Then, as $S \equiv 0 \pmod 4$ by (1.12), we obtain

$$h(-2p) \equiv h(2p)\frac{S}{2} + 8 \pmod{16},$$

which completes the proof of the theorem in this case.

The authors would like to thank Mr. Lee-Jeff Bell, who did some computing for them in connection with preparation of this paper.

References

[1] Pierre Kaplan, *Divisibilité par 8 du nombre des classes des corps quadratiques dont le 2-groupe des classes est cyclique, et réciprocité biquadratique*, J. Math. Soc. Japan 25 (1973), pp. 596–608.

[2] Edmund Landau, *Elementary number theory*, Chelsea Publishing Company, New York 1958.

[3] Bernard Oriat, *Sur la divisibilité par 8 et 16 des nombres de classes d'idéaux des corps quadratiques* $Q(\sqrt{2p})$ *at* $Q(\sqrt{-2p})$, J. Math. Soc. Japan 30 (1978), pp. 279–285.

[4] Kenneth S. Williams, *On the class number of* $Q(\sqrt{-p})$ *modulo 16, for* $p \equiv 1 \pmod 8$ *a prime*, Acta Arith. 39 (1981), pp. 381–398.

10 Allée Jacques Offenbach
54420 — Saulxures les Nancy
France

DEPARTMENT OF MATHEMATICS AND STATISTICS
CARLETON UNIVERSITY
Ottawa, Ontario, Canada
K1S 5B6

Received on 27.3.1979
and in revised form on 21.12.1979 (1151)

ACTA ARITHMETICA
XL (1982)

A weighted sieve of Brun's type

by

G. Greaves (Cardiff)

To the memory of Viggo Brun

1. Introduction and statement of theorems. We study a class of integer sequences $\mathscr{A} = \mathscr{A}_X$ depending on a real parameter $X \geqslant 2$. With possible applications in mind, assume the sequence satisfies certain general conditions of the type introduced by Halberstam and Richert [3] (see also Ankeny and Onishi [1]). These authors supply many interesting examples of such sequences. The object of the exercise is to deduce, for a suitable integer $R \geqslant 2$, that the sequence $\mathscr{A}$ contains one (indeed, many) numbers having no more than R prime factors.

In the first place we assume

$$\sum_{\substack{a \in \mathscr{A} \\ a \equiv 0 \bmod l}} 1 = \frac{X}{l}\gamma(l) + R(X, l) \quad \text{if} \quad \mu^2(l) = 1, \tag{1.1}$$

μ being the Möbius function. The function γ is assumed to be multiplicative and to satisfy

$$0 \leqslant \gamma(p) < p, \quad -L < \sum_{w \leqslant p < z} \frac{\gamma(p)\log p}{p - \gamma(p)} - \log\frac{z}{w} \leqslant A_1 \quad (2 \leqslant w < z). \tag{A$_1$}$$

Here and below the constants $g, c, A_1, A_2, \ldots$ are absolute; this means independent of the real variables X, y, z, w. With the possible applications in mind, however, L will be allowed to depend on X as in [3].

Also assume, when $a \in \mathscr{A}$,

$$1 \leqslant a < y^g; \quad L \leqslant \log y, \tag{A$_2$}$$

where the significance of y will appear below. The further condition on L is added for our convenience.

The results of this paper are stated in such a way that they are independent of hypotheses relating to the "error term" R in (1.1). For applications, however, some knowledge of the following type would be needed. One could use, for example:

(A_3) *for the function* $f(l)$ *appearing in Theorem* 1 (*this satisfies* $0 \leqslant f(l) \leqslant 1$) *the number* $y = y(X)$ *appearing in* (A_2) *has the property*

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

We are interested in results of this type: *assume* (A_1), (A_2), (A_3). *Then, if* $g \leqslant R - \delta_R$, *there exists* a *in* $\mathscr{A}$ *with at most* R *prime factors.*

If it be required that repeated prime factors of a should be counted once only, then an additional hypothesis is required, e.g.

(A_4) *there exists a constant* c *such that*

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

Richert [9] (see also Chapter 9 of [3]) proved a result of the type described above, showing that we may take

$$\delta_R = \left[\log\left\{\frac{4}{3(1+3^{-R})}\right\}\right]\Big/\log 3, \tag{1.2}$$

so that, for all R, $\delta_R \leqslant \delta_\infty = 0.262\ldots$, while (in the perhaps most interesting case) $\delta_2 = 0.167\ldots$ In this work certain parameters were chosen so that δ_R should be an elementary function; if numerical integrations are invoked the author finds that an optimal choice of parameters in this method leads to the permissible values

$$\delta_\infty = 0.178\ldots, \qquad \delta_2 = 0.136\ldots \tag{1.3}$$

In this paper we describe a method which leads to improved values for δ_R. If R is large (and numerical work indicates that $R \geqslant 3$ is large enough) the method leads to the value

$$\delta_R = \left[\log\left\{\frac{4}{3(1+3^{-R}e^{-\beta})}\right\} - \alpha\right]\Big/\log 3 \tag{1.4}$$

for certain positive constants α, β (to be described). For $R = 2$ the description of δ_R is somewhat different (see Corollary 2 below).

Preliminary numerical estimations of α, β lead to estimations of δ_R from below which indicate

$$\delta_\infty < 0.131, \qquad \delta_2 < 0.068. \tag{1.5}$$

It is hoped to report on a more accurate computation at a later date, when discussing some applications.

As will become clear, the method used is an analogue of Rosser's version of Brun's sieve (for which see e.g. Iwaniec [4], [5]) in which the weighting device of Kuhn [7], [8], Ankeny and Onishi [1] and Richert [9] is introduced *ab initio*. The "combinatorial" contents of Section 3 would retain some relevance to the "k-dimensional" context (in which the term $\log(z/w)$ in (A_1) would be multiplied by k). However, our analysis of the resulting "main term" in the subsequent sections contains some features peculiar to the context $k = 1$ in which Brun's ideas appear to operate best.

We work with parameters satisfying

$$V < \tfrac{1}{4}; \quad \tfrac{1}{2} < U < 1; \quad 0 < T_0 < \tfrac{1}{4}, \tag{1.6}$$

$$V + RU \geqslant g, \tag{1.7}$$

where g is as in (A_2). The notation

$$u = 1/U \tag{1.8}$$

would accord with that of [3], [9]. The parameter V, however, may be positive, zero or negative.

A "weight" function w will be given by

$$w(1) = W(1) > 0; \quad w(p) = W(\log p/\log y), \tag{1.9}$$

where p is prime and $p \leqslant y$. In the case when

$$3V + U \leqslant 1, \tag{1.10}$$

the increasing function W is specified by

$$W(t) = \begin{cases} U - V & \text{if} \quad U < t \leqslant 1, \\ t - V & \text{if} \quad \tfrac{1}{4} < t \leqslant U, \\ t - (1-U)/3 & \text{if} \quad T_0 < t \leqslant \tfrac{1}{4}, \text{ and } t > (1-U)/3, \\ 0 & \text{if} \quad 0 \leqslant t \leqslant T_0 \text{ or if } 0 \leqslant t \leqslant (1-U)/3. \end{cases} \tag{1.11}$$

In the contrary case when

$$3V + U > 1, \tag{1.12}$$

then

$$W(t) = \begin{cases} U - V & \text{if} \quad U < t \leqslant 1, \\ t - V & \text{if} \quad T_0 < t \leqslant U, \text{ and } t > V, \\ 0 & \text{if} \quad 0 \leqslant t \leqslant T_0, \text{ or if } 0 \leqslant t \leqslant V. \end{cases} \tag{1.13}$$

For convenience introduce the notation

$$\nu_{y,U}(a) = \sum_{\substack{p \mid a \\ p < y^U}} 1 + \sum_{\substack{p, \alpha \\ p^\alpha \mid a,\, p \geqslant y^U}} 1 \tag{1.14}$$

for the number of prime factors of a, where multiple prime factors p (of a) are counted multiply only if $p \geqslant y^U$. The principal result of this paper is then as follows.

THEOREM 1. *There exist functions H, h such that $h(x,s) = H(x,s) = 0$ if $x > 1$ and satisfying the integral equations*

$$h(x,s) = \int_{t>s} H\left(\frac{xt}{t-1},\ t-1\right)\frac{dt}{t-1} \qquad (s>1,\ 0<x\leqslant 1),$$

(1.15)

$$H(x,s) = \int_{\substack{t>s\\ t>3}} h\left(\frac{xt}{t-1},\ t-1\right)\frac{dt}{t-1} + h_1(x,s) \qquad (s\geqslant 0,\ 0<x\leqslant 1),$$

$$h_1(x,s) = \int_{\substack{x<x_1<1/s\\ 1/3<x_1<u\\ x+x_1+u=1}} \frac{dx_1}{x_1 u} \qquad (s\geqslant 0,\ 0<x\leqslant 1),$$

together with the continuity condition

$$h(x,1) = \lim_{s\to 1+} h(x,s) \qquad (0<x\leqslant 1). \tag{1.16}$$

Assume (A_1), (A_2). *Let q_a denote the least prime factor of a, and μ the Möbius function. Then there exists $f(l)$ with $0\leqslant f(l)\leqslant 1$ (to be described) such that*

$$\sum_{\substack{a\in\mathscr{A}\\ \nu_{y,U}(a)\leqslant R}} w(q_a) \geqslant 2e^{\gamma_0}X\left\{\prod_{p<y}\left(1-\frac{\gamma(p)}{p}\right)\right\}\left\{\mathscr{M}(W)+O\left(\frac{L^{1/5}}{\log^{1/5}y}\right)\right\} - \left|\sum_{l\leqslant y}\mu(l)f(l)R(X,l)\right|, \tag{1.17}$$

where $\nu_{y,U}(a)$, w, W are as described above, γ_0 is Euler's constant, and

$$\mathscr{M}(W) = -\int_{1/2}^{1}\frac{W(1)-W(t)}{1-t}\frac{dt}{t} + \int_{1/4}^{1/2}\frac{W(t)}{1-t}\frac{dt}{t} + \int_0^{1/4} W(t)\left\{\frac{1}{1-t} - h(t,1)\right\}\frac{dt}{t}. \tag{1.18}$$

The O-constant may depend on the parameter U introduced in (1.6).

Next, introduce the numbers α, β referred to in (1.4). With h as in Theorem 1 define

$$\psi(t) = \frac{1}{1-t} - h(t,1) \qquad (0<t<1/4) \tag{1.19}$$

and

$$\psi^+(t) = \begin{cases}\psi(t) & \text{if} \quad \psi(t)\geqslant 0,\\ 0 & \text{if} \quad \psi(t)<0.\end{cases} \tag{1.20}$$

It will appear that ψ is monotone and that $\psi(t)<0$ for t near enough to 0. Thus we can define

$$\alpha = \int_0^{1/4}\psi^+(t)\,dt, \qquad \beta = \int_0^{1/4}\psi^+(t)\,dt/t. \tag{1.21}$$

The number T_0 in (1.6) will be the unique zero of ψ:

$$\psi(T_0) = 0. \tag{1.22}$$

We shall also see that $h(t,1)>0$ when $0<t<1/4$, so that

$$0\leqslant \alpha < \log(4/3). \tag{1.23}$$

The following two Corollaries of Theorem 1 follow by appropriate choices of the parameters U, V.

COROLLARY 1. *Let*

$$U_R = (1+3^{-R}e^{-\beta/3})^{-1}, \qquad g_R = R-\Delta_R, \qquad \Delta_R = \log\left(\frac{4U_R}{3e^{\alpha}}\right)\Big/\log 3. \tag{1.24}$$

Suppose the parameter g of (A_2) satisfies

$$g\leqslant g_R-\eta_R, \tag{1.25}$$

where

$$\eta_R>0, \qquad \eta_R\geqslant N_R = (1-U_R)(R-\tfrac{1}{3})-\Delta_R. \tag{1.26}$$

Then the result (1.17) *of Theorem 1 holds with*

$$\mathscr{M}(W)\geqslant \eta_R\log 3. \tag{1.27}$$

Observe that, in the notation used earlier, Corollary 1 leads to the permissible choice

$$\delta_R = \max\{\Delta_R,\ \Delta_R+N_R\}. \tag{1.28}$$

We may remark that because of (1.23) we have

$$\Delta_\infty = \lim_{R\to\infty}\Delta_R = \{\log(4/3)-\alpha\}/\log 3 > 0,$$

whence $\lim_{R\to\infty} N_R<0$, so that $\delta_R = \Delta_R>0$ for all large enough R. Numerical work indicates that $R\geqslant 3$ is large enough for this statement to hold.

