Thus,

$$\sum_{\substack{m_1, m_2 \leqslant y \\ m_1, m_2 \in M_k}} \frac{(m_1, m_2)}{m_1 m_2} \leqslant \sum_{\substack{m_1 \leqslant y \\ m_1 \in M_k}} \frac{1}{m_1} \sum_{d \mid m_1} d \sum_{\substack{m_2 \leqslant y \\ d \mid m_2 \\ m_2 \in M_k}} \frac{1}{m_2}. \tag{4.3}$$

If we majorize the inner sum trivially by $(\log y)/d$, we obtain

$$\sum_{\substack{m_1, m_2 \leqslant y \\ m_1, m_2 \in M_k}} \frac{(m_1, m_2)}{m_1 m_2} \leqslant \log y \sum_{\substack{m \leqslant y \\ m \in M_k}} \frac{d(m)}{m} \leqslant (\log y)^{1+\log 2/(1-\beta)} \sum_{\substack{m \leqslant y \\ m \in M_k}} \frac{1}{m},$$

using (4.2). But from (1.1) this last sum is $(\log y)^{o(1)}$. Thus if we choose β so that

$$\frac{1}{\beta} = 1 + \frac{\log 2}{1-\beta}$$

we have (4.1). This leads to the value

$$\beta = \tfrac{1}{2}\left(2 + \log 2 - \sqrt{4 \log 2 + \log^2 2}\right)$$

and establishes (1.3) for every $A < 0.617122930\ldots$.

References

[1] N. G. de Bruijn, *On the number of positive integers $\leqslant x$ and free of prime factors $> y$*, Nederl. Akad. Wetensch. Proc. Ser. A 54 (1951), pp. 50–60.

[2] P. Erdős, *On the normal number of prime factors of $p-1$ and some related questions concerning Euler's φ-function*, Quart. J. Math. (Oxford Ser.) 6 (1935), pp. 205–213.

[3] P. Erdős and R. R. Hall, *On the values of Euler's φ-function*, Acta Arith. 22 (1973), pp. 201–206.

[4] — — *Distinct values of Euler's φ-function*, Mathematika 23 (1976), pp. 1–3.

[5] G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number n*, Quart. J. Math. 48 (1917), pp. 76–92.

[6] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Fourth Ed., Oxford, London 1965.

[7] J. L. Nicolas, *Sur la distribution des nombres entiers ayant une quantité fixée de facteurs premiers*, Acta Arith. 44 (1984), pp. 191–200.

[8] I. Niven and H. S. Zuckerman, *An Introduction to the Theory of Numbers*, John Wiley and Sons, New York 1960.

BELL COMMUNICATIONS RESEARCH, INC.
MORRISTOWN, NEW JERSEY 07960, USA

CURRENT ADDRESS
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GEORGIA
ATHENS, GEORGIA 30602, USA

Received on 10.12.1984
and in revised form on 11.2.1985 (1477)

ACTA ARITHMETICA
XLVII (1986)

The weighted linear sieve and Selberg's λ^2-method

by

G. Greaves (Cardiff)

1. Introduction. In the weighted linear sieve we study sequences $\mathscr{A} = \mathscr{A}_X$, depending on a real parameter $X \geqslant 2$, which satisfy certain general conditions of the type described by Halberstam and Richert [5]. The conditions specified in this paper are labelled (Ω_1), (Ω_2), (R), (D) below. As usual these are chosen with due regard to applicability on one hand (cf. the examples provided in [5], for example), and on the other hand to the requirements of a workable proof of a result of the type established in this paper. The object of the exercise is to deduce, for a suitably small integer $R \geqslant 2$, that the sequence $\mathscr{A}$ contains many numbers having no more than R prime factors.

In [2] the present author obtained an improvement on the results previously known on this problem via a study of expressions of the type

$$\sum_{d \mid a} \mu(d) \chi_y^-(d) \{W(1) - \sum_{p \mid d} w(p)\}, \tag{1.1}$$

where $w(p)$ was a 'weight' function of the type appearing in earlier approaches to this problem (see Chapter 9 of [5], for example) and $\chi_y^-(d)$ was the 'characteristic' function appearing in Iwaniec's and Rosser's version (see [7], [9]) of Brun's sieve. Thus in the case $w(p) = 0$ the expression (1.1) reduces to the corresponding expression studied in [7], [9]. In this "unweighted" context an essentially equivalent result had been obtained by Jurkat and Richert [10] (see also Chapter 8 of [5]), using a method in which the well-known λ^2-device of Selberg played a significant rôle. In this paper we replace the expression $\chi_y^-(d)$ in (1.1) by the expression implicit in the paper of Jurkat and Richert. We shall see that in the problem of the weighted linear sieve the expressions are not equivalent, in that we shall obtain an improvement, in certain cases, upon the result in the author's earlier paper [2]. At the same time the result obtained falls short of that which seems to be generally conjectured to be true, and which would be best possible.

The properties we require of the set $\mathscr{A} = \mathscr{A}_X$ are as follows. We write

$$\sum_{\substack{a \in \mathscr{A} \\ a \equiv 0 \bmod d}} 1 = \frac{X}{d}\varrho(d) + R(\mathscr{A}, d). \tag{1.2}$$

The function ϱ will be assumed to be multiplicative and to satisfy an assumption of the type usual in the theory of the linear sieve:

$$0 \leqslant \frac{\varrho(p)}{p} \leqslant 1 - \frac{1}{A_1}, \tag{Ω_1}$$

$$-L < \sum_{w \leqslant p < z} \frac{\varrho(p)\log p}{p} - \log\frac{z}{w} \leqslant A_2 \quad \text{if} \quad 2 \leqslant w \leqslant z. \tag{Ω_2}$$

These conditions imply (cf. § 3 of Ch. 2 in [5]) the form of (Ω_2) used by some authors, in which the denominator of the summand is replaced by $p - \varrho(p)$. As usual, we shall keep explicit track of the dependence of our estimates on the parameter L, which will be supposed to satisfy

$$1 \leqslant L \leqslant \log y. \tag{1.3}$$

We introduce a "level of distribution" y, in terms of which our results will be stated. Theorem 1 below is independent of any hypothesis concerning the "remainder" term $R(\mathscr{A}, d)$ in (1.2). For applications, however, some knowledge of $R(\mathscr{A}, d)$ would of course be required. We would require

(R) *for the expression $f(d)$ appearing in Theorem* 1, *the number* $y = y(X)$ *satisfies*

$$\Big|\sum_{d \leqslant y} \mu(d) f(d) R(\mathscr{A}, d)\Big| < A_3 X/\log^2 X.$$

We will see that $f(d)$ satisfies

$$|f(d)| \leqslant 3^{\nu(d)},$$

where, as elsewhere in this paper,

$$\nu(d) = \sum_{p \mid d} 1$$

is the number of distinct prime factors of d.

The "degree" g will satisfy the hypothesis

(D) $\qquad 1 \leqslant a < y^g \quad$ *when* $\quad a \in \mathscr{A}$.

Then, as in [2], we aim to obtain results of the type: *assume* (Ω_1), (Ω_2), (R), (D). *Then, if* $g \leqslant R - \delta_R$, *there exists* a *in* $\mathscr{A}$ *with at most* R *prime factors.*

As usual, if repeated prime factors of a are to be counted once only in such a result, then an extra property of $\mathscr{A}$, such as

there exists a constant c such that

$$\sum_{w \leqslant p < z} \log p \sum_{\substack{a \in \mathscr{A} \\ a \equiv 0 \bmod p^2}} 1 \leqslant A_4 X/w \quad \textit{if} \quad z \leqslant X/(\log X)^c,$$

would be required.

It remains the case that results of the desired type only follow from Theorem 1 for certain positive values of δ_R. Numerical work indicates that values of δ_R better (i.e. smaller) than those obtained in [2] follow only for certain small values of R, which may, however, be regarded as being the most interesting. The author's calculations indicate that the following values are accessible:

$$\delta_2 = 0.044560, \quad \delta_3 = 0.074267, \quad \delta_4 = 0.103974. \tag{1.4}$$

Theorem 1 involves a "weight" function w which has to satisfy certain properties. Following Richert [12] (but with some change of notation) we use parameters U, V satisfying

$$V < 1/4, \quad 1/2 < U < 1, \quad V + RU \geqslant g. \tag{1.5}$$

The notation $U = 1/u$ would accord with [12]; our V may, however, be positive, zero or negative. The function w will be written as

$$w(p) = W(\log p/\log y), \tag{1.6}$$

where p is prime and $p \leqslant y$. Define

$$W(1) = U - V.$$

Write

$$m = \max\{V, (1-U)/2\}. \tag{1.7}$$

We will require

$$0 \leqslant W(t) \leqslant \begin{cases} W(1) & \text{if} \quad U < t \leqslant 1, \\ t - V & \text{if} \quad 1/3 \leqslant t \leqslant U, \\ t - m & \text{if} \quad 0 \leqslant t \leqslant 1/3 \text{ and } t > m, \\ 9(U - \frac{1}{3})t^2 & \text{if} \quad 0 \leqslant t. \end{cases} \tag{1.8}$$

As in [2] we notate

$$\nu_{y,U}(a) = \sum_{p \mid a;\, p < y^U} 1 + \sum_{\substack{p,\alpha \\ p^\alpha \mid a;\, p \geqslant y^U}} 1$$

for the number of prime factors of a, where multiple prime factors p of a are counted multiply only if $p \geqslant y^U$.

