

Bombieri–Vinogradov type theorems for sparse sets of moduli

by

STEPHAN BAIER (Kingston) and LIANGYI ZHAO (Toronto)

1. Introduction and history. The classical Bombieri–Vinogradov theorem [7, 22] touches upon the distribution of primes in arithmetic progressions on average. More precisely, the theorem states the following.

Let $A > 0$ be fixed. Then

$$(1.1) \quad \sum_{q \leq x^{1/2}/(\log x)^{A+5}} \max_{y \leq x} \max_{\substack{a \\ \gcd(a,q)=1}} \left| \psi(y; q, a) - \frac{y}{\varphi(q)} \right| \ll \frac{x}{(\log x)^A}.$$

Here and in what follows, $\varphi(n)$ is the Euler function and

$$\psi(x; q, a) = \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \Lambda(n),$$

where $\Lambda(n)$ is the von Mangoldt function.

Moreover, theorems concerning the mean square error in the prime number theorem for arithmetic progressions were initiated by Barban [6] and Davenport and Halberstam [10]. Their results were sharpened by Gallagher [13]. The theorem in question states the following. *For any $A > 0$ fixed, we have*

$$(1.2) \quad \sum_{q \leq x/(\log x)^{A+1}} \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^q \left(\psi(x; q, a) - \frac{x}{\varphi(q)} \right)^2 \ll \frac{x^2}{(\log x)^A}.$$

Montgomery [18] showed that (1.2) is the best possible.

In this paper, we aim to prove theorems of the types of Bombieri–Vinogradov and Barban–Halberstam–Davenport for sparse sets of moduli.

2000 *Mathematics Subject Classification*: 11B25, 11L20, 11L40, 11N05, 11N14, 11N32, 11N35, 11N36.

Key words and phrases: large sieve, estimates for character sums, primes in arithmetic progressions, special moduli, primes in sparse sets, primes represented by polynomials.

Theorems of both types are also proved for square moduli. As in the classical theorems, one of the key ideas in the proof is the large sieve. We shall need versions of the large sieve for these sparse sets of moduli. Works in the direction of large sieve for special moduli can be found in [2–4, 23, 24].

It is also noteworthy that the implied constants in both (1.1) and (1.2) are ineffective because of the use of the Siegel–Walfisz theorem in the estimates. Such ineffectiveness is also present in Theorems 2 and 4 of this paper for the same reason.

2. Statements of results. Throughout this paper, we assume without loss of generality that $x \geq 3$. We start with the following notation and definition. Let S be a set of natural numbers. For $t \in \mathbb{N}$ we set

$$(2.1) \quad S_t = \{q \in \mathbb{N} : qt \in S\}, \quad S_t(R) = \{q \in S_t : R < q \leq 2R\}$$

and

$$S(R) = S_1(R).$$

We say that S is *well-distributed* if, for $\varepsilon > 0$, $t \in \mathbb{N}$, $R \leq x < x + y \leq 2R$ and $\gcd(k, l) = 1$, we have

$$(2.2) \quad |\{q \in S_t : x \leq q \leq x + y, q \equiv l \pmod k\}| \ll \left(\frac{|S_t(R)|y}{kR} + 1\right)(Rt)^\varepsilon,$$

where the implied constant depends at most on ε and S . It should be noted that the above is a natural definition of a set being “well-distributed” in arithmetic progressions as the majorant in (2.2) gives essentially the “expected” cardinality of the left-hand side of (2.2).

We are now ready to state the results of this paper. First we have the following Barban–Davenport–Halberstam type theorem for sparse sets of moduli.

THEOREM 1. *Let S be a well-distributed set. Assume that*

$$(2.3) \quad |S_q(Q/q)| \ll \frac{|S(Q)|}{q^\varepsilon}$$

for any fixed $\varepsilon > 0$ and all $q \in \mathbb{N}$ with the implied constant depending at most on ε and S . Suppose further that

$$(2.4) \quad |S(Q)| \gg \sqrt{Q} \quad \text{as } Q \rightarrow \infty.$$

Then, if $Q \leq x^{1-\varepsilon}$, we have, for any $A > 0$ fixed,

$$(2.5) \quad \sum_{q \in S(Q)} \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^q \left| \psi(x; q, a) - \frac{x}{\varphi(q)} \right|^2 \ll \frac{|S(Q)|}{Q} \frac{x^2}{(\log x)^A}.$$

As a corollary to Theorem 1, we have the following theorem for square moduli.

THEOREM 2. For any $\varepsilon > 0$ and fixed $A > 0$, we have

$$(2.6) \quad \sum_{q \leq x^{1/2-\varepsilon}} q \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^{q^2} \left| \psi(x; q^2, a) - \frac{x}{\varphi(q^2)} \right|^2 \ll \frac{x^2}{(\log x)^A}.$$

Next, we have a Bombieri–Vinogradov type theorem for sparse sets of moduli.

THEOREM 3. Let S be a well-distributed set. Assume that condition (2.3) holds and

$$|S(Q)| \gg Q^{3/4} \quad \text{as } Q \rightarrow \infty.$$

Then, if $Q \leq x^{1/2-\varepsilon}$, we have, for any fixed $A > 0$,

$$(2.7) \quad \sum_{q \in S(Q)} \max_{\substack{a \\ \gcd(a,q)=1}} \left| \psi(x; q, a) - \frac{x}{\varphi(q)} \right| \ll \frac{|S(Q)|}{Q} \frac{x}{(\log x)^A}.$$

We shall also prove the following Bombieri–Vinogradov type theorem for square moduli.

