On the second moment for primes in
an arithmetic progression

by

D. A. GOLDSTON (San Jose, CA) and C. Y. YILDIRIM (Ankara)

1. Introduction and statement of results. In this paper we calculate a lower bound, of the same order of magnitude as conjectured, for the second moment of primes in an arithmetic progression. Specifically we examine

$$(1.1) \qquad I(x, h, q, a) := \int_{x}^{2x} \left(\psi(y + h; q, a) - \psi(y; q, a) - \frac{h}{\phi(q)} \right)^2 dy$$

where

$$(1.2) \qquad \psi(x; q, a) = \sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n),$$

and Λ is the von Mangoldt function. We will take

$$(1.3) \qquad (a, q) = 1, \quad x \geq 2, \quad 1 \leq q \leq h \leq x$$

(other ranges not being interesting). We shall assume the truth of the Generalized Riemann Hypothesis (GRH), which implies, in particular,

$$(1.4) \qquad E(x; q, a) := \psi(x; q, a) - \frac{x}{\phi(q)} \ll x^{1/2} \log^2 x \quad (q \leq x).$$

The idea of our method originates from the work of Goldston [3] for the case of all primes, corresponding in the present formulation to $q = 1$. An improved and generalized version of this result appeared in [6] as

THEOREM A. *Assume GRH. Then for any $\varepsilon > 0$ and*

$$1 \leq \frac{h}{q} \ll \frac{x^{1/3}}{q^{\varepsilon} \log^3 x}$$

we have

$$(1.5) \qquad \sum_{\substack{a \pmod q \\ (a,q)=1}} I(x,h,q,a) \geq \frac{1}{2} xh \log\left(\left(\frac{q}{h}\right)^3 x \right) - O(xh(\log\log x)^3).$$

Moreover, for almost all q with $h^{3/4} \log^5 x \leq q \leq h$ we have

$$(1.6) \qquad \sum_{\substack{a \pmod q \\ (a,q)=1}} I(x,h,q,a) \sim xh \log\left(\frac{xq}{h}\right).$$

For an individual arithmetic progression Özlük [10] proved unconditionally

THEOREM B. *For $1 \leq q \leq (\log x)^{1-\delta}$ and $h \leq (\log x)^c$ (δ and c are any fixed positive numbers) satisfying $q \leq h$, we have*

$$(1.7) \qquad I(x,h,q,a) > \left(\frac{1}{2} - \varepsilon\right) \frac{xh}{\phi(q)} \log x$$

for any ε and $x \geq X(\varepsilon, c)$.

We shall see below that the GRH implies a result of the type in Theorem B for much wider ranges of q and h. An asymptotic estimate for $I(x,h,q,a)$ in certain ranges was shown by Yıldırım [11] to be implied by GRH and a pair correlation conjecture for the zeros of Dirichlet's L-functions.

THEOREM C. *Assume GRH. Let α_1, α_2, η be fixed and satisfying $0 < \eta < \alpha_1 \leq \alpha_2 \leq 1$, and let $\delta = x^{-\alpha}$ where $\alpha_1 \leq \alpha \leq \alpha_2$. Assume, as $x \to \infty$, uniformly for*

$$(1.8) \qquad q \leq \min(x^{1/2}\delta^{1/2} \log^A x, \delta^{-1} x^{-\eta}) \quad (q \text{ prime or } 1)$$

and

$$(1.9) \qquad \frac{x^{\alpha_1}}{\phi(q)} \log^{-3} x \leq T \leq \phi(q) x^{\alpha_2} \log^3 x$$

that, for $(a,q) = 1$,

$$(1.10) \qquad \sum_{\chi_1,\chi_2 \pmod q} \overline{\chi}_1(a)\chi_2(a) \sum_{\substack{0 < \gamma_1,\gamma_2 \leq T \\ L(1/2+i\gamma_1,\chi_1)=0 \\ L(1/2+i\gamma_2,\chi_2)=0}} x^{i(\gamma_1-\gamma_2)} \frac{4}{4 + (\gamma_1 - \gamma_2)^2}$$

$$\sim \phi(q) \frac{T}{2\pi} \log qT.$$

Then

$$(1.11) \qquad \int_x^{2x} \left(\psi(u+u\delta;q,a) - \psi(u;q,a) - \frac{u\delta}{\phi(q)} \right)^2 du \sim \frac{3}{2} \cdot \frac{\delta x^2}{\phi(q)} \log \frac{q}{\delta}$$

uniformly for $x^{-\alpha_2} \leq \delta \leq x^{-\alpha_1}$ and q as in (1.8).

It was also shown in [11] that the left-hand side of (1.10) is

$$\sim \phi(q)\frac{T}{2\pi}\log x \quad \text{for } 1 \leq q \leq x^{1/2}\log^{-3} x$$

when $(x/q)\log x \leq T \leq e^{x^{1/4}}$. These asymptotic values are what the diagonal terms ($\chi_1 = \chi_2$) would contribute, so the assumption (1.10) is a way of expressing that the zeros of different Dirichlet L-functions are uncorrelated. Theorem C is a generalization of one half of a result of Goldston and Montgomery [5] for the case $q = 1$, where an equivalence between the pair correlation conjecture for $\zeta(s)$ and the second moment for primes was established. Since the argument in [5] works reversibly, a suitable converse to Theorem C is also provable. The restriction to prime q was made in order to avoid the presence of imprimitive characters. The formula (1.11) involving differences $u\delta$ which vary with u can be converted to a formula involving a fixed-difference h.

Our main result is the following theorem.

THEOREM 1. *Assume GRH. Then for any $\varepsilon > 0$ and*

$$(1.12) \qquad q \leq h \leq (xq)^{1/3-\varepsilon}$$

we have as $x \to \infty$,

$$(1.13) \qquad I(x,h,q,a) \geq \frac{1}{2} \cdot \frac{xh}{\phi(q)}\log\left(\frac{xq}{h^3}\right) - O\left(\frac{xh}{\phi(q)}(\log\log x)^3\right).$$

Notice that the conditions in (1.12) imply that both h and q are $\ll x^{1/2-\varepsilon}$.

The proof of the theorem uses some new results on the function $\lambda_R(n)$ used as an approximation for the von Mangoldt function in our earlier work. Propositions 2–4 embody these results, and we expect they will have further applications to other problems.

