PRIMES IN BEATTY SEQUENCES IN SHORT INTERVALS

GLYN HARMAN

December 3, 2015

Abstract

In this paper we show that sieve methods used previously to investigate primes in short intervals and corresponding Goldbach type problems can be modified to obtain results on primes in Beatty sequences in short intervals.

1 Introduction

Let $[\cdot]$ denote the integer part function. The sequence $[\xi n + \eta]$ is called a *Beatty sequence*. In some sense it is a generalisation of an arithmetic progression, and for integer values of ξ that is precisely what it is. There has been a flurry of interest shown recently in prime values of Beatty sequences, for example see [6, 13, 16]. We write $\pi(x; \xi, \eta)$ for the number of primes of the form $[\xi n + \eta] \leq x$. In [7] the following result is proved.

Theorem 1. *Let an irrational $\xi > 1$ be given, and for $x > 1$ write $y = x^\theta$ with $\theta > 5/8$. Then*

$$\pi(x + y; \xi, \eta) - \pi(x; \xi, \eta) = \frac{y}{\xi \log x} (1 + o(1)) \tag{1}$$

as $x \to \infty$.

The authors note that if ξ is rational the question collapses to the well-known problem of primes in an arithmetic progression. To discuss this case, suppose that $\xi = q/a$ as a reduced fraction and let $N(\xi, \eta)$ be the number of solutions in m to

$$0 < m + a(1 - \eta) \leq a, \quad (m, q) = 1. \tag{2}$$

Write $\delta(\xi, \eta) = N(\xi, \eta)/\phi(q)$. In the rational case it is necessary to have $\delta(\xi, \eta) > 0$ (see §2 below, the primes p counted satisfy $pa \equiv m \bmod q$). This number effectively replaces the $1/\xi$ factor in (1) for the rational case. We note that if a and q are both 'large' then $\delta(\xi, \eta) = (1+o(1))\xi^{-1}$. However, for a 'small' in terms of q we can easily get $\delta(\xi, \eta) = 0$ (the case $a = 1, (q, [\eta]) > 1$ is of course trivial). For example, if $q = 15, a = 2, \eta = 3$ then $[\xi n + \eta]$ is alternately divisible by 3 and 5. More generally, by the Chinese Remainder Theorem we can have $\delta(\xi, \eta) = 0$ for any a by choosing q, η suitably. In the rational case the work of Huxley and Iwaniec [11] establishes (1) (with ξ^{-1} replaced by $\delta(\xi, \eta)$) for $\theta > 7/12$. The exponent $7/12$ is the well-known limit for currently known results on primes in short intervals when trying to establish an asymptotic formula. In recent years the size of y required to ensure that the interval $[x, x + y)$ contains a prime has been reduced using sieve methods, culminating in the result from [4] that $y > x^{21/40}$ suffices. It is the purpose of this paper to show that the techniques developed for this and related results enable us to replace (1) by a lower bound of the correct order of magnitude for smaller values of y. We here prove the following.

Theorem 2. *If $\xi > 1$ is irrational and $y = x^\theta$ with $\theta > 5/9$ then*

$$\pi(x + y; \xi, \eta) - \pi(x; \xi, \eta) > \frac{y}{10\xi \log x} (1 + o(1)) \tag{3}$$

as $x \to \infty$. In the case ξ is the rational $\frac{q}{a}$ and $\delta(\xi, \eta) > 0$ this becomes

$$\pi(x + y; \xi, \eta) - \pi(x; \xi, \eta) > \frac{99y}{100 \log x} (1 + o(1)) \, \delta(\xi, \eta) \, . \tag{4}$$

In the ξ rational case one can reduce y further to $x^{21/40}$ with the constant $99/100$ weakened to $1/100$. The reader should note that our exponent $5/9$ is smaller than the $7/12$ 'classical' exponent for obtaining an asymptotic formula for the number of primes in a short interval, but larger than the exponents obtained for primes in short intervals since the work of Heath-Brown and Iwaniec [10]. The reason for this will become clear when we see how little Type I and Type II information is available for the exponential sums we need to use. The techniques we require to prove our new result are essentially all contained in [2] and [3], although we will require some adaptation especially in the final case of the proof (see Lemma 7, without which our results would have been somewhat weaker). We shall set up the proof in a simpler way than done in [7] (more like the earlier work of Banks and Shparlinski [6]) to reduce our task to applying a sieve method in tandem with estimates for double exponential sums over short intervals and Dirichlet polynomial techniques.

