

SOME PROBLEMS OF ANALYTIC NUMBER THEORY ON ARITHMETIC SEMIGROUPS

GLYN HARMAN & KAISA MATOMÄKI

Abstract: Let \mathcal{E} be a set of primes with density $\tau > 0$ in the set of primes. Write \mathcal{A} for the set of positive integers composed solely of primes from \mathcal{E} . We discuss the distribution of the integers from \mathcal{A} in short intervals, and whether for fixed $k \in \mathbb{Z}$ there are solutions to $m + k = p$ with $m \in \mathcal{A}$, where p denotes a prime, or $m + k = n$ where n has a large prime factor ($> n^\xi$ for $\xi > \frac{1}{2}$).

Keywords: greatest prime factors, distribution in short intervals.

1. Introduction

Throughout this paper we suppose that \mathcal{E} is a set of primes which usually will be assumed to satisfy

$$\sum_{\substack{p \in \mathcal{E} \\ p \leq x}} \frac{\log p}{p} = \tau \log x + O(1). \quad (1)$$

Let \mathcal{A} be the multiplicative semigroup generated by primes from \mathcal{E} , that is

$$\mathcal{A} = \{m : p|m \Rightarrow p \in \mathcal{E}\}.$$

Write

$$\mathcal{A}(y) = |\{n \in \mathcal{A} : n \leq y\}|.$$

It is well-known that (1) gives rise to the formula

$$\mathcal{A}(y) \sim \frac{e^{-\gamma\tau} y}{\Gamma(\tau) \log y} \prod_{\substack{p \leq y \\ p \in \mathcal{E}}} \left(1 - \frac{1}{p}\right)^{-1}. \quad (2)$$

See [26], and also compare the results in [4], [7]. Alternatively, see [21, §2.5] where the result is more general and is attributed to Bredikhin. One can obtain (2) under the alternative assumption that

$$\sum_{\substack{p \in \mathcal{E} \\ p \leq x}} 1 = (\tau + o(1)) \frac{x}{\log x}, \quad (3)$$

but we shall sometimes need the stronger expression (1) in order to sieve by a set of primes.

Given \mathcal{D} , a set of integers which satisfy some “natural” property, the tools of analytic number theory are often applied to determine the distribution of members of \mathcal{D} in short intervals, or to determine whether the shifted set $k + \mathcal{D}$ contains primes or prime-like elements. For examples where \mathcal{D} consists of the set of primes, see [3] and [2]. If \mathcal{E} consists of all primes in an arithmetic progression, such problems can often be tackled successfully for \mathcal{A} . For example, the case \mathcal{E} consists of all primes not congruent to $3 \pmod{4}$ is well-known. In this case $\tau = \frac{1}{2}$ and we know $1 + \mathcal{A}$ contains infinitely many primes [16], and the problem can even be considered in short intervals [18]. There does not seem to have been a systematic treatment of these problems for the general case when there is no further information given about the arithmetic structure of \mathcal{E} beyond the statement (1). It is the purpose of this paper to begin such an investigation. The motivation for this originally came from questions posed by Christian Elsholtz in the context of the additive and multiplicative decomposition of sets [8,9], but the problems appear to be of independent interest and so we have gathered a selection of results here. First we consider the short intervals problem.

Theorem 1.1. *Let $1 \geq \tau > \frac{1}{2}$ and suppose that (1) holds. Then there exists $\delta = \delta(\tau) > 0$ and $C(\mathcal{E}) > 0$ such that*

$$|\{n \in \mathcal{A} : x \leq n \leq x + x^{1-\delta}\}| > \frac{C(\mathcal{E})x^{1-\delta}}{(\log x)^{1-\tau}}, \quad (4)$$

for all large x . Moreover, if \mathcal{E} includes all sufficiently large primes in some arithmetic progression \pmod{q} then (4) holds for all $\tau > 0, \delta < \frac{1}{2}$ with $C(\mathcal{E})$ replaced by $C(\tau, q)$.

Remark. It may seem disappointing that we require $\tau > \frac{1}{2}$ here (unless we suppose \mathcal{E} includes all sufficiently large primes in some arithmetic progression), but this restriction arises naturally when using sieves to give non-trivial lower bounds when there is little other arithmetic structure known in a set.

Next we consider the possibility that $(\mathcal{A} + k) \cap \mathcal{P}$ has infinitely many members, where \mathcal{P} is the set of all primes. Here we need τ close to 1 to obtain the result.

Theorem 1.2. *Let k be an even non-zero integer. There exists $\tau_0 < 1$ such that if $\tau_0 \leq \tau \leq 1$, and \mathcal{E} satisfies (1), then there exists $C(\mathcal{E}) > 0$ such that*

$$|\{p + k \in \mathcal{A} : p \leq x\}| > \frac{C(\mathcal{E})x}{(\log x)^{2-\tau}} \quad (5)$$

for all sufficiently large x . The result also holds for odd k provided that $2 \in \mathcal{E}$.

The following gives a satisfactory solution (with certain applications in mind) for shifted sets having integers with large prime factors whenever $\tau > 0$.

Theorem 1.3. *Suppose that (3) holds for some $\tau > 0$. Let k be a non-zero integer and $\epsilon > 0$. Then there exists $C(\mathcal{E}) > 0$ such that there are $> C(\mathcal{E})x(\log x)^{\tau-1}$ solutions to*

$$m \in \mathcal{A}, \quad m + k = rp \leq x, \quad p > x^{\frac{3}{5}-\epsilon}.$$

For a variant of this problem, see [22].

It is possible to use the method of proof for Theorem 1.3 even in cases when $\tau = 0$. To discuss this case we suppose that there is a constant $B > 1$ such that for all large y we have

$$\sum_{\substack{p \in \mathcal{E} \\ y \leq p < 2y}} 1 \gg \frac{y}{(\log y)^B}. \quad (6)$$

Theorem 1.4. *Suppose that (6) holds for \mathcal{E} . Let k be a non-zero integer and $\epsilon > 0$. Then there exists $C'(\mathcal{E}) > 0$ such that there are $> C'(\mathcal{E})x(\log x)^{-3B}$ solutions to*

$$m \in \mathcal{A}, \quad m + k = rp \leq x, \quad p > x^{\frac{3}{5}-\epsilon}.$$

Indeed, the values of m counted have exactly 3 prime factors.