2. Preliminaries. We shall need the following in our calculations. Let

$$(2.1) \qquad f(n,x,h) = \int\limits_{[x,2x]\cap[n-h,n)} 1\, dy$$

$$= \begin{cases} n - x & \text{for } x \leq n < x + h, \\ h & \text{for } x + h \leq n \leq 2x, \\ 2x - n + h & \text{for } 2x < n \leq 2x + h, \\ 0 & \text{elsewhere.} \end{cases}$$

LEMMA 1. *For real numbers a_n and b_n we have*

$$\int\limits_{x}^{2x} \Big(\sum_{y<n\leq y+h} a_n \Big)\Big(\sum_{y<m\leq y+h} b_m \Big) dy$$

$$= \sum_{x<n\leq 2x+h} a_n b_n f(n,x,h)$$

$$+ \sum_{0<k\leq h} \Big(\sum_{x<n\leq 2x+h-k} (a_n b_{n+k} + a_{n+k}b_n)f(n,x,h-k) \Big).$$

LEMMA 2. *Let $C(x) = \sum_{n\leq x} c_n$. Then*

$$\sum_{x<n\leq 2x+h} c_n f(n,x,h) = \int\limits_{2x}^{2x+h} C(u)\,du - \int\limits_{x}^{x+h} C(u)\,du.$$

Lemmas 1 and 2 were proved in [6]. We take this opportunity to correct a minor error in Lemma 1 of [6]. In that lemma an extraneous term $h(c_{x+h} - c_{2x})$ was incorrectly included and should be removed. This term then contributed an unnecessary error term in equations (2.7), (2.14), and (2.15) of [6]. However these same error terms correctly occurred for a different reason in equation (2.9) so that starting with equation (2.16) these error terms were correctly included in the rest of [6].

Calling $\Delta\psi = \psi(y+h;q,a) - \psi(y;q,a)$ for brevity, we have, from (1.1),

$$I(x,h,q,a) = \int\limits_{x}^{2x} (\Delta\psi)^2\,dy - \frac{2h}{\phi(q)} \int\limits_{x}^{2x} (\Delta\psi)\,dy + \frac{h^2 x}{\phi^2(q)}.$$

By the above lemmas and (1.4) we obtain

$$\int\limits_{x}^{2x} (\Delta\psi)\,dy = \sum_{\substack{x<n\leq 2x+h \\ n\equiv a\,(\mathrm{mod}\,q)}} \Lambda(n)f(n,x,h)$$

$$= \frac{xh}{\phi(q)} + \int\limits_{2x}^{2x+h} E(y;q,a)\,dy - \int\limits_{x}^{x+h} E(y;q,a)\,dy$$

$$= \frac{xh}{\phi(q)} + O(hx^{1/2}\log^2 x),$$

so that

$$(2.2) \qquad I(x,h,q,a) = \int\limits_{x}^{2x} (\Delta\psi)^2\,dy - \frac{xh^2}{\phi^2(q)} + O\Big(\frac{x^{1/2}h^2\log^2 x}{\phi(q)}\Big).$$

The integral $\int(\Delta\psi)^2$ leads to sums of the sort $\sum \Lambda(n)\Lambda(n+k)$ which are in the territory of the twin prime conjecture. In the uninteresting case $1 \leq$

$h \leq q$, only the sum $\sum \Lambda^2(n)$ is present, giving easily the evaluation

$$(2.3) \quad I(x, h, q, a) = \frac{xh}{\phi(q)} \log x - \frac{xh}{\phi(q)} - \frac{xh^2}{\phi^2(q)} + O(x^{1/2} h \log^3 x)$$

$$(h \leq q \leq x).$$

Now let $\lambda_R(n)$ be any arithmetical function, and set

$$(2.4) \qquad \psi_R(y; q, a) = \sum_{\substack{n \leq y \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n);$$

$$\Delta \psi_R = \psi_R(y + h; q, a) - \psi_R(y; q, a).$$

Trivially $\int (\Delta \psi - \Delta \psi_R)^2 \geq 0$, so that

$$(2.5) \qquad \int_x^{2x} (\Delta \psi)^2 \, dy \geq 2 \int_x^{2x} (\Delta \psi) \cdot (\Delta \psi_R) \, dy - \int_x^{2x} (\Delta \psi_R)^2 \, dy.$$

By Lemma 1 we have

$$(2.6) \qquad \int_x^{2x} (\Delta \psi) \cdot (\Delta \psi_R) \, dy$$

$$= \sum_{\substack{x < n \leq 2x+h \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n) \lambda_R(n) f(n, x, h)$$

$$+ \sum_{\substack{0 < k \leq h \\ q \mid k}} \sum_{\substack{x < n \leq 2x+h-k \\ n \equiv a \,(\mathrm{mod}\, q)}} [\lambda_R(n) \Lambda(n+k) + \Lambda(n) \lambda_R(n+k)] f(n, x, h-k).$$

By Lemma 2 we see that

$$\sum_{\substack{x < n \leq 2x+h \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n) \lambda_R(n) f(n, x, h) = \left(\int_{2x}^{2x+h} - \int_x^{x+h} \right) \sum_{\substack{n \leq u \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n) \lambda_R(n) \, du.$$

In the second term on the right-hand side of (2.6) we write $k = jq$, and by Lemma 2 we have for the first part of this term

$$\sum_{\substack{0 < j \leq h/q \\ n \equiv a \,(\mathrm{mod}\, q)}} \sum_{\substack{x < n \leq 2x+h-jq}} \lambda_R(n) \Lambda(n+jq) f(n, x, h-jq)$$

$$= \sum_{0 < j \leq h/q} \left(\int_{2x}^{2x+h-jq} - \int_x^{x+h-jq} \right) \sum_{\substack{n \leq u \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n) \Lambda(n+jq) \, du$$

$$= \int_{2x}^{2x+h} \sum_{0<j\le(u-2x)/q} \sum_{\substack{n\le u-jq \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda_R(n)\Lambda(n+jq)\,du$$

$$- \int_{x}^{x+h} \sum_{0<j\le(u-x)/q} \sum_{\substack{n\le u-jq \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda_R(n)\Lambda(n+jq)\,du,$$

wherein to see the last equality observe that

$$\sum_{0<j\le h/q} \int_{x}^{x+h-jq} \sum_{\substack{n\le u \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda_R(n)\Lambda(n+jq)\,du$$

$$= \sum_{0<j\le h/q} \int_{x+jq}^{x+h} \sum_{\substack{n\le u-jq \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda_R(n)\Lambda(n+jq)\,du$$

$$= \left(\int_{x+q}^{x+2q} + \ldots + \int_{x+\lfloor h/q\rfloor q}^{x+h} \right) \sum_{0<j\le(u-x)/q} \sum_{\substack{n\le u-jq \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda_R(n)\Lambda(n+jq)\,du$$

$$= \int_{x}^{x+h} \sum_{0<j\le(u-x)/q} \sum_{\substack{n\le u-jq \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda_R(n)\Lambda(n+jq)\,du.$$

