.9) and $\mathcal{P}_0$ is given by (2.10).

Lemma 4. *Suppose that $d > D_1^2$, $d|\mathcal{P}$ and $(d, \mathcal{P}_0) < D_1$. For any R^* satisfying*

$$D_1^2 < R^* < d, \tag{3.7}$$

there is a factorization $d = rq$ such that $D_1^{-1}R^* < r < R^*$ and $(q, \mathcal{P}_0) = 1$.

Proof. Since d is square-free and $d/(d, \mathcal{P}_0) > D_1$, we may write $d/(d, \mathcal{P}_0)$ as

$$\frac{d}{(d, \mathcal{P}_0)} = \prod_{j=1}^{n} p_j \quad \text{with} \quad D_0 < p_1 < p_2 < \cdots < p_n < D_1, \quad n \geq 2.$$

By (3.7), there is a $n' < n$ such that

$$(d, \mathcal{P}_0) \prod_{j=1}^{n'} p_j < R^* \quad \text{and} \qquad (d, \mathcal{P}_0) \prod_{j=1}^{n'+1} p_j \geq R^*.$$

The assertion follows by choosing

$$r = (d, \mathcal{P}_0) \prod_{j=1}^{n'} p_j, \qquad q = \prod_{j=n'+1}^{n} p_j$$

and noting that $r \geq (1/p_{n'+1})R^*$. □

Lemma 5. *Suppose that $1 \leq i \leq k_0$ and $|\mu(qr)| = 1$. There is a bijection*

$$\mathcal{C}_i(qr) \to \mathcal{C}_i(r) \times \mathcal{C}_i(q), \qquad c \mapsto (a, b)$$

such that $c (\bmod qr)$ is a common solution to $c \equiv a (\bmod r)$ and $c \equiv b (\bmod q)$.

Proof. By the Chinese remainder theorem. □

The next lemma is a special case of the combinatorial identity due to Heath-Brown [9].

Lemma 6. *Suppose that* $x^{1/10} \le x^* < \eta x^{1/10}$. *For* $n < 2x$ *we have*

$$\Lambda(n) = \sum_{j=1}^{10} (-1)^{j-1} \binom{10}{j} \sum_{m_1,\dots,m_j \le x^*} \mu(m_1)\dots\mu(m_j) \sum_{n_1\dots n_j m_1\dots m_j = n} \log n_1.$$

The next lemma is a truncated Poisson formula.

Lemma 7. *Suppose that* $\eta \le \eta^* \le \eta^{19}$ *and* $x^{1/4} < M < x^{2/3}$. *Let* f *be a function of* $C^\infty(-\infty,\infty)$ *class such that* $0 \le f(y) \le 1$,

$$f(y) = 1 \quad \textit{if} \quad M \le y \le \eta^* M,$$

$$f(y) = 0 \quad \textit{if} \quad y \notin [(1 - M^{-\varepsilon})M,\, (1 + M^{-\varepsilon})\eta^* M],$$

and

$$f^{(j)}(y) \ll M^{-j(1-\varepsilon)}, \quad j \ge 1,$$

the implied constant depending on ε *and* j *at most. Then we have*

$$\sum_{m \equiv a(d)} f(m) = \frac{1}{d} \sum_{|h|<H} \hat{f}(h/d) e_d(-ah) + O(x^{-2})$$

for any $H \ge dM^{-1+2\varepsilon}$, *where* $\hat{f}$ *is the Fourier transform of* f; *i.e.*,

$$\hat{f}(z) = \int_{-\infty}^{\infty} f(y) e(yz)\, dy.$$

Lemma 8. *Suppose that* $1 \le N < N' < 2x$, $N' - N > x^\varepsilon d$ *and* $(c,d) = 1$. *Then for* $j, \nu \ge 1$ *we have*

$$\sum_{\substack{N \le n \le N' \\ n \equiv c(d)}} \tau_j(n)^\nu \ll \frac{N' - N}{\varphi(d)} \mathcal{L}^{j^\nu - 1},$$

the implied constant depending on ε, j *and* ν *at most.*

Proof. See [12, Th. 1]. □

The next lemma is (essentially) contained in the proof of [6, Th. 4].

Lemma 9. *Suppose that* $H, N \ge 2$, $d > H$ *and* $(c,d) = 1$. *Then we have*

$$\sum_{\substack{n \le N \\ (n,d)=1}} \min\left\{H,\, ||c\bar{n}/d||^{-1}\right\} \ll (dN)^\varepsilon (H + N). \tag{3.8}$$

Proof. We may assume $N \ge H$ without loss of generality. Write $\{y\} = y - [y]$, and assume $\xi \in [1/H,\, 1/2]$. Note that $\{c\bar{n}/d\} \le \xi$ if and only if $bn \equiv c(\operatorname{mod} d)$ for some $b \in (0, d\xi]$, and $1 - \xi \le \{c\bar{n}/d\}$ if and only if $bn \equiv$

$-c(\bmod d)$ for some $b \in (0, d\xi]$, Thus, the number of the n satisfying $n \le N$, $(n, d) = 1$ and $||c\bar{n}/d|| \le \xi$ is bounded by

$$\sum_{\substack{q \le dN\xi \\ q \equiv \pm c(d)}} \tau(q) \ll d^{\varepsilon} N^{1+\varepsilon} \xi.$$

Hence, for any interval I of the form

$$I = (0, 1/H], \quad I = [1 - 1/H, 1), \quad I = [\xi, \xi'] \quad \text{or} \quad I = [1 - \xi', 1 - \xi]$$

with $1/H \le \xi < \xi' \le 1/2$, $\xi' \le 2\xi$, the contribution from the terms on the left side of (3.8) with $\{c\bar{n}/d\} \in I$ is $\ll d^{\varepsilon} N^{1+\varepsilon}$. This completes the proof. □

Lemma 10. *Suppose that $\beta = (\beta(n))$ satisfies (A_2) and $R \le x^{-\varepsilon} N$. Then for any q we have*

$$\sum_{r \sim R} \varrho_2(r) \sideset{}{^*}\sum_{l (\bmod r)} \Bigg| \sum_{\substack{n \equiv l(r) \\ (n,q)=1}} \beta(n) - \frac{1}{\varphi(r)} \sum_{(n,qr)=1} \beta(n) \Bigg|^2 \ll \tau(q)^B N^2 \mathcal{L}^{-100A}.$$

Proof. Since the inner sum is $\ll \varphi(r)^{-1} N^2 \mathcal{L}^B$ by Lemma 8, the assertion follows by Cauchy's inequality and [1, Th. 0]. □

Lemma 11. *Suppose that $N \ge 1$, $d_1 d_2 > 10$ and $|\mu(d_1)| = |\mu(d_2)| = 1$. Then we have, for any c_1, c_2 and l,*

$$\sum_{\substack{n \le N \\ (n,d_1)=1 \\ (n+l,d_2)=1}} e\left(\frac{c_1 \bar{n}}{d_1} + \frac{c_2 \overline{(n+l)}}{d_2} \right) \ll (d_1 d_2)^{1/2+\varepsilon} + \frac{(c_1, d_1)(c_2, d_2)(d_1, d_2)^2 N}{d_1 d_2}. \tag{3.9}$$

Proof. Write $d_0 = (d_1, d_2)$, $t_1 = d_1/d_0$, $t_2 = d_2/d_0$ and $d = d_0 t_1 t_2$. Let

$$K(d_1, c_1; d_2, c_2; l, m) = \sum_{\substack{n \le d \\ (n,d_1)=1 \\ (n+l,d_2)=1}} e\left(\frac{c_1 \bar{n}}{d_1} + \frac{c_2 \overline{(n+l)}}{d_2} + \frac{mn}{d} \right).$$

We claim that

$$|K(d_1, c_1; d_2, c_2; l, m)| \le d_0 |S(m, b_1; t_1) S(m, b_2; t_2)| \tag{3.10}$$

for some b_1 and b_2 satisfying

$$(b_i, t_i) \le (c_i, d_i), \tag{3.11}$$

where $S(m, b; t)$ denotes the ordinary Kloosterman sum.

Note that d_0, t_1 and t_2 are pairwise coprime. Assume that

$$n \equiv t_1 t_2 n_0 + d_0 t_2 n_1 + d_0 t_1 n_2 \pmod d$$

and

$$l \equiv t_1 t_2 l_0 + d_0 t_1 l_2 \pmod{d_2}.$$

The conditions $(n, d_1) = 1$ and $(n + l, d_2) = 1$ are equivalent to

$$(n_0, d_0) = (n_1, t_1) = 1 \quad \text{and} \quad (n_0 + l_0, d_0) = (n_2 + l_2, t_2) = 1$$

respectively. Letting $a_i(\bmod d_0)$, $b_i(\bmod t_i)$, $i = 1, 2$ be given by

$$a_1 t_1^2 t_2 \equiv c_1 (\bmod d_0), \quad a_2 t_1 t_2^2 \equiv c_2 (\bmod d_0),$$
$$b_1 d_0^2 t_2 \equiv c_1 (\bmod t_1), \quad b_2 d_0^2 t_1 \equiv c_2 (\bmod t_2),$$

so that (3.11) holds, by the relation

$$\frac{1}{d_i} \equiv \frac{\bar{t}_i}{d_0} + \frac{\bar{d}_0}{t_i} \pmod 1$$

we have

$$\frac{c_1 \bar{n}}{d_1} + \frac{c_2 \overline{(n+l)}}{d_2} \equiv \frac{a_1 \bar{n}_0 + a_2 \overline{(n_0 + l_0)}}{d_0} + \frac{b_1 \bar{n}_1}{t_1} + \frac{b_2 \overline{(n_2 + l_2)}}{t_2} \pmod 1.$$

Hence,

$$\begin{aligned} &\frac{c_1 \bar{n}}{d_1} + \frac{c_2 \overline{(n+l)}}{d_2} + \frac{mn}{d} \\ &\quad \equiv \frac{a_1 \bar{n}_0 + a_2 \overline{(n_0 + l_0)} + mn_0}{d_0} \\ &\qquad + \frac{b_1 \bar{n}_1 + mn_1}{t_1} + \frac{b_2 \overline{(n_2 + l_2)} + m(n_2 + l_2)}{t_2} - \frac{ml_2}{t_2} (\bmod 1). \end{aligned}$$

From this we deduce, by the Chinese remainder theorem, that

$$\begin{aligned} &K(d_1, c_1; d_2, c_2; l, m) \\ &\quad = e_{t_2}(-ml_2) S(m, b_1; t_1) S(m, b_2; t_2) \sum_{\substack{n \le d_0 \\ (n, d_0) = 1 \\ (n + l_0, d_0) = 1}} e_{d_0}\Big(a_1 \bar{n} + a_2 \overline{(n + l_0)} + mn\Big), \end{aligned}$$

whence (3.10) follows.

By (3.10) with $m = 0$ and (3.11), for any $k > 0$ we have

$$\Bigg| \sum_{\substack{k \le n < k + d \\ (n, d_1 = 1 \\ (n + l, d_2) = 1}} e\bigg(\frac{c_1 \bar{n}}{d_1} + \frac{c_2 \overline{(n+l)}}{d_2}\bigg) \Bigg| \le (c_1, d_1)(c_2, d_2) d_0.$$

It now suffices to prove (3.9) on assuming $N \le d - 1$. By standard Fourier techniques, the left side of (3.9) may be rewritten as

$$\sum_{-\infty < m < \infty} u(m) K(d_1, c_1; d_2, c_2; l, m)$$

with

$$u(m) \ll \min \left\{ \frac{N}{d}, \frac{1}{|m|}, \frac{d}{m^2} \right\}. \tag{3.12}$$

By (3.10) and Weil's bound for Kloosterman sums, we find that the left side of (3.9) is

$$\ll d_0\biggl(|u(0)|(b_1,t_1)(b_2,t_2)+(t_1t_2)^{1/2+\varepsilon}\sum_{m\neq 0}|u(m)|(m,b_1,t_1)^{1/2}(m,b_2,t_2)^{1/2}\biggr).$$

This leads to (3.9) by (3.12) and (3.11). □

Remark. In the case $d_2=1$, (3.9) becomes

$$\sum_{\substack{n\le N\\(n,d_1)=1}} e_{d_1}(c_1\bar{n})\ll d_1^{1/2+\varepsilon}+\frac{(c_1,d_1)N}{d_1}. \tag{3.13}$$

This estimate is well known (see [2, Lemma 6], for example), and it will find application somewhere.

LEMMA 12. *Let*

$$T(k;m_1,m_2;q)=\sideset{}{'}\sum_{l(\bmod q)}\ \sideset{}{^*}\sum_{t_1(\bmod q)}\ \sideset{}{^*}\sum_{t_2(\bmod q)} e_q\bigl(\bar{l}t_1-\overline{(l+k)}t_2+m_1\bar{t}_1-m_2\bar{t}_2\bigr),$$

where $\sum'$ *is restriction to* $(l(l+k),q)=1$. *Suppose that* q *is square-free. Then we have*

$$T(k;m_1,m_2;q)\ll (k,q)^{1/2}q^{3/2+\varepsilon}.$$

Proof. By [6, eq. (1.26)], it suffices to show that

$$T(k;m_1,m_2;p)\ll (k,p)^{1/2}p^{3/2}.$$

In the case $k\not\equiv 0(\bmod p)$, this follows from the Birch-Bombieri result in the appendix to [6] (the proof is straightforward if $m_1m_2\equiv 0(\bmod p)$); in the case $k\equiv 0(\bmod p)$, this follows from Weil's bound for Kloosterman sums. □

4. Upper bound for S_1

Recall that S_1 is given by (2.1) and $\lambda(n)$ is given by (2.11). The aim of this section is to establish an upper bound for S_1 (see (4.20) below).

Changing the order of summation we obtain

$$S_1=\sum_{d_1|\mathcal{P}}\sum_{d_2|\mathcal{P}}\mu(d_1)g(d_1)\mu(d_2)g(d_2)\sum_{\substack{n\sim x\\P(n)\equiv 0([d_1,d_2])}}1.$$

By the Chinese remainder theorem, for any square-free d, there are exactly $\varrho_1(d)$ distinct residue classes $(\bmod d)$ such that $P(n)\equiv 0(\bmod d)$ if and only if n lies in one of these classes, so the innermost sum above is equal to

$$\frac{\varrho_1([d_1,d_2])}{[d_1,d_2]}x+O(\varrho_1([d_1,d_2])).$$

It follows that

$$S_1 = \mathcal{T}_1 x + O(D^{2+\varepsilon}), \tag{4.1}$$

where

$$\mathcal{T}_1 = \sum_{d_1|\mathcal{P}} \sum_{d_2|\mathcal{P}} \frac{\mu(d_1)g(d_1)\mu(d_2)g(d_2)}{[d_1, d_2]} \varrho_1([d_1, d_2]).$$

Note that $\varrho_1(d)$ is supported on square-free integers. Substituting $d_0 = (d_1, d_2)$ and rewriting d_1 and d_2 for d_1/d_0 and d_2/d_0 respectively, we deduce that

$$\mathcal{T}_1 = \sum_{d_0|\mathcal{P}} \sum_{d_1|\mathcal{P}} \sum_{d_2|\mathcal{P}} \frac{\mu(d_1 d_2)\varrho_1(d_0 d_1 d_2)}{d_0 d_1 d_2} g(d_0 d_1) g(d_0 d_2). \tag{4.2}$$

We need to estimate the difference $\mathcal{T}_1 - \mathcal{T}_1^*$. We have

$$\mathcal{T}_1^* = \Sigma_1 + \Sigma_{31},$$

where

$$\Sigma_1 = \sum_{d_0 \le x^{1/4}} \sum_{d_1} \sum_{d_2} \frac{\mu(d_1 d_2)\varrho_1(d_0 d_1 d_2)}{d_0 d_1 d_2} g(d_0 d_1) g(d_0 d_2),$$

$$\Sigma_{31} = \sum_{x^{1/4} < d_0 < D} \sum_{d_1} \sum_{d_2} \frac{\mu(d_1 d_2)\varrho_1(d_0 d_1 d_2)}{d_0 d_1 d_2} g(d_0 d_1) g(d_0 d_2).$$