Finally we give an additional application of the method of proof of Theorem 1.3 to another problem that has not been considered in the literature before: shifted very smooth numbers with a large prime factor. Although not strictly a problem on arithmetic semigroups, it may have applications to questions of the type studied in [8,9].

Theorem 1.5. *Let $P(n)$ denote the greatest prime factor of a positive integer n . Let $k \in \mathbb{Z}, k \neq 0, \epsilon > 0$. Then there are infinitely many solutions to*

$$m + k = n, \quad P(m) < g(m, \epsilon), \quad P(n) > n^{\frac{3}{5}-\epsilon}, \quad (7)$$

where

$$g(m, \epsilon) = \exp\left((\log m)^{\frac{1}{2}+\epsilon}\right).$$

Remark. We note that the proof of Theorem 1.5 actually supplies $\gg x^{1-\epsilon}$ (and this lower bound can be improved slightly) solutions with $m \leq x$.

The proofs of these final three theorems depend on an adaptation of a deep result of Bombieri, Friedlander and Iwaniec given in [5]. The reader may wonder what may be achieved with less sophisticated technology, so in section 6 we prove a weaker version of Theorem 1.3 by completely elementary means.

2. Preliminary results

Let \mathcal{U} be the complement of \mathcal{E} in the set of primes and write $u = 1 - \tau$. So (1) gives

$$\sum_{\substack{p \in \mathcal{U} \\ p \leq x}} \frac{\log p}{p} = u \log x + O(1). \quad (8)$$

The important thing is that one can sieve by the primes in \mathcal{U} if they obey (8). Given any set of primes \mathcal{Q} we write

$$V(z, \mathcal{Q}) = \prod_{\substack{p < z \\ p \in \mathcal{Q}}} \left(1 - \frac{1}{p}\right).$$

More generally, for a multiplicative function $\omega(n)$ satisfying $0 \leq \omega(p) < p$ for $p \in \mathcal{Q}$, we put

$$V(z, \mathcal{Q}, \omega) = \prod_{\substack{p < z \\ p \in \mathcal{Q}}} \left(1 - \frac{\omega(p)}{p}\right).$$

We note that (2) with Mertens' prime number theorem gives

$$\mathcal{A}(y) \sim c_2(\tau) y V(y, \mathcal{U}), \quad (9)$$

where

$$c_2(\tau) = \frac{e^{u\gamma}}{\Gamma(\tau)}.$$

It will be important that the “constants” depending on τ in (9) and (2) are “well-behaved” for τ near 1. Of course, there is a dependence on the small primes in \mathcal{E} , but this affects all the terms that arise equally; asymptotically it is the relation (1) that is crucial. We thus write

$$c_1(\mathcal{E}) = \lim_{y \rightarrow \infty} (\log y)^u V(y, \mathcal{U}),$$

which is a positive real number by (1), and gives

$$V(z, \mathcal{U}) = \frac{c_1(\mathcal{E})(1 + o(1))}{(\log z)^u}.$$

We also write

$$c_3(\mathcal{E}) = \prod_{\substack{p \notin \mathcal{E} \\ p > 2}} \left(1 - \frac{1}{(p-1)^2}\right).$$

This is related to the twin prime constant (with equality if $\mathcal{E} = \emptyset$), and so we obtain $0.66 < c_3(\mathcal{E}) < 1$ for all possible \mathcal{E} .

Given a set of integers \mathcal{D} with a set of primes \mathcal{Q} and a positive parameter z we write

$$S(\mathcal{D}, z, \mathcal{Q}) = |\{n \in \mathcal{D} : p|n \Rightarrow p \geq z \text{ or } p \notin \mathcal{Q} \text{ or both}\}|, \quad (10)$$

and $|\mathcal{D}_d| = |\{n \in \mathcal{D} : n \equiv 0 \pmod{d}\}|$. We shall need the following sieve upper and lower bounds which can be found in [10, Chapter 4] or [17].

Lemma 2.1. *Let \mathcal{Q} be a set of primes, and $\omega(d)$ a multiplicative function for which*

$$\prod_{\substack{p \in \mathcal{Q} \\ w \leq p < z}} \left(1 - \frac{\omega(p)}{p}\right)^{-1} < \left(\frac{\log z}{\log w}\right)^\kappa \left(1 + \frac{K}{\log w}\right), \quad (11)$$

where $\kappa \geq 0$, and let \mathcal{D} be a set of positive integers. Write

$$I(t) = I(t, \kappa) = \int_0^\infty \frac{e^{-y}}{y^\kappa} \exp\left(t\kappa \int_y^\infty \frac{e^{-v}}{v} dv\right) dy,$$

and

$$P(z) = \prod_{\substack{p \leq z \\ p \in \mathcal{Q}}} p.$$

Suppose for $d|P(z)$ we have

$$|\mathcal{D}_d| = \frac{X\omega(d)}{d} + R_d.$$

Let $y \geq z \geq 2$, $s = (\log y)/(\log z)$, $1 \leq s \leq 2$. Then

$$S(\mathcal{D}, z, \mathcal{Q}) \leq \frac{AX}{s^\kappa} V(z, \mathcal{Q}, \omega)(1 + o_K(1)) + \sum_{\substack{d \leq y \\ d|P(z)}} |R_d|. \quad (12)$$

Also, if $\kappa < \frac{1}{2}$, $s = 1$, we have

$$S(\mathcal{D}, z, \mathcal{Q}) \geq \frac{BX}{s^\kappa} V(z, \mathcal{Q}, \omega)(1 + o_K(1)) - \sum_{\substack{d \leq y \\ d|P(z)}} |R_d|. \quad (13)$$

Here

$$B = B(\kappa) = \frac{2e^{\gamma\kappa}}{I(1, \kappa) + I(-1, \kappa)}$$

and

$$A = A(\kappa) = \begin{cases} G(\kappa) & \text{if } 1 \leq \kappa \leq 2, \\ C(\kappa) & \text{if } \kappa < \frac{1}{2}, \end{cases}$$

where

$$C = \frac{BI(1, \kappa)}{\Gamma(1 - \kappa)},$$

and $G(\kappa) \leq K_1$ for an absolute constant K_1 .

The following well-known result will be needed in the next section.

