ELEMENTARY AND ANALYTIC THEORY OF NUMBERS
BANACH CENTER PUBLICATIONS, VOLUME 17
PWN POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

A WEIGHTED SIEVE OF GREAVES' TYPE II

H. HALBERSTAM

Department of Mathematics, University of Illinois
Urbana Illinois, U.S.A.

H.-E. RICHERT

Department of Mathematics, University of Ulm, Ulm, F.R.G.

1. Introduction

In this paper, the sequel to part I [8], we continue our study of the weighted sifting function

$$H(\mathscr{A}, z_1, z_2) = H(\mathscr{A}, \mathscr{P}, z_1, z_2) = \sum_{a \in \mathscr{A}} \gamma\big((a, P(z_2))\big)$$

in the context of sieve problems of dimension 1, characterized by

$$\varkappa = 1, \quad \beta = \beta(\varkappa) = 2. \tag{1.1}$$

We maintain the notation of I. (When we refer to a formula from I we shall prefix its number by I.) Thus

$$\gamma(n) = \{1 - \sum_{\substack{p \mid n \\ p \in \mathscr{P}}} (1 - w(p))\}^+, \quad \{x\}^+ = \max(x, 0), \tag{1.2}$$

where, for $p \mid P(z_2) = \prod_{p < z_2, p \in \mathscr{P}} p$,

$$w(p) = \begin{cases} \dfrac{1}{T-E}\left(\dfrac{\log p}{\log y} - E\right), & y^{1/4} \leqslant p < y^U, \\ \dfrac{1}{T-E}\left(\dfrac{\log p}{\log y} - E_0\right), & y^V \leqslant p < y^{1/4}, \\ 0, & p < y^V; \end{cases} \tag{1.3}$$

recall (I, Section 5) that y is as before, our basic parameter (controlling the size of the remainder term) – we require y to be large and indicate this by

$$y \geqslant y_0 - \tag{1.4}$$

that

$$z_1 = y^V, \quad z_2 = y^U \tag{1.5}$$

and that the four constants T, U, V, E together with

$$E_0 = \max\left(E, \tfrac{1}{3}(1-T)\right)$$

satisfy the conditions

$$E_0 \leqslant V \leqslant 1/4, \quad 1/2 \leqslant U \leqslant T < 1, \tag{1.6}$$

and

$$U + 3V \geqslant 1, \quad V \geqslant V_1 > 0. \tag{1.7}$$

Remember that, always,

$$0 \leqslant w(p) \leqslant \frac{1}{T-E}\left(\frac{\log p}{\log y} - E\right) < 1, \quad p < y^U. \tag{1.8}$$

As usual, we have, whenever $d \mid P(z_2)$,

$$|\mathscr{A}_d| = |\{a \in \mathscr{A}: \ a \equiv 0 \bmod d\}| = \frac{\omega(d)}{d} X + R_d,$$

where $\omega(\cdot)$ is a non-negative multiplicative arithmetic function such that $\omega(p) = 0$ when $p \notin \mathscr{P}$,

$$0 \leqslant \omega(p) < p, \quad p \in \mathscr{P}, \tag{A_0}$$

and

$$\left| \sum_{z_1 \leqslant p < z} \frac{\omega(p)}{p} \log p - \log \frac{z}{z_1} \right| \leqslant A, \quad 2 \leqslant z_1 \leqslant z; \tag{$\Omega(1)$}$$

in what follows, O- and $\ll$-constants will depend at most on A (and on U).

Our objective in this paper is to show that the remainder term in the weighted sieve of Greaves' type may be given the structure of an Iwaniec bilinear form (see Theorem A in Section 7 below). Then, in Section 9 we indicate some applications.

Our method derives from Motohashi's elegant version [12] of the original Iwaniec treatment [10] of the remainder term in the unweighted linear sieve.

2. The dissection

Let

$$\delta = \delta(z_2) = \frac{\log\log\log z_2}{\log\log z_2} \tag{2.1}$$

and introduce

$$z_0 = z_2^{(1+\delta)^{-J}} \tag{2.2}$$

where

$$J = \left[\frac{\log\log\log z_2}{5\log(1+\delta)}\right]. \tag{2.3}$$

Then

$$(\log\log z_2)^{-1/5}\log z_2 \leqslant \log z_0 \leqslant (1+\delta)(\log\log z_2)^{-1/5}\log z_2; \tag{2.4}$$

since $z_2 = y^U \geqslant y_0^{1/2}$ is large, we may assume that

$$z_0 \leqslant y^V$$

and hence that, by (1.3),

$$w(p) = 0 \quad \text{for} \quad p \mid P(z_0). \tag{2.5}$$

We subdivide the interval $[z_0, z_2)$ into disjoint intervals

$$I_\nu = [z_2^{(1+\delta)^{-\nu}}, z_2^{(1+\delta)^{-\nu+1}}), \quad 1 \leqslant \nu \leqslant J. \tag{2.6}$$

Thus, if $1 \leqslant \nu_1 < \nu_2 \leqslant J$, I_{ν_1} is to the right of I_{ν_2} and we indicate this by writing $I_{\nu_1} > I_{\nu_2}$ or $I_{\nu_2} < I_{\nu_1}$.

Let

$$d = p_1 \dots p_r \quad (p_1 > \dots > p_r), \quad d \mid P(z_0, z_2).$$

Particular importance attaches to those divisors d of $P(z_0, z_2)$ that are *well-separated* by our dissection in the sense that the prime factors of d lie in *distinct* sub-intervals I_ν, that is,

$$p_j \in I_{\nu_j} \quad (1 \leqslant j \leqslant r) \quad \text{with} \quad \nu_1 < \dots < \nu_r.$$

Accordingly we form the direct product

$$D = \bigcup_{j=1}^{r} I_{\nu_j}$$

and write

$$d \in D.$$

Let $\mathscr{D} = \{D\}$ denote the set of all such 'square-free' direct products, including the "empty" singleton set consisting of 1 only, to be denoted by $\emptyset$. Given a particular direct product D, all integers $d \in D$ have the same number of prime factors, and we may speak unambiguously of this common number as $\nu(D)$. It is quite natural to introduce also the 'Möbius function' on $\mathscr{D}$ by

$$\mu(D) = \begin{cases} 1, & D = \emptyset, \\ (-1)^r, & D \doteq \bigcup_{j=1}^{r} I_{\nu_j} \quad (\nu_1 < \ldots < \nu_r); \end{cases}$$

and when it is convenient to speak of 'non-squarefree' $D = \bigcup_{j=1}^{r} I_{\nu_j}$ ($I_{\nu_l} = I_{\nu_k}$ for some pair k, l with $k \neq l$) then, of course, we put $\mu(D) = 0$. One may observe that $\hat{\mathscr{D}}$, the set of all direct products D, squarefree and non-squarefree, is a partially ordered set, even a lattice, with respect to inclusion. We record here also the obvious fact (cf. (2.6))

$$\text{card } \mathscr{D} = 2^J. \tag{2.7}$$

For any interval I from (2.6), let i denote the right-hand end point of I. To simplify the notation, write a typical member D of $\mathscr{D}$ in the form

$$D = I_1 \ldots I_r \quad (I_1 > \ldots > I_r) \tag{2.8}$$

(so that here $I_1, \ldots, I_r$ are not necessarily the first r intervals (2.6) counting from the right), and introduce the notations

$$\Delta = \Delta(D) = i_1 i_2 \ldots i_r,$$

$$I < D \quad \text{if and only if} \quad I < I_r,$$

and

$$D < I \quad \text{if and only if} \quad I_1 < I.$$

Clearly, equivalent ways of writing $I < D$, $D < I$ are, respectively, $i < i_r$ and $i_1 < i$.

By analogy with (I.3.5) and (I.3.6) we now introduce on $\mathscr{D}$ the Buchstab–Rosser–Iwaniec functions $\chi_x^{\pm}(D)$, where

$$x = yz_0^{-L}, \quad L = \log \log \log z_2 \tag{2.9}$$

(for technical reasons we have to use x in place of the expected y) as follows: Let $\chi_x^{\pm}(\emptyset) = 1$ and, if $D \neq \emptyset$ is as in (2.8), define

$$\chi_x^{+}(D) = 1 \quad \text{if} \quad i_{2s-1}^3 i_{2s-2} \ldots i_1 < x \quad (1 \leqslant s \leqslant \tfrac{1}{2}(r+1)), \tag{2.10}$$

$$\chi_x^{-}(D) = 1 \quad \text{if} \quad i_{2s}^3 i_{2s-1} \ldots i_1 < x \quad (1 \leqslant s \leqslant \tfrac{1}{2}r), \tag{2.11}$$

and

$$\chi_x^{\pm}(D) = 0$$

otherwise.

Furthermore, let $\bar{\chi}_x^{\pm}(\emptyset) = 0$ and, if $D \neq \emptyset$ is as in (2.8), define

$$\bar{\chi}_x^{\pm}(D) = \chi_x^{\pm}(I_1 \ldots I_{r-1}) - \chi_x^{\pm}(I_1 \ldots I_r).$$

Then

$$\bar{\chi}_x^{-}(D) = 0 \quad \text{if} \quad 2 \nmid \nu(D); \qquad \bar{\chi}_x^{+}(D) = 0 \quad \text{if} \quad 2 \mid \nu(D), \tag{2.12}$$

so that

$$\mu(D)\bar{\chi}_x^{+}(D) \leqslant 0 \leqslant \mu(D)\bar{\chi}_x^{-}(D).$$

We record the following simple consequences of these definitions: first, since $\chi_x^{\pm}(D)$ is defined in terms of the right-hand endpoints of the intervals making up D, we see that

$$\chi_x^{\pm}(D) = 1 \quad \text{implies that } \chi_x^{\pm}(d) = 1 \text{ whenever } d \in D. \tag{2.13}$$

Next,

$$\bar{\chi}_x^{-}(D) = 1 \quad \text{if} \quad r = \nu(D) \text{ is positive and even and}$$
$$i_{2s}^3 i_{2s-1} \ldots i_1 < x \ (s = 1, \ldots, \tfrac{1}{2}r - 1), \ i_r^3 i_{r-1} \ldots i_1 \geqslant x. \tag{2.14}$$

Finally,

$$\chi_x^{-}(D) \neq 0, \ \nu(D) \geqslant 2 \quad \text{together imply that} \quad i_r \Delta(D) < x. \tag{2.15}$$

With these definitions and basic properties we are in a position to prove the following combinatorial identity.

LEMMA 1. *For any arithmetic function $\psi(\cdot)$ we have*

$$\sum_{d \mid P(z_0, z_2)} \mu(d)\psi(d) = \sum_{D \in \mathcal{D}} \mu(D)\chi_x^{-}(D) \sum_{d \in D} \psi(d) +$$
$$+ \sum_{I} \sum_{\substack{D \in \mathcal{D} \\ I < D}} \mu(D)\chi_x^{-}(ID) \sum_{\substack{p', p \in I \\ p' < p}} \sum_{d \in D} \sum_{t \mid P(z_0, p')} \mu(t)\psi(p' p d t) +$$
$$+ \sum_{D \in \mathcal{D}} \bar{\chi}_x^{-}(D) \sum_{d \in D} \sum_{t \mid P(z_0, p(d))} \mu(t)\psi(dt).$$

Note. We are using here the notation $p(d)$ for the least prime factor of d, and below we shall use also, when $\nu(d) \geqslant 2$, $p_2(d)$ for the next to least prime factor of d.

