PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 115, Number 1, May 1992

THE DISTRIBUTION OF SMOOTH NUMBERS IN ARITHMETIC PROGRESSIONS

ANTAL BALOG AND CARL POMERANCE

(Communicated by William Adams)

ABSTRACT. We estimate the number of integers n up to x in the arithmetic progression $a \pmod q$ with n free of prime factors exceeding y. For a wide range of the variables x, y, q, and a we show that this number is about $x/(qu^u)$, where $u = \log x/\log y$.

1. INTRODUCTION

We say a natural number n is y-smooth if every prime factor p of n satisfies $p \le y$. Let $\psi(x, y)$ denote the number of y-smooth integers up to x. Thanks to many researchers in this century, we now know a great deal about the function $\psi(x, y)$. Less studied is the function $\psi(x, y; q, a)$ which denotes the number of y-smooth integers $n \le x$ with $n \equiv a \pmod q$. It is the object of this paper to get reasonable estimate for $\psi(x, y; q, a)$ for a wide range of the four variables.

Before stating the precise result we obtain, we briefly review some prior work. It is known that

$$\psi(x, y) \sim \rho(u)x, \qquad u = \log x/\log y, \tag{1.1}$$

uniformly in a large portion of the x, y plane. Here, ρ is the continuous solution to the differential difference equation $u\rho'(u) = -\rho(u-1)$ for $u > 1$, with the initial condition $\rho(u) = 1$ for $u \le 1$. Results of this nature are due to Dickman, Buchstab and de Bruijn. The largest x, y region for which we know (1.1) is due to Hildebrand [Hi]:

$$y \ge \exp\{(\log\log x)^{5/3+\varepsilon}\}$$

for any fixed $\varepsilon > 0$. For (1.1) to be useful, it would be good to know estimates for $\rho(u)$. This is provided by de Bruijn [de B]:

$$\rho(u) = \exp\{-u(\log u + \log\log(u+1) + O(1))\} \tag{1.2}$$

for $u \ge 1$. (In the cited paper, a more precise estimate of $\rho(u)$ is given.)

Received by the editors October 12, 1990.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11N25.
Supported in part by an NSF grant.

What do we expect for $\psi(x, y; q, a)$? For this quantity to be nonzero it is necessary that no prime factor of (a, q) exceed y. If this occurs, then

$$\psi(x, y; q, a) = \psi\left(\frac{x}{(a, q)}, y; \frac{q}{(a, q)}, \frac{a}{(a, q)}\right).$$

Thus we may as well assume that $(a, q) = 1$. Let $\psi_q(x, y)$ denote the number of y-smooth integers $n \le x$ with $(n, q) = 1$. Then one might expect these $\psi_q(x, y)$ integers n to be roughly equally distributed in the $\varphi(q)$ residue classes $a \pmod q$ with $(a, q) = 1$. Thus a natural function with which to compare $\psi(x, y; q, a)$ is $\frac{1}{\varphi(q)}\psi_q(x, y)$. Recently, Fouvry and Tenenbaum [F-T] have shown that

$$\psi(x, y; q, a) = \frac{1}{\varphi(q)}\psi_q(x, y)(1 + O_A(e^{-c_1(\log y)^{1/2}})) \tag{1.3}$$

uniformly for

(1.4)

$$x \ge 3, \qquad \exp\{c_2(\log\log x)^2\} \le y \le x, \qquad 1 \le q \le (\log x)^A, \qquad (a, q) = 1$$

where c_1, c_2 are positive, absolute constants and $A > 0$ is arbitrary. (They also have a similar result when $q \le e^{c_1(\log y)^{1/2}}$.) Further, they show that

$$\psi_q(x, y) = \frac{\varphi(q)}{q}\psi(x, y)\left(1 + O_\varepsilon\left(\frac{\log\log(qy)\log\log x}{\log y}\right)\right) \tag{1.5}$$

if

$$\begin{gathered} x \ge x_0(\varepsilon), \quad \exp\{(\log\log x)^{5/3+\varepsilon}\} \le y \le x, \\ \log\log(q+2) \le \left(\frac{\log y}{\log(u+1)}\right)^{1-\varepsilon}. \end{gathered} \tag{1.6}$$

In particular, (1.5) holds under hypothesis (1.4).