Using a standard notation we say ξ is of *finite type* if there is some $A >$ such that for all positive integers n we have

$$||n\xi|| \gg n^{-A}. \tag{5}$$

Here $||x||$ denotes the distance from x to a nearest integer. We remark that if ξ is of finite type then the implied constant in (3) is effectively computable. Otherwise there is an implicit appeal to Siegel's theorem in the proof of our result in the case $q < (\log x)^{40}$ later (this case does not arise when ξ is of finite type). It is of some interest to consider if Theorem 1 can be improved for certain ξ. We shall give in the final section a brief proof of the following improvement of Theorem 1 when ξ is of finite type.

Theorem 3. *If $\xi > 1$ and is irrational of finite type and $y = x^\theta$ with $\theta > 3/5$ then*

$$\pi(x + y; \xi, \eta) - \pi(x; \xi, \eta) = \frac{y}{\xi \log x}(1 + o(1)) \tag{6}$$

as $x \to \infty$.

We note further recent work on Beatty primes in short intervals, namely that R.C. Baker and L. Zhao [5] have shown that there are infinitely many bounded gaps between primes in Beatty sequences by adapting Maynard's method [14] to this situation. Also, the present author [9] has studied necessary and sufficient conditions for the intersection of two Beatty sequences to contain infinitely many primes. Combining the techniques in [9] with the methods we shall explore here would give results on primes in short intervals which lie simultaneously in two Beatty sequences satisfying certain necessary compatibility conditions. The exponent $\frac{5}{9}$ would have to be increased to $\frac{4}{7}$ for certain cases, however.

The author thanks the referee for their careful reading of the paper and subsequent comments.

2 Method outline

As previous authors have observed, $p = [n\xi + \eta]$ is equivalent to

$$0 < \{p\alpha + \beta\} \le \alpha, \tag{7}$$

where

$$\alpha = 1/\xi, \quad \beta = (1-\eta)/\xi,$$

and $\{x\}$ represents the fractional part of x. We thus wish to count solutions to (7) with $x < p \le x + y$. We suppose that x is sufficiently large and the proof takes three different directions depending on Diophantine approximations to α. Henceforth we suppose that $y = x^{\frac{5}{9}+\epsilon}$ with ϵ "small". Write $\Xi = x^{0.551}, \nu = \Xi^{-1}$. The crucial point is that $\frac{11}{20} < 0.551 < \frac{5}{9}$. By Dirichlet's theorem in Diophantine approximation there is a reduced fraction a/q such that $\alpha = a/q + \gamma$ with

$$q < y^2 x^{-1} (\log x)^{-40}, \quad |q\gamma| < xy^{-2}(\log x)^{40}.$$

If

$$q < (\log x)^{40} \quad \text{and} \quad |q\gamma| < \nu \tag{8}$$

(and this case includes the rational case for all sufficiently large x) we can split the interval $[x, x+y]$ into $y\nu + O(1)$ intervals $\mathcal{I}_j$ of length Ξ on each of which the set $\mathcal{M}_j$ of integers m with

$$0 < p\gamma q + \beta q + m \le \alpha q, \quad p \in \mathcal{I}_j, \quad (m, q) = 1,$$

is unchanged. We then need to solve $pa \equiv m \pmod q$, $p \in \mathcal{I}_j, m \in \mathcal{M}_j$ for each j (in the rational case $\gamma = 0$ and the above is just (2) rewritten). For each m coprime to q we obtain $> 0.99\Xi/(\phi(q)\log x)$ solutions by [3, Theorem 3]. If α is irrational then $q \to \infty$ and $|\mathcal{M}_j| = (1 + o(1))\alpha\phi(q)$. This gives

$$|\{p \in \mathcal{I}_j : p = [n\xi + \eta]\}| \ge \frac{99\alpha\Xi}{100\log x}(1 + o(1)).$$

Hence (3) follows in this case. If ξ is rational then for all large x there is just one set $\mathcal{M}_1$ with $|\mathcal{M}_1| = N(\xi, \eta)$ and so we obtain (4) for this case.