Lemma 2.2 (Bombieri-Vinogradov Theorem). *Let $A > 0$ be given. Suppose that $x^{1/2}(\log x)^{-A} \leq Q \leq x^{1/2}$. Then*

$$\sum_{q \leq Q} \max_{y \leq x} \max_{(a,q)=1} \left| \sum_{\substack{p \leq y \\ p \equiv a \pmod{q}}} 1 - \frac{\text{Li}(x)}{\phi(q)} \right| \ll_A x^{\frac{1}{2}} Q (\log x)^5. \quad (14)$$

Here, as usual,

$$\text{Li}(x) = \int_2^x \frac{1}{\log v} dv \sim \frac{x}{\log x}.$$

Proof. See [6, Chapter 28], where we have removed the von Mangoldt function by partial summation. ■

We now state a result that enables us to go beyond the “one half” barrier of the Bombieri-Vinogradov theorem in certain contexts (see [1] for an application). This is a variant of [5, Theorem 3]. Since we need to make several modifications to the work of Bombieri, Friedlander and Iwaniec, we devote the final section of this paper to giving an outline of the proof.

Lemma 2.3. *Let $\epsilon > 0$, $t \in \mathbb{Z}$, $t \neq 0$, and $K, L, M > x^\epsilon$. Let z be a real parameter satisfying*

$$2 \leq z \leq \exp\left((\log x)^{\frac{1}{2}}\right).$$

Suppose that $KLM = x$, and write $n \sim N$ to mean $\frac{1}{2}N < n \leq N$. Let $x^{\frac{1}{2}} < Q < x^{\frac{2}{3}-\epsilon}$, with

$$\begin{aligned} Q &< KLx^{-\epsilon}, \\ K^2L^3 &< Qx^{1-\epsilon}, \end{aligned}$$

and

$$K^4L^2(K+L) < x^{2-\epsilon}.$$

Then there exists $A = A(\epsilon)$ such that for any sequences $\lambda_\ell, \mu_m, \kappa_k$ satisfying

$$0 \leq \kappa_k, \lambda_\ell, \mu_m \leq 1,$$

and

$$p|k\ell \Rightarrow p \geq z \quad \text{if } \kappa_k \lambda_\ell \neq 0, \quad (15)$$

we have

$$\sum_{\substack{q \sim Q \\ q \text{ prime}}} \Delta(q) \ll \frac{x(\log x)^A}{z^{\frac{1}{4}}} \quad (16)$$

where

$$\Delta(q) = \left| \sum_{k \sim K} \kappa_k \sum_{\ell \sim L} \lambda_\ell \sum_{\substack{m \sim M \\ k\ell m \equiv t \pmod{q}}} \mu_m - \frac{1}{\phi(q)} \sum_{k \sim K} \kappa_k \sum_{\ell \sim L} \lambda_\ell \sum_{m \sim M} \mu_m \right|.$$

The implied constant in (16) depends only on ϵ and t .

Remark. We have made two significant changes to [5, Theorem 3]. First, we have removed the necessity for one or more of the coefficients to satisfy a Siegel-Walfisz type condition, that is, for small q , a formula of the type

$$\sum_{\substack{\ell \sim L \\ \ell \equiv a \pmod{q}}} \lambda_\ell = \frac{1 + O((\log x)^{-A})}{\phi(q)} \sum_{\substack{\ell \sim L \\ (\ell, q) = 1}} \lambda_\ell.$$

Clearly we cannot assume such a condition with our weak hypothesis on the sets under consideration. Second, we have increased the saving on the trivial estimate. That is, we can save more than the traditional power of a logarithm. To compensate, we have assumed (all that we need in our present application) that the sum over q in (16) is only over prime values. This could be relaxed to

$$p|q \Rightarrow p > \exp\left((\log x)^{\frac{1}{2}}\right),$$

which is what we actually use later.

3. Proof of Theorem 1.2

In the following $K_j, j = 2, 3, \dots$ denote absolute constants. We note that $I(-1)$ and $I(1)$ are bounded above and below by positive absolute constants if (as will be the case in this section's application) $0 \leq \kappa \leq \frac{1}{4}$. Write $S(x)$ for the left hand side of (5). We only treat the case $k = 2$, and so we can assume $2 \in \mathcal{U}$ since $p + 2$ will be odd for $p \geq 3$. For general k various expressions involving products over primes in the following will need to be altered to either include or exclude the prime divisors of k . Now let

$$\mathcal{D} = \{p + 2 : p \leq x\},$$

and put $\mathcal{Q} = \mathcal{U} \setminus \{2\}$ in the definition (10) (since we automatically have $p + 2$ odd).

In Lemma 2.1 we can take the parameter y , commonly referred to as the “distribution level”, to be $x^{1/2}(\log x)^{-10}$ by Lemma 2.2 and we also take this to be the value for z . Here we have $X = \text{Li}(x)$ and $\omega(d) = d/\phi(d)$ giving $\omega(p) = p/(p-1)$. From (13) we deduce that

$$\begin{aligned} S(\mathcal{D}, z, \mathcal{U}) &> \frac{Bx}{\log x} (1 + o(1)) V(z, \mathcal{Q}, \omega) \\ &= \frac{2^u B c_1(\mathcal{E}) c_3(\mathcal{E}) (1 + o(1)) x}{(\log x)^{1+u}}. \end{aligned} \tag{17}$$

We note that B is bounded above and below by absolute constants under our assumption on u .

Now

$$S(x) = S(\mathcal{D}, z, \mathcal{U}) - Y,$$

where Y counts the number of solutions to

$$p + 2 = mn, \quad p \leq x \tag{18}$$

with

$$q|m \Rightarrow q \geq z, q \in \mathcal{U}, \quad \text{and} \quad n \in \mathcal{A}.$$

Clearly m is either a prime or has exactly two prime factors, both of which must exceed z . Write \mathcal{F} for the set of such integers up to x . We consider the case $n \leq x^{3\xi}$, say, where ξ will be determined later, separately. For each such n write

$$\mathcal{F}(n) = \{m \in \mathcal{F} : mn = p + 2, p \leq x\}.$$

We can bound the number of elements in $\mathcal{F}(n)$ for each n using Theorem 3.12 from [11]. In doing this we discount the condition $q|m \Rightarrow q \in \mathcal{U}$. For each n we thus obtain an upper bound

$$8c_3(\emptyset) \frac{x(1+o(1))}{n(\log(x/n))^2} \prod_{2 < p|n} \left(1 - \frac{1}{p-1}\right)^{-1}.$$

This gives an upper bound which summed over n becomes

$$\leq K_2 \frac{x}{(\log x)^2} \sum_{\substack{n \in \mathcal{A} \\ n \leq x^{3\xi}}} \frac{1}{\phi(n)} \leq K_3 c_1(\mathcal{E}) \frac{\xi^\tau}{(\log x)^{1+u}}, \tag{19}$$

where K_2, K_3 are absolute constants (assuming $0 \leq u \leq \frac{1}{4}$).