The other part is

$$\sum_{0<j\le h/q} \sum_{\substack{x<n\le 2x+h-k \\ n\equiv a\,(\mathrm{mod}\,q)}} \Lambda(n)\lambda_R(n+jq)f(n,x,h-jq)$$

$$= \int_{2x}^{2x+h} \sum_{0<j\le(u-2x)/q} \sum_{\substack{n\le u-jq \\ n\equiv a\,(\mathrm{mod}\,q)}} \Lambda(n)\lambda_R(n+jq)\,du$$

$$- \int_{x}^{x+h} \sum_{0<j\le(u-x)/q} \sum_{\substack{n\le u-jq \\ n\equiv a\,(\mathrm{mod}\,q)}} \Lambda(n)\lambda_R(n+jq)\,du$$

$$= \int_{2x}^{2x+h} \sum_{-(u-2x)/q\le j<0} \sum_{\substack{-jq<m\le u \\ m\equiv a\,(\mathrm{mod}\,q)}} \Lambda(m+jq)\lambda_R(m)\,du$$

$$- \int_{x}^{x+h} \sum_{-(u-x)/q\le j<0} \sum_{\substack{-jq<m\le u \\ m\equiv a\,(\mathrm{mod}\,q)}} \Lambda(m+jq)\lambda_R(m)\,du,$$

where we have substituted $m = n + jq$ and then used $-j$ in place of j to get to the last member.

Hence (2.6) has now been cast into

$$(2.7) \quad \int_x^{2x} (\Delta\psi) \cdot (\Delta\psi_R) \, dy$$

$$= \Big(\int_{2x}^{2x+h} - \int_x^{x+h} \Big) \sum_{\substack{n \leq u \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n)\lambda_R(n) \, du$$

$$+ \int_{2x}^{2x+h} \sum_{0 < |j| \leq (u-2x)/q} \sum_{\substack{N_1 < n \leq N_2 \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n)\Lambda(n+jq) \, du$$

$$- \int_x^{x+h} \sum_{0 < |j| \leq (u-x)/q} \sum_{\substack{N_1 < n \leq N_2 \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n)\Lambda(n+jq) \, du,$$

with $N_1 = \max(0, -jq)$ and $N_2 = \min(u, u - jq)$. Similarly

$$(2.8) \quad \int_x^{2x} (\Delta\psi_R)^2 \, dy = \Big(\int_{2x}^{2x+h} - \int_x^{x+h} \Big) \sum_{\substack{n \leq u \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R^2(n) \, du$$

$$+ 2 \int_{2x}^{2x+h} \sum_{0 < j \leq (u-2x)/q} \sum_{\substack{n \leq u-jq \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n)\lambda_R(n+jq) \, du$$

$$- 2 \int_x^{x+h} \sum_{0 < j \leq (u-x)/q} \sum_{\substack{n \leq u-jq \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n)\lambda_R(n+jq) \, du.$$

3. The choice of $\lambda_R(n)$ and some number-theoretic sums. As the auxiliary function we use

$$(3.1) \quad \lambda_R(n) := \sum_{r \leq R} \frac{\mu^2(r)}{\phi(r)} \sum_{d|(r,n)} d\mu(d).$$

This function is known ([3], [7]) to exhibit behavior similar to $\Lambda(n)$ when considered on average in arithmetic progressions, and it has been employed in related problems ([1], [2], [4], [6]). An upper bound for $\lambda_R(n)$ is

$$(3.2) \qquad |\lambda_R(n)| \leq \sum_{d|n} d \sum_{\substack{r \leq R \\ d|r}} \frac{1}{\phi(r)} \leq \max_{r \leq R} \left(\frac{r}{\phi(r)} \right) \sum_{d|n} d \sum_{\substack{r \leq R \\ d|r}} \frac{1}{r}$$

$$\ll d(n) \log R \log \log R.$$

To evaluate the sums which arise when (3.1) is used in (2.7) and (2.8) we shall need some lemmas. In the following p will denote a prime number.

LEMMA 3 (Hildebrand [8]). *We have for each positive integer k, uniformly for $R \geq 1$,*

$$(3.3) \qquad L_k(R) := \sum_{\substack{n \leq R \\ (n,k)=1}} \frac{\mu^2(n)}{\phi(n)} = \frac{\phi(k)}{k}(\log R + c + v(k)) + O\left(\frac{w(k)}{\sqrt{R}} \right),$$

where

$$c := \gamma + \sum_p \frac{\log p}{p(p-1)}; \quad v(k) := \sum_{p|k} \frac{\log p}{p};$$

$$(3.4)$$

$$w(k) := \sum_{d|k} \frac{\mu^2(d)}{\sqrt{d}} = \prod_{p|k} \left(1 + \frac{1}{\sqrt{p}} \right); \quad v(1) = 0, \ w(1) = 1.$$

LEMMA 4. *We have*

$$(3.5) \qquad\qquad\qquad v(k) \ll \log \log 3k,$$

$$(3.6) \qquad\qquad\qquad \sum_{p|k} \frac{1}{\sqrt{p}} \ll \frac{\sqrt{\log k}}{\log \log 3k},$$

and

$$(3.7) \qquad\qquad g(k) := \prod_{p|k} \left(1 + \frac{p}{p-1} \right) \ll 2^{\nu(k)}(\log \log 3k).$$

Proof. We show (3.7); the other inequalities can be proved similarly. Let $\nu(k)$ be the number of distinct prime factors of k, which satisfies the bound $\nu(k) \ll (\log k)/\log \log k$. We have

$$\log g(k) = \sum_{p|k} \log \left(2 + \frac{1}{p-1} \right)$$

$$< \nu(k) \log 2 + \sum_{p|k} \frac{1}{p} < \nu(k) \log 2 + \sum_{p \leq 2 \log 2k} \frac{1}{p}$$

$$= \nu(k) \log 2 + \log \log \log 21k + O(1),$$

where the prime number theorem and Mertens' theorem have been employed. Exponentiating both sides we obtain (3.7).