The constraint $\eta_R>N_R$ arises because Corollary 1 assumes there is an optimal choice of the parameters U, V satisfying $3V+U<1$. When $R=2$ it is advantageous to select equality in (1.10). This leads to:

COROLLARY 2. *Let U_0 be the solution U of the equation*

$$U\log\left(\frac{1}{U}-1\right)+\log\left(\frac{3}{4(1-U)}\right)+\alpha-\frac{1-U}{3}\log(3e^{\beta}) = 0 \tag{1.29}$$

that satisfies $\frac{1}{2} < U_0 < 1$. *Define* g_R *by*

$$g_R = R - \Delta_R; \quad \Delta_R = (1-U_0)(R-\tfrac{1}{3}). \tag{1.30}$$

Suppose

$$g \leqslant g_R - \eta_R, \quad \textit{where} \quad \eta_R > 0. \tag{1.31}$$

Then the result of Theorem 1 *holds with*

$$\mathscr{M}(W) \geqslant \eta_R \log(3e^{\beta}). \tag{1.32}$$

In making use of Theorem 1, the problem arises of determining when $\mathscr{M}(W) > 0$. In the Corollaries this question has been transferred to a corresponding problem of determining bounds for the constants α, β. We might contemplate replacing the function W by a polynomial. This remark shows that one attack on these questions consists of studying the moments

$$h_n(s) = \int_0^1 x^n h(x, s)\,dx, \quad H_n(s) = \int_0^1 x^n H(x, s)\,dx \quad (s \geqslant 1). \tag{1.33}$$

The integral equations (1.15) induce difference-differential equations for h_n, H_n of the type more familiar in this subject, which we analyse using ideas due to de Bruijn [2]. The result is:

Theorem 2. *For integers* n *define*

$$j_n = \frac{1}{n!}\int_0^{\infty} x^n \exp\left\{-x - \int_x^{\infty} \frac{e^{-t}}{t}\,dt\right\} dx \quad \textit{if} \quad n \geqslant 0, \tag{1.34}$$

$$i_n = \frac{1}{n!}\int_0^{\infty} x^n \exp\left\{-x + \int_x^{\infty} \frac{e^{-t}}{t}\,dt\right\} dx \quad \textit{if} \quad n \geqslant 1. \tag{1.35}$$

Let

$$\begin{aligned} B_n &= 2^{n+1}(n+1)(j_n - j_{n+1}); \quad D_n = 2^{n+1}(n+1)(i_n - i_{n+1});\\ A_n &= j_n - c_{1,n}(1)B_n; \quad C_n = i_n - c_{1,n}(1)D_n, \end{aligned} \tag{1.36}$$

where

$$c_{1,n}(1) = \int_0^{1/2} \frac{x^n}{1-x}\,dx. \tag{1.37}$$

Then

$$\begin{aligned} A_n\{H_n(3) + k_{1,n}(2) + 1\} + B_n h_n(1) &= 1 \quad \textit{if} \quad n \geqslant 0,\\ C_n\{H_n(3) + k_{1,n}(2) + 1\} + D_n h_n(1) &= 1 \quad \textit{if} \quad n \geqslant 1, \end{aligned} \tag{1.38}$$

where h_n, H_n *are as in* (1.33) *and*

$$k_{1,n}(2) = \int_0^1 x^n h_1(x, 2)\,dx, \tag{1.39}$$

$h_1(x, s)$ *being as in Theorem* 1. *For* $n = 0$ *the second of the equations* (1.38) *should be replaced by*

$$(1 - 2\log 2)\{H_n(3) + h_{1,n}(2) + 1\} + 2h_n(1) = 0 \tag{1.40}$$

(*as results from formally substituting* $i_0 = \infty$).

The significance of Theorem 2 is that it provides a transformation of, for example, the question of estimating the constants α, β in (1.21). Originally relating to the somewhat recondite $h(x, s)$ in (1.15), this becomes a question about the numbers i_n and j_n. These arise from two functions of a single complex variable, regular in a half-plane, and satisfying a simple difference-differential equation.

As a result some (at least) of the computational questions involved are reduced to problems not beyond the practicable reach of a (programmable) pocket calculator. We summarise the results of such a calculation in § 8.

2. Arithmetical identities. The identities of this section will be crucial in the sequel. Except where indicated, the exposition is independent of the assumptions of § 1.

Lemma 2.1. *Let* A *be squarefree, and suppose real* $w(1)$ *and* $w(p)$, *for primes* p, *are given. For real* b *define*

$$\Lambda\{A, b\} = \begin{cases} b & \textit{if} \quad A = 1,\\ w(p) & \textit{if} \quad A = p,\\ 0 & \textit{if} \quad A \textit{ is composite}. \end{cases} \tag{2.1}$$

Suppose l *divides* A. *Then*

$$\sum_{\substack{d \mid A\\ d \equiv 0 \bmod l}} \mu(d)\left\{w(1) - \sum_{p \mid d} w(p)\right\} = \mu(l)\Lambda\left\{A/l, w(1) - \sum_{p \mid l} w(p)\right\}. \tag{2.2}$$

Proof. Because A is squarefree the sum on the left of (2.2) becomes (on setting $d = ld_1$)

$$\mu(l)\sum_{d_1 \mid (A/l)} \mu(d_1)\left\{w(1) - \sum_{p \mid l} w(p) - \sum_{p \mid d_1} w(p)\right\}$$

$$= \mu(l)\left\{w(1) - \sum_{p \mid l} w(p)\right\}\sum_{d_1 \mid (A/l)} \mu(d_1) - \mu(l)\sum_{p \mid (A/l)} w(p)\mu(p)\sum_{d_2 \mid (A/(lp))} \mu(d_2),$$

and (2.2) follows by the characteristic property of the Möbius function μ.

The next piece of notation will be standard in this memoir. For each squarefree d write

$$d = p_1 p_2 \ldots p_t; \quad p_1 > p_2 > \ldots > p_t, \tag{2.3}$$

where p_i is prime. For $i = 1, 2, \ldots, I$ let $B_{2i} = B_{2i}(p_1, \ldots, p_{2i-1})$ be, for the moment, arbitrary. Let $\mathscr{B}$ be the set of those d such that

$$p_{2i} \leqslant B_{2i} \quad \text{if} \quad i \leqslant I \text{ and } t \geqslant 2i. \tag{2.4}$$

Following Brun, consider

$$\mu_{\mathscr{B}}(d) = \begin{cases} \mu(d) & \text{if } d \in \mathscr{B}, \\ 0 & \text{otherwise}. \end{cases} \tag{2.5}$$

Let $\mathscr{S}_{2j}$ be the set of those squarefrees d that have

$$p_{2j} > B_{2j}; \quad p_{2i} \leqslant B_{2i} \quad \text{when} \quad 1 \leqslant i < j$$

(so that the label t in (2.3) satisfies $t \geqslant 2j$). Conventionally, $\mathscr{S}_0$ is the set of all squarefrees and $\mathscr{S}_{2(I+1)}$ is the set of squarefrees having $p_{2i} \leqslant B_{2i}$ for $1 \leqslant i \leqslant I$, any reference to $B_{2(I+1)}$ being vacuous. Then Brun's methods rest on the observations:

the sets $\mathscr{S}_2, \ldots, \mathscr{S}_{2(I+1)}$ are disjoint; $\mathscr{S}_0$ is the union of sets $\mathscr{S}_0 = \mathscr{B} \cup \cup \mathscr{S}_2 \cup \ldots \cup \mathscr{S}_{2(I+1)}$.

This is employed in the form

$$\sum_{\substack{d \in \mathscr{B} \\ d | A}} \xi(d) = \sum_{\substack{d \in \mathscr{S}_0 \\ d | A}} \xi(d) - \sum_{1 \leqslant j \leqslant I+1} \sum_{\substack{d \in \mathscr{S}_{2j} \\ d | A}} \xi(d) \tag{2.6}$$

where ξ was the Möbius function μ in Brun's work, but in ours will be a weighted analogue, of the type appearing in Lemma 2.1:

$$\xi(d) = \mu(d) \Big\{ w(1) - \sum_{p | d} w(p) \Big\}.$$

Lemma 2.2. *Define*

$$\Sigma_{\mathscr{B}}(A, w) = \sum_{d | A} \mu_{\mathscr{B}}(d) \Big\{ w(1) - \sum_{p | d} w(p) \Big\}. \tag{2.7}$$

Then, with A as in Lemma 2.1, and using the notation introduced above, we have

$$\Sigma_{\mathscr{B}}(A, w) = \Lambda\{A, w(1)\} - \sum_{1 \leqslant j \leqslant I+1} \sum_{\substack{p_1 p_2 \ldots p_{2j} \in \mathscr{S}_{2j} \\ p_1 p_2 \ldots p_{2j} | A}} \Lambda \Big\{ \big(A, \Pi(p_j)\big), w(1) - \sum_{1 \leqslant i \leqslant 2j} w(p_i) \Big\}, \tag{2.8}$$

where

$$\Pi(u) = \prod_{p < u} p \tag{2.9}$$

denotes the product of all primes strictly less than u.

Proof. The d in $\mathscr{S}_{2j}$ are precisely those expressible as

$$d = p_1 p_2 \ldots p_{2j} d'; \quad d' | \Pi(p_{2j})$$

with the product $p_1 p_2 \ldots p_{2j}$ in $\mathscr{S}_{2j}$. Now use Lemma 2.1, with $l = p_1 \ldots p_{2j}$ and A replaced by

$$A_1 = \big(A, p_1 \ldots p_{2j} \Pi(p_{2j})\big),$$

so that $A_1/l = \big(A, \Pi(p_{2j})\big)$. This gives

$$\sum_{\substack{d \in \mathscr{S}_{2j} \\ d | A}} \mu(d) \Big\{ w(1) - \sum_{p | d} w(p) \Big\} = \sum_{\substack{p_1 p_2 \ldots p_{2j} \in \mathscr{S}_{2j} \\ p_1 p_2 \ldots p_{2j} | A}} \Lambda \Big\{ \big(A, \Pi(p_j)\big), w(1) - \sum_{1 \leqslant i \leqslant 2j} w(p_i) \Big\},$$

and (2.6) now yields (2.8).

3. The inequalities of the weighted sieve. It will become clear in this section why it is desirable to specify the function w as was done in § 1. Meantime the reader may, if he prefers, regard this choice as not yet having been made. The assumptions (A_1) and (A_2), where relevant, are supposed to hold.

We shall use the identities of § 2, with

$$A = \big(a, \Pi(y)\big), \tag{3.1}$$

where $\Pi(y)$ is as in (2.9) and y is as in (A_2).

Our object is to choose the functions B_{2i} and w in such a way that

$$\text{if } \mu_{\mathscr{B}}(d) \neq 0 \text{ then } d \leqslant y, \tag{3.2}$$

$$\text{if } \Sigma_{\mathscr{B}}(A, w) > 0 \text{ then } \nu_{y,U}(a) \leqslant R, \tag{3.3}$$

where $\nu_{y,U}$ is as in (1.14), and

$$S_{\mathscr{B}}(y, w) > 0. \tag{3.4}$$

Here $\Sigma_{\mathscr{B}}(A, w)$ is as in (2.7) and $S_{\mathscr{B}}(y, w)$ is the "main term"

$$S_{\mathscr{B}}(y, w) = \sum_{d \leqslant y} \mu_{\mathscr{B}}(d) \frac{\gamma(d)}{d} \Big\{ w(1) - \sum_{p | d} w(p) \Big\}. \tag{3.5}$$

The rôle of U in (3.3) is simply that $w(p) = w(1)$ if $y^U \leqslant p \leqslant y$.

