Integers, without large prime factors, in arithmetic progressions, II

Andrew Granville †

University of Georgia, Athens GA 30602, USA

and

The Isaac Newton Institute for the Mathematical Sciences,

20 Clarkson Road, Cambridge CB2 OEH, UK.

Abstract: We show that, for any fixed $\varepsilon > 0$, there are asymptotically the same number of integers up to x, that are composed only of primes $\leq y$, in each arithmetic progression $(\bmod\, q)$, provided that $y \geq q^{1+\varepsilon}$ and $\log x / \log q \to \infty$ as $y \to \infty$: this improves on previous estimates.

† An Alfred P. Sloan Research Fellow. Supported, in part, by the National Science Foundation

Integers, without large prime factors, in arithmetic progressions, II

Andrew Granville

1. Introduction.

The study of the distribution of integers with only small prime factors arises naturally in many areas of number theory; for example, in the study of large gaps between prime numbers, of values of character sums, of Fermat's Last Theorem, of the multiplicative group of integers modulo m, of S–unit equations, of Waring's problem, and of primality testing and factoring algorithms. For over sixty years this subject has received quite a lot of attention from analytic number theorists and we have recently begun to attain a very precise understanding of their distribution.

Let $\Psi(x,y)$ be the number of integers $\leq x$ that are free of prime factors $> y$, $\Psi_q(x,y)$ be the number of such integers that are also coprime to q, and $\Psi(x,y;a,q)$ be the number of such integers in the congruence class $a \pmod{q}$. Our goal is to prove that the asymptotic formula

$$\Psi(x,y;a,q) \sim \frac{\Psi_q(x,y)}{\phi(q)} \tag{1.1}$$

holds whenever a is coprime to q, in as wide a range as possible. Currently the widest such range is given in (Granville, 1993), where the estimate

$$\Psi(x,y;a,q) = \frac{\Psi_q(x,y)}{\phi(q)} \left\{1 + O\left(\frac{\log q}{\log y}\right)\right\} \tag{1.2}$$

is shown to hold uniformly in the range

$$(a,q)=1, \qquad x \geq y \geq 2, \qquad q \leq \min\{x, y^N\}, \tag{1.3}$$

for any fixed $N > 0$; this implies (1.1) only if $\log y / \log q \to \infty$ as $y \to \infty$.

In this paper we will consider what happens when q is a small power of y, and x is at least a large power of y. Fouvry and Tenenbaum (1991) gave a 'Bombieri–Vinogradov type result' for such a range but, currently, the best estimates for individual progressions are given in (Granville, 1993), building on work of Friedlander (1981), and of Balog and Pomerance (1992): For any fixed N in the range $0 < N < 4/3$ and $\varepsilon > 0$, we have the estimate

$$\Psi(x, y; a, q) \asymp \frac{\Psi_q(x, y)}{\phi(q)}$$

uniformly in the range

$$(a, q) = 1,\ y \geq 2,\ q \leq y^N,\ x \geq \max\{y^{3/2+\varepsilon}, yq^{3/4+\varepsilon}\}.$$

Our main result gives a stronger estimate for any $q \leq y^{1-\varepsilon}$:

Theorem. *For any given $\varepsilon > 0$, there exists a constant $c > 0$ such that the estimate*

$$\Psi(x, y; a, q) = \frac{\Psi_q(x, y)}{\phi(q)} \left\{1 + O\left(\frac{1}{u^c}\frac{\log q}{\log y} + \frac{1}{\log y}\right)\right\} \tag{1.4}$$

holds uniformly in the range

$$(a, q) = 1, \qquad x \geq y \geq q^{1+\varepsilon}, \qquad q \geq 2. \tag{1.5}$$

From this we deduce that (1.1) holds uniformly as $y \to \infty$ for $y \geq q^{1+\varepsilon}$ provided $\log x/\log q \to \infty$.

It looks hopeless to prove (1.1) uniformly when y is a small fixed power of q; for this would imply improvements on what is known for the famous open problem of proving that there exists a prime $\ll_\varepsilon q^\varepsilon$ which is not a quadratic residue $\pmod q$. There may be similar difficulties in proving (1.1) when x and q are both *any* given fixed powers of y.

The proof of the Theorem is built up of a number of increasingly complicated ideas. In order to make these more accessible we have chosen to first present a stronger result with an easier proof in section 4, but which only works for those q for which the primes are 'well–distributed' modulo q. Then, in section 5, we modify our proof so that it works for those q for which the 'P2s' (integers with no more than two prime factors) are 'well–distributed' modulo q. Finally, in section 6, we modify this further, and complete the proof of the Theorem.

As preparation, we discuss, in section 2, estimates for $\Psi_q(x, y)$ which we will need in sections 4, 5 and 6 (we save the proofs of these estimates, which are straightforward given the ideas of (Hildebrand and Tenenbaum, 1986), until Appendix One), and in section 3 we present a number of functional equations that will be useful. Finally, in Appendix Two, we present modifications of work of Mikawa (1989), to prove a result that we will need on 'P2s' in arithmetic progressions.

Notation: Throughout c and ε are taken to be absolute positive constants; however, they may change value from one proof to another. Writing $\Psi(x, y; a/d, q)$ is notationally consistent with the definition in the second paragraph.

Acknowledgements: I'd like to especially thank Henryk Iwaniec with whom I had a number of very helpful conversations, Hiroshi Mikawa for his correspondance and also Antal Balog, Brian Conrey, John Friedlander, Amit Ghosh and Dan Goldston for their comments.

2. Estimates involving $\Psi_q(x, y)$.

In 1930, Dickman showed that for any fixed $u > 0$,

$$\Psi(x, y) \sim x\rho(u) \qquad (x \to \infty, \quad y = x^{1/u})$$

where $\rho(u)$, the *Dickman function*, equals 1 for $0 \le u \le 1$, and is the continuous solution of the differential difference equation $u\rho'(u) + \rho(u-1) = 0$ for $u > 1$. Hildebrand (1986) proved this estimate uniformly for all $y \ge \exp\left(c(\log\log x)^{5/3+\varepsilon}\right)$ (for any fixed $\varepsilon > 0$), using the functional equation

$$\Psi(x, y)\log x \;=\; \int_1^x \frac{\Psi(t, y)}{t}\mathrm{d}t + \sum_{\substack{p^m \le x \\ p \le y}} \Psi\Big(\frac{x}{p^m}, y\Big)\log p. \tag{2.1}$$