In another recent paper, Granville [G] has obtained (1.3) with a weaker error term, but in a range much wider than (1.4). In particular, he shows

$$\psi(x, y; q, a) = \frac{1}{\varphi(q)}\psi_q(x, y)\left(1 + O\left(\frac{\log q}{\log y}\right)\right) \tag{1.7}$$

uniformly holds in the range

$$2 \le y \le x, \qquad 1 \le q \le \min\{x, y^\alpha\}, \qquad (a, q) = 1.$$

Here, α is any fixed positive quantity. Thus (1.7) contains the asymptotic result

$$\psi(x, y; q, a) \sim \frac{1}{\varphi(q)}\psi_q(x, y)$$

as $x \to \infty$, $\log q/\log y \to 0$.

In this paper, we do not get an asymptotic estimate for $\psi(x, y; q, a)$, but rather, upper and lower bounds. Note that by combining (1.1)–(1.3) and (1.5), one has

$$\psi(x, y; q, a) = \frac{x}{q}\exp\{-u(\log u + \log\log(u+1) + O(1))\} \tag{1.8}$$

in the range (1.4). By (1.7), this result continues to hold for values of q satisfying (1.6) and $q \le y^\varepsilon$. We are able to show that the double inequality of (1.8) holds for a still larger range.

Theorem. *For each* $\varepsilon > 0$ *and all* x, y, q, a *satisfying*

$$x \geq 2, \qquad \exp\{(\log\log x)^2\} \leq y \leq x^{2/3-\varepsilon}, \tag{1.9}$$
$$1 \leq q \leq \min\{y^{4/3-\varepsilon}, (x/y)^{4/3-\varepsilon}\}, \qquad (a, q) = 1,$$

we have

$$\psi(x, y; q, a) = \frac{x}{q}\exp\{-u(\log u + \log\log u + O(1))\}, \tag{1.10}$$

where $u = \log x/\log y$. *The constants implied by the* O *notation depend at most on the choice of* ε.

The proof falls naturally into two cases: the upper bound implicit in (1.10) and the lower bound. The upper bound result rests strongly on [F2] in which a stronger result is proved for a somewhat narrower range. Our proof of the lower bound implicit in (1.10) bears some similarity to the method in [F1]. In particular, we too show that the case $u \geq 2$ follows from the case $u < 2$. However, our proof also introduces new elements including the Weyl-Hooley estimation of incomplete Kloosterman sums, the lower bound in the fundamental lemma of the sieve and a combinatorial argument reminiscent of [CEP].

Finally, we mention that in [G], Granville, using the theorem above, has shown that

$$\psi(x, y; q, a) \asymp \frac{1}{\varphi(q)}\psi_q(x, y)$$

for x, y, q, and a satisfying (1.9) and, in fact, in the expanded region where the lower bound on y is replaced with $y \geq 2$.

2. The upper bound

In this section we establish the upper bound implicit in the theorem. We begin by stating the following result from Friedlander [F2].

Lemma 2.1. *If* $(a, q) = 1$, $x \geq q^2y^5$, *and* $y \geq \exp\{(\log x)^{4/5}\}$, *then*

$$\psi(x, y; q, a) \ll \frac{x}{q}\rho\left(u - 4 - \frac{\log q}{\log y}\right),$$

where $u = \log x/\log y$.

Using (1.2) we have that

$$\rho\left(u - 4 - \frac{\log q}{\log y}\right) = \exp\{-u(\log u + \log\log u + O(1))\}$$

for q, y satisfying (1.9). Thus Lemma 2.1 gives us the upper bound in the theorem for

$$\exp\{(\log x)^{4/5}\} \leq y \leq x^{1/8}.$$

Our task in the remainder of this section is to expand this interval to $\exp\{(\log\log x)^2\} \leq y \leq x^{2/3-\varepsilon}$.

The range $x^{1/8} \leq y \leq x^{2/3-\varepsilon}$ is basically trivial since $\psi(x, y; q, a)$ is majorized by the number of integers $n \leq x$ in the residue class $a \pmod q$. Thus

$$\psi(x, y; q, a) \leq \frac{x}{q} + 1,$$

which gives the upper bound in the theorem when u is bounded from infinity and from 1.