THEOREM 4. For any $\varepsilon > 0$ and fixed $A > 0$, we have

$$(2.8) \quad \sum_{q \leq x^{2/9-\varepsilon}} q \max_{\substack{a \\ \gcd(a,q)=1}} \left| \psi(x; q^2, a) - \frac{x}{\varphi(q^2)} \right| \ll \frac{x}{(\log x)^A}.$$

A better version of Theorem 4 can be obtained under the assumption of a conjecture, as will be noted in a later section.

It should also be noted that we achieve a saving of an arbitrary power of logarithm beyond the trivial bound in our results, as in the classical results. Moreover, the weight q that is present in both (2.6) and (2.8) is suitable for square moduli to obtain results analogous to the classical theorems.

As an application of Theorem 4, we have the following. It is an old but hitherto unresolved problem if there exist infinitely many primes of the form $m^2 + 1$, where m is an integer. G. H. Hardy and J. E. Littlewood [14] gave a conjectural asymptotic estimate for the number of primes not exceeding x of this form. One may find several results on approximations to this problem in the literature. For example, Ankeny [1] and Kubilius [17] showed independently that under the Riemann hypothesis for Hecke L -functions there exist infinitely many primes of the form $p = m^2 + n^2$ with $n < c \log p$, where c is some positive constant. For some unconditional results in this direction see [5] and [15]. Moreover, the prime number theorem of Pyatetskiĭ-Shapiro [20] can also be viewed as an approximation of the problem of representation of primes by quadratic polynomials.

Here, we approximate the problem of representation of primes by $m^2 + 1$ in the following way. For a natural number n let $s(n)$ be the unique square-free number a such that $\mathbb{Q}(\sqrt{n}) = \mathbb{Q}(\sqrt{a})$. In other words, $s(n) = n/m^2$,

where m^2 is the largest square dividing n . We note that $s(n) = 1$ if and only if n is a perfect square. We consider primes of the form $n + 1$, where $s(n)$ is small. More in particular, we shall prove the following.

THEOREM 5. *Let $\varepsilon > 0$. Then there exist infinitely many primes p such that $s(p - 1) \leq p^{5/9+\varepsilon}$.*

Equivalently, the above result may be formulated as follows.

THEOREM 5'. *Let $\varepsilon > 0$. Then there exist infinitely many primes of the form $p = am^2 + 1$ such that $a \leq p^{5/9+\varepsilon}$.*

We also note that the set of integers of the form $m^2 + 1$ is very sparse. Friedlander and Iwaniec [12] proved the celebrated result that there exist infinitely many primes of the form $m^2 + n^4$ (with an asymptotic formula). The set of integers of the form $m^2 + n^4$ contains the set of integers of the form $m^2 + 1$ but is still very sparse. The number of such integers not exceeding x is $O(x^{3/4})$. It is noteworthy that the set of natural numbers n with $s(n) \leq n^{5/9+\varepsilon}$ is also very sparse. More in particular, the number of $n \leq x$ with $s(n) \leq n^{5/9+\varepsilon}$ is $O(x^{7/9+\varepsilon/2})$ as the following calculation shows:

$$\begin{aligned} |\{n \leq x : s(n) \leq n^{5/9+\varepsilon}\}| &\leq |\{(a, m) \in \mathbb{N}^2 : a \leq x^{5/9+\varepsilon}, am^2 \leq x\}| \\ &= \sum_{a \leq x^{5/9+\varepsilon}} \sum_{m \leq \sqrt{x/a}} 1 = O(x^{7/9+\varepsilon/2}). \end{aligned}$$

3. Preliminary lemmas. In this section, we quote lemmas that we shall need in the proofs of our theorems. We begin with the following.

LEMMA 1 (Siegel–Walfisz). *Let B be any positive constant. Then there exist positive numbers $C_1(B)$ and $C_2(B)$ depending only on N , such that if q satisfies*

$$q \leq (\log x)^B$$

then

$$\psi(x; q, a) = \frac{x}{\varphi(q)} + O(x \exp(-C_1(B)\sqrt{\log x})),$$

uniformly in q , and

$$\psi(x, \chi) \ll x \exp(-C_2(B)\sqrt{\log x}), \quad \text{where } \psi(x, \chi) = \sum_{n \leq x} \Lambda(n)\chi(n)$$

for any non-principal character $\chi \pmod{q}$.

Proof. See Section 22 of [9]. ■

We shall also need the following identity.

LEMMA 2 (Vaughan identity). *Let $U \geq 1$, $V \geq 1$, $UV \leq x$. Then for every arithmetic function $f(n)$, we have*

$$\sum_{n \leq x} f(n)\Lambda(n) \ll U(\log x) \max_{n \leq x} |f(n)| + (\log x)T_1(f) + T_2(f) + T_3(f)$$

with

$$T_1(f) = \sum_{l \leq \max(U,V)} \max_w \left| \sum_{w < k \leq x/l} f(kl) \right|$$

and

$$T_i(f) = \left| \sum_{U < m \leq \max(x/V, UV)} \sum_{k \leq x/m} a_i(m)b_i(k)f(mk) \right|$$

for $i = 2$ and 3 , where $a_i(m)$, $b_i(k)$ are arithmetic functions which depend only on U and V and satisfy the inequalities $|b_i(k)| \leq \tau(k)$ and $|a_i(k)| \leq \log k$ for all $k \in \mathbb{N}$.