Of the four lines appearing on the right of (1.8), the first two appear for reasons already featuring in [12], the third for considerations of a type appearing in [2], and the last because of requirements imposed by our use of the λ^2-method.

Our Theorem 1 involves a function h somewhat smaller than the function playing a corresponding rôle in [2]. For integers $i \geqslant 1$ define

$$h_i(x, s) = \int_{R_i} \cdots \int \frac{dx_1 \ldots dx_i}{x_1 \ldots x_i u},$$

where the region R_i is given by the conditions:

$$x < x_i < \ldots < x_1 < 1/s,$$
$$3x_j + x_{j-1} + \ldots + x_1 \leqslant 1 \quad \text{if} \quad 1 \leqslant j < i,$$
$$1 < 3x_i + x_{i-1} + \ldots + x_1, \quad x_i < u,$$
$$x + x_i + \ldots + x_1 + u = 1.$$

Thus, as in [2],

$$h_1(x, s) = \int\limits_{\substack{x < x_1 < 1/s \\ 1/3 < x_1 < u, x + x_1 + u = 1}} \frac{dx_1}{x_1 u},$$

but for $i > 1$ we now have

$$h_i(x, s) = \int\limits_{s < t; 3 < t} h_{i-1}\left(\frac{tx}{t-1}, t-1\right) \frac{dt}{t-1}$$

(the condition $3 < t$ applying regardless of the parity of i). Now define

$$h(x, s) = \sum_{k \geqslant 1} h_{2k}(x, s) \quad \text{when} \quad 0 < x \leqslant 1 \leqslant s,$$

the summation being finite.

If we specify

$$H(x, s) = \sum_{k \geqslant 1} h_{2k-1}(x, s) \quad \text{when} \quad 0 < x \leqslant 1 \leqslant s$$

and

$$H(x, s) = h(x, s) = 0 \quad \text{when} \quad x > 1,$$

then these functions are characterized by the integral equations

$$h(x, s) = \int\limits_{s < t; 3 < t} H\left(\frac{xt}{t-1}, t-1\right) \frac{dt}{t-1},$$

$$H(x, s) = h_1(x, s) + \int\limits_{s < t; 3 < t} h\left(\frac{xt}{t-1}, t-1\right) \frac{dt}{t-1}.$$

The statement of our principal theorem is now, verbally although not in content, similar to Theorem 1 in [2]. In the theorem, and elsewhere, $p(a)$ denotes the least prime factor of a.

THEOREM 1. *Assume* (Ω_1), (Ω_2), (D). *Then there exists* $f(d)$ *with* $0 \leqslant f(d) \leqslant 3^{\nu(d)}$ *(to be described) such that*

$$\sum_{\substack{a \in \mathscr{A} \\ \nu_{y,U}(a) \leqslant R}} w(p(a))$$
$$\geqslant 2e^{\gamma} X \prod_{p < y} \left(1 - \frac{\varrho(p)}{p}\right) \left\{ \mathscr{M}(W) + O\left(\frac{L^{1/5}}{\log^{1/5} y}\right) \right\} - \Big| \sum_{d \leqslant y} \mu(d) f(d) R(\mathscr{A}, d) \Big|,$$

where $\nu_{y,U}(a)$, w, W *are as described above,* γ *is Euler's constant, and*

$$\mathscr{M}(W) = -\int_{1/2}^{1} \frac{W(1) - W(t)}{1-t} \frac{dt}{t} + \int_{0}^{1/2} W(t) \left\{ \frac{1}{1-t} - h(t, 1) \right\} \frac{dt}{t}.$$

The O*-constant may depend on* U *as well as on* A_1 *and* A_2.

The proof of Theorem 1 is an assembly of the results of Lemmas 1, 5 and 6.

In making an application of Theorem 1 it will be necessary to estimate an expression

$$\int_0^{1/4} W(t) h(t, 1) \, dt/t$$

(the function $h(t, 1)$ vanishing when $t > 1/4$). We postpone our account of these estimations until Section 5.

The author's approach to the questions involved in handling the function h is by passing to the moments

$$h_n(s) = \int_0^1 x^n h(x, s) \, dx, \qquad H_n(s) = \int_0^1 x^n H(x, s) \, dx. \tag{1.9}$$

As in [2], these can be related to the solutions of the differential-difference problems

$$\begin{aligned} &\frac{d}{ds}\{sJ(s)\} + J(s+1) = 0 \quad (\mathrm{Re}(s) > 0), \qquad sJ(s) \to 1 \quad \text{as} \quad s \to 0, \\ &\frac{d}{ds}\{s^2 I'(s)\} - sI'(s+1) = 0 \quad (\mathrm{Re}(s) > 0), \qquad s^2 I'(s) \to -1 \quad \text{as} \quad s \to 0, \end{aligned} \tag{1.10}$$

where the function I' is not asserted to be the derivative of a function I. We write

$$j_n = (-1)^n J^{(n)}(1)/n!, \qquad i_n = (-1)^n I^{(n)}(1)/n!,$$

so that i_n, j_n are representable by integrals as in [2]. The moments (1.9) are then determined as follows.

THEOREM 2. *Let i_n, j_n be as above and let h_n, H_n be defined as in* (1.9). *Define*

$$B_n = 2^{n+1}(n+1)(j_n - j_{n+1}), \qquad D_n = 2^{n+1}(n+1)(i_n - i_{n+1}), \qquad c_n = \int_0^{1/2} \frac{x^n}{1-x}\,dx.$$

Then

$$j_n + B_n\big(h_n(3) - c_n - H_n\big) = 1 \quad \text{if} \quad n \geqslant 0,$$
$$i_n + D_n\big(h_n(3) - c_n + H_n\big) = 1 \quad \text{if} \quad n \geqslant 1,$$
$$2\big(h_0(3) - c_0 + H_0\big) + 1 = 0.$$

The proof of Theorem 2 is by the technique used in [2], and is therefore omitted. A rather simpler application of similar ideas leads to the following fact about the function $h(x, s)$ itself. Here, we should define

$$h(0, s) = \lim_{x \to 0+} h(x, s).$$

THEOREM 3. *At the point $x = 0$ the quantity $h(x, s)$ satisfies*

$$h(0, s) = 1 - 1/2J(1) \quad \text{if} \quad s < 3,$$

where $J(s)$ is as in (1.10).

This result, though not without interest, appears to have little application to our theme other than providing a check on any numerical method purporting to evaluate $h(x, s)$ for small x and s.

The constant $J(1) = 0.62432\,99885\ldots$ has been computed by a variety of methods and authors: see [1], [3], [11], for example.

We adopt several notations which are becoming standard in this subject, some of which have been mentioned already and others of which will be defined as they appear. In particular we denote $P(z) = \prod_{p<z} p$ for the product of primes less than z. The least and greatest prime factors of an integer $n > 1$ are denoted by $p(n)$, $q(n)$ respectively: our convention is $q(1) = 1$, $p(1) = \infty$. We set

(1.11) $$V(z) = \prod_{p<z} \left\{1 - \frac{\varrho(p)}{p}\right\}.$$

2. Basic identities and inequalities. For squarefree d we adopt a standard notation

(2.1) $$d = p_1 p_2 \ldots p_\nu; \qquad p_1 > p_2 > \ldots > p_\nu,$$

p_i being prime. Let

$$B_i = B_i(p_{i-1}, \ldots, p_1)$$

be, for the moment, arbitrary. Then we define a "characteristic" function $\chi_{\mathscr{B}}$ by

(2.2) $$\chi_{\mathscr{B}}(d) = 1 \quad \text{if} \quad p_i < B_i \text{ when } 1 \leqslant i \leqslant \nu,$$

so that in particular $\chi_{\mathscr{B}}(1) = 1$. Following a notation used by Halberstam and Richert [6] define an associated function $\bar{\chi}_{\mathscr{B}}$ by

(2.3) $$\bar{\chi}_{\mathscr{B}}(1) = 0; \quad \bar{\chi}_{\mathscr{B}}(d) = \chi\big(d/p(d)\big) - \chi(d) \quad \text{when} \quad d > 1.$$

Here $p(d) = p_\nu$, the smallest prime factor of d. Observe that

(2.4) $$\bar{\chi}_{\mathscr{B}}(d) = \begin{cases} 1 & \text{if } B_\nu \leqslant p_\nu \text{ and } p_i < B_i \text{ when } 1 \leqslant i \leqslant \nu - 1, \\ 0 & \text{otherwise.} \end{cases}$$

The 'Fundamental Identity' of combinatorial sieve methods may now be written, for squarefree A, as

(2.5) $$\sum_{d|A} \mu(d)\varphi(d) = \sum_{d|A} \mu(d)\chi_{\mathscr{B}}(d)\varphi(d) + \sum_{\delta|A} \mu(d)\bar{\chi}_{\mathscr{B}}(\delta) \sum_{t|A;\, q(t) < p(\delta)} \mu(t)\varphi(\delta t),$$

where $q(t)$ denotes the greatest prime factor of t, and φ is an arbitrary arithmetic function.