We now consider the interval $\exp\{(\log\log x)^2\} \le y \le \exp\{(\log x)^{4/5}\}$. We shall use an argument similar to that of Friedlander's for Lemma 2.1. Our proof, in fact, works for y up to $x^{1/\log\log x}$, but not beyond. So we still need the more precise Lemma 2.1, at least for the range $x^{1/\log\log x} \le y \le x^{1/8}$.

Let $P(n)$ denote the largest prime factor of n. If $x/y^3 < n \le x$ and $P(n) \le y$, then n has at least one factorization as ml where $x/y^3 < m \le x/y^2$. Thus

$$\begin{aligned}\psi(x, y; q, a) &\le \psi\left(\frac{x}{y^3}, y; q, a\right) + \sum_{\substack{\frac{x}{y^3}<m\le\frac{x}{y^2}\\ P(m)\le y\\ (m,q)=1}} \sum_{\substack{l\le x/m\\ ml\equiv a \pmod q}} 1 \\ &\le \psi\left(\frac{x}{y^3}, y\right) + \sum_{\substack{x/y^3<m\le x/y^2\\ P(m)\le y}} \left(\frac{x}{mq}+1\right).\end{aligned} \tag{2.1}$$

Now

$$\begin{aligned}\sum_{\substack{x/y^3<m\le x/y^2\\ P(m)\le y}} \left(\frac{x}{mq}+1\right) &\le 2\frac{x}{q} \sum_{\substack{x/y^3<m\le x/y^2\\ P(m)\le y}} \frac{1}{m} \\ &\le 2\frac{x}{q}\left\{\frac{\psi(x/y^2, y)}{x/y^2} + \int_{x/y^3}^{x/y^2} \frac{1}{t^2}\psi(t, y)\,dt\right\}.\end{aligned} \tag{2.2}$$

Using the fact that (1.2) and the Hildebrand region of validity of (1.1) imply

$$\psi(t, y) = t \cdot \exp\{-u(\log u + \log\log u + O(1))\}$$

uniformly for $x/y^3 \le t \le x/y^2$, $\exp\{(\log\log x)^2\} \le y \le x^{1/4}$ (where $u = \log x/\log y$), we have from (2.1) and (2.2) that

$$\begin{aligned}\psi(x, y; q, a) &\le \left(\frac{x}{y^3} + 2\frac{x}{q} + 2\frac{x}{q}\int_{x/y^3}^{x/y^2} \frac{dt}{t}\right)\exp\{-u(\log u + \log\log u + O(1))\} \\ &= \frac{x\log y}{q}\exp\{-u(\log u + \log\log u + O(1))\}.\end{aligned}$$

This inequality gives the upper bound we are looking for provided $\log y = \exp\{O(u)\}$, which holds if $y \le x^{1/\log\log x}$.

3. The lower bound

We may suppose ε in the theorem satisfies $0 < \varepsilon < 1/10$. Fix δ as $\varepsilon/4$ and assume $y^{2-\delta} \le x$. Consider a number $ml \le x$ where

$$\frac{x}{y^{2-\delta}} \le m \le \frac{x}{y^{2-2\delta}} \tag{3.1}$$

and every prime $p|m$ satisfies $y^{\delta/2} \le p \le y$, $p \nmid q$.

The number of representations of $n \le x$ in the form ml, where m satisfies (3.1), is at most

$$\left(1 + \frac{2u}{\delta}\right)^{4/\delta} \le u^{c_\delta} \tag{3.2}$$

for some constant c_δ. Indeed the factor l of n has all of the prime factors of n below $y^{\delta/2}$, so it is determined by its prime factors exceeding $y^{\delta/2}$. But n has at most $2u/\delta$ such prime factors and l can have at most $4/\delta$ of them, since $l \le x/m < y^2$.

From (3.2) we have

$$\psi(x, y; q, a) \ge u^{-c_\delta} \sum_m \psi\left(\frac{x}{m}, y; q, a\overline{m}\right), \tag{3.3}$$

where m satisfies (3.1) and $m\overline{m} \equiv 1 \pmod q$. For any fixed m let $w = x/m$, $b = a\overline{m}$. Note that (3.1) implies $y^{2-2\delta} \le w \le y^{2-\delta}$. Thus

$$y \le w^{(2-2\delta)^{-1}} < w^{(2-2\varepsilon)^{-1}} < w^{2/3-\varepsilon},$$

and

$$(w/y)^{(4/3)(1-\delta)} \ge (y^{1-2\delta})^{(4/3)(1-\delta)} > y^{4/3-4\delta} = y^{4/3-\varepsilon} \ge q.$$

From these inequalities we see that w, y, q, and b satisfy the hypothesis of the theorem (with ε replaced with $\varepsilon' = \varepsilon/3$) and $y^{2-\delta} \ge w$. But if $\delta' = \varepsilon'/4$, we have $y^{2-\delta'} > y^{2-\delta} \ge w$.