Now suppose that (8) does not hold. Then solving (7) can be quickly transformed into a problem of bounding exponential sums over primes in short intervals by well-known methods. The method for bounding the sums will depend on whether $q \geq (\log x)^{40}$ or $q < (\log x)^{40}$ and $|\gamma q| > \nu$. In the former case we can use elementary methods to bound the sums (see Lemmas 2 and 3 below). These correspond to the *minor arcs* in the Circle Method. In the latter case we use *major arc* techniques. The reader should note though that we never deal with the two cases at once which was necessary in [2] and also arises in [7].

We write, as is customary, $e(x) = \exp(2\pi i x)$. Before proceeding further we must introduce our sieve method. We wish to count primes by using a function, say $\rho(k)$, which takes the value 1 at primes and takes a non-positive integer value at composite numbers. This function must satisfy certain properties depending on which case we are considering. For example, when $q \geq (\log x)^{40}$, we must have

(i) $\sum \rho(k) e(\gamma k)$ can be decomposed into multiple sums that can be estimated by known techniques for exponential sums;

(ii) $\sum \rho(k)$ can be decomposed into multiple sums that can be estimated by Dirichlet polynomial techniques leading to a positive lower bound for the sum.

If we were content to settle for the larger exponent 4/7 the required function is essentially given as $A_0(k)$ by the first display on [2, p.778] and would lead to the 1/10 in the lower bound being increased to 1/4. The value we obtain here improves on 4/7 for two reasons. One: we only need to sieve one variable whereas in [2] two variables are sieved using a vector sieve. Two: we do not have to have the same decomposition for both cases as was required in [2]. If we were to combine our present techniques with the method in [9] this second assertion might not always hold.

We quote [8, Chapter 3] to introduce the function

$$\rho(n, z) = \begin{cases} 1 & \text{if } n \in \mathbb{N},\ p|n \Rightarrow p \geq z, \\ 0 & \text{otherwise.} \end{cases}$$

As there, it will be important that $\rho(n, z) = 0$ if $n \notin \mathbb{N}$. We note that

$$\rho(n, z) = \sum_{\substack{d|n \\ d|P(z)}} \mu(d), \quad \text{where} \quad P(z) = \prod_{p<z} p\,,$$

and Buchstab's identity gives, for $1 \le w < z$,

$$\rho(n, z) = \rho(n, w) - \sum_{w \le p < z} \rho(n/p, p). \tag{9}$$

Also, $\rho(n, z(n))$ is the characteristic function of the set of primes if $n^{1/2} < z(n) \le n$. It will be our goal to decompose $\rho(n, X)$ where $X = \sqrt{2x}$ into multiple sums using Buchstab's identity and discard certain non-negative sums which cannot be evaluated with present knowledge in order to form our desired lower bound function $\rho(n)$.

3 Exponential sums

The following result converts the problem of counting fractional parts in an interval to the estimation of sums of exponential sums.

Lemma 1. *Suppose* $0 < \alpha < 1, L\alpha > 1$. *Let* θ_n *and* a_n *be sequences of real numbers. Then*

$$\sum_{\substack{n \le N \\ 0 < \{\theta_n\} \le \alpha}} a_n = \alpha \sum_{n \le N} a_n + O(E) \tag{10}$$

where

$$E = \frac{1}{L} \sum_{n \le N} |a_n| + O\left(\sum_{1 \le \ell \le L} \min(\alpha, \ell^{-1}) \left| \sum_{n \le N} a_n e(\ell \theta_n) \right| \right).$$

Proof. See [1, pp. 18-21] □

We take $\alpha = \xi^{-1}$ as above and put $L = (\log x)^{10}$. Henceforth we write $I = [x, x + y]$. Also, we use $s \sim S$ to denote $S \le s < 2S$. We then have the following results on multiple sums over short intervals. These results are given in [2] but with less precision for the $\log x$ terms and without the additional summation over ℓ.