In 1986 Hildebrand and Tenenbaum obtained precise estimates for $\Psi(x, y)$ for all $x \ge y \ge 2$. Their starting point was an old observation of Rankin that, for any $\sigma > 0$,

$$\Psi(x, y) \le \sum_{\substack{n \ge 1 \\ p|n \Rightarrow p \le y}} \left(\frac{x}{n}\right)^{\sigma} = x^{\sigma}\zeta(\sigma, y),$$

where $\zeta(s, y) = \prod_{p \le y}(1 - p^{-s})^{-1}$. The right-hand side of this equation is minimized for $\sigma = \alpha = \alpha(x, y)$, the (unique) real solution of the equation

$$\sum_{p \le y} \frac{\log p}{p^{\alpha} - 1} \;=\; \log x, \tag{2.2}$$

so that $\Psi(x,y) \leq x^\alpha \zeta(\alpha, y)$. Using the saddle point method, Hildebrand and Tenenbaum gave an asymptotic formula for $\Psi(x,y)$ in terms of α (which showed that Rankin's upper bound is too large (asymptotically) by only a small factor). From this they were able to deduce a very accurate 'local' result, which gives the estimate

$$\Psi(cx,y) = c^{\alpha(x,y)}\Psi(x,y)\{1+O(1/u)\} \tag{2.3}$$

uniformly in the range $x \geq y \geq \log x$ with $1 \leq c \leq y$.

In Appendix One we will sketch a proof of a similar result for Ψ_q; that the estimate

$$\Psi_q(cx,y) = c^{\alpha(x,y)}\Psi_q(x,y)\{1+O(1/u)\} \tag{2.4}$$

holds uniformly in the range

$$x \geq y \geq y_0,\ y \geq \log^3 x \quad \text{with} \quad 1 \leq c \leq y \quad \text{and} \quad q \leq y^N \tag{2.5}$$

for any fixed $N > 0$. (Note that $\alpha > 3/5$ in this range (by (2.4) of (Hildebrand and Tenenbaum, 1986)).)

Furthermore, we will deduce the following two easy consequences of (2.4):

First that the bound

$$\int_1^x \frac{\Psi(t,y;a,q)}{t}\mathrm{d}t \ + \sum_{p\leq y,\ p\nmid q,\ m\geq 2} \Psi\Big(\frac{x}{p^m}, y; \frac{a}{p^m}, q\Big)\log p \ \ll \ \frac{\Psi_q(x,y)}{\phi(q)} \tag{2.6}$$

holds uniformly in the range

$$(a,q)=1, \qquad x \geq y \geq q^B, \qquad q \geq 2, \qquad \text{and} \qquad y \geq \log^3 x, \tag{2.7}$$

with $x \geq y^4$, for any fixed $B > 0$.

Secondly, for δ in the range $0 < \delta < 1$ and for integer N and positive reals y and z such that $(1+1/z)^N = y^\delta$, define $w_i = y^{1-\delta}(1+1/z)^i$, so that $w_N = y$. For any $\Delta > 0$ we have

$$\sum_{i=0}^{N-1} w_i \left\{ \Psi_q\Big(\frac{x}{w_{i+1}}, y\Big) - (1-\Delta)\Psi_q\Big(\frac{x}{w_i}, y\Big)\right\} \ \gg \ (\Delta - 1/z)z\ \Psi_q(x,y)\log x, \tag{2.8}$$

uniformly in the range (2.7).

We shall prove all of our results on $\Psi(x,y;a,q)$ in the range (2.7). In order to extend our estimates to all of (1.5), we need the following result, which is a straightforward consequence of Proposition 1 of (Granville, 1993) (using (1.2) and (2.4) above):

Proposition 0. *Given $y \geq q$, define $x^* = x^*(y)$ so that $y = \log^4 x^*$. Suppose that there exists Δ such that*

$$\left| \Psi(x,y;a,q) \; - \; \frac{\Psi_q(x,y)}{\phi(q)} \right| \; \leq \; \Delta \frac{\Psi_q(x,y)}{\phi(q)},$$

for all $(a,q)=1$ and x in the range $x^*/y^4 \leq x \leq x^*$. Then

$$\Psi(x,y;a,q) \; = \; \frac{\Psi_q(x,y)}{\phi(q)} \left\{ 1 + O\left(\Delta + \frac{1}{y} \right) \right\},$$

uniformly for all $(a,q)=1$ and $x \geq x^*$.

3. Functional Equations.

We start by giving functional equations for $\Psi(x,y;a,q)$ and $\Psi_q(x,y)$, analogous to Hildebrand's equation (2.1): The idea is to evaluate $\sum\limits_{\substack{n \leq x, \; n \equiv a \pmod q \\ p|n \Rightarrow p \leq y, p \nmid q}} \log n$ in two different ways. First by partial summation, and second by writing each $\log n$ as $\sum\limits_{p^m|n} \log p$, and then swapping the order of summation. This leads to the identity

$$\Psi(x,y;a,q) \log x \; = \; \int_1^x \frac{\Psi(t,y;a,q)}{t} \mathrm{d}t + \sum_{\substack{p^m \leq x \\ p \leq y, p \nmid q}} \Psi\Big(\frac{x}{p^m}, y; \frac{a}{p^m}, q\Big) \log p \; . \tag{3.1}$$

Summing (3.1) over all integers $(a,q)=1$ we get

$$\Psi_q(x,y) \log x \; = \; \int_1^x \frac{\Psi_q(t,y)}{t} \mathrm{d}t + \sum_{\substack{p^m \leq x \\ p \leq y, p \nmid q}} \Psi_q\Big(\frac{x}{p^m}, y\Big) \log p \; . \tag{3.2}$$

Notice that (3.2) is just (3.1) with each term of the form $\Psi(t,y;b,q)$ replaced by $\Psi_q(t,y)$.

(2.6) allows us to modify these functional equations to

$$\Psi(x,y;a,q) \log x \; = \; \sum_{p \leq y, p \nmid q} \Psi\Big(\frac{x}{p}, y; \frac{a}{p}, q\Big) \log p \; + \; O\left(\frac{\Psi_q(x,y)}{\phi(q)} \right) \; , \tag{3.3}$$

and

$$\Psi_q(x,y) \log x \; = \; \sum_{p \leq y, p \nmid q} \Psi_q\Big(\frac{x}{p}, y\Big) \log p \; + \; O(\Psi_q(x,y)) \; , \tag{3.4}$$

uniformly in the range (2.7) with $x \geq y^4$.