In the case $d_0 > x^{1/4}$, $d_0 d_1 < D$, $d_0 d_2 < D$ and $|\mu(d_1 d_2)| = 1$, the conditions $d_i|\mathcal{P}$, $i = 1, 2$ are redundant. Hence,

$$\mathcal{T}_1 = \Sigma_2 + \Sigma_{32},$$

where

$$\Sigma_2 = \sum_{\substack{d_0 \le x^{1/4} \\ d_0|\mathcal{P}}} \sum_{d_1|\mathcal{P}} \sum_{d_2|\mathcal{P}} \frac{\mu(d_1 d_2)\varrho_1(d_0 d_1 d_2)}{d_0 d_1 d_2} g(d_0 d_1) g(d_0 d_2),$$

$$\Sigma_{32} = \sum_{\substack{x^{1/4} < d_0 < D \\ d_0|\mathcal{P}}} \sum_{d_1} \sum_{d_2} \frac{\mu(d_1 d_2)\varrho_1(d_0 d_1 d_2)}{d_0 d_1 d_2} g(d_0 d_1) g(d_0 d_2).$$

It follows that

$$|\mathcal{T}_1 - \mathcal{T}_1^*| \le |\Sigma_1| + |\Sigma_2| + |\Sigma_3|, \tag{4.3}$$

where

$$\Sigma_3 = \sum_{\substack{x^{1/4} < d_0 < D \\ d_0 \nmid \mathcal{P}}} \sum_{d_1} \sum_{d_2} \frac{\mu(d_1 d_2)\varrho_1(d_0 d_1 d_2)}{d_0 d_1 d_2} g(d_0 d_1) g(d_0 d_2).$$

First we estimate Σ_1. By Möbius inversion, the inner sum over d_1 and d_2 in Σ_1 is equal to

$$\frac{\varrho_1(d_0)}{d_0}\sum_{(d_1,d_0)=1}\sum_{(d_2,d_0)=1}\frac{\mu(d_1)\varrho_1(d_1)\mu(d_2)\varrho_1(d_2)}{d_1d_2}g(d_0d_1)g(d_0d_2)\Bigg(\sum_{q|(d_1,d_2)}\mu(q)\Bigg)$$
$$=\frac{\varrho_1(d_0)}{d_0}\sum_{(q,d_0)=1}\frac{\mu(q)\varrho_1(q)^2}{q^2}\mathcal{A}_1(d_0q)^2.$$

It follows that

$$\Sigma_1=\sum_{d_0\le x^{1/4}}\sum_{(q,d_0)=1}\frac{\varrho_1(d_0)\mu(q)\varrho_1(q)^2}{d_0q^2}\mathcal{A}_1(d_0q)^2. \tag{4.4}$$

The contribution from the terms with $q\ge D_0$ above is $\ll D_0^{-1}\mathcal{L}^B$. Thus, substituting $d_0q=d$, we deduce that

$$\Sigma_1=\sum_{d<x^{1/4}D_0}\frac{\varrho_1(d)\vartheta^*(d)}{d}\mathcal{A}_1(d)^2+O(D_0^{-1}\mathcal{L}^B), \tag{4.5}$$

where

$$\vartheta^*(d)=\sum_{\substack{d_0q=d\\ d_0<x^{1/4}\\ q<D_0}}\frac{\mu(q)\varrho_1(q)}{q}.$$

By the simple bounds

$$\mathcal{A}_1(d)\ll\mathcal{L}^{l_0}(\log\mathcal{L})^B, \tag{4.6}$$

which follows from (3.3),

$$\vartheta^*(d)\ll(\log\mathcal{L})^B$$

and

$$\sum_{x^{1/4}\le d<x^{1/4}D_0}\frac{\varrho_1(d)}{d}\ll\mathcal{L}^{k_0+1/k_0-1}, \tag{4.7}$$

the contribution from the terms on the right side of (4.5) with $x^{1/4}\le d<x^{1/4}D_0$ is $o(\mathcal{L}^{k_0+2l_0})$. On the other hand, assuming $|\mu(d)|=1$ and noting that

$$\sum_{q|d}\frac{\mu(q)\varrho_1(q)}{q}=\vartheta_1(d)^{-1}, \tag{4.8}$$

for $d<x^{1/4}$ we have

$$\vartheta^*(d)=\vartheta_1(d)^{-1}+O\Big(\tau_{k_0+1}(d)D_0^{-1}\Big),$$

so that, by (3.3),

$$\vartheta^*(d)\mathcal{A}_1(d)^2=\frac{1}{(l_0!)^2}\mathfrak{S}^2\vartheta_1(d)\bigg(\log\frac{D}{d}\bigg)^{2l_0}+O\Big(\tau_{k_0+1}(d)D_0^{-1}\mathcal{L}^B\Big)+O(\mathcal{L}^{2l_0-1+\varepsilon}).$$

Inserting this into (4.5) we obtain

$$\Sigma_1 = \frac{1}{(l_0!)^2}\mathfrak{S}^2 \sum_{d \le x^{1/4}} \frac{\varrho_1(d)\vartheta_1(d)}{d}\left(\log\frac{D}{d}\right)^{2l_0} + o(\mathcal{L}^{k_0+2l_0}). \tag{4.9}$$

Together with (3.5), this yields

$$|\Sigma_1| \le \frac{\delta_1}{k_0!(l_0!)^2}\mathfrak{S}(\log D)^{k_0+2l_0} + o(\mathcal{L}^{k_0+2l_0}), \tag{4.10}$$

where

$$\delta_1 = (1+4\varpi)^{-k_0}.$$

Next we estimate Σ_2. Similar to (4.4), we have

$$\Sigma_2 = \sum_{\substack{d_0 \le x^{1/4} \\ d_0|\mathcal{P}}} \sum_{\substack{(q,d_0)=1 \\ q|\mathcal{P}}} \frac{\varrho_1(d_0)\mu(q)\varrho_1(q)^2}{d_0 q^2}\mathcal{A}_1^*(d_0 q)^2,$$

where

$$\mathcal{A}_1^*(d) = \sum_{\substack{(r,d)=1 \\ r|\mathcal{P}}} \frac{\mu(r)\varrho_1(r)g(dr)}{r}.$$

In a way similar to the proof of (4.5), we deduce that

$$\Sigma_2 = \sum_{\substack{d < x^{1/4}D_0 \\ d|\mathcal{P}}} \frac{\varrho_1(d)\vartheta^*(d)}{d}\mathcal{A}_1^*(d)^2 + O(D_0^{-1}\mathcal{L}^B). \tag{4.11}$$

Assume $d|\mathcal{P}$. By Möbius inversion we have

$$\mathcal{A}_1^*(d) = \sum_{(r,d)=1} \frac{\mu(r)\varrho_1(r)g(dr)}{r} \sum_{q|(r,\mathcal{P}^*)} \mu(q) = \sum_{q|\mathcal{P}^*} \frac{\varrho_1(q)}{q}\mathcal{A}_1(dq),$$

where

$$\mathcal{P}^* = \prod_{D_1 \le p < D} p.$$

Noting that

$$\vartheta_1(q) = 1 + O(D_1^{-1}) \quad \text{if} \quad q|\mathcal{P}^* \quad \text{and} \quad q < D, \tag{4.12}$$

by (3.3) we deduce that

$$|\mathcal{A}_1^*(d)| \le \frac{1}{l_0!}\mathfrak{S}\vartheta_1(d)\left(\log\frac{D}{d}\right)^{l_0} \sum_{\substack{q|\mathcal{P}^* \\ q<D}} \frac{\varrho_1(q)}{q} + O(\mathcal{L}^{l_0-1+\varepsilon}). \tag{4.13}$$

If $q|\mathcal{P}^*$ and $q < D$, then q has at most 292 prime factors. In addition, by the prime number theorem we have

$$\sum_{D_1 \le p < D} \frac{1}{p} = \log 293 + O(\mathcal{L}^{-A}). \tag{4.14}$$

It follows that

$$\sum_{\substack{q|\mathcal{P}^*\\ q<D}} \frac{\varrho_1(q)}{q} \le 1 + \sum_{\nu=1}^{292} \frac{((\log 293)k_0)^\nu}{\nu!} + O(\mathcal{L}^{-A}) = \delta_2 + O(\mathcal{L}^{-A}), \qquad \text{say.}$$

Inserting this into (4.13) we obtain

$$|\mathcal{A}_1^*(d)| \le \frac{\delta_2}{l_0!}\mathfrak{S}\vartheta_1(d)\biggl(\log\frac{D}{d}\biggr)^{l_0} + O(\mathcal{L}^{l_0-1+\varepsilon}).$$

Combining this with (4.11), in a way similar to the proof of (4.9) we deduce that

$$|\Sigma_2| \le \frac{\delta_2^2}{(l_0!)^2}\mathfrak{S}^2 \sum_{d<x^{1/4}} \frac{\varrho_1(d)\vartheta_1(d)}{d}\biggl(\log\frac{D}{d}\biggr)^{2l_0} + o(\mathcal{L}^{k_0+2l_0}).$$

Together with (3.5), this yields

$$|\Sigma_2| \le \frac{\delta_1\delta_2^2}{k_0!(l_0!)^2}\mathfrak{S}(\log D)^{k_0+2l_0} + o(\mathcal{L}^{k_0+2l_0}). \tag{4.15}$$

We now turn to Σ_3. In a way similar to the proof of (4.5), we deduce that

$$\Sigma_3 = \sum_{x^{1/4}<d<D} \frac{\varrho_1(d)\tilde{\vartheta}(d)}{d}\mathcal{A}_1(d)^2, \tag{4.16}$$

where

$$\tilde{\vartheta}(d) = \sum_{\substack{d_0q=d\\ x^{1/4}<d_0\\ d_0\nmid\mathcal{P}}} \frac{\mu(q)\varrho_1(q)}{q}.$$

By (4.6) and (4.7), we find that the contribution from the terms with $x^{1/4} < d \le x^{1/4}D_0$ in (4.16) is $o(\mathcal{L}^{k_0+2l_0})$.

Now assume that $x^{1/4}D_0 < d < D$, $|\mu(d)| = 1$ and $d \nmid \mathcal{P}$. Noting that the conditions $d_0|d$ and $x^{1/4} < d_0$ together imply $d_0 \nmid \mathcal{P}$, by (4.8) we obtain

$$\tilde{\vartheta}(d) = \sum_{\substack{d_0q=d\\ x^{1/4}<d_0}} \frac{\mu(q)\varrho_1(q)}{q} = \vartheta_1(d)^{-1} + O\Bigl(\tau_{k_0+1}(d)D_0^{-1}\Bigr).$$

Together with (3.3), this yields

$$\tilde{\vartheta}(d)\mathcal{A}_1(d)^2 = \frac{1}{(l_0!)^2}\mathfrak{S}^2\vartheta_1(d)\biggl(\log\frac{D}{d}\biggr)^{2l_0} + O\Bigl(\tau_{k_0+1}(d)D_0^{-1}\mathcal{L}^B\Bigr) + O(\mathcal{L}^{2l_0-1+\varepsilon}).$$

Combining these results with (4.16) we obtain

$$\Sigma_3 = \frac{1}{(l_0!)^2}\mathfrak{S}^2 \sum_{\substack{x^{1/4}D_0<d<D\\ d\nmid\mathcal{P}}} \frac{\varrho_1(d)\vartheta_1(d)}{d}\biggl(\log\frac{D}{d}\biggr)^{2l_0} + o(\mathcal{L}^{k_0+2l_0}). \tag{4.17}$$

By (4.12), (4.14) and (3.5) we have

$$\begin{aligned}\sum_{\substack{x^{1/4}<d<D\\ d\nmid\mathcal{P}}}\frac{\varrho_1(d)\vartheta_1(d)}{d}&\le\sum_{d<D}\frac{\varrho_1(d)\vartheta_1(d)}{d}\sum_{p|(d,\mathcal{P}^*)}1\\&\le\sum_{D_1\le p<D}\frac{\varrho_1(p)\vartheta_1(p)}{p}\sum_{d<D/p}\frac{\varrho_1(d)\vartheta_1(d)}{d}\\&\le\frac{(\log 293)\delta_1}{(k_0-1)!}\mathfrak{S}^{-1}(\log D)^{k_0}+o(\mathcal{L}^{k_0}).\end{aligned}$$

Together with (4.17), this yields

$$|\Sigma_3|\le\frac{(\log 293)\delta_1}{(k_0-1)!(l_0!)^2}\mathfrak{S}(\log D)^{k_0+2l_0}+o(\mathcal{L}^{k_0+2l_0}). \tag{4.18}$$

Since

$$\frac{1}{k_0!(l_0!)^2}=\frac{1}{(k_0+2l_0)!}\binom{k_0+2l_0}{k_0}\binom{2l_0}{l_0},$$

it follows from (4.3), (4.10), (4.15) and (4.18) that

$$|\mathcal{T}_1-\mathcal{T}_1^*|\le\frac{\kappa_1}{(k_0+2l_0)!}\binom{2l_0}{l_0}\mathfrak{S}(\log D)^{k_0+2l_0}+o(\mathcal{L}^{k_0+2l_0}), \tag{4.19}$$

where

$$\kappa_1=\delta_1(1+\delta_2^2+(\log 293)k_0)\binom{k_0+2l_0}{k_0}.$$

Together with (3.1), this implies that

$$\mathcal{T}_1\le\frac{1+\kappa_1}{(k_0+2l_0)!}\binom{2l_0}{l_0}\mathfrak{S}(\log D)^{k_0+2l_0}+o(\mathcal{L}^{k_0+2l_0}).$$

Combining this with (4.1), we deduce that

$$S_1\le\frac{1+\kappa_1}{(k_0+2l_0)!}\binom{2l_0}{l_0}\mathfrak{S}x(\log D)^{k_0+2l_0}+o(x\mathcal{L}^{k_0+2l_0}). \tag{4.20}$$

We conclude this section by giving an upper bound for κ_1. By the inequality

$$n!>(2\pi n)^{1/2}n^ne^{-n}$$

and simple computation, we have

$$1+\delta_2^2+(\log 293)k_0<2\left(\frac{((\log 293)k_0)^{292}}{292!}\right)^2<\frac{1}{292\pi}(185100)^{584}$$

and

$$\binom{k_0+2l_0}{k_0}<\frac{2k_0^{2l_0}}{(2l_0)!}<\frac{1}{\sqrt{180\pi}}(26500)^{360}.$$

It follows that

$$\log \kappa_1 < -3500000 \log \frac{293}{292} + 584 \log(185100) + 360 \log(26500) < -1200.$$

This gives

$$\kappa_1 < \exp\{-1200\}. \tag{4.21}$$

5. Lower bound for S_2

Recall that S_2 is given by (2.2). The aim of this section is to establish a lower bound for S_2 on assuming Theorem 2 (see (5.6) below), which together with (4.20) leads to (2.3).