Proof. We employ the Fundamental Identity from I, Lemma 2. There let $\varphi(d) = 0$ if $(d, P(z_0)) > 1$ and $\varphi(d) = \psi(d)$ otherwise. Next, choose χ in

Lemma 2 of I in accordance with

$$\chi(d) = \begin{cases} 0, & d \in D \text{ and } \mu(D) = 0, \\ \chi_x^-(D), & d \in D \text{ and } \mu(D) \neq 0, \end{cases} \tag{2.16}$$

so that $\chi(d)$ is 0 whenever d is not well-separated by our dissection. Clearly the first sum on the right of (I.3.2) gives exactly the first sum on the right of the lemma.

In the second sum on the right of (I.3.2) our concern is with $\bar{\chi}(d)$, and only integers $d > 1$ need to be considered. Suppose the typical $d \in D$ with $d/p(d) \in D_1$. If $d/p(d)$ is not well separated, neither is d and so $\bar{\chi}(d) = 0$. Suppose then that $d/p(d)$ is well separated. Here $\bar{\chi}(d) = \bar{\chi}_x^-(D)$ if $\mu(D) \neq 0$ (i.e. d is also well separated) and $\bar{\chi}(d) = \chi_x^-(D_1)$ if $\mu(D) = 0$ (i.e. d is not well separated). The second sum on the right of (I.3.2) splits into two sums according to these two possibilities: the third sum on the right of the lemma derives from precisely those $d > 1$ with both $d/p(d)$ and d well separated – note that $\mu(D) = 1$ when $\bar{\chi}_x^-(D) \neq 0$ by (2.12) – and the second sum from those with $d/p(d)$ but not d well separated. These latter divisors d necessarily have at least two prime factors and the smallest two prime factors of d lie in the same interval I of our dissection. This completes the proof of the lemma.

3. Preparation of the sifting function $H(\mathscr{A}, z_1, z_2)$ – a combinatorial inequality

Recall from (I.3.20) that

$$H_q(\mathscr{A}, z_1, z) = \sum_{d \mid P(z)} \mu(d) W(qd) |\mathscr{A}_{qd}|, \quad (q, P(z)) = 1, \tag{3.1}$$

where (cf. (I.3.17))

$$W(d) = \sum_{t \mid d} \mu(t) \gamma(t), \quad d \mid P(z). \tag{3.2}$$

The corresponding formula for the classical sifting function S is

$$S(\mathscr{A}_q, z) = \left|\{a \in \mathscr{A}_q : (a, P(z)) = 1\}\right| = \sum_{d \mid P(z)} \mu(d) |\mathscr{A}_{qd}|, \quad (q, P(z)) = 1. \tag{3.3}$$

Let z' be any number satisfying $2 \leqslant z' \leqslant z$, and, on the right of (3.1), put $d = d_1 d_2$ where $d_1 \mid P(z', z)$ and $d_2 \mid P(z')$. Then

$$\begin{aligned} H_q(\mathscr{A}, z_1, z) &= \sum_{d_1 \mid P(z', z)} \mu(d_1) \sum_{d_2 \mid P(z')} \mu(d_2) W(qd_1 d_2) |\mathscr{A}_{qd_1 d_2}| \\ &= \sum_{d \mid P(z', z)} \mu(d) . H_{qd}(\mathscr{A}, z_1, z'), \end{aligned} \tag{3.4}$$

$$2 \leqslant z' \leqslant z, \ (q, P(z)) = 1.$$

By (1.2) and (2.5) we see that $\gamma(n) = 0$ if $(n, P(z_0)) > 1$, so that, by (3.2),

$$W(d) = W\big((d, P(z_0, z))\big), \tag{3.5}$$

and hence, by (3.1) and (3.3),

$$H_q(\mathscr{A}, z_1, z_0) = W(q) S(\mathscr{A}_q, z_0), \quad (q, P(z_0)) = 1.$$

Therefore (3.4) with $z' = z_0$ gives

$$(3.6) \qquad H_q(\mathscr{A}, z_1, z) = \sum_{d \mid P(z_0, z)} \mu(d) W(qd) S(\mathscr{A}_{qd}, z_0), \quad (q, P(z)) = 1,$$

and, in particular,

$$(3.7) \qquad H(\mathscr{A}, z_1, z_2) = \sum_{d \mid P(z_0, z_2)} \mu(d) W(d) S(\mathscr{A}_d, z_0).$$

We now apply Lemma 1 with $\psi(d) = W(d) S(\mathscr{A}_d, z_0)$ in (3.7), making use also of (3.6). We obtain

$$\begin{aligned} H(\mathscr{A}, z_1, z_2) = &\sum_{D \in \mathscr{S}} \mu(D) \chi_x^-(D) \sum_{d \in D} W(d) S(\mathscr{A}_d, z_0) + \\ &+ \sum_I \sum_{\substack{D \in \mathscr{D} \\ I < D}} \mu(D) \chi_x^-(ID) \sum_{\substack{p', p \in I \\ p' < p}} \sum_{d \in D} H_{dp'p}(\mathscr{A}, z_1, p') + \\ &+ \sum_{D \in \mathscr{S}} \bar{\chi}_x^-(D) \sum_{d \in D} H_d(\mathscr{A}, z_1, p(d)). \end{aligned}$$

In the third sum on the right we see from rewriting (3.1) that

$$H_d(\mathscr{A}, z_1, p(d)) = \sum_{a \in \mathscr{A}_d} \sum_{\substack{t \mid a \\ t \mid P(p(d))}} \mu(t) W(dt) = \sum_{a \in \mathscr{A}_d} \gamma_d\big((a, P(p(d)))\big)$$

in the notation of (I.3.19) and (I.3.21); since any D counted in that sum has $\nu(D) \geqslant 2$ and even, Lemma 7 of I tells us that $H_d(\mathscr{A}, z_1, p(d)) \geqslant 0$ whenever $d \in D$ and $\bar{\chi}_x^-(D) = 1$. Since our aim is a lower bound for $H(\mathscr{A}, z_1, z_2)$ we may now, and we shall, drop the third sum.

In the second sum we follow an idea of Motohashi [12]: $\chi_x^-(ID) \neq 0$ implies, by (2.11), condition (I.5.10) (obviously $\nu(dpp') > 2$ when $\nu(d) \geqslant 1$) so that the terms $H_{dp'p}(\mathscr{A}, z_1, p')$ are non-negative by Lemma 7 of I, as before. Hence those terms in the second sum with $\mu(D) = 1$ may also be dropped. For the remaining terms, with $\mu(D) = -1$, we have $\nu(dpp') > 2$ and here we use (I.5.11) of Lemma 7 of I, together with (1.8) to conclude that

$$H_{dp'p}(\mathscr{A}, z_1, p') \leqslant S(\mathscr{A}_{dp'p}, z_0).$$

Combining all these remarks we arrive at the lower estimate

$$(3.8) \qquad \begin{aligned} H(\mathscr{A}, z_1, z_2) \geqslant &\sum_{D \in \mathscr{D}} \mu(D) \chi_x^-(D) \sum_{d \in D} W(d) S(\mathscr{A}_d, z_0) - \\ &- \sum_I \sum_{\substack{D \in \mathscr{D} \\ I < D \\ 2 \nmid \nu(D)}} \chi_x^-(ID) \sum_{\substack{p', p \in I \\ p' < p}} \sum_{d \in D} S(\mathscr{A}_{dp'p}, z_0). \end{aligned}$$

This is the combinatorial inequality referred to in the section heading. The dominant term in the eventual lower bound for $H(\mathscr{A}, z_1, z_2)$ (see Theorem A below) derives entirely from the first sum on the right of (3.8). The second sum on the right will be absorbed into the remainder term, thanks to the summation over primes in I.

4. Analysis of the inequality for $H(\mathscr{A}, z_1, z_2)$

We replace the functions $S(\mathscr{A}_q, z_0)$ on the right of (3.8) by means of the following version of a Fundamental Lemma:

LEMMA 2 (Friedlander–Iwaniec [1]). *Let*

$$L \geqslant 2, \quad z_0 \geqslant 2, \quad \mu(q) \neq 0, \quad (q, P(z_0)) = 1.$$

Then there exist functions $\beta^{\pm}(\cdot)$ *satisfying*

$$\beta^{\pm}(m) = 0 \text{ or } 1 \tag{4.1}$$

such that

$$\sum_{\substack{m < z_0^L \\ m | P(z_0)}} \mu(m)\beta^{-}(m) R_{qm} \leqslant S(\mathscr{A}_q, z_0) - X\frac{\omega(q)}{q} V(z_0)\{1 + O(e^{-L})\}$$

$$\leqslant \sum_{\substack{m < z_0^L \\ m | P(z_0)}} \mu(m)\beta^{+}(m) R_{qm},$$

and the O-constant depends at most on the constant A from $(\Omega(1))$.

We take z_0 and L in Lemma 2 to be as given by (2.2) and (2.9) respectively, and apply the Lemma in (3.8). By (2.13), $\chi_x^-(D) = 1$ implies that $\chi_x^-(d) = 1$ when $d \in D$, and since $x < y$ this means that also $\chi_y^-(d) = 1$ when $d \in D$; also, as is shown in Lemmas 7 and 8 of I,

$$d | P(z_2) \text{ and } \chi_y^-(d) = 1 \text{ imply that } W(d) = W_0(d) \geqslant 0, \tag{4.2}$$

so that in these circumstances

$$1 \geqslant W(d) = W_0(d) = 1 - \sum_{p|d} w(p) \geqslant 0, \tag{4.3}$$

and therefore

(4.4) $H(\mathscr{A}, z_1, z_2)$

$$\geqslant XV(z_0) \sum_{D \in \mathscr{D}} \mu(D)\chi_x^-(D) \sum_{d \in D} W_0(d)\frac{\omega(d)}{d} + O\left(XV(z_0)e^{-L} \sum_{d | P(z_0, z_2)} \frac{\omega(d)}{d}\right) +$$

$$+ \sum_{D \in \mathscr{D}} \mu(D)\chi_x^-(D) \sum_{d \in D} W_0(d) \sum_{\substack{m < z_0^L \\ m | P(z_0)}} \mu(m)\beta^{(-)^{\nu(D)+1}}(m) R_{dm} +$$

$$+O\left(XV(z_0)\left(\max_I \sum_{p\in I}\frac{\omega(p)}{p}\right)\sum_{z_0\leqslant p<z_2}\frac{\omega(p)}{p}\sum_{d|P(z_0,z_2)}\frac{\omega(d)}{d}\right)-$$

$$-\sum_I \sum_{\substack{D\in\mathscr{D}\\ I<D\\ 2\nmid\nu(D)}} \chi_x^-(ID)\sum_{\substack{p',p\in I\\ p'<p}}\sum_{d\in D}\sum_{\substack{m<z_0^L\\ m|P(z_0)}}\mu(m)\beta^+(m)R_{p'pdm}.$$

Let

$$G^* = \sum_{D\in\mathscr{D}}\mu(D)\chi_x^-(D)\sum_{d\in D}W_0(d)\frac{\omega(d)}{d} \tag{4.5}$$

and

$$R = \sum_{D\in\mathscr{D}}\mu(D)\chi_x^-(D)\sum_{d\in D}W_0(d)\sum_{\substack{m<z_0^L\\ m|P(z_0)}}\mu(m)\beta^{(-)^{\nu(D)+1}}(m)R_{dm}- \tag{4.6}$$

$$-\sum_I \sum_{\substack{D\in\mathscr{D}\\ I<D\\ 2\nmid\nu(D)}} \chi_x^-(ID)\sum_{\substack{p',p\in I\\ p'<p}}\sum_{d\in D}\sum_{\substack{m<z_0^L\\ m|P(z_0)}}\mu(m)\beta^+(m)R_{p'pdm}$$

$$=\Sigma_1-\Sigma_2,$$

say. We shall devote most of the rest of this section to the 'leading' sum G^*. The O-terms are easy to estimate, and the remainder sum R will be investigated in the next section.