The identity (2.5) is almost obvious. The divisors d of A either satisfy $p_i < B_i$ for all i, or else there is a least j for which $p_j \geqslant B_j$, in which case we write $d = \delta t$ with

$$\delta = p_1 p_2 \ldots p_j; \qquad t = p_{j+1} \ldots p_\nu.$$

Now (2.5) follows. Alternatively, see the discussion in [6].

In this paper, as in [2], [6], the choice of the function φ in the identity (2.5) will be of the type

(2.6) $$\varphi(d) = W(1) - \sum_{p|d} w(p).$$

In the unweighted sieve, where $\varphi(d) = 1$, the quantity in (2.5) is

(2.7) $$\delta\{A\} = \sum_{d|A} \mu(d) = \begin{cases} 1 & \text{if} \quad A = 1, \\ 0 & \text{if} \quad A > 1. \end{cases}$$

In the case when $A = (a, P(z))$ we shall also use the abbreviation

$$S(a, z) = \delta\{(a, P(z))\}, \tag{2.8}$$

which is 1 if $a = 1$ or if a has no prime factors less than z, and is 0 for other values of a. In the weighted sieve, the analogous quantities are, following a notation of [2],

$$\Lambda\{A, W(1), w\} = \sum_{d|A} \mu(d)\varphi(d) = \begin{cases} W(1) & \text{if } A = 1, \\ w(p) & \text{if } A = p, \\ 0 & \text{otherwise,} \end{cases} \tag{2.9}$$

the last identity following by the characteristic property of the Möbius function μ. Note the equation

$$\Lambda\{(a, P(z)), W(1), w\} = W(1)S(a, z) + \sum_{p|(a,P(z))} w(p)S(a/p, z), \tag{2.10}$$

which follows directly using (2.6), (2.7), (2.8).

The familiar "Buchstab" identity

$$S(a, z) = 1 - \sum_{p|(a,P(z))} S(a/p, p), \tag{2.11}$$

obtainable by writing $d = pd'$, where $q(d') < p$, whenever $d > 1$ in (2.7), is a special case of the "fundamental" (2.5), which may in turn be considered as following from multiple applications of (2.11).

We shall also need to refer to the identity

$$\sum_{t|A/\delta} \mu(t)\varphi(t\delta) = \Lambda\{A/\delta, W(1) - \sum_{p|\delta} w(p), w\} \tag{2.12}$$

when $\delta | A$; this is an immediate extension of (2.9).

Our initial construction is closely related to that of [2], but we give an essentially self-contained treatment here. Using the Buchstab identity (2.11) we obtain, for the quantity in (2.10), the identity

$$\Lambda\{(a, P(z)), W(1), w\}$$
$$= W(1) - \sum_{p|a;p<z} \{W(1) - w(p)\} S\left(\frac{a}{p}, p\right) - \sum_{pq|a;p<q<z} w(p) S\left(\frac{a}{pq}, q\right),$$

wherein q as well as p denotes a prime. Define

$$\Lambda_2\{(a, P(z)), W(1), w\} = \sum_{\substack{p_1p_2|a \\ p_2<p_1<z;y\leqslant p_2^2p_1}} \left[\{W(1) - w(p_1)\} S\left(\frac{a}{p_1p_2}, p_2\right) - w(p_2) S\left(\frac{a}{p_1p_2}, p_1\right)\right], \tag{2.13}$$

the departure from [2] lying in the use of the exponent 2 in the conditions of summation; this will be related to the quadratic nature of the λ^2-method employed in this paper. Since

$$S\left(\frac{a}{p}, p\right) = S\left(\frac{a}{p}, 1\right) - \sum_{q|a/p;q<p} S\left(\frac{a}{pq}, q\right),$$

the quantity

$$\Sigma(a, z) = \Lambda\{(a, P(z)), W(1), w\} - \Lambda_2\{(a, P(z)), W(1), w\} \tag{2.14}$$

satisfies the relation

$$\Sigma(a, z) = \Sigma_1 + \Sigma_2, \tag{2.15}$$

where

$$\Sigma_1 = W(1) - \sum_{p|a;p<z} \{W(1) - w(p)\},$$
$$\Sigma_2 = \sum_{\substack{p_1p_2|a \\ p_2<p_1<z;p_2^2p_1<y}} \left[\{W(1) - w(p_1)\} S\left(\frac{a}{p_1p_2}, p_2\right) - w(p_2) S\left(\frac{a}{p_1p_2}, p_1\right)\right]. \tag{2.16}$$

The basic inequality of the method is as follows.

LEMMA 1. *Suppose w satisfies the conditions* (1.8) *and suppose $z \geqslant y^U$. Then if $\Sigma(a, z) > 0$ we have*

$$\nu_{y,u}(a) \leqslant R; \quad \Sigma(a, z) \leqslant w(p(a)).$$

Let $p(a) = Q_1(a) < Q_2(a)$ denote the two smallest prime factors of a. If $Q_1^2(a)Q_2(a) \geqslant y$ then the sum Σ_2 is empty (because no two prime factors of a can satisfy $p_2^2 p_1 < y$). The result of Lemma 1 is then inherited via (2.15) from the properties of Richert's expression (2.16); this point was discussed *ab initio* in [2] in spite of having been fully covered in [12], [5]. If $Q_1^2(a)Q_2(a) < y$ then the terms $S(a/p_1p_2, p_1)$ in (2.13) are zero (because otherwise we should have $p_2 = Q_1(a)$, $p_1 = Q_2(a)$, contrary to the conditions of summation $y \leqslant p_2^2 p_1$). Because of (2.14) we obtain

$$\Sigma(a, z) \leqslant \Lambda\{(a, P(z)); W(1)\},$$

and the conclusion of Lemma 1 follows because of the definition (2.9) of Λ and the fact that, from (1.8), (1.7), (1.6), $w(p) > 0$ implies $p > y^V$, so if $\nu_{y,u}(a) > R$ we should have from (1.5) that

$$a \geqslant y^{V+RU} \geqslant y^{\theta},$$

contrary to (D).

Combinatorial sieves rest on an application of the "fundamental" identity (2.5) in the case when

$$A = (a, P(z)).$$

In [2] the present author followed Rosser and Iwaniec in specifying, for squarefree $d > 1$,

$$\chi_y^-(d) = 1 \Leftrightarrow hp^2(h) < y \quad \text{when} \quad h|d \text{ and } 2|\nu(h). \tag{2.17}$$

This may appear more familiar in the form

$$\chi_y^-(d) = 1 \Leftrightarrow p_{2i}^3 p_{2i-1} \ldots p_1 < y \quad \text{when} \quad 1 \leqslant i \leqslant \nu, \tag{2.18}$$

where d is expressed as in (2.1). In this paper we use ideas from the paper [10] of Jurkat and Richert in specifying χ as follows: define $\chi(1) = 1$, and for squarefree $d > 1$ say $\chi(d) = 1$ if the two conditions

$$\text{(a)} \quad h|d,\ 2|\nu(h) \Rightarrow hp^2(h) < y, \tag{2.19}$$

$$\text{(b)} \quad h|d,\ \nu(h) \leqslant 1/m \Rightarrow hp^2(h) < y \tag{2.20}$$

both hold. In all other cases define $\chi(d) = 0$. Thus the numbers B_i of (2.2) are given by

$$p_i < B_i \Leftrightarrow p_i^3 p_{i-1} \ldots p_1 < y \quad \text{if} \quad i \leqslant 1/m \text{ or if } 2|i, \tag{2.21}$$

$$B_i = +\infty \quad \text{if} \quad i > 1/m \text{ and } 2 \nmid i. \tag{2.22}$$

Here the parameter $m > 0$ is that of (1.7).

The next lemma sets up a decomposition of $\Sigma(a, z)$ appropriate to our approach.