We have

$$S_2 = \sum_{1\le i\le k_0} \sum_{n\sim x} \theta(n)\lambda(n-h_i)^2 + O(x^\varepsilon). \tag{5.1}$$

Assume that $1 \le i \le k_0$. Changing the order of summation we obtain

$$\sum_{n\sim x} \theta(n)\lambda(n-h_i)^2 = \sum_{d_1|\mathcal{P}} \sum_{d_2|\mathcal{P}} \mu(d_1)g(d_1)\mu(d_2)g(d_2) \sum_{\substack{n\sim x \\ P(n-h_i)\equiv 0([d_1,d_2])}} \theta(n).$$

Now assume $|\mu(d)| = 1$. To handle the innermost sum we first note that the condition

$$P(n-h_i) \equiv 0 \pmod d \quad \text{and} \quad (n,d) = 1$$

is equivalent to $n \equiv c (\operatorname{mod} d)$ for some $c \in \mathcal{C}_i(d)$. Further, for any p, the quantity $|\mathcal{C}_i(p)|$ is equal to the number of distinct residue classes $(\operatorname{mod} p)$ occupied by the $h_i - h_j$ with $h_j \not\equiv h_i (\operatorname{mod} p)$, so $|\mathcal{C}_i(p)| = \nu_p - 1$. This implies $|\mathcal{C}_i(d)| = \varrho_2(d)$ by Lemma 5. Thus the innermost sum above is equal to

$$\sum_{c\in\mathcal{C}_i([d_1,d_2])} \sum_{\substack{n\sim x \\ n\equiv c([d_1,d_2])}} \theta(n) = \frac{\varrho_2([d_1,d_2])}{\varphi([d_1,d_2])} \sum_{n\sim x} \theta(n) + \sum_{c\in\mathcal{C}_i([d_1,d_2])} \Delta(\theta; [d_1,d_2], c).$$

Since the number of the pairs $\{d_1, d_2\}$ such that $[d_1, d_2] = d$ is equal to $\tau_3(d)$, it follows that

$$\sum_{n\sim x} \theta(n)\lambda(n-h_i)^2 = \mathcal{T}_2 \sum_{n\sim x} \theta(n) + O(\mathcal{E}_i), \tag{5.2}$$

where

$$\mathcal{T}_2 = \sum_{d_1|\mathcal{P}} \sum_{d_2|\mathcal{P}} \frac{\mu(d_1)g(d_1)\mu(d_2)g(d_2)}{\varphi([d_1,d_2])} \varrho_2([d_1,d_2])$$

which is independent of i, and

$$\mathcal{E}_i = \sum_{\substack{d<D^2 \\ d|\mathcal{P}}} \tau_3(d)\varrho_2(d) \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta; d, c)|.$$

By Cauchy's inequality and Theorem 2 we have

$$\mathcal{E}_i \ll x\mathcal{L}^{-A}. \tag{5.3}$$

It follows from (5.1)–(5.3) and the prime number theorem that

$$S_2 = k_0\mathcal{T}_2 x + O(x\mathcal{L}^{-A}). \tag{5.4}$$

Similar to (4.2), we may rewrite $\mathcal{T}_2$ as

$$\mathcal{T}_2 = \sum_{d_0|\mathcal{P}}\sum_{d_1|\mathcal{P}}\sum_{d_2|\mathcal{P}} \frac{\mu(d_1d_2)\varrho_2(d_0d_1d_2)}{\varphi(d_0d_1d_2)} g(d_0d_1)g(d_0d_2).$$

In a way much similar to the proof of (4.19), from the second assertions of Lemmas 2 and 3 we deduce that

$$|\mathcal{T}_2 - \mathcal{T}_2^*| < \frac{\kappa_2}{(k_0+2l_0+1)!}\binom{2l_0+2}{l_0+1}\mathfrak{S}(\log D)^{k_0+2l_0+1} + o(\mathcal{L}^{k_0+2l_0+1}), \tag{5.5}$$

where

$$\kappa_2 = \delta_1(1+4\varpi)(1+\delta_2^2+(\log 293)k_0)\binom{k_0+2l_0+1}{k_0-1}.$$

Together with (3.2), this implies that

$$\mathcal{T}_2 \geq \frac{1-\kappa_2}{(k_0+2l_0+1)!}\binom{2l_0+2}{l_0+1}\mathfrak{S}(\log D)^{k_0+2l_0+1} + o(\mathcal{L}^{k_0+2l_0+1}).$$

Combining this with (5.4), we deduce that

$$S_2 \geq \frac{k_0(1-\kappa_2)}{(k_0+2l_0+1)!}\binom{2l_0+2}{l_0+1}\mathfrak{S}x(\log D)^{k_0+2l_0+1} + o(x\mathcal{L}^{k_0+2l_0+1}). \tag{5.6}$$

We are now in a position to prove Theorem 1 on assuming Theorem 2. By (4.20), (5.6) and the relation

$$\mathcal{L} = \frac{4}{1+4\varpi}\log D,$$

we have

$$S_2 - (\log 3x)S_1 \geq \omega\mathfrak{S}x(\log D)^{k_0+2l_0+1} + o(x\mathcal{L}^{k_0+2l_0+1}), \tag{5.7}$$

where

$$\omega = \frac{k_0(1-\kappa_2)}{(k_0+2l_0+1)!}\binom{2l_0+2}{l_0+1} - \frac{4(1+\kappa_1)}{(1+4\varpi)(k_0+2l_0)!}\binom{2l_0}{l_0},$$

which may be rewritten as

$$\omega = \frac{1}{(k_0+2l_0)!}\binom{2l_0}{l_0}\left(\frac{2(2l_0+1)}{l_0+1}\frac{k_0(1-\kappa_2)}{k_0+2l_0+1} - \frac{4(1+\kappa_1)}{1+4\varpi}\right).$$

Note that

$$\frac{\kappa_2}{\kappa_1} = \frac{k_0(k_0+2l_0+1)(1+4\varpi)}{(2l_0+1)(2l_0+2)} < 10^8.$$

Thus, by (4.21), both of the constants κ_1 and κ_2 are extremely small. It follows by simple computation that

$$\omega > 0. \tag{5.8}$$

Finally, from (5.7) and (5.8) we deduce (2.3), whence Theorem 1 follows.

Remark. The bounds (4.19) and (5.5) are crude and there may be some ways to improve them considerably. It is even possible to evaluate $\mathcal{T}_1$ and $\mathcal{T}_2$ directly. Thus one might be able to show that (2.3) holds with a considerably smaller k_0.

6. **Combinatorial arguments**

The rest of this paper is devoted to proving Theorem 2. In this and the next six sections we assume that $1 \le i \le k_0$. Write

$$D_2 = x^{1/2-\varepsilon}.$$

On the left side of (2.12), the contribution from the terms with $d \le D_2$ is $\ll x\mathcal{L}^{-A}$ by the Bombieri-Vinogradov Theorem. Recalling that D_1 and $\mathcal{P}_0$ are given by (2.9) and (2.10) respectively, by trivial estimation, for $D_2 < d < D^2$ we may also impose the constraint $(d, \mathcal{P}_0) < D_1$ and replace $\theta(n)$ by $\Lambda(n)$. Thus Theorem 2 follows from the following:

$$\sum_{\substack{D_2<d<D^2 \\ d|\mathcal{P} \\ (d,\mathcal{P}_0)<D_1}} \sum_{c\in\mathcal{C}_i(d)} |\Delta(\Lambda; d, c)| \ll x\mathcal{L}^{-A}. \tag{6.1}$$

The aim of this section is to reduce the proof of (6.1) to showing that

$$\sum_{\substack{D_2<d<D^2 \\ d|\mathcal{P} \\ (d,\mathcal{P}_0)<D_1}} \sum_{c\in\mathcal{C}_i(d)} |\Delta(\gamma; d, c)| \ll x\mathcal{L}^{-41A} \tag{6.2}$$

for γ being of Type I, II or III.

Let L be given by $L(n) = \log n$. By Lemma 6, for $n \sim x$ we have $\Lambda(n) = \Lambda_1(n)$, where

$$\Lambda_1 = \sum_{j=1}^{10} (-1)^{j-1} \binom{10}{j} \sum_{M_j,\ldots,M_1,N_j,\ldots,N_1} (\mu\varkappa_{M_j})*\cdots*(\mu\varkappa_{M_1})*(\varkappa_{N_j})*\cdots*(L\varkappa_{N_1}).$$

Here $M_j, \ldots, M_1, N_j, \ldots, N_1 \ge 1$ run over the powers of η satisfying

$$M_t \le x^{1/10}, \tag{6.3}$$

$$[M_j \ldots M_1 N_j \ldots N_1, \eta^{20} M_j \ldots M_1 N_j \ldots N_1) \cap [x, 2x) \ne \phi. \tag{6.4}$$

Let Λ_2 have the same expression as Λ_1 but with the constraint (6.4) replaced by

$$[M_j \dots M_1 N_j \dots N_1, \eta^{20} M_j \dots M_1 N_j \dots N_1) \subset [x, 2x). \tag{6.5}$$

Since $\Lambda_1 - \Lambda_2$ is supported on $[\eta^{-20}x, \eta^{20}x] \cup [2\eta^{-20}x, 2\eta^{20}x]$ and $(\Lambda_1 - \Lambda_2)(n) \ll \tau_{20}(n)\mathcal{L}$, by Lemma 8 we have

$$\sum_{\substack{D_2 < d < D^2 \\ d|\mathcal{P} \\ (d,\mathcal{P}_0) < D_1}} \sum_{c \in \mathcal{C}_i(d)} |\Delta(\Lambda_1 - \Lambda_2; d, c)| \ll x\mathcal{L}^{-A}.$$

Further, let

$$\Lambda_3 = \sum_{j=1}^{10} (-1)^{j-1} \binom{10}{j} (\log N_1) \times \sum_{M_j, \dots, M_1, N_j, \dots, N_1} (\mu \varkappa_{M_j}) * \dots * (\mu \varkappa_{M_1}) * (\varkappa_{N_j}) * \dots * (\varkappa_{N_1}), \tag{6.6}$$

with $M_j, \dots, M_1, N_j, \dots, N_1$ satisfying (6.3) and (6.5). Since $(\Lambda_2 - \Lambda_3)(n) \ll \tau_{20}(n)\mathcal{L}^{-2A}$, by Lemma 8 we have

$$\sum_{\substack{D_2 < d < D^2 \\ d|\mathcal{P} \\ (d,\mathcal{P}_0) < D_1}} \sum_{c \in \mathcal{C}_i(d)} |\Delta(\Lambda_2 - \Lambda_3; d, c)| \ll x\mathcal{L}^{-A}.$$

Now assume that $1 \le j' \le j \le 10$. Let γ be of the form

$$\gamma = (\log N_{j'})(\mu \varkappa_{M_j}) * \dots * (\mu \varkappa_{M_1}) * (\varkappa_{N_j}) * \dots * (\varkappa_{N_1}),$$

with $M_j, \dots, M_1, N_j, \dots, N_1$ satisfying (6.3) and (6.5), and $N_j \le \dots \le N_1$. We claim that either the estimate

$$\Delta(\gamma; d, c) \ll \frac{x^{1-\varpi+\varepsilon}}{d} \tag{6.7}$$

trivially holds for $d < D^2$ and $(c, d) = 1$, or γ is of Type I, II or III.

Write $M_t = x^{\mu_t}$ and $N_t = x^{\nu_t}$. We have

$$0 \le \mu_t \le \frac{1}{10}, \quad 0 \le \nu_j \le \dots \le \nu_1, \quad 1 \le \mu_j + \dots + \mu_1 + \nu_j + \dots + \nu_1 < 1 + \frac{\log 2}{\mathcal{L}}.$$

In the case $3/8 + 8\varpi < \nu_1 \le 1/2$, γ is of Type I or II by choosing $\beta = \varkappa_{N_1}$; in the case $1/2 < \nu_1 \le 1/2 + 3\varpi$, γ is of Type II by choosing $\alpha = \varkappa_{N_1}$; in the case $1/2 + 3\varpi < \nu_1$, the estimate (6.7) trivially holds.

Since $\nu_1 \ge 2/5$ if $j = 1, 2$, it remains to deal with the case

$$j \ge 3, \qquad \nu_1 \le \frac{3}{8} + 8\varpi.$$

Write

$$\nu^* = \mu_j + \dots + \mu_1 + \nu_j + \dots + \nu_4.$$

(The partial sum $\nu_j + \cdots + \nu_4$ is void if $j = 3$.) In the case $\nu^* + \nu_1 \le 3/8 + 8\varpi$, γ is obviously of Type III. Further, if ν^* has a partial sum, say ν', satisfying

$$\frac{3}{8} + 8\varpi < \nu' + \nu_1 < \frac{5}{8} - 8\varpi,$$

then γ is of Type I or II. For example, if

$$\frac{3}{8} + 8\varpi < \mu_j + \cdots + \mu_1 + \nu_1 \le \frac{1}{2},$$

we choose $\beta = (\mu\varkappa_{M_j}) * \cdots * (\mu\varkappa_{M_1}) * (\varkappa_{N_1})$; if

$$\frac{1}{2} < \mu_j + \cdots + \mu_1 + \nu_1 < \frac{5}{8} - 8\varpi,$$

we choose $\alpha = (\mu\varkappa_{M_j}) * \cdots * (\mu\varkappa_{M_1}) * (\varkappa_{N_1})$.

It now suffices to assume that

$$\nu^* + \nu_1 \ge \frac{5}{8} - 8\varpi, \tag{6.8}$$

and every partial sum ν' of ν^* satisfies either

$$\nu' + \nu_1 \le \frac{3}{8} + 8\varpi \qquad \text{or} \qquad \nu' + \nu_1 \ge \frac{5}{8} - 8\varpi.$$

Let ν_1' be the smallest partial sum of ν^* such that $\nu_1' + \nu_1 \ge 5/8 - 8\varpi$ (the existence of ν_1' follows from (6.8), and there may be more than one choice of ν_1'), and let $\tilde{\nu}$ be a positive term in ν_1'. Since $\nu_1' - \tilde{\nu}$ is also a partial sum of ν^*, we must have

$$\nu_1' - \tilde{\nu} + \nu_1 \le 3/8 + 8\varpi,$$

so that

$$\tilde{\nu} \ge \frac{1}{4} - 16\varpi.$$

This implies that $\tilde{\nu}$ must be one of the ν_t, $t \ge 4$ (that arises only if $j \ge 4$). In particular, we have $\nu_4 \ge 1/4 - 16\varpi$. Now, the conditions

$$\frac{1}{4} - 16\varpi \le \nu_4 \le \nu_3 \le \nu_2 \le \nu_1, \qquad \nu_4 + \nu_3 + \nu_2 + \nu_1 < 1 + \frac{\log 2}{\mathcal{L}}$$

together imply that

$$\frac{1}{2} - 32\varpi \le \nu_3 + \nu_4 < \frac{1}{2} + \frac{\log 2}{2\mathcal{L}}.$$

It follows that γ is of Type I or II by choosing $\beta = \varkappa_{N_3} * \varkappa_{N_4}$.

It should be remarked, by the Siegel-Walfisz theorem, that for all the choices of β above, the Siegel-Walfisz assumption in (A_2) holds. Noting that the sum in (6.6) contains $O(\mathcal{L}^{40A})$ terms, by the above discussion we conclude that (6.2) implies (6.1).

7. The dispersion method

In this and the next three sections we treat the Type I and II estimates simultaneously via the methods in [5] and [1, §§3–7]. We henceforth assume that $\gamma = \alpha * \beta$ satisfies (A_1), (A_2) and (A_3). Recall that x_1 and x_2 are given by (2.13). We shall apply Lemma 4 with

$$R^* = x^{-\varepsilon} N \tag{7.1}$$

if γ is of Type I ($x_1 < N \le x_2$), and

$$R^* = x^{-3\varpi} N \tag{7.2}$$

if γ is of Type II ($x_2 < N < 2x^{1/2}$).

Note that $D_1^2 < R^* < D_2$. By Lemma 4 and Lemma 5, the proof of (6.2) is reduced to showing that

$$\sum_{R^*/D_1 < r < R^*} |\mu(r)| \sum_{a \in \mathcal{C}_i(r)} \sum_{\substack{D_2/r < q < D^2/r \\ q \mid \mathcal{P} \\ (q, r\mathcal{P}_0)=1}} \sum_{b \in \mathcal{C}_i(q)} \Big| \Delta(\gamma; r, a; q, b) \Big| \ll x\mathcal{L}^{-41A}$$

where, for $|\mu(qr)| = (a,r) = (b,q) = 1$,

$$\Delta(\gamma; r, a; q, b) = \sum_{\substack{n \equiv a(r) \\ n \equiv b(q)}} \gamma(n) - \frac{1}{\varphi(qr)} \sum_{(n,qr)=1} \gamma(n).$$

It therefore suffices to prove that

$$\mathcal{B}(\gamma; Q, R) := \sum_{r \sim R} |\mu(r)| \sum_{a \in \mathcal{C}_i(r)} \sum_{\substack{q \sim Q \\ q \mid \mathcal{P} \\ (q, r\mathcal{P}_0)=1}} \sum_{b \in \mathcal{C}_i(q)} \Big| \Delta(\gamma; r, a; q, b) \Big| \ll x\mathcal{L}^{-43A}, \tag{7.3}$$

subject to the conditions

$$x^{-\varpi} R^* < R < R^* \tag{7.4}$$

and

$$\frac{1}{2} x^{1/2-\varepsilon} < QR < x^{1/2+2\varpi}, \tag{7.5}$$

which are henceforth assumed.