LEMMA 5. *We have*

$$(3.8) \qquad \sum_{l|k} \frac{\mu^2(l)}{\phi(l)} \log l = \frac{k\,v(k)}{\phi(k)}.$$

Proof. Observe that

$$(3.9) \qquad \sum_{l|k} \frac{\mu^2(l)}{\phi(l)} \log l = \sum_{p|k} (\log p) \sum_{l|k,\,p|l} \frac{\mu^2(l)}{\phi(l)}.$$

For the inner sum we see on writing $l = pm$ that

$$\sum_{l|k,\,p|l} \frac{\mu^2(l)}{\phi(l)} = \frac{1}{\phi(p)} \sum_{\substack{m|k/p \\ (m,p)=1}} \frac{\mu^2(m)}{\phi(m)}$$

$$= \frac{1}{\phi(p)} \cdot \frac{1}{1+1/\phi(p)} \prod_{p'|k} \left(1 + \frac{1}{\phi(p')}\right) = \frac{k}{p\phi(k)},$$

and using this in (3.9) gives the result. We note for later use that in the previous calculation we can drop the condition $p\,|\,l$ and obtain

$$(3.10) \qquad \sum_{m|k} \frac{\mu^2(m)}{\phi(m)} = \frac{k}{\phi(k)}.$$

LEMMA 6 (Goldston [3]). *We have*

$$(3.11) \qquad \sum_{r \le R} \frac{\mu^2(r)}{\phi(r)} \sum_{\substack{d|r \\ (d,k)=1}} \frac{d\mu(d)}{\phi(d)} = \mathfrak{S}(k) + O\left(\frac{kd(k)}{R\phi(k)}\right),$$

where

$$(3.12) \qquad \mathfrak{S}(k) = \begin{cases} 2C \prod_{\substack{p|k \\ p>2}} \left(\dfrac{p-1}{p-2}\right) & \text{if } k \text{ is even, } k \ne 0, \\[2ex] 0 & \text{if } k \text{ is odd,} \end{cases}$$

with

$$(3.13) \qquad C = \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right).$$

Proof. The proof can be found in [3]; we just note that

$$(3.14) \qquad \sum_{\substack{d|r \\ (d,k)=1}} \frac{d\mu(d)}{\phi(d)} = \frac{\mu(r)}{\phi(r)}\mu((r,k))\phi((r,k)),$$

so the left-hand side of (3.11) may be expressed as

$$(3.15) \qquad \sum_{r=1}^{\infty} \frac{\mu(r)\mu((r,k))\phi((r,k))}{\phi^2(r)} + O\left(\sum_{r>R} \frac{\mu^2(r)\mu^2((r,k))\phi((r,k))}{\phi^2(r)}\right).$$

Here the first sum is $\mathfrak{S}(k)$ and the error term is $\ll$ the O-term in (3.11).

Lemma 7 (Friedlander and Goldston [1]). *We have*

$$(3.16) \qquad \sum_{0<j\leq h/q} (h-jq)\mathfrak{S}(jq) = \frac{h^2}{2\phi(q)} - \frac{h}{2}\log\frac{h}{q} + O(h(\log\log 3q)^3).$$

Lemma 8 (Hooley [9]). *Assuming GRH, we have*

$$(3.17) \qquad \sum_{\substack{a\,(\mathrm{mod}\,q)\\(a,q)=1}} \max_{u\leq x} |E(u;q,a)|^2 \ll x\log^4 x \quad \text{for } q \leq x.$$

Lemma 9. *We have*

$$(3.18) \qquad \sum_{r\leq R} \frac{\mu^2(r)\sigma(r)}{\phi(r)} \ll R,$$

$$(3.19) \qquad \sum_{r\leq R} \frac{\mu^2(r)\sigma_{1/2}(r)}{\phi(r)} \ll \sqrt{R},$$

$$(3.20) \qquad \sum_{0<r\leq R} \frac{rd(r)}{\phi(r)} \ll R\log 2R.$$

Proof. To prove (3.18), note

$$\sum_{r\leq R} \frac{\mu^2(r)\sigma(r)}{\phi(r)} = \sum_{r\leq R} \mu^2(r) \prod_{p|r}\left(1+\frac{2}{p-1}\right) = \sum_{r\leq R} \mu^2(r) \sum_{d|r} \frac{2^{\nu(d)}}{\phi(d)}$$

$$= \sum_{d\leq R} \frac{\mu^2(d)2^{\nu(d)}}{\phi(d)} \sum_{l\leq R/d} \mu^2(l) \leq R \sum_{d\leq R} \frac{\mu^2(d)2^{\nu(d)}}{d\phi(d)}$$

$$\leq R \prod_{p}\left(1+\frac{2}{p(p-1)}\right) \ll R.$$

The proof of (3.19) is similar, and (3.20) was considered in [3].

4. The proof of the Theorem. In this section we calculate the right-hand sides of (2.7) and (2.8), and so obtain our result.

Proposition 1. *Assuming GRH, we have*

$$(4.1) \qquad \sum_{\substack{n\leq N\\n\equiv a\,(\mathrm{mod}\,q)}} \Lambda(n)\lambda_R(n) = \frac{N\log R}{\phi(q)} + \frac{cN}{\phi(q)} + O\left(\frac{N}{\phi(q)\sqrt{R}}\right)$$

$$+ O(N^{1/2}\log^3 N) + O(R\log N).$$

Proof. Starting from the definition (3.1) and recalling $L_k(R)$ from (3.3), we have

$$\sum_{\substack{n \leq N \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n)\lambda_R(n)$$

$$= \sum_{r \leq R} \frac{\mu^2(r)}{\phi(r)} \sum_{d|r} d\mu(d) \sum_{\substack{n \leq N \\ n \equiv a \,(\mathrm{mod}\, q) \\ d|n}} \Lambda(n)$$

$$= L_1(R)\psi(N; q, a) - \sum_{r \leq R} \frac{\mu^2(r)}{\phi(r)} \sum_{p|r} p \log p \sum_{\substack{k \geq 1 \\ p^k \leq N \\ p^k \equiv a \,(\mathrm{mod}\, q)}} 1.$$

Here the sum over k is trivially of size $O(\log N/\log p)$, so that by Lemma 3

$$\sum_{r \leq R} \frac{\mu^2(r)}{\phi(r)} \sum_{p|r} p \log p \sum_{\substack{k \geq 1 \\ p^k \leq N \\ p^k \equiv a \,(\mathrm{mod}\, q)}} 1 \ll (\log N) \sum_{r \leq R} \frac{\mu^2(r)}{\phi(r)} \sum_{p|r} p$$

$$= (\log N) \sum_{p \leq R} \frac{p}{\phi(p)} \sum_{\substack{m \leq R/p \\ (m,p)=1}} \frac{\mu^2(m)}{\phi(m)}$$

$$\ll (\log N) \sum_{p \leq R} \left(1 + \log \frac{R}{p}\right) \ll R \log N.$$

By the prime number theorem we obtain (4.1). In order for the main term to dominate the error terms in (4.1) we will require that

$$(4.2) \qquad qR \leq \frac{N}{\log N}, \qquad q \leq \frac{N^{1/2}}{\log^3 N}.$$

Hence the relevant contribution to (2.7) will be

$$(4.3) \qquad \left(\int_{2x}^{2x+h} - \int_{x}^{x+h}\right) \sum_{\substack{n \leq u \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n)\lambda_R(n) \, du$$

$$= \frac{xh}{\phi(q)}(\log R + c) + O\left(\frac{xh}{\phi(q)\sqrt{R}}\right) + O(x^{1/2}h\log^3 x) + O(Rh\log x).$$