Proof. This is Theorem 6.1.1 in [8] and has its origin in [21]. ■

We shall also need the following lemma due to Pólya and Vinogradov.

LEMMA 3. *Let $\chi(n)$ be a non-principal character modulo q and $M \in \mathbb{Z}$ and $N \in \mathbb{N}$. Then*

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \leq 6\sqrt{q} \log q.$$

Proof. This is Theorem 12.5 in [16]. ■

We shall not need the classical large sieve inequality in its full generality as given by Davenport and Halberstam [11], but only the inequality touching upon Dirichlet characters.

LEMMA 4 (Classical large sieve). *Let $\{a_n\}$ be a sequence of complex numbers. Suppose that $M \in \mathbb{Z}$, $N, Q \in \mathbb{N}$. Then*

$$(3.1) \quad \sum_{q=1}^Q \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^* \left| \sum_{n=M+1}^{M+N} a_n \chi(n) \right|^2 \leq (Q^2 + N) \sum_{n=M+1}^{M+N} |a_n|^2,$$

where \sum^* henceforth denotes the sum over primitive characters to the specified modulus.

Proof. See [19] for the proof of this lemma. ■

We also need the following lemma.

LEMMA 5. *For $1 \leq Q_0 < Q$ we have*

$$\sum_{Q_0 < q \leq Q} \frac{1}{\varphi(q)} \sum_{\chi \bmod q}^* \max_{y \leq x} |\psi'(y, \chi)| \ll (x/Q_0 + x^{5/6} \log Q + x^{1/2} Q)(\log Qx)^4.$$

Proof. See Section 28 of [9]. ■

We also need the following for the proof of Theorem 5.

LEMMA 6. *We have*

$$(3.2) \quad \sum_{y < q \leq 2y} \frac{1}{\varphi(q^2)} = \frac{1}{2\zeta(2)y} + O\left(\frac{\log y}{y^2}\right).$$

Proof. We write

$$\begin{aligned} \sum_{y < q \leq 2y} \frac{1}{\varphi(q^2)} &= \sum_{y < q \leq 2y} \frac{1}{q\varphi(q)} = \sum_{y < q \leq 2y} \frac{1}{q^2} \sum_{d|q} \frac{\mu(d)}{d} \\ &= \sum_{d \leq 2y} \frac{\mu(d)}{d} \sum_{\substack{y < q \leq 2y \\ q \equiv 0 \pmod d}} \frac{1}{q^2} = \sum_{d \leq 2y} \frac{\mu(d)}{d} \sum_{y/d < r \leq 2y/d} \frac{1}{(dr)^2} \\ &= \sum_{d \leq 2y} \frac{\mu(d)}{d^3} \left(\frac{d}{2y} + O\left(\frac{d^2}{y^2}\right) \right) = \frac{1}{2y} \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + O\left(\frac{\log y}{y^2}\right). \end{aligned}$$

From this and

$$\sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} = \frac{1}{\zeta(2)},$$

we obtain (3.2). ■

4. Large sieve inequality with sparse sets of moduli. As stated before, large sieve is one of the key ideas in the proofs of our theorems. We have the following lemma.

LEMMA 7. *Suppose that S is a well-distributed set. Let $\{a_n\}$ be a sequence of complex numbers, $M \in \mathbb{Z}$, $Q, t \in \mathbb{N}$ and $t \leq Q$. Then*

$$(4.1) \quad \sum_{q \in S_t(Q/t)} \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^q \left| \sum_{n=M+1}^{M+N} a_n e\left(\frac{a}{q}n\right) \right|^2 \ll \left(N + \frac{Q}{t} (QN)^\varepsilon (\sqrt{N} + |S_t(Q/t)|) \right) \sum_{n=M+1}^{M+N} |a_n|^2,$$

where the implied constant depends at most on ε and S .

Proof. We observe from (2.1) that $(S_{t_1})_{t_2} = S_{t_1 t_2}$. Using this and (2.2), we deduce the lemma from Theorem 2 in [2] with \mathcal{S} replaced by $S_t(Q/t)$ and $X = Q^\varepsilon$ after a linear change of variables in n . ■

LEMMA 8. *Under the same conditions as in Lemma 7, we have*

$$(4.2) \quad \sum_{q \in S_t(Q/t)} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^* \left| \sum_{n=M+1}^{M+N} a_n \chi(n) \right|^2 \ll \left(N + \frac{Q}{t} (QN)^\varepsilon (\sqrt{N} + S_t(Q/t)) \right) \sum_{n=M+1}^{M+N} |a_n|^2,$$

where the implied constant depends at most on ε and S .