We apply Lemma 1, this time with $\psi(d)=\chi_x^-(d)W_0(d)\dfrac{\omega(d)}{d}$ to G^*. By virtue of (2.13) we obtain

$$G^* = \sum_{d|P(z_0,z_2)}\mu(d)\chi_x^-(d)W_0(d)\frac{\omega(d)}{d}- \tag{4.7}$$

$$-\sum_I\sum_{\substack{D\in\mathscr{D}\\ I<D}}\mu(D)\chi_x^-(ID)\times$$

$$\times\sum_{\substack{p',p\in I\\ p'<p}}\sum_{d\in D}\sum_{t|P(z_0,p')}\mu(t)\chi_x^-(p'pdt)W_0(p'pdt)\frac{\omega(p'pdt)}{p'pdt}-$$

$$-\sum_{D\in\mathscr{D}}\bar{\chi}_x^-(D)\sum_{d\in D}\sum_{t|P(z_0,p(d))}\mu(t)\chi_x^-(dt)W_0(dt)\frac{\omega(dt)}{dt}.$$

By (4.2) and (4.3) the second sum on the right of (4.7) is, in absolute value, at most

$$\left(\max_I\sum_{p\in I}\frac{\omega(p)}{p}\right)\sum_{z_0\leqslant p<z_2}\frac{\omega(p)}{p}\sum_{\substack{d|P(z_0,z_2)\\ p(d)>p}}\frac{\omega(d)}{d}\sum_{t|P(z_0,p)}\frac{\omega(t)}{t}$$

$$\leqslant\left(\max_I\sum_{p\in I}\frac{\omega(p)}{p}\right)\left(\sum_{z_0\leqslant p<z_2}\frac{\omega(p)}{p}\right)\sum_{d|P(z_0,z_2)}\frac{\omega(d)}{d}.$$

Hence this portion of G^*, together with the factor $XV(z_0)$, can be absorbed into the second O-term on the right of (4.4). In the third sum on the right of (4.7), (4.2) and (4.3) together again imply that $0 \leqslant W_0(dt) \leqslant 1$ whenever $\chi_x^-(dt) \neq 0$. Also, by (I.4.8) (with $\nu = 1$),

$$\chi_x^-(dt) = \chi_x^-(d)\chi_{x/d}^{(-)^{\nu(d)+1}}(t)$$

so that the last sum on the right of (4.7) is, in absolute value, at most

$$\sum_{\substack{D \in \mathscr{D} \\ 2|\nu(D)}} \bar{\chi}_x^-(D) \sum_{d \in D} \chi_x^-(d) \frac{\omega(d)}{d} \sum_{t|P(z_0,z_2)} \frac{\omega(t)}{t}. \tag{4.8}$$

Here let

$$D = I_1 \dots I_r, \qquad I_1 > \dots > I_r, \qquad r \text{ even},$$

so that any d in D has the form $d = p_1 \dots p_r$ $(p_j \in I_j, j = 1, \dots, r)$. Then, by (I.3.6) (with $\nu = r$ and $\beta = 2$) and (2.14),

$$x > p_1 \dots p_r^3 \geqslant (i_1 \dots i_r^3)^{1/(1+\delta)} \geqslant x^{1/(1+\delta)};$$

writing $d = d_1 p$ with $p(d_1) > p$ (so that, in fact, $d_1 = p_1 \dots p_{r-1}$, $p = p_r$), we have

$$x^{1/(1+\delta)}/d_1 \leqslant p^3 < x/d_1,$$

and in any case $z_0 \leqslant p < z_2$, of course. Hence the sum (4.8) is at most

$$\left(\sum_{t|P(z_0,z_2)} \frac{\omega(t)}{t}\right)\left(\sum_{d_1|P(z_0,z_2)} \frac{\omega(d_1)}{d_1} \sum_{u_1 \leqslant p < u_2} \frac{\omega(p)}{p}\right)$$

where

$$u_1 = u_1(d_1) = \max\left(z_0, (x^{1/(1+\delta)}/d_1)^{1/3}\right), \qquad u_2 = u_2(d_1) = \min\left(z_2, (x/d_2)^{1/3}\right).$$

By (I.2.4), (2.9) and (1.6), the innermost sum is no larger than

$$\log \frac{\log u_2}{\log u_1} + \frac{A}{\log u_1} \leqslant \frac{\delta}{3} \frac{\log x}{\log z_0} + \frac{A}{\log z_0} \leqslant \delta \frac{\log z_2}{\log z_0},$$

so that the sum (4.8) is at most

$$\delta \frac{\log z_2}{\log z_0}\left(\sum_{t|P(z_0,z_2)} \frac{\omega(t)}{t}\right)^2.$$

We feed all this information back into (4.7) to obtain

$$G^* \geqslant \sum_{d|P(z_0,z_2)} \mu(d)\chi_x^-(d) W_0(d) \frac{\omega(d)}{d} -$$

$$-\left(\max_I \sum_{p \in I} \frac{\omega(p)}{p}\right)\left(\sum_{z_0 \leqslant p < z_2} \frac{\omega(p)}{p}\right) \sum_{d|P(z_0,z_2)} \frac{\omega(d)}{d} - \delta \frac{\log z_2}{\log z_0}\left(\sum_{d|P(z_0,z_2)} \frac{\omega(d)}{d}\right)^2;$$

and this, in combination with (4.4), (4.5) and (4.6) gives

$$(4.9)\qquad H(\mathscr{A}, z_1, z_2) \geqslant XV(z_0) \sum_{d|P(z_0,z_2)} \mu(d)\chi_x^-(d) W_0(d) \frac{\omega(d)}{d} + R +$$
$$+O\Bigg(XV(z_0) \sum_{d|P(z_0,z_2)}' \frac{\omega(d)}{d} \times$$
$$\times \Bigg\{ e^{-L} + \Big(\max_I \sum_{p\in I} \frac{\omega(p)}{p}\Big) \sum_{z_0\leqslant p<z_2} \frac{\omega(p)}{p} +$$
$$+\delta \frac{\log z_2}{\log z_0} \sum_{d|P(z_0,z_2)} \frac{\omega(d)}{d} \Bigg\}\Bigg).$$

By (I.2.6) (with $\varkappa = 1$) we have

$$V(z_0) \sum_{d|P(z_0,z_2)} \frac{\omega(d)}{d} \leqslant \left(\frac{V(z_0)}{V(z_2)}\right)^2 V(z_2) \ll \left(\frac{\log z_2}{\log z_0}\right)^2 V(z_2).$$

By (I.2.4) (also with $\varkappa = 1$) together with (2.6), (2.1), (2.2) and (2.4),

$$\max_I \sum_{p\in I} \frac{\omega(p)}{p} \leqslant \log(1+\delta) + \frac{A}{\log z_0} \ll \delta$$

and

$$\sum_{z_0\leqslant p<z_2} \frac{\omega(p)}{p} \ll \log\log\log z_2.$$

Hence the error term in (4.9) is at most of order

$$XV(z_2)\left(\frac{\log z_2}{\log z_0}\right)^2 \left\{\frac{1}{\log\log z_2} + \delta \log\log\log z_2 + \delta \left(\frac{\log z_2}{\log z_0}\right)^2\right\}$$
$$\ll XV(z_2) Le^{-L/5}$$

by (2.4), so that (4.9) yields

$$(4.10)\qquad H(\mathscr{A}, z_1, z_2)$$
$$\geqslant XV(z_0) \sum_{d|P(z_0,z_2)} \mu(d)\chi_x^-(d) W_0(d) \frac{\omega(d)}{d} + R + O\big(XV(z_2) Le^{-L/5}\big)$$

where L is given in (2.9). We simplify (4.10) in one further respect, by means of the following general observation.

LEMMA 3. *Suppose that $z_2 \geqslant z \geqslant z_0$. Let $\psi(\cdot)$ be any arithmetic function satisfying*

$$(4.11)\qquad \psi(d) = \psi\big((d, P(z_0, z))\big)$$

and

(4.12) $\qquad \psi(d) \geqslant 0 \quad$ *when* $\quad \bar{\chi}_x^{(-)^\nu}(d) \neq 0$ *and* $(d, P(z_0)) = 1$.

Then for $\nu = 0$ *and for* $\nu = 1$

$$(-1)^{\nu+1} V(z_0) \sum_{d|P(z_0,z)} \mu(d)\chi_x^{(-)^\nu}(d) \frac{\omega(d)}{d} \psi(d)$$

$$\geqslant (-1)^{\nu+1} \sum_{d|P(z)} \mu(d)\chi_x^{(-)^\nu}(d) \frac{\omega(d)}{d} \psi(d).$$

Proof. By (I.3.5) and (I.3.6) we know that $(-1)^{\nu+1}\mu(d)\bar{\chi}_x^{(-)^\nu}(d) \geqslant 0$ for any $d | P(z_2)$. Also, if $t > 1$, $q(t) < p(d_1)$ we have, using (I.4.8),

$$\bar{\chi}_x^{(-)^\nu}(d_1 t) = \chi_x^{(-)^\nu}(d_1)\bar{\chi}_{x/d_1}^{(-)^{\nu(d_1)+\nu}}(t) \qquad (\nu = 0, 1).$$

Hence, by (4.11) and (4.12),

$$0 \leqslant (-1)^{\nu+1} \sum_{\substack{n|P(z) \\ (n,P(z_0))>1}} \mu(n)\bar{\chi}_x^{(-)^\nu}(n) \frac{\omega(n)}{n} V(p(n))\psi(n)$$

$$= (-1)^{\nu+1} \sum_{d|P(z_0,z)} \mu(d)\chi_x^{(-)^\nu}(d) \frac{\omega(d)}{d} \psi(d) \times$$

$$\times \sum_{1<t|P(z_0)} \mu(t)\bar{\chi}_{x/d}^{(-)^{\nu(d)+\nu}}(t) \frac{\omega(t)}{t} V(p(t))$$

on first writing $n = d_1 t$, $d_1 | P(z_0, z)$ and $t | P(z_0)$, and then dropping the suffix in d_1. We now apply Lemma 2 of I — the Fundamental Identity — to the inner sum (the sum over t) with $z = z_0$, $\chi = \chi_{x/d}^{(-)^{\nu(d)+\nu}}$ and $\varphi(d) = \omega(d)/d$ (so that $\sum_{t|P(p(d))} \mu(t)\varphi(dt) = \omega(d)V(p(d))/d$ in Lemma 2 of I). It is then clear that the inner sum is equal to

$$V(z_0) - \sum_{t|P(z_0)} \mu(t)\chi_{x/d}^{(-)^{\nu(d)+\nu}}(t) \frac{\omega(t)}{t}$$

and we have

$$0 \leqslant V(z_0)(-1)^{\nu+1} \sum_{d|P(z_0,z)} \mu(d)\chi_x^{(-)^\nu}(d) \frac{\omega(d)}{d} \psi(d) -$$

$$-(-1)^{\nu+1} \sum_{d|P(z_0,z)} \mu(d)\chi_x^{(-)^\nu}(d) \frac{\omega(d)}{d} \psi(d) \sum_{t|P(z_0)} \mu(t)\chi_{x/d}^{(-)^{\nu(d)+\nu}}(t) \frac{\omega(t)}{t}.$$

But by (I.4.8), $\chi_x^{(-)^\nu}(d)\chi_{x/d}^{(-)^{\nu(d)+\nu}}(t) = \chi_x^{(-)^\nu}(dt)$. Hence, combining d and t and invoking (4.11), again, the lemma follows at once.