PROPOSITION 2. *Assuming GRH, we have for*

$$(4.4) \qquad 1 \leq q \leq h \leq x, \qquad qR \leq x,$$

that

$$(4.5) \quad \sum_{\substack{0<k\leq h \\ q|k}} \sum_{\substack{x<n\leq 2x+h-k \\ n\equiv a \,(\mathrm{mod}\, q)}} [\Lambda(n)\lambda_R(n+k) + \Lambda(n+k)\lambda_R(n)]f(n,x,h-k)$$

$$= \frac{xh^2}{\phi^2(q)} - \frac{xh}{\phi(q)}\log\frac{h}{q} + O\left(\frac{xh}{\phi(q)}(\log\log 3q)^3\right) + O\left(\frac{xh^2 d(q)}{\phi^2(q)R}\log\frac{2h}{q}\right)$$

$$+ O\left(\frac{x^{1/2}h^{3/2}R\log^2 x}{q^{1/2}}\right) + O\left(\frac{x^{1/2}h^2 R^{1/2}\log^2 x}{q}\right).$$

Proof. We have

$$\sum_{\substack{N_1<n\leq N_2 \\ n\equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n)\Lambda(n+jq) = \sum_{r\leq R}\frac{\mu^2(r)}{\phi(r)}\sum_{d|r}d\mu(d)\sum_{\substack{N_1<n\leq N_2 \\ n\equiv a \,(\mathrm{mod}\, q) \\ d|n}}\Lambda(n+jq).$$

We may write the innermost sum as

$$\sum_{\substack{N_1+jq<m\leq N_2+jq \\ m\equiv a \,(\mathrm{mod}\, q) \\ m\equiv jq \,(\mathrm{mod}\, d)}}\Lambda(m).$$

Here $m - jq = ld$ for some integer l, and so $a \equiv ld \,(\mathrm{mod}\, q)$. Since $(a,q)=1$, we can include only those d's such that $(d,q)=1$. Then there is a unique b, $0<b<qd$, such that $m\equiv b \,(\mathrm{mod}\, qd)$. We know $(m,q)=1$, so that $(m,d)=1$ if and only if $(j,d)=1$. Hence the innermost sum is equal to

$$\psi(N_2+jq;qd,b) - \psi(N_1+jq;qd,b)$$

$$= \frac{N_2-N_1}{\phi(qd)}E_{qd,b} + E(N_2+jq;qd,b) - E(N_1+jq;qd,b),$$

where $E_{qd,b} = 1$ if $(qd,b)=1$, and $E_{qd,b}=0$ if $(qd,b)>1$. Thus

$$(4.6) \quad \sum_{\substack{N_1<n\leq N_2 \\ n\equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n)\Lambda(n+jq) = \frac{u-|j|q}{\phi(q)}\sum_{r\leq R}\frac{\mu^2(r)}{\phi(r)}\sum_{\substack{d|r \\ (d,jq)=1}}\frac{d\mu(d)}{\phi(d)}$$

$$+ \sum_{r\leq R}\frac{\mu^2(r)}{\phi(r)}\sum_{\substack{d|r \\ (d,q)=1}}d\mu(d)[E(N_2+jq;qd,b) - E(N_1+jq;qd,b)],$$

where the first term on the right-hand side is the main term, its value settled

by Lemma 6, and its contribution to (2.7) will be

$$
(4.7) \quad \int\limits_{2x}^{2x+h} \sum_{0<|j|\leq(u-2x)/q} \frac{u-|j|q}{\phi(q)} \left[\mathfrak{S}(jq) + O\left(\frac{jqd(jq)}{R\phi(jq)} \right) \right] du
$$

$$
- \int\limits_{x}^{x+h} \sum_{0<|j|\leq(u-x)/q} \frac{u-|j|q}{\phi(q)} \left[\mathfrak{S}(jq) + O\left(\frac{jqd(jq)}{R\phi(jq)} \right) \right] du
$$

$$
= \frac{2x}{\phi(q)} \sum_{0<j\leq h/q} (h-jq)\mathfrak{S}(jq) + O\left(\frac{xhqd(q)}{R\phi^2(q)} \sum_{0<j\leq h/q} \frac{jd(j)}{\phi(j)} \right)
$$

$$
= \frac{xh^2}{\phi^2(q)} - \frac{xh}{\phi(q)} \log \frac{h}{q} + O\left(\frac{xh}{\phi(q)} (\log\log 3q)^3 \right) + O\left(\frac{xh^2 d(q)}{R\phi^2(q)} \log \frac{2h}{q} \right),
$$

by Lemma 7 and (3.20). For the second term on the right-hand side of (4.6), if we use (1.4) directly, we will get the upper bound $Rx^{1/2}\log^2 x$, by (3.18). This will lead to a contribution of $O\left(\frac{x^{1/2}h^2R}{q}\log^2 x\right)$ in (2.7). Instead, in view of the averaging over j in (2.7), we will use Hooley's estimate quoted as Lemma 8 above. To do this note that some of the d's may not be coprime to b, but we can discard them (from the j- and n-summations) with an error

$$
\ll \sum_{\substack{0<|j|\leq h/q \\ (j,d)>1}} \psi(3x;qd,b) \ll \sum_{0<|j|\leq h/q} \sum_{p|d} \sum_{\substack{n\leq 3x \\ p|n}} \Lambda(n) \ll \frac{h}{q}\log^2 x,
$$

and this leads to an error of $O\left(\frac{h^2R}{q}\log^2 x\right)$ in (2.7). Hence the contribution to (2.7) from the second term on the right-hand side of (4.6) is

$$
(4.8) \quad \ll \frac{h^2R}{q}\log^2 x + \left(\int\limits_{2x}^{2x+h} - \int\limits_{x}^{x+h} \right) \sum_{r\leq R} \frac{\mu^2(r)}{\phi(r)}
$$

$$
\times \sum_{\substack{d|r \\ (d,q)=1}} d \sum_{\substack{0<|j|\leq h/q \\ (j,d)=1}} \max_{u\leq 2x+h} |E(u;qd,b)|\, du
$$

$$
\ll \frac{h^2R}{q}\log^2 x + h \sum_{r\leq R} \frac{\mu^2(r)}{\phi(r)} \sum_{\substack{d|r \\ (d,q)=1}} d \left(\frac{h}{q} \right)^{1/2}
$$

$$
\times \left(\sum_{\substack{0<|j|\leq h/q \\ (j,d)=1}} \max_{u\leq 2x+h} |E(u;qd,b)|^2 \right)^{1/2}
$$

$$\ll \frac{h^2 R}{q} \log^2 x + \frac{h^{3/2}}{q^{1/2}} \sum_{r \le R} \frac{\mu^2(r)}{\phi(r)} \sum_{d \mid r} d \left(1 + \left(\frac{h}{qd}\right)\right)^{1/2}$$

$$\times \left(\sum_{\substack{j \,(\mathrm{mod}\, d) \\ (j,d)=1}} \max_{u \le 3x} |E(u; qd, b)|^2 \right)^{1/2}.$$