Lemma 2 (Type I information)**.** *Suppose that*

$$|q\alpha - a| < q^{-1} \quad \text{with} \quad (a, q) = 1. \tag{11}$$

Then, for any $M \ge 1$, $L \ge 1$, *we have*

$$\sum_{\ell \le L} \sum_{s \le M} \left| \sum_{st \in I} e(\alpha \ell st) \right| \ll L(\log x) \left(\frac{y}{q} + LM + q \right). \tag{12}$$

Proof. This follows from Lemma 2.2 of [15] with the trivial bound L for the number of ways an integer h can be represented as $\ell s, \ell \le L$. □

Corollary. *In the notation of Lemmas 1 and 2 we have*

$$\sum_{1\le \ell\le L} \min(\alpha, \ell^{-1}) \sum_{s\le M} \left| \sum_{st\in I} e(\alpha \ell st) \right| \ll (\log\log x)(\log x) \left(\frac{y}{q} + LM + q \right). \tag{13}$$

Lemma 3 (Type II information)**.** *Suppose that* (11) *holds and* a_s, b_t *are sequences such that, for any* M*,*

$$\sum_{s\le M} |a_s|^2 \ll M(\log M)^A, \quad \sum_{t\le M} |b_t|^2 \ll M(\log M)^B. \tag{14}$$

Then

$$\sum_{\ell\le L} \sum_{s\sim M} \sum_{st\in I} a_s b_t e(\alpha \ell st) \ll (\log x)^{1+(A+B)/2} y L \theta, \tag{15}$$

with

$$\theta^2 = \frac{ML}{y} + \frac{L}{q} + \frac{Lx}{yM} + \frac{qx}{y^2}.$$

Proof. This follows from the Cauchy-Schwarz inequality together with [15, Lemma 2.2]. See [2, p.767] for details. □

Corollary. *In the notation of Lemmas 1 and 3 we have*

$$\sum_{1\le \ell\le L} \min(\alpha, \ell^{-1}) \left| \sum_{s\sim M} \sum_{st\in I} a_s b_t e(\alpha st) \right| \ll (\log\log x)(\log x)^{1+(A+B)/2} y\theta. \tag{16}$$

Now to obtain our results we only need a small saving on the trivial bound for the exponential sums, that is we require an upper bound $\ll y(\log x)^{-C}$ for a suitably large C. For Lemma 2 Corollary this reduces to

$$M \ll y(\log x)^{-C-12}, \quad (\log x)^{C+2} \ll q \ll y(\log x)^{-C-2}.$$

For Lemma 3 Corollary this reduces to

$$(x/y)(\log x)^D \ll M \ll y(\log x)^{-D}, \quad (\log x)^D \ll q \ll y^2 x^{-1}(\log x)^{-D},$$

where $D = 23 + 2C + A + B$. We should remark first that these results are near to best possible since we cannot get cancellation in a sum with $\ll 1$ terms on average. Second, we comment that the range for q is quite restricted in Lemma 3. Third, we note that the Type I information is poorer than usual in problems reducing sums over primes to double sums. This is why Vaughan's identity cannot give a value of y less than $x^{2/3}$, and we run into problems with sieve methods as y reduces below $x^{5/9}$. The method of Zhan [17] can work with y reduced to $x^{3/5}$ for q in a certain range, and $x^{5/8}$ in another by using Heath-Brown's generalised Vaughan identity in tandem with bounds for exponential sums with two Type I ranges whose product is sufficiently large. Appropriate mean-value results for Dirichlet polynomials are then needed to complete the proof. The results we have given above only suffice to prove our main theorem in the case $q \geq (\log x)^{40}$. This would suffice to prove the theorem when ξ is of finite type. We too shall need estimates that depend on bounds for Dirichlet polynomials. These are found in [3] and we shall use them as in [2]. In the following $\chi(n)$ denotes a Dirichlet character to the modulus q where $q \leq (\log x)^{40}$. We also set $T_0 = \exp((\log x)^{1/3}))$.