We now consider the case $n \geq x^{3\xi}$. We need to apply an upper bound sieve to detect that p is a prime in (18) and that $n \in \mathcal{A}$. To do this we define a new set

$$\mathcal{G} = \{(nm - 2)(\Pi_1 n + \Pi_2) : mn - 2 \leq x\},$$

where

$$\Pi_1 = \prod_{\substack{p \in \mathcal{E} \\ p \leq \xi}} p, \quad \Pi_2 = \prod_{\substack{p \in \mathcal{U} \\ p \leq x^\xi}} p.$$

Now suppose that m, n give a solution to (18). This certainly gives $p|(mn - 2) \Rightarrow p > x^\xi$. But also, since $n \in \mathcal{A}$, $p|(\Pi_1 n + \Pi_2) \Rightarrow p > x^\xi$. We can therefore bound the number of solutions to (18) for each m by estimating $S(\mathcal{G}, x^\xi, \mathcal{P})$. Write $\omega(d)$ for the number of solutions to

$$(nm - 2)(\Pi_1 n + \Pi_2) \equiv 0 \pmod{d}.$$

Then $\omega(d)$ is a multiplicative function with

$$\omega(p) = \begin{cases} 1 & \text{if } p \notin \mathcal{U} \text{ or } p = 2, \\ 2 & \text{if } p \in \mathcal{U}, p \neq 2. \end{cases}$$

We obtain an upper bound for $S(\mathcal{G}, x^\xi, \mathcal{P})$ from Lemma 2.1 with $z = x^\xi, y = z^2$. This bound is

$$\leq K_4 c_1(\mathcal{E}) c_3(\mathcal{E}) \frac{x(1+o(1))}{m(\xi \log x)^{1+u}} + \sum_{\substack{d \leq y \\ d|P(z)}} |R_d|.$$

To estimate the remainder term we note that $|\mathcal{G}_d|$ is the sum over $d_1 d_2 = d$ of the number of solutions in n to the simultaneous congruences

$$\begin{aligned} nm - 2 &\equiv 0 \pmod{d_1} \\ \Pi_1 n + \Pi_2 &\equiv 0 \pmod{d_2}. \end{aligned}$$

There will always be exactly one solution to the first congruence $\pmod{d_1}$ (since $p|m \Rightarrow p > x^\xi$), while the second congruence has a solution $\pmod{d_2}$ if and only if d_2 has all its prime factors from \mathcal{U} . Hence

$$|\mathcal{G}_d| = \sum_{\substack{d_1 d_2 = d \\ p|d_2 \Rightarrow p \in \mathcal{U}}} \left(\frac{x}{m d} + O(1) \right) = \frac{x}{m d} \omega(d) + O(\tau(d)).$$

It follows that

$$\sum_{\substack{d \leq y \\ d|P(z)}} |R_d| \ll x^{2\xi} \log x.$$

Since

$$\sum_{m \in \mathcal{F}} \frac{1}{m} \ll u,$$

we eventually obtain an upper bound

$$\leq K_5 u \xi^{-1-u} c_1(\mathcal{E}) c_3(\mathcal{E}) \frac{x(1+o(1))}{(\log x)^{1+u}}.$$

Comparing this with (19) we should take:

$$\xi = \left(\frac{K_5 c_3(\mathcal{E}) u}{K_3} \right)^{\frac{1}{2}}.$$

We then have

$$Y \leq K_7 u^{\tau/2} c_1(\mathcal{E}) \frac{x}{(\log x)^{1+u}},$$

where K_7 is bounded independently of τ and \mathcal{E} so long as $\tau \geq \frac{3}{4}$. It follows that, for all sufficiently small u , we have, for sufficiently large x ,

$$S(x) > \frac{C(\mathcal{E})x}{(\log x)^{2-\tau}},$$

which completes the proof.

4. Proof of Theorem 1.1

Let $w = x^{1-\delta}$. We now put

$$\mathcal{D} = \mathbb{N} \cap [x, x+w].$$

We take y , the distribution level in the lower bound sieve (13), to be $w/(\log x)^{10}$, with this also being the value for z , and $\omega(p) \equiv 1$. A satisfactory estimation of the remainder term in the sieve is then completely elementary. We thus get a lower bound

$$S(\mathcal{D}, z, \mathcal{U}) > \frac{Bc_1(\mathcal{E})w}{(\log z)^u} (1 + o(1)). \quad (20)$$

We note that B depends on τ , but not on δ . It remains to obtain an upper bound for Z , the number of solutions to

$$mp \in [x, x+w], \quad m \in \mathcal{A}, \quad p \in \mathcal{U}, p > z.$$

Let Z^* be the number of solutions to this with the condition $p \in \mathcal{U}$ removed. Assuming, as we may, that $\delta < \frac{5}{24}$, we can appeal to Huxley's prime number theorem [15] to obtain an asymptotic formula

$$Z^* = \frac{w(1+o(1))}{\log z} \sum_{m \leq g} \frac{1}{m} \leq K_8 c_1(\mathcal{E}) \delta^\tau \frac{w}{(\log x)^{1-\tau}}. \quad (21)$$

Here $g = (x+w)/z$. Comparing (20) with (21) we can find $\delta > 0$ such that $Z \leq \frac{1}{2} S(\mathcal{D}, z, \mathcal{U})$ and the proof of the first part of Theorem 1.1 is complete.