By (4.10) and Lemma 3 (with $\nu = 1$, $z = z_2$ and $\psi(d) = W(d)$), we arrive finally at

$$H(\mathscr{A}, z_1, z_2) \geqslant X \sum_{d|P(z_2)} \mu(d)\chi_x^-(d) W_0(d)\frac{\omega(d)}{d} + R + O(XV(z_2)Le^{-L/5}) \tag{4.13}$$

$$= XG + R + O(XV(z_2)Le^{-L/5}),$$

say, where

$$G = \sum_{d|P(z_2)} \mu(d)\chi_x^-(d) W_0(d)\frac{\omega(d)}{d}. \tag{4.14}$$

5. The remainder sum R

We show in this section that the problem of estimating R may be transmuted to the estimation of a certain bilinear form.

PROPOSITION 1. *Let M and N be any two real numbers satisfying*

$$M > z_2, \quad N > 1, \quad MN = y.$$

Then there exist real coefficients a_m $(1 \leqslant m < M)$, b_n $(1 \leqslant n < N)$ satisfying

$$|a_m| \leqslant 1, \quad |b_n| \leqslant 1$$

such that

$$|R| < (\log y)^{1/3} \Big| \sum_{\substack{m<M \\ mn|P(z_2)}} \sum_{n<N} a_m b_n R_{mn} \Big|.$$

The proof requires some preparation.

In Σ_1 (cf. (4.6), where $R = \Sigma_1 - \Sigma_2$ is defined) we first separate out the term arising from $D = \emptyset$, when $d = 1$ is the only integer belonging to D, and obtain

$$\Sigma_1^{(1)} = \sum_{\substack{m<z_0^L \\ m|P(z_0)}} \mu(m)\beta^-(m) R_m. \tag{5.1}$$

In the remaining terms of Σ_1 we have $\nu(D) \geqslant 1$. We write each $d \in D$ uniquely in the form

$$d = qd_1, \quad q(d) = q \quad \text{and} \quad q(d_1) < q, \quad q \in I_0, \tag{5.2}$$

say, where, unambiguously,

$$D = I_0 D_1, \quad D_1 < I_0. \tag{5.3}$$

Accordingly we may write (after dropping the suffices in D_1 and d_1 at the last)

$$\begin{aligned}(5.4)\qquad \Sigma_1^{(2)} &= \Sigma_1 - \Sigma_1^{(1)} \\ &= \sum_{I_0} \sum_{\substack{D\in\mathscr{D}\\ D<I_0}} \mu(I_0 D)\chi_x^-(I_0 D) \sum_{q\in I_0} \sum_{d\in D} W_0(qd) \sum_{\substack{m<z_0^L\\ m|P(z_0)}} \mu(m)\beta^{(-)^{\nu(D)}}(m) R_{qdm};\end{aligned}$$

we note for future reference that, whenever $\chi_x^-(I_0 D) \neq 0$, (4.3) may be invoked to give

$$(5.5)\qquad 0 \leqslant W_0(qd) \leqslant 1.$$

In Σ_2 (see (4.6) again for the definition) $\nu(D)$ is always odd and therefore at least 1. We may therefore apply decompositions (5.2) and (5.3) again, and when we do we obtain

$$(5.6)\qquad \Sigma_2 = \sum_I \sum_{I_0} \sum_{\substack{D\in\mathscr{D}\\ I<D<I_0\\ 2|\nu(D)}} \chi_x^-(II_0 D) \sum_{q\in I_0} \sum_{\substack{p',p\in I\\ p'<p}} \sum_{d\in D} \sum_{\substack{m<z_0^L\\ m|P(z_0)}} \mu(m)\beta^+(m) R_{p'pqdm}.$$

To take $\Sigma_1^{(2)}$ further we require a combinatorial result characteristic of the original Iwaniec approach. To state this result we need to recall that $\Delta(D)$ denotes, for $D\in\mathscr{D}$ and D as given in (2.8), the product of the right hand end points of $I_1, \ldots, I_r$.

LEMMA 4 (cf. Iwaniec [10]). *Suppose that*

$$(5.7)\qquad \chi_x^-(I_0 D) = 1, \quad D < I_0,$$

and let real numbers Y_1, Y_2 be given to satisfy

$$(5.8)\qquad Y_1 > z_2, \quad Y_2 > 1, \quad Y_1 Y_2 = x.$$

Then there exists a decomposition

$$I_0 D = D_1 D_2$$

such that

$$\Delta(D_1) \leqslant Y_1 \quad \text{and} \quad \Delta(D_2) \leqslant Y_2.$$

It follows for every integer qd in $I_0 D$, with q prime and in I_0 and d in D, that qd may be written in the form

$$qd = d_1 d_2, \quad d_1 \in D_1 \text{ and } d_2 \in D_2, \quad d_1 \leqslant Y_1 \text{ and } d_2 \leqslant Y_2.$$

Proof. When $D = \emptyset$, we may take $D_1 = I_0$, $D_2 = \emptyset$, so that $\Delta(D_1) = i_0 < z_2$ and $\Delta(D_2) = 1$. This is a decomposition of the type required. (Here, if $qd \in I_0 D$, $q \in I_0$ and is prime, and $d = 1$; put $d_1 = q$, $d_2 = 1$.)

Now suppose $\nu(D) \geqslant 1$. When $\nu(D) = 1$, (5.7) implies that

$$\Delta(D)^3 i_0 < x. \tag{5.9}$$

If $\Delta(D) i_0 < Y_1$, take $D_1 = I_0 D$, $D_2 = \emptyset$ (so that $d_1 = qd$, $d_2 = 1$). Otherwise take $D_1 = I_0$, $D_2 = D$ (so that $d_1 = q$, $d_2 = d$); here $\Delta(D_1) = i_0 \leqslant z_2$, and

$$\Delta(D_2) = \Delta(D) < \frac{x}{\Delta(D)^2 i_0} < \frac{x}{\Delta(D) i_0} \leqslant \frac{x}{Y_1} = Y_2,$$

as required.

Now let $\nu(D) = r$, $r \geqslant 2$, and write

$$D = I_1 \ldots I_r, \quad I_1 > \ldots > I_r. \tag{5.10}$$

We proceed by induction on r and assume that the result has already been proved for all $I_0 D$ with $\nu(D) = r-1$, so that, in particular, one has

$$I_0 I_1 \ldots I_{r-1} = D_1' D_2', \quad \Delta(D_j') \leqslant Y_j \ (j = 1, 2). \tag{5.11}$$

By (5.7) and (2.15) we have

$$i_r \Delta(D) i_0 < x,$$

so that by (5.11) and (5.10)

$$i_r^2 \Delta(D_1') \Delta(D_2') = i_0 i_1 \ldots i_{r-1} i_r^2 < x.$$

Either $i_r \Delta(D_1') \leqslant Y_1$, in which case we put

$$D_1 = D_1' I_r, \quad D_2 = D_2';$$

or $i_r \Delta(D_1') > Y_1$, in which case $i_r \Delta(D_2') < x/Y_1 = Y_2$ and so we put

$$D_1 = D_1', \quad D_2 = D_2' I_r.$$

In either case the inductive step has been accomplished and the lemma proved.

We now apply Lemma 4 to $\Sigma_1^{(2)}$, as given in (5.4). For each set $I_0 D$ occurring in this sum we may assume (5.7) to be satisfied. In accordance with Lemma 4, therefore,

$$|\Sigma_1^{(2)}| \leqslant \sum_{D_1 \in \mathscr{D}} \sum_{D_2 \in \mathscr{D}} \chi_x^-(D_1 D_2) \times \tag{5.12}$$

$$\times \Big| \sum_{\substack{d_1 \in D_1 \\ d_1 < Y_1}} \sum_{\substack{d_2 \in D_2 \\ d_2 < Y_2 \\ d_1 d_2 | P(z_0, z_2)}} W_0(d_1 d_2) \sum_{\substack{m < z_0^L \\ m | P(z_0)}} \mu(m) \beta^{(-)^{\nu(D_1, D_2)+1}}(m) R_{d_1 d_2 m} \Big|;$$

note that, by (5.5) (from Lemma 4, $qd = d_1 d_2$)

$$0 \leqslant W_0(d_1 d_2) \leqslant 1 \tag{5.13}$$

in each term on the right. Now let N_1, N_2 be any two real numbers satisfying

$$N_1 > z_2, \quad N_2 > 1, \quad N_1 N_2 = y. \tag{5.14}$$

Then the innermost triple sum on the right of (5.12) is, in absolute value, at most $2R^{(0)}$ where $R^{(0)}$ is given by

$$R^{(0)} = \max \Big| \sum_{n_1 < N_1} \sum_{\substack{n_2 < N_2 \\ n_1 n_2 | P(z_2)}} a_{n_1} b_{n_2} R_{n_1 n_2} \Big| \tag{5.15}$$

where the maximum is with respect to all sets of coefficients a_{n_1}, b_{n_2} satisfying

$$|a_{n_1}| \leqslant 1, \quad |b_{n_2}| \leqslant 1. \tag{5.16}$$

To see this we argue as follows. If $N_2 \geqslant z_0^L$ take

$$n_1 = d_1, \quad n_2 = md_2, \quad Y_1 = N_1, \quad Y_2 = N_2 z_0^{-L}$$

in (5.12), so that $n_1 < N_1$ and $n_2 < N_2$. If $N_2 < z_0^L$, (5.14) implies that $N_1 > yz_0^{-L} > z_0^L$ (also $N_1 z_0^{-L} > z_2$), and this time we take

$$n_1 = d_1 m, \quad n_2 = d_2, \quad Y_1 = N_1 z_0^{-L}, \quad Y_2 = N_2$$

in (5.12) so that again $n_1 < N_1$ and $n_2 < N_2$. In both cases note that $Y_1 Y_2 = yz_0^{-L} = x$ by (2.9), so that (5.8) in Lemma 4 is satisfied; note also, that the indicated representations of n_1, n_2 are unique, with m dividing n_1 or n_2 and $n_1 n_2 = d_1 d_2 m$ so that, by construction, $n_1 n_2 | P(z_2)$. As for the coefficients, for a given pair D_1, D_2, $\nu(D_1 D_2)$ is the same for all the terms in the triple sum and, depending only on the size of N_2, $\mu(m)\beta^{(-)^{\nu(D_1 D_2)+1}}(m)$ is therefore associated wholely with n_1 or with n_2. As for W_0, (5.13) and (2.5) imply that

$$0 \leqslant W_0(d_1 d_2) = W_0(d_1 d_2 m) = W_0(n_1 n_2) \leqslant 1,$$

and since $0 \leqslant w(p) \leqslant 1$ it follows that $0 \leqslant W_0(n_1) \leqslant 1$ and $0 \leqslant W_0(n_2) \leqslant 1$.