Suppose that $x \geq y^5$. If we replace each term $\Psi(t,y;b,q)$ on the right side of (3.3) with the expression given by (3.3) for $\Psi(t,y;b,q)$, then the error term is

$$\ll \frac{1}{\phi(q)} \left(\Psi_q(x,y) + \frac{1}{\log x} \sum_{p \leq y, p \nmid q} \Psi_q\left(\frac{x}{p}, y\right) \right) \log p \ll \frac{\Psi_q(x,y)}{\phi(q)}$$

by (3.4). Therefore we have

(3.5)

$$\Psi(x,y;a,q) \log x = \sum_{p_1,p_2 \leq y,\ p_1,p_2 \nmid q} \frac{\log p_1 \log p_2}{\log (x/p_1)} \Psi\left(\frac{x}{p_1p_2}, y; \frac{a}{p_1p_2}, q\right) + O\left(\frac{\Psi_q(x,y)}{\phi(q)}\right),$$

and similarly

$$\Psi_q(x,y) \log x = \sum_{p_1,p_2 \leq y,\ p_1,p_2 \nmid q} \frac{\log p_1 \log p_2}{\log (x/p_1)} \Psi_q\left(\frac{x}{p_1p_2}, y\right) + O(\Psi_q(x,y)), \tag{3.6}$$

uniformly in the range (2.7) with $x \geq y^5$.

4. Proof when the primes are well–distributed modulo q.

In this section we will prove a strong form of the theorem for the case when we know that the primes are well distributed among all the arithmetic progressions modulo q; specifically, when we know that

$$\sum_{\substack{x \leq p \leq x + x/z \\ p \equiv a \pmod q}} \log p \gg \frac{x}{z\phi(q)}, \tag{4.1}$$

for all $(a,q)=1$, and some given $z \geq 1$, provided $x > q^A$.
(By suitably modifying the proof of Linnik's theorem given in (Bombieri, 1987, pgs.54–55), one can prove such a result provided $x > (qz)^B$ for some fixed $B > 0$, and $L(s,\chi) \neq 0$ for all $\sigma \geq 1 - \frac{c}{\log T}$, $|t| \leq T$, with $T = qz \log^2 x$, for all primitive characters χ of modulus q.)

Under the assumption of (4.1) for $y^{1-\delta} \leq x \leq y$, with z a fixed power of y, we shall prove that for any given $\varepsilon > 0$, there exists $c > 0$ such that the estimate

$$\Psi(x,y;a,q) = \frac{\Psi_q(x,y)}{\phi(q)} \left\{ 1 + O\left(\frac{1}{e^{cu}} \frac{\log q}{\log y} + \frac{1}{y^{1/5}} \right) \right\} \tag{4.2}$$

holds uniformly in the range (2.7), for $B = A(1+\varepsilon)$. This may then be extended to all of $x \ge y \ge q^B$ by using Proposition 0.

Remark 1: The error term here is better than that in the Theorem by an exponentiation.

Remark 2: With a little more care the $O(1/y^{1/5})$ in the error term may be replaced by $O(\log^3 y/y)$.

Proof of (4.2): The result holds for $1 \le u \le u_0$ by (1.2), for fixed u_0, so assume henceforth that $u \ge u_0$.

Choose $\delta > 0$ such that $(1+\varepsilon)(1-\delta) > 1$, and $z = y^{1/5} + O(1)$, so that there exists an integer N such that $(1+1/z)^N = y^\delta$. Define $w_i = y^{1-\delta}(1+1/z)^i$ for each integer $i \ge 0$. Choose x^* so that $y = \log^3 x^*$, and define

$$\Delta(x) := \max_{xy^2 \ge x' \ge x} \max_{(a,q)=1} \left| 1 - \left\{ \Psi(x', y; a, q) \Big/ \frac{\Psi_q(x', y)}{\phi(q)} \right\} \right| ,$$

so that

$$(1+\Delta(x)) \frac{\Psi_q(x', y)}{\phi(q)} \ \ge\ \Psi(x', y; a, q) \ \ge\ (1-\Delta(x)) \frac{\Psi_q(x', y)}{\phi(q)}$$

for $xy^2 \ge x' \ge x$. Using the functional equation (3.1) we obtain, for any $(a,q)=1$,

$$\begin{aligned} \Psi(x, y; a, q) \log x \ &\ge\ \int_{x/y^2}^{x} \frac{\Psi(t, y; a, q)}{t} \mathrm{d}t + \sum_{\substack{p^m \le y^2 \\ p \le y,\, p \nmid q}} \Psi\Big(\frac{x}{p^m}, y; \frac{a}{p^m}, q\Big) \log p \\ &\ge\ \frac{(1-\Delta(x/y^2))}{\phi(q)} \left(\Psi_q(x, y) \log x - G_q(x, y) \right) \\ &\quad + \sum_{\substack{y^{1-\delta} < p^m \le y \\ p \nmid q}} \left\{ \Psi\Big(\frac{x}{p^m}, y; \frac{a}{p^m}, q\Big) - (1-\Delta(x/y^2)) \frac{\Psi_q\big(\frac{x}{p^m}, y\big)}{\phi(q)} \right\} \log p \,, \end{aligned} \tag{4.3}$$

using (3.2), where $G_q(x, y)$ is the contribution of those terms $\Psi_q(t, y)$ in (3.2) with $t \le x/y^2$. Using the trivial inequality $\Psi_q(t, y) \le \Psi_q(x/y^2, y)$ for all $t \le x/y^2$, we see that $G_q(x, y) \ll \Psi_q(x/y^2, y) y \log x / \log y$ provided $x \ge y^3$.

Cutting the sum of (4.3) into intervals $(w_i, w_{i+1}]$ and into arithmetic progressions modulo q, and choosing $v = v_{i,b} \in (w_i, w_{i+1}]$ so as to minimize

$$\left\{ \Psi\Big(\frac{x}{v}, y; \frac{a}{b}, q\Big) - (1-\Delta(x/y^2)) \frac{\Psi_q\big(\frac{x}{v}, y\big)}{\phi(q)} \right\} ,$$

we obtain

$$\sum_{i=0}^{N-1} \sum_{(b,q)=1} \sum_{\substack{w_i<p^m\leq w_{i+1}\\ p^m\equiv b \pmod q}} \left\{ \Psi\Big(\frac{x}{p^m}, y; \frac{a}{p^m}, q\Big) - (1-\Delta(x/y^2))\frac{\Psi_q\Big(\frac{x}{p^m}, y\Big)}{\phi(q)} \right\} \log p$$

$$(4.4) \geq \sum_{i=0}^{N-1} \sum_{(b,q)=1} \left\{ \Psi\Big(\frac{x}{v_{i,b}}, y; \frac{a}{b}, q\Big) - (1-\Delta(x/y^2))\frac{\Psi_q\Big(\frac{x}{v_{i,b}}, y\Big)}{\phi(q)} \right\} \sum_{\substack{w_i<p^m\leq w_{i+1}\\ p^m\equiv b \pmod q}} \log p$$

$$\gg \frac{1}{z\phi(q)} \sum_{i=0}^{N-1} w_i \sum_{(b,q)=1} \left\{ \Psi\Big(\frac{x}{w_{i+1}}, y; \frac{a}{b}, q\Big) - (1-\Delta(x/y^2))\frac{\Psi_q\Big(\frac{x}{w_i}, y\Big)}{\phi(q)} \right\}$$

$$= \frac{1}{z\phi(q)} \sum_{i=0}^{N-1} w_i \left\{ \Psi_q\Big(\frac{x}{w_{i+1}}, y\Big) - (1-\Delta(x/y^2))\Psi_q\Big(\frac{x}{w_i}, y\Big) \right\}$$

$$(4.5) \gg (\Delta(x/y^2) - 1/z)\frac{\Psi_q(x,y)}{\phi(q)}\log x,$$

by (4.1) and then (2.8) (for $x \leq x^*$).