For notational simplicity, in some expressions the subscript i will be omitted even though they depend on it. In what follows we assume that

$$r \sim R, \quad |\mu(r)| = 1, \quad a \in \mathcal{C}_i(r). \tag{7.6}$$

Let $c(r, a; q, b)$ be given by

$$c(r, a; q, b) = \operatorname{sgn} \Delta(\gamma; r, a; q, b)$$

if

$$q\sim Q,\quad q|\mathcal{P},\quad (q,r\mathcal{P}_0)=1,\quad b\in\mathcal{C}_i(q),$$

and

$$c(r,a;q,b)=0\qquad\text{otherwise.}$$

Changing the order of summation we obtain

$$\sum_{\substack{q\sim Q\\ q|\mathcal{P}\\ (q,r\mathcal{P}_0)=1}}\ \sum_{b\in\mathcal{C}_i(q)}\Big|\Delta(\gamma;r,a;q,b)\Big|=\sum_{(m,r)=1}\alpha(m)\mathcal{D}(r,a;m),$$

where

$$\mathcal{D}(r,a;m)=\sum_{(q,m)=1}\sum_{b}c(r,a;q,b)\Bigg(\sum_{\substack{mn\equiv a(r)\\ mn\equiv b(q)}}\beta(n)-\frac{1}{\varphi(qr)}\sum_{(n,qr)=1}\beta(n)\Bigg).$$

It follows by Cauchy's inequality that

$$\mathcal{B}(\gamma;Q,R)^2\ll MR\mathcal{L}^B\sum_{r\sim R}|\mu(r)|\sum_{a\in\mathcal{C}_i(r)}\ \sum_{(m,r)=1}f(m)\mathcal{D}(r,a;m)^2,\tag{7.7}$$

where $f(y)$ is as in Lemma 7 with $\eta^*=\eta^{19}$. We have

$$\sum_{(m,r)=1}f(m)\mathcal{D}(r,a;m)^2=\mathcal{S}_1(r,a)-2\mathcal{S}_2(r,a)+\mathcal{S}_3(r,a),\tag{7.8}$$

where $\mathcal{S}_j(r,a)$, $j=1,2,3$ are defined by

$$\mathcal{S}_1(r,a)=\sum_{(m,r)=1}f(m)\Bigg(\sum_{(q,m)=1}\sum_{b}c(r,a;q,b)\sum_{\substack{mn\equiv a(r)\\ mn\equiv b(q)}}\beta(n)\Bigg)^2,$$

$$\mathcal{S}_2(r,a)=\sum_{q_1}\sum_{b_1}\sum_{q_2}\sum_{b_2}\frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{\varphi(q_2r)}$$
$$\times\sum_{n_1}\sum_{(n_2,q_2r)=1}\beta(n_1)\beta(n_2)\sum_{\substack{mn_1\equiv a(r)\\ mn_1\equiv b_1(q_1)\\ (m,q_2)=1}}f(m),$$

$$\mathcal{S}_3(r,a)=\sum_{q_1}\sum_{b_1}\sum_{q_2}\sum_{b_2}\frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{\varphi(q_1r)\varphi(q_2r)}$$
$$\times\sum_{(n_1,q_1r)=1}\ \sum_{(n_2,q_2r)=1}\beta(n_1)\beta(n_2)\sum_{(m,q_1q_2r)=1}f(m).$$

By (7.7) and (7.8), the proof of (7.3) is reduced to showing that

$$\sum_{r}\sum_{a}\Big(\mathcal{S}_1(r,a)-2\mathcal{S}_2(r,a)+\mathcal{S}_3(r,a)\Big)\ll xNR^{-1}\mathcal{L}^{-87A}\tag{7.9}$$

on assuming $A\geq B$. Here we have omitted the constraints given in (7.6) for notational simplicity, so they have to be remembered in the sequel.

8. **Evaluation of** $\mathcal{S}_3(r,a)$

In this section we evaluate $\mathcal{S}_3(r,a)$. We shall make frequent use of the trivial bound

$$\hat{f}(z) \ll M. \tag{8.1}$$

By Möbius inversion and Lemma 7, for $q_j \sim Q$, $j = 1, 2$ we have

$$\sum_{(m,q_1q_2r)=1} f(m) = \frac{\varphi(q_1q_2r)}{q_1q_2r}\hat{f}(0) + O(x^{\varepsilon}). \tag{8.2}$$

This yields

$$\begin{aligned}\mathcal{S}_3(r,a) = \hat{f}(0)\sum_{q_1}\sum_{b_1}\sum_{q_2}\sum_{b_2}&\frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{\varphi(q_1)\varphi(q_2)\varphi(r)}\frac{\varphi(q_1q_2)}{q_1q_2r}\\ \times \sum_{(n_1,q_1r)=1}\ \sum_{(n_2,q_2r)=1}&\beta(n_1)\beta(n_2) + O(x^{\varepsilon}N^2R^{-2}).\end{aligned}$$

In view of (2.10), if $(q_1q_2, \mathcal{P}_0) = 1$, then either $(q_1, q_2) = 1$ or $(q_1, q_2) > D_0$. Thus, on the right side above, the contribution from the terms with $(q_1, q_2) > 1$ is, by (8.1) and trivial estimation,

$$\ll xND_0^{-1}R^{-2}\mathcal{L}^B.$$

It follows that

$$\mathcal{S}_3(r,a) = \hat{f}(0)X(r,a) + O(xND_0^{-1}R^{-2}\mathcal{L}^B), \tag{8.3}$$

where

$$\begin{aligned}X(r,a) = \sum_{q_1}\sum_{b_1}&\sum_{(q_2,q_1)=1}\\ \times\sum_{b_2}&\frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{q_1q_2r\varphi(r)}\sum_{(n_1,q_1r)=1}\ \sum_{(n_2,q_2r)=1}\beta(n_1)\beta(n_2).\end{aligned}$$

9. **Evaluation of** $\mathcal{S}_2(r,a)$

The aim of this section is to show that

$$\mathcal{S}_2(r,a) = \hat{f}(0)X(r,a) + O(xND_0^{-1}R^{-2}\mathcal{L}^B). \tag{9.1}$$

Assume $c(r,a;q_1,b_1)c(r,a;q_2,b_2) \neq 0$. Let $\nu(\operatorname{mod} q_1r)$ be a common solution to

$$\nu \equiv a(\operatorname{mod} r), \qquad \nu \equiv b_1(\operatorname{mod} q_1).$$

Substituting $mn_1 = n$ and applying Lemma 8 we obtain

$$\sum_{n_1}\beta(n_1)\sum_{\substack{mn_1\equiv a(r)\\ mn_1\equiv b_1(q_1)\\ (m,q_2)=1}} f(m) \ll \sum_{\substack{n<2x\\ n\equiv\nu(q_1r)}}\tau_{20}(n) \ll \frac{x\mathcal{L}^B}{q_1r}.$$

It follows that the contribution from the terms with $(q_1, q_2) > 1$ in $\mathcal{S}_2(r, a)$ is

$$\ll xND_0^{-1}R^{-2}\mathcal{L}^B,$$

so that,

(9.2)

$$\mathcal{S}_2(r,a) = \sum_{q_1}\sum_{b_1}\sum_{(q_2,q_1)=1}\sum_{b_2} \frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{\varphi(q_2 r)} \times \sum_{n_1}\sum_{(n_2,q_2r)=1} \beta(n_1)\beta(n_2) \sum_{\substack{mn_1\equiv a(r)\\ mn_1\equiv b_1(q_1)\\ (m,q_2)=1}} f(m) + O(xND_0^{-1}R^{-2}\mathcal{L}^B).$$

Note that the innermost sum in (9.2) is void unless $(n_1, q_1 r) = 1$. For $|\mu(q_1q_2r)| = 1$ and $(q_2, \mathcal{P}_0) = 1$ we have

$$\frac{q_2}{\varphi(q_2)} = 1 + O(\tau(q_2)D_0^{-1})$$

and, by Lemma 8,

$$\sum_{(n_1,q_1r)=1} \beta(n_1) \sum_{\substack{mn_1\equiv a(r)\\ mn_1\equiv b_1(q_1)\\ (m,q_2)>1}} f(m) \ll \sum_{\substack{n<2x\\ n\equiv\nu(q_1r)\\ (n,q_2)>1}} \tau_{20}(n) \ll \frac{\tau_{20}(q_2)x\mathcal{L}^B}{q_1 r D_0}.$$

Thus the relation (9.2) remains valid if the constraint $(m, q_2) = 1$ in the innermost sum is removed and the denominator $\varphi(q_2 r)$ is replaced by $q_2\varphi(r)$. Namely we have

(9.3)

$$\mathcal{S}_2(r,a) = \sum_{q_1}\sum_{b_1}\sum_{(q_2,q_1)=1}\sum_{b_2} \frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{q_2\varphi(r)} \times \sum_{n_1}\sum_{(n_2,q_2r)=1} \beta(n_1)\beta(n_2) \sum_{\substack{mn_1\equiv a(r)\\ mn_1\equiv b_1(q_1)}} f(m) + O(xND_0^{-1}R^{-2}\mathcal{L}^B).$$

By Lemma 7, for $(n_1, q_1 r) = 1$ we have

$$\sum_{\substack{mn_1\equiv a(r)\\ mn_1\equiv b_1(q_1)}} f(m) = \frac{1}{q_1 r}\sum_{|h|<H_2} \hat{f}\left(\frac{h}{q_1 r}\right) e_{q_1 r}(-h\mu) + O(x^{-2}),$$

where

$$H_2 = 4QRM^{-1+2\varepsilon},$$

and $\mu(\bmod q_1 r)$ is a common solution to

$$\mu n_1 \equiv a(\bmod r), \qquad \mu n_1 \equiv b_1(\bmod q_1). \tag{9.4}$$

Inserting this into (9.3) we deduce that

$$\mathcal{S}_2(r,a) = \hat{f}(0)X(r,a) + \mathcal{R}_2(r,a) + O(xND_0^{-1}R^{-2}\mathcal{L}^B), \tag{9.5}$$

where

$$\mathcal{R}_2(r,a) = \sum_{q_1}\sum_{b_1}\sum_{(q_2,q_1)=1}\sum_{b_2}\frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{q_1q_2r\varphi(r)}\bigg(\sum_{(n_2,q_2r)=1}\beta(n_2)\bigg)$$
$$\times \sum_{(n_1,q_1r)=1}\beta(n_1)\sum_{1\le|h|<H_2}\hat{f}\bigg(\frac{h}{q_1r}\bigg)e_{q_1r}(-h\mu).$$

The proof of (9.1) is now reduced to estimating $\mathcal{R}_2(r,a)$. First we note that the second inequality in (7.5) implies

$$H_2 \ll x^{-1/2+2\varpi+2\varepsilon}N < 2x^{2\varpi+2\varepsilon} \tag{9.6}$$

since $M^{-1} \ll x^{-1}N$. (Here and in what follows, we use the second inequality in (7.5) only.) This implies that $\mathcal{R}_2(r,a) = 0$ if γ is of Type I.

Now assume that γ is of of Type II. Noting that

$$\frac{\mu}{q_1r} \equiv \frac{a\overline{q_1n_1}}{r} + \frac{b_1\overline{rn_1}}{q_1} \pmod 1$$

by (9.4), we have

$$\mathcal{R}_2(r,a) \ll N^{1+\varepsilon}R^{-2}\sum_{\substack{n\sim N\\(n,r)=1}}|\mathcal{R}^*(r,a;n)|, \tag{9.7}$$

where

$$\mathcal{R}^*(r,a;n) = \sum_{(q,n)=1}\sum_{b}\frac{c(r,a;q,b)}{q}\sum_{1\le|h|<H_2}\hat{f}\bigg(\frac{h}{qr}\bigg)e\bigg(\frac{-ah\overline{qn}}{r} - \frac{bh\overline{rn}}{q}\bigg).$$

To estimate the sum of $|\mathcal{R}^*(r,a;n)|$ we observe that

$$|\mathcal{R}^*(r,a;n)|^2 = \sum_{(q,n)=1}\sum_{b}\sum_{(q',n)=1}\sum_{b'}\frac{c(r,a;q,b)c(r,a;q',b')}{qq'}$$
$$\times \sum_{1\le|h|<H_2}\sum_{1\le|h'|<H_2}\hat{f}\bigg(\frac{h}{qr}\bigg)\overline{\hat{f}\bigg(\frac{h'}{q'r}\bigg)}e\bigg(\frac{a(h'\bar{q}'-h\bar{q})\bar{n}}{r} - \frac{bh\overline{rn}}{q} + \frac{b'h'\overline{rn}}{q'}\bigg).$$

It follows, by changing the order of summation and applying (8.1), that

(9.8)
$$M^{-2}\sum_{\substack{n\sim N\\(n,r)=1}}|\mathcal{R}^*(r,a;n)|^2 \ll \sum_{q}\sum_{b}\sum_{q'}\sum_{b'}\frac{|c(r,a;q,b)c(r,a;q',b')|}{qq'}$$
$$\times \sum_{1\le|h|<H_2}\sum_{1\le|h'|<H_2}|\mathcal{W}(r,a;q,b;q',b';h,h')|,$$

where

$$\mathcal{W}(r,a;q,b;q',b';h,h') = \sum_{\substack{n\sim N \\ (n,qq'r)=1}} e\left(\frac{a(h'\bar{q}'-h\bar{q})\bar{n}}{r} - \frac{bh\overline{rn}}{q} + \frac{b'h'\overline{rn}}{q'}\right).$$

Since $M^{-1} \ll N^{-1}$, by the second inequality in (7.4) and (7.2) we have

$$H_2 Q^{-1} \ll x^{-3\varpi+\varepsilon}. \tag{9.9}$$

It follows that, on the right side of (9.8), the contribution from the terms with $h'q = hq'$ is

$$\ll NQ^{-2} \sum_{1\le h<H_2} \sum_{q<2Q} \tau(hq) \ll x^{-3\varpi+\varepsilon} N. \tag{9.10}$$

Now assume that $c(r,a;q,b)c(r,a;q',b') \neq 0$, $1 \le |h| < H_2$, $1 \le |h'| < H_2$ and $h'q \neq hq'$. Letting $d = [q,q']r$, we have

$$\frac{a(h'\bar{q}'-h\bar{q})}{r} - \frac{bh\bar{r}}{q} + \frac{b'h\bar{r}}{q'} \equiv \frac{c}{d} \pmod 1$$

for some c with

$$(c,r) = (h'\bar{q}'-h\bar{q}, r).$$

It follows by the estimate (3.13) that

$$\mathcal{W}(r,a;q,b;q',b';h,h') \ll d^{1/2+\varepsilon} + \frac{(c,d)N}{d}. \tag{9.11}$$

Since $N > x_2$, by the first inequality in (7.4), (7.2) and (2.13) we have

$$R^{-1} < x^{4\varpi} N^{-1} < x^{-1/2+8\varpi}. \tag{9.12}$$

Together with (7.5), this implies that

$$Q \ll x^{10\varpi}. \tag{9.13}$$

By (9.13) and (7.5) we have

$$d^{1/2} \ll (Q^2R)^{1/2} \ll x^{1/4+6\varpi}.$$

On the other hand, noting that

$$h'\bar{q}' - h\bar{q} \equiv (h'q - hq')\overline{qq'} \pmod r,$$

we have

$$(c,d) \le (c,r)[q,q'] \ll [q,q']H_2 Q. \tag{9.14}$$

Together with (9.6), (9.12) and (9.13), this yields

$$\frac{(c,d)N}{d} \ll H_2 NQR^{-1} \ll x^{16\varpi+\varepsilon}.$$

Combining these estimates with (9.11) we deduce that

$$\mathcal{W}(r,a;q,b;q',b';h,h') \ll x^{1/4+7\varpi}.$$

Together with (9.6), this implies that, on the right side of (9.8), the contribution from the terms with $h'q \neq hq'$ is $\ll x^{1/4+12\varpi}$, which is sharper than the right side of (9.10). Combining these estimates with (9.8) we conclude that

$$\sum_{\substack{n\sim N\\(n,r)=1}} |\mathcal{R}^*(r,a;n)|^2 \ll x^{1-3\varpi+\varepsilon}M.$$

This yields, by Cauchy's inequality,

$$\sum_{\substack{n\sim N\\(n,r)=1}} |\mathcal{R}^*(r,a;n)| \ll x^{1-3\varpi/2+\varepsilon}.$$

Inserting this into (9.7) we obtain

$$\mathcal{R}_2(r,a) \ll x^{1-\varpi}NR^{-2}, \tag{9.15}$$

which is sharper than the O term in (9.5).

The relation (9.1) follows from (9.5) and (9.15) immediately.

10. A truncation of the sum of $\mathcal{S}_1(r,a)$

We are unable to evaluate each $\mathcal{S}_1(r,a)$ directly. However, we shall establish a relation of the form

$$\sum_r\sum_a \mathcal{S}_1(r,a) = \sum_r\sum_a \Big(\hat{f}(0)X(r,a) + \mathcal{R}_1(r,a)\Big) + O(xNR^{-1}\mathcal{L}^{-87A}) \tag{10.1}$$

with $\mathcal{R}_1(r,a)$ to be specified below in (10.10). In view of (8.3) and (9.1), the proof of (7.9) will be reduced to estimating $\mathcal{R}_1(r,a)$.