LEMMA 2. *The quantity $\Sigma(a, z)$ defined in (2.14) satisfies the identity*

$$\Sigma(a, z) = \Sigma_1(a, z) + \Sigma_2(a, z) + \Sigma_3(a, z) + \Sigma_4(a, z), \tag{2.23}$$

where

$$\Sigma_1(a, z) = \sum_{d|(a,P(z))} \mu(d)\chi_y(d)\varphi(d), \tag{2.24}$$

$$\Sigma_2(a, z) = \sum_{p|(a,P(z))} \mu(p)\bar{\chi}_y(p)\left[\{W(1) - w(p)\} S\left(\left(\frac{a}{p}\right), \left(\frac{y}{p}\right)^{1/2}\right) + \sum_{\substack{q|a/p \\ q<p;\, q^2p<y}} w(q) S\left(\frac{a}{pq}, p\right)\right], \tag{2.25}$$

$$\Sigma_3(a, z) = \sum_{\substack{\delta|(a,P(z)) \\ 2\nmid\nu(\delta);\, 3\leqslant\nu(\delta)\leqslant 1/m}} \mu(\delta)\bar{\chi}_y(\delta) \sum_{t|(a,P(p(\delta)))} \mu(t)\varphi(t\delta), \tag{2.26}$$

$$\Sigma_4(a, z) = \sum_{\substack{\delta|(a,P(z)) \\ 2|\nu(\delta);\, \nu(\delta)\geqslant 2}} \mu(\delta)\bar{\chi}_y(\delta) \sum_{t|(a,P(p(\delta)))} \mu(t)\varphi(t\delta). \tag{2.27}$$

The definition (2.14) of $\Sigma(a, z)$ and the fundamental identity (2.5) for the quantity Λ defined in (2.9) give

$$\Sigma(a, z) = \Sigma_1(a, z) + \sum_{\delta|(a,P(z))} \mu(\delta)\bar{\chi}_y(\delta) \sum_{t|(a,P(p(\delta)))} \mu(t)\varphi(t\delta) - \Lambda_2\{(a, P(t)), W(z), w\}. \tag{2.28}$$

Here, the terms with $\nu(\delta) = 1$ sum to $-S_1$, say, where, because of (2.4), (2.18), (2.12), (2.10), we have

$$S_1 = \sum_{\substack{p_1|(a,P(z)) \\ y\leqslant p_1^3}} \left[\{W(1) - w(p_1)\} S\left(\frac{a}{p_1}, p_1\right) + \sum_{\substack{p_2|a/p_1 \\ p_2<p_1}} w(p_2) S\left(\frac{a}{p_1p_2}, p_1\right)\right].$$

In the Lemma, the term $\Sigma_2(a, z)$ is provided by the identity

$$-\Sigma_2(a, z) = \Lambda_2\{(a, P(z)), W(1), w\} + S_1, \tag{2.29}$$

where Λ_2 is as in (2.13). The conditions of summation therein imply $y \leqslant p_1^3$, so (2.29) follows from an application

$$S\left(\frac{a}{p}, \left(\frac{y}{p}\right)^{1/2}\right) = S\left(\frac{a}{p}, p\right) + \sum_{\sqrt{y/p}\leqslant q<p} S\left(\frac{a}{pq}, q\right)$$

of the Buchstab identity, valid when $y < p^3$.

The terms in (2.28) with $\nu(\delta) > 1$ provide the terms $\Sigma_3(a, z)$, $\Sigma_4(a, z)$ in Lemma 2, whose proof is now complete.

3. Application of the λ^2-method. To obtain a lower bound for

$$\sum_{a\in\mathcal{A}} \Sigma(a, z)$$

from Lemma 2, we shall derive, from the λ^2-method, suitable bounds for the second and third sums in (2.23).

It will be convenient to denote (as usual)

$$\mathcal{A}_\delta = \{a/\delta:\ a\in\mathcal{A},\ a \equiv 0 \bmod \delta\}.$$

We shall need to refer to it when $\delta p^2(\delta) > y$. In this case we have, when $d \leqslant \sqrt{y/\delta}$,

$$\sum_{\substack{a_\delta\in\mathcal{A}_\delta \\ a_\delta\equiv 0 \bmod d}} 1 = (X\varrho(\delta)/\delta)\frac{\varrho(d)}{d} + R(\mathcal{A}, d\delta) \tag{3.1}$$

as a consequence of (1.2), since ϱ is multiplicative, and $(d, \delta) = 1$ in our case.

We shall refer to a multiplicative function ϱ_p, given by

$$\varrho_p(d) = \varrho(d) \quad \text{if} \quad (d, p) = 1, \qquad \varrho_p(p) = 0. \tag{3.2}$$

This satisfies

(3.3) $$-A_2-L<\sum_{w\leqslant q<z}\frac{\varrho_p(q)\log q}{q}-\log\frac{z}{w}\leqslant A_2 \quad \text{if} \quad 2\leqslant w\leqslant z,$$

since (cf. (2.3.8) in [5])

(3.4) $$-L\leqslant\frac{\varrho(p)\log p}{p}\leqslant A_2$$

follows from (Ω_2).

LEMMA 3. *Suppose that a "weight" $f(p)$ satisfies*

$$f(p)\leqslant f(1)=1 \quad \text{if} \quad p<y/\delta,$$
$$f(p)\leqslant 4\left(\frac{\log p}{\log(y/\delta)}\right)^2 \quad \text{if} \quad p<\sqrt{y/\delta},$$
$$f(p)=0 \quad \text{if} \quad p<y^{\varepsilon}$$

and suppose that $\delta<y^{1-\varepsilon}$, where $\varepsilon>0$. For brevity denote, when $\delta|a$,

$$A_\delta=(a/\delta,\ P(\sqrt{y/\delta})),$$

and suppose that

$$\sqrt{y/\delta}<p(\delta),$$

so that (3.1) *applies. Denote*

(3.5) $$A\{A_\delta,1,f\}=S(a/\delta,\sqrt{y/\delta})+\sum_{p|(a/\delta);\,p<\sqrt{(y/\delta)}}f(p)S(a/\delta p,\sqrt{y/\delta})$$

as in (2.10). *Then*

$$\sum_{\substack{a\in\mathscr{A}\\ a\equiv 0 \bmod \delta}}A\{A_\delta,1,f\}\leqslant\frac{X\varrho(\delta)}{\delta}e^{\gamma}V\left(\left(\frac{y}{\delta}\right)^{1/2}\right)\left\{1+O_\varepsilon\left(\frac{L}{\log y}\right)\right\}+E(\delta),$$

where $V(\cdot)$ is as in (1.11) *and*

$$E(\delta)=\sum_{d_1,d_2}\lambda_{y/\delta}(d_1)\lambda_{y/\delta}(d_2)R(\mathscr{A},[d_1,d_2]\delta)$$
$$=O\Big(\sum_{d\leqslant y/\delta}\mu^2(d)3^{\nu(d)}|R(\mathscr{A},d\delta)|\Big),$$

the symbol λ being as defined below. The constant implied by the symbol O_ε may depend on ε.

We use the non-negative expression

(3.6) $$\Big\{\sum_{d|(a/\delta)}\lambda_{y/\delta}(d)\Big\}^2\geqslant 0,$$

where, as usual in Selberg's method (cf. Chapter 3 of [5]),

(3.7) $$\lambda_x(d)=C_x\frac{\mu(d)d}{\varrho^*(d)}\sum_{\substack{n\leqslant\sqrt{x}/d\\ (n,d)=1}}\mu^2(n)g(n),$$

with

$$g(n)=\varrho(n)/\varrho^*(n),$$

ϱ^* being the multiplicative function "conjugate" to ϱ:

$$\varrho^*(n)=\prod_{p|n}\{p-\varrho(p)\},$$

and with the usual normalisation

(3.8) $$C_x^{-1}=\sum_{n\leqslant\sqrt{x}}\mu^2(n)g(n),$$

so that

$$\lambda(1)=1.$$

When $y^{\varepsilon}<p\leqslant\sqrt{x}$ we need to study the expression

$$\lambda_x(1)+\lambda_x(p)=C_x\sum_{n\leqslant\sqrt{x}}\mu^2(n)g(n)-\frac{C_xp}{\varrho^*(p)}\sum_{\substack{n\leqslant\sqrt{x}/p\\ (n,p)=1}}\mu^2(n)g(n)$$
$$=C_x\sum_{\substack{n\leqslant\sqrt{x}\\ (n,p)=1}}\mu^2(n)g(n)+C_x\frac{\varrho(p)-p}{\varrho^*(p)}\sum_{\substack{m\leqslant\sqrt{x}/p\\ (m,p)=1}}\mu^2(m)g(m)$$
$$=C_x\sum_{\substack{\sqrt{x}/p<m\leqslant\sqrt{x}\\ (m,p)=1}}\mu^2(m)g(m).$$

In this sum, the condition $(m,p)=1$ is expressible by replacing ϱ by ϱ_p, as defined in (3.2). Because of (3.3) we infer from the last equation in § 5.3 of [5] that

$$C_x^{-1}=\prod_q\left\{\frac{q-1}{q-\varrho(q)}\right\}\log(\sqrt{x})\left\{1+O\left(\frac{L}{\log x}\right)\right\}$$

and (we specified $L\geqslant 1$ in (1.3))

$$\sum_{\substack{\sqrt{x}/p<m\leqslant\sqrt{x}\\ (m,p)=1}}\mu^2(m)g(m)=\prod_q\left\{\frac{q-1}{q-\varrho_p(q)}\right\}\log p\left\{1+O\left(\frac{L}{\log p}\right)\right\},$$

whence

$$\lambda_x(1)+\lambda_x(p)=2\left\{1-\frac{\varrho(p)}{p}\right\}\frac{\log p}{\log x}\left\{1+O\left(\frac{L}{\log p}\right)\right\}=2\frac{\log p}{\log x}\left\{1+O\left(\frac{L}{\log p}\right)\right\},$$

where we have used (3.4).