Plugging this and the bound for $G_q(x,y)$ into (4.3) and dividing through by $\frac{\Psi_q(x,y)}{\phi(q)}\log x$, we obtain, as $\Delta(x/y^2) \leq 1$, and where $c_1 > 0$ is the absolute constant implied in (4.5),

$$\begin{aligned} 1 - \frac{\Psi(x,y;a,q)}{\Psi_q(x,y)/\phi(q)} &\leq (1-c_1)\Delta(x/y^2) + O\left(\frac{1}{z} + \frac{y}{\log y}\frac{\Psi_q(x/y^2,y)}{\Psi_q(x,y)}\right) \\ &\leq (1-c_1)\Delta(x/y^2) + O(1/y^{1/5}), \end{aligned}$$

by (2.4) as $\alpha > 3/5$ in our range.

When we proceed in a similar fashion to examine an upper bound for $\Psi(x,y;a,q)$, we obtain the inequality

$$\frac{\Psi(x,y;a,q)}{\Psi_q(x,y)/\phi(q)} - 1 \leq (1-c_1)\Delta(x/y^2) + O(1/y^{1/5}).$$

(The proof is exactly analogous to that before except that we get $(\Delta(x/y^2)(1-1/z)-1/z)$ instead of $(\Delta(x/y^2)-1/z)$ in (4.5); however, as $\Delta(x/y^2) \ll 1$ by (1.2), this only affects the constant implied by the 'O' term above. A similar (unimportant) difference occurs in the 'analogous arguments' in both the next two sections.)

So, taking $x = x'$ in the two equations immediately above, for x' in the range $xy^2 \geq x' \geq x$, we obtain the inequality

$$\Delta(x) \ \leq \ (1 - c_1) \max_{xy^2 \geq x' \geq x} \Delta(x'/y^2) \ + \ O(1/y^{1/5}).$$

Thus, as $\Delta(x'/y^2) \leq \max \{\Delta(x/y^2), \Delta(x)\}$ for any x' in this range, we see that

$$\text{(4.6)} \quad \text{either (a)} \ \ \Delta(x) \leq (1 - c_1)\Delta(x/y^2) \ + \ O(1/y^{1/5}), \ \ \text{or (b)} \ \ \Delta(x) \ = \ O(1/y^{1/5}).$$

Finally, suppose that u_0 is fixed to be sufficiently large, $x = y^u \leq x^*$ and let $r = [(u - u_0)/2]$. A straightforward induction hypothesis using (4.6) gives

$$\Delta(x_0 y^{2r}) \ \leq \ (1 - c_1)^r \Delta(x_0) \ + \ O(1/y^{1/5}),$$

which implies (4.2) as $\Delta(x_0) \ll \log q/\log y$ by (1.2), and as $r \asymp u$.

Remark: Fouvry and Tenenbaum (1991) proved a much stronger estimate than either (1.2) or (1.4) for the range $q \leq \log^N y$, for fixed $N > 0$; specifically that

$$\Psi(x, y; a, q) = \frac{\Psi_q(x, y)}{\phi(q)} \left\{ 1 + O(\exp(-c\sqrt{\log y})) \right\} \tag{4.7}$$

for $x \geq y \geq \exp(c(\log\log x)^2)$. It is a simple exercise to prove that (4.1) holds, with $z = y^c$, for any x in the range $y^{1/2} \leq x \leq y$, for such q. It is then a straightforward task to modify the proof above (but taking $\Delta(x_0) \ll \exp(-c\sqrt{\log y})$), to prove that for any fixed $N > 0$, the estimate

$$\Psi(x, y; a, q) = \frac{\Psi_q(x, y)}{\phi(q)} \left\{ 1 + O\left(\frac{1}{e^{c(u + \sqrt{\log y})}} \ + \ \frac{1}{y^c} \right) \right\} \tag{4.8}$$

holds uniformly in the range $x \geq y \geq 2$, $q \leq \log^N y$. This is an improvement of (4.7) for $y \ll \exp(c(\log x)^{2/3})$.

5. Modifying the proof to use $P2$'s.

The main problem with the method described in the previous section is that one needs a strong result, such as (4.1), on the distribution of primes in arithmetic progressions in order to make it work. To obtain such results in a wide range seems to be beyond the scope of what is possible in the foreseeable future. However, it is possible to prove results of similar strength in a much wider range if one only requires an understanding of the distribution of $P2$'s, that is integers with at most two (not necessarily distinct) prime factors. So, define

$$\Lambda_2(n) = \begin{cases} \log^2 p & \text{if } n = p \text{ is prime,} \\ \log p_1 \log p_2 & \text{if } n = p_1 p_2 \text{ where } p_1 \text{ and } p_2 \text{ are primes,} \\ 0 & \text{otherwise .} \end{cases}$$

It requires only straightforward modifications of the proof of Theorem 13 in (Iwaniec, 1982) to prove that, for $A = 1.845$, we have

$$\sum_{\substack{x \le n \le x + x/z \\ n \equiv a \pmod q}} \Lambda_2(n) \gg \frac{x \log x}{z \phi(q)} , \tag{5.1}$$

for all $(a, q) = 1$, and any $1 \le z \le \log^2 x$, provided $x \ge q^A$. Using this we shall prove that for any $\varepsilon > 0$, there exists $c > 0$ such that (1.4) holds uniformly in the range (2.7) with $B = A(1 + \varepsilon)$. This may then be extended to all of $x \ge y \ge q^B$ using Proposition 0.

Proof of (1.4) in the range (2.7): Define x^* and $\Delta(x)$ as in the previous section. We choose $z = \log y + O(1)$ so that N is an integer.