Proof. This lemma follows from Lemma 7 in the usual way by considering Gauss sums. ■

LEMMA 9. *Let $\{a_m\}$ and $\{b_n\}$ be two sequences of complex numbers. Then under the same conditions as in Lemma 7, we have*

$$(4.3) \quad \sum_{q \in S_t(Q/t)} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^* \max_X \left| \sum_{\substack{m \leq M \\ mn \leq X}} \sum_{n \leq N} a_m b_n \chi(mn) \right| \ll \log(2MN) (\Delta(M, Q, t) \Delta(N, Q, t) Z_a Z_b)^{1/2},$$

where the implied constant depends at most on ε and S ,

$$\Delta(Y, Q, t) = Y + \frac{Q}{t} (QY)^\varepsilon (\sqrt{Y} + |S_t(Q/t)|)$$

and

$$Z_a = \sum_{m \leq M} |a_m|^2, \quad Z_b = \sum_{n \leq N} |b_n|^2.$$

Proof. This lemma follows from Lemma 8 in the standard way, using the techniques of Section 28 in [9]. ■

5. Proof of Theorem 1. Following the arguments in Section 29 of [9], we have

$$(5.1) \quad \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^q \left(\psi(x; q, a) - \frac{x}{\varphi(q)} \right)^2 \ll (\log qx)^2 + \frac{1}{\varphi(q)} \sum_{\chi} |\psi'(x, \chi_1)|^2,$$

where

$$\psi'(y, \chi) = \begin{cases} \psi(y, \chi) & \text{if } \chi \neq \chi_0, \\ \psi(y, \chi_0) - y & \text{if } \chi = \chi_0, \end{cases}$$

and χ_1 is the primitive character that induces χ . Therefore, we have

$$(5.2) \quad \sum_{q \in S(Q)} \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^q \left(\psi(x; q, a) - \frac{x}{\varphi(q)} \right)^2 \\ \ll \sum_{\substack{t,q \\ Q \leq tq \leq 2Q \\ tq \in S(Q)}} \frac{1}{\varphi(q)\varphi(t)} \sum_{\chi \bmod q}^* |\psi'(x, \chi)|^2 + \sum_{q \in S(Q)} (\log qx)^2.$$

The second term on the right-hand side of (5.2) is

$$(5.3) \quad \ll |S(Q)|(\log Qx)^2.$$

It now remains to estimate the first term on the right-hand side of (5.2). If $Q \leq (\log x)^B$ for a large fixed $B > 0$, then we apply the Siegel–Walfisz theorem (Lemma 1) for the estimate. Otherwise, we divide the sum of interest into the following two sums:

$$\Sigma_1 = \sum_{q \leq 2Q(\log x)^C / |S(Q)|} \frac{1}{\varphi(q)} \sum_{t \in S_q(Q/q)} \frac{1}{\varphi(t)} \sum_{\chi \bmod q}^* |\psi'(x, \chi)|^2, \\ \Sigma_2 = \sum_{t \leq |S(Q)| / (\log x)^C} \frac{1}{\varphi(t)} \sum_{q \in S_t(Q/t)} \frac{1}{\varphi(q)} \sum_{\chi \bmod q}^* |\psi'(x, \chi)|^2,$$

where $2 < C < B/2$. Note that χ is never trivial in the innermost sum of Σ_2 since $q > 1$. Using (2.3), we have

$$(5.4) \quad \sum_{t \in S_q(Q/q)} \frac{1}{\varphi(t)} \ll (\log \log Q) |S_q(Q/q)| \frac{q}{Q} \ll (\log \log Q) |S(Q)| \frac{q^{1-\varepsilon}}{Q}.$$

Therefore,

$$\Sigma_1 \ll \frac{(\log \log Q) |S(Q)|}{Q} \sum_{q \leq 2Q(\log x)^C / |S(Q)|} \frac{q^{1-\varepsilon}}{\varphi(q)} \sum_{\chi \bmod q}^* |\psi'(x, \chi)|^2.$$

Using the Siegel–Walfisz theorem (Lemma 1) for $q \leq (\log x)^D$ with an appropriate $D > 0$ and the classical large sieve in Lemma 4 for $q > (\log x)^D$, we obtain

$$\Sigma_1 \ll \frac{|S(Q)|}{Q} \left(\frac{Q^2 (\log x)^{2C}}{|S(Q)|^2} + x \right) \frac{x}{(\log x)^A} \ll \frac{|S(Q)|}{Q} \frac{x^2}{(\log x)^A},$$

where we have used (2.4) and $Q \leq x^{1-\varepsilon}$ for the second inequality above. To estimate Σ_2 , we use Lemma 8, the fact that $|S_t(Q/t)| \leq |S(Q)|$ and the

conditions of the theorem to get

$$\begin{aligned} \Sigma_2 &\ll \sum_{t \leq |S(Q)|/(\log x)^C} \frac{t}{\varphi(t)Q} \left(x + \frac{Q}{t} (Qx)^\varepsilon (\sqrt{x} + |S_t(Q/t)|) \right) x \log x \\ &\ll \frac{|S(Q)|x^2}{Q(\log x)^A} + x^{3/2+\varepsilon} + x^{1+\varepsilon}|S(Q)| \ll \frac{|S(Q)|x^2}{Q(\log x)^A}, \end{aligned}$$

where $A = C - 2$. Combining the estimates for Σ_1 , Σ_2 and (5.3), we get the desired result.