Suppose the theorem holds when $y^{2-\delta} \ge x$. In this case the theorem asserts that $\psi(x, y; q, a) \gg_\varepsilon x/q$. The remaining case $y^{2-\delta} \le x$ would thus follow from using this estimate on the right side of (3.3) and a lower bound for $\sum 1/m$. That is, we need only show two things: the case $y^{2-\delta} \ge x$ and the following lemma.

Lemma 3.1. *If* $0 < \delta < 1$, $\exp\{(\log\log x)^2\} \le y$, $y^{2-\delta} \le x$, *then*

$$\sum_m \frac{1}{m} \ge \exp\{-u(\log u + \log\log u + O_\delta(1))\},$$

where m *satisfies* (3.1) *and* $u = \log x/\log y$.

4. Proof of Lemma 3.1

We divide the proof into two cases: $2 - \delta \le u \le e^{3/\delta}$, $u > e^{3/\delta}$. In the first case, let

$$k := \left[\frac{2(u-2)}{\delta}\right] + 3 \ge 1.$$

Suppose m is a product of k not necessarily distinct primes p with $y^{\delta/2} < p \le y^{(1+1/k)\delta/2}$ and $p \nmid q$. Then

$$m > y^{k\delta/2} > y^{u-2+\delta} = \frac{x}{y^{2-\delta}},$$

$$m \le y^{(k+1)\delta/2} \le y^{u-2+2\delta} = \frac{x}{y^{2-2\delta}}.$$

That is, such a number m satisfies (3.1). Thus

$$\sum 1/m \ge S^k/k!, \tag{4.1}$$

where S is the sum $\sum 1/p$ where p runs over the primes in $(y^{\delta/2}, y^{(1+1/k)\delta/2}]$ which do not divide q. Since $u \le e^{3/\delta}$ implies $k = O_\delta(1)$, we have $S \gg_\delta 1$, so that (4.1) gives

$$\sum 1/m \gg_\delta 1 \gg \exp\{-u(\log u + \log\log u + O(1))\}.$$

We now assume $u > e^{3/\delta}$. Consider integers $m = m_1 m_2$ where m_1 is composed of $v_1 := [u-2]$ not necessarily distinct primes p with

$$y^{1-1/\log u} < p \le y, \quad p \nmid q, \tag{4.2}$$

and m_2 is composed of v_2 not necessarily distinct primes p with

$$y^{\delta/2} < p \le y^{(1+1/u)\delta/2}, \quad p \nmid q, \tag{4.3}$$

where $v_2 = v_2(m_1)$ is that integer such that

$$y^{(v_2-1)\delta/2} < \frac{x}{m_1 y^{2-\delta}} \le y^{v_2\delta/2}. \tag{4.4}$$

We have to check that such integers $m = m_1 m_2$ satisfy (cf. (3.1))

$$\frac{x}{y^{2-\delta}} < m \le \frac{x}{y^{2-2\delta}}.$$

We first note that

$$2 \le v_2 < 3 + \frac{2}{\delta} + \frac{2u}{\delta \log u} < u. \tag{4.5}$$

To see this note that $m_1 \le y^{u-2} = x/y^2$ by (4.2), so that $v_2 \ge 2$ follows at once from (4.4). On the other hand,

$$m_1 > y^{v_1(1-1/\log u)} > y^{(u-3)(1-1/\log u)} > \frac{x}{y^{3+u/\log u}},$$

so that (4.4) implies

$$\begin{aligned} v_2 &< 1 + \frac{\log(x/m_1 y^{2-\delta})}{\log(y^{\delta/2})} < 1 + \frac{1+\delta+u/\log u}{\delta/2} \\ &= 3 + \frac{2}{\delta} + \frac{2u}{\delta \log u} < 3 + \frac{2}{\delta} + \frac{2u}{3} < u, \end{aligned}$$

using $0 < \delta < 1$, $u > e^{3/\delta}$. Thus we have (4.5).