From the functional equation (3.3), we obtain that $\Psi(x, y; a, q) \log x =$

$$\begin{aligned} &= \frac{(1 - \Delta(x/y^2))}{\phi(q)} \sum_{p \le y, p \nmid q} \Psi_q\Big(\frac{x}{p}, y\Big) \log p \; + \; O\left(\frac{\Psi_q(x, y)}{\phi(q)}\right) \\ &\quad + \sum_{p \le y, p \nmid q} \left\{ \Psi\Big(\frac{x}{p}, y; \frac{a}{p}, q\Big) \; - \; \frac{(1 - \Delta(x/y^2))}{\phi(q)} \Psi_q\Big(\frac{x}{p}, y\Big) \right\} \log p \\ &\ge \frac{\Psi_q(x, y)}{\phi(q)} ((1 - \Delta(x/y^2)) \log x - O(1)) \\ &\quad + \frac{1}{\log x} \sum_{\substack{y^{1-\delta} < p \le y \\ p \nmid q}} \left\{ \Psi\Big(\frac{x}{p}, y; \frac{a}{p}, q\Big) \; - \; \frac{(1 - \Delta(x/y^2))}{\phi(q)} \Psi_q\Big(\frac{x}{p}, y\Big) \right\} \log^2 p \end{aligned}$$

using (3.4). Next, from the functional equation (3.5), we obtain that $\Psi(x,y;a,q)\log x =$

$$\begin{aligned}
&= \frac{(1-\Delta(x/y^2))}{\phi(q)} \sum_{p_1,p_2\le y,\ p_1,p_2\nmid q} \Psi_q\Big(\frac{x}{p_1p_2},y\Big)\frac{\log p_1\log p_2}{\log (x/p_1)} + O\left(\frac{\Psi_q(x,y)}{\phi(q)}\right)\\
&\quad + \sum_{p_1,p_2\le y,\ p_1,p_2\nmid q} \left\{\Psi\Big(\frac{x}{p_1p_2},y;\frac{a}{p_1p_2},q\Big) - \frac{(1-\Delta(x/y^2))}{\phi(q)}\Psi_q\Big(\frac{x}{p_1p_2},y\Big)\right\}\frac{\log p_1\log p_2}{\log (x/p_1)}\\
&\ge \frac{\Psi_q(x,y)}{\phi(q)}((1-\Delta(x/y^2))\log x - O(1))\\
&\quad + \frac{1}{\log x}\sum_{\substack{y^{1-\delta}<p_1,p_2\le y\\ p_1,p_2\nmid q}} \left\{\Psi\Big(\frac{x}{p_1p_2},y;\frac{a}{p_1p_2},q\Big) - \frac{(1-\Delta(x/y^2))}{\phi(q)}\Psi_q\Big(\frac{x}{p_1p_2},y\Big)\right\}\log p_1\log p_2
\end{aligned}$$

using (3.6).

Now, if we add the two equations above together, and, as in the previous section, cut the sum into intervals (w_i,w_{i+1}), and into arithmetic progressions modulo q, we get that,

$$\begin{aligned}
\Psi(x,y;a,q) \ \ge\ & \frac{\Psi_q(x,y)}{\phi(q)}((1-\Delta(x/y^2)) - O(1/\log x)) + \frac{1}{2\log^2 x}\times\\
\times & \sum_{i=0}^{N-1}\sum_{(b,q)=1}\left\{\Psi\Big(\frac{x}{v_{i,b}},y;\frac{a}{b},q\Big) - (1-\Delta(x/y^2))\frac{\Psi_q\left(\frac{x}{v_{i,b}},y\right)}{\phi(q)}\right\}\sum_{\substack{w_i<n\le w_{i+1}\\ n\equiv b \pmod q}}\Lambda_2(n).
\end{aligned}\tag{5.2}$$

Using (5.1) (instead of (4.1)) and (2.8), and otherwise proceeding in the manner of the previous section, we find that the sum here is

$$\gg\ (\Delta(x/y^2)-1/z)\ u\log^2 y\ \frac{\Psi_q(x,y)}{\phi(q)};$$

and so, for some $c>0$,

$$\begin{aligned}
1-\Psi(x,y;a,q)\Big/\frac{\Psi_q(x,y)}{\phi(q)} \ &\le\ \Delta(x/y^2)+O(1/\log x) - \frac{3c}{u}(\Delta(x/y^2)-1/z)\\
&\le\ (1-3c/u)\Delta(x/y^2) + O\left(\frac{1}{u\log y}\right).
\end{aligned}$$

An analogous argument gives the same upper bound for $\Psi(x,y;a,q)/\frac{\Psi_q(x,y)}{\phi(q)} - 1$. So, proceeding as in the previous section, we deduce that

$$\text{either (a)}\ \ \Delta(x)\ \le\ \left(1-\frac{2c}{u}\right)\Delta(x/y^2)+O\left(\frac{1}{u\log y}\right),\ \text{ or (b)}\ \ \Delta(x)=O\left(\frac{1}{\log y}\right).\tag{5.3}$$

Suppose that u_0 is fixed to be sufficiently large, $x = y^u \leq x^*$ and let $r = [(u-u_0)/2]$. A straightforward induction hypothesis using (5.3) gives that

$$\Delta(x_0 y^{2r}) \ \leq \ \Delta(x_0) \prod_{i=1}^{r} \left(1 - \frac{2c}{u_0+2i}\right) \ + \ O\left(\frac{1}{\log y}\right) \ \ll \ \left(\frac{u_0}{u}\right)^c \Delta(x_0) + \frac{1}{\log y},$$

which implies (1.4) as $\Delta(x_0) \ll \log q/\log y$ by (1.2).

6. Proof of the Theorem.

Recently Mikawa (1989) has shown that, for any fixed $\varepsilon > 0$, there exists a P2, in almost all reduced residue classes modulo q, which is $\ll q^{1+\varepsilon}$. We now modify the method of the previous section so that we can use such an 'almost all' result in our proof. In fact we will only need that (5.1) holds for 'most' arithmetic progressions modulo q, in order to prove our Theorem.

In Appendix Two, we will modify Mikawa's proof for our needs. An immediate consequence of Proposition 1 (of Appendix Two) is

Corollary. *Fix $\varepsilon > 0$. Then (5.1) holds uniformly for at least $2/3$rd's of the arithmetic progressions $a \pmod q$, for any $1 \leq z \leq 2\log x$, and for any $x \geq q^{1+\varepsilon}$.*

Using this we now prove that (1.4) holds uniformly in the range (2.7), with $B = 1+\varepsilon$. Then, as before, we use Proposition 0 to extend this to the whole of (1.5).