We compare the expression $\Lambda\{A_\delta, 1, f\}$, as given in (3.5), with the expression (3.6). If $A_\delta = 1$ then the two expressions are equal (to 1), for the familiar reason that in this case the sume defining $\lambda_{y/\delta}(d)$ in (3.7) is empty when $d > 1$. If A_δ is a prime p then

$$\Lambda\{A_\delta, 1, f\} = f(p),$$

while the expression (3.6) is

$$\{\lambda_{y/\delta}(1)+\lambda_{y/\delta}(p)\}^2 = \frac{4\log^2 p}{\log^2(y/\delta)}\left\{1+O_\varepsilon\left(\frac{L}{\log y}\right)\right\}$$

in the case when $p > y^\varepsilon$. In all other cases the Λ expression on the left of (3.5) is zero, so in all cases we have

$$\Lambda\{A_\delta, 1, f\} \leqslant \{\sum_{d|(a/\delta)} \lambda_{y/\delta}(d)\}^2\left\{1+O_\varepsilon\left(\frac{1}{\log y}\right)\right\}.$$

It is well known (see § 5.3 of [5]) how (1.2), (Ω_1), (Ω_2) now lead via (3.1) and

$$\sum_{\substack{a\in\mathscr{A}\\ a\equiv 0 \bmod \delta}} \{\sum_{d|a} \lambda_x(d)\}^2 = \sum_d \{\sum_{[d_1,d_2]=d} \lambda_x(d_1)\lambda_x(d_2)\}\left\{\frac{X\varrho(d\delta)}{d\delta}+R(\mathscr{A}, d\delta)\right\}$$

to the result of Lemma 3.

Next, we verify that the conditions (1.8) specified for the function w bring the sums $\Sigma_2(a, z)$, $\Sigma_3(a, z)$ of Lemma 2 into the ambit of Lemma 3, and also ensure that $\Sigma_4(a, z)$ is non-negative.

LEMMA 4. *When w satisfies* (1.8) *we have*

(i) *if* $p_2^2 p_1 < y$, $p_2 < p_1$, $y \leqslant p_1^3$ *then*

$$w(p_2) \leqslant 4\frac{\log^2 p_2}{\log^2(y/p_1)}\{W(1)-w(p_1)\}, \tag{3.9}$$

(ii) *when* $2 \nmid \nu(\delta) \geqslant 3$ *and* $\bar{\chi}_y(\delta) = 1$, *in the sense described in* (2.4), (2.21), (2.22), *and when a prime* q *satisfies* $q < p(\delta)$, *we have*

$$w(q) \leqslant W(1) - \sum_{p|\delta} w(p), \tag{3.10}$$

$$w(q) \leqslant \frac{4\log^2 q\,\{W(1)-\sum_{p|\delta} w(p)\}}{\{\log y - \sum_{p|\delta}\log p\}^2} \quad \text{if} \quad q^2 < y/\delta, \tag{3.11}$$

(iii) *when* $2 | \nu(\delta) \geqslant 2$ *and* $\bar{\chi}_y(\delta) = 1$ *we have*

$$\sum_{p|\delta} w(p) \leqslant W(1).$$

Observe first that (i) holds in the case when $p_1 \geqslant y^U$, since the condition $p_2^2 p_1 < y$ gives $p_2 < y^{(1-U)/2}$, and the required inequality is satisfied since (1.8) shows that the expressions on both sides equal zero.

In the case when $p_1 < y^U$ write

$$x = \log p_1/\log y = \log_y p_1,$$

the logarithm of p_1 to the base y. Because of (1.8), to prove (3.9) it suffices to show

$$w(p_2) \leqslant \frac{U-x}{(1-x)^2}\cdot 4\log_y^2 p_2 \quad \text{when} \quad p_2 < y^{1/3} \leqslant p_1 < y/p_2^2. \tag{3.12}$$

Since

$$\frac{d}{dx}\left\{\frac{U-x}{(1-x)^2}\right\} = \frac{2(U-x)-(1-x)}{(1-x)^3}\begin{cases} <0 & \text{if} \quad 2U-1 < x < 1,\\ >0 & \text{if} \quad x < 2U-1,\end{cases}$$

the function of x in (3.12) has no local minima and it suffices to check (3.12) in the cases $p_1 = y^{1/3}$ and $p_2 p_1^2 = y$. When $p_2^2 p_1 = y$ the inequality (3.12) reduces to

$$w(p_2) \leqslant U - x = W(1) - w(p_1). \tag{3.13}$$

In fact whenever $p_2^2 p_1 < y$ we have, from (1.8),

$$2w(p_2)+w(p_1) \leqslant 2\left(\log_y p_2 - \frac{1-U}{2}\right)+(\log_y p_1 - V) \leqslant U - V = W(1),$$

so (3.13) holds as required. On the other hand when $p_1 = y^{1/3}$ the requirement (3.12) reduces to one specified in (1.8).

The structure of our proof of (ii) is similar. We may write, as in (2.1),

$$\delta = p_1 p_2 \ldots p_\nu, \quad p_1 > p_2 > \ldots > p_\nu > q. \tag{3.14}$$

Since $\bar{\chi}_y(\delta) = 1$ and $2 \nmid \nu = \nu(\delta) \geqslant 3$ we have by (2.4), (2.21), (2.22) that $\nu \leqslant 1/m$ and

$$p_1^3 < y; \quad p_1 p_2^3 < y; \quad p_1 p_2 \ldots p_{\nu-2} p_{\nu-1}^3 < y \leqslant p_1 \ldots p_{\nu-1} p_\nu^3, \tag{3.15}$$

which conditions need not all be distinct.

If $p_1 > y^U$ we have $w(p_1) = W(1)$ from (1.8), but $p_2 < y^{(1-U)/2}$ from (3.15), whence (1.8) gives $w(p_j) = 0$ when $j \geqslant 2$, and (3.10) follows in this case.

Observe that (3.15), (3.14) imply, when $q^2 < y/\delta$, that

$$\frac{y}{p_1 \ldots p_\nu} < p_\nu^2 < (p_1 \ldots p_\nu)^{2/\nu} < \left(\frac{y}{q^2}\right)^{2/\nu},$$

so that

$$\delta > y^{1-2/\nu} q^{4/\nu}. \tag{3.16}$$

If $p_1 \leqslant y^U$ then write

$$\log_y \delta = \sum_{p|\delta} \log_y p = x.$$

Because of (3.15), (1.8) we have

$$\sum_{p|\delta} w(p) \leqslant x - \nu m \tag{3.17}$$

where $m = \max\{V, (1-U)/2\}$ as in (1.7). To establish (3.11) it therefore suffices to show

$$w(q) \leqslant \frac{U-V+\nu m - x}{(1-x)^2} \cdot 4\log_y^2 q \tag{3.18}$$

when $q^2 < y/\delta$ and (3.16) holds. As before it suffices to check (3.18) when δ takes the extreme values

$$y/q^2, \quad y(y/q^2)^{-2/\nu}.$$

When $\delta = y/q^2$ the inequality (3.18) reduces to

$$w(q) \leqslant U - V + \nu m - x.$$

Because of (3.17) the proof of this inequality under the conditions (3.15) will also establish (3.10). In fact we have, using (3.15),

$$w(q) + x \leqslant \log_y q - m + \log_y(p_1 \ldots p_\nu) \leqslant 1 - m \leqslant U - V + \nu m,$$

because $m \geqslant V$ and $\nu m \geqslant 2m \geqslant 1 - U$.

It remains to check (3.18) when $y/\delta = (y/q^2)^{2/\nu}$. We require

$$w(q) \leqslant \frac{U-V+\nu m - 1 + (1-x)}{(1-x)^2} \cdot 4\log_y^2 q$$

when

$$1 - x = \frac{2}{\nu}(1 - 2\log_y q).$$

Since $m \geqslant (1-U)/2$, $m \geqslant V$, $\nu \geqslant 3$ we have

$$U - V + \nu m - 1 \geqslant U + (\nu-1)m - 1 \geqslant U - 1 + 2m \geqslant 0.$$

The inequality (3.11) will thus follow if

$$w(q) \leqslant \frac{4\nu \log_y^2 q}{2(1 - 2\log_y q)}.$$

This follows from (1.8) since

$$9(U - \tfrac{1}{3}) \leqslant 6 \leqslant 2\nu.$$

The proof of (iii) is related to that of (ii). Since $\nu = \nu(\delta) \geqslant 2$ we have

$$p_1^3 < y; \quad p_1 \ldots p_{\nu-3} p_{\nu-2}^3 < y.$$

Consequently when $\nu > 2$ we have

$$\sum_{p|\delta} w(p) \leqslant 1 - \nu m \leqslant U - V,$$

because $m \geqslant V$ and $(\nu - 1)m \geqslant 2m \geqslant 1 - U$. If $\nu = 2$ then

$$w(p_1) + w(p_2) \leqslant \tfrac{2}{3} - 2m \leqslant U - V,$$

because $m \geqslant V$ and $m + U \geqslant (1+U)/2 > 2/3$ when $U \geqslant 1/2$. This completes the proof of Lemma 4.