Lemma 4 (Analytic information)**.** *Suppose that $b(k)$ is a real sequence such that for any $A > 0$*

$$\int_{T'}^{T'+T} \left| \sum_{k \leq x} \frac{b(k)\chi(k)}{k^{1/2+it}} \right| dt \ll Tyx^{-1/2}(\log x)^{-A} \tag{17}$$

whenever

$$T \in [xy^{-1}, x], \; T_0 \leq T' \ll T^2, \; T' + T \leq x.$$

Then if for some $A > 0$ we have

$$\alpha = \frac{a}{q} + \gamma, \quad (a, q) = 1, \quad |\gamma| < y^{-2}x(\log x)^A, \quad q \leq (\log x)^A$$

we have

$$\sum_{n \in I} b(n)e(n\alpha) = \frac{\mu(q)}{y\phi(q)} \sum_{k \in I} e(k\gamma) \sum_{m \in I} b(m)\chi_0(m) + O\left(y(\log x)^{-A}\right).$$

In particular, if

$$\sum_{m \in I} b(m) \ll \frac{y}{\log x},$$

then
$$\left|\sum_{n\in I} b(n)e(n\alpha)\right| \ll |q\gamma|^{-1} + y(\log x)^{-A}\,. \tag{18}$$

Proof. This is established in [2, pp. 767-771]. □

We note that in the case $q < (\log x)^{40}, |q\gamma| > \nu$ we can take $A = 50$ to turn (18) into the bound

$$\sum_{1\le \ell\le L} \min(\alpha, \ell^{-1}) \left|\sum_{n\in I} b(n)e(\ell n\alpha)\right| \ll y(\log x)^{-49}\,. \tag{19}$$

Here we have applied (18) with α substituted with $\alpha\ell$ for each ℓ, replacing a, q, γ by $b = b(\ell), r = r(\ell), \gamma' = \gamma'(\ell)$ respectively, and so

$$\ell\alpha = \frac{b}{r} + \gamma', \quad r \le (\log x)^{40} < (\log x)^{50}, \quad \nu q^{-1} < |\gamma'| < y^{-2}x(\log x)^{50}\,.$$

4 The case $q \ge (\log x)^{40}$

For convenience we work with the standard sifting function defined for any set of integers $\mathcal{E}$ by
$$S(\mathcal{E}, z) = \sum_{n\in\mathcal{E}} \rho(n, z)\,.$$
We also put
$$\mathcal{E}_n = \{m : mn \in \mathcal{E}\}\,.$$
With this notation, Buchstab's identity becomes
$$S(\mathcal{E}, z) = S(\mathcal{E}, w) - \sum_{w\le p<z} S(\mathcal{E}_p, p)\,.$$

We shall take $\mathcal{E} = \mathcal{A}, \mathcal{B}$ or $\mathcal{C}$ where

$$\mathcal{B} = \{n : x \le n \le x + y\}, \quad \mathcal{A} = \{n \in \mathcal{B} : 0 < \{n\alpha + \beta\} \le \alpha\},$$
$$\mathcal{C} = \{n : x \le n \le x + x\delta\},$$
with
$$\delta = \exp\left(-(\log x)^{\frac{1}{3}}\right).$$

We note that by the Prime Number Theorem we have

$$S\left(\mathcal{C}, x^{1/u}\right) = \frac{\omega(u)\delta x}{\log x}(1 + o(1)),$$

where $\omega(u)$ is Buchstab's function (see [8, pp 339-342]). We require the following result which can be found as [8, Theorem 3.1].

Lemma 5. *Suppose that for any sequences of complex numbers a_m, b_n that satisfy $|a_m| \leq 1, |b_n| \leq 1$ we have, for some $\lambda > 0$, $\mu > 0$, $\kappa \leq \frac{1}{2}$, $M \geq 1$, that*

$$\sum_{\substack{mn\in\mathcal{A} \\ m\leq M}} a_m = \lambda \sum_{\substack{mn\in\mathcal{B} \\ m\leq M}} a_m + O(Y) \tag{20}$$

and

$$\sum_{\substack{mn\in\mathcal{A} \\ x^{\mu}\leq m\leq x^{\mu+\kappa}}} a_m b_n = \lambda \sum_{\substack{mn\in\mathcal{B} \\ x