In this paper we follow Rosser's (see [10]) and Iwaniec's [4] work on the unweighted sieve by specifying in (2.4)

$$p_{2i} \leqslant B_{2i} \quad \text{if and only if} \quad p_{2i}^3 p_{2i-1} \ldots p_1 \leqslant y. \tag{3.6}$$

It then follows easily by (2.5) that (3.2) is satisfied.

In the current section we describe some choices of w that satisfy (3.3) for a suitable integer R, when a satisfies (A_2).

To satisfy the conclusion of (3.3) we might seek, via (2.8), a suitable lower bound for $\Lambda\{A, w(1)\}$, Λ being defined in (2.1). To this end we describe conditions (Lemma 3.2) which guarantee that the sum of the other Λ terms in (2.8) is not negative, except in certain circumstances when (Lemma 3.1) the desired conclusion $\nu_{y,U}(a) \leqslant R$ follows for other reasons.

To this end we consider only functions w satisfying

(3.7) $$w(1) \geqslant 0; \quad w(p) \geqslant 0 \text{ if } p \leqslant y,$$

(3.8) for each $j \geqslant 2$ the condition $p_{2i} \leqslant B_{2i}$ whenever $1 \leqslant i \leqslant j-1$ implies

$$\sum_{i=1}^{2j} w(p_i) \leqslant w(1).$$

Because of (2.1) the only Λ in (2.8) that can now assume negative values are those with the label $j = 1$.

We shall achieve control over these possibly negative values by borrowing an idea from [9]. Suppose in addition

(3.9) for primes p the function $w(p)$ is non-decreasing,

(3.10) U, V satisfy (1.6), (1.7),

(3.11) $$w(p) \leqslant \log p/\log y - V \quad \text{if} \quad \log p/\log y \leqslant U,$$

(3.12) $$w(p) = w(1) \quad \text{if} \quad U < \log p/\log y \leqslant 1,$$

where we normalise (as in § 1) to

(3.13) $$w(1) = U - V,$$

consistent with (3.7) because of (1.6). Note the consequence

(3.14) $$w(p) \leqslant w(1) \quad \text{for all } p \leqslant y.$$

LEMMA 3.1. *Suppose that* (3.1), (3.6) *and the seven conditions* (3.7)–(3.13) *hold. Then*

(3.15) *the statement* (3.3) *holds,*

(3.16) *again under the hypothesis in* (3.3) *we have*

$$\Sigma_{\mathscr{B}}(A, w) \leqslant w(q_a),$$

where q_a *is as in Theorem* 1 *and* $\Sigma_{\mathscr{B}}$ *is as in Lemma* 2.2.

Proof. The easy case is when $\Lambda\{A, w(1)\} > 0$. Then from (2.1), (3.1), (3.11) the least prime factor q_a of a satisfies $q_a > y^V$ (this being trivially true if $V < 0$) and all other prime factors of a satisfy $p \geqslant y^U$. If there were R or more of these others, multiple factors being counted multiply, then we should have $a \geqslant y^{RU+V}$ contrary to (A_2), (3.10). Thus $\nu_{y,U}(a) \leqslant R$ as required by (3.15). For (3.16) observe that, in the present case, (2.8) reduces to

$$\Sigma_{\mathscr{B}}(A, w) = \Lambda\{A, w(1)\},$$

while, by (2.1), (3.9), (3.12),

$$\Lambda\{A, w(1)\} \leqslant w(q_a).$$

Now consider the contrary case when

(3.17) $$\Lambda\{A, w(1)\} = 0.$$

Because of (3.8), (2.1) the identity (2.8) gives

(3.18) $$\Sigma_{\mathscr{B}}(A, w) \leqslant -\sum_{\substack{p_1p_2|A \\ p_2<p_1 \\ y<p_2^3p_1}} \Lambda\{(A, \Pi(p_2)), w(1)-w(p_1)-w(p_2)\}$$

where we have explicitly written the consequence of the definitions (3.6).

Let $Q_1 < Q_2 < \ldots < Q_r$ be the prime factors of A. Because of (2.1) and the hypothesis in (3.3), viz.

$$\Sigma_{\mathscr{B}}(A, w) > 0,$$

the inequality (3.18) requires that at least one value of Λ in the sum is $w(1)-w(p_2)-w(p_1)$, where necessarily $p_2 = Q_1$. If now $p_1 = Q_i$ for $i > 2$ then this value of Λ may be paired with one arising from $p_1 = Q_i$, $p_2 = Q_2$; this latter value, by (2.1), is $w(Q_1)$. Then the sum of these values is expressible as $w(1)-w(p_1)$, which is not negative because of (3.1), (3.14). This proves that since $\Sigma_{\mathscr{B}}(A, w) > 0$ there is in (3.18) a summand with $p_1 = Q_2$, $p_2 = Q_1$. Consequently

$$y < Q_1^3 Q_2,$$

so that $y < p_2^3 p_1$ whenever $p_2 < p_1$, $p_1 p_2 | A$. Thus the condition $y < p_2^3 p_1$ may be omitted from the conditions of summation in (3.18).

Now it follows that

(3.19) $$\Sigma_{\mathscr{B}}(A, w) \leqslant -\sum_{\substack{p_1p_2|A \\ p_2<p_1}} \Lambda\{(A, \Pi(p_2)), w(1)-w(p_1)-w(p_2)\}$$
$$= w(1) - \sum_{p|A} \{w(1)-w(p)\},$$

the last equality being that special case of (2.8) in which $B_2 = 0$, because of (3.17).

Hence, because of (3.19), (3.11), (3.13),

$$\sum_{p|a} \{U - \log p/\log y\} \leqslant U - V,$$

where we may extend the summation over those $p > y^U$, taking account even of the multiplicity with which they divide a, because the added summands are negative. Thus, using (A_2), we find

$$U - V \geqslant U\nu_{y,U}(a) - g,$$

with $\nu_{y,U}$ as in (1.14). If now $\nu_{y,U}(a) > R$ we should obtain

$$U - V \geqslant U(R+1) - g,$$

a contradiction of (3.10).

It remains to prove (3.16) when $\Lambda\{A, w(1)\} = 0$. Again we use (3.19). The right side is

$$w(q_a) - \sum_{\substack{q_a < p \\ p|A}} \{w(1) - w(p)\} \leqslant w(q_a)$$

because of (3.14). This completes the proof of Lemma 3.1.

The next lemma shows how to satisfy a key hypothesis of Lemma 3.1.

LEMMA 3.2. *Suppose an increasing function* m *satisfies*

(3.20) $\qquad 0 \leqslant m(\lambda t) \leqslant \lambda m(t)$ *when* $0 \leqslant \lambda \leqslant 1,\ 0 \leqslant t \leqslant \frac{1}{4}$,

(3.21) $\qquad m(\frac{1}{4}) = \frac{1}{4}$,

and that for $\frac{1}{4} < t \leqslant 1$ *the value* $m(t)$ *is then specified by*

(3.22) $\qquad m(\frac{1}{4} + 3s) - \frac{1}{4} = 3\{\frac{1}{4} - m(\frac{1}{4} - s)\} \quad (0 < s \leqslant \frac{1}{4})$.

Suppose the function w *satisfies* (1.9) *with*

(3.23) $\qquad W(t) \leqslant W(1)m(t) \quad (0 \leqslant t \leqslant 1)$.

Then (3.8) *holds.*

Proof. Of the conditions granted in (3.8) we need use only $p_2 \leqslant B_2$ and $p_{2(j-1)} \leqslant B_{2(j-1)}$ (which need not be distinct). By the choice (3.6) of B these imply

(3.24) $\qquad p_2^3 p_1 \leqslant y, \quad p_2 \leqslant y^{1/4}, \quad p_1 p_2 \ldots p_{2j} \leqslant y,$

where $p_1 > p_2 > \ldots > p_{2j}$ as in (2.3).

It may happen that $p_1 \leqslant y^{1/4}$. Now (3.20), (3.21), (3.23) imply $W(t) \leqslant tW(1)$ when $0 \leqslant t \leqslant \frac{1}{4}$. Then the last of the conditions (3.24) yields the conclusions required in (3.8).

Let us write $\log p_i/\log y = t_i$. In the remaining case when $y^{1/4} < p_i \leqslant y$ write $t_1 = \frac{1}{4} + 3s_1$. Then the first condition of (3.24) says $3t_2 + \frac{1}{4} + 3s_1 \leqslant 1$, i.e. $t_2 \leqslant \frac{1}{4} - s_1$. By (3.20) we now have when $2 \leqslant i \leqslant 2j$ that

$$m(t_i) \leqslant t_i \frac{m(\frac{1}{4} - s_1)}{\frac{1}{4} - s_1},$$

while (3.24) gives

$$t_2 + t_3 + \ldots + t_{2j} \leqslant 1 - t_1 = 3(\tfrac{1}{4} - s_1).$$

Consequently

$$m(t_2) + m(t_3) + \ldots + m(t_{2j}) \leqslant 3m(\tfrac{1}{4} - s_1),$$

while by (3.22)

$$3m(\tfrac{1}{4} - s_1) = 1 - m(\tfrac{1}{4} + 3s_1) = 1 - m(t_1),$$

so that the conclusion required in (3.8) again holds.

We now specialise our function w to that described in § 1.

LEMMA 3.3. *Let* W *be as described in* (1.10), (1.11), (1.12), (1.13). *Then the hypotheses* (*and conclusions*) *of Lemmas* 3.1 *and* 3.2 *hold.*

Proof. Deal with Lemma 3.2 first. With $\frac{1}{2} < U < 1$ as in (1.6) the function $m(t)$ will be given by

$$m(t) = \frac{t - V^*}{U - V^*} \text{ if } \tfrac{1}{4} < t < U; \quad m(t) = \frac{t - (1-U)/3}{U - V^*} \text{ if } \frac{1-U}{3} < t < \frac{1}{4};$$

$$m(t) = 1 \text{ if } U \leqslant t \leqslant 1; \quad m(\tfrac{1}{4}) = \tfrac{1}{4}; \quad m(t) = 0 \text{ if } 0 < t < (1-U)/3,$$

where $3V^* + U \leqslant 1$, so that

$$m(\tfrac{1}{4} - 0) = \frac{U/3 - 1/12}{U - V^*} \leqslant \frac{U - \frac{1}{4}}{3U - (1-U)} = \frac{1}{4}.$$

It is straightforward to verify that (3.20), (3.21), (3.22) hold.

When $3V + U \leqslant 1$ take $V^* = V$. Then (3.23) follows directly from (1.11).

When $3V + U > 1$ take V^* so that $3V^* + U = 1$, so that $V > V^* > 0$. Now $m(t) = (t - V^*)/(U - V^*)$ whenever $V^* \leqslant t \leqslant U$, and $m(\frac{1}{4}) = \frac{1}{4}$. Then (3.20), (3.21), (3.22) again follow straightforwardly and (3.23) follows from (1.13).

It remains to verify that the hypotheses of Lemma 3.1 hold. Of these, (3.1), (3.6) are definitions and (3.8) is the result of Lemma 3.2. The remaining conditions (3.7), (3.9)–(3.13) follow at once from (1.11) and (1.13).