Lemma 5. *When $\mathscr{A}$ satisfies* (1.2) *and w satisfies* (1.6), (1.8) *we have*

$$\sum_{a \in \mathscr{A}} \Sigma(a, z) \geqslant XM(z) + R(z), \tag{3.19}$$

where

$$M(z) = M_1(z) - M_2(z) - M_3(z), \quad R(z) = R_1(z) - R_2(z) \tag{3.20}$$

with

$$M_1(z) = \sum_{d|P(z)} \frac{\mu(d)\chi_y(d)\varphi(d)\varrho(d)}{d}, \tag{3.21}$$

$$M_2(z) = \sum_{y^{1/3} \leqslant p < z} e^\gamma V\left(\left(\frac{y}{p}\right)^{1/2}\right) \frac{\varrho(p)\varphi(p)}{p} \left\{1 + O\left(\frac{L}{\log y}\right)\right\}, \tag{3.22}$$

(3.23)

$$M_3(z) = \sum_{\substack{\delta | P(z) \\ 2 \nmid \nu(\delta);\, 3 \leqslant \nu(\delta) \leqslant 1/m}} e^\gamma V\left(\left(\frac{y}{\delta}\right)^{1/2}\right) \frac{\bar{\chi}_y(\delta)\varrho(\delta)\varphi(\delta)}{\delta} \left\{1 + O\left(\frac{L}{\log y}\right)\right\}$$

and

$$R_1(z) = \sum_{d|P(z)} \mu(d)\chi_y(d)\varphi(d)R(\mathscr{A}, d), \tag{3.24}$$

$$R_2(z) = \sum_{\substack{\delta | P(z) \\ 2 \nmid \nu(\delta) \leqslant 1/m}} \bar{\chi}_y(\delta)\varphi(\delta) \sum_{d_1, d_2} \lambda_{y/\delta}(d_1)\lambda_{y/\delta}(d_2) R(\mathscr{A}, [d_1, d_2]\delta). \tag{3.25}$$

Here $V(\cdot)$ is as defined in (1.11), *φ as in* (2.6), *and χ as in* (2.19), (2.20).

In particular we have

$$|R(z)| \leqslant W(1) \sum_{d|P(z)} 3^{\nu(d)} |R(\mathscr{A}, d)|. \tag{3.26}$$

The proof of this lemma uses the decomposition of $\Sigma(a, z)$ provided by Lemma 2. Application of (1.2) gives

$$\sum_{a \in \mathcal{A}} \Sigma_1(a, z) = XM_1(z) + R_1(z). \tag{3.27}$$

For $\Sigma_2(a, z)$, part (i) of Lemma 4 shows that Lemma 3 applies to give

$$\Sigma_2(a, z) \geqslant -XM_2(z) + R_2^{(1)}(z). \tag{3.28}$$

Here we have used the fact that $p \geqslant y^{1/3}$ when $\bar{\chi}_y(p) \neq 0$, so that $q^2 p < y$ implies $q < p$. The O-constant depends on the parameter U in (1.8). The entry $R_2^{(1)}(z)$ represents the contribution to $R_2(z)$, as defined in (3.25), from the terms with $\delta = p$.

For $\Sigma_3(a, z)$ we show that the inner sum in (2.26) satisfies

$$\sum_{t|(a, P(p(\delta)))} \mu(t)\varphi(t\delta) \leqslant \{W(1) - \sum_{p|\delta} w(p)\} S\left(\frac{a}{\delta}, \left(\frac{y}{\delta}\right)^{1/2}\right) + \sum_{q|a;\, q < \sqrt{y/\delta}} f(q) S\left(\frac{a}{q\delta}, \left(\frac{y}{\delta}\right)^{1/2}\right), \tag{3.29}$$

where

$$f(q) = w(q)/\{W(1) - \sum_{p|\delta} w(p)\}. \tag{3.30}$$

The expression on the right of (3.30) is non-negative because of Lemma 4. The inequality (3.29) follows because the expression on the left is expressible via (2.12) and (2.10) as

$$\{W(1) - \sum_{p|\delta} w(p)\} S\left(\frac{a}{\delta}, z_1\right) + \sum_{q|(a/\delta);\, q < z_1} w(q) S\left(\frac{a}{\delta q}, z_1\right) \tag{3.31}$$

with $z_1 = p(\delta)$; the change to $z_1 = \sqrt{y/\delta}$ increases the right side of (3.31), as asserted, because in (2.26) we have $\bar{\chi}_y(\delta) = 1$, so that $p(\delta) > \sqrt{y/\delta}$. Furthermore part (ii) of Lemma 4 applies to show that $f(q)$ satisfies the hypotheses of Lemma 3, which now gives

$$\sum_{a \in \mathcal{A}} \Sigma_3(a, z) \geqslant -XM_3(z) + R_2^{(2)}(z), \tag{3.32}$$

where $R_2^{(2)}(z)$ is the contribution to $R_2(z)$ from those δ with $\nu(\delta) \geqslant 3$. Here, for the O-term in the expression (3.23) for $M_3(z)$, we use the fact that when $\bar{\chi}_y(\delta) \neq 0$ and when δ is written as in (3.14) we have (cf. the argument leading to (8.4.11) in [5])

$$\log \frac{y}{p_1 \ldots p_\nu} > \frac{1}{2} \log \frac{y}{p_1 \ldots p_{\nu-1}} > \frac{1}{2}\left(\frac{2}{3}\right)^{\nu-1} \log y.$$

Since (see (2.20)) ν is bounded by $1/m$ when $2 \nmid \nu(\delta)$, this shows that the parameter ε in the current application of Lemma 3 depends only on $U < 1$.

For $\Sigma_4(a, z)$ in Lemma 2 part (iii) of Lemma 4 gives, via (2.12), that

$$\Sigma_4(a, z) \geqslant 0.$$

With (3.27), (3.28), (3.32), this completes the proof of (3.19).

To deduce (3.26) we use the fact (cf. § 3.1 in [5]) that

$$|R_2(z)| \leqslant W(1) \sum_{\substack{\delta | P(z) \\ 2 \nmid \nu(\delta) \leqslant 1/m}} \bar{\chi}_y(\delta) \sum_{\substack{d \leqslant y/\delta \\ d | P(\sqrt{y/\delta})}}{}' 3^{\nu(d)} |R(\mathcal{A}, d\delta)|,$$

together with the facts that the expression of a given d ad $\delta d'$ with $\bar{\chi}_y(\delta) = 1$, $d | P(\sqrt{y/\delta})$ is, if it exists, unique and implies $\chi_y(d) = 0$.

4. The approximate identity for the main term. In this section we prove the following result.

Lemma 6. *Suppose that the multiplicative function ϱ satisfies (Ω_1), (Ω_2), and that the weight function w satisfies (1.6), (1.8), where U, V satisfy (1.5). Then the quantity $M(z)$ defined in (3.20) is estimated by*

$$M(y^U) = 2e^\gamma V(y)\left\{\mathcal{M}(W) + O\left(\frac{L^{1/5}}{\log^{1/5} y}\right)\right\}, \tag{4.1}$$

where $\mathcal{M}(W)$ is as stated in Theorem 1 and $V(\cdot)$ is as in (1.11).

Actually, the estimate (4.1) holds good for $M(z)$ for all z satisfying

$$y^{\max\{U, 1-2V\}} \leqslant z \leqslant y,$$

as an examination of the proof would show.

The proof of Lemma 6 will follow from Lemmas 8 and 12 below.

We will handle $M(z)$ by replacing the quantities $M_2(z)$, $M_3(z)$, appearing in (3.22), (3.23), by certain essentially equivalent expressions. This equivalence is based on the fact that, in the range of the relevant parameters allowed in Lemma 7 below, the (unweighted) upper bound sieve of Rosser and Iwaniec gives the same bound as does the λ^2-method of Selberg employed in the proof of Lemma 3.

The Rosser–Iwaniec functions $\chi_y^{\pm}$ are defined as usual, by specifying $\chi_y^{\pm}(1) = 1$ and, when $d > 1$, $|\mu(d)| = 1$,

$$\chi_y^+(d) = \begin{cases} 1 & \text{if } hp^2(h) < y \text{ when } h|d \text{ and } 2 \nmid \nu(h), \\ 0 & \text{otherwise}, \end{cases} \tag{4.2}$$

$$\chi_y^-(d) = \begin{cases} 1 & \text{if } hp^2(h) < y \text{ when } h|d \text{ and } 2|\nu(h), \\ 0 & \text{otherwise}. \end{cases} \tag{4.3}$$

For our convenience we also define a function $\chi_y^*(d)$ by $\chi_y^*(1) = 1$ and, when $d > 1$, $|\mu(d)| = 1$,

$$\chi_y^*(d) = \begin{cases} 1 & \text{if } hp^2(h) < y \text{ when } h|d \text{ and } 2 \nmid \nu(h) \leqslant 1/m, \\ 0 & \text{otherwise}. \end{cases} \tag{4.4}$$

Then the function χ_y defined in (2.19), (2.20) is expressible as

(4.5) $$\chi_y(d) = \chi_y^-(d)\,\chi_y^*(d).$$

Note that

(4.6) $$\overline{\chi}_y(\delta) = \chi_y^-(\delta)\,\overline{\chi}_y^*(\delta) \quad \text{if} \quad 2 \nmid \nu(\delta),$$

for then $\chi_y^-(\delta/p(\delta)) = \chi_y^-(\delta)$, and (4.6) follows using (2.3). Also

$$\overline{\chi}_y^*(\delta) \neq 0 \Rightarrow 2 \nmid \nu(\delta) \leqslant 1/m.$$

The property of χ_y^+ that we invoke is as follows.