6. Proof of Theorem 3. We will now prove Theorem 3. Following the arguments in Section 28 of [9], we have

$$(6.1) \quad \max_{\substack{a \\ \gcd(a,q)=1}} \left| \psi(x; q, a) - \frac{x}{\varphi(q)} \right| \ll (\log qx)^2 + \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\psi'(x, \chi)|,$$

where $\psi'(x, \chi)$ is as defined in the previous section and χ_1 is the primitive character that induces χ . Therefore,

$$\begin{aligned} (6.2) \quad \sum_{q \in S(Q)} \max_{\substack{a \\ \gcd(a,q)=1}} \left| \psi(x; q, a) - \frac{x}{\varphi(q)} \right| &\ll \sum_{\substack{t,q \\ Q \leq tq \leq 2Q \\ tq \in S(Q)}} \frac{1}{\varphi(q)\varphi(t)} \sum_{\chi \bmod q}^* |\psi'(x, \chi)| + \sum_{q \in S(Q)} (\log qx)^2. \end{aligned}$$

The second term on the right-hand side of (6.2) is

$$(6.3) \quad \ll |S(Q)|(\log Qx)^2.$$

It still remains to estimate the first term on the right-hand side of (6.2). If $Q \leq (\log x)^B$ for a large fixed $B > 0$, then we apply the Siegel–Walfisz theorem (Lemma 1) for the estimate. Otherwise, we divide the sum of interest into the following two sums:

$$(6.4) \quad \Sigma_1 = \sum_{q \leq 2Q/Q_1} \frac{1}{\varphi(q)} \sum_{t \in S_q(Q/q)} \frac{1}{\varphi(t)} \sum_{\chi \bmod q}^* |\psi'(x, \chi)|,$$

$$(6.5) \quad \Sigma_2 = \sum_{t \leq Q_1} \frac{1}{\varphi(t)} \sum_{q \in S_t(Q/t)} \frac{1}{\varphi(q)} \sum_{\chi \bmod q}^* |\psi(x, \chi)|,$$

where $Q_1 < Q$ is to be chosen later. Note again here that the innermost sum of Σ_2 never contains the trivial character. By virtue of (5.4), we have

$$\Sigma_1 \ll \frac{(\log \log Q)|S(Q)|}{Q} \sum_{q \leq 2Q/Q_1} \frac{q^{1-\varepsilon}}{\varphi(q)} \sum_{\chi \bmod q}^* |\psi'(x, \chi)|.$$

Using the Siegel–Walfisz theorem (Lemma 1) for $q \leq (\log x)^C$ with an appropriate $C > 0$ and Lemma 5 for $q > (\log x)^C$, we obtain

$$(6.6) \quad \Sigma_1 \ll \frac{|S(Q)|}{Q} \left(\frac{x}{(\log x)^A} + x^{5/6} \left(\frac{Q}{Q_1} \right)^{1-\varepsilon} + \sqrt{x} \left(\frac{Q}{Q_1} \right)^{2-\varepsilon} \right) (\log xQ)^5.$$

To estimate Σ_2 , we majorize Σ_2 in the following way:

$$(6.7) \quad \Sigma_2 \ll \frac{1}{Q} \sum_{t \leq Q_1} \frac{t}{\varphi(t)} \Sigma_2(t), \quad \text{where}$$

$$\Sigma_2(t) = \sum_{q \in S_t(Q/t)} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^* |\psi(x, \chi)|.$$

To estimate $\Sigma_2(t)$, we proceed as in the proof of the classical Bombieri–Vinogradov theorem (1.1). Using the Vaughan identity (Lemma 2), we have

$$(6.8) \quad \Sigma_2(t) \ll U(\log x) \frac{Q}{t} |S_t(Q/t)| + (\log x) \Sigma_{2,1}(t) + \Sigma_{2,2}(t) + \Sigma_{2,3}(t),$$

where

$$(6.9) \quad \Sigma_{2,1}(t) = \sum_{q \in S_t(Q/t)} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^* \sum_{l \leq \max(U,V)} \max_w \left| \sum_{w < k \leq x/l} \chi(kl) \right|$$

and

$$(6.10) \quad \Sigma_{2,i}(t) = \sum_{q \in S_t(Q/t)} \frac{q}{\varphi(q)} \times \sum_{\chi \bmod q}^* \left| \sum_{U \leq m \leq \max(UV, x/V)} \sum_{k \leq x/m} a_i(m) b_i(k) \chi(km) \right|$$

with $i = 2, 3$, $|a_i(m)| \leq \log x$ and $|b_i(k)| \leq \tau(k)$. We estimate (6.9) by using the Pólya–Vinogradov theorem (Lemma 3). We have

$$(6.11) \quad \Sigma_{2,1}(t) \ll \max(U, V) |S_t(Q/t)| (Q/t)^{3/2} \log x.$$

We split the sum over m in (6.10) into dyadic intervals and estimate the following:

$$(6.12) \quad \Sigma_{2,i}(t, M) = \sum_{q \in S_t(Q/t)} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^* \left| \sum_{M < m \leq 2M} \sum_{k \leq x/m} a_i(m) b_i(k) \chi(km) \right|.$$

By Lemma 9, we get

$$(6.13) \quad \Sigma_{2,i}(t, M) \ll (\log x)^4 \sqrt{x} \left(\frac{Q}{t} x^\varepsilon (|S_t(Q/t)| + \sqrt{M}) + M \right)^{1/2} \times \left(\frac{Q}{t} x^\varepsilon (|S_t(Q/t)| + \sqrt{\frac{x}{M}}) + \frac{x}{M} \right)^{1/2}.$$