Proof: Note that, for any i and b, we have

$$\left\{ \Psi\left(\frac{x}{v_{i,b}}, y; \frac{a}{b}, q\right) - (1-\Delta(x/y^2)) \frac{\Psi_q\left(\frac{x}{v_{i,b}}, y\right)}{\phi(q)} \right\} \ \leq \ 2\Delta(x/y^2) \frac{\Psi_q\left(\frac{x}{v_{i,b}}, y\right)}{\phi(q)}$$

$$\leq \ 2\Delta(x/y^2) \frac{\Psi_q\left(\frac{x}{w_i}, y\right)}{\phi(q)}. \tag{6.1}$$

We follow the proof of the previous section up to (5.2), which still remains valid. However, using (6.1), the sum in (5.2) is now

$$\gg \frac{\log y}{z\phi(q)} \sum_{i=0}^{N-1} w_i \Bigg\{ \Psi_q\Big(\frac{x}{w_{i+1}}, y\Big) - (1-\Delta(x/y^2))\Psi_q\Big(\frac{x}{w_i}, y\Big) - \sum_{\substack{(b,q)=1 \\ (5.1) \text{ fails}}} 2\Delta(x/y^2) \frac{\Psi_q\Big(\frac{x}{w_i}, y\Big)}{\phi(q)} \Bigg\}$$

$$\geq \frac{\log y}{z\phi(q)} \sum_{i=0}^{N-1} w_i \left\{ \Psi_q\Big(\frac{x}{w_{i+1}}, y\Big) - \left(1 - \frac{1}{3}\Delta(x/y^2)\right) \Psi_q\Big(\frac{x}{w_i}, y\Big) \right\},$$

by the Hypothesis. From here we proceed exactly as in the previous section (but this time taking $\Delta = \frac{1}{3}\Delta(x/y^2)$ in (2.8)), to deduce (1.4).

Remark: Using the methods of this section, it is possible to obtain a result like (1.4) in a wider range, but it would be necessary to use the full strength of Proposition 1, and to prove a stronger version of (2.8). Specifically if, for a given value of q, we take $y_0 = C\phi(q)\log^7 q$ (with C as in Proposition 1), then for any $y \geq 2y_0$ we have

$$\Psi(x, y; a, q) = \frac{\Psi_q(x, y)}{\phi(q)} \left\{ 1 + O\left(\frac{1}{\exp\left(c\frac{\log(y/y_0)}{\log y}\log^2 u\right)} + \frac{1}{u^c} + \frac{1}{\log y} \right) \right\}.$$

However it does not seem worth going into details here, as we are as yet unable to reach our true goal of having such an estimate for some range of values of $y \leq q$.

Appendix One. The details of section 2.

Proof of (2.4):

Fouvry and Tenenbaum (1991) gave an estimate for $\Psi_q(x, y)$ in terms of $\Psi(x, y)$ which implies that

$$\Psi_q(x, y) = \frac{\phi(q)}{q}\Psi(x, y)\left\{1 + O\left(\frac{(\log\log 2y)^2}{\log y}\right)\right\} \tag{A.1}$$

in the range (1.3) with $u \le \log y$. By writing

$$\frac{\Psi_q(cx,y)}{c^{\alpha(x,y)}\Psi_q(x,y)} = \frac{\Psi_q(cx,y)}{\frac{\phi(q)}{q}\Psi(cx,y)} \, \frac{\Psi(cx,y)}{c^{\alpha(x,y)}\Psi(x,y)} \, \frac{\frac{\phi(q)}{q}\Psi(x,y)}{\Psi_q(x,y)},$$

and then applying (A.1), (2.3) and (A.1), respectively, to the terms on the right side, we obtain the estimate

$$\Psi_q(cx,y) = c^{\alpha(x,y)}\Psi_q(x,y)\left\{1 + O\left(\frac{1}{u} + \frac{(\log\log 2y)^2}{\log y}\right)\right\}$$

uniformly in the range (1.3) with $u \le \log y$ and $1 \le c \le y$. This implies (2.4) in the range (2.5) provided $u \le \sqrt{\log y}$.

We may henceforth assume that $\log y \le u^2 \le y$. Our proof proceeds almost exactly as in (Hildebrand and Tenenbaum, 1986, Theorem 3); we shall only expound upon the one non–trivial change needed. Define $\xi = \xi(u)$ as the non–zero solution of the equation $e^\xi = 1 + u\xi$; $\zeta_q(s,y) = \prod_{p\le y,\ p\nmid q}(1-p^{-s})^{-1}$; and $\alpha = \alpha_q(x,y)$ to be the (unique) real solution of the equation

$$\sum_{p\le y,\ p\nmid q} \frac{\log p}{p^\alpha - 1} = \log x.$$

The proof of (Hildebrand and Tenenbaum, 1986, Theorem 3) (which implies (2.3)) is easily modified to give the estimate

$$\Psi_q(cx,y) = c^{\alpha_q(x,y)}\Psi_q(x,y)\{1 + O(1/u)\} \tag{A.2}$$

in the range (2.5), with $u \ge \sqrt{\log y}$, provided we can prove the analogue to (Hildebrand and Tenenbaum, 1986, (3.5)) in this range: that is, the estimate

$$(1-\alpha_q(x,y))\log y = \xi(u) + O\left(\frac{1}{u} + e^{-\sqrt{\log y}}\right). \tag{A.3}$$

We will prove that

$$\alpha_q(x,y) = \alpha(x,y) + O\left(\frac{1}{\log x}\right) \tag{A.4}$$

holds uniformly in this range. This immediately implies (A.3) (as (A.3) is known to hold for $q = 1$, see (Hildebrand and Tenenbaum, 1986, (3.5))); and also shows that (2.4) follows from (A.2) in this range. It thus remains to give a

Proof of (A.4): It is easily shown that $\alpha = \alpha_q(x,y) \gg 1$ in this range, and so any $p^\alpha - 1 \geq \alpha \log p \gg \log p$. Also note that q has no more than $O(\log y)$ prime factors as $q \leq y^N$. Thus

$$\sum_{p\leq y} \frac{\log p}{p^\alpha - 1} = \sum_{p\leq y,\ p\nmid q} \frac{\log p}{p^\alpha - 1} + \sum_{p\leq y,\ p\mid q} \frac{\log p}{p^\alpha - 1}$$
$$= \log x + O(\log y) = \{u + O(1)\}\log y.$$

Define v so that $\alpha(y^v, y) = \alpha$; therefore $v = u + O(1)$ by comparing the equation above with (2.2). (Hildebrand and Tenenbaum, 1986, (6.6)) implies that

$$\frac{\partial}{\partial t}\alpha(y^t, y) \asymp -\frac{1}{t\log y} \tag{A.5}$$

for $y/\log y \geq t \geq t_0$. Thus

$$\alpha_q(x,y) - \alpha(x,y) = \int_u^v \left\{\frac{\partial}{\partial t}\alpha(y^t, y)\right\} \mathrm{d}t \asymp -(v-u)\frac{1}{\log x} \ll \frac{1}{\log x},$$

and (A.4) follows.