By definition we have

$$\mathcal{S}_1(r,a) = \sum_{q_1}\sum_{b_1}\sum_{q_2}\sum_{b_2} c(r,a;q_1,b_1)c(r,a;q_2,b_2) \times \sum_{n_1}\sum_{n_2\equiv n_1(r)} \beta(n_1)\beta(n_2) \sum_{\substack{mn_1\equiv a(r)\\ mn_1\equiv b_1(q_1)\\ mn_2\equiv b_2(q_2)}} f(m). \tag{10.2}$$

Let $\mathcal{U}(r,a;q_0)$ denote the sum of the terms in (10.2) with $(q_1,q_2)=q_0$. Clearly we have $\mathcal{U}(r,a;q_0)=0$ unless

$$q_0 < 2Q, \qquad q_0|\mathcal{P}, \qquad (q_0, r\mathcal{P}_0) = 1,$$

which are henceforth assumed. We first claim that

$$\sum_r\sum_a\sum_{q_0>1} \mathcal{U}(r,a;q_0) \ll xN(D_0R)^{-1}\mathcal{L}^B. \tag{10.3}$$

Assume that, for $j=1,2$,

$$q_j \sim Q, \quad q_j|\mathcal{P}, \quad (q_j, r\mathcal{P}_0)=1, \quad b_j \in \mathcal{C}_i(q_j)$$

and $(q_1, q_2) = q_0$. Write $q_1' = q_1/q_0$, $q_2' = q_2/q_0$. By Lemma 5, there exist $t_1, t_2 \in \mathcal{C}_i(q_0)$, $b_1' \in \mathcal{C}_i(q_1')$ and $b_2' \in \mathcal{C}_i(q_2')$ such that

$$b_j \equiv t_j (\operatorname{mod} q_0), \quad b_j \equiv b_j' (\operatorname{mod} q_j').$$

Note that the conditions $mn_1 \equiv t_1 (\operatorname{mod} q_0)$ and $mn_2 \equiv t_2 (\operatorname{mod} q_0)$ together imply that

$$t_2 n_1 \equiv t_1 n_2 (\operatorname{mod} q_0). \tag{10.4}$$

Thus the innermost sum in (10.2) is void if (10.4) fails to hold for any $t_1, t_2 \in \mathcal{C}_i(q_0)$. On the other hand, if (10.4) holds for some $t_1, t_2 \in \mathcal{C}_i(q_0)$, the innermost sum in (10.2) may be rewritten as

$$\sum_{\substack{mn_1 \equiv a_1 (q_0 r) \\ mn_1 \equiv b_1' (q_1') \\ mn_2 \equiv b_2' (q_2')}} f(m),$$

where $a_1 (\operatorname{mod} q_0 r)$ is a common solution to $a_1 \equiv a (\operatorname{mod} r)$ and $a_1 \equiv t_1 (\operatorname{mod} q_0)$. Hence, changing the order of summation we obtain

$$\begin{aligned} &\sum_{b_1} \sum_{b_2} c(r, a; q_1, b_1) c(r, a; q_2, b_2) \sum_{n_1} \sum_{n_2 \equiv n_1 (r)} \beta(n_1) \beta(n_2) \sum_{\substack{mn_1 \equiv a(r) \\ mn_1 \equiv b_1 (q_1) \\ mn_2 \equiv b_2 (q_2)}} f(m) \\ &\quad \ll \sum_{t_1 \in \mathcal{C}_i(q_0)} \sum_{t_2 \in \mathcal{C}_i(q_0)} \sum_{n_1} \sum_{\substack{n_2 \equiv n_1 (r) \\ t_1 n_2 \equiv t_2 n_1 (q_0)}} |\beta(n_1) \beta(n_2)| \\ &\quad \times \sum_{mn_1 \equiv a_1 (q_0 r)} f(m) \mathcal{J}(mn_1, q_1') \mathcal{J}(mn_2, q_2'), \end{aligned}$$

where

$$\mathcal{J}(n, q') = \sum_{\substack{b_j' \in \mathcal{C}_i(q') \\ b_j' \equiv n (q')}} 1.$$

This yields, by summing over q_1 and q_2 with $(q_1, q_2) = q_0$ and changing the order of summation,

$$\begin{aligned} \mathcal{U}(r, a; q_0) \ll &\sum_{t_1 \in \mathcal{C}_i(q_0)} \sum_{t_2 \in \mathcal{C}_i(q_0)} \sum_{n_1} \sum_{\substack{n_2 \equiv n_1 (r) \\ t_1 n_2 \equiv t_2 n_1 (q_0)}} |\beta(n_1) \beta(n_2)| \\ &\times \sum_{mn_1 \equiv a_1 (q_0 r)} f(m) \mathcal{X}(mn_1) \mathcal{X}(mn_2), \end{aligned} \tag{10.5}$$

where

$$\mathcal{X}(n) = \sum_{q' \sim Q/q_0} |\mu(q')| \mathcal{J}_i(n, q').$$

We may assume that $(n_1, q_0r) = 1$, since the innermost sum in (10.5) is void otherwise. Let $a_2(\bmod q_0r)$ be a common solution to $a_2 \equiv a_1(\bmod r)$ and $a_2 \equiv t_2(\bmod q_0)$. In the case

$$n_2 \equiv n_1(\bmod r), \quad t_1n_2 \equiv t_2n_1(\bmod q_0),$$

the condition $mn_1 \equiv a_1(\bmod q_0r)$ is equivalent to $mn_2 \equiv a_2(\bmod q_0r)$. Thus the innermost sum in (10.5) is

$$\leq \sum_{mn_1\equiv a_1(q_0r)} f(m)\mathcal{X}(mn_1)^2 + \sum_{mn_2\equiv a_2(q_0r)} f(m)\mathcal{X}(mn_2)^2.$$

Since

$$\mathcal{J}(n,q') = \begin{cases} 1 & \text{if} \quad q'|P(n-h_i),\ (q',n)=1, \\ 0 & \text{otherwise,} \end{cases}$$

it follows that

$$\mathcal{X}(n) \leq \sum_{\substack{q'|P(n-h_i)\\(q',n)=1}} |\mu(q')|. \tag{10.6}$$

Assume that $j = 1, 2$, $1 \leq \mu \leq k_0$ and $\mu \neq i$. Write

$$n_{j\mu} = \frac{n_j}{(n_j, h_\mu - h_i)}, \qquad h^*_{j\mu} = \frac{h_\mu - h_i}{(n_j, h_\mu - h_i)}.$$

Noting that the conditions $p|(mn_j + h_\mu - h_i)$ and $p \nmid n_j$ together imply that $p|(mn_{j\mu} + h^*_{j\mu})$, by (10.6) we have

$$\mathcal{X}(mn_j) \leq \prod_{\substack{1\leq\mu\leq k_0\\ \mu\neq i}} \tau(mn_{j\mu} + h^*_{j\mu}) \leq \sum_{\substack{1\leq\mu\leq k_0\\ \mu\neq i}} \tau(mn_{j\mu} + h^*_{j\mu})^{k_0-1}.$$

Since $(n_{j\mu}, h^*_{j\mu}) = (n_{j\mu}, q_0r) = 1$, it follows by Lemma 8 that

$$\sum_{mn_j\equiv a_j(q_0r)} f(m)\mathcal{X}(mn_j)^2 \ll \frac{M\mathcal{L}^B}{q_0r} + x^{\varepsilon/3}.$$

(Here the term $x^{\varepsilon/3}$ is necessary when $q_0r > x^{-\varepsilon/4}M$.) Combining these estimates with (10.5), we deduce that

$$\begin{aligned}\mathcal{U}(r,a;q_0) \ll \left(\frac{M\mathcal{L}^B}{q_0r} + x^{\varepsilon/3}\right) &\sum_{(n_1,q_0r)=1} |\beta(n_1)| \\ \times \sum_{t_1\in\mathcal{C}_i(q_0)} \sum_{t_2\in\mathcal{C}_i(q_0)} &\sum_{\substack{n_2\equiv n_1(r)\\ t_1n_2\equiv t_2n_1(q_0)}} |\beta(n_2)|.\end{aligned}$$

Using Lemma 8 again, we find that the innermost sum is

$$\ll \frac{N\mathcal{L}^B}{q_0r} + x^{\varepsilon/3}.$$

It follows that

$$\mathcal{U}(r,a;q_0) \ll \varrho_2(q_0)^2\left(\frac{xN\mathcal{L}^B}{(q_0r)^2} + \frac{x^{1+\varepsilon/2}}{q_0r} + x^\varepsilon N\right).$$

This leads to (10.3), since $NR^{-1} > x^\varepsilon$ and

$$NQR \ll x^{1/2+2\varpi}N \ll x^{1-\varpi}NR^{-1}$$

by (7.5), (7.1), (7.2) and the second inequality in (7.4).

We now turn to $\mathcal{U}(r,a;1)$. Assume $|\mu(q_1q_2r)| = 1$. In the case $(n_1, q_1r) = (n_2, q_2r) = 1$, the innermost sum in (10.2) is, by Lemma 7, equal to

$$\frac{1}{q_1q_2r}\sum_{|h|<H_1}\hat{f}\left(\frac{h}{q_1q_2r}\right)e_{q_1q_2r}(-\mu h) + O(x^{-2}),$$

where

$$H_1 = 8Q^2RM^{-1+2\varepsilon} \tag{10.7}$$

and $\mu(\operatorname{mod} q_1q_2r)$ is a common solution to

$$\mu n_1 \equiv a(\operatorname{mod} r), \quad \mu n_1 \equiv b_1(\operatorname{mod} q_1), \quad \mu n_2 \equiv b_2(\operatorname{mod} q_2). \tag{10.8}$$

It follows that

$$\mathcal{U}(r,a;1) = \hat{f}(0)X^*(r,a) + \mathcal{R}_1(r,a) + O(1), \tag{10.9}$$

where

$$X^*(r,a) = \sum_{q_1}\sum_{b_1}\sum_{(q_2,q_1)=1}\sum_{b_2}\frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{q_1q_2r} \times \sum_{(n_1,q_1r)=1}\sum_{\substack{n_2\equiv n_1(r)\\(n_2,q_2)=1}}\beta(n_1)\beta(n_2)$$

and

$$\mathcal{R}_1(r,a) = \sum_{q_1}\sum_{b_1}\sum_{(q_2,q_1)=1}\sum_{b_2}\frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{q_1q_2r} \times \sum_{(n_1,q_1r)=1}\sum_{\substack{n_2\equiv n_1(r)\\(n_2,q_2)=1}}\beta(n_1)\beta(n_2)\sum_{1\le|h|<H}\hat{f}\left(\frac{h}{q_1q_2r}\right)e_{q_1q_2r}(-\mu h). \tag{10.10}$$

By (10.2), (10.3) and (10.9) we conclude that

$$\sum_r\sum_a \mathcal{S}_1(r,a) = \sum_r\sum_a\left(\hat{f}(0)X^*(r,a) + \mathcal{R}_1(r,a)\right) + O(xN(D_0R)^{-1}\mathcal{L}^B).$$

In view of (8.1), the proof of (10.1) is now reduced to showing that

$$\sum_r\sum_a(X^*(r,a) - X(r,a)) \ll N^2R^{-1}\mathcal{L}^{-87A}. \tag{10.11}$$

We have

$$X^*(r,a) - X(r,a) = \sum_{q_1} \sum_{b_1} \sum_{(q_2,q_1)=1} \sum_{b_2} \frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{q_1q_2r} \mathcal{V}(r;q_1,q_2),$$

with

$$\mathcal{V}(r;q_1,q_2) = \sum_{(n_1,q_1r)=1} \sum_{\substack{n_2 \equiv n_1 (r) \\ (n_2,q_2)=1}} \beta(n_1)\beta(n_2) - \frac{1}{\varphi(r)} \sum_{(n_1,q_1r)=1} \sum_{(n_2,q_2r)=1} \beta(n_1)\beta(n_2),$$

which is independent of a. It follows that

$$\sum_r \sum_a (X^*(r,a) - X(r,a)) \ll \frac{1}{R} \sum_{q_1 \sim Q} \sum_{q_2 \sim Q} \frac{\varrho_2(q_1q_2)}{q_1q_2} \sum_{\substack{r \sim R \\ (r,q_1q_2)=1}} \varrho_2(r)|\mathcal{V}(r;q_1,q_2)|. \tag{10.12}$$

Noting that

$$\mathcal{V}(r;q_1,q_2) = \sideset{}{^*}\sum_{l \bmod r} \Bigg(\sum_{\substack{n \equiv l(r) \\ (n,q_1)=1}} \beta(n) - \frac{1}{\varphi(r)} \sum_{(n,q_1r)=1} \beta(n) \Bigg) \times \Bigg(\sum_{\substack{n \equiv l(r) \\ (n,q_2)=1}} \beta(n) - \frac{1}{\varphi(r)} \sum_{(n,q_2r)=1} \beta(n) \Bigg),$$

by Cauchy's inequality, the condition (A_2) and Lemma 10, we find that the innermost sum in (10.12) is

$$\ll \tau(q_1q_2)^B N^2 \mathcal{L}^{-100A},$$

whence (10.11) follows.

A combination of (8.3), (9.1) and (10.1) leads to

$$\sum_r \sum_a \big(\mathcal{S}_1(r.a) - 2\mathcal{S}_2(r,a) + \mathcal{S}_3(r,a) \big) = \sum_r \sum_a \mathcal{R}_1(r,a) + O(xNR^{-1}\mathcal{L}^{-87A}). \tag{10.13}$$

Note that

$$\frac{\mu}{q_1q_2r} \equiv \frac{a\overline{q_1q_2n_1}}{r} + \frac{b_1\overline{q_2rn_1}}{q_1} + \frac{b_2\overline{q_1rn_2}}{q_2} \pmod 1$$

by (10.8). Hence, on substituting $n_2 = n_1 + kr$, we may rewrite $\mathcal{R}_1(r,a)$ as

$$\mathcal{R}_1(r,a) = \frac{1}{r} \sum_{|k|<N/R} \mathcal{R}_1(r,a;k), \tag{10.14}$$

where

$$\mathcal{R}_1(r,a;k) = \sum_{q_1}\sum_{b_1}\sum_{(q_2,q_1)=1}\sum_{b_2}\frac{c(r,a;q_1,b_1)c(r,a;q_2,b_2)}{q_1q_2}\sum_{1\le|h|<H_1}\hat{f}\Big(\frac{h}{q_1q_2r}\Big)$$
$$\times \sum_{\substack{(n,q_1r)=1\\(n+kr,q_2)=1}}\beta(n)\beta(n+kr)e(-h\xi(r,a;q_1,b_1;q_2,b_2;n,k)),$$

with

$$\xi(r,a;k;q_1,b_1;q_2,b_2;n) = \frac{a\overline{q_1q_2n}}{r}+\frac{b_1\overline{q_2rn}}{q_1}+\frac{b_2\overline{q_1r(n+kr)}}{q_2}.$$

Recall that, in the Type I and II cases, we have reduced the proof of (6.2) to proving (7.9) at the end of Section 7. Now, by (10.13) and (10.14), the proof of (7.9) is in turn reduced to showing that

$$\mathcal{R}_1(r,a;k) \ll x\mathcal{L}^{-88A}$$

for $|k| < NR^{-1}$. In fact, we shall prove the sharper bound

$$\mathcal{R}_1(r,a;k) \ll x^{1-\varpi/2} \tag{10.15}$$

in the next two sections.

We conclude this section by showing that the gap between (10.15) and some trivial bounds is not too large. It trivially follows from (8.1) that

$$\mathcal{R}_1(r,a;k) \ll x^{1+\varepsilon}H_1.$$

On the other hand, in view of (2.13), since

$$H_1 \ll x^{\varepsilon}(QR)^2(MN)^{-1}NR^{-1}$$

and, by the first inequality in (7.4), (7.1) and (7.2),

$$NR^{-1} < \begin{cases} x^{\varpi+\varepsilon} & \text{if} \quad x_1 < N \le x_2,\\ x^{4\varpi} & \text{if} \quad x_2 < N < 2x^{1/2},\end{cases} \tag{10.16}$$

it follows from (7.5) that

$$H_1 \ll \begin{cases} x^{5\varpi+2\varepsilon} & \text{if} \quad x_1 < N \le x_2,\\ x^{8\varpi+\varepsilon} & \text{if} \quad x_2 < N < 2x^{1/2}.\end{cases} \tag{10.17}$$

Thus, in order to prove (10.15), we need only to save a small power of x from the trivial estimate.

The bounds (10.16) and (10.17) will find application in the next two sections.