4. Analysis of the main term. We proceed by applying results from the theory of the unweighted sieve (see [4], [5]) to the "main term"

$S_{\mathscr{B}}(y, w)$ defined in (3.5). Of this, the "unweighted part" is

$$S^-(y, 1) = \sum_d \mu_{\mathscr{B}}(d)\gamma(d)/d, \tag{4.1}$$

where $\mu_{\mathscr{B}}$ is as in (2.5) and we denote

$$S^-(z, s) = \sum_{t\geqslant 0} \sum_{\substack{p_t<p_{t-1}<\ldots<p_1<z^{1/s} \\ p_1\ldots p_{2j-1}p_{2j}^3\leqslant z \text{ if } 2\leqslant 2j\leqslant t}} \mu(d)\gamma(d)/d,$$

$$S^+(z, s) = \sum_{t\geqslant 0} \sum_{\substack{p_t<p_{t-1}<\ldots<p_1<z^{1/s} \\ p_1\ldots p_{2j-2}p_{2j-1}^3\leqslant z \text{ if } 1\leqslant 2j-1\leqslant t}} \mu(d)\gamma(d)/d,$$

when $z\geqslant 2$, $s\geqslant 1$, d being expressed as in (2.3). The specification of $\mathscr{B}$ was given in (3.6).

LEMMA 4.1. *The main term satisfies*

$$S_{\mathscr{B}}(y, w) = w(1)S^-(y, 2) - \sum_{y^{1/2}\leqslant p<y} \{w(1)-w(p)\}\frac{\gamma(p)}{p} S^+\left(\frac{y}{p}, \frac{\log(y/p)}{\log p}\right) + \sum_{p<y^{1/2}} w(p)\sum_{r\geqslant 1}(-1)^{r+1}\Sigma_r(y, p), \tag{4.2}$$

where

$$\Sigma_r(y, p) = \sum_{\substack{p=p_r<p_{r-1}<\ldots<p_1<y \\ p_1\ldots p_{2j-1}p_{2j}^3<y \text{ if } 2\leqslant 2j\leqslant 2r}} \frac{\gamma(p_1\ldots p_r)}{p_1\ldots p_r} \times S^{(-)^r}\left(\frac{y}{p_1\ldots p_r}, \frac{\log\{y/(p_1\ldots p_r)\}}{\log p_r}\right). \tag{4.3}$$

Proof. Use the recurrences

$$S^-(z, s) = 1 - \sum_{p_1<z^{1/s}} \frac{\gamma(p_1)}{p_1} S^+\left(\frac{z}{p_1}, \frac{\log(z/p_1)}{\log p_1}\right), \tag{4.4}$$

$$S^+(z, s) = 1 - \sum_{\substack{p_1<z^{1/s} \\ p_1^3\leqslant z}} \frac{\gamma(p_1)}{p_1} S^-\left(\frac{z}{p_1}, \frac{\log(z/p_1)}{\log p_1}\right) \tag{4.5}$$

for $z\geqslant 2$, $s\geqslant 1$. These give firstly

$$S^-(y, 1) = S^-(y, 2) - \sum_{y^{1/2}\leqslant p_1<y} \frac{\gamma(p_1)}{p_1} S^+\left(\frac{y}{p_1}, \frac{\log(y/p_1)}{\log p_1}\right).$$

Hence the coefficient of $w(1)$ in (4.2) is as stated.

The conditions of summation in (3.5) give (see (3.24)) that $p_j\geqslant y^{1/2}$ implies $p_j = p_1$. Now multiple application of the recurrences (4.4), (4.5) gives that the coefficient of $w(p)$ in (4.2) is also as stated.

We shall use the following estimates of $S^{\pm}(z, s)$.

LEMMA 4.2. *For $s\geqslant 2$, $z\geqslant 2$ we have because of* (A_1) *that*

$$S^{(-)^r}(z, s) = \left\{\prod_{p<z^{1/s}}\left(1-\frac{\gamma(p)}{p}\right)\right\}\left\{G^{(-)^r}(s) + O\left(\frac{L^{1/5}}{\log^{1/5} z}\right)\right\},$$

where $G^+(s) = F(s)$ and $G^-(s) = f(s)$ are as described below. The estimate for $S^+(z, s)$ holds good if $s\geqslant 1$.

Here the functions F, f satisfy the well-known (see [6], [3]) difference-differential equations

$$\frac{d}{ds}\{sF(s)\} = f(s-1) \quad (s\geqslant 1); \quad \frac{d}{ds}\{sf(s)\} = F(s-1) \quad (s\geqslant 2), \tag{4.6}$$

with boundary conditions

$$F(s) = 1+O(e^{-s}), \quad f(s) = 1+O(e^{-s}) \quad (s\geqslant 1). \tag{4.7}$$

A superficially rather special case of Lemma 4.2 would be a consequence of the principal theorem of [4] on the unweighted sieve, where some trouble was taken to obtain an error term considerably superior to that quoted above. As stated, the result can be established by the somewhat simpler technique described, also by Iwaniec, in the "half-dimensional" context in [5].

The paper [4] describes how techniques introduced by de Bruijn [2] may be used to infer

$$sF(s) = 2e^{\gamma_0} \quad (1\leqslant s\leqslant 3), \tag{4.8}$$

$$sf(s) = 2e^{\gamma_0}\log(s-1) \quad (2\leqslant s\leqslant 4), \tag{4.9}$$

γ_0 being Euler's constant. We shall again use de Bruijn's ideas in our proof of Theorem 2.

In what follows we take advantage of the fact that the weight function $w(p)$ specified in (1.9) has the property

$$w(p) = 0 \quad \text{if} \quad p<y^{T_0}, \tag{4.10}$$

where T_0 is the absolute constant specified in (1.22). This implies that in (4.2) the suffix r is bounded:

$$r = O(1). \tag{4.11}$$

Lemma 4.2 gives the following estimate for the quantity in (4.3):

$$\Sigma_r(y,p) = \Big\{\prod_{q<p}\Big(1-\frac{\gamma(q)}{q}\Big)\Big\} \sum_{\substack{p=p_r<p_{r-1}<\ldots<p_1 \\ p_1\cdots p_{2j-1}p_{2j}^3\leqslant y \text{ if } 2\leqslant 2j\leqslant r}} \frac{\gamma(p_1\cdots p_r)}{p_1\cdots p_r} \times \tag{4.12}$$

$$\times\Big\{G^{(-)^r}\Big(\frac{\log\{y/(p_1\cdots p_r)\}}{\log p}\Big) + O\Big(\frac{L^{1/5}}{\log^{1/5} p}\Big)\Big\}.$$

For the conditions of summation imply (via (3.24)) that

$$y/(p_1\cdots p_r) \geqslant p,$$

so that the O-term is indeed as stated.

To make progress we deal with

$$h_r(p,z,s,G) = \sum_{\substack{p=p_r<p_{r-1}<\ldots<p_1<z^{1/s} \\ p_1\cdots p_{2j-1}p_{2j}^3\leqslant y \text{ if } 2\leqslant 2j\leqslant r}} \frac{\gamma(p_1\cdots p_r)}{p_1\cdots p_r} G\Big(\frac{\log\{y/(p_1\cdots p_r)\}}{\log p_r}\Big),$$

$$H_r(p,z,s,G) = \sum_{\substack{p=p_r<\ldots<p_1<z^{1/s} \\ p_1\cdots p_{2j-2}p_{2j-1}^3\leqslant y \text{ if } 1\leqslant 2j-1\leqslant r}} \frac{\gamma(p_1\cdots p_r)}{p_1\cdots p_r} G\Big(\frac{\log\{y/(p_1\cdots p_r)\}}{\log p_r}\Big) \tag{4.13}$$

for $z\geqslant 2$, $s\geqslant 1$, $r\geqslant 1$, where the function G may be $G^+ = F$ or $G^- = f$ (see (4.6)) or $G(x) = 1$ for all x. In each case these satisfy the recurrences

$$h_r(p,z,s,G) = \sum_{p<p_1<z^{1/s}} \frac{\gamma(p_1)}{p_1} H_{r-1}\Big(p, \frac{z}{p_1}, \frac{\log z}{\log p_1}-1, G\Big),$$

$$H_r(p,z,s,G) = \sum_{\substack{p<p_1<z^{1/s} \\ p_1^3\leqslant z}} \frac{\gamma(p_1)}{p_1} h_{r-1}\Big(p, \frac{z}{p_1}, \frac{\log z}{\log p_1}-1, G\Big), \tag{4.14}$$

for $z\geqslant 2$, $s\geqslant 1$, $r\geqslant 1$.

We estimate these quantities in terms of

$$k_r(x,s,G) = \int\limits_{\substack{x=x_r<x_{r-1}<\ldots<x_1<1/s \\ x_1+\ldots+x_{2j-1}+3x_{2j}\leqslant 1 \text{ if } 2\leqslant 2j\leqslant r}} \!\!\cdots\!\! \int G\Big(\frac{1-x-x_{r-1}-\ldots-x_1}{x}\Big)\prod_{i=1}^{r-1}\frac{dx_i}{x_i},$$

$$K_r(x,s,G) = \int\limits_{\substack{x=x_r<x_{r-1}<\ldots<x_1<1/s \\ x_1+\ldots+x_{2j-2}+3x_{2j-1}<1 \text{ if } 1\leqslant 2j-1\leqslant r}} \!\!\cdots\!\! \int G\Big(\frac{1-x-x_{r-1}-\ldots-x_1}{x}\Big)\prod_{i=1}^{r-1}\frac{dx_i}{x_i}, \tag{4.15}$$

where $0<x\leqslant 1\leqslant s$. These satisfy the recurrences

$$k_r(x,s,G) = \int\limits_{x<x_1<1/s} K_{r-1}\Big(\frac{x}{1-x_1}, \frac{1}{x_1}-1, G\Big)\frac{dx_1}{x_1},$$

$$K_r(x,s,G) = \int\limits_{\substack{x<x_1<1/s \\ 3x_1\leqslant 1}} k_{r-1}\Big(\frac{x}{1-x_1}, \frac{1}{x_1}-1, G\Big)\frac{dx_1}{x_1}, \tag{4.16}$$

when $0<x\leqslant 1\leqslant s$.

It is immediate that for the functions G specified at (4.13) we have

$$0\leqslant k_r(x,s,G)\leqslant 1,\quad 0\leqslant K_r(x,s,G)\leqslant 1 \quad \text{if} \quad 0<x\leqslant 1\leqslant s. \tag{4.17}$$

We shall often use partial summation in the following form (cf. Lemma 8 of [5]):

Lemma 4.3. *Let $B(x)$ be positive, continuous and monotone in the interval $w\leqslant x\leqslant z$, where $w\geqslant 2$. Then, because of* (A_1),

$$\frac{-LB_0}{\log w} \leqslant \sum_{w\leqslant p<z}\frac{\gamma(p)}{p}B\Big(\frac{\log p}{\log y}\Big) - \int\limits_{\log w/\log y}^{\log z/\log y} B(l)\frac{dl}{l} \leqslant \frac{A_1 B_0}{\log z},$$

where

$$B_0 = \max\Big\{B\Big(\frac{\log z}{\log y}\Big), B\Big(\frac{\log w}{\log y}\Big)\Big\}.$$

As in § 5 of [5] it also follows from (A_1) that

$$\prod_{w\leqslant q<z}\Big(1-\frac{\gamma(q)}{q}\Big)^{-1} \leqslant \frac{\log z}{\log w}\Big\{1+O\Big(\frac{1}{\log w}\Big)\Big\}. \tag{4.18}$$

The proof of the next lemma is by induction on r.