To prove the second part of Theorem 1.1 we first suppose that \mathcal{E} contains all sufficiently large primes $\equiv a \pmod{q}$. We write χ for a character \pmod{q} , $\eta = (\log x)^{-\frac{3}{4}}$, and note that for all sufficiently large x , no L -function $L(s, \chi)$ has a zero in the region

$$1 - 2\eta \leq \operatorname{Re} s < 1, \quad -x < \operatorname{Im} s < x. \quad (22)$$

Here we have merely noted that the standard Vinogradov zero-free region for the Riemann zeta-function [24, p.135] holds for $L(s, \chi)$ since we are treating q as a constant. Let $\epsilon > 0$ and put

$$\mathcal{B} = \left[x^{\frac{1}{2}-2\epsilon}, x^{\frac{1}{2}-\epsilon} \right].$$

We then consider numbers of the form $p mn \in [x, x + w] = \mathcal{J}$, say, with

$$p \in \mathcal{E} \cap \mathcal{B}, \quad m \equiv a \pmod{q}, \quad n \in \mathcal{A} \cap \mathcal{B}.$$

Let

$$V = \sum_{p \in \mathcal{E} \cap \mathcal{B}} \sum_{n \in \mathcal{A} \cap \mathcal{B}} \sum_{\substack{m \equiv a \pmod{q} \\ pmn \in \mathcal{J}}} \Lambda(m).$$

Write

$$P(s) = \sum_{p \in \mathcal{E} \cap \mathcal{B}} p^{-s}, \quad A(s) = \sum_{n \in \mathcal{A} \cap \mathcal{B}} n^{-s}.$$

Then, by the truncated Perron formula [24, Lemma 3.19] we have

$$\begin{aligned} V = & -\frac{1}{2\pi i \phi(q)} \sum_{\chi \pmod{q}} \bar{\chi}(a) \int_{c-iW}^{c+iW} \frac{L'}{L}(s, \chi) P(s) A(s) \frac{(x+w)^s - x^s}{s} ds \\ & + O(w(\log x)^{-10}). \end{aligned}$$

Here $W = (x/w)(\log x)^{20}$ and $c = 1 + (\log x)^{-1}$. In time-honoured fashion we move the line of integration back to $\operatorname{Re} s = d = 1 - \eta$ to pick up the main term from the pole of $L(s, \chi_0)$ at $s = 1$ which contributes

$$\frac{wA(1)P(1)}{\phi(q)} \gg_{\epsilon, q} w(\log x)^\tau.$$

The contribution from the horizontal line segments is negligible, while on $\operatorname{Re} s = d$ we have

$$\frac{L'}{L}(s, \chi) \ll (\log x)^2$$

by [6, Chapter 16] using (22). The integral on $\operatorname{Re} s = d$ is therefore

$$\begin{aligned} & \ll_q (\log x)^2 w \exp\left(-(\log x)^{\frac{1}{4}}\right) \times \\ & \times \left(\int_{d-iW}^{d+iW} |A(s)|^2 |ds| \right)^{\frac{1}{2}} \left(\int_{d-iW}^{d+iW} |P(s)|^2 |ds| \right)^{\frac{1}{2}}. \end{aligned}$$

Using the mean value theorem of Montgomery and Vaughan [20] this is

$$\ll w(\log x)^2 \left(W x^{(\frac{1}{2}-\epsilon)(1-2d)-(1-d)} + x^{(\frac{1}{2}-\epsilon)(2-2d)-(1-d)} \right).$$

If we choose $\epsilon = \frac{1}{3}(\frac{1}{2} - \delta)$ this leads to an error

$$\ll w(\log x)^2 \exp\left(-2\epsilon(\log x)^{\frac{1}{4}}\right) \ll w(\log x)^{-10},$$

which is of a suitable size.

The proof is completed by noting that the terms with $m = p^r$ contribute a smaller order term,

$$\Lambda(m) \leq 4\epsilon \log x,$$

and any number of the required form in \mathcal{J} is counted $\ll_{\epsilon} 1$ times.

5. Proof of Theorem 1.3 and related results

The trick is to count solutions coming from only a certain subset of \mathcal{A} designed to satisfy the hypotheses of Lemma 2.3. To be precise, consider

$$p_1 p_2 m + k = r q, \tag{23}$$

with $m \in \mathcal{A}$ and p_1, p_2, q primes satisfying

$$p_j \in \mathcal{E}, \quad p_1 \sim K, \quad p_2 \sim L, \quad q \sim Q.$$

Here we take $K = 2^{-j} x^{\frac{1}{5}}, L = 2^{-\ell} x^{\frac{2}{5} - 2\delta}$, where $\delta = \epsilon/6$. The values taken by j, ℓ are $0, 1, \dots, J$, with

$$J = [\delta \log_2 x]$$

and \log_2 here represents logarithm to base 2. We shall apply Lemma 2.3 with the ϵ there as the present parameter δ . The bound $Q < KLx^{-\delta}$ is then satisfied whenever $Q < x^{\frac{3}{5} - 4\delta}$. The bound $K^2 L^3 < Qx^{1-\delta}$ will hold provided $Q > x^{\frac{3}{5} - 5\delta}$. Finally the bound $K^4 L^2 (K + L) < x^{2-\delta}$ is established independently of Q . Let $G(K, L, q)$ denote the number of solutions to (23) for each fixed q , and write

$$H(K, L) = \sum_{\substack{p_1 \in \mathcal{E} \\ p_1 \sim K}} \sum_{\substack{p_2 \in \mathcal{E} \\ p_2 \sim L}} \sum_{\substack{m \in \mathcal{A} \\ m \sim x/KL}} 1.$$

We note that, with $M = x/(KL)$, we have

$$H(K, L) = (1 + o(1)) \frac{1}{8} \tau^2 c_2(\tau) c_1(\mathcal{E}) \frac{K}{\log K} \frac{L}{\log L} \frac{M}{(\log M)^u}.$$

Then, by (16), we have

$$\sum_{\substack{q \sim Q \\ q \text{ prime}}} \left| G(K, L, q) - \frac{1}{q} H(K, L) \right| \ll x(\log x)^A \exp\left(-\frac{1}{4}(\log x)^{\frac{1}{2}}\right),$$

for some A . The right hand side above is then clearly $\ll x(\log x)^{-10}$. Summing over the possible values for K, L and Q we obtain that the number of solutions to (23) with

$$x^{\frac{1}{5} - \delta} \leq p_1 \leq x^{\frac{1}{5}}, \quad x^{\frac{2}{5} - 3\delta} \leq p_2 \leq x^{\frac{2}{5} - 2\delta}, \quad x^{\frac{3}{5} - 5\delta} < q < x^{\frac{3}{5} - 4\delta},$$

is $> C(\mathcal{E})x(\log x)^{\tau-1}$, as required.

To prove Theorem 1.4 we simply replace (23) with $p_1 p_2 p_3 + k = r q$ and use (6) to give a lower bound on the numbers being counted.