LEMMA 7. *When ϱ satisfies (Ω_1), (Ω_2) and $0 < s \leqslant 3$ we have*

$$\sum_{d \mid P(y^{1/s})} \frac{\mu(d)\,\chi_y^+(d)\,\varrho(d)}{d} = e^{\gamma} V(\sqrt{y})\left\{1+O\left(\frac{L^{1/5}}{\log^{1/5} y}\right)\right\}.$$

The upper bound in Lemma 7 is, even with the better error term $O(\log^{-1/3} y)$, a special case of the result of [9]. As stated, the result follows using the techniques described in [8]. For our convenience we have taken $s = 2$ in an expression $F(s)\,V(y^{1/s})$ on the right, the sum on the left being in fact independent of s when $0 < s \leqslant 3$.

In the next lemma and elsewhere we decompose d, when $p \mid d$ and $|\mu(d)| = 1$, as

(4.7) $$d = d_1 p d_2 \quad \text{with} \quad p(d_1) > p > q(d_2).$$

LEMMA 8. *The quantity $M(z)$ defined in (3.20) satisfies the approximate identity*

(4.8) $$M(z) = M^*(z)\left\{1+O\left(\frac{L^{1/5}}{\log^{1/5} y}\right)\right\},$$

where

(4.9) $$M^*(z) = \sum_{d \mid P(z)} \frac{\mu(d)\,\varrho(d)\,\chi_y^-(d)}{d}\Big\{W(1) - \sum_{p \mid d} w(p)\,\chi_y^*(d_1)\Big\},$$

with d_1 as in (4.7) and χ^* as in (4.4).

The first stage in proving Lemma 8 is to use Lemma 7 (with y replaced by y/δ) in the expressions for $M_2(z)$ and $M_3(z)$ in Lemma 5. We obtain

$$M_2(z)+M_3(z) = \sum_{\substack{d \mid P(z) \\ 2 \nmid \nu(d) \leqslant 1/m}} \frac{\overline{\chi}_y(d)\,\varrho(d)\,\varphi(d)}{d} \sum_{t \mid P(p(d))} \frac{\mu(t)\,\chi_{y/d}^+(t)\,\varrho(t)}{t}\left\{1+O\left(\frac{L^{1/5}}{\log^{1/5} y}\right)\right\}.$$

The condition on t appears first as $t \mid P(\xi)$ with $\xi = \sqrt{y/d}$. However, the sum over ξ is then independent of ξ when $\xi \geqslant \sqrt{y/d}$, so when $\overline{\chi}_y(d) \neq 0$ we may

take $\xi = p(d)$ as stated. This establishes (4.8) with, however, $M^*(z)$ expressed as

(4.10) $$M^*(z) = \sum_{d \mid P(z)} \frac{\mu(d)\,\chi_y(d)\,\varphi(d)\,\varrho(d)}{d} - \sum_{\substack{d \mid P(z) \\ 2 \nmid \nu(d) \leqslant 1/m}} \frac{\overline{\chi}_y(d)\,\varrho(d)\,\varphi(d)}{d} \sum_{t \mid P(p(d))} \frac{\mu(t)\,\chi_{y/d}^+(t)\,\varrho(t)}{t},$$

where

(4.11) $$\varphi(d) = W(1) - \sum_{p \mid d} w(p)$$

as in (2.6).

Because of (4.5), (4.6), (4.2), (4.3), this can be rewritten as

(4.12) $$M^*(z) = \sum_{d \mid P(z)} \frac{\mu(d)\,\chi_y^*(d)\,\varphi(d)\,\chi_y^-(d)\,\varrho(d)}{d} + \sum_{\substack{d \mid P(z) \\ 2 \nmid \nu(d) \leqslant 1/m}} \mu(d)\,\overline{\chi}_y^*(d)\,\varphi(d) \sum_{t \mid P(p(d))} \frac{\mu(t)\,\chi_y^-(dt)\,\varrho(dt)}{dt}.$$

This can be expressed as

(4.13) $$M^*(z) = W(1) \sum_{d \mid P(z)} g_1(d) + \sum_{p < z} w(p) \sum_{p \mid d \mid P(z)} g_p(d).$$

It remains to show that this expression is identical to that appearing in (4.9).

In the expression (4.12), the coefficient of $W(1)$ is obtained by replacing $\varphi(d)$ by $W(1)$. Then we apply the "fundamental identity" (2.5), in which we take

$$\chi_{\mathscr{B}} = \chi_y^*; \quad A = P(z); \quad \varphi(d) = \chi_y^-(d)\,\varrho(d)/d.$$

This shows that in (4.13) we have

(4.14) $$g_1(d) = \frac{\mu(d)\,\chi_y^-(d)\,\varrho(d)}{d}.$$

We use similar principles to show that in (4.13) we have

(4.15) $$g_p(d) = -\mu(d)\,\varrho(d)\,\chi_y^-(d)\,\chi_y^*(d_1)/d,$$

with d_1 as in (4.7). It is almost immediate that this is admissible for those pairs p, d with $d_1 = 1$ (so that p is the largest prime factor of $d = pd_2$). To see this, observe that in (4.12) there is a solitary term with $d = 1$; in every other term we can write $d = pd_2$ (possibly with $d_2 = 1$). In each of these terms in (4.12) there is a summand $-w(p)$ corresponding to each occurrence of a summand $W(1)$, because of (4.11). Consequently we have, in (4.13), the equation

$$g_1(d) = -g_p(d) \quad \text{when} \quad d_1 = 1.$$

This establishes (4.15) in the case $d_1 = 1$.

An argument similar to that just used shows that it is sufficient to establish (4.15) in the case when $2 \nmid \nu(d_1)$, since the result in the contrary case will follow from it. In those terms in (4.12) with $\nu(d_1) = 2k$, $d \neq d_1$ write $d = d_1 p d_2$ as in (4.7). Corresponding to each summand $W(1) - \sum_{q|d_1} w(q)$ in (4.12) there is a summand $-w(p)$ (there being no terms with $\bar{\chi}_y^*(d_1) = 1$). Then (4.14) will follow in the case when $\nu(d_1) = 2k+1$.

When $2 \nmid \nu(d_1)$ we have

$$\bar{\chi}_y^*(d) = \chi_y^*(pd_1)\,\bar{\chi}_{y/(pd_1)}^*(d_2).$$

In (4.13) we now have, from (4.12), (4.11),

$$g_p(d) = \mu(d_1)\chi_y^*(pd_1)\psi(p, d_1)w(p)$$

where

$$\psi(p, d_1) = \sum_{\substack{d_2|P(p)\\ d=d_1pd_2}} \mu(d_2)\chi_{y/(pd_1)}^*(d_2)\chi_y^-(d)\varrho(d)/d$$

$$+ \sum_{\substack{d_2|P(p)\\ d=d_1pd_2}} \mu(d_2)\bar{\chi}_{y/(pd_1)}^*(d_2) \sum_{t|P(p(d_2))} \frac{\mu(t)\chi_y^-(dt)\varrho(dt)}{dt}$$

$$= \sum_{\substack{d_2|P(p)\\ d=d_1pd_2}} \mu(d_2)\chi_y^-(d)\varrho(d)/d,$$

by the "fundamental identity" (2.5). Hence

$$g_p(d) = -\mu(d)\chi_y^*(pd_1)\chi_y^-(d)\varrho(d)/d,$$

and (4.15) follows since

$$\chi_y^*(pd_1) = \chi_y^*(d_1) \quad \text{when} \quad 2 \nmid \nu(d_1).$$

This completes the proof of Lemma 8.

In the remainder of this section our treatment is closely related to the corresponding details in the article [6] of Halberstam and Richert. We follow them in denoting

$$T^{(-)^\nu}(y, y^{1/s}) = \sum_{d|P(y^{1/s})} \frac{\mu(d)\varrho(d)\chi_y^{(-)^\nu}(d)}{d}. \tag{4.16}$$

The following "Reduction Lemma" is a case of Lemma 5 in [6]. A continuous analogue of Lemma 9 was provided by the present author as Lemma 5.1 in [2]. The ensuing arguments in this section could also be arranged as in [2], instead of as below.

LEMMA 9. *The sum defined in* (4.16) *satisfies*

$$T^{(-)^\nu}(y, y^{1/s}) = \sum_{\substack{y^{1/s} < p(t), q(t) < y/t\\ \nu(t) \equiv \nu \bmod 2}} \frac{\mu^2(t)\varrho(t)}{t} T^+\left(\frac{y}{t}, \left(\frac{y}{t}\right)^{1/3}\right) + O\left(V(y^{1/s})\log^{-1/5} y\right),$$

where the O-constant may depend upon $s \geq 1$.

Essentially, the proof of Lemma 9 provided in [6] rests upon repeated use of identities of Buchstab's type to express the left side in terms of T-functions for which the estimates are of the elementary type provided by Lemma 7 (and by the analogous result for $T^-(y, \sqrt{y})$). There are some complications, relating to the occurrence of non-squarefree moduli in the resulting expressions, which are, it is shown in [6], absorbed by the error term in our Lemma 9.

We will need the following identity.