Proof of (2.6): By summing (2.6) over the arithmetic progressions (mod q), we obtain the estimate

$$\int_1^x \frac{\Psi_q(t,y)}{t}\mathrm{d}t + \sum_{p\leq y,\ p\nmid q,\ m\geq 2} \Psi_q\Big(\frac{x}{p^m}, y\Big)\log p \ll \Psi_q(x,y) \tag{A.6}$$

uniformly for the range (2.7) with $x \geq y^4$. We shall, in fact, first prove (A.6), and then deduce (2.6) as a consequence:

Proof of (A.6): Let $\alpha = \alpha(x,y)$. By (A.5), we have that $\alpha = \alpha(x/c, y) + O(1/\log x)$ for $1 \leq c \leq y$, so that, by (2.4),

$$\Psi_q(x/c, y) = c^{-\alpha(x,y)}\Psi_q(x,y)\{1 + O(1/u)\} \tag{A.7}$$

in the range (2.5); thus $\Psi_q(x/d, y) \ll \Psi_q(x,y)/d^\alpha$ uniformly for $1 \leq d \leq y^3$.

For any given prime $p \leq y$ define $k = [3\log y/\log p]$; by the above we see that

$$\sum_{m\geq 2} \Psi_q(x/p^m, y) = \sum_{m=2}^{k} \Psi_q(x/p^m, y) + \sum_{m=k+1}^{[2\log x]} \Psi_q(x/p^m, y)$$
$$\ll \sum_{m=2}^{k} \Psi_q(x,y)/p^{m\alpha} + \log x\,\Psi_q(x,y)/y^{3\alpha} \ll \Psi_q(x,y)/p^{2\alpha}$$

as $\alpha > 3/5$. Therefore

$$\sum_{p\le y,\ p\nmid q,\ m\ge 2} \Psi_q\Big(\frac{x}{p^m}, y\Big)\log p \ \ll\ \Psi_q(x,y)\sum_{p\le y}\frac{\log p}{p^{2\alpha}} \ \ll\ \Psi_q(x,y),$$

by the Prime Number Theorem. Finally note that

$$\int_1^x \frac{\Psi_q(t,y)}{t}\mathrm{d}t \ \le\ \sum_{j\ge 0}\int_{x/2^{j+1}}^{x/2^j}\frac{\Psi_q(x/2^j,y)}{t}\mathrm{d}t \ \le\ \sum_{j\ge 0}\Psi_q(x/2^j,y)\log 2 \ \ll\ \Psi_q(x,y).$$

Completion of the Proof of (2.6): To bound the expression on the left side of (2.6) we use the bounds

$$\Psi(t,y;a,q)\ \begin{cases} \ll\ \Psi_q(t,y)/\phi(q) & \text{for } t>y \\ \le\ \Psi_q(y,y)/\phi(q) & \text{for } t\le y; \end{cases}$$

the first of these bounds comes from (1.2), the second from the trivial inequality $\Psi(t,y;a,q)\le \Psi(y,y;a,q)$ and then (1.2). Thus the left side of (2.6) is bounded by $\frac{c}{\phi(q)}$ times (A.6) (for some absolute $c>0$) plus $\frac{\Psi_q(y,y)}{\phi(q)}$ times

$$\int_1^y \frac{\mathrm{d}t}{t}\ +\ \sum_{p\le y,\ p\nmid q,\ x\ge p^m\ge x/y}\log p \ \ll\ y,$$

by the Prime Number Theorem. Now $y\Psi_q(y,y)\ll \Psi_q(y^2,y)$ by (A.1), and so the left side of (2.6) is

$$\ll \frac{1}{\phi(q)}\left(\Psi_q(x,y)\ +\ y\Psi_q(y,y)\right)\ \ll\ \frac{\Psi_q(x,y)}{\phi(q)},$$

as $x\ge y^4$.

Proof of (2.8): Start by noting that (A.3) implies that

$$y^{1-\alpha}\ =\ u\xi(u)\left\{1+O\left(\frac{1}{u}\ +\ \frac{1}{\log y}\right)\right\} \tag{A.8}$$

for $\alpha=\alpha(x,y)$. Thus, by this and (A.7), we get that

$$\begin{aligned} w_N\Psi_q\Big(\frac{x}{w_N},y\Big)\ -\ w_0\Psi_q\Big(\frac{x}{w_0},y\Big)\ &=\ \Psi_q(x,y)\left(w_N^{1-\alpha}\ -\ w_0^{1-\alpha}\ +\ O\left(\frac{w_N^{1-\alpha}}{u}\right)\right) \\ &=\ \Psi_q(x,y)u\xi(u)\left\{1+O\left(\frac{1}{u^\delta}\ +\ \frac{1}{\log y}\right)\right\}, \qquad (A.9a) \end{aligned}$$

and

$$\begin{aligned}
\sum_{i=0}^{N-1} w_i \Psi_q\Big(\frac{x}{w_i}, y\Big) &= \Psi_q(x,y) \sum_{i=0}^{N-1} w_i^{1-\alpha}\{1+O(1/u)\} \\
&= \Psi_q(x,y) \frac{w_N^{1-\alpha} - w_0^{1-\alpha}}{(1+1/z)^{1-\alpha}-1}\{1+O(1/u)\} \\
&= \Psi_q(x,y) \frac{z}{(1-\alpha)} u\xi(u) \left\{1 + O\left(\frac{1}{u^\delta} + \frac{1}{\log y} + \frac{1}{z}\right)\right\} \\
&= z\Psi_q(x,y) \log x \left\{1 + O\left(\frac{1}{u^\delta} + \frac{1}{\log y} + \frac{1}{z}\right)\right\}, \qquad (A.9b)
\end{aligned}$$

using (A.3).