Finally, to prove Theorem 1.5 we replace (23) with $m_1 m_2 m_3 + k = r q$ where now

$$p | m_j \Rightarrow \exp\left((\log x)^{\frac{1}{2}}\right) < p \leq \exp\left((\log w)^{\frac{1}{2} + \epsilon}\right) \quad (24)$$

with $w = x^{\frac{1}{2}}$. This immediately gives $P(m_1 m_2 m_3) < g(w, \epsilon)$ for all large x , of course. Since $m_1 m_2 m_3 > w$ we can safely conclude that $P(m) < g(m, \epsilon)$ with $m = m_1 m_2 m_3$. For a satisfactory completion of the proof we need the following result.

Lemma 5.1. *Let $\Theta(x, y, z)$ denote the number of integers from x to $2x$, all of whose prime factors are in the range $z \leq p \leq y$. Then, writing $u = (\log x)/(\log y)$, and assuming that $(\log 2u)(\log z) \leq \log y$ we have*

$$\Theta(x, y, z) = \frac{e^{-\gamma}(1 + o(1))}{\log z} \Psi(x, y), \quad (25)$$

where $\Psi(x, y)$ denotes the number of integers from x to $2x$ all of whose prime factors are $\leq y$. Moreover, if

$$y = \exp((\log x)^\alpha),$$

with $0 < \alpha < 1$, then

$$\Psi(x, y) = x \exp\left(-(\log x)^{1-\alpha+o(1)}\right) \quad (26)$$

Proof. See [14] and [23]. Here we have changed the standard notation slightly to count integers from x to $2x$ rather than the usual 1 to x . This makes no difference to the estimates for the parameters in the ranges we require. ■

To finish the proof, given $\epsilon > 0$ we replace ϵ by $\delta = \epsilon/6$ as above in the proof of Lemma 2.3. The coefficients in that lemma now are the characteristic functions of the set of integers satisfying (24). In that lemma we can therefore take

$$z = \exp\left((\log x)^{\frac{1}{2}}\right).$$

The reader can quickly verify that all the conditions are then satisfied to complete the proof. That is, we save a factor $z^{-\frac{1}{4}}(\log x)^A$ by Lemma 2.3, while the number of integers counted is

$$\gg x \exp\left(-(\log x)^{\frac{1}{2} - \epsilon + o(1)}\right)$$

by Lemma 5.1, which is of a larger order than $x z^{-\frac{1}{4}}(\log x)^A$.

6. A different approach to a Theorem 1.3 type problem

Here we prove a weaker version of Theorem 1.3. For this we do not need the deep results of [5]. The proof is based on the large sieve inequality and other elementary sieve results. Note that we do not need to use the Siegel-Walfisz theorem.

Theorem 6.1. *Suppose that (3) holds for some $\tau > 0$. Let k be a non-zero integer. Then there exists $\xi(\tau) > \frac{1}{2}$, $C(\mathcal{E}) > 0$ such that there are $> C(\mathcal{E})x(\log x)^{\tau-1}$ solutions to*

$$m \in \mathcal{A}, \quad m + k = rp \leq x, \quad p > x^\xi.$$

The major problem we encounter in adapting previous work on large prime factors of shifted sets is that we cannot prove a Bombieri-Vinogradov theorem for \mathcal{A} . To circumvent this difficulty we work instead with

$$\mathcal{A}' = \{rq : q \in \mathcal{E}, r \in \mathcal{A}, x^{\frac{2}{3}} < q < x^{\frac{3}{4}}, \frac{1}{2}x \leq rq \leq x\}.$$

The exponents $\frac{2}{3}, \frac{3}{4}$ are two arbitrary values strictly between $\frac{1}{2}$ and 1. Of course $m \in \mathcal{A}' \Rightarrow m \in \mathcal{A}$, and m could be counted only once in \mathcal{A}' . Also

$$|\mathcal{A}'| \gg |\mathcal{A}(x)|.$$

We go back to an old idea of Chebychev that has been used to great effect in finding numbers with a large prime factor in certain sequences. See, in particular, [12], which has certain features in common with our approach. We have

$$\begin{aligned} \sum_{m \in \mathcal{A}'} \sum_{p|(m+k)} \log p &= \sum_{m \in \mathcal{A}'} \sum_{t|(m+k)} \Lambda(t) + O\left(\sum_{m \in \mathcal{A}'} \sum_{p^r|(m+k), r>1} \log p\right) \\ &= \sum_{m \in \mathcal{A}'} \sum_{t|(m+k)} \Lambda(t) + O(x) \\ &= \sum_{m \in \mathcal{A}'} \log(m+k) + O(x) \\ &= |\mathcal{A}'| \log x + O(x). \end{aligned}$$

Since $|\mathcal{A}'| \gg_\tau c_1(\mathcal{E})x(\log x)^{\tau-1}$, we have

$$\sum_{m \in \mathcal{A}'} \sum_{p|(m+k)} \log p \geq (1 - \epsilon)|\mathcal{A}'| \log x$$

for any $\epsilon > 0$ for all sufficiently large x . It now only remains to show that for some $\xi > \frac{1}{2}$ we have

$$\sum_{p \leq x^\xi} \log p \sum_{\substack{m \in \mathcal{A}' \\ m \equiv -k \pmod{p}}} 1 < (1 - \epsilon)|\mathcal{A}'|.$$

Let $E = (\log x)^{10}$, $\eta = E^{-1}$. We initially consider

$$S_1 = \sum_{E < p < x^{\frac{1}{2}\eta}} \log p \sum_{\substack{m \in \mathcal{A}' \\ m \equiv -2 \pmod{p}}} 1.$$

To deal with this sum we first establish the following Bombieri-Vinogradov type theorem for \mathcal{A}' .

Lemma 6.2. *With the above notation we have*

$$\sum_{E < p < x^{\frac{1}{2}}\eta} \left| \sum_{\substack{m \in \mathcal{A}' \\ m \equiv -k \pmod{p}}} 1 - \frac{|\mathcal{A}'|}{p-1} \right| \ll x(\log x)^{-5}. \quad (27)$$

Proof. We go back to Vaughan's elementary proof of the Bombieri-Vinogradov theorem [25]. The crucial fact is that we have forced \mathcal{A}' to be a Type II sum of the correct shape for (27) to follow. Indeed, (27) is an almost immediate consequence of the large sieve inequality for character sums expressed as (see [6, p.164])

$$\sum_{q \leq Q} \frac{q}{\phi(q)} \sum_{\chi \text{ primitive (mod } q)} \left| \sum_{\ell \leq L} a_\ell \chi(\ell) \right|^2 \ll (N + Q^2) \sum_{\ell \leq L} |a_\ell|^2.$$