Lemma 4.4. *With h_r, H_r, k_r, K_r as defined in* (4.13), (4.15) *we have for $z\geqslant 2$, $s\geqslant 1$ that*

$$h_r(p,z,s,G) = \frac{\gamma(p)}{p}\Big\{k_r\Big(\frac{\log p}{\log z}, s, G\Big) + O\Big(\frac{L}{\log p}\Big)\Big\},$$

$$H_r(p,z,s,G) = \frac{\gamma(p)}{p}\Big\{K_r\Big(\frac{\log p}{\log z}, s, G\Big) + O\Big(\frac{L}{\log p}\Big)\Big\}.$$

Proof. When $r=1$ the result is trivially true since for example (4.13), (4.15) give

$$h_1(p,z,s,G) = \frac{\gamma(p)}{p}G\Big(\frac{\log(y/p)}{\log p}\Big),\qquad k_1(x,s,G) = G\Big(\frac{1-x}{x}\Big).$$

Also, use of the inductive hypothesis and partial summation gives, from (4.14),

$$h_r(p, z, s, G) = \sum_{p<p_1<z^{1/s}} \frac{\gamma(p_1)}{p_1} K_{r-1}\left(\frac{\log p}{\log z - \log p_1}, \frac{\log z}{\log p_1} - 1, G\right) + O\left(\frac{L}{\log p}\right)$$
$$= \int_{x_1<1/s} K_{r-1}\left(\frac{\log p/\log z}{1-x_1}, \frac{1}{x_1} - 1, G\right)\frac{dx_1}{x_1} + O\left(\frac{L}{\log p}\right),$$

where we have used (4.17), Lemma 4.3, and (for the O-term) the fact that

$$\sum_{p<p_1<y^{1/U}} \frac{\gamma(p_1)}{p_1} = O(1) \tag{4.19}$$

because of (4.10).

The other half of the inductive step follows similarly.

Now proceed to the last result of this section.

LEMMA 4.5. *The expression in* (3.5) *satisfies*

$$S_{\mathscr{A}}(y, w) = \left\{\prod_{q<y}\left(1-\frac{\gamma(q)}{q}\right)\right\}\log y\left\{-\sum_{y^{1/2}<p\leqslant y^U} \frac{w(1)-w(p)}{p\log p}\gamma(p)F\left(\frac{\log(y/p)}{\log p}\right) + \right.$$
$$\left. + \sum_{p\leqslant y^{1/2}} \frac{w(p)\gamma(p)}{p\log p}\sum_r (-1)^{r+1}k_r\left(\frac{\log p}{\log y}, 1, G^{(-)^r}\right) + O\left(\frac{w(1)}{\log y}\right)\right\}.$$

Proof. For the first term in (4.2), Lemma 4.2 and (4.9) give

$$S^-(y, 2) = O\left\{\prod_{q<y}\left(1-\frac{\gamma(q)}{q}\right)L^{1/5}/\log^{1/5}y\right\}, \tag{4.20}$$

where we have used (4.18).

For the second term we similarly find, when $y^{1/2}<p_1\leqslant y^U$, that

$$S^+\left(\frac{y}{p_1}, \frac{\log(y/p_1)}{\log p_1}\right) = \left\{\prod_{q<p}\left(1-\frac{\gamma(q)}{q}\right)\right\}\left\{F\left(\frac{\log(y/p_1)}{\log p_1}\right) + O\left(\frac{L^{1/5}}{\log^{1/5}y}\right)\right\}$$
$$= \left\{\prod_{q<y}\left(1-\frac{\gamma(q)}{q}\right)\right\}\frac{\log y}{\log p_1}\left\{F\left(\frac{\log(y/p_1)}{\log p_1}\right) + O\left(\frac{L^{1/5}}{\log^{1/5}y}\right)\right\}$$

the O-terms being as stated because (see (1.6)) $U<1$ is an absolute constant. Consequently

$$\sum_{y^{1/2}<p_1\leqslant y^U} \{w(1)-w(p_1)\}\frac{\gamma(p_1)}{p_1}S^+\left(\frac{y}{p_1}, \frac{\log(y/p_1)}{\log p_1}\right) \tag{4.21}$$
$$= \left\{\prod_{q<y}\left(1-\frac{\gamma(q)}{q}\right)\right\}\left\{\sum_{y^{1/2}<p_1\leqslant y^U}\frac{\log y}{\log p_1}\,\frac{w(1)-w(p_1)}{p_1}\gamma(p_1)F\left(\frac{\log(y/p_1)}{\log p_1}\right)\right\} +$$
$$+ O\left(\frac{w(1)L^{1/5}}{\log^{1/5}y}\right),$$

where at the last step (4.19) has again been used.

For the remaining term in (4.2) we have, from (4.12) and Lemma 4.4,

$$\Sigma_r(y, p) = \left\{\prod_{q<p}\left(1-\frac{\gamma(q)}{q}\right)\right\}\left\{k_r\left(\frac{\log y}{\log p}, 1, G^{(-)^r}\right) + O\left(\frac{L^{1/5}}{\log^{1/5}y}\right) + O\left(\frac{L}{\log p}\right)\right\},$$

where (4.17) was used to estimate the O-term.

Consequently, and by (4.10), (4.11),

$$\sum_{p\leqslant y^{1/2}} w(p)\sum_{r\geqslant 1}(-1)^{r+1}\Sigma_r(y, p) \tag{4.22}$$
$$= \left\{\prod_{q<y}\left(1-\frac{\gamma(q)}{q}\right)\right\}\sum_{p<y^{1/2}}\frac{\log y}{\log p}\,\frac{w(p)\gamma(p)}{p} \times$$
$$\times\left\{\sum_r(-1)^{r+1}k_r\left(\frac{\log p}{\log y}, 1, G^{(-)^r}\right) + O\left(\frac{L^{1/5}}{\log^{1/5}y}\right)\right\},$$

where we have used $L\leqslant\log y$ as given in (A_2).

The result of the lemma now follows from (4.20), (4.21), (4.22) and Lemma 4.1, on again invoking (4.17).

5. An identity for the leading term. This identity, which appears in Lemma 5.2, involves a quantity arising (see Lemma 6.1) from the sums over primes p appearing in the estimate (Lemma 4.5) for the "main term" $S_{\mathscr{A}}(y, w)$. Introduce the abbreviation

$$k_r(x, G^{(-)^r}) = k_r(x, 1, G^{(-)^r}) \tag{5.1}$$

for the quantity defined in (4.15) and appearing in Lemma 4.5.

The following representation of the functions $G^{\pm}$ is very convenient.

LEMMA 5.1. *Let* $G^+ = F$ *and* $G^- = f$ *be defined* (*as usual*) *by* (4.6), (4.7). *Define*

$$I_r(s) = \int\limits_{\substack{1<u_r<u_{r-1}<\ldots<u_1\\ u_r+u_{r-1}+\ldots+u_1=s}}\cdots\int \frac{du_1\ldots du_{r-1}}{u_1\ldots u_{r-1}u_r} \quad \textit{if} \quad s\geqslant 1,\ r\geqslant 1.$$

Then

$$F(s) = 2e^{\gamma_0} \sum_{r\ \mathrm{odd}} I_r(s) \quad (s \geqslant 1),$$

$$f(s) = 2e^{\gamma_0} \sum_{r\ \mathrm{even}} I_r(s) \quad (s \geqslant 2),$$

where γ_0 *is Euler's constant (as in* (4.8)).

Proof. By the change of variable $u_i = sv_i$ we have

$$sI_r(s) = \int\limits_{\substack{1/s<v_r<\ldots<v_1\\ v_r+\ldots+v_1=1}} \ldots \int \frac{dv_1 \ldots dv_{r-1}}{v_1 \ldots v_{r-1}v_r}.$$

Hence, for $r \geqslant 2$,

$$\frac{d}{ds} sI_r(s) = \frac{1}{s} \int\limits_{\substack{1/s<v_{r-1}<\ldots<v_1\\ v_r+\ldots+v_1=1-1/s}} \ldots \int \frac{dv_1 \ldots dv_{r-2}}{v_1 \ldots v_{r-2}v_{r-1}}$$

$$= \frac{1}{s-1} \int\limits_{\substack{1/(s-1)<z_{r-1}<\ldots<z_1\\ z_{r-1}+\ldots+z_1=1}} \ldots \int \frac{dz_1 \ldots dz_{r-2}}{z_1 \ldots z_{r-2}z_{r-1}}$$

$$= I_r(s-1)$$

if $s > 1$, where we set $sv_i = (s-1)z_i$. But $I_1(s) = 1/s$ for $s > 1$ and $I_2(2) = 0$, whence $sI_2(s) = \log(s-1)$ if $s \geqslant 2$. Also $I_r(s) = 0$ if $r \geqslant s$. Hence the expressions of this lemma for F and f satisfy (4.6), (4.8), (4.9), which specify F, f uniquely. This proves Lemma 5.1.

We will define the functions h, H of Theorem 1 as follows. For integers $i \geqslant 1$ define

(5.2) $$h_i(x, s) = \int\limits_{\mathscr{R}_i} \ldots \int \frac{dx_1 \ldots dx_{2k}}{x_1 \ldots x_{2k}u} \quad (0 < x \leqslant 1 \leqslant s)$$

where the sets of conditions $\mathscr{R}_i$ are given by

(5.3) $$\mathscr{R}_{2k}: \begin{cases} x < x_{2k} < \ldots < x_1 < 1/s, \\ 3x_{2j} + x_{2j-1} + \ldots + x_1 \leqslant 1 \text{ if } j < k, \\ 1 < 3x_{2k} + x_{2k-1} + \ldots + x_1, \\ x_{2k} < u, \\ x + x_{2k} + \ldots + x_1 + u = 1; \end{cases}$$

(5.4) $$\mathscr{R}_{2k-1}: \begin{cases} x < x_{2k-1} < \ldots < x_1 < 1/s, \\ 3x_{2j-1} + x_{2j-2} + \ldots + x_1 \leqslant 1 \text{ if } j < k, \\ 1 < 3x_{2k-1} + x_{2k} + \ldots + x_1, \\ x_{2k-1} < u, \\ x + x_{2k-1} + \ldots + x_1 + u = 1. \end{cases}$$

Then, on using the change of variable

$$x_i = (1-x_1)y_i' \quad (i \geqslant 2); \quad u = (1-x_1)u',$$

we see

(5.5) $$h_{2k}(x, s) = \int\limits_{x_1<1/s} \frac{dx_1}{x_1(1-x_1)} h_{2k-1}\left(\frac{x}{1-x_1}, \frac{1-x_1}{x_1}\right)$$

$$= \int\limits_{s<t} \frac{dt}{t-1} h_{2k-1}\left(\frac{tx}{t-1}, t-1\right)$$

if $0 < x \leqslant 1 < s$, and similarly

(5.6) $$h_{2k-1}(x, s) = \int\limits_{s<t;3<t} \frac{dt}{t-1} h_{2k-2}\left(\frac{tx}{t-1}, t-1\right)$$

if $k > 1$ and $0 \leqslant x \leqslant 1$, $s \geqslant 0$, the trivial extension to $0 \leqslant s \leqslant 1$ being made for convenience. Also $h_1(x, s)$ is as stated in Theorem 1. Now specify

(5.7) $$h(x, s) = \sum_{k\geqslant 1} h_{2k}(x, s) \quad (0 < x \leqslant 1 \leqslant s)$$
$$H(x, s) = \sum_{k\geqslant 1} h_{2k-1}(x, s) \quad (0 < x \leqslant 1;\ s \geqslant 0).$$

Since $x > 0$ these summations are finite. Then the relations (1.15), (1.16) stated in Theorem 1 follow at once.

Now proceed to the main result of this section.

Lemma 5.2. *Let* $k_r(x, G^{(-)^r})$ *be as in* (5.1). *Then*

$$\sum_{r\geqslant 1} (-1)^{r+1} k_r(x, G^{(-)^r}) = 2e^{\gamma_0} x \left\{\frac{1}{1-x} - h(x, 1)\right\},$$

where $h(x, 1)$ *is as in Theorem* 1.