We are now in a position to confirm that $m = m_1 m_2$ satisfies (3.1). Indeed, from (4.3) and (4.4) we have

$$m > m_1 y^{v_2\delta/2} \ge \frac{x}{y^{2-\delta}},$$

while from (4.3)–(4.5) we have

$$\begin{aligned} m &\le m_1 y^{v_2(1+1/u)\delta/2} = m_1 y^{(v_2-1)\delta/2} y^{\delta/2} y^{(v_2/u)\delta/2} \\ &< \frac{x}{y^{2-\delta}} y^{\delta/2} y^{\delta/2} = \frac{x}{y^{2-2\delta}}. \end{aligned}$$

Since $u > e^{3/\delta} > 10$, the intervals (4.2) and (4.3) are disjoint. Thus

$$\sum \frac{1}{m} \ge \sum \frac{1}{m_1 m_2} = \sum \frac{1}{m_1} \sum \frac{1}{m_2} \tag{4.6}$$

where m_1, m_2 run over the numbers described above. Let S_2 denote the sum $\sum 1/p$ where p satisfies (4.3). Then

$$S_2 \asymp 1/u,$$

using $\exp\{(\log\log x)^2\} \le y$ and $u > e^{3/\delta}$. Indeed,

$$S_2 = \log\log(y^{(1+1/u)\delta/2}) - \log\log(y^{\delta/2}) + O(\exp\{-(\log y)^{3/5-\varepsilon}\}) + O\left(\frac{1}{\delta y^{\delta/2}}\right)$$
$$= \log\left(1+\frac{1}{u}\right) + O(\exp\{-(\log\log x)^{6/5-2\varepsilon}\}) + O\left(\frac{1}{\delta y^{\delta/2}}\right).$$

We conclude from (4.5) that uniformly for each m_1 we have

$$\sum \frac{1}{m_2} \ge \frac{S_2^{v_2}}{v_2!} = \exp\{-O(u/\delta)\} = \exp\{-O_\delta(u)\}. \tag{4.7}$$

Further, if S_1 is the sum $\sum 1/p$ where p satisfies (4.2), we have

$$S_1 \sim \frac{1}{\log u}$$

uniformly for $\exp\{(\log\log x)^2\} \le y$. Thus

$$\sum \frac{1}{m_1} \ge \frac{S_1^{v_1}}{v_1!} = \exp\{-u(\log u + \log\log u + O(1))\}. \tag{4.8}$$

Combining (4.6)–(4.8), we have our result.

5. The case $y^{2-\delta} \ge x$

We begin by stating three technical results.

Lemma 5.1 (the lower bound in the fundamental lemma of the sieve). *Let D, $z \ge 1$ and let q be a positive integer. Let*

$$P_q = \prod_{\substack{p<z \\ p\nmid q}} p, \qquad s = \frac{\log D}{\log z}.$$

Then there is a sequence $\{\lambda_d\}_{d=1}^\infty$ of "sieving weights" such that for each positive integer l,

$$\sum_{d|l} \lambda_d = 1 \quad \text{if } (l, P_q) = 1, \qquad \sum_{d|l} \lambda_d \le 0 \quad \text{if } (l, P_q) > 1,$$
$$|\lambda_d| \le 1 \quad \text{for all } d, \qquad \lambda_d = 0 \quad \text{for all } d > D \text{ and for all } d \nmid P_q,$$
$$\sum \frac{\lambda_d}{d} = \left(\prod_{p|P_q}(1-p^{-1})\right)(1+O(e^{-s})).$$

For example, see Lemma 5 in [F-I]. Actually this result is implicit in [H-R] with the better error term $O(s^{-s})$.

Lemma 5.2 (the Erdös-Turán inequality). *Let $H \ge 1$ and let $\{a_n\}_{n=1}^\infty$, $\{x_n\}_{n=1}^\infty$ be real sequences with each $a_n \ge 0$ and $\sum a_n < \infty$. Then*

$$\left|\sum_n a_n \psi(x_n)\right| \ll \frac{1}{H}\sum_n a_n + \sum_{h\le H}\frac{1}{h}\left|\sum_n a_n e(hx_n)\right|$$

holds uniformly, where $\psi(\theta) := \theta - [\theta] - \frac{1}{2}$ and $e(\theta) := e^{2\pi i\theta}$.