The identity

$$W_0(n_1 n_2) = W_0(n_1) - \big(1 - W_0(n_2)\big)$$

therefore enables us to conclude that the innermost triple sum on the right of (5.12) is, in absolute value, at most $2R^{(0)}$.

It follows at once, from (5.12) and (2.7), that

$$|\Sigma_1^{(2)}| \leqslant 2^{2J+1} R^{(0)}. \tag{5.17}$$

It remains to deal with Σ_2, now to be thought of in the form (5.6). For this we require the following companion to Lemma 4.

LEMMA 5 (cf. Motohashi [12]). *Suppose that*

$$I < D < I_0, \quad 2 \mid \nu(D), \quad D \in \mathscr{D}, \tag{5.18}$$

and

$$\chi_x^- (I_0 D I) = 1. \tag{5.19}$$

Let Y_1, Y_2 *be any two given real numbers satisfying*

$$Y_1 > z_2, \quad Y_2 > 1, \quad Y_1 Y_2 = x. \tag{5.20}$$

Then there exists a decomposition (1)

$$I_0 D I^2 = D_1 D_2$$

such that

$$\Delta(D_1) \leqslant Y_1 \quad \text{and} \quad \Delta(D_2) \leqslant Y_2.$$

It follows for every integer $qdpp'$ *in* $I_0 D I^2$, *with* q *prime and in* I_0, d *in* D, *and* p, p' *in* $I (p' < p)$ *that* $qdpp'$ *may be written in the form*

$$qdpp' = d_1 d_2, \quad d_1 \in D_1, \quad d_2 \in D_2, \quad d_1 \leqslant Y_1, \quad d_2 \leqslant Y_2,$$

where $D_j \in \mathscr{D}$ *or* $D_j = D_j' I^2$ *with* $D_j' \in \mathscr{D}$ *and* $I < D_j'$.

Proof. Since (5.19) implies (5.7) we may take advantage of Lemma 4 and conclude that there exists a decomposition

$$I_0 D = D_1' D_2' \tag{5.21}$$

with

$$\Delta(D_1') \leqslant Y_1 \quad \text{and} \quad \Delta(D_2') \leqslant Y_2. \tag{5.22}$$

Also, since $2 \mid \nu(D)$, (5.19) implies that

$$i_0 \Delta(D) i^3 < x$$

and hence, by (5.21), that

$$\Delta(D_1') \Delta(D_2') i^3 < x. \tag{5.23}$$

Either $\Delta(D_1') i^2 \leqslant Y_1$, in which case choose

$$D_1 = D_1' I^2, \quad D_2 = D_2' \quad (\text{so that } d_1 = d_1' pp' \text{ with } p, p' \in I,\ d_2 = d_2')$$

or $\Delta(D_2') i^2 \leqslant Y_2$, in which case choose

$$D_1 = D_1', \quad D_2 = D_2' I^2 \quad (\text{so that } d_1 = d_1',\ d_2 = d_2' pp' \text{ with } p, p' \in I);$$

---

(1) Here I^2 is to denote only all distinct pairs p', p of primes in I.

and if neither of these possibilities holds, then, by (5.23), necessarily

$$\Delta(D_1')\, i < Y_1 < \Delta(D_1')\, i^2$$

and

$$\Delta(D_2')\, i < Y_2 < \Delta(D_2')\, i^2.$$

In this case choose

$$D_1 = D_1' I, \qquad D_2 = D_2' I$$

(so that $d_1 = d_1' p$, $d_2 = d_2' p'$ with $p, p' \in I$ and $p' < p$).

The stated decomposition (5.21) together with (5.22) has now been established in each of the three cases which exhaust all possibilities.

Now apply Lemma 5 to Σ_2, as given by (5.6). Conditions (5.18) and (5.19) of Lemma 5 are satisfied so that

$$|\Sigma_2| \leqslant \sum_{D_1 \in \mathscr{D}^*} \sum_{D_2 \in \mathscr{D}^*} \Big| \sum_{\substack{d_1 \in D_1 \\ d_1 < Y_1}} \sum_{\substack{d_2 \in D_2 \\ d_2 < Y_2 \\ d_1 d_2 | P(z_0, z_2)}} \sum_{\substack{m < z_0^L \\ m | P(z_0)}} \mu(m) \beta^+(m) R_{d_1 d_2 m} \Big|,$$

where $\mathscr{D}^* = \mathscr{D} \cup \{DI^2 \colon D \in \mathscr{D},\ I < D\}$.

Arguing as before in the case of $\Sigma_1^{(2)}$ (except that here there is no W to worry about), we have, by (2.6),

$$|\Sigma_2| \leqslant (J+1)^2\, 2^{2J} R^{(0)}.$$

Since $z_2 > z_0^L$, $\Sigma_1^{(1)}$ is trivially dominated by $R^{(0)}$, so that, by (5.17), altogether

$$|R| \leqslant 2J^2\, 2^{2J} R^{(0)}. \tag{5.24}$$

The proof of Proposition 1 is all but complete. We have only to estimate the factor $2J^2 2^{2J}$ on the right of (5.24). For this purpose note first that, by (1.5) and (1.6),

$$y^{1/2} \leqslant z_2 = y^U < y$$

so that (1.4) implies that also z_2 is sufficiently large. By (2.3) and (2.1) we have therefore

$$J \leqslant \frac{\log\log\log z_2}{5 \log(1+\delta)} \leqslant \frac{\log\log\log z_2}{5\delta}(1+\delta) = \frac{1+\delta}{5} \log\log z_2$$

and hence

$$2J^2\, 2^{2J} < (\log\log z_2)^2 (\log z_2)^{3/10} < (\log y)^{1/3}.$$

This completes the proof of Proposition 1; the coefficients a_m, b_n in Proposition 1 are simply taken to be those giving the maximum of $R^{(0)}$, with $N_1 = M$ and $N_2 = N$.

6. The main term

We now turn to the main term G on the right of (4.13), given by (4.14). This term differs from the leading term G of (I.6.2) only in that it involves χ_x^- rather than χ_y^-. The error term induced by this difference is the price we pay here for having the remainder sum R in a more flexible form. Nevertheless, the preparatory analysis of I, Section 6 applies here also and leads us in much the same way to (cf. I.6.15)

$$G = T^-(x, x^{1/4}) - \sum_{x^{1/4} \leqslant p < z_2} (1 - w(p)) \frac{\omega(p)}{p} \sigma^+\left(\frac{x}{p}\right) + \tag{6.1}$$
$$+ \sum_{y^V \leqslant p < x^{1/4}} w(p) \frac{\omega(p)}{p} \left\{\sigma^+\left(\frac{x}{p}\right) - \sum_{\substack{n \mid P(p^+, z_2) \\ p(n) < x/(pn)}} \bar{\chi}_x^-(n) \frac{\omega(n)}{n} \sigma^+\left(\frac{x}{pn}\right)\right\} +$$
$$+ O\left(V(x) \log^{-1/3} x\right)$$

where we have omitted (as we shall continue to do from here on) the subscript $\varkappa$ ($= 1$ now) from T and σ^+ and set $\beta = \beta(1) = 2$. We apply (I.7.1) to the first term on the right of (6.1); since $A_\varkappa = A_1 = 2e^\gamma$ ([9], p. 176), where γ is Euler's constant, we obtain

$$T^-(x, x^{1/4}) \geqslant V(x) 2e^\gamma \{\log 3 + O(\log^{-1/3} x)\}. \tag{6.2}$$

Our O- and $\ll$- constants depend from now on at most on A and U.

In order to apply (I.7.2) to the second expression on the right of (6.1) we have first to allow for the difference in the summation conditions. By (1.8), (I.4.4) and (I.2.4) we have (remember that $z_2 = y^U$)

$$\sum_{x^U \leqslant p < y^U} (1 - w(p)) \frac{\omega(p)}{p} \sigma^+\left(\frac{x}{p}\right) \ll V(x)\left(\log \frac{\log y}{\log x} + \frac{1}{\log x}\right) \tag{6.3}$$
$$\ll V(x) \frac{\log \log \log y}{(\log \log y)^{1/5}}$$

since, by (2.9) and (2.4),

$$\frac{\log y}{\log x} = \left(1 - LU \frac{\log z_0}{\log z_2}\right)^{-1} \leqslant \left(1 - U(1+\delta) \frac{L}{(\log \log z_2)^{1/5}}\right)^{-1} \tag{6.4}$$

so that, by (2.1),

$$\log \frac{\log y}{\log x} \ll \frac{\log \log \log y}{(\log \log y)^{1/5}}. \tag{6.5}$$

Hence

$$\sum_{x^{1/4} \leqslant p < z_2} (1-w(p)) \frac{\omega(p)}{p} \sigma^+\left(\frac{x}{p}\right) = \sum_{x^{1/4} \leqslant p < x^U} (1-w(p)) \frac{\omega(p)}{p} \sigma^+\left(\frac{x}{p}\right) + $$
$$+ O\left(V(x) \frac{\log\log\log y}{(\log\log y)^{1/5}}\right).$$

We are still not quite ready to apply (I.7.2) to the sum on the right because the definition of $w(\cdot)$ involves $\log y$. However, the replacement of $\log y$ by $\log x$ in this context yields an error at most of order

$$V(x) \sum_{x^{1/4} \leqslant p < x^U} \left(\frac{\log p}{\log x} - \frac{\log p}{\log y}\right) \frac{\omega(p)}{p} \leqslant V(x) \left(\frac{1}{\log x} - \frac{1}{\log y}\right) \left((U-\tfrac{1}{4}) \log x + A\right)$$
$$\ll V(x) \frac{\log\log\log y}{(\log\log y)^{1/5}}$$

by $(\Omega(1))$ and (6.4). Therefore we may modify the sum on the right in this way and then apply (I.7.2); and we then obtain

$$\text{(6.6)} \quad \sum_{x^{1/4} \leqslant p < z_2} (1-w(p)) \frac{\omega(p)}{p} \sigma^+\left(\frac{x}{p}\right)$$
$$= V(x) \frac{2e^\gamma}{T-E} \int_{1/U}^{4} \left(T - \frac{1}{t}\right) \frac{dt}{t-1} + O\left(V(x) \frac{\log\log\log y}{(\log\log y)^{1/5}}\right)$$
$$= V(x) \frac{2e^\gamma}{T-E} \left\{T \log 3 - T \log \frac{1}{U} - (1-T) \log \frac{1}{1-U} + \log \frac{4}{3} + \right.$$
$$\left. + O\left(\frac{\log\log\log y}{(\log\log y)^{1/5}}\right)\right\}.$$