Proof. When $\mathscr{D}$ is an appropriate set of inequalities denote by $\{\mathscr{D}\}$ the value of

$$\int\limits_{\mathscr{D}} \frac{1}{x_1 u_1} \prod_{i\geqslant 2} \frac{dx_i}{x_i} \prod_{j\geqslant 2} \frac{du_j}{u_j},$$

where Δ denotes the region described by the inequalities $\mathscr{D}$ together with the equation

$$\sum_{i\geqslant 1} x_i + \sum_{j\geqslant 1} u_j = 1. \tag{5.8}$$

Begin by inserting the representation of Lemma 5.1 into the expression (4.15) for k_r. This shows

$$k_r(x, G^{(-)^r}) = k_r(x, 1, G^{(-)^r})$$

$$= 2e^{\gamma_0} \sum_{\substack{t\geqslant 1\\ t\equiv r \bmod 2}} \int_{\substack{x=x_r<\ldots<x_1\\ 3x_{2j}+x_{2j-1}+\ldots+x_1<1 \text{ if } 2\leqslant 2j\leqslant r}} \cdots \int \prod_{i=1}^{r-1} \frac{dx_i}{x_i} \times$$

$$\times \int_{\substack{1<v_t<\ldots<v_1\\ v_t+\ldots+v_1=(1-x_1-\ldots-x_{r-1}-x)/x}} \cdots \int \frac{1}{v_t} \prod_{i=1}^{t-1} \frac{dv_i}{v_i}$$

$$= 2e^{\gamma_0} x \sum_{\substack{t\geqslant 1\\ t\equiv r \bmod 2}} \begin{Bmatrix} x = x_r < \ldots < x_1 \\ 3x_{2j} + x_{2j-1} + \ldots + x_1 \leqslant 1 \text{ if } 2j \leqslant r \\ x < u_t < \ldots < u_1 \end{Bmatrix},$$

in the notation described above; this follows on substituting $xv_i = u_i$. Next, invoke a combinatorial identity analogous to (2.6). This shows

$$\sum_{r\geqslant 1} (-1)^{r+1} k_r(x, G^{(-)^r}) \tag{5.9}$$

$$= 2e^{\gamma_0} x \sum_{\substack{t\geqslant 1, r\geqslant 1\\ t\equiv r \bmod 2}} (-1)^{r+1} \left[\begin{Bmatrix} x = x_r < \ldots < x_1 \\ x < u_t < \ldots < u_1 \end{Bmatrix} - \sum_k \begin{Bmatrix} x = x_r < \ldots < x_1 \\ x < u_t < \ldots < u_1 \end{Bmatrix}_k\right],$$

where, here and below, the subscript k denotes that the variables in braces $\{\ldots\}_k$ are subject to the further restrictions

$$\begin{aligned} 3x_{2j} + x_{2j-1} + \ldots + x_1 \leqslant 1 \quad &\text{if} \quad j < k, \\ 1 < 3x_{2k} + x_{2k-1} + \ldots + x_1. & \end{aligned} \tag{5.10}$$

Next, employ many times a further identity. This states: for $0 < x \leqslant w < 1$ we have

$$\begin{Bmatrix} x = x_{\alpha+1} < x_\alpha < \ldots < x_1 < w \\ x < u_\beta < \ldots < u_1 < w \end{Bmatrix} \tag{5.11}$$

$$= \binom{\alpha+\beta}{\beta} \{x = x_{\alpha+\beta+1} < x_{\alpha+\beta} < \ldots < x_1 < w\}.$$

To see this, one may observe (for example) that, on the right, β of the variables $x_1, x_2, \ldots, x_{\alpha+\beta}$ may be chosen to be the u_j on the left.

The terms in (5.9) not having a label k thus sum to

$$2e^{\gamma_0} x \sum_{\substack{s\geqslant 1\\ s \text{ odd}}} \sum_{\substack{t+r=s+1\\ t\geqslant 1; r\geqslant 1}} (-1)^{t+1} \binom{s}{t} \{x = x_{s+1} < \ldots < x_1\} \tag{5.12}$$

$$= 2e^{\gamma_0} x \sum_{\substack{s\geqslant 1\\ s \text{ odd}}} \{x = x_{s+1} < \ldots < x_1\}.$$

For the terms in (5.9) having a suffix k observe first that the conditions of summation imply

$$u_2 \leqslant x_{2k}. \tag{5.13}$$

For if $x_{2k} < u_2$ then (because $u_2 < u_1$)

$$1 \leqslant 3x_{2k} + x_{2k-1} + \ldots + x_1 < x_{2k} + \ldots + x_1 + u_2 + u_1 \leqslant \sum x_i + \sum u_j = 1,$$

where (5.8), (5.10) have been used.

Consider first the contribution to (5.9) from the subregion in which the condition

$$u_1 < x_{2k} \tag{5.14}$$

is imposed. For each fixed k this contribution is

$$-2e^{\gamma_0} \sum_{\substack{t\geqslant 1, r\geqslant 1\\ t\equiv r \bmod 2}} (-1)^{r+1} \begin{Bmatrix} x = x_r < \ldots < x_1 \\ x < u_t < \ldots < u_1 < x_{2k} \end{Bmatrix}_k \tag{5.15}$$

$$= -2e^{\gamma_0} x \sum_{\substack{s\geqslant 1\\ s \text{ odd}}} \sum_{\substack{t+r=2k+s+1\\ 1\leqslant t\leqslant s}} (-1)^{t+1} \binom{s}{t} \{x = x_{s+2k+1} < \ldots < x_1\}_k,$$

by the identity (5.11) applied to the variables u_j, x_i $(i > 2k)$. For the restrictions implied by the suffix k apply only to x_i for $i \leqslant 2k$. Computing the alternating sum of binomial coefficients shows that the expression (5.15) is

$$-2e^{\gamma_0} x \sum_{\substack{s\geqslant 1\\ s \text{ odd}}} \{x = x_{s+2k+1} < \ldots < x_1\}_k. \tag{5.16}$$

Similarly the remaining contribution to (5.8), where (5.14) is false, is (because of (5.13))

$$(5.17)\qquad -2e^{\gamma_0}\sum_{\substack{t\geq 1;r\geq 1\\ t\equiv r \bmod 2}}(-1)^{r+1}\begin{Bmatrix}x=x_r<\ldots<x_1\\ x<u_t<\ldots<u_2<x_{2k}<u_1\end{Bmatrix}_k$$

$$= -2e^{\gamma_0}\sum_{\substack{s\geq 1\\ s\text{ odd}}}\sum_{\substack{t+r=2k+s+1\\ 1\leq t}}(-1)^{t+1}\binom{s-1}{t-1}\begin{Bmatrix}x=x_{s+2k}<\ldots<x_1\\ x_{2k}\leq u_1\end{Bmatrix}_k$$

$$= 2e^{\gamma_0}x\begin{Bmatrix}x=x_{2k+1}<\ldots<x_1\\ x_{2k}<u_1\end{Bmatrix}_k,$$

the summation over t this time reducing to 0 except when $s=1$.

The expressions in (5.12), (5.16) sum to

$$(5.18)\qquad 2e^{\gamma_0}x\Big[\sum_{\substack{s\geq 1\\ s\text{ odd}}}\{x=x_{s+1}<\ldots<x_1\}-\sum_k\sum_{\substack{s>2k\\ s\text{ odd}}}\{x=x_{s+1}<\ldots<x_1\}_k\Big].$$

The suffix k has the significance described in (5.10). Hence a further use of the combinatorial principle previously employed at (2.6), (5.9) shows that this expression (5.18) is

$$2e^{\gamma_0}x\sum_{\substack{s\geq 1\\ s\text{ odd}}}\begin{Bmatrix}x=x_{s+1}<\ldots<x_1\\ 3x_{2j}+x_{2j-1}+\ldots+x_1\leq 1 \text{ if } 2j<s\end{Bmatrix}.$$

If $s>1$ the summand is 0 because the inequality with label $2j=s-1$ gives

$$1\geq x_1+\ldots+x_{2j-1}+3x_{2j}>x_1+\ldots+x_{2j}+x_s+x=1.$$

Thus (5.18) reduces, because of (5.8), to

$$2e^{\gamma_0}x\{x=x_2<x_1\}=x/x_1=x/(1-x).$$

This provides the first term on the right of Lemma 5.2. The other is provided by (5.17), since the implied equation and inequalities are as listed in (5.3).

6. Theorem 1 and its Corollaries. The first step is to apply partial summation (Lemma 4.3) to the result of Lemma 4.5, using also Lemma 6.2.

LEMMA 6.1. *Let $S_{\mathscr{B}}(y,w)$ be as in (3.5). Then*

$$S_{\mathscr{B}}(y,w)=2e^{\gamma_0}\Big\{\prod_{q<y}\Big(1-\frac{\gamma(q)}{q}\Big)\Big\}\Big\{\mathscr{M}(W)+O\Big(\frac{L^{1/5}}{\log^{1/5}y}\Big)\Big\},$$

where $\mathscr{M}(W)$ is as in Theorem 1.

Proof. For the sum over $p>y^{1/2}$ in Lemma 4.5 a direct application of Lemma 4.3 gives, because of (4.8),

$$\sum_{y^{1/2}<p\leq yU}\frac{w(1)-w(p)}{p\log p}\gamma(p)F\Big(\frac{\log y-\log p}{\log p}\Big)$$

$$=\frac{2e^{\gamma_0}}{\log y}\Big\{\int_{1/2}^{U}\frac{W(1)-W(t)}{1-t}\frac{dt}{t}+O\Big(\frac{L}{\log y}\Big)\Big\},$$

W being related to w by (1.9). For the sum over smaller primes p we similarly obtain

$$\sum_{p\leq y^{1/2}}\frac{w(p)\gamma(p)}{p\log p}\sum_r(-1)^{r+1}k_r\Big(\frac{\log p}{\log y},1,G^{(-)^r}\Big)$$

$$=\frac{1}{\log y}\Big\{\int_0^{1/2}\frac{W(t)}{t}\sum_r(-1)^{r+1}k_r(t,1,G^{(-)^r})\frac{dt}{t}+O\Big(\frac{L}{\log y}\Big)\Big\},$$

where (4.10) has been used. Because of Lemma 5.2 this expression is

$$\frac{2e^{\gamma_0}}{\log y}\int_0^{1/2}W(t)\Big\{\frac{1}{1-t}-h(t,1)\Big\}\frac{dt}{t}+O\Big(\frac{L}{\log^2 y}\Big),$$

where h is as in Theorem 1.

The required result now follows from Lemma 4.5.

We can now complete the proof of Theorem 1. By Lemma 3.3 we can invoke the conclusions of Lemma 3.1. Hence

$$\sum_{\substack{a\in\mathscr{A}\\ v_{y,U}(a)\leq R}}w(q_a)\geq\sum_{a\in\mathscr{A}}\Sigma_{\mathscr{B}}(A,w).$$

Here $\Sigma_{\mathscr{B}}(A,w)$ is as given first in (2.7), so that we may write

$$\Sigma_{\mathscr{B}}(A,w)=w(1)\sum_d\mu(d)f(d),$$

with

$$\mu(d)f(d)=\mu_{\mathscr{B}}(d)\Big(1-\sum_{p|d}\{w(p)/w(1)\}\Big),$$

$\mu_{\mathscr{B}}$ being as in (2.5). Thus

$$0\leq f(d)\leq 1,$$

as required in Theorem 1.

Now invoke (1.1). Because of (3.1), (3.2) we obtain

$$\sum_{\substack{a\in\mathcal{A}\\ v_{y,U}(a)\leqslant R}} w(q_a) \geqslant w(1)\sum_{a\in\mathcal{A}}\sum_{d\mid a}\mu(d)f(d)$$
$$= w(1)\sum_{d}\mu(d)f(d)\sum_{\substack{a\in\mathcal{A}\\ a\equiv 0 \bmod d}} 1$$
$$\geqslant XS_{\mathcal{B}}(y, w) - w(1)\Big|\sum_{d\leqslant y}\mu(d)f(d)R(X,d)\Big|,$$

where $S_{\mathcal{B}}$ is as in (3.5). Theorem 1 now follows using Lemma 6.1.