For example, see Iwaniec [I, Lemma 6] or Laborde [L].

Lemma 5.3 (incomplete Kloosterman sums). *If $A < B$ and q is a positive integer, then*

$$\sum_{\substack{A<m\le B\\(m,q)=1}} e\left(\frac{b\overline{m}}{q}\right) \ll_\eta (b,q)^{1/2}q^{1/2+\eta} + (b,q)\frac{B-A}{q}$$

for any $\eta > 0$, where $\overline{m}$ denotes the inverse of $m \pmod q$.

For example, see Hooley [Ho, Lemma 1].

Finally, we need the following simple inequalities:

$$\sum_{h\le H}(h,q) \le \sum_{d|q} d \sum_{\substack{h\le H\\ d|h}} 1 \le \sum_{d|q} H = \tau(q)H, \tag{5.1}$$

$$\sum_{h\le H}\frac{(h,q)}{h} \le \tau(q) + \int_1^H \frac{\tau(q)}{t}\,dt = \tau(q)(1+\log H), \tag{5.2}$$

where $\tau(q)$ denotes the number of natural divisors of q.

We can now proceed with the proof of the lower bound in the theorem when $y^{2-\delta} \ge x$ (where $\delta = \varepsilon/4$). Let $s_0 \ge 1$ be an absolute constant so large that the factor $1 + O(e^{-s_0})$ in Lemma 5.1 is at least $1/2$. Let P_q denote the product of the primes $p < y^{\delta/s_0^2}$, $p \nmid q$. Consider integers $n = ml \le x$, where

$$x/y < m \le y, \qquad (m, P_q) = 1, \qquad ml \equiv a \pmod q. \tag{5.3}$$

Note that these conditions imply $P(n) \le y$. The number of representations of $n \le x$ as ml in this fashion is at most the number of divisors of n free of prime factors below y^{δ/s_0^2}. But $n \le x < y^2$, so that n has at most $2s_0^2/\delta$ prime factors that can possibly be used to make up m. Thus n has at most $2^{2s_0^2/\delta} = O_\delta(1)$ representations as ml. Thus

$$\psi(x,y;q,a) \gg_\delta \sum_{ml\le x}{}^{*}\, 1 \tag{5.4}$$

where the star indicates the sum is over pairs m, l satisfying (5.3). By Lemma 5.1, with $D = y^{\delta/s_0}$ and $z = y^{\delta/s_0^2}$, we have

$$\begin{aligned}
\sum_{ml\le x}{}^{*}\, 1 &\ge \sum_{\substack{ml\le x\\ x/y<m\le y\\ ml\equiv a(q)}} \sum_{d|m} \lambda_d = \sum_{(d,q)=1} \lambda_d \sum_{\substack{kl\le x/d\\ kl\equiv a\overline{d}(q)\\ x/(dy)<k\le y/d}} 1 \\
&= \sum_{(d,q)=1} \lambda_d \sum_{\substack{x/(dy)<k\le y/d\\ (k,q)=1}} \left(\left[\frac{x}{dkq} - \frac{a\overline{d}\,\overline{k}}{q}\right] - \left[-\frac{a\overline{d}\,\overline{k}}{q}\right]\right) \\
&= \frac{x}{q}\sum_{(d,q)=1}\frac{\lambda_d}{d}\sum_k{}^{(d)}\frac{1}{k} \\
&\quad - \sum_{(d,q)=1}\lambda_d \sum_k{}^{(d)}\left\{\psi\left(\frac{x}{dkq} - \frac{a\overline{d}\,\overline{k}}{q}\right) - \psi\left(-\frac{a\overline{d}\,\overline{k}}{q}\right)\right\} \\
&= M - R, \quad \text{say},
\end{aligned} \tag{5.5}$$

where a bar over a number denotes the inverse modulo q and $\sum^{(d)}$ denotes a sum over integers k satisfying $x/(dy) < k \le y/d$ and $(k, q) = 1$. We view M as the main term in (5.5) and R as the error term.