A similar procedure applies to the third expression on the right of (6.1). We want to alter the range of summation to $[x^V, x^{1/4})$, and this we may do, by virtue of (I.4.4) and (1.8), at a cost of an error at most of order

$$V(x) \sum_{x^V \leqslant p < y^V} \frac{\omega(p)}{p} \left\{1 + \sum_{n \mid P(y^V, y^U)} \frac{\omega(n)}{n}\right\} \ll V(x) \frac{\log\log\log y}{(\log\log y)^{1/5}}$$

by (I.2.6) and (6.5). Then, as before, we must replace $\log y$ in the definition of $w(p)$ by $\log x$, and this costs us an error of order at most

$$V(x) \sum_{y^V \leqslant p < y^{1/4}} \left(\frac{\log p}{\log x} - \frac{\log p}{\log y}\right) \frac{\omega(p)}{p} \left\{1 + \sum_{n \mid P(y^V, y^U)} \frac{\omega(n)}{n}\right\}$$
$$\ll V(x) \left(\frac{1}{\log x} - \frac{1}{\log y}\right) \left((\tfrac{1}{4} - V) \log y + A\right) \ll V(x) \frac{\log\log\log y}{(\log\log y)^{1/5}}$$

by (6.4). With the third expression on the right of (6.1) modified in these ways, and invoking also the remark that follows (I.7.3) – this allows us now to replace the condition $n|P(p^+, z_2)$ in the inner sum by $p(n) > p$ and insertion in the term being summed a factor $\mu^2(n)$ – we have by (6.1), (6.2), (6.6) and (I.4.4) that (cf. (I.7.7))

$$(6.7)\quad G \geqslant V(x)\frac{2e^\gamma}{T-E}\Big\{T\log\frac{1}{U}+(1-T)\log\frac{1}{1-U}-\log\frac{4}{3}-E\log 3+ +O\left(\frac{\log\log\log y}{(\log\log y)^{1/5}}\right)+ + \sum_{x^V \leqslant p < x^{1/4}}\left(\frac{\log p}{\log x}-E_0\right)\frac{\omega(p)}{p}\times \times\left(\frac{\log x}{\log(x/p)}-\sum_{p<p(n)<x/(pn)}\mu^2(n)\overline{\chi}_x^-(n)\frac{\omega(n)}{n}\frac{\log x}{\log(x/(pn))}\right)\Big\}.$$

For the moment let Σ denote the innermost sum on the right, so that

$$(6.8)\quad \Sigma = \sum_{\substack{n\\ p<p(n)<x/(pn)}}\mu^2(n)\overline{\chi}_x^-(n)\frac{\omega(n)}{n}\frac{\log x}{\log(x/(pn))} = \sum_{r\geqslant 1}\textstyle\sum_{2r}^-(x, p),$$

say, where

$$\textstyle\sum_k^{(-)^\nu}(\xi, p) = \displaystyle\sum_{\substack{p<p_k<\dots<p_1\\ p_k<x/(pp_1\dots p_k)}}\overline{\chi}_\xi^{(-)^\nu}(p_1\dots p_k)\frac{\omega(p_1\dots p_k)}{p_1\dots p_k}\frac{\log\xi}{\log(\xi/(p_1\dots p_k p))}.$$

By (I.4.8) (see also the beginning of the proof of Lemma 3 above) we have

$$\overline{\chi}_\xi^{(-)^\nu}(p_1\dots p_k) = \chi_\xi^{(-)^\nu}(p_1)\chi_{\xi/p_1}^{(-)^{\nu+1}}(p_2\dots p_k)$$

so that

$$(6.9)\quad \textstyle\sum_k^{(-)^\nu}(\xi, p) = \displaystyle\sum_{p_1}\frac{\omega(p_1)}{p_1}\chi_\xi^{(-)^\nu}(p_1)\textstyle\sum_{k-1}^{(-)^{\nu+1}}\displaystyle\left(\frac{\xi}{p_1}, p\right),\quad k\geqslant 1.$$

This recursion formula enables us to proceed by induction on k via the use of Lemma 6 of I for converting sums to integrals. In each application of Lemma 6 the quantity B in that result is $\ll 1$, and Σ contains only finitely many terms since the condition $p < p(n) < x/(pn)$ implies that

$$(6.10)\quad x > p^{2+\nu(n)} > x^{(2+\nu(n))V}$$

so that for each n counted in Σ, $\nu(n) < V^{-1}-2$. A straightforward calculation leads to

$$\sum\nolimits_{2r}^{-}(x,p) = \int\limits_{\substack{p<\tau_{2r}<\ldots<\tau_1 \\ \tau_{2r}<x/(p\tau_1\ldots\tau_{2r}) \\ \tau_{2i}^3\ldots\tau_1<x(i<r) \\ \tau_{2r}^3\ldots\tau_1\geqslant x}}\!\!\cdots\!\int \frac{d\tau_1\ldots d\tau_{2r}}{\tau_1\log\tau_1\ldots\tau_{2r}\log\tau_{2r}}\,\frac{\log x}{\log\left(x/(p\tau_1\ldots\tau_{2r})\right)}+O\left(\frac{1}{\log x}\right)$$

where all but the first two sets of integration conditions derive from the inequalities implicit in $\bar{\chi}_x^-$. The change of variables

$$p = x^t, \quad \tau_i = y^{t_i}, \quad 1 \leqslant i \leqslant 2r,$$

leads directly to

$$\sum\nolimits_{2r}^{-}(x,p) = h_{2r}\left(\frac{\log p}{\log x}\right)+O\left(\frac{1}{\log y}\right), \tag{6.11}$$

where

$$h_{2r}(t) = \int\limits_{\substack{t<t_{2r}<\ldots<t_1 \\ 3t_{2i}+\ldots+t_1<1(i=1,\ldots,r-1) \\ 3t_{2r}+\ldots+t_1\geqslant 1 \\ t_{2r}<1-t-t_1-\ldots-t_{2r}}}\!\!\cdots\!\int \frac{dt_1\ldots dt_{2r}}{t_1\ldots t_{2r}}\,\frac{1}{1-t-t_1-\ldots-t_{2r}}. \tag{6.12}$$

We write

$$h(t) = \sum_{r\geqslant 1} h_{2r}(t), \tag{6.13}$$

so that, by (6.8),

$$\Sigma = h\left(\frac{\log p}{\log x}\right)+O\left(\frac{1}{\log y}\right).$$

Hence

$$\sum_{x^V\leqslant p<x^{1/4}}\left(\frac{\log p}{\log x}-E_0\right)\frac{\omega(p)}{p}\left(\frac{\log x}{\log(x/p)}-\Sigma\right) \tag{6.14}$$

$$= \sum_{x^V\leqslant p<x^{1/4}}\left(\frac{\log p}{\log x}-E_0\right)\frac{\omega(p)}{p}\left(\frac{1}{1-\dfrac{\log p}{\log x}}-h\left(\frac{\log p}{\log x}\right)\right)+O\left(\frac{1}{\log y}\right)$$

by $(\Omega(1))$. The function

$$\psi(t) = \frac{1}{1-t}-h(t), \quad 0<t\leqslant 1/4, \tag{6.15}$$

is, as Greaves [2] proved, increasing and has a unique zero V_0 given by (Greaves [2], p. 331)

$$V_0 = 0.074\,368\ldots\ (= 1/13.446\ldots) \tag{6.16}$$

so that

$$\psi(t) \geqslant 0 \quad \text{for} \quad t \geqslant V_0.$$

Hence Lemma 6 of I may be applied to the sum (6.14). We have $\max\limits_{V \leqslant t \leqslant 1/4} \psi(t) = \psi(1/4) = 4/3$ since $h(1/4) = 0$ by (6.10). Thus the quantity B in Lemma 6 is $\ll 1$, and the sum (6.14) is equal to

$$\int\limits_{V}^{1/4} \frac{t-E_0}{t}\,\psi(t)\,dt + O\left(\frac{1}{\log y}\right). \tag{6.17}$$

Following Greaves [2], we put

$$\alpha(V) = \int\limits_{V}^{1/4} \psi(t)\,dt, \qquad \beta(V) = \int\limits_{V}^{1/4} \psi(t)\,\frac{dt}{t}, \tag{6.18}$$

so that the sum (6.14) is equal to

$$\alpha(V) - E_0\beta(V) + O\left(\frac{1}{\log y}\right).$$

It follows from (6.7) and (6.14) that

$$G \geqslant V(x)\,\frac{2e^{\gamma}}{T-E}\left\{T\log\frac{1}{U} + (1-T)\log\frac{1}{1-U} - \left(\log\tfrac{4}{3} - \alpha(V)\right) - \right. \\ \left. - E\log 3 - E_0\beta(V) + O\left(\frac{\log\log\log y}{(\log\log y)^{1/5}}\right)\right\}. \tag{6.19}$$

It seems at first, from the form (6.17) of the sum (6.14), that the right choice of V is V_0, the point where $\psi(t)$ changes from being negative as V increases. However, it turns out that the optimal choice of V depends on the application in view; for example, in the problem of P_2's in short intervals the best choice of V lies closer to 1/6. Obviously the pairs $\alpha(V)$, $\beta(V)$ require tabulation. We record here only the values

$$\alpha(V_0) = 0.150\,5528\ldots, \qquad \beta(V_0) = 0.876\,95\ldots,$$

(Greaves [2], pp. 301, 331) and

$$\alpha(1/6) = 0.098\,580\,030\ldots, \qquad \beta(1/6) = 0.474\,533\,776\ldots$$

(Grupp–Richert [4]). The choice $V \geqslant 1/6$ is interesting because here, by (6.10), $h(t)$ reduces to the single term $h_2(t)$ and it is not hard to derive from (6.12) that (cf. (I.7.12) and (I.7.13))

$$h(t) = h_2(t) = \frac{t}{1-t} \int_2^{1/t-2} \frac{\log(u-1)}{2-(u+2)t}\,du, \qquad 1/6 \leqslant t \leqslant 1/4. \tag{6.20}$$

7. A lower bound for the sifting function

For convenience and quotation later on we summarize our results so far in the form of a theorem.

THEOREM A. *We postulate conditions* (A_0) *and* $(\Omega(1))$. *The constants* T, U, V, E *and*

$$E_0 = \max\left(E, \tfrac{1}{3}(1-T)\right)$$

are to satisfy

$$E_0 \leqslant V \leqslant 1/4, \qquad 1/2 \leqslant U \leqslant T < 1, \qquad U+3V \geqslant 1, \qquad V \geqslant V_0 \tag{7.1}$$

where V_0 *is given in* (6.16). *Then, as* $y \to \infty$ *we have*

$$\begin{aligned} H(\mathscr{A}, y^V, y^U) \geqslant XV(y)\frac{2e^\gamma}{T-E}\Bigg\{ & T\log\frac{1}{U} + (1-T)\log\frac{1}{1-U} - \\ & -\left(\log\tfrac{4}{3} - \alpha(V)\right) - E\log 3 - E_0\beta(V) + \\ & + O\left(\frac{\log\log\log y}{(\log\log y)^{1/5}}\right)\Bigg\} - (\log y)^{1/3}\Bigg|\sum_{m<M}\ \sum_{\substack{n<N \\ mn|P(y^U)}} a_m b_n R_{mn}\Bigg|, \end{aligned} \tag{7.2}$$

where $\alpha(V)$ and $\beta(V)$ are defined in (6.18), M and N are any two real numbers satisfying

$$M > y^U, \qquad N > 1, \qquad MN = y, \tag{7.3}$$

and a_m, b_n are certain real numbers satisfying

$$|a_m| \leqslant 1, \qquad |b_n| \leqslant 1.$$

The O-constant depends at most on A (from $(\Omega(1))$) and on U.