Proceed to the inference to Corollary 1. When W is given by (1.11) we have in Theorem 1

(6.1) $$\mathcal{M}(W) \geqslant \mathcal{M}^*(W),$$

where

(6.2) $$\mathcal{M}^*(W) = -\int_{1/2}^{U}\frac{U-t}{1-t}\frac{dt}{t} + \int_{1/4}^{1/2}\frac{t-V}{1-t}\frac{dt}{t} + a - \frac{\beta}{3}(1-U),$$

because of (1.19), (1.20), (1.21). Equality holds in (6.1) if $(1-U)/3 \leqslant T_0$, where T_0 is as in (1.22). Accordingly define g_R, U_R, V_R by the "Lagrange multiplier" equations

(6.3) $$\mathcal{M}^*(W) = 0; \quad g_R = RU_R + V_R,$$

(6.4) $$\lambda R = -\int_{1/2}^{U_R}\frac{dt}{t(1-t)} + \frac{\beta}{3}; \quad \lambda = -\int_{1/4}^{1/2}\frac{dt}{t(1-t)}.$$

The solution of (6.4) is

$$\lambda = -\log 3, \quad \frac{1}{U_R} = 1 + 3^{-R}e^{-\beta/3},$$

so that U_R is as in (1.24). Thus

$$-g_R\log 3 = \lambda g_R = U_R\Big[-\int_{1/2}^{U_R}\frac{dt}{t(1-t)} + \frac{\beta}{3}\Big] - V_R\int_{1/4}^{1/2}\frac{dt}{t(1-t)}.$$

With (6.2), (6.3) this gives

$$0 = \mathcal{M}^*(W) = -g_R\log 3 + a - \beta/3 + \int_{1/4}^{U_R}\frac{dt}{1-t},$$

wherein

$$\int_{1/4}^{U_R}\frac{dt}{1-t} = \log\Big\{\frac{3}{4U_R(U_R^{-1}-1)}\Big\} = R\log 3 + \beta/3 + \log\Big(\frac{3}{4U_R}\Big).$$

Consequently

$$g_R = R - \frac{1}{\log 3}[\log(4U_R/3) - a],$$

so that all the requirements of (1.24) are satisfied.

We choose U, V, η_R, g so that

(6.5) $$U = U_R, \quad g \leqslant g_R - \eta_R = RU + V;$$

thus (1.25) holds as required. Then (6.2), (6.3) show, because $V = V_R - \eta_R$, that

(6.6) $$\mathcal{M}(W) \geqslant \mathcal{M}^*(W) = \eta_R\log 3,$$

as required in (1.27). These choices are consistent with the constraints (1.7), (1.10) of Theorem 1 provided

$$\eta_R > 0$$

and

$$1 \geqslant 3V + U = U_R + 3(g_R - RU_R) - 3\eta_R.$$

Because of (1.24) this says

$$3\eta_R \geqslant U_R - 1 + 3R(1-U_R) - \frac{3}{\log 3}\log\Big(\frac{4U_R}{3e^a}\Big),$$

i.e.

$$\eta_R \geqslant \Big(R - \frac{1}{3}\Big)(1-U_R) - \frac{1}{\log 3}\log\Big(\frac{4U_R}{3e^a}\Big).$$

Because of (1.26), this completes the proof of Corollary 1.

The proof of Corollary 2 is shorter. When equality holds in (1.10) and W is as in (1.11) we have in Theorem 1

$$\mathcal{M}(W) \geqslant \mathcal{M}^*(W)$$

where $\mathcal{M}^*(W)$ is given by (6.2) with $V = (1-U)/3$. Accordingly define g_R, U_0, V_0 by the equations

$$\mathcal{M}^*(W) = 0, \quad g_R = RU_0 + V_0, \quad V_0 = (1-U_0)/3.$$

The first two of these are precisely (1.29), (1.30) respectively. When (1.31) holds choose U, V so that

$$U = U_0, \quad g = RU + V;$$

then $V \leqslant V_0 - \eta_R$. Now (6.2) shows

$$\mathcal{M}(W) \geqslant \mathcal{M}^*(W) + \eta_R \log(3e^\beta),$$

and the stated result follows.

7. Proof of Theorem 2. We commence the study of the moments $h_n(s)$ introduced in (1.33) by defining

$$k_{i,n}(s) = \int_{x>0} x^n h_i(x, s)\,dx, \tag{7.1}$$

where h_i is as in (5.2). The recurrences (5.5), (5.6) induce

$$k_{i,n}(s) = \int_{x>0} x^n \int_{\substack{s<t\\ 3<t \text{ if } 2\nmid i}} h_{i-1}\left(\frac{tx}{t-1},\ t-1\right)\frac{dt}{t-1}$$

$$= \int_{\substack{s<t\\ 3<t \text{ if } 2\nmid i}} \left(\frac{t-1}{t}\right)^{n+1} \frac{dt}{t-1} \int_{x>0} \left(\frac{tx}{t-1}\right)^n h_{i-1}\left(\frac{tx}{t-1},\ t-1\right)\frac{d}{dx}\left(\frac{tx}{t-1}\right)dx,$$

so that

$$k_{i,n}(s) = \int_{\substack{s<t\\ 3<t \text{ if } 2\nmid i}} \left(\frac{t-1}{t}\right)^{n} k_{i-1,n}(t-1)\frac{dt}{t}, \tag{7.2}$$

for $s > 1$ initially, then for $s = 1$ by continuity.

We need the following bound for $k_{i,n}(s)$.

LEMMA 7.1. *There exist $\beta < 1$, $c > 0$ such that*

$$k_{i,n}(s) \leqslant c\beta^i (s+1)e^{-s} \quad (s \geqslant 1,\ n \geqslant 0).$$

We can deduce this result from the argument of Lemma 3 of [4]. Because of (7.1), (7.2), (1.15) it follows that

$$k_{i,n}(s) \leqslant j_i(s) \quad \text{when} \quad s \geqslant 1,\ n \geqslant 0,$$

where

$$j_1(s) = 1; \qquad j_i(s) = \int_{\substack{s<t\\ 2<t \text{ if } 2\nmid i}} j_{i-1}(t-1)\frac{dt}{t}, \tag{7.3}$$

the unnecessary change from 3 to 2 in the conditions of integration merely making the problem harder. However, on writing $j_i(s) = l_i(s+1)$, (7.3) gives the recursion studied in [4]. Thus it follows that

$$l_i(w) \ll \beta^i w e^{-w} \quad (w \geqslant 2),$$

which implies the result of our lemma.

We may now define

$$h_n(s) = \sum_{\substack{i \text{ even}\\ i\geqslant 2}} k_{i,n}(s) \quad (s \geqslant 1);$$

$$H_n(s) = \sum_{\substack{i \text{ odd}\\ i\geqslant 1}} k_{i,n}(s) \quad (s \geqslant 0), \tag{7.4}$$

convergence being assured by Lemma 7.1. This does not contradict (1.33) because of the following consequence of Lemma 7.1:

LEMMA 7.2. *The functions defined in* (7.4) *satisfy* (1.33) *and the recursions*

$$H_n(s) = \int_{\substack{s<t\\ 3<t}} \left(1 - \frac{1}{t}\right)^n h_n(t-1)\frac{dt}{t} + k_{1,n}(s) \quad (s \geqslant 0,\ n \geqslant 0), \tag{7.5}$$

$$h_n(s) = \int_{s<t} \left(1 - \frac{1}{t}\right)^n H_n(t-1)\frac{dt}{t} \quad (s \geqslant 1,\ n \geqslant 0), \tag{7.6}$$

where

$$k_{1,n}(s) = \iint_{\substack{0<x;\, x_1<u\\ 1/3<x_1<1/s\\ x_1+x+u=1}} \frac{x^n\,dx\,dx_1}{x_1 u}. \tag{7.7}$$

Also

$$H_n(s) \ll (s+1)e^{-s}; \quad h_n(s) \ll (s+1)e^{-s}. \tag{7.8}$$

Proof. The equation (7.7) for $k_{1,n}(s)$ is an assembly from (7.1) and (1.15). The inequalities (7.8) follow from (7.4) and Lemma 7.1. For (1.33) argue from (7.1) that, for $a = 0$ or 1,

$$\sum_{i\equiv a \bmod 2} k_{i,n}(s) = \sum_i \int = \int_{x>0} \sum_{i\equiv a \bmod 2} x^n h_i(x, s)\,dx \quad (s \geqslant 1,\ n \geqslant 0),$$

the interchange of order of summation and integration being permissible because

$$\sum_i \int \ll \sum_i \beta^i (s+1)e^{-s} \leqslant 1/(1-\beta).$$

Similarly, from (7.4), (7.2),

$$\int_{s<t} \left(1 - \frac{1}{t}\right)^n h_n(t-1)\frac{dt}{t} = \int_{s<t} \left(1 - \frac{1}{t}\right)^n \sum_{i \text{ even}} h_i(x,\ t-1)\frac{dt}{t}$$

$$= \sum_{\substack{i \text{ even}\\ i\geqslant 2}} h_{i+1,n}(s) = H_n(s) - k_{1,n}(s),$$

which is (7.5); (7.6) follows similarly.

To continue with the proof of Theorem 2, introduce the function

$$(7.9)\qquad J(s) = \int_0^\infty \exp\left\{-sx - \int_x^\infty \frac{e^{-t}}{t}\,dt\right\}dx \qquad (\mathrm{Re}(s) > 0),$$

which is analytic in the indicated domain. It is easy to verify (cf. [2]) that it then satisfies

$$\frac{d}{ds}\{sJ(s)\} + J(s+1) = 0 \qquad (\mathrm{Re}(s) > 0).$$

Hence

$$\frac{d}{ds}\{s^{n+1}J^{(n)}(s)\} + s^n J^{(n)}(s+1) = 0 \qquad (\mathrm{Re}(s) > 0).$$

Also from (7.9) we see

$$(7.10)\qquad J^n(s) = (-1)^n \int_0^\infty x^n \exp\left\{-sx - \int_x^\infty \frac{e^{-t}}{t}\,dt\right\}dx.$$

Hence

$$(7.11)\qquad (-1)^n J^{(n)}(1)/n! = j_n,$$

where j_n is as stated in (1.34).

In a similar way we may write

$$(7.12)\qquad i_n = (-1)^n I^{(n)}(1)/n! \qquad (n \geqslant 1),$$

$$(7.13)\qquad \frac{d}{ds}\{s^{n+1}I^{(n)}(s)\} - s^n I^{(n)}(s+1) = 0,$$

$$(7.14)\qquad I^{(n)}(s) = (-1)^n \int_0^\infty x^n \exp\left\{-sx + \int_x^\infty \frac{e^{-t}}{t}\,dt\right\}dx,$$

although the integral defining $I^{(n)}(s)$ does not converge if $n = 0$. From the point of view of classical analysis, we have introduced a function $I' = I^{(1)}$ and its derivatives $I^{(n)}$ $(n \geqslant 1)$ without introducing a function I.

The equation

$$\frac{d}{ds}\{sI(s)\} - I(s+1) = 0$$

does however have a solution

$$I(s) = 1 \qquad (\text{all } s).$$

This fact will be used later.

From the integral representations (7.10), (7.14) it is evident that

$$(7.15)\qquad J^{(n)}(s) = O(1)\ (s \geqslant 1,\ n \geqslant 0); \qquad I^{(n)}(s) = O(1)\ (s \geqslant 1,\ n \geqslant 1),$$

where the O-constant might depend upon n. We shall also need

$$(7.16)\qquad \lim_{s\to 0+} s^{n+1}I^{(n)}(s) = (-1)^n/n! \qquad (n \geqslant 1),$$

$$(7.17)\qquad \lim_{s\to 0+} s^{n+1}J^{(n)}(s) = (-1)^n/n! \qquad (n \geqslant 0).$$

We may deduce these from standard results on Laplace transforms, or *ad hoc* as follows. Rewrite (7.9) as

$$J(s) = \frac{1}{s} + f_1(s),$$

where

$$f_1(s) = \int_0^\infty e^{-(s+1)x}e^x\left\{-1 + \exp\left(-\int_x^\infty \frac{e^{-t}}{t}\,dt\right)\right\}dx.$$

Then $f_1(s)$ is analytic in $\mathrm{Re}(s) > -1$, because

$$\int_x^\infty \frac{e^{-t}}{t}\,dt = \frac{e^{-x}}{x} - \int_x^\infty \frac{e^{-t}}{t^2}\,dt = e^{-x}\left\{\frac{1}{x} - \int_x^\infty \frac{e^{x-t}}{t^2}\,dt\right\} = O(e^{-x}),$$

which implies

$$\exp\left\{-\int_x^\infty \frac{e^{-t}}{t}\,dt\right\} = 1 + O(e^{-x}).$$

Now (7.17) follows; (7.16) can be proved by a similar method.