To estimate $\sum^{(d)} 1/k$ in the main term, note first that by sieving with the prime factors of q, we have

$$\sum_{\substack{k \le t \\ (k,q)=1}} 1 = \frac{\varphi(q)}{q} t + O(\tau(q)).$$

Thus by partial summation we have, uniformly for every $d \le D$,

$$\sum_k^{(d)} \frac{1}{k} = \frac{\varphi(q)}{q} \log \frac{y^2}{x} + O\left(\frac{\tau(q)\, dy}{x}\right).$$

Using this together with Lemma 5.1 (and $\log D/\log z = s_0$), we get

$$\begin{aligned} M &= \frac{x}{q} \frac{\varphi(P_q)}{P_q}(1 + O(e^{-s_0})) \frac{\varphi(q)}{q} \log \frac{y^2}{x} + O\left(\frac{y\tau(q)}{q} \sum_d |\lambda_d|\right) \\ &\ge \frac{1}{2} \frac{x}{q} \frac{\varphi(qP_q)}{qP_q} \log \frac{y^2}{x} + O\left(\frac{yD\tau(q)}{q}\right) \gg \frac{x}{q} \frac{\varphi(qP_q)}{qP_q} \log \frac{y^2}{x}. \end{aligned}$$

Using Mertens' theorem to estimate $\varphi(qP_q)/qP_q$, we have

$$M \gg \frac{x}{q} \frac{\delta \log y}{s_0^{-2} \delta \log y} \gg \frac{x}{q}. \tag{5.6}$$

Thus to complete the proof it will be sufficient to show that the error term R in (5.5) is $o(x/q)$.

If $q \le y^{1/2}$, we have trivially that for any H with $D \le H \le y$

$$\begin{aligned} R &\ll \sum_{d \le D} \frac{y}{d} \ll y \log D < y \log x \\ &= \frac{yq \log x}{q} \le \frac{y^{3/2} \log x}{q} \le \frac{x^{1-3\varepsilon/2} \log x}{q} = o\left(\frac{x}{q}\right). \end{aligned}$$

We thus may assume that $q > y^{1/2}$.

From Lemma 5.2 we have that for any H with $D \le H \le y$

$$\begin{aligned} |R| &\le \sum_{\substack{d \le D \\ (d,q)=1}} \left| \sum_k^{(d)} \left\{ \psi\left(\frac{x}{dkq} - \frac{a\bar{d}\,\bar{k}}{q}\right) - \psi\left(-\frac{a\bar{d}\,\bar{k}}{q}\right) \right\} \right| \\ &\ll \sum_{\substack{d \le D \\ (d,q)=1}} \left\{ \frac{y}{H} + \sum_{h \le H/d} \frac{1}{h} \left| \sum_k^{(d)} e\left(\frac{xh}{dkq} - \frac{a\bar{d}\,\bar{k}h}{q}\right) \right| \right. \\ &\qquad \left. + \sum_{h \le H/d} \frac{1}{h} \left| \sum_k^{(d)} e\left(-\frac{a\bar{d}\,\bar{k}h}{q}\right) \right| \right\}. \end{aligned} \tag{5.7}$$

The factor $e(xh/dkq)$ is smooth in the variable k and so partial summation is appropriate for the first inner sum:

$$\begin{aligned}
&\sum_k^{(d)} e\left(\frac{xh}{dkq}\right) e\left(-\frac{a\overline{d}\,\overline{k}h}{q}\right) \\
&\quad = e\left(\frac{xh}{yq}\right) \sum_k^{(d)} e\left(-\frac{a\overline{d}\,\overline{k}h}{q}\right) \\
&\qquad + 2\pi i \int_{x/dy}^{y/d} \frac{xh}{dt^2 q} e\left(\frac{xh}{dtq}\right) \sum_{k\le t}^{(d)} e\left(-\frac{a\overline{d}\,\overline{k}h}{q}\right) dt \\
&\quad \ll \left(1+\frac{yh}{q}\right) \max_{t\le y/d} \left| \sum_{k\le t}^{(d)} e\left(-\frac{a\overline{d}\,\overline{k}h}{q}\right) \right| \\
&\quad \ll_\eta \left(1+\frac{yh}{q}\right)(h,q)\left(q^{1/2+\eta}+\frac{y}{dq}\right),
\end{aligned}$$

for any $\eta > 0$, using Lemma 5.3 for the last step.