LEMMA 10. *Define*

$$b(n) = \sum_{\substack{d|n\\ q(n/d) < y/(pn)}} \mu(d)\chi_y(d).$$

Then $b(1) = 1$, *and if* $n > 1$ *then*

$$b(n) = \begin{cases} -\bar{\chi}_y(n) & \text{if } np(n) < y/p, \\ 0 & \text{otherwise.} \end{cases}$$

The proof of this identity is almost identical to that of a corresponding identity for χ_y^- established in [6], which may in turn be regarded as an arithmetisation of Lemma 5.2 in [2]. The function χ_y shares the properties of the function χ_y^- necessary to make the proof of Lemma 10 a straightforward exercise to the reader of [6].

In the next lemma we adopt the abbreviation, used in [6],

$$\sigma^+(\eta) = T^+(\eta, \eta^{1/3}) \quad (\eta > 1).$$

LEMMA 11. *When* $z = y^U$ *the quantity* $M^*(z)$ *appearing in Lemma* 8 *is expressible as*

$$M^*(y^U) = W(1)T^-(y, y^{1/2}) - \sum_{y^{1/2} \leq p \leq y^U} \{W(1) - w(p)\}\frac{\varrho(p)}{p}\sigma^+\left(\frac{y}{p}\right)$$

$$+ \sum_{p < y^{1/2}} \frac{w(p)\varrho(p)}{p}\left\{\sigma^+\left(\frac{y}{p}\right) - \sum_{\substack{n|P(y);\, 2|\nu(n)\\ p < p(n) < y/(pn)}} \frac{\bar{\chi}_y(n)\varrho(n)}{n}\sigma^+\left(\frac{y}{pn}\right)\right\}$$

$$+ O\left(V(y)\log^{-1/5} y\right).$$

The proof of this lemma is analogous to that of equation (6.15) in [6], to which we refer the reader for discussion of details. In this proof, the coefficient of $W(1)$ is exactly as in [6]. The remaining contribution to $M^*(z)$

is expressible as

$$H(z) = \sum_{d|P(z)} \mu(d)\varrho(d)\chi_y^-(d)w(p)\chi^*(d_1)/d,$$

where d has been expressed as $d = d_1 p d_2$ as in (4.7). This can be rewritten as

$$\sum_{p<z} \frac{w(p)\varrho(p)}{p} \sum_{\substack{d_1|P(z)\\ p(d_1)>p}} \frac{\mu(d_1)\varrho(d_1)\chi_y(d_1)\chi_y^-(d_1 p)}{d_1} \sum_{d_2|P(p)} \frac{\mu(d_2)\varrho(d_2)\chi_{y/(pd_1)}^{(-)^{\nu(d_1)}}(d_2)}{d_2}.$$

The inner sum over d_2 is expressible via Lemma 9 as

$$T^{(-)^{\nu(d_1)}}\left(\frac{y}{pd_1}, p\right) = \sum_{\substack{p(t)>p\\ tq(t)<y/(d_1 p)\\ \nu(t)\equiv\nu(d_1) \bmod 2}} \frac{\mu^2(t)\varrho(t)}{t}\sigma^+\left(\frac{y}{d_1 pt}\right) + O\left(V(p)\log^{-1/5} p\right).$$

The conditions of summation imply $\chi_y^-(d_1 p) = 1$, hence

$$H(z) = \sum_{p<z} \frac{w(p)\varrho(p)}{p} \sum_{\substack{d|P(z)\\ p<p(d)}} \frac{\mu(d)\varrho(d)\chi_y(d)}{d} \sum_{\substack{p<p(t)\\ tq(t)<y/(dp)\\ \nu(t)\equiv\nu(d) \bmod 2}} \frac{\mu^2(t)\varrho(t)}{t}\sigma^+\left(\frac{y}{dpt}\right) + O\left(V(y)\log^{-1/5} y\right),$$

wherein t also necessarily divides $P(z)$, since $z = y^U$. Writing $dt = n$, we then obtain

$$H(z) = \sum_{p<z} \frac{w(p)\varrho(p)}{p} \sum_{\substack{p<p(n);\, n|P(z)\\ 2|\nu(n)}} \frac{\mu^2(n)\varrho(n)}{n}\sigma^+\left(\frac{y}{pn}\right) \sum_{\substack{d|n\\ q(n/d)\leqslant y/(pn)}} \mu(d)\chi_y(d) + O\left(V(y)\log^{-1/5} y\right),$$

the terms with $(d, t) > 1$ being dealt with as in [6], and Lemma 11 follows after an application of Lemma 10.

In Lemma 11, the factor $\bar{\chi}_y(n)$ is 0 if $\nu(n) > 1/m$; this is a consequence of (2.4) with our specification (2.22), of which we have already taken advantage in our treatment of the O-term in (3.23). This fact is, however, no longer relevant, because from (1.8) we have that $w(p)$ is zero if $p < y^{1/m}$. Accordingly, on estimating sums by integrals in the established way (cf. Lemma 4.4 in [2], for example) we obtain the concluding result of this section, as follows.

LEMMA 12. *The quantity* $M^*(y^U)$ *appearing in Lemma 11 is expressible as*

$$M^*(y^U) = 2e^\gamma V(y)\left\{\mathcal{M}(W) + O\left(\frac{L^{1/5}}{\log^{1/5} y}\right)\right\}$$

where $V(y)$ *is as in* (1.11) *and* $\mathcal{M}(W)$ *is as in Theorem* 1.

As indicated previously, Lemma 6 follows from Lemmas 8 and 12, and with Lemmas 1 and 5 completes the proof of Theorem 1.

5. Numerical estimations. For applications, e.g. to the result stated in (1.4), it is necessary to make an appropriate choice of U, V and then estimate the expression $\mathcal{M}(W)$ of Theorem 1 (from below). The considerations relating to the choice of parameters differ from those in [2], because (cf. Theorem 1) it is not now the case that $h(t, 1)$ exceeds $1/(1-t)$ for small t; in fact Theorem 3 indicates that $h(t)$ is always less than 1/5.

For $R = 2, 3, 4, \ldots$ we attempt to maximise g subject to the constraints required in Section 1, denoting the optimal value by

$$\Lambda_R = R - \delta_R.$$

The constraints are that $\mathcal{M}(W) > 0$, with U, V chosen subject to (1.5) and W satisfying (1.8). The calculations of $\mathcal{M}(W)$ were carried out using Theorem 2 and the methods outlined in [4] for the numerical inversion of the moment map, using information derived from an implementation of the algorithms described in [3]. A large computer was employed, and a search conducted for the best choice of U, V for each R. Equality was chosen in the upper bound for W specified in (1.8), save that in the (unimportant) interval where

$$0 < t - m < 9(U - \tfrac{1}{3})t^2$$

we took $W(t) = 0$, for simplicity. It appeared that only for $R \leqslant 4$ does an improvement results upon that derived by the methods of [2]. For these R the optimal choice appears to equal (or to lie very close to) those given by the equations

$$2V + U = 1; \quad \mathcal{M}(W) = 0.$$

The solution U of this system was calculated and found to satisfy

$$U = 0.9702933\ldots$$

This leads to the values of δ_R stated in (1.4).

It appears to be difficult to make any particularly perceptive comment on the fact that when $R \geqslant 5$ our calculations indicate that the method of this paper does not improve upon that of [2]. It is certainly the case that the methods of Section 3 yield better bounds for the expressions $\Sigma_2(a, z)$, $\Sigma_3(a, z)$ of Lemma 2 than do the combinatorial methods of [2], provided the constraints (1.8) are satisfied; when R is large we are however led to values of the relevant parameters for which these constraints, in particular the requirement

$$W(t) \leqslant t - (1-U)/2 \quad \text{if} \quad (1-U)/2 \leqslant t \leqslant 1/3,$$

are significantly more stringent than the corresponding requirements in [2].

References

[1] K. Dickman, *On the frequency of numbers containing prime factors of a certain relati magnitude*, Ark. Matematik Astron. och Fysik 22A, No. 10 (1930), pp. 1–14.
[2] G. Greaves, *A weighted sieve of Brun's type*, Acta. Arith. 40 (1982), pp. 297–332.
[3] — *An algorithm for the solution of certain differential-difference equations of advanced ty* Maths. of Comp. 38 No. 157 (1982), pp. 237–247.
[4] — *An algorithm for the Hausdorff moment problem*, Numerische Math. 39 (1982), pp. 23 238.
[5] H. Halberstam and H.-E. Richert, *Sieve Methods*, London 1974.
[6] — — *A weighted sieve of Greaves' type I*; in: *Elementary and Analytic Theory of Numbe* Banach Center Publications, Vol. 17, Polish Scientific Publishers, Warszawa 1985, p 155–182.
[7] H. Iwaniec, *On the error term in the linear sieve*, Acta Arith. 19 (1971), pp. 1–30.
[8] — *The half dimensional sieve*, ibid. 29 (1976), pp. 69–95.
[9] — *Rosser's sieve*, ibid. 36 (1980), pp. 171–202.
[10] W. B. Jurkat and H.-E. Richert, *An improvement of Selberg's sieve method. I*, ibid. (1965), pp. 217–240.
[11] W. C. Mitchell, *An evaluation of Golomb's constant*, Math. Comp. 22 (1968), pp. 41 415.
[12] H.-E. Richert, *Selberg's sieve with weights*, Mathematika 16 (1969), pp. 1–22.

Received on 18.2.1985 (149