For brevity, write for $\varepsilon = \pm 1$

$$(7.18)\qquad K_{n,\varepsilon}(s) = \begin{cases} 1 & \text{if } \varepsilon = 1,\ n = 0,\\ I^{(n)}(s) & \text{if } \varepsilon = 1,\ n \geqslant 1,\\ J^{(n)}(s) & \text{if } \varepsilon = -1,\ n \geqslant 0.\end{cases}$$

Thus

$$(7.19)\qquad \frac{d}{ds}\{s^{n+1}K_{n,\varepsilon}(s)\} = \varepsilon s^n K_{n,\varepsilon}(s+1) \qquad (s > 0).$$

Form the functions

$$(7.20)\qquad \alpha_{n,\varepsilon}(s) = H_n(s) + \varepsilon h_n(s) \qquad (s \geqslant 2).$$

Because of Lemma 7.2 they satisfy

$$(7.21)\qquad s\frac{d}{ds}a_{n,\varepsilon}(s) = -\varepsilon\left(\frac{s-1}{s}\right)^n a_{n,\varepsilon}(s-1)\qquad (s>3).$$

For further brevity omit the suffices n, ε *ad lib.*

We shall continue the functions a into $s>0$ by requiring that (7.21) hold for $s>1$, if $s\neq 2$, $s\neq 3$. The continued a will have discontinuities at 1, 2. Note that because of (7.5), (7.7)

$$(7.22)\qquad H(s) = H(3)+k_1(s) = H(3)+k_1(2)\quad \text{if}\quad 0<s<2.$$

For $2<s<3$ (7.21) requires, because of (7.20), (7.5), (7.6),

$$-\varepsilon\left(\frac{s-1}{s}\right)^n a(s-1) = sk_1'(s)-\varepsilon\left(\frac{s-1}{s}\right)^n\{H(3)+k_1(2)\}.$$

So for $1<s<2$ we define

$$a(s) = H(3)+k_1(2)-\varepsilon(s+1)^{n+1}s^{-n}k_1'(s+1).$$

For $2<t<3$, (7.7) shows

$$k_1'(t) = -\frac{1}{t}\int\limits_{\substack{0<x;1/t<u\\ x+u=1-1/t}}\frac{x^n}{u}\,dx = -\frac{(t-1)^n}{t^{n-1}}\int\limits_{\substack{0<x;1/(t-1)<u\\ x+u=1}}\frac{x^n}{u}\,dx.$$

Thus

$$(7.23)\qquad a(s) = H(3)+k_1(2)+\varepsilon\int\limits_{0<x<1-1/s}\frac{x^n}{1-x}\,dx\qquad (1<s<2).$$

For $1<s<2$ we now require

$$-\varepsilon\left(\frac{s-1}{s}\right)^n a(s-1) = \varepsilon\left(\frac{s-1}{s}\right)^n,$$

so define

$$(7.24)\qquad a(s) = -1\quad (0<s<1).$$

The equation (7.21) now holds for $s>0$. With its adjoint (7.19) we obtain

$$\frac{d}{ds}\{s^{n+1}K(s)a(s)\} = -\varepsilon(s-1)^nK(s)a(s-1)+\varepsilon s^nK(s+1)a(s)$$

for $s>1$, $s\neq 2$, $s\neq 3$. Because of (7.15), (7.20), (7.8) this integrates to

$$2^{n+1}K(2)\{a(2+)-a(2-)\}+K(1)a(1+) = \varepsilon\int_0^1 t^nK(t+1)a(t)\,dt.$$

Thus, and by (7.19), (7.23), (7.24),

$$2^{n+1}\varepsilon K(2)\left\{h(2)-\int_0^{1/2}\frac{x^n}{1-x}\,dx\right\}+K(1)\{H(3)+k_1(2)\}+$$
$$+K(1)-\lim_{t\to 0}t^{n+1}K(t) = 0,$$

the limit existing because of (7.16), (7.17), (7.18).

From (7.6), (7.22) we also have

$$h(1)-h(2) = \int_1^2\left(\frac{t-1}{t}\right)^n\{H(3)+k_1(2)\}\frac{dt}{t}$$
$$= \{H(3)+k_1(2)\}\int_{1/2}^1(1-y)^n\,dy/y$$
$$= \{H(3)+k_1(2)\}\int_0^{1/2}\frac{x^n}{1-x}\,dx.$$

Thus

$$2^{n+1}\varepsilon K(2)h(1)+\{H(3)+k_1(2)+1\}\left\{K(1)-2^{n+1}\varepsilon K(2)\int_0^{1/2}\frac{x^n}{1-x}\,dx\right\}$$
$$= \lim_{t\to 0+}t^{n+1}K(t).$$

But (7.19) gives

$$2^{n+1}\varepsilon K(2) = K'(1)+(n+1)K(1),$$

and the notation is as described by (7.18), (7.11), (7.12). Because of (7.16), (7.17) this completes the proof of Theorem 2.

8. Numerical approximations. For applications of Theorem 1 and its corollaries it will be necessary to obtain satisfactory bounds for the numbers α, β. We outline an approach to this question which, although not the best possible, is easy to implement. Numerical results already obtained indicate that Theorem 1 does, as was claimed in the introduction, represent an improvement upon earlier results of its type.

The first stage is to obtain approximation to the numbers i_n, j_n of Theorem 2. This amounts to estimating the derivatives at the point $s=1$ of the functions I', J defined in (7.9), (7.14). The author has developed an algorithm for this which takes advantage of the fact that these functions, being analytic in the half-plane $\operatorname{Re}(s) > 0$, are representable as sums of their Taylor series.

Next, use of Theorem 2 gives numerical values for $h_n(1)$. Together with a rather easier calculation involving the term $1/(1-t)$ in (1.18) this leads to numerical values for the moments

$$\mu_n = \int_0^1 t^n \varphi(t)\,dt = 4^{n+1} \int_0^{1/4} \psi(t)\,dt,$$

where for convenience we have written

$$\varphi(t) = \psi(t/4) \qquad (0 < t < 1),$$

the function ψ being as in (1.19). These numerical approximations are believed to be correct to within a few units of the rounding error imposed by the machine employed.

It is now possible to derive a lower bound for the positive constant α defined in (1.21), as follows. By differencing the μ_n we reach numerical approximations to

$$\mu_{n,i} = \int_0^1 t^i (1-t)^{n-i} \varphi(t)\,dt.$$

Then observe

$$\int\limits_{\substack{0<t<1\\ \varphi(t)>0}} \varphi(t)\,dt = \int\limits_{\substack{0<t<1\\ \varphi(t)>0}} \sum_{i=1}^{n} \binom{n}{i} t^i (1-t)^{n-i} \varphi(t)\,dt$$
$$\geqslant \sum_{i=1}^{n} \binom{n}{i} \max\left\{0, \int_0^1 t^i (1-t)^{n-i} \varphi(t)\,dt\right\}.$$

In the numerical work based upon this inequality, care must be taken that the effect of the relatively small errors in the assumed values for μ_n does not become too large. In practice this limits the value of n which may be used if useful information is to follow.

Similarly, of course, we could obtain lower bounds for the positive number β in (1.21), but this is not what is required. Inspection of the proof of Corollaries 1 and 2 shows that what is required is a lower bound for

$$\int\limits_{\substack{0<t<1\\ \varphi(t)>0}} \left(1 - \frac{c}{t}\right) \varphi(t)\,dt \tag{8.1}$$

where, because of (1.11), (1.13), $c = 4(1-U)/3 > 0$ for Corollary 1 and $c = 4V > 0$ for Corollary 2.

We use the estimate

$$\sum_{i=1}^{n} \binom{n}{i} \max\left\{0, \int_0^1 \left(1 - \frac{c}{t}\right) t^i (1-t)^{n-i} \varphi(t)\,dt\right\} \tag{8.2}$$

for the quantity (8.1). It is not now as clear as in the case $c = 0$ that this estimate is from below, though this holds for sufficiently small c because of a continuity consideration. Our claim that it holds for the values of c actually used therefore rests meanwhile on the observation that the estimate (8.2) increases with n for all values for which it has been computed, together with the fact that it can be proved to tend to the desired quantity (8.1) as $n \to \infty$.

To obtain our bounds on the numbers δ_R described in the introduction, we use Corollary 1 (see (1.28)) when $R \geqslant 3$, since our computations then indicate $N_R < 0$. For $R = 2$ we use Corollary 2. Corollary 1 remains applicable when $R = 2$, but only for a somewhat restricted range of g, because we find not only $\Delta_2 < 0.056$ but also $N_2 > 0.074$.

In this way we find $\delta_2 < 0.068$, $\delta_3 < 0.106$, and $\delta_R < 0.131$ for all integers $R \geqslant 4$. It is hoped to report on a more definitive computation of α, β when dealing with some applications of Theorem 1 in a subsequent publication.

Note added in proof, 21. 10. 1981. Since this article was written the author has undertaken a computation of the constants α, β, T_0 using a large computer and methods more refined than those described in Section 8 above. It was found that

$$\alpha = 0.1505528\ldots, \qquad \beta = 0.87695\ldots, \qquad T_0 = 0.074368\ldots$$

The values of α, β lead to

$$\delta_2 = 0.063734\ldots, \qquad \delta_3 = 0.099995\ldots, \qquad \delta_\infty = 0.124820\ldots$$

The value of $1/T_0 = 13.446\ldots$ means that the method used attaches weights $w(p_i)$ to the eleven largest prime factors p_i $(1 \leqslant p_i \leqslant 11)$ of some of the unwanted a in $\mathscr{A}$.

A more extensive table of values of δ_R has been supplied in the author's article (introductory to the present paper) *Rosser's Sieve with Weights*, in [H. Halberstam and C. Hooley (editors), *Progress in Number Theory*. I, pp. 61–68, London 1981]. The author's algorithm for the calculation of the numbers j_n, i_n, of Theorem 2 will be described (for the numbers j_n) in [G. Greaves, *An Algorithm for the Solution of Certain Differential-Difference Equations of Advanced Type*, to appear in Math. of Comp. 38 (1982)]. The methods used for the more accurate computation of α, β have been outlined in [G. Greaves, *An Algorithm for the Hausdorff Moment Problem*, submitted to Numerische Math.].

References

[1] N. C. Ankeny and H. Onishi, *The general sieve*, Acta. Arith. 10 (1964), pp. 31–62.

[2] N. G. de Bruijn, *On the number of uncancelled elements in the sieve of Eratosthenes*, Nederl. Akad. Wetensch. Proc. 53 (1950), pp. 803–812 = Indag. Math. 12 (1950), pp. 247–256.

[3] H. Halberstam and H. E. Richert, *Sieve methods*, London 1974.
[4] H. Iwaniec, *On the error term in the linear sieve*, Acta. Arith. 19 (1971), pp. 1–
[5] — *The half dimensional sieve*, ibid. 29 (1976), pp. 69–95.
[6] W. B. Jurkat and H. E. Richert, *An improvement of Selberg's sieve meth I*, ibid. 11 (1965), pp. 217–240.
[7] P. Kuhn, *Zur Viggo Brunschen Siebmethode. I*, Norske Vid. Selsk. Fo Trondhjem, 14 (1941), pp. 145–148.
[8] — *Neue Abschätzungen auf Grund der Viggo Brunschen Siebmethode*, 12. Ska Mat. Kongr., Lund (1953), pp. 160–168.
[9] H. E. Richert, *Selberg's sieve with weights*, Mathematika 16 (1969), pp. 1–
[10] A. Selberg, *Sieve methods*, Proc. Sympos. Pure Math. 20 (1971), pp. 311–3

Received on 17.5.1979 (11