From (5.7) we thus have for any $\eta > 0$ (using Lemma 5.3 for the second inner sum),

$$\begin{aligned}
R &\ll_\eta \sum_{d\le D} \left\{ \frac{y}{H} + \sum_{h\le H/d} \frac{1}{h}\left(1+\frac{yh}{q}\right)(h,q)\left(q^{1/2+\eta}+\frac{y}{dq}\right) \right\} \\
&\le \frac{yD}{H} + \sum_{d\le D}\left(1+\log\frac{H}{d}+\frac{yH}{dq}\right)\tau(q)\left(q^{1/2+\eta}+\frac{y}{dq}\right),
\end{aligned}$$

by (5.1) and (5.2). Thus for any $\eta > 0$,

$$R \ll_\eta \frac{yD}{H} + \tau(q)\left\{ Dq^{1/2+\eta}\log H + \frac{y}{q}\log D \log H + \frac{yH}{q^{1/2-\eta}}\log D + \frac{y^2 H}{q^2} \right\}.$$

Recall that $D = y^{\delta/s_0} \le y^\delta$. Using $\tau(q)\log^2 y \ll_\eta y^\eta$ for any $\eta > 0$, we have

$$\begin{aligned}
R &\ll_\eta \frac{y^{1+\delta}}{H} + y^{2\eta}\left\{ y^\delta q^{1/2} + \frac{y}{q} + \frac{yH}{q^{1/2}} + \frac{y^2 H}{q^2} \right\} \\
&< y^{2\eta+\delta}\left\{ \frac{y}{H} + q^{1/2} + \frac{y}{q} + \frac{yH}{q^{1/2}} + \frac{y^2 H}{q^2} \right\}.
\end{aligned}$$

We now choose $\eta = \delta/4 = \varepsilon/16$ and

$$H = \begin{cases} q^{1/4}, & \text{if } q > y^{2/3}, \\ qy^{-1/2}, & \text{if } y^{1/2} < q \le y^{2/3}. \end{cases}$$

Recalling the hypotheses of the theorem, it is now easily seen that $R = o(x/q)$. This result with (5.4)–(5.6) completes the proof.

Acknowledgment

We thank the referee for several helpful suggestions and the Institute for Advanced Study for its hospitality during the completion of our work on this paper. The first author would like to thank the University of Georgia for its hospitality during the commencement of the work on this paper.

REFERENCES

[de B] N. G. de Bruijn, *The asymptotic behavior of a function occurring in the theory of primes*, J. Indian Math. Soc. (NS) **15** (1951), 25–32.

[CEP] E. R. Canfield, P. Erdös, and C. Pomerance, *On a problem of Oppenheim concerning "Factorisatio Numerorum,"* J. Number Theory **17** (1983), 1–28.

[F-T] E. Fouvry and G. Tenenbaum, *Entiers sans grand facteur premier en progressions arithmétiques*, Proc. London Math. Soc. (2) (to appear).

[F1] J. B. Friedlander, *Integers without large prime factors*, Nederl. Akad. Wetensch. Proc. Ser A **76** (1973), 443–451.

[F2] ——, *Integers without large prime factors.* II, Acta Arith. **39** (1981), 53–57.

[F-I] J. B. Friedlander and H. Iwaniec, *On Bombieri's asymptotic sieve*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **5** (1978), 719–756.

[G] A. Granville, *Integers, without large prime factors, in arithmetic progressions*, Acta Math. (to appear).

[H-R] H. Halberstam and H. E. Richert, *Sieve methods*, Academic Press, London and New York, 1974.

[Hi] A. Hildebrand, *On the number of positive integers* $\leq x$ *and free of prime factors* $> y$, J. Number Theory **22** (1986), 289–307.

[Ho] C. Hooley, *On the Brun-Titchmarsh theorem*, J. Reine Angew. Math. **255** (1972), 60–79.

[I] H. Iwaniec, *On the Brun-Titchmarsh theorem*, J. Math. Soc. Japan **34** (1982), 95–123.

[L] M. Laborde, *Nombres presque-premiers dans de petits intervalles*, Sém. de Théorie des Nombres, Bordeaux, 1977–1978, exp. no. 15.

MATHEMATICAL INSTITUTE, HUNGARIAN ACADEMY OF SCIENCES, REÁLTANODA U. 13–15, H-1053 BUDAPEST, HUNGARY

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF GEORGIA, ATHENS, GEORGIA 30602