11. Estimation of $\mathcal{R}_1(r,a;k)$: The Type I case

In this and the next sections we assume that $|k| < NR^{-1}$, and we abbreviate

$$\mathcal{R}_1, \quad c(q_1,b_1), \quad c(q_2,b_2) \quad \text{and} \quad \xi(q_1,b_1;q_2,b_2;n)$$

for

$$\mathcal{R}_1(r,a;k), \quad c(r,a;q_1,b_1), \quad c(r,a;q_2,b_2) \quad \text{and} \quad \xi(r,a,k;q_1,b_1;q_2,b_2;n)$$

respectively, with the aim of proving (10.15). The variables r, a and k may also be omitted somewhere else for notational simplicity. The proof is somewhat analogous to the estimation of $\mathcal{R}_2(r,a)$ in Section 9; the main tool we need is Lemma 11.

Assume that $x_1 < N \le x_2$ and R^* is as in (7.1). We have

$$\mathcal{R}_1 \ll N^{\varepsilon} \sum_{q_1} \sum_{b_1} \frac{|c(q_1,b_1)|}{q_1} \sum_{\substack{n\sim N \\ (n,q_1r)=1}} |\mathcal{F}(q_1,b_1;n)|, \tag{11.1}$$

where

$$\mathcal{F}(q_1,b_1;n) = \sum_{1\le|h|<H_1} \sum_{(q_2,q_1(n+kr))=1} \times \sum_{b_2} \frac{c(q_2,b_2)}{q_2} \hat{f}\left(\frac{h}{q_1q_2r}\right) e(-h\xi(q_1,b_1;q_2,b_2;n)).$$

In what follows assume $c(q_1,b_1) \ne 0$. To estimate the sum of $|\mathcal{F}(q_1,b_1;n)|$ we observe that, similar to (9.8),

(11.2)

$$M^{-2} \sum_{\substack{n\sim N \\ (n,q_1r)=1}} |\mathcal{F}(q_1,b_1;n)|^2 \ll \sum_{(q_2,q_1)=1} \sum_{(q_2',q_1)=1} \sum_{b_2} \sum_{b_2'} \frac{|c(q_2,b_2)c(q_2',b_2')|}{q_2q_2'} \times \sum_{1\le|h|<H_1} \sum_{1\le|h'|<H_1} |\mathcal{G}(h,h';q_1,b_1,q_2,b_2;q_2',b_2')|,$$

where

$$\mathcal{G}(h,h';q_1,b_1,q_2,b_2;q_2',b_2') = \sum_{\substack{n\sim N \\ (n,q_1r)=1 \\ (n+kr,q_2q_2')=1}} e\Big(h'\xi(q_1,b_1;q_2',b_2';n) - h\xi(q_1,b_1;q_2,b_2;n)\Big).$$

The condition $N \le x_2$ is essential for bounding the terms with $h'q_2 = hq_2'$ in (11.2). By (7.5) we have

$$H_1Q^{-1} \ll x^{\varepsilon}(QR)(MN)^{-1}N \ll x^{-2\varpi+\varepsilon}.$$

It follows that, on the right side of (11.2), the contribution from the terms with $h'q_2 = hq_2'$ is

$$\ll NQ^{-2} \sum_{1\le h<H_1} \sum_{q\sim Q} \tau(hq)^B \ll x^{-2\varpi+\varepsilon}N. \tag{11.3}$$

Now assume that $c(q_2, b_2)c(q_2', b_2') \neq 0$, $(q_2q_2', q_1) = 1$ and $h'q_2 \neq hq_2'$. We have

$$\begin{aligned} &h'\xi(q_1, b_1; q_2', b_2'; n) - h\xi(q_1, b_1; q_2, b_2; n) \\ &\quad\equiv \frac{(h'\overline{q_2'} - h\bar{q}_2)a\overline{q_1 n}}{r} \\ &\qquad + \frac{(h'\overline{q_2'} - h\bar{q}_2)b_1\overline{rn}}{q_1} + \frac{h'b_2'\overline{q_1 r(n+kr)}}{q_2'} - \frac{hb_2\overline{q_1 r(n+kr)}}{q_2} \pmod 1. \end{aligned}$$

Letting $d_1 = q_1 r$ and $d_2 = [q_2, q_2']$, we may write

$$\frac{(h'\overline{q_2'} - h\bar{q}_2)a\bar{q}_1}{r} + \frac{(h'\overline{q_2'} - h\bar{q}_2)b_1\bar{r}}{q_1} \equiv \frac{c_1}{d_1} \pmod 1$$

for some c_1 with

$$(c_1, r) = (h'\overline{q_2'} - h\bar{q}_2, r),$$

and

$$\frac{h'b_2'\overline{q_1 r}}{q_2'} - \frac{hb_2\overline{q_1 r}}{q_2} \equiv \frac{c_2}{d_2} \pmod 1$$

for some c_2, so that

$$h'\xi(q_1, b_1; q_2', b_2'; n) - h\xi(q_1, b_1; q_2, b_2; n) \equiv \frac{c_1\bar{n}}{d_1} + \frac{c_2\overline{(n+kr)}}{d_2} \pmod 1.$$

Since $(d_1, d_2) = 1$, it follows by Lemma 11 that

$$\mathcal{G}(h, h'; q_1, b_1, q_2, b_2; q_2', b_2') \ll (d_1d_2)^{1/2+\varepsilon} + \frac{(c_1, d_1)N}{d_1}. \tag{11.4}$$

We appeal to the condition $N > x_1$ that gives, by (10.16),

$$R^{-1} < x^{\varpi+\varepsilon}N^{-1} < x^{-3/4-15\varpi+\varepsilon}N. \tag{11.5}$$

Together with (7.5), this yields

$$(d_1d_2)^{1/2} \ll (Q^3R)^{1/2} \ll x^{3/4+3\varpi}R^{-1} \ll x^{-12\varpi+\varepsilon}N.$$

A much sharper bound for the second term on the right side of (11.4) can be obtained. In a way similar to the proof of (9.14), we find that

$$(c_1, d_1) \le (c_1, r)q_1 \ll H_1Q^2.$$

It follows by (10.17), (7.5) and the first inequality in (11.5) that

$$\frac{(c_1, d_1)}{d_1} \ll H_1(QR)R^{-2} \ll x^{1/2+9\varpi+4\varepsilon}N^{-2} \ll x^{-1/4-6\varpi}.$$

Here we have used the condition $N > x_1$ again. Combining these estimates with (11.4) we deduce that

$$\mathcal{G}(h, h'; q_1, b_1, q_2, b_2; q_2', b_2') \ll x^{-12\varpi+\varepsilon} N.$$

Together with (10.17), this implies that, on the right side of (11.2), the contribution from the terms with $h'q_2 \neq hq_2'$ is

$$\ll x^{-12\varpi+\varepsilon} H_1^2 N \ll x^{-2\varpi+5\varepsilon} N,$$

which has the same order of magnitude as the right side of (11.3) essentially. Combining these estimates with (11.2) we obtain

$$\sum_{\substack{n\sim N\\(n,q_1r)=1}} |\mathcal{F}(h; q_1, b_1; n)|^2 \ll x^{1-2\varpi+5\varepsilon} M.$$

This yields, by Cauchy's inequality,

$$\sum_{\substack{n\sim N\\(n,q_1r)=1}} |\mathcal{F}(h; q_1, b_1; n)| \ll x^{1-\varpi+3\varepsilon}. \tag{11.6}$$

The estimate (10.15) follows from (11.1) and (11.6) immediately.

12. Estimation of $\mathcal{R}_1(r, a; k)$: The Type II case

Assume that $x_2 < N < 2x^{1/2}$ and R^* is as in (7.2). We have

$$\mathcal{R}_1 \ll N^\varepsilon \sum_{\substack{n\sim N\\(n,r)=1}} |\mathcal{K}(n)|, \tag{12.1}$$

where

$$\mathcal{K}(n) = \sum_{(q_1,n)=1} \sum_{b_1} \sum_{(q_2,q_1(n+kr))=1} \sum_{b_2} \frac{c(q_1,b_1)c(q_2,b_2)}{q_1q_2} \times \sum_{1\le|h|<H_1} \hat{f}\left(\frac{h}{q_1q_2r}\right) e(-h\xi(q_1, b_1; q_2, b_2; n)).$$

Let $\sum^{\#}$ stand for a summation over the 8-tuples $(q_1, b_1; q_2, b_2; q_1', b_1'; q_2', b_2')$ with

$$(q_1, q_2) = (q_1', q_2') = 1.$$

To estimate the sum of $|\mathcal{K}(n)|$ we observe that, similar to (9.8),

(12.2)

$$M^{-2} \sum_{\substack{n\sim N\\(n,r)=1}} |\mathcal{K}(n)|^2 \ll \sum\nolimits^{\#} \frac{|c(q_1,b_1)c(q_2,b_2)c(q_1',b_1')c(q_2',b_2')|}{q_1q_2q_1'q_2'} \times \sum_{1\le|h|<H_1} \sum_{1\le|h'|<H_1} |\mathcal{M}(h, h'; q_1, b_1; q_2, b_2; q_1', b_1'; q_2', b_2')|,$$

where

$$\mathcal{M}(h, h'; q_1, b_1; q_2, b_2; q_1', b_1'; q_2', b_2') \\ = \sum_{n\sim N}{}' e(h'\xi(q_1', b_1'; q_2', b_2'; n) - h\xi(q_1, b_1; q_2, b_2; n)).$$

Here $\sum'$ is restriction to $(n, q_1q_1'r) = (n + kr, q_2q_2') = 1$.

Similar to (9.9), we have

$$H_1Q^{-2} \ll x^{-3\varpi+\varepsilon}.$$

Hence, on the right side of (12.2), the contribution from the terms with $h'q_1q_2 = hq_1'q_2'$ is

$$\ll NQ^{-4} \sum_{1\le h<H_1} \sum_{q\sim Q} \sum_{q'\sim Q} \tau(hqq')^B \ll x^{-3\varpi+\varepsilon}N. \tag{12.3}$$

Note that the bounds (9.12) and (9.13) are valid in the present situation. Since R is near to $x^{1/2}$ in the logarithmic scale and Q is small, it can be shown via Lemma 11 that the terms with $h'q_1q_2 \neq hq_1'q_2'$ on the right side of (12.2) make a small contribution in comparison with (12.3). Assume that

$$c(q_1, b_1)c(q_2, b_2)c(q_1', b_1')c(q_2', b_2') \neq 0, \quad (q_1, q_2) = (q_1', q_2') = 1, \quad h'q_1q_2 \neq hq_1'q_2'.$$

We have

$$h'\xi(q_1', b_1'; q_2', b_2'; n) - h\xi(q_1, b_1; q_2, b_2; n) \\ \equiv \frac{s\bar{n}}{r} + \frac{t_1\bar{n}}{q_1} + \frac{t_1'\bar{n}}{q_1'} + \frac{t_2\overline{(n+kr)}}{q_2} + \frac{t_2'\overline{(n+kr)}}{q_2'} \pmod 1 \tag{12.4}$$

with

$$s \equiv a(h'\overline{q_1'q_2'} - h\overline{q_1q_2}) (\mathrm{mod}\, r),\ \ t_1 \equiv -b_1h\overline{q_2r} (\mathrm{mod}\, q_1),\ \ t_1' \equiv b_1'h'\overline{q_2'r} (\mathrm{mod}\, q_1'),$$
$$t_2 \equiv -b_2h\overline{q_1r} (\mathrm{mod}\, q_2), \quad t_2' \equiv b_2'h'\overline{q_1'r} (\mathrm{mod}\, q_2').$$

Letting $d_1 = [q_1, q_1']r$, $d_2 = [q_2, q_2']$, we may rewrite (12.4) as

$$h'\xi(q_1', b_1'; q_2', b_2'; n) - h\xi(q_1, b_1; q_2, b_2; n) \equiv \frac{c_1\bar{n}}{d_1} + \frac{c_2\overline{(n+kr)}}{d_2} \pmod 1$$

for some c_1 and c_2 with

$$(c_1, r) = (h'\overline{q_1'q_2'} - h\overline{q_1q_2}, r).$$

It follows by Lemma 11 that

$$\mathcal{M}(h, h'; q_1, b_1; q_2, b_2; q_1', b_1'; q_2', b_2') \ll (d_1d_2)^{1/2+\varepsilon} + \frac{(c_1, d_1)(d_1, d_2)^2N}{d_1}. \tag{12.5}$$

By (7.5) and (9.13) we have

$$(d_1d_2)^{1/2} \ll (Q^4R)^{1/2} \ll x^{1/4+16\varpi}.$$

On the other hand, we have $(d_1, d_2) \leq (q_1q_1', q_2q_2') \ll Q^2$, since $(q_2q_2', r) = 1$, and, similar to (9.14),

$$(c_1, d_1) \leq (c_1, r)[q_1, q_1'] \ll [q_1, q_1']H_1Q^2.$$

It follows by (10.16), (9.13) and the first inequality in (9.12) that

$$\frac{(c_1, d_1)(d_1, d_2)^2 N}{d_1} \ll H_1 N Q^6 R^{-1} \ll x^{72\varpi}.$$

Combining these estimates with (12.5) we deduce that

$$\mathcal{M}(h, h'; q_1, b_1; q_2, b_2; q_1', b_1'; q_2', b_2') \ll x^{1/4+16\varpi+\varepsilon}.$$

Together with (10.16), this implies that, on the right side of (12.2), the contribution from the terms with $h'q_1q_2 \neq hq_1'q_2'$ is

$$\ll x^{1/4+16\varpi+\varepsilon}H_1^2 \ll x^{1/4+33\varpi},$$

which is sharper than the right side of (12.3). Combining these estimates with (12.2) we obtain

$$\sum_{\substack{n\sim N\\(n,r)=1}} |\mathcal{K}(n)|^2 \ll x^{1-3\varpi+\varepsilon}M.$$

This yields, by Cauchy's inequality,

$$\sum_{\substack{n\sim N\\(n,r)=1}} |\mathcal{K}(n)| \ll x^{1-\varpi}. \tag{12.6}$$

The estimate (10.15) follows from (12.1) and (12.6) immediately.

13. The Type III estimate: Initial steps

Assume that $\gamma = \alpha * \varkappa_{N_3} * \varkappa_{N_2} * \varkappa_{N_1}$ is of Type III. Our aim is to prove that

$$\Delta(\gamma; d, c) \ll \frac{x^{1-\varepsilon/2}}{d} \tag{13.1}$$

for any d and c satisfying

$$(d, c) = 1, \quad x^{1/2-\varepsilon} < d < x^{1/2+2\varpi}, \quad d|\mathcal{P}, \quad (d, \mathcal{P}_0) < D_1,$$

which are henceforth assumed. This leads to (6.2).