In the last sum as j runs through the reduced residues modulo d, b runs through those elements of the set $\{a, a+q, \ldots, a+(d-1)q\}$ which are relatively prime to d (note that $a \equiv ld \,(\mathrm{mod}\, q)$ and $(a,q) = 1$ implies $(a,d) = 1$), and this correspondence is one-to-one. This is because $m \equiv a \,(\mathrm{mod}\, q)$ and $m \equiv jq \,(\mathrm{mod}\, d)$ if and only if $m \equiv n_1 da + n_2 jq^2 \,(\mathrm{mod}\, qd)$ where n_i satisfy $dn_1 \equiv 1 \,(\mathrm{mod}\, q)$, $qn_2 \equiv 1 \,(\mathrm{mod}\, d)$, and we have $n_1 da + n_2 jq^2 \equiv a + tq \,(\mathrm{mod}\, qd)$ if and only if $j - t \equiv an_2 \,(\mathrm{mod}\, d)$. Hence we may replace the j-sum in (4.8) by

$$(4.9) \qquad \sum_{\substack{t \,(\mathrm{mod}\, d) \\ (a+tq,d)=1}} \max_{1 \le u \le 3x} |E(u; qd, a+tq)|^2.$$

To put an upper bound on (4.9) we resort to Hooley's estimate. In view of (1.4) one hardly hopes to do better even though the sum in (4.9) is over only $1/\phi(q)$ of the reduced residue classes modulo qd. A run through Hooley's proof [9] suggests that to examine (4.9) itself, one has to calculate certain integrals involving some pairs of Dirichlet L-functions and this does not seem to be feasible without unproved assumptions. Recall that Theorem C, which gives an asymptotic estimate for our integral, already rests upon such an assumption, viz. (1.10), about Dirichlet L-functions. So we proceed by taking $x \log^4 x$ as upper bound for (4.9) on the condition that $qR \le x$, and on applying Lemma 9 we conclude that the expression in (4.8) is

$$(4.10) \qquad \ll \frac{x^{1/2} h^{3/2} R}{q^{1/2}} \log^2 x + \frac{x^{1/2} h^2 R^{1/2}}{q} \log^2 x.$$

This completes the proof of Proposition 2.

PROPOSITION 3. *For $(a,q) = 1$ we have*

$$(4.11) \qquad \sum_{\substack{n \le N \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R^2(n)$$

$$= \frac{N}{\phi(q)} \left[\log R + c + O(v(q)) + O\left(\frac{\prod_{p \mid q} \left(1 + \frac{1}{\sqrt{p}}\right)\left(1 + \frac{\sqrt{p}}{p-1}\right)}{\sqrt{R}} \right) \right] + O(R^2)$$

$$= \frac{N}{\phi(q)} \left[\log R + c + O(\log \log 3q) + O\left(\frac{\exp\left[c' \frac{\sqrt{\log q}}{\log \log 3q}\right]}{\sqrt{R}} \right) \right] + O(R^2).$$

PROPOSITION 4. *For $(a, q) = 1$ and $j \neq 0$ we have*

(4.12)
$$\sum_{\substack{n \leq N \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n)\lambda_R(n + jq)$$

$$= \frac{N}{\phi(q)}\mathfrak{S}(jq) + O\left(\frac{Ng(q)}{\phi(q)R} \cdot \frac{jd(j)}{\phi(j)}\right) + O(R^2).$$

Proof. The beginning of the proof of Proposition 3 may be incorporated into that of Proposition 4 upon a notational stipulation for the case $j = 0$. When the positive integer t satisfies $t \mid j$, if $j = 0$ we will understand that t can be any positive integer; and we will take $(t, 0) = t$.

By definition (3.1),

(4.13)
$$\sum_{\substack{n \leq N \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n)\lambda_R(n + jq)$$

$$= \sum_{r,\, r' \leq R} \frac{\mu^2(r)\mu^2(r')}{\phi(r)\phi(r')} \sum_{\substack{d \mid r \\ e \mid r'}} d\mu(d)e\mu(e) \sum_{\substack{n \leq N \\ n \equiv a \,(\mathrm{mod}\, q) \\ d \mid n,\, e \mid n + jq}} 1.$$

In the innermost sum, the conditions on n, d, and e imply $(q, de) = 1$, $(d, e) \mid j$, and thus n belongs to a unique residue class modulo $[q, d, e]$. Hence we have

(4.14)
$$\sum_{\substack{n \leq N \\ n \equiv a \,(\mathrm{mod}\, q) \\ d \mid n,\, e \mid n + jq}} 1 = \frac{N}{[q, d, e]} + O(1).$$

The contribution of the $O(1)$-term in (4.14) to (4.13) is

(4.15)
$$\ll \sum_{r,\, r' \leq R} \frac{\mu^2(r)\mu^2(r')}{\phi(r)\phi(r')} \sum_{\substack{d \mid r \\ e \mid r'}} de = \left(\sum_{r \leq R} \frac{\mu^2(r)\sigma(r)}{\phi(r)}\right)^2 \ll R^2$$

by (3.18), and this is where the $O(R^2)$-terms in (4.11) and (4.12) come from. Hence

(4.16)
$$\sum_{\substack{n \leq N \\ n \equiv a \,(\mathrm{mod}\, q)}} \lambda_R(n)\lambda_R(n + jq)$$

$$= \frac{N}{q} \sum_{r,\, r' \leq R} \frac{\mu^2(r)\mu^2(r')}{\phi(r)\phi(r')} \sum_{\substack{d \mid \frac{r}{(r,q)} \\ e \mid \frac{r'}{(r',q)} \\ (d,e) \mid j}} \mu(d)\mu(e)(d, e) + O(R^2).$$

Let $(d, e) = \delta$, $d = d'\delta$, $e = e'\delta$, so that $(d', e') = 1$. The inner sums over d and e become

$$
(4.17) \qquad \sum_{\substack{\delta | j \\ \delta | (\frac{r}{(r,q)}, \frac{r'}{(r',q)})}} \delta \sum_{d' | \frac{r}{\delta(r,q)}} \mu(d') \sum_{\substack{e' | \frac{r'}{\delta(r',q)} \\ (e', d') = 1}} \mu(e').
$$

Here the innermost sum is

$$
(4.18) \qquad \sum_{\substack{e' | \frac{r'}{\delta(r',q)} \\ (e', d') = 1}} \mu(e') = \prod_{\substack{p | \frac{r'}{\delta(r',q)} \\ p \nmid d'}} (1 + \mu(p)) = \begin{cases} 1 & \text{if } \dfrac{r'}{\delta(r', q)} \,\Big|\, d', \\ 0 & \text{otherwise.} \end{cases}
$$