This result is merely a matter of combining (4.13), (6.19) and Proposition 1; recall that $z_1 = y^V$, $z_2 = y^U$, $L = \log\log\log z_2$, that

$$V(x) = V(y)\left(1 + O\left(\frac{\log\log\log y}{(\log\log y)^{1/5}}\right)\right)$$

by (I.2.2) (with $\varkappa = 1$) and (6.5), and that

$$V(z_2)\,Le^{-L/5} \ll V(y)\,\frac{\log\log\log y}{(\log\log y)^{1/5}}$$

by (I.2.2) (with $\varkappa = 1$) and (7.1).

In the leading term on the right of (7.2), when it comes to applications, T, U, V, E and E_0 are numerical constants and only the expression $\alpha(V)-E_0\beta(V)$ presents a computational problem.

Greaves [3] has developed a method for computing these numbers via certain moments (for a more direct approach see Grupp–Richert [4]), but it comes as a very pleasant surprise that heavy computing may after all be avoided at the cost of only a slight loss in accuracy. Let (cf. (6.15) and (6.13))

$$\psi_k(t) = \frac{1}{1-t} - h_2(t) - \ldots - h_{2k}(t), \qquad 0 < t \leqslant 1/4, \tag{7.4}$$

so that

$$\psi_1(t) \geqslant \psi_2(t) \geqslant \ldots \geqslant \psi(t) \geqslant 0 \quad \text{for} \quad V_0 \leqslant t \leqslant 1/4 \tag{7.5}$$

and, by (6.10) or (6.12) (cf. (6.20))

$$\psi(t) = \psi_k(t), \qquad V_0 \leqslant 1/(2k+4) \leqslant t \leqslant 1/4. \tag{7.6}$$

For each $k = 1, 2, 3, 4$ we may view $\psi_k(t)$ as an approximation to $\psi(t)$ on the whole interval $[V_0, \frac{1}{4}]$, and (cf. (6.18)) the functions

$$\alpha_k(V) = \int_V^{1/4} \psi_k(t)\,dt, \qquad \beta_k(V) = \int_V^{1/4} \psi_k(t)\,\frac{dt}{t} \tag{7.7}$$

as approximations to $\alpha(V)$ and $\beta(V)$. Define the error functions

$$\bar{\alpha}_k(V) = \alpha_k(V) - \alpha(V), \qquad \bar{\beta}_k(V) = \beta_k(V) - \beta(V), \tag{7.8}$$

and let

$$\Delta_k(V) = \bar{\alpha}_k(V) - V\bar{\beta}_k(V), \tag{7.9}$$

$$\varepsilon_k(V) = \Delta_k(V) + (V - E_0)\bar{\beta}_k(V), \tag{7.10}$$

so that

$$\alpha(V) - E_0\beta(V) = \alpha_k(V) - E_0\beta_k(V) - \varepsilon_k(V). \tag{7.11}$$

It is clear from (7.5) that each of $\bar{\alpha}_k(V)$, $\bar{\beta}_k(V)$ is non-negative and non-increasing on $[V_0, \frac{1}{4}]$ and is actually 0 for $1/(2k+4) \leqslant V \leqslant 1/4$. In particular, for each $k = 1, \ldots, 4$, (7.12) $\bar{\beta}_k(V) = 0, \quad \dfrac{1}{2k+4} \leqslant V \leqslant \tfrac{1}{4}; \quad \bar{\beta}_k(V) \leqslant \bar{\beta}_k(V_0), \quad V_0 \leqslant V \leqslant \dfrac{1}{k+4}.$

We now establish the following companion result.

LEMMA 6. *For each* $k = 1, \ldots, 4$, $\Delta_k(V) = 0$ *for* $1/(2k+4) \leqslant V \leqslant 1/4$ *and*

$$\Delta_k(V) \leqslant \frac{1-(2k+4)V}{1-(2k+4)V_0}\Delta_k(V_0) \quad \textit{for} \quad V_0 \leqslant V \leqslant \frac{1}{2k+4}.$$

Proof. The first statement is obvious from what was said earlier. As for the second statement, it is easy to check from (7.8) and (7.7) that

$$\bar{\alpha}_k'(V) = \psi(V) - \psi_k(V) \quad \text{and} \quad \bar{\beta}_k'(V) = \frac{1}{V}(\psi(V) - \psi_k(V)).$$

Hence, by (7.9),

$$\Delta_k'(V) = \bar{\alpha}_k'(V) - V\bar{\beta}_k'(V) - \bar{\beta}_k(V) = -\bar{\beta}_k(V)$$

and

$$\Delta_k''(V) = -\bar{\beta}_k'(V) = \frac{1}{V}(\psi_k(V) - \psi(V)) \geqslant 0$$

by (7.5). This means that $\Delta_k(V)$ is convex, and the inequality now follows since $\Delta_k\left(\dfrac{1}{2k+4}\right) = 0.$

It follows from (7.10), (7.12) and Lemma 6 that, for each $k = 1, \ldots, 4$, $\varepsilon_k(V) = 0$ when $1/(2k+4) \leqslant V \leqslant 1/4$ and

(7.13) $\varepsilon_k(V) \leqslant \big(1-(2k+4)V\big)\delta_k + (V - E_0)\delta_k', \quad V_0 \leqslant V \leqslant 1/(2k+4),$

where

(7.14) $$\delta_k = \frac{\Delta_k(V_0)}{1-(2k+4)V_0} \quad \text{and} \quad \delta_k' = \bar{\beta}_k(V_0).$$

It turns out that the simplest case $k = 1$ already gives very small values of $\varepsilon_1(V)$ for $V_0 \leqslant V \leqslant 1/6$. By (7.7), (7.4) with $k = 1$, and (6.20) a straightforward calculation leads to

(7.15) $$\alpha_1(V) = \log\frac{4}{3}(1-V) - \int_2^{1/V-2} \frac{\log(u-1)}{u} \times$$

$$\times \left\{\frac{2}{u+2}\log\big(2-(u+2)V\big) + \log\frac{u+1}{(1-V)(u+2)}\right\}du, \quad 0 < V \leqslant 1/4,$$

and

$$(7.16)\qquad \beta_1(V) = \log\frac{1-V}{3V} - \int_2^{1/V-2} \frac{\log(u-1)}{u} \times$$
$$\times\left\{\log\left(2-(u+2)V\right)+\log\frac{u+1}{(1-V)(u+2)}\right\}du,\quad 0 < V \leqslant 1/4,$$

and from these formulae $\alpha_1(V)$ and $\beta_1(V)$ are easily computed by numerical integration. It turns out that, in (7.14), using (6.16),

$$\delta_1 \leqslant 0.000706, \qquad \delta_1' \leqslant 0.026756,$$
$$\delta_2 \leqslant 0.00000244, \qquad \delta_2' \leqslant 0.00016975.$$

In practice, these bounds show $\varepsilon_k(V)$ to be very small indeed; for the first term on the right of (7.13) is obviously small since δ_k is so small, and the second term is controlled less by δ_k' than by $V-E_0$ being small in applications and often actually 0. By way of a simple numerical illustration, take $V = 0.1$ and $E_0 = 0.09$. Here

$$\varepsilon_1(0.1) \leqslant 0.00031, \qquad \varepsilon_2(0.1) \leqslant 0.0000022.$$

8. An arithmetical interpretation

Let $\mathscr{P}^c$ denote the complement of $\mathscr{P}$. In any sieve problem it is reasonable (but not essential) to choose $\mathscr{P}$ so that the primes making up the elements a of $\mathscr{A}$ come exclusively from $\mathscr{P}$; and when $\mathscr{P}$ is so chosen we indicate this by the symbolic requirement

$$(\mathrm{A}_1)\qquad (\mathscr{A}, \mathscr{P}^c) = 1.$$

To simplify matters, choose $U = T$ in (7.1)([2]). Also, define, for a in $\mathscr{A}$,

$$(8.1)\qquad \nu(a, y^T) = \nu(a) + \sum_{\substack{p,m\geqslant 2\\ p^m|a,\,p\geqslant y^T}} 1.$$

Then

THEOREM B. *Suppose that conditions* (A_0), (A_1) *and* $(\Omega(1))$ *hold, and that there exist numerical constants* T, V, E *and* $E_0 = \max\left(E, \frac{1}{3}(1-T)\right)$ *satisfying*

$$(8.2)\qquad E_0 \leqslant V \leqslant 1/4, \quad 1/2 \leqslant T < 1, \quad V \geqslant V_0,$$

([2]) Note that then (1.6) implies that $T+3V \geqslant 1$.

where V_0 *is given in* (6.16), *such that*

$$H(\mathscr{A}, \mathscr{P}, y^V, y^T) > 0.$$

Suppose r *is a positive integer such that*

$$\max_{a \in \mathscr{A}} |a| \leqslant y^{rT+E}. \tag{8.3}$$

Then there exists an a *in* $\mathscr{A}$ *such that* $\nu(a, y^T) \leqslant r$.

Proof. By (I.1.4) and (A_1), the fact that $H(\mathscr{A}, \mathscr{P}, y^V, y^T) > 0$ implies the existence of an element a in $\mathscr{A}$ having no prime factor less than y^V such that

$$0 < \{1 - \sum_{\substack{p|a \\ p<y^T}} (1-w(p))\}^+.$$

Then, since $E_0 \geqslant E$ and $T - \log p/\log y \leqslant 0$ when $p \geqslant y^T$, we have by (1.3)

$$0 < T-E- \sum_{\substack{p|a \\ p<y^T}} \left(T-\frac{\log p}{\log y}\right) \leqslant T-E-\sum_{p|a}\left(T-\frac{\log p}{\log y}\right) - \sum_{\substack{p,m\geqslant 2 \\ p^m|a, p\geqslant y^T}} \left(T-\frac{\log p}{\log y}\right)$$

$$\leqslant T-E-T\nu(a, y^T)+\frac{\log |a|}{\log y} \leqslant T-E-T\nu(a, y^T)+rT+E$$

$$= T(r-\nu(a, y^T)),$$

whence the result.

We use Theorem A to show that, subject to a suitable choice of parameters, $H(\mathscr{A}, \mathscr{P}, y^V, y^T) > 0$; but, of course, that theorem will tell us even that then $H(\mathscr{A}, \mathscr{P}, y^V, y^T) \gg XV(y)$. If, as is so often the case, the number

$$\sum_{y^V \leqslant p < y^T} |\mathscr{A}_{p^2}| \tag{8.4}$$

of elements of $\mathscr{A}$ that are not squarefree with respect to the primes between y^V and y^T, is small compared with $H(\mathscr{A}, \mathscr{P}, y^V, y^T)$, then Theorem B leads in the usual way to the stronger conclusion that $\mathscr{A}$ *contains elements having at most* r *prime factors counted according to multiplicity.*

The best results (at least in some applications) derive from taking U and T different; but then the method of Iwaniec–Laborde [11], or some analogous idea will have to be incorporated.