We first derive some lower bounds for the N_j from (A_4) and (A_5). We have

$$N_1 \geq N_2 \geq \left(\frac{x}{MN_1}\right)^{1/2} \geq x^{5/16-4\varpi} \tag{13.2}$$

and

$$N_3 \geq \frac{x}{MN_1N_2} \geq x^{1/4-16\varpi}M \geq x^{1/4-16\varpi}. \tag{13.3}$$

Let f be as in Lemma 7 with $\eta^* = \eta$ and with N_1 in place of M. Note that the function $\varkappa_{N_1} - f$ is supported on $[N_1^-, N_1] \cup [\eta N_1, \eta N_1^+]$ with $N_1^\pm = (1 \pm N_1^{-\varepsilon})N_1$. Letting $\gamma^* = \alpha * \varkappa_{N_3} * \varkappa_{N_2} * f$, we have

$$\frac{d}{\varphi(d)} \sum_{(n,d)=1} (\gamma - \gamma^*)(n) \ll x^{1-\varepsilon/2}$$

and

$$\sum_{n \equiv c(d)} (\gamma - \gamma^*)(n) \ll \mathcal{L} \sum_{\substack{N_1^- \le q \le N_1 \\ (q,d)=1}} \sum_{\substack{1 \le l < 3x/q \\ lq \equiv c(d)}} \tau_{19}(l)$$

$$+ \mathcal{L} \sum_{\substack{\eta N_1 \le q \le \eta N_1^+ \\ (q,d)=1}} \sum_{\substack{1 \le l < 3x/q \\ lq \equiv c(d)}} \tau_{19}(l) \ll \frac{x^{1-\varepsilon/2}}{d}.$$

It therefore suffices to prove (13.1) with γ replaced by γ^*. In fact, we shall prove the sharper bound

$$\Delta(\gamma^*; d, c) \ll \frac{x^{1-\varpi/3}}{d}, \tag{13.4}$$

In a way similar to the proof of (8.2), we obtain

$$\sum_{(n,d)=1} f(n) = \frac{\varphi(d)}{d} \hat{f}(0) + O(x^\varepsilon).$$

This yields, by (13.2),

$$\frac{1}{\varphi(d)} \sum_{(n,d)=1} \gamma^*(n) = \frac{\hat{f}(0)}{d} \sum_{(m,d)=1} \sum_{\substack{n_3 \simeq N_3 \\ (n_3,d)=1}} \sum_{\substack{n_2 \simeq N_2 \\ (n_2,d)=1}} \alpha(m) + O(d^{-1}x^{3/4}).$$

Here and in what follows, $n \simeq N$ stands for $N \le n < \eta N$. On the other hand, we have

$$\sum_{n \equiv c(d)} \gamma^*(n) = \sum_{(m,d)=1} \sum_{\substack{n_3 \simeq N_3 \\ (n_3,d)=1}} \sum_{\substack{n_2 \simeq N_2 \\ (n_2,d)=1}} \alpha(m) \sum_{mn_3n_2n_1 \equiv c(d)} f(n_1).$$

The innermost sum is, by Lemma 7, equal to

$$\frac{1}{d} \sum_{|h| < H^*} \hat{f}(h/d) e_d\big(-ch\overline{mn_3n_2}\big) + O(x^{-2}),$$

where

$$H^* = dN_1^{-1+2\varepsilon}.$$

It follows that

$$\Delta(\gamma^*; d, c) = \frac{1}{d} \sum_{\substack{m \simeq M \\ (m,d)=1}} \sum_{\substack{n_3 \simeq N_3 \\ (n_3,d)=1}} \sum_{\substack{n_2 \simeq N_2 \\ (n_2,d)=1}} \alpha(m) \times \sum_{1 \le |h| < H^*} \hat{f}(h/d) e_d\Big(- ch\overline{mn_3n_2}\Big) + O(d^{-1}x^{3/4}).$$

The proof of (13.4) is therefore reduced to showing that

$$\sum_{1 \le h < H^*} \sum_{\substack{n_3 \simeq N_3 \\ (n_3,d)=1}} \sum_{\substack{n_2 \simeq N_2 \\ (n_2,d)=1}} \hat{f}(h/d) e_d\Big(ah\overline{n_3n_2}\Big) \ll x^{1-\varpi/2+2\varepsilon} M^{-1} \tag{13.5}$$

for any a with $(a, d) = 1$.

On substituting $d_1 = d/(h, d)$ and applying Möbius inversion, the left side of (13.5) may be rewritten as

$$\begin{aligned} &\sum_{d_1 | d} \sum_{\substack{1 \le h < H \\ (h,d_1)=1}} \sum_{\substack{n_3 \simeq N_3 \\ (n_3,d)=1}} \sum_{\substack{n_2 \simeq N_2 \\ (n_2,d)=1}} \hat{f}(h/d_1) e_{d_1}\Big(ah\overline{n_3n_2}\Big) \\ &= \sum_{d_1 d_2 = d} \sum_{b_3 | d_2} \sum_{b_2 | d_2} \mu(b_3)\mu(b_2) \sum_{\substack{1 \le h < H \\ (h,d_1)=1}} \sum_{\substack{n_3 \simeq N_3/b_3 \\ (n_3,d_1)=1}} \\ &\quad \times \sum_{\substack{n_2 \simeq N_2/b_2 \\ (n_2,d_1)=1}} \hat{f}(h/d_1) e_{d_1}\Big(ah\overline{b_3b_2n_3n_2}\Big), \end{aligned}$$

where

$$H = d_1 N_1^{-1+2\varepsilon}. \tag{13.6}$$

It therefore suffices to show that

$$\sum_{\substack{1 \le h < H \\ (h,d_1)=1}} \sum_{\substack{n_3 \simeq N_3' \\ (n_3,d_1)=1}} \sum_{\substack{n_2 \simeq N_2' \\ (n_2,d_1)=1}} \hat{f}(h/d_1) e_{d_1}\Big(bh\overline{n_3n_2}\Big) \ll x^{1-\varpi/2+\varepsilon} M^{-1} \tag{13.7}$$

for any d_1, b, N_3' and N_2' satisfying

$$d_1 | d, \quad (b, d_1) = 1, \quad \frac{d_1 N_3}{d} \le N_3' \le N_3, \quad \frac{d_1 N_2}{d} \le N_2' \le N_2, \tag{13.8}$$

which are henceforth assumed. Note that (13.2) implies

$$H \ll x^{3/16+6\varpi+\varepsilon}. \tag{13.9}$$

In view of (13.6), the left side of (13.7) is void if $d_1 \le N_1^{1-2\varepsilon}$, so we may assume $d_1 > N_1^{1-2\varepsilon}$. By the trivial bound

$$\hat{f}(z) \ll N_1, \tag{13.10}$$

and (3.13), we find that the left side of (13.7) is

$$\ll H N_3 N_1 \Big(d_1^{1/2+\varepsilon} + d_1^{-1} N_2 \Big) \ll d_1^{3/2+\varepsilon} N_1^{2\varepsilon} N_3.$$

In the case $d_1 \le x^{5/12-6\varpi}$, the right side is $\ll x^{1-\varpi+3\varepsilon}M^{-1}$ by (A_4) and (2.13). This leads to (13.7). Thus we may further assume

$$d_1 > x^{5/12-6\varpi}. \tag{13.11}$$

We appeal to the Weyl shift and the factorization (2.8) with d_1 in place of d. By Lemma 4, we can choose a factor r of d_1 such that

$$x^{44\varpi} < r < x^{45\varpi}. \tag{13.12}$$

Write

$$\mathcal{N}(d_1,k) = \sum_{\substack{1\le h<H\\(h,d_1)=1}} \sum_{\substack{n_3\simeq N_3'\\(n_3,d_1)=1}} \sum_{\substack{n_2\simeq N_2'\\(n_2+hkr,d_1)=1}} \hat{f}(h/d_1)e_{d_1}\Big(bh\overline{(n_2+hkr)n_3}\Big),$$

so that the left side of (13.7) is just $\mathcal{N}(d_1,0)$. Assume $k>0$. We have

$$\mathcal{N}(d_1,k) - \mathcal{N}(d_1,0) = \mathcal{Q}_1(d_1,k) - \mathcal{Q}_2(d_1,k), \tag{13.13}$$

where

$$\mathcal{Q}_i(d_1,k) = \sum_{\substack{1\le h<H\\(h,d_1)=1}} \sum_{\substack{n_3\simeq N_3'\\(n_3,d_1)=1}} \sum_{\substack{l\in\mathcal{I}_i(h)\\(l,d_1)=1}} \hat{f}(h/d_1)e_{d_1}\Big(bh\overline{ln_3}\Big), \qquad i=1,2,$$

with

$$\mathcal{I}_1(h) = \Big[\eta N_2',\, \eta N_2' + hkr\Big), \qquad \mathcal{I}_2(h) = \Big[N_2',\, N_2'+hkr\Big).$$

To estimate $\mathcal{Q}_i(d_1,k)$ we first note that, by Möbius inversion,

$$\mathcal{Q}_i(d_1,k) = \sum_{st=d_1} \mu(s) \sum_{1\le h<H/s} \sum_{\substack{n_3\simeq N_3'\\(n_3,d_1)=1}} \sum_{\substack{l\in\mathcal{I}_i(h)\\(l,d_1)=1}} \hat{f}(h/t)e_t\Big(bh\overline{ln_3}\Big).$$

The inner sum is void unless $s<H$. Since $H^2 = o(d_1)$ by (13.9) and (13.11), it follows, by changing the order of summation, that

$$|\mathcal{Q}_i(d_1,k)| \le \sum_{\substack{st=d_1\\t>H}} \sum_{\substack{n_3\simeq N_3'\\(n_3,d_1)=1}} \sum_{\substack{l\in\mathcal{I}_i(H)\\(l,d_1)=1}} \Bigg| \sum_{h\in J_i(s,l)} \hat{f}(h/t)e_t\Big(bh\overline{ln_3}\Big)\Bigg|,$$

where $J_i(s,l)$ is a certain interval of length $<H$ and depending on s and l. Noting that, by integration by parts,

$$\frac{d}{dz}\hat{f}(z) \ll \min\Big\{N_1^2, |z|^{-2}N_1^{\varepsilon}\Big\},$$

by partial summation and (13.10) we obtain

$$\sum_{h\in J_i(s,l)} \hat{f}(h/t)e_t\Big(bh\overline{ln_3}\Big) \ll N_1^{1+\varepsilon}\min\Big\{H,\, ||b\overline{ln_3}/t||^{-1}\Big\}.$$

It follows that

$$\mathcal{Q}_i(d_1,k) \ll N_1^{1+\varepsilon} \sum_{\substack{t|d_1\\ t>H}} \sum_{\substack{l\in\mathcal{I}_i(H)\\ (l,d_1)=1}} \sum_{\substack{n_3<2N_3\\ (n_3,d_1)=1}} \min\left\{H,\, ||b\overline{ln_3}/t||^{-1}\right\}.$$

Since $H = o(N_3)$ by (13.3) and (13.9), the innermost sum is $\ll N_3^{1+\varepsilon}$ by Lemma 9. In view of (13.6), this leads to

$$\mathcal{Q}_i(d_1,k) \ll d_1^{1+\varepsilon} krN_3. \tag{13.14}$$

We now introduce the parameter

$$K = [x^{-1/2-48\varpi} N_1 N_2], \tag{13.15}$$

which is $\gg x^{1/8-56\varpi}$ by (13.2). By (A_5) and the second inequality in (13.12), we see that the right side of (13.14) is $\ll x^{1-\varpi+\varepsilon}M^{-1}$ if $k < 2K$. Hence, by (13.13), the proof of (13.7) is reduced to showing that

$$\frac{1}{K}\sum_{k\sim K} \mathcal{N}(d_1,k) \ll x^{1-\varpi/2+\varepsilon}M^{-1}. \tag{13.16}$$

14. The Type III estimate: Completion

The aim of this section is to prove (13.16), which will complete the proof of Theorem 2.

We start with the relation

$$h\overline{(n_2+hkr)} \equiv \overline{l+kr} \pmod{d_1}$$

for $(h,d_1) = (n_2+hkr, d_1) = 1$, where $l \equiv \bar{h}n_2 \,(\bmod\, d_1)$. Thus we may rewrite $\mathcal{N}(d_1,k)$ as

$$\mathcal{N}(d_1,k) = \sum_{\substack{l(\bmod d_1)\\ (l+kr,d_1)=1}} \nu(l;d_1) \sum_{\substack{n_3\simeq N_3'\\ (n_3,d_1)=1}} e_{d_1}\Big(b\overline{(l+kr)n_3}\Big),$$

with

$$\nu(l;d_1) = \sideset{}{'}\sum_{\bar{h}n_2\equiv l(d_1)} \hat{f}(h/d_1).$$

Here $\sum'$ is restriction to $1\le h<H$, $(h,d_1)=1$ and $n_2\simeq N_2'$. It follows by Cauchy's inequality that

$$\Bigg|\sum_{k\sim K}\mathcal{N}(d_1,k)\Bigg|^2 \le P_1P_2, \tag{14.1}$$

where

$$P_1 = \sum_{l(\bmod d_1)} |\nu(l;d_1)|^2, \quad P_2 = \sum_{l(\bmod d_1)} \Bigg| \sum_{\substack{k\sim K\\ (l+kr,d_1)=1}} \sum_{\substack{n\simeq N_3'\\ (n,d_1)=1}} e_{d_1}\Big(b\overline{(l+kr)n}\Big)\Bigg|^2.$$

The estimation of P_1 is straightforward. By (13.10) we have

$$P_1 \ll N_1^2 \,\#\{(h_1, h_2; n_1, n_2) :\, h_2 n_1 \equiv h_1 n_2 (\mathrm{mod}\, d_1),\, 1 \le h_i < H,\, n_i \simeq N_2'\}.$$

The number of the 4-tuples $(h_1, h_2; , n_1, n_2)$ satisfying the above conditions is

$$\ll \sum_{l(\bmod d_1)} \Bigg(\sum_{\substack{1 \le m < 2HN_2 \\ m \equiv l(d_1)}} \tau(m) \Bigg)^2.$$

Since $HN_2 \ll d_1^{1+\varepsilon}$ by (13.6), it follows that

$$P_1 \ll d_1^{1+\varepsilon} N_1^2. \tag{14.2}$$

The estimation of P_2 is more involved. We claim that

$$P_2 \ll d_1 x^{3/16+52\varpi+\varepsilon} K^2. \tag{14.3}$$

Write $d_1 = rq$. Note that

$$\frac{N_3'}{r} \gg x^{1/6-69\varpi} \tag{14.4}$$

by (13.8), (13.11), (13.3) and the second inequality in (13.12). Since

$$\sum_{\substack{n \simeq N_3' \\ (n,d_1)=1}} e_{d_1}\Big(b\overline{(l+kr)n}\Big) = \sum_{\substack{0 \le s < r \\ (s,r)=1}} \sum_{\substack{n \simeq N_3'/r \\ (nr+s,q)=1}} e_{d_1}\Big(b\overline{(l+kr)(nr+s)}\Big) + O(r),$$

it follows that

$$\sum_{\substack{k \sim K \\ (l+kr,d_1)=1}} \sum_{\substack{n \simeq N_3' \\ (n,d_1)=1}} e_{d_1}\Big(b\overline{(l+kr)n}\Big) = U(l) + O(Kr),$$

where

$$U(l) = \sum_{\substack{0 \le s < r \\ (s,r)=1}} \sum_{\substack{k \sim K \\ (l+kr,d_1)=1}} \sum_{\substack{n \simeq N_3'/r \\ (nr+s,q)=1}} e_{d_1}\Big(b\overline{(l+kr)(rn+s)}\Big).$$

Hence,

$$P_2 \ll \sum_{l(\bmod d_1)} |U(l)|^2 + d_1(Kr)^2. \tag{14.5}$$

The second term on the right side is admissible for (14.3) by the second inequality in (13.12). On the other hand, we have

$$\sum_{l(\bmod d_1)} |U(l)|^2 = \sum_{k_1 \sim K} \sum_{k_2 \sim K} \sum_{\substack{0 \le s_1 < r \\ (s_1,r)=1}} \sum_{\substack{0 \le s_2 < r \\ (s_2,r)=1}} V(k_2 - k_1; s_1, s_2), \tag{14.6}$$

where

$$V(k;s_1,s_2)=\sum_{\substack{n_1\simeq N_3'/r\\(n_1r+s_1,q)=1}}\sum_{\substack{n_2\simeq N_3'/r\\(n_2r+s_2,q)=1}}\times\sideset{}{'}\sum_{l(\bmod d_1)}e_{d_1}\Big(b\overline{l(n_1r+s_1)}-b\overline{(l+kr)(n_2r+s_2)}\Big).$$

Here $\sum'$ is restriction to $(l,d_1)=(l+kr,d_1)=1$.

To handle the right side of (14.6) we first note that if $l\equiv l_1r+l_2q(\bmod d_1)$, then the condition $(l(l+kr),d_1)=1$ is equivalent to $(l_1(l_1+k),q)=(l_2,r)=1$. In this situation, by the relation

$$\frac{1}{d_1}\equiv\frac{\bar r}{q}+\frac{\bar q}{r}\pmod 1,$$

we have

$$\begin{aligned}&\frac{\overline{l(n_1r+s_1)}-\overline{(l+kr)(n_2r+s_2)}}{d_1}\\&\equiv\frac{\overline{r^2l_1(n_1r+s_1)}-\overline{r^2(l_1+k)(n_2r+s_2)}}{q}+\frac{\overline{q^2s_1s_2l_2}(s_2-s_1)}{r}\pmod 1.\end{aligned}$$

Thus the innermost sum in the expression for $V(k;s_1,s_2)$ is, by the Chinese remainder theorem, equal to

$$C_r(s_2-s_1)\sum_{\substack{l(\bmod q)\\(l(l+k),q)=1}}e_q\Big(b\overline{r^2l(n_1r+s_1)}-b\overline{r^2(l+k)(n_2r+s_2)}\Big).$$

It follows that

$$V(k;s_1,s_2)=W(k;s_1,s_2)C_r(s_2-s_1),\tag{14.7}$$

where

$$W(k;s_1,s_2)=\sum_{\substack{n_1\simeq N_3'/r\\(n_1r+s_1,q)=1}}\sum_{\substack{n_2\simeq N_3'/r\\(n_2r+s_2,q)=1}}\times\sideset{}{'}\sum_{l(\bmod q)}e_q\Big(b\overline{r^2l(n_1r+s_1)}-b\overline{r^2(l+k)(n_2r+s_2)}\Big).$$

Here $\sum'$ is restriction to $(l(l+k),q)=1$.