Next the sum over d' becomes

$$
(4.19) \qquad \sum_{\substack{d' | \frac{r}{\delta(r,q)} \\ \frac{r'}{\delta(r',q)} | d'}} \mu(d') = \begin{cases} \mu\left(\dfrac{r}{\delta(r, q)}\right) & \text{if } \dfrac{r'}{(r', q)} = \dfrac{r}{(r, q)}, \\ 0 & \text{otherwise,} \end{cases}
$$

so the main term of (4.16) is

$$
(4.20) \qquad \frac{N}{q} \sum_{\substack{r, r' \leq R \\ \frac{r}{(r,q)} = \frac{r'}{(r',q)}}} \frac{\mu^2(r)\mu^2(r')}{\phi(r)\phi(r')} \mu\left(\frac{r}{(r, q)}\right) \sum_{\delta | (\frac{r}{(r,q)}, j)} \delta\mu(\delta).
$$

Since

$$
(4.21) \qquad \sum_{\delta | (\frac{r}{(r,q)}, j)} \delta\mu(\delta) = \mu\left(\left(\frac{r}{(r, q)}, j\right)\right) \phi\left(\left(\frac{r}{(r, q)}, j\right)\right),
$$

the main term is

$$
(4.22) \quad \frac{N}{q} \sum_{\substack{r, r' \leq R \\ \frac{r}{(r,q)} = \frac{r'}{(r',q)}}} \frac{\mu^2(r)\mu^2(r')}{\phi(r)\phi(r')} \mu\left(\frac{r}{(r, q)}\right) \mu\left(\left(\frac{r}{(r, q)}, j\right)\right) \phi\left(\left(\frac{r}{(r, q)}, j\right)\right).
$$

Writing $(r, q) = l$, $(r', q) = m$, $r = ls$, $r' = ms$ where $(s, q) = 1$, (4.22) takes the form

$$
(4.23) \qquad \frac{N}{q} \sum_{l | q} \frac{\mu^2(l)}{\phi(l)} \sum_{m | q} \frac{\mu^2(m)}{\phi(m)} \sum_{\substack{s \leq \min(R/l, R/m) \\ (s, q) = 1}} \frac{\mu(s)}{\phi^2(s)} \mu((s, j)) \phi((s, j)).
$$

The $j = 0$ case. We rewrite (4.23) as

$$
(4.24) \qquad \frac{N}{q} \sum_{l | q} \frac{\mu^2(l)}{\phi(l)} \sum_{m | q} \frac{\mu^2(m)}{\phi(m)} \sum_{\substack{s \leq \min(R/l, R/m) \\ (s, q) = 1}} \frac{\mu^2(s)}{\phi(s)}.
$$

It is convenient to regard (4.24) as

$$(4.25) \quad \frac{N}{q} \sum_{l|q} \frac{\mu^2(l)}{\phi(l)} \sum_{m|q} \frac{\mu^2(m)}{\phi(m)} \sum_{\substack{s \leq R/l \\ (s,q)=1}} \frac{\mu^2(s)}{\phi(s)}$$

$$- \frac{N}{q} \sum_{l|q} \frac{\mu^2(l)}{\phi(l)} \sum_{\substack{m|q \\ m>l}} \frac{\mu^2(m)}{\phi(m)} \sum_{\substack{R/m<s \leq R/l \\ (s,q)=1}} \frac{\mu^2(s)}{\phi(s)}.$$

For the first term of (4.25), by (3.10), Lemma 3 and Lemma 5 we have

$$(4.26) \quad \frac{N}{q} \sum_{l|q} \frac{\mu^2(l)}{\phi(l)} \sum_{m|q} \frac{\mu^2(m)}{\phi(m)} \sum_{\substack{s \leq R/l \\ (s,q)=1}} \frac{\mu^2(s)}{\phi(s)}$$

$$= \frac{N}{\phi(q)}(\log R + c + v(q)) - \frac{N}{q} \sum_{l|q} \frac{\mu^2(l)}{\phi(l)} \log l$$

$$+ O\left(\frac{Nw(q)}{\phi(q)\sqrt{R}} \sum_{l|q} \frac{\mu^2(l)\sqrt{l}}{\phi(l)}\right)$$

$$= \frac{N}{\phi(q)}(\log R + c) + O\left(\frac{N}{\phi(q)\sqrt{R}} \prod_{p|q} \left(1 + \frac{1}{\sqrt{p}}\right)\left(1 + \frac{\sqrt{p}}{p-1}\right)\right).$$

The logarithm of the last product is

$$(4.27) \quad \sum_{p|q} \log\left(1 + \frac{1}{\sqrt{p}}\right) + \log\left(1 + \frac{\sqrt{p}}{p-1}\right) \leq 2\sum_{p|q} \frac{1}{\sqrt{p}} + O(1)$$

$$\ll \frac{\sqrt{\log q}}{\log\log 3q},$$

by (3.6). Hence the first term of (4.25) is

$$(4.28) \quad \frac{N}{\phi(q)}(\log R + c) + O\left(\frac{N \exp\left[c' \frac{\sqrt{\log q}}{\log\log 3q}\right]}{\phi(q)\sqrt{R}}\right)$$

for any arbitrarily small and fixed $\varepsilon > 0$. If we use

$$(4.29) \quad \sum_{\substack{R/m<s \leq R/l \\ (s,q)=1}} \frac{\mu^2(s)}{\phi(s)} = \frac{\phi(q)}{q} \log\frac{m}{l} + O\left(w(q)\sqrt{\frac{m}{R}}\right),$$

which is implied by Lemma 3, the second term of (4.25) is expressed as

(4.30) $$\frac{N\phi(q)}{q^2} \sum_{l|q} \frac{\mu^2(l)}{\phi(l)} \sum_{\substack{m|q \\ m>l}} \frac{\mu^2(m)}{\phi(m)} \log m$$

$$- \frac{N\phi(q)}{q^2} \sum_{l|q} \frac{\mu^2(l)}{\phi(l)} (\log l) \sum_{\substack{m|q \\ m>l}} \frac{\mu^2(m)}{\phi(m)}$$

$$+ O\left(\frac{Nw(q)}{q\sqrt{R}} \sum_{l|q} \frac{\mu^2(l)}{\phi(l)} \sum_{\substack{m|q \\ m>l}} \frac{\mu^2(m)}{\phi(m)} \sqrt{m} \right).$$

Each term of (4.30) is majorized by deleting the restriction $m > l$. Then, by (3.8) and (3.10), the first two terms are each $\ll Nv(q)/\phi(q)$, and the error term is the same as that of (4.26). Hence we have proved Proposition 3.