9. An application to almost-primes in short intervals

From now on let x denote a sufficiently large positive real number. We shall try out Theorem B on the problem of almost-primes in short intervals, where the technical difficulties are least and the result of Halberstam, Heath-Brown and Richert [5] affords a basis for comparison. Let

$$\mathscr{A} = \{n:\ x - x^{\theta} < n \leqslant x\}.$$

Then the problem is to show that, for a suitable $\theta = \theta_r$ $(r \geqslant 2)$ and $x \geqslant x_r$, there exist almost-primes P_r of order r in $\mathscr{A}$. Here $\mathscr{P}$ is the set of all primes, so that $\bar{\mathscr{P}}$ is empty and (A_1) is automatically true. Here also $X = x^{\theta}$ and $\omega(p) = 1$ for all primes p, so that (A_0) is trivially true and $(\Omega(1))$ follows from Mertens prime number theory. Finally, we have

$$XV(y)\,2e^{\gamma} = \frac{2x^{\theta}}{\log y}\left(1 + O\left(\frac{1}{\log y}\right)\right),$$

and

$$R_d = \left[\frac{x}{d}\right] - \left[\frac{x - x^{\theta}}{d}\right] - \frac{x^{\theta}}{d}.$$

To deal with the remainder sum in Theorem A, (7.2) we quote Lemma 3 of [5]:

LEMMA 7. *Let A and B be positive numbers satisfying*

$$AB^2 \leqslant x, \tag{9.1}$$

and, given an absolute constant $c_0 \geqslant 2$, let A' and B' satisfy

$$A < A' \leqslant c_0 A, \qquad B < B' \leqslant c_0 B.$$

Then

$$\sum_{A < m \leqslant A'} \sum_{B < n \leqslant B'} a(m)\, b(n)\, R_{mn} \ll x^{\theta - 2\eta} + (A^{1/2} x^{\theta/2} + A^{(\lambda - \varkappa + 1)/2} B x^{\varkappa(1-\theta)/2})\, x^{6\eta} \tag{9.2}$$

where $|a(m)| \leqslant 1$ $(A < m \leqslant A')$, $|b(n)| \leqslant 1\,(B < n \leqslant B')$, $(\varkappa, \lambda)$ is an exponent pair in the sense of van der Corput–Phillips, and $\eta = c_1/\log\log x$ with some positive numerical constant c_1.

The remainder sum in (7.2) is clearly the sum of $\ll (\log M)\,(\log N)$ expressions such as on the left of (9.2) and therefore, if we require M and N to satisfy

$$MN^2 \leqslant x, \tag{9.3}$$

we deduce from Lemma 7 that

$$\sum_{\substack{m < M \\ mn \mid P(y^U)}} \sum_{n < N} a_m b_n R_{mn} \ll x^{\theta - \eta} + (M^{1/2} x^{\theta/2} + M^{(1 + \lambda - \varkappa)/2} N x^{\varkappa(1-\theta)/2})\, x^{7\eta}. \tag{9.4}$$

We write

$$\varrho = ((3+2\varkappa-\lambda)\theta-\varkappa)/2,$$

and require that

$$\theta < \varrho \leqslant (1+\theta)/2. \tag{9.5}$$

We choose

$$M = x^{\theta-16\eta}, \quad N = x^{\varrho-\theta},$$

so that (9.3) holds by virtue of (9.5); and with this choice of M and N the expression on the right of (9.4) is $\ll x^{\theta-\eta}$. Our choice of M and N has also to satisfy (7.3), if Theorem A is to be invoked: therefore we require that

$$y = x^{\varrho-16\eta} \tag{9.6}$$

and that

$$U < \theta/\varrho. \tag{9.7}$$

To summarize,

LEMMA 8. *If*

$$\varrho = ((3+2\varkappa-\lambda)\theta-\varkappa)/2$$

and (9.5), (9.6) *and* (9.7) *are satisfied, then there exist numbers* M *and* N *satisfying* (7.3) *such that*

$$\sum_{\substack{m<M \\ mn|P(y^U)}} \sum_{n<N} a_m b_n R_{mn} \ll x^{\theta-\eta}.$$

We now apply Theorem A with a view to an application of Theorem B. Taking $U = T$, we deduce from (7.2) and Lemma 8 that

$$H(\mathscr{A}, y^V, y^T)$$
$$\geqslant \frac{x^\theta}{\log y}\frac{2}{T-E}(1+o(1))\{\Phi(T)-(\log\tfrac{4}{3}-\alpha(V))-E\log 3-E_0\beta(V))\}$$

where y is given by (9.6) and (9.5), (9.7) hold, the latter with $U \doteq T$; here

$$\Phi(T) = T\log\frac{1}{T}+(1-T)\log\frac{1}{1-T}.$$

It follows that if T, V, E and E_0 satisfy (8.2) and also

$$\Phi(T) > \log\tfrac{4}{3}-\alpha(V)+E\log 3+E_0\beta(V),$$

then

$$H(\mathscr{A}, y^V, y^T) \gg x^\theta/\log x$$

and, in particular, $H(\mathscr{A}, y^V, y^T) > 0$. Therefore Theorem B yields

THEOREM C. *Suppose that* T, V, E *and* $E_0 = \max(E, \frac{1}{3}(1-T))$ *are numerical constants satisfying*

$$E_0 \leqslant V \leqslant 1/4, \quad 1/2 \leqslant T < 1, \quad V \geqslant V_0, \tag{9.8}$$

where V_0 *is given in* (6.16), *and*

$$T \log \frac{1}{T} + (1-T) \log \frac{1}{1-T} > \log \tfrac{4}{3} - \alpha(V) + E \log 3 + E_0 \beta(V) \tag{9.9}$$

where $\alpha(V)$, $\beta(V)$ *are defined in* (6.18) (*in association with* (6.15), (6.13) *and* (6.12)).

For any exponent pair $(\varkappa, \lambda)$ *define*

$$\varrho = \varrho(\theta) = ((3+2\varkappa-\lambda)\theta - \varkappa)/2, \quad 0 < \theta < 1,$$

and require of it that it satisfy

$$\theta < \varrho(\theta) \leqslant \tfrac{1}{2}(1+\theta), \quad T < \theta/\varrho(\theta). \tag{9.10}$$

Suppose r *is a positive integer such that*

$$1/\varrho(\theta) < rT + E. \tag{9.11}$$

Then if $x \geqslant x_r$, *the interval* $(x-x^\theta, x]$ *contains* $\gg x^\theta/\log x$ *almost-primes* P_r *of order* r.

Proof. We have only to check that the conclusion of Theorem C, which is slightly stronger than that of Theorem B, is justified; and, following the discussion towards the end of Section 8, it suffices to check that the sum (8.4) is small compared with $x^\theta/\log x$. But this sum is at most

$$\sum_{\substack{m,p \\ x-x^\theta < mp^2 \leqslant x \\ y^V \leqslant p < y^T}} 1 \leqslant \sum_{y^V \leqslant p < y^T} \left(\frac{x^\theta}{p^2} + 1\right) \ll x^\theta y^{-V} + y^T \ll x^{\theta-\eta}$$

since, by (9.8), $V \geqslant V_0$ and, by (9.6) and (9.10), $y^T = x^{\varrho T - 16\eta T} < x^{\theta - 8\eta}$.

We illustrate the quality of Theorem C with some specific results.

(i) $r = 2$: Take $E_0 = E$, $V = 0.1672$, $(\varkappa, \lambda)$ as in [5], equation (8.19). One can then show that

$$\theta = \theta_2 \leqslant 0.4545.$$

(ii) $r = 3$: Take $E_0 = E$, $V = V_0$, and (³), $(\varkappa, \lambda) = AC_2C_1C_2(\frac{1}{2}, \frac{1}{2})$. Then

$$\theta = \theta_3 < 0.3257.$$

(iii) $r = 4$: Take $E_0 = (1-T)/3$, $V = V_0$, $(\varkappa, \lambda) = A^2C_2(\frac{1}{2}, \frac{1}{2})$. Then

$$\theta = \theta_4 < 0.2496.$$

(³) $C_n := BA^n$ and A, B are the usual Weyl and Van der Corput steps respectively.

(iv) $r = 5$: Take $E_0 = (1-T)/3$, $V = V_0$, $(\varkappa, \lambda) = A^3 C_2(\frac{1}{2}, \frac{1}{2})$. Then

$$\theta = \theta_5 < 0.202.$$

The improvement (over [5]) in the case $r = 2$ is not impressive. However, it so happens that the Fourier analysis of the remainder sum in [5] was subsequently improved by Iwaniec and Laborde [11]. It is implicit in their argument that the conclusion of Lemma 8 above holds with ϱ replaced by the superior

$$\varrho' = \big((6+5\lambda-3\varkappa)\theta + \varkappa - 3\lambda\big)/4.$$

When ϱ in Theorem C is replaced by ϱ', and $(\varkappa, \lambda) = C_2 C_1^2 C_3(\frac{1}{2}, \frac{1}{2})$ then Theorem C (with $V = 0.1742$) yields

$$\theta_2 \leqslant 0.4523.$$

As we remarked at the end of Section 8, the choice $U = T$ is, sometimes at least, not the best, and for the case $r = 2$ of Theorem C (probably also for $r = 3$) $U < T$ is superior. In this case we use the inequality (see [7], (3.7))

$$H(\mathscr{A}, y^V, y^T) \geqslant H(\mathscr{A}, y^V, y^U) - \sum_{y^U \leqslant p < y^T} (1-w(p)) S(\mathscr{A}_p, y^V);$$

Theorem A is then applied to $H(\mathscr{A}, y^V, y^U)$ and the method of Iwaniec–Laborde [11] to the second expression on the right. In this way

$$\theta_2 \leqslant 0.4476$$

can be reached, which is a little better than the result of [11]. These and other applications will be discussed elsewhere. To make our objectives clear, we remark in conclusion that we search above all for superior sieve methods, and list applications such as those described earlier merely to test the quality of these methods. It is clear that the estimations of the remainder sum in various applications pose problems of independent interest and of great importance.

References

[1] J. Friedlander and H. Iwaniec, *On Bombieri's asymptotic sieve*, Annali Pisa (4) 4 (1978), 719–756.

[2] G. Greaves, *A weighted sieve of Brun's type*, Acta Arith. 40 (1982), 297–332.

[3] —, *An algorithm for the Hausdorff moment problem*, Numer. Math. 39 (1982), 231–238.

[4] F. Grupp and H.-E. Richert, *The functions of the linear sieve*, J. Number Theory, to appear.

[5] H. Halberstam and D. R. Heath-Brown and H.-E. Richert, *Almost-primes in short intervals*, in: *Recent progress in Analytic Number Theory*, Vol. I, Academic Press, London 1981, 69–101.

[6] H. Halberstam and H.-E. Richert, *Sieve methods*, London 1974.

[7] —, —, *Weighted sieves*, Publications Mathématiques d'Orsay. 83.04. Colloque Hubert Delanqe 7&8 Juin 1982, 97–114.

[8] H. Halberstam and H.-E. Richert, *A weighted sieve of Greaves' type I*, these Proceedings, pp. 155–182.
[9] H. Iwaniec, *Rosser's sieve*, Acta Arith. 36 (1980), 171–202.
[10] —, *A new form of the error term in the linear sieve*, ibid. 37 (1980), 307–320.
[11] H. Iwaniec and M. Laborde, P_2 *in short intervals*, Ann. Inst. Fourier (Grenoble), 31 (1981), 37–56.
[12] Y. Motohashi, *Lectures on sieve methods and prime number theory*, Tata Institute of Fundamental Research, Bombay 1981.

Presented to the Semester
Elementary and Analytic Theory of Numbers
September 1–November 13, 1982