By virtue of (14.7), we estimate the contribution from the terms with $k_1=k_2$ on the right side of (14.6) as follows. For $(n_1r+s_1,q)=(n_2r+s_2,q)=1$, we have

$$\sideset{}{^*}\sum_{l(\bmod q)}e_q\Big(b\overline{r^2l(n_1r+s_1)}-b\overline{r^2l(n_2r+s_2)}\Big)=C_q((n_1-n_2)r+s_1-s_2).$$

On the other hand, since $N_3' \ll x^{1/3}$, by (13.11) and the second inequality in (13.12) we have

$$\frac{N_3'}{d_1} \ll x^{-1/12+6\varpi} \ll r^{-1}. \tag{14.8}$$

This implies $N_3'/r = o(q)$, so that

$$\sum_{n \simeq N_3'/r} |C_q(nr+m)| \ll q^{1+\varepsilon}$$

for any m. It follows that

$$W(0; s_1, s_2) \ll q^{1+\varepsilon} r^{-1} N_3'.$$

Inserting this into (14.7) and using the simple estimate

$$\sum_{0 \le s_1 < r} \sum_{0 \le s_2 < r} |C_r(s_2 - s_1)| \ll r^{2+\varepsilon},$$

we deduce that

$$\sum_{\substack{0 \le s_1 < r \\ (s_1, r)=1}} \sum_{\substack{0 \le s_2 < r \\ (s_2, r)=1}} V(0; s_1, s_2) \ll d_1^{1+\varepsilon} N_3.$$

It follows that the contribution from the terms with $k_1 = k_2$ on the right side of (14.6) is $\ll d_1^{1+\varepsilon} K N_3$, which is admissible for (14.3), since

$$K^{-1} N_3 \ll x^{1/2+48\varpi} N_1^{-1} \ll x^{3/16+52\varpi}$$

by (13.15) and (13.2). The proof of (14.3) is therefore reduced to showing that

$$\sum_{k_1 \sim K} \sum_{\substack{k_2 \sim K \\ k_2 \ne k_1}} \sum_{\substack{0 \le s_1 < r \\ (s_1, r)=1}} \sum_{\substack{0 \le s_2 < r \\ (s_2, r)=1}} V(k_2 - k_1; s_1, s_2) \ll d_1 x^{3/16+52\varpi+\varepsilon} K^2. \tag{14.9}$$

In view of (14.4) and (14.8), letting

$$n' = \min\{n : n \simeq N_3'/r\}, \qquad n'' = \max\{n : n \simeq N_3'/r\},$$

we may rewrite $W(k; s_1, s_2)$ as

$$W(k; s_1, s_2) = \sum_{\substack{n_1 \le q \\ (n_1 r + s_1, q)=1}} \sum_{\substack{n_2 \le q \\ (n_2 r + s_2, q)=1}} \sideset{}{'}\sum_{l (\bmod q)} F(n_1/q) F(n_2/q)$$
$$\times e_q\Big(b\overline{r^2 l (n_1 r + s_1)} - b\overline{r^2 (l+k)(n_2 r + s_2)}\Big),$$

where $F(y)$ is a function of $C^2[0,1]$ class such that

$$0 \le F(y) \le 1,$$

$$F(y) = 1 \quad \text{if} \quad \frac{n'}{q} \le y \le \frac{n''}{q},$$

$$F(y) = 0 \quad \text{if} \quad y \notin \left[\frac{n'}{q} - \frac{1}{2q}, \ , \frac{n''}{q} + \frac{1}{2q}\right],$$

and such that the Fourier coefficient

$$\kappa(m) = \int_0^1 F(y)e(-my)\,dy$$

satisfies

$$\kappa(m) \ll \kappa^*(m) := \min\left\{\frac{1}{r}, \frac{1}{|m|}, \frac{q}{m^2}\right\}. \tag{14.10}$$

Here we have used (14.8). By the Fourier expansion of $F(y)$ we obtain

$$W(k;s_1;s_2) = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \kappa(m_1)\kappa(m_2)Y(k;m_1,m_2;s_1,s_2), \tag{14.11}$$

where

$$\begin{aligned} &Y(k;m_1,m_2;s_1,s_2) \\ &\quad = \sum_{\substack{n_1\le q \\ (n_1r+s_1,q)=1}} \sum_{\substack{n_2\le q \\ (n_2r+s_2,q)=1}} \sideset{}{'}\sum_{l(\bmod q)} e_q\big(\delta(l,k;m_1,m_2;n_1,n_2;s_1,s_2)\big), \end{aligned}$$

with

$$\begin{aligned} \delta(l,k;m_1,m_2;n_1,n_2;s_1,s_2) &= b\overline{r^2 l(n_1r+s_1)} \\ &\quad - b\overline{r^2(l+k)(n_2r+s_2)} + m_1n_1 + m_2n_2. \end{aligned}$$

Moreover, if $n_jr + s_j \equiv t_j (\bmod q)$, then $n_j \equiv \bar{r}(t_j - s_j)(\bmod q)$, so that

$$m_1n_1 + m_2n_2 \equiv \bar{r}(m_1t_1 + m_2t_2) - \bar{r}(m_1s_1 + m_2s_2) \pmod q.$$

Hence, on substituting $n_jr + s_j = t_j$, we may rewrite $Y(k;m_1,m_2;s_1,s_2)$ as

$$Y(k;m_1,m_2;s_1,s_2) = Z(k;m_1,m_2)e_q\big(-\bar{r}(m_1s_1 + m_2s_2)\big), \tag{14.12}$$

where

$$\begin{aligned} &Z(k;m_1,m_2) \\ &\quad = \sideset{}{^*}\sum_{t_1(\bmod q)} \sideset{}{^*}\sum_{t_2(\bmod q)} \sideset{}{'}\sum_{l(\bmod q)} e_q\big(b\overline{r^2lt_1} - b\overline{r^2(l+k)t_2} + \bar{r}(m_1t_1 + m_2t_2)\big). \end{aligned}$$

It follows from (14.7), (14.11) and (14.12) that

$$\begin{aligned} &\sum_{\substack{0\le s_1<r \\ (s_1,r)=1}} \sum_{\substack{0\le s_2<r \\ (s_2,r)=1}} V(k;s_1,s_2) \\ &\qquad = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \kappa(m_1)\kappa(m_2)Z(k;m_1,m_2)J(m_1,m_2), \end{aligned} \tag{14.13}$$

where

$$J(m_1,m_2) = \sum_{\substack{0\le s_1<r \\ (s_1,r)=1}} \sum_{\substack{0\le s_2<r \\ (s_2,r)=1}} e_q\big(-\bar{r}(m_1s_1 + m_2s_2)\big)C_r(s_2 - s_1).$$

We now appeal to Lemma 12. By simple substitution we have

$$Z(k;m_1,m_2) = T(k, bm_1\bar{r}^3, -bm_2\bar{r}^3; q),$$

so Lemma 12 gives

$$Z(k;m_1,m_2) \ll (k,q)^{1/2} q^{3/2+\varepsilon},$$

the right side being independent of m_1 and m_2. On the other hand, we have the following estimate, which will be proved later:

$$\sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \kappa^*(m_1)\kappa^*(m_2)|J(m_1,m_2)| \ll r^{1+\varepsilon}. \tag{14.14}$$

Combining these two estimates with (14.13) we obtain

$$\sum_{\substack{0\le s_1<r\\(s_1,r)=1}} \sum_{\substack{0\le s_2<r\\(s_2,r)=1}} V(k;s_1,s_2) \ll (k,q)^{1/2} q^{3/2+\varepsilon} r^{1+\varepsilon}.$$

This leads to (14.9), since

$$q^{1/2} = (d_1/r)^{1/2} < x^{1/4-21\varpi} = x^{3/16+52\varpi}$$

by the first inequality in (13.12), and

$$\sum_{k_1\sim K} \sum_{\substack{k_2\sim K\\k_2\ne k_1}} (k_2-k_1,q)^{1/2} \ll q^{\varepsilon} K^2,$$

whence (14.3) follows.

The estimate (13.16) follows from (14.1)–(14.3) immediately since

$$N_1 \le x^{3/8+8\varpi} M^{-1}, \qquad d_1 < x^{1/2+2\varpi}, \qquad \frac{31}{32} + 36\varpi = 1 - \frac{\varpi}{2}.$$

It remains to prove (14.14). The left side of (14.14) may be rewritten as

$$\frac{1}{r} \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \sum_{0\le k<r} \kappa^*(m_1)\kappa^*(m_2+k)|J(m_1,m_2+k)|.$$

In view of (14.10), we have

$$\sum_{m=-\infty}^{\infty} \kappa^*(m) \ll \mathcal{L},$$

and $\kappa^*(m+k) \ll \kappa^*(m)$ for $0 \le k < r$, since $r < q$ by (13.11) and the second inequality in (13.12). Thus, in order to prove (14.14), it suffices to show that

$$\sum_{0\le k<r} |J(m_1,m_2+k)| \ll r^{2+\varepsilon} \tag{14.15}$$

for any m_1 and m_2.

Substituting $s_2 - s_1 = t$ and applying Möbius inversion we obtain

$$J(m_1, m_2) = \sum_{|t|<r} C_r(t) \sum_{\substack{s \in I_t \\ (s(s+t), r)=1}} e_q\Big(-\bar{r}(m_2 t + (m_1 + m_2)s)\Big) \tag{14.16}$$

$$\ll \sum_{|t|<r} |C_r(t)| \sum_{r_1 | r} \Bigg| \sum_{\substack{s \in I_t \\ s(s+t) \equiv 0 (r_1)}} e_q\Big(\bar{r}(m_1 + m_2)s\Big)\Bigg|,$$

where I_t is a certain interval of length $< r$ and depending on t. For any t and square-free r_1, there are exactly $\tau(r_1/(t, r_1))$ distinct residue classes $(\bmod\, r_1)$ such that

$$s(s+t) \equiv 0 \pmod{r_1}$$

if and only if s lies in one of these classes. On the other hand, if $r = r_1 r_2$, then

$$\sum_{\substack{s \in I_t \\ s \equiv a(r_1)}} e_q\Big(\bar{r}(m_1 + m_2)s\Big) \ll \min\Big\{r_2,\, ||\bar{r}_2(m_1 + m_2)/q||^{-1}\Big\}$$

for any a. Hence the inner sum on the right side of (14.16) is

$$\ll \tau(r) \sum_{r_2 | r} \min\Big\{r_2,\, ||\bar{r}_2(m_1 + m_2)q||^{-1}\Big\},$$

which is independent of t. Together with the simple estimate

$$\sum_{|t|<r} |C_r(t)| \ll \tau(r) r,$$

this yields

$$J(m_1, m_2) \ll \tau(r)^2 r \sum_{r_2 | r} \min\Big\{r_2,\, ||\bar{r}_2(m_1 + m_2)/q||^{-1}\Big\}.$$

It follows that the left side of (14.15) is

$$\ll \tau(r)^2 r \sum_{r_1 r_2 = r} \sum_{0 \le k_1 < r_1} \sum_{0 \le k_2 < r_2} \min\Big\{r_2,\, ||\bar{r}_2(m_1 + m_2 + k_1 r_2 + k_2)/q||^{-1}\Big\}. \tag{14.17}$$

Assume $r_2 | r$. By the relation

$$\frac{\bar{r}_2}{q} \equiv -\frac{\bar{q}}{r_2} + \frac{1}{q r_2} \pmod 1,$$

for $0 \le k < r_2$ we have

$$\frac{\bar{r}_2(m+k)}{q} \equiv \frac{\bar{r}_2 m}{q} - \frac{\bar{q}k}{r_2} + O\bigg(\frac{1}{q}\bigg) \pmod 1.$$

This yields

$$\sum_{0 \le k < r_2} \min\Big\{r_2,\, ||\bar{r}_2(m+k)/q||^{-1}\Big\} \ll r_2 \mathcal{L} \tag{14.18}$$

for any m. The estimate (14.15) follows from (14.17) and (14.18) immediately.

Acknowledgments. The author is grateful to the referees for carefully reading the manuscript and making useful suggestions.

References

[1] E. Bombieri, J. B. Friedlander, and H. Iwaniec, Primes in arithmetic progressions to large moduli, *Acta Math.* **156** (1986), 203–251. MR 0834613. Zbl 0588.10042. http://dx.doi.org/10.1007/BF02399204.

[2] E. Bombieri, J. B. Friedlander, and H. Iwaniec, Primes in arithmetic progressions to large moduli. II, *Math. Ann.* **277** (1987), 361–393. MR 0891581. Zbl 0625.10036. http://dx.doi.org/10.1007/BF01458321.

[3] E. Bombieri, J. B. Friedlander, and H. Iwaniec, Primes in arithmetic progressions to large moduli. III, *J. Amer. Math. Soc.* **2** (1989), 215–224. MR 0976723. Zbl 0674.10036. http://dx.doi.org/10.2307/1990976.

[4] P. Deligne, La conjecture de Weil. I, *Inst. Hautes Études Sci. Publ. Math.* **43** (1974), 273–307. MR 0340258. Zbl 0287.14001. http://dx.doi.org/10.1007/BF02684373.

[5] E. Fouvry and H. Iwaniec, Primes in arithmetic progressions, *Acta Arith.* **42** (1983), 197–218. MR 0719249. Zbl 0517.10045. Available at http://pldml.icm.edu.pl/pldml/element/bwmeta1.element.bwnjournal-article-aav42i2p197bwm?q=bwmeta1.element.bwnjournal-number-aa-1982-1983-42-2;5&qt=CHILDREN-STATELESS.

[6] J. B. Friedlander and H. Iwaniec, Incomplete Kloosterman sums and a divisor problem, *Ann. of Math.* **121** (1985), 319–350. MR 0786351. Zbl 0572.10029. http://dx.doi.org/10.2307/1971175.

[7] D. A. Goldston, J. Pintz, and C. Y. Yildirim, Primes in tuples. I, *Ann. of Math.* **170** (2009), 819–862. MR 2552109. Zbl 1207.11096. http://dx.doi.org/10.4007/annals.2009.170.819.

[8] D. A. Goldston, J. Pintz, and C. Y. Yildirim, Primes in tuples. II, *Acta Math.* **204** (2010), 1–47. MR 2600432. Zbl 1207.11097. http://dx.doi.org/10.1007/s11511-010-0044-9.

[9] D. R. Heath-Brown, Prime numbers in short intervals and a generalized Vaughan identity, *Canad. J. Math.* **34** (1982), 1365–1377. MR 0678676. Zbl 0478.10024. http://dx.doi.org/10.4153/CJM-1982-095-9.

[10] D. R. Heath-Brown, The divisor function $d_3(n)$ in arithmetic progressions, *Acta Arith.* **47** (1986), 29–56. MR 0866901. Zbl 0549.10034. Available at http://pldml.icm.edu.pl/pldml/element/bwmeta1.element.bwnjournal-article-aav47i1p29bwm?q=bwmeta1.element.bwnjournal-number-aa-1986-47-1;2&qt=CHILDREN-STATELESS.

[11] H. Iwaniec, A new form of the error term in the linear sieve, *Acta Arith.* **37** (1980), 307–320. MR 0598883. Zbl 0444.10038. Available at http://pldml.icm.edu.pl/pldml/element/bwmeta1.element.bwnjournal-article-aav37i1p307bwm?q=bwmeta1.element.bwnjournal-number-aa-1980-37-1;26&qt=CHILDREN-STATELESS.

[12] P. Shiu, A Brun-Titchmarsh theorem for multiplicative functions, *J. Reine Angew. Math.* **313** (1980), 161–170. MR 0552470. Zbl 0412.10030. http://dx.doi.org/10.1515/crll.1980.313.161.

[13] K. Soundararajan, Small gaps between prime numbers: the work of Goldston-Pintz-Yıldırım, *Bull. Amer. Math. Soc.* **44** (2007), 1–18. MR 2265008. Zbl 1193.11086. http://dx.doi.org/10.1090/S0273-0979-06-01142-6.

(Received: April 17, 2013)
(Revised: May 16, 2013)

University of New Hampshire, Durham, NH
E-mail: yitangz@unh.edu