Of course, all the non-principal characters are primitive in our application since we are dealing with prime moduli only. \blacksquare

It follows from (27) that

$$\begin{aligned} S_1 &\leq O(x(\log x)^{-5}) + |\mathcal{A}'| \sum_{E < p < x^{\frac{1}{2}}\eta} \frac{\log p}{p-1} \\ &= \frac{1}{2} |\mathcal{A}'| (\log x + O(\log \log x)). \end{aligned}$$

Let S_2, S_3 denote the remaining sums from $p \leq E$ and $x^{\frac{1}{2}}\eta \leq p \leq x^\epsilon$ respectively. We need to bound the number of solutions to

$$rq \equiv -k \pmod{p}, \quad x^{\frac{2}{3}} < q < x^{\frac{3}{4}}, \quad \frac{1}{2}x \leq rq \leq x$$

for each fixed $r \in \mathcal{A}$ and each fixed p . Here we discard the condition $q \in \mathcal{E}$ and so we can then appeal to the well-known Brun-Titchmarsh inequality in the form given by Montgomery and Vaughan [19] (since we are not aiming to get the best possible result here one could use earlier forms). We thus obtain an upper bound

$$\leq \frac{2x}{r(p-1) \log(x/(2rp))}.$$

We then have

$$\begin{aligned} S_2 &\leq \sum_{p \leq E} \log p \sum_{\substack{r \in \mathcal{A} \\ \frac{1}{2}x^{\frac{1}{4}} < r < x^{\frac{1}{3}}}} \frac{2x}{r(p-1) \log(x/(2rp))} \\ &\ll x(\log E)(\log x)^{\tau-1} = o(|\mathcal{A}'| \log x). \end{aligned}$$

We can therefore neglect this error.

The bound for S_3 sets the maximum permissible value for ξ . We have

$$\begin{aligned} S_3 &\leq \sum_{x^{\frac{1}{2}}\eta < p \leq x^\xi} \log p \sum_{\substack{r \in \mathcal{A} \\ \frac{1}{2}x^{\frac{1}{4}} < r < x^{\frac{1}{3}}} } \frac{2x}{r(p-1)\log(x/(2rp))} \\ &\leq (\xi - \frac{1}{2})c'(\tau)c_1(\mathcal{E})x(\log x)^\tau, \end{aligned}$$

for all large x , where $c'(\tau) > 0$. By choosing ξ sufficiently close to $\frac{1}{2}$ we can ensure that this is $< (\frac{1}{2} - \epsilon)|\mathcal{A}'|\log x$ and the proof is complete.

7. Proof of Lemma 2.3

In the following we shall assume for simplicity that

$$z = \exp\left((\log x)^{\frac{1}{2}}\right).$$

Let

$$\beta_n = \sum_{k\ell=n} \kappa_k \lambda_\ell, \quad \Omega = \{q \sim Q : p|q \Rightarrow p > z\}.$$

Let $\rho(n)$ be the characteristic function of the set Ω . Then, if $n \sim Q$ we write

$$\rho(n) = \rho_1(n) + \rho_2(n), \tag{28}$$

where

$$\rho_1(n) = \sum_{\substack{d|n \\ d|P(z) \\ d \leq D}} \mu(d) \tag{29}$$

and

$$D = x^{\frac{1}{2}\epsilon}, \quad P(z) = \prod_{p < z} p.$$

By Fundamental Lemma type results (see, for example, [13, p. 71]) we obtain

$$|\rho_2(n)| \leq \sum_{\substack{d|n \\ d|P(z) \\ D \leq d \leq Dz}} 1 = \rho_3(n), \quad \text{say,}$$

and

$$\sum_{q \sim Q} \rho_3(q) \ll Q \exp\left(-\frac{\epsilon}{10}(\log x)^{\frac{1}{2}}(\log \log x)\right). \tag{30}$$

Clearly

$$\sum_{\substack{q \sim Q \\ q \text{ prime}}} \Delta(q) \leq \sum_q \rho(q)\Delta(q) = S(Q), \quad \text{say.}$$

The first move in [5] is an application of Cauchy's inequality to yield (see (4.3) there)

$$S^2(Q) \leq M^2 Q S(M, N; Q),$$

with

$$S(M, N; Q) = \sum_{m \sim M} \sum_{(q, tm)=1} \rho(q) \left\{ \sum_{n \equiv t\bar{m} \pmod{q}} \beta_n - \frac{1}{\phi(q)} \sum_{(n, q)=1} \beta_n \right\}^2.$$

The authors then introduce a smooth weight function $f(m)$ to obtain an upper bound to $S(M, N; Q)$. At this stage we replace $\rho(q)$ by $\rho_1(q)$. This introduces an additional error which is of a satisfactory size in view of (30). We now follow [5] in squaring out the resulting expression to obtain $S_1 - 2S_2 + S_3$ (see [5, (4.8)]). The required estimate is then established by proving that all the S_j are equal up to an acceptable error $O(E)$. In our case we need to take

$$E = \frac{xN(\log x)^A}{Qz^{\frac{1}{2}}}.$$

Here N is of magnitude KL and so satisfies $Qx^\epsilon \leq N \leq x^{\frac{3}{5}}$ by the hypotheses of Lemma 2.3.

In Section 5 of [5] S_3 is evaluated. In our modified situation their argument gives

$$S_3 = \hat{f}(0)X + O(E),$$

with

$$X = \sum_{(q, t)=1} \rho_1(q) \frac{1}{q\phi(q)} \left(\sum_{(n, q)=1} \beta_n \right)^2.$$

In Section 6 of [5] they evaluate S_2 via a well-known estimate for the incomplete Kloosterman sum. In their argument q is the variable in the Kloosterman sum so they require it to run over consecutive integers with no weight attached. In our context

$$\begin{aligned} S_2 &= \sum_{n_1} \sum_{n_2} \beta_{n_1} \beta_{n_2} \sum_{(q, tn_1 n_2)=1} \frac{\rho_1(q)}{\phi(q)} \sum_{m \equiv t\bar{n}_1 \pmod{q}} f(m) \\ &= \sum_{\substack{d|P(z) \\ d \leq D}} \mu(d) \sum_{(n_1, d)=1} \sum_{(n_2, d)=1} \beta_{n_1} \beta_{n_2} \sum_{\substack{qd \sim Q \\ (q, tn_1 n_2)=1}} \frac{1}{\phi(qd)} \sum_{m \equiv t\bar{n}_1 \pmod{qd}} f(m). \end{aligned}$$

The rest of the argument in that section of [5] can be executed in a satisfactory manner since the addition of the variable d into the argument makes little difference as there is plenty of headroom at this point (in their work they only assume $N \leq x^{\frac{2}{3}-\epsilon}$ at this stage).