As $w_i = w_{i+1}/(1+1/z)$ for all i, we see that the first term in (2.8) is $(1+1/z)^{-1}$ times

$$\sum_{j=1}^{N} w_j \Psi_q\Big(\frac{x}{w_j}, y\Big) = \sum_{i=0}^{N-1} w_i \Psi_q\Big(\frac{x}{w_i}, y\Big) + \left(w_N \Psi_q\Big(\frac{x}{w_N}, y\Big) - w_0 \Psi_q\Big(\frac{x}{w_0}, y\Big)\right).$$

Therefore the sum in (2.8) equals

$$\begin{aligned}
&= \left(\Delta - \frac{1}{z+1}\right) \sum_{i=0}^{N-1} w_i \Psi_q\Big(\frac{x}{w_i}, y\Big) + \frac{1}{1+1/z}\left(w_N \Psi_q\Big(\frac{x}{w_N}, y\Big) - w_0 \Psi_q\Big(\frac{x}{w_0}, y\Big)\right) \\
&= \Psi_q(x,y) \log x \left\{\left(\Delta - \frac{1}{z+1}\right) z + \frac{\xi(u)}{\log y}\right\} \left\{1 + O\left(\frac{1}{u^\delta} + \frac{1}{\log y} + \frac{1}{z}\right)\right\}.
\end{aligned}$$

by (A.9), which implies (2.8), for $u \ge u_0$. For $2 \le u \le u_0$ we prove (2.8) exactly as above, except that we use the estimate (A.1) (together with $\Psi(x,y) \sim x\rho(u)$) instead of (A.7).

Appendix Two. Almost primes in short intervals of arithmetic progressions.

Recently, Mikawa (1989) proved that if $g(x)$ is any function that $\to \infty$ as $x \to \infty$ then, for any given q, there exists a P2 in almost all reduced residue classes modulo q, which is $\le g(q) q \log^5 q$. We modify his proof to obtain:

Proposition 1. *There exists $\varepsilon > 0$ (sufficiently small) and a constant $c > 0$ such that the estimate*

$$(B.1) \qquad \sum_{\substack{x-y\le n\le x \\ n\equiv a \pmod{q}}} \Lambda_2(n) \ \ge\ c\frac{y\log x}{\phi(q)}$$

fails for

$$\ll\ \phi(q)\left\{\frac{1}{x^{\varepsilon}} + \sqrt{\frac{\phi(q)\log^5 x}{y}}\right\}$$

reduced residue classes modulo q, provided $x \ge y \ge x^{1-\varepsilon}$.

(To be consistent with (Mikawa, 1989), and what is described below, it is necessary to replace ε by $\varepsilon/4$ in the statement of Proposition 1.)

The proof is based on ‘Richert weights’ and the Hooley–Weil estimates for incomplete Kloosterman sums, as in (Mikawa, 1989, pp. 390–392). The main difference here is that our sums are all in the short range $(x-y, x]$ rather than $(x, 2x]$, and also that we estimate the sums S_4 differently so as to allow a wider range. We only give here the modifications necessary to prove Proposition 1; for a complete argument we refer the reader to (Mikawa, 1989). Henceforth we use the notation of (Mikawa, 1989):

First, let $D = MN = x^{\alpha},\ y = x^{1/u},\ z = x^{1/v}$ instead of (Mikawa, 1989, p.390). The most difficult part of the argument is (Mikawa, 1989, Lemma 4). This should be altered here to give the conclusion

$$\sum_{(a,q)=1}\left|\sum_{\substack{x-y<n\le x \\ n\equiv a \pmod{q}}}\left(\sum_{\substack{d|n \\ (d,q)=1}}\lambda_d\right) - \left(\sum_{\substack{d \\ (d,q)=1}}\frac{\lambda_d}{d}\right)\frac{y}{q}\right|^2 \ll\ y\log^3 x + \frac{y}{q}x^{1-4\varepsilon}.$$

The proof is essentially the same, though with a few modifications:

Given our change of range ($(x-y, x]$ rather than $(x, 2x]$), the terms x/q in (Mikawa, 1989) change to y/q; error terms such as $O(x\log^3 x)$ change to $O(y\log^3 x)$; and in (Mikawa, 1989, (4.2), (4.3) and beyond) the term $x_1 = x$ becomes $x_1 = x - y$, and $x_2 = 2x - ql = 2x - q\delta k$ becomes $x_2 = x - ql = x - q\delta k$. However we may use the integral in (5.7), as it stands, as an upper bound for R_2, as this entails no significant loss.

In order to allow Mikawa's proof to work for any value of x (and not just those very close to q), we must improve the estimate for S_4 given in (Mikawa, 1989, (3.2) and the line below). Mikawa informed me of the following elegant argument due to Balog:

$$\sum_{(a,q)=1} S_4(x;q,a) = \sum_{(a,q)=1} \sum_{\substack{x-y<n\le x \\ n\equiv a \pmod q}} \sum_{\substack{p^2|n \\ x^{2\varepsilon}<p<x^{1/2-\varepsilon}}} 1$$
$$\le \sum_{x^{2\varepsilon}<p<x^{1/2-\varepsilon}} \sum_{\substack{x-y<n\le x \\ p^2|n}} 1 \ll \frac{y}{x^{2\varepsilon}} + x^{1/2-\varepsilon}.$$

Using Cauchy's inequality this, together with Lemma 4, leads to the bound

$$\sum_{(a,q)=1} |E(x;q,a)| \ll \sqrt{\phi(q) y \log^3 x} + \sqrt{\phi(q)\frac{y}{q}x^{1-4\varepsilon}} + \frac{y}{x^{2\varepsilon}} + x^{1/2-\varepsilon},$$

where

$$\#\{P2:\ x-y<P2\le x,\ P2\equiv a \pmod q,\ p|P2 \Rightarrow p>z\} \ \ge\ \frac{C}{\log z}\frac{y}{\phi(q)} \ -\ |E(x;q,a)|,$$

(in place of (Mikawa, 1989, (3.5))). Proposition 1 then follows.

References

Balog, A. & Pomerance, C. 1992 The distribution of smooth numbers in arithmetic progressions. *Proc. Amer. Math. Soc.* **115**, 33–43.

Bombieri, E. 1987/1974 Le grand crible dans la théorie analytique des nombres. *Astérisque* **18**.

Fouvry, E. & Tenenbaum, G. 1991 Entiers sans grand facteur premier en progressions arithmétiques. *Proc. London Math. Soc.* **63**, 449–494.

Friedlander, J.B. 1981 Integers without large prime factors, II. *Acta Arith.* **39**, 53–57.

Granville, A. 1993 Integers, without large prime factors, in arithmetic progressions, I. *Acta Mathematica* (to appear).

Hildebrand, A. 1986 On the number of positive integers $\le x$ and free of large prime factors $> y$. *J. Number Theory.* **22**, 289–307.

Hildebrand, A. & Tenenbaum, G. 1986 On integers free of large prime factors. *Trans. Amer. Math. Soc.* **296**, 265–290.

Iwaniec, H. 1982 On the Brun–Titchmarsh theorem, *J. Math. Soc. Japan.* **34**, 95–123.

Mikawa, H. 1989 Almost–primes in arithmetic progressions and short intervals. *Tsukuba J. Math.*. **13**, 387–401.