Expanding (6.13), summing up the contributions from the dyadic intervals, setting $U = V$ and using (6.11), (6.8) and (6.7), we obtain

$$(6.14) \quad \Sigma_2 \ll x^\varepsilon Q^{1/2} U \Omega_{1,3/2} + x^{1/2+\varepsilon} \Omega_{1,1} + \left(\frac{x^{3/4+\varepsilon}}{U^{1/4}} + x^{1/2+\varepsilon} U^{1/2} \right) \Omega_{1/2,1} \\ + x^{3/4+\varepsilon} + \left(\frac{x^{1/2+\varepsilon} U}{Q^{1/2}} + \frac{x^{1+\varepsilon}}{Q^{1/2} U^{1/2}} \right) \Omega_{1/2,1/2} \\ + \left(\frac{x^{3/4+\varepsilon} U^{1/2}}{Q^{1/2}} + \frac{x^{1+\varepsilon}}{Q^{1/2} U^{1/4}} \right) Q_1^{1/2} + \frac{x(\log x)^5 Q_1}{Q},$$

where

$$\Omega_{a,b} = \sum_{t \leq Q_1} \frac{|S_t(Q/t)|^a}{t^b}.$$

Since $|S_t(Q/t)| \leq |S(Q)|$, we have

$$\Omega_{1,3/2} \ll |S(Q)|, \quad \Omega_{1,1} \ll |S(Q)| \log x, \quad \Omega_{1/2,1} \ll |S(Q)|^{1/2} \log x.$$

Moreover,

$$\Omega_{1/2,1/2} \leq \left(\sum_{t \leq Q_1} \frac{1}{t} \right)^{1/2} \left(\sum_{t \leq Q_1} |S_t(Q/t)| \right)^{1/2} \ll (\log x)^{1/2} \left(\sum_{q \in S} \sum_{t|q} 1 \right)^{1/2} \\ \ll x^\varepsilon |S(Q)|^{1/2}.$$

Consequently, the right-hand side of (6.14) is

$$(6.15) \quad \ll x^\varepsilon Q^{1/2} U |S(Q)| + x^{1/2+\varepsilon} |S(Q)| + \left(\frac{x^{3/4+\varepsilon}}{U^{1/4}} + x^{1/2+\varepsilon} U^{1/2} \right) |S(Q)|^{1/2} \\ + x^{3/4+\varepsilon} + \left(\frac{x^{1/2+\varepsilon} U}{Q^{1/2}} + \frac{x^{1+\varepsilon}}{Q^{1/2} U^{1/2}} \right) |S(Q)|^{1/2} \\ + \left(\frac{x^{3/4+\varepsilon} U^{1/2}}{Q^{1/2}} + \frac{x^{1+\varepsilon}}{Q^{1/2} U^{1/4}} \right) Q_1^{1/2} + \frac{x(\log x)^5 Q_1}{Q}.$$

Choosing $U = x^{1/4}$ and using (6.15) and the conditions of the theorem, we get

$$(6.16) \quad \Sigma_2 \ll \frac{|S(Q)| x^{1-\varepsilon}}{Q} + \frac{x^{15/16+\varepsilon}}{Q^{1/2}} Q_1^{1/2} + \frac{x(\log x)^5 Q_1}{Q}.$$

Now we choose

$$Q_1 = \min \left\{ \frac{x^{1/8-3\varepsilon} |S(Q)|^2}{Q}, \frac{|S(Q)|}{(\log x)^{A+5}} \right\}.$$

Using (6.2), (6.3), (6.6), (6.16) and the conditions of the theorem, we get the desired result.

7. Proof of Theorem 4. Let S be the set of squares. It suffices to estimate (6.4) and (6.5) with square moduli under the assumption $Q > (\log x)^B$ as was done in the previous section.

Using the first inequality in (5.4) and the estimate

$$|S_q(Q/q)| \ll \frac{|S(Q)|}{q^{1/2}} \ll \left(\frac{Q}{q}\right)^{1/2},$$

which follows from Section 6 of [2], we obtain

$$\Sigma_1 \ll \frac{\log \log Q}{Q^{1/2}} \sum_{q \leq 2Q/Q_1} \frac{q^{1/2}}{\varphi(q)} \sum_{\chi \bmod q}^* |\psi'(x, \chi)|.$$

Using the Siegel–Walfisz theorem (Lemma 1) for small moduli $q \leq (\log x)^B$ with an appropriate $B > 0$ and Lemma 5 for $q > (\log x)^B$, we obtain

$$(7.1) \quad \Sigma_1 \ll \frac{(\log xQ)^5}{Q^{1/2}} \left(\frac{x}{(\log x)^A} + x^{5/6} \left(\frac{Q}{Q_1}\right)^{1/2} + x^{1/2} \left(\frac{Q}{Q_1}\right)^{3/2} \right).$$

To estimate Σ_2 , we use (6.15). Choosing $U = x^{1/3}$ and using $|S(Q)| \ll Q^{1/2}$, we obtain

$$(7.2) \quad \Sigma_2 \ll x^\varepsilon (Qx^{1/3} + Q^{1/4}x^{2/3} + x^{3/4} + x^{5/6}Q^{-1/4} + x^{1/2}Q^{1/2} + x^{11/12}Q^{-1/2}Q_1^{1/2}) + x(\log x)^5 Q_1 Q^{-1}.$$

Now we choose $Q_1 = Q^{3/8-\varepsilon}$. Then if $(\log x)^C < Q \leq x^{4/9-\varepsilon}$ with a large $C > 0$, we obtain

$$\Sigma_1 + \Sigma_2 \ll \frac{x}{Q^{1/2}(\log x)^A}$$

from (7.1) and (7.2). We have therefore proved, in a manner similar to the previous section, that

$$(7.3) \quad \sum_{Q < q^2 \leq 2Q} \max_{\gcd(a,q)=1} \left| \psi(x; q^2, a) - \frac{x}{\varphi(q^2)} \right| \ll \frac{x}{Q^{1/2}(\log x)^A}$$

if $(\log x)^C < Q \leq x^{4/9-\varepsilon}$ with $C > 0$ chosen suitably. (7.3) also holds if $Q < (\log x)^C$ by the Siegel–Walfisz theorem (Lemma 1). Now, (7.3) implies the desired estimate of the theorem.