Sections 7–10 of [5] deal with S_1 . In Section 7 our hypothesis (15) is required to dismiss the case $(n_1, n_2) > 1$ in a satisfactory manner (see the first paragraph on page 375). Nothing else in this section is altered for our proof: the proof has been reduced to estimating two terms \mathcal{X} and \mathcal{R}_1 .

Section 8 is the next part of the argument that requires some work to be adapted to our situation. We have, from (7.2) of [5], that

$$\mathcal{X} = \sum_{(q,t)=1} \frac{\rho_1(q)}{q} \sum_{\substack{(n_1, n_2)=1 \\ n_1 \equiv n_2 \pmod{q} \\ (n_1 n_2, q)=1}} \beta_{n_1} \beta_{n_2}.$$

We now replace $\rho_1(q)$ by $\rho(q)$ at the expense of an acceptably small error using (30) again. In this way we can show that $\mathcal{X} = X$ plus a suitable error from another switch from ρ to ρ_1 plus an error \mathcal{X}_1 corresponding to the term with the same nomenclature in [5, §8]. We note that the estimation of \mathcal{X}_1 in [5] is the only point where the Siegel-Walfisz hypothesis is used, and this is to furnish an analogue of the Barban-Davenport-Halberstam theorem. However, the Siegel-Walfisz hypothesis is only needed for those q with small divisors, and so we can take the method of [6, Chapter 9] for q without small divisors (and this is where we need $q \in \mathcal{Q}$) to obtain the required bound. To be precise, we follow the argument of [6, Chapter 9] down to the choice of Q_1 on page 171, where we can pick $Q_1 = z$. We have switched back to $\rho(q)$ to ensure that no q counted has a smaller divisor than z , and so the final part of the argument of [6, Chapter 9] is then redundant. By this stage of the argument Cauchy's inequality has been applied twice, which is why our saving is $z^{\frac{1}{4}}$ and not z .

Finally, in section 9 of [5] where \mathcal{R}_1 is discussed, we make use of (29) much as we did for the section 6 argument. There is little headroom in this part of the argument, and that is why we have taken D to be $x^{\epsilon/2}$. The extra summation over d introduced in this way does not then greatly perturb the argument which can subsequently be pushed to a satisfactory conclusion.

References

- [1] R.C. Baker and G. Harman, *The Brun-Titchmarsh Theorem on Average*, Progress in Mathematics 138, Birkhäuser, Boston, 1996.
- [2] R.C. Baker and G. Harman, *Shifted primes without large prime factors*, Acta Arith. **83** (1998), 331–361.
- [3] R.C. Baker, G. Harman, and J. Pintz, *The difference between consecutive primes II*, Proc. London Math. Soc. (3) **83** (2001), 532–562.
- [4] A. Beurling, *Analyse de la loi asymptotique de la distribution des nombres premiers généralisés*, Acta Math. **68** (1937) 255–291.
- [5] E. Bombieri, J.B. Friedlander, and H. Iwaniec, *Primes in Arithmetic Progressions to Large Moduli. II*, Math. Ann. **227** (1987), 361–393.
- [6] H. Davenport. *Multiplicative Number Theory, 2nd ed.* (revised by H. L. Montgomery), Springer, New York, 1980.

- [7] H. Delange, Sur la distribution des entiers ayant certaines propriétés, *Ann. Sci. Ecole Norm. Sup.* (3) **73** (1956), 15–74.
- [8] C. Elsholtz, *Additive decomposability of multiplicatively defined sets*, *Funct. Approx.* **35** (2006), 61–77.
- [9] C. Elsholtz, *Multiplicative decomposability of additively shifted primes*, to appear.
- [10] G. Greaves, *Sieves in Number Theory*, Springer-Verlag, Berlin 2001.
- [11] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.
- [12] G. Harman, *On the greatest prime factor of $p - 1$ with effective constants*, *Math. Comp.* **74** (2005), 2035–2041.
- [13] G. Harman, *Prime-Detecting Sieves*, LMS Monographs (New Series) 33, Princeton University Press, 2007.
- [14] A. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, *Journal de Théorie des Nombres de Bordeaux*, **5** (1993), 411–484.
- [15] M.N. Huxley, *On the difference between consecutive primes*, *Invent. Math.* **15** (1972), 164–170.
- [16] H. Iwaniec, *Primes of the type $\phi(x, y) + A$ where ϕ is a quadratic form*, *Acta Arithmetica* **21** (1972), 203–234.
- [17] H. Iwaniec, *Rosser's Sieve*, *Acta Arithmetica* **36** (1980), 171–202.
- [18] K. Matomäki, *Prime numbers of the form $m^2 + n^2 + 1$ in short intervals*, *Acta Arithmetica* **128** (2007), 193–200.
- [19] H.L. Montgomery and R.C. Vaughan, *The large sieve*, *Mathematika* **20** (1973), 119–134.
- [20] H.L. Montgomery and R.C. Vaughan, *Hilbert's Inequality*, *J. London Math. Soc.* (2), **8** (1974), 73 – 82.
- [21] A.G. Postnikov, *Introduction to Analytic Number Theory*, Translations of AMS Monographs **68**, American Mathematical Society, Providence, Rhode Island, 1988.
- [22] C.L. Stewart, *On the greatest prime factors of integers of the form $ab + 1$* , *Period. Math. Hungar.* **43** (2001), 81–91.
- [23] G. Tenenbaum, *Cribler les entiers sans grand facteur premier*, *Philos. Trans. Roy. Soc. London Ser. A* **345** (1993), 377–384.
- [24] E.C. Titchmarsh, *The Theory of the Riemann Zeta-function*, (revised D.R. Heath-Brown), Oxford 1986.
- [25] R.C. Vaughan, *An elementary method in prime number theory*, *Acta Arith.* **37** (1980), 111–115.
- [26] E. Wirsing, *Über die Zahlen, deren Primteiler einer gegebenen Menge angehören*, *Arch. Math.* **7** (1956), 263–272.

Address: Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX UK

E-mail: g.harman@rhul.ac.uk; K.S.Matomaki@rhul.ac.uk

Received: 31 July 2007; **revised:** 4 September 2007