The $j \neq 0$ case. Since

(4.31) $$\sum_{\substack{s=1 \\ (s,q)=1}}^{\infty} \frac{\mu(s)}{\phi^2(s)} \mu((s,j))\phi((s,j)) = \prod_{\substack{p|j \\ p\nmid q}} \left(1 + \frac{1}{p-1}\right) \prod_{p\nmid jq} \left(1 - \frac{1}{(p-1)^2}\right)$$

$$= \mathfrak{S}(jq)\frac{\phi(q)}{q},$$

by (3.12) and (3.13), (4.23) is

(4.32) $$\frac{N}{\phi(q)}\mathfrak{S}(jq)$$

$$+ O\left(\frac{N}{q} \sum_{l|q} \frac{\mu^2(l)}{\phi(l)} \sum_{m|q} \frac{\mu^2(m)}{\phi(m)} \sum_{s>\min(R/l,R/m)} \frac{\mu^2(s)\mu^2((s,j))\phi((s,j))}{\phi^2(s)} \right).$$

The sum over s was encountered before in (3.15) and majorized as in (3.11), so the O-term in (4.32) is

(4.33) $$\ll \frac{N}{qR} \cdot \frac{jd(j)}{\phi(j)} \sum_{l|q} \frac{\mu^2(l)}{\phi(l)} \sum_{m|q} \frac{\mu^2(m)}{\phi(m)} \max(l,m) \ll \frac{Ng(q)}{\phi(q)R} \cdot \frac{jd(j)}{\phi(j)}.$$

This completes the proof of Proposition 4.

The relevant contributions to (2.8) are

(4.34) $$\sum_{\substack{x<n\leq 2x+h \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda_R^2(n) f(n,x,h)$$

$$= \frac{xh}{\phi(q)}(\log R + c + O(v(q))) + O\left(\frac{\exp\left[c'\frac{\sqrt{\log q}}{\log\log 3q}\right]}{\sqrt{R}} \right) + O(hR^2),$$

and

$$(4.35) \quad \sum_{\substack{0<k\leq h \\ q|k}} \sum_{\substack{x<n\leq 2x+h-k \\ n\equiv a\,(\mathrm{mod}\,q)}} \lambda_R(n)\lambda_R(n+k)f(n,x,h-k)$$

$$= \frac{x}{\phi(q)} \sum_{0<j\leq h/q} \mathfrak{S}(jq) + O\left(\frac{xh}{R\phi(q)}g(q)\sum_{j\leq h/q}\frac{jd(j)}{\phi(j)}\right)$$

$$= \frac{1}{2}\cdot\frac{xh^2}{\phi^2(q)} - \frac{1}{2}\cdot\frac{xh}{\phi(q)}\log\frac{h}{q} + O\left(\frac{xh}{\phi(q)}(\log\log 3q)^3\right)$$

$$+ O\left(\frac{xh^2 g(q)}{Rq\phi(q)}\log\frac{2h}{q}\right) + O\left(\frac{h^2 R^2}{q}\right).$$

To finish the proof of Theorem 1 we now put together equations (2.2), (2.5), (2.7), (2.8), (4.3), (4.5), (4.34), (4.35), subject to (1.3) and (4.2) with $N = x$, to obtain

$$(4.36) \quad I(x,h,q,a)$$

$$\geq \frac{xh}{\phi(q)}\log\frac{Rq}{h} + O\left(\frac{xh}{\phi(q)}(\log\log 3q)^3\right) + O\left(\frac{x^{1/2}h^{3/2}R\log^2 x}{q^{1/2}}\right)$$

$$+ O\left(\frac{x^{1/2}h^2 R^{1/2}\log^2 x}{\phi(q)}\right) + O\left(\frac{xh^2}{R\phi(q)}\left(\frac{d(q)}{\phi(q)}+\frac{g(q)}{q}\right)\log\frac{2h}{q}\right)$$

$$+ O\left(\frac{xh}{\phi(q)\sqrt{R}}\exp\left[c'\frac{\sqrt{\log q}}{\log\log 3q}\right]\right) + O\left(\frac{h^2 R^2}{q}\right).$$

Recall that $d(q)$ and $g(q)$ are both $\ll q^\varepsilon$. Here we pick

$$(4.37) \quad R = \left(\frac{x}{hq\log^4 x}\right)^{1/2},$$

which makes all the error terms $O\left(\frac{xh}{q}(\log\log x)^3\right)$ provided that $h \leq (xq)^{1/3-\varepsilon}$.

References

[1] J. B. Friedlander and D. A. Goldston, *Variance of distribution of primes in residue classes*, Quart. J. Math. Oxford (2) 47 (1996), 313–336.

[2] —, —, *Note on a variance in the distribution of primes*, in: Number Theory in Progress (Proc. of the Schinzel conference, Zakopane, 1997), Vol. 2, de Gruyter, 1999, 841–848.

[3] D. A. Goldston, *A lower bound for the second moment of primes in short intervals*, Exposition. Math. 13 (1995), 366–376.

[4] D. A. Goldston, S. M. Gonek , A. E. Özlük and C. Snyder, *On the pair correlation of zeros of the Riemann zeta-function*, Proc. London Math. Soc. 80 (2000), 31–49.

[5]   D. A. Goldston and H. L. Montgomery, *Pair correlation of zeros and primes in short intervals*, in: Analytic Number Theory and Diophantine Problems, Birkhäuser, Boston, 1987, 183–203.

[6]   D. A. Goldston and C. Y. Yıldırım, *Primes in short segments of arithmetic progressions*, Canad. J. Math. 50 (1998), 563–580.

[7]   D. R. Heath-Brown, *The ternary Goldbach problem*, Rev. Mat. Iberoamericana 1 (1985), 45–59.

[8]   A. Hildebrand, *Über die punktweise Konvergenz von Ramanujan-Entwicklungen zahlentheoretischer Funktionen*, Acta Arith. 44 (1984), 109–140.

[9]   C. Hooley, *On the Barban–Davenport–Halberstam theorem: VI*, J. London Math. Soc. (2) 13 (1976), 57–64.

[10] A. E. Özlük, *On the irregularity of distribution of primes in an arithmetic progression over short intervals*, Bull. Tech. Univ. Istanbul 40 (1987), 255–264.

[11] C. Y. Yıldırım, *The pair correlation of zeros of Dirichlet L-functions and primes in arithmetic progressions*, Manuscripta Math. 72 (1991), 325–334.

Department of Mathematics and Computer Science
San Jose State University
San Jose, CA 95192, U.S.A.
E-mail: goldston@mathcs.sjsu.edu

Department of Mathematics
Bilkent University
Ankara 06533, Turkey
E-mail: yalcin@fen.bilkent.edu.tr