8. Proof of Theorem 5. We shall derive Theorem 5 from the following estimate.

LEMMA 10. *Let ε and A be arbitrary positive constants. If $(\log x)^{A+1} \leq y \leq x^{2/9-\varepsilon}$, then*

$$(8.1) \quad \sum_{x < n \leq 2x} \Lambda(n+1) \sum_{\substack{y < q \leq 2y \\ q^2 | n}} 1 = \frac{x}{2\zeta(2)y} + O\left(\frac{x}{y(\log x)^A}\right),$$

where the O -constant depends only on ε and A .

Proof. To establish (8.1), we write

$$\begin{aligned}
 (8.2) \quad & \sum_{x < n \leq 2x} \Lambda(n+1) \sum_{\substack{y < q \leq 2y \\ q^2 | n}} 1 = \sum_{y < q \leq 2y} \sum_{\substack{x+1 < k \leq 2x+1 \\ k \equiv 1 \pmod{q^2}} \Lambda(k) \\
 & = \sum_{y < q \leq 2y} (\psi(2x+1; q^2, 1) - \psi(x+1; q^2, 1)) \\
 & = x \sum_{y < q \leq 2y} \frac{1}{\varphi(q^2)} + O\left(\frac{1}{y} \sum_{q \leq 2y} q \left| \psi(2x+1; q^2, 1) - \frac{2x+1}{\varphi(q^2)} \right| \right. \\
 & \quad \left. + \frac{1}{y} \sum_{q \leq 2y} q \left| \psi(x+1; q^2, 1) - \frac{x+1}{\varphi(q^2)} \right| \right).
 \end{aligned}$$

Now (8.1) follows from (8.2), Theorem 4 and Lemma 6. ■

Let f be the characteristic function of the set of primes, that is,

$$f(n) = \begin{cases} 1 & \text{if } n \text{ is a prime,} \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, let $\tau(n)$ be the number of divisors of n . We have $s(n) \leq n^{5/9+\varepsilon}$ if and only if there exists $q \in \mathbb{N}$ such that $q^2 | n$ and $q \geq n^{2/9-\varepsilon/2}$. Therefore the number of primes $p \leq x$ such that $s(p-1) \leq p^{5/9+\varepsilon}$ is

$$(8.3) \quad \geq \sum_{n \leq x-1} f(n+1) \frac{1}{\tau(n)} \sum_{\substack{q \geq n^{2/9-\varepsilon/2} \\ q^2 | n}} 1.$$

Using $\tau(n) \ll n^{\varepsilon/2}$ and the definition of the von Mangoldt function $\Lambda(n)$, the expression in (8.3) is

$$(8.4) \quad \geq \left(c_1 x^{-\varepsilon/2} \sum_{n \leq x-1} \Lambda(n+1) \sum_{\substack{q \geq n^{2/9-\varepsilon/2} \\ q^2 | n}} 1 \right) - c_2 x^{1/2+\varepsilon}$$

for some positive constants c_1, c_2 . Moreover, we have

$$(8.5) \quad \sum_{n \leq x-1} \Lambda(n+1) \sum_{\substack{q \geq n^{2/9-\varepsilon/2} \\ q^2 | n}} 1 \geq \sum_{(x-1)/2 < n \leq x-1} \Lambda(n+1) \sum_{\substack{y < q \leq 2y \\ q^2 | n}} 1,$$

where $y := x^{2/9-\varepsilon/2}$. From (8.3)–(8.5) and Lemma 10, we deduce that the number of primes $p = n+1 \leq x$ with $s(n) \leq n^{5/9+\varepsilon}$ is $\gg x^{7/9}$. This implies the result of Theorem 5.

9. Implication of a conjecture. We start with the following conjecture, which can be viewed as a generalization of the conjecture given in [23].

CONJECTURE 1. Let S be the set of squares and $S_t(Q/t)$ as defined before. Let $M \in \mathbb{Z}$, $N \in \mathbb{N}$ and $\{a_n\}$ be a complex sequence. Then

$$(9.1) \quad \sum_{q \in S_t(Q/t)} \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^q \left| \sum_{n=M+1}^{M+N} a_n e\left(\frac{a}{q}n\right) \right|^2 \\ \ll Q^\varepsilon \left(\frac{Q}{t} |S_t(Q/t)| + N\right) \sum_{n=M+1}^{M+N} |a_n|^2,$$

where the implied constant depends on ε .

Under the truth of this conjecture, the range of summation for the outermost sum in (2.8) can be extended to $q \leq x^{1/4-\varepsilon}$. This better range would give the complete analogue of the classical Bombieri–Vinogradov theorem, up to an ε -factor, for square moduli.

Furthermore, from (2.8) with the extended range for q with $q \leq x^{1/4-\varepsilon}$, it would follow that there exist infinitely many primes p such that $s(p-1) \leq p^{1/2+\varepsilon}$. We note that the set of n such that $s(n) \leq n^{1/2+\varepsilon}$ is “almost” as sparse as the set of numbers $m^2 + n^4$ considered by Friedlander and Iwaniec [12]. Indeed, the number of $n \leq x$ such that $s(n) \leq n^{1/2+\varepsilon}$ is $O(x^{3/4+\varepsilon/2})$.

It is conceivable that an Elliott–Halberstam type hypothesis holds for primes in arithmetic progressions to square moduli, i.e., that (2.8) holds with the exponent $1/2 - \varepsilon$ in place of $2/9 - \varepsilon$. This would imply that there exist infinitely many primes p such that $s(p-1) \leq p^\varepsilon$. A result of this kind comes very close to the conjecture that there exist infinitely many primes of the form $n^2 + 1$ since the number of $n \leq x$ such that $s(n) \leq n^\varepsilon$ is $O(x^{1/2+\varepsilon/2})$.

Acknowledgements. This paper was written when the first- and second-named authors held postdoctoral fellowships at the Department of Mathematics and Statistics at Queen’s University and the Department of Mathematics at the University of Toronto, respectively. The authors wish to thank these institutions for their financial support.

References

- [1] N. C. Ankeny, *Representations of primes by quadratic forms*, Amer. J. Math. 74 (1952), 913–919.
- [2] S. Baier, *On the large sieve with sparse sets of moduli*, arXiv:math.NT/0512228, 2005.
- [3] S. Baier and L. Zhao, *An improvement for the large sieve for square moduli*, arXiv:math.NT/0512271, 2005.
- [4] —, —, *Large sieve inequality with characters for powerful moduli*, Int. J. Number Theory 1 (2005), 265–279.
- [5] A. Balog, *On the fractional part of p^θ* , Arch. Math. (Basel) 40 (1983), 434–440.

- [6] M. B. Barban, *On the distribution of primes in arithmetic progressions “on average”*, Dokl. Akad. Nauk UzSSR 1964, no. 5, 5–7 (in Russian).
- [7] E. Bombieri, *On the large sieve*, Mathematika 12 (1965), 201–225.
- [8] J. Brüdern, *Einführung in die analytische Zahlentheorie*, Springer, Berlin, 1995.
- [9] H. Davenport, *Multiplicative Number Theory*, 3rd ed., Grad. Texts in Math. 74, Springer, Barcelona, 2000.
- [10] H. Davenport and H. Halberstam, *Primes in arithmetic progressions*, Michigan Math. J. 13 (1966), 485–489; corrigendum, *ibid.* 15 (1968), 505.
- [11] —, —, *The values of a trigonometrical polynomial at well spaced points*, Mathematika 13 (1966), 91–96; corrigendum and addendum, *ibid.* 14 (1967), 229–232.
- [12] J. Friedlander and H. Iwaniec, *The polynomial $x^2 + y^4$ captures its primes*, Ann. of Math. 148 (1998), 945–1040.
- [13] P. X. Gallagher, *The large sieve*, Mathematika 14 (1967), 14–20.
- [14] G. H. Hardy and J. E. Littlewood, *Some problems of “Partitio Numerorum” III: On the expression of a number as a sum of primes*, Acta Math. 44 (1923), 1–70.
- [15] G. Harman and P. Lewis, *Gaussian primes in narrow sectors*, Mathematika 48 (2001), 119–135.
- [16] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Amer. Math. Soc. Colloq. Publ. 53, Amer. Math. Soc., Providence, RI, 2004.
- [17] J. P. Kubilius, *On a problem in the n -dimensional analytic theory of numbers*, Vilniaus Valst. Univ. Mokslo Darbai Mat. Fiz. Chem. Mokslu Ser. 4 (1955), 5–43 (in Lithuanian; Russian summary).
- [18] H. L. Montgomery, *Primes in arithmetic progressions*, Michigan Math. J. 17 (1970), 33–39.
- [19] H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika 20 (1973), 119–134.
- [20] I. I. Pyateckii–Šapiro [I. I. Pyatetskii–Shapiro], *On the distribution of prime numbers in sequences of the form $[f(n)]$* , Mat. Sb. N.S. 33 (75) (1953), 559–566 (in Russian).
- [21] R. C. Vaughan, *Mean value theorems in prime number theory*, J. London Math. Soc. 10 (1975), 153–162.
- [22] A. I. Vinogradov, *The density hypothesis for Dirichet L -series*, Izv. Akad. Nauk SSSR Ser. Mat. 29 (1965), 903–934 (in Russian); correction, *ibid.* 30 (1966), 719–720 (in Russian).
- [23] L. Zhao, *Large sieve inequality with characters to square moduli*, Acta Arith. 112 (2004), 297–308.
- [24] —, *Large sieve inequalities for special characters to prime square moduli*, Funct. Approx. Comment. Math. 32 (2004), 99–106.

Department of Mathematics and Statistics
 Queen’s University
 99 University Ave.
 Kingston, ON K7L 3N6, Canada
 E-mail: sbaier@mast.queensu.ca

Department of Mathematics
 University of Toronto
 40 Saint George Street
 Toronto, ON M5S 2E4, Canada
 E-mail: lzhao@math.toronto.edu

*Received on 14.2.2006
 and in revised form on 15.5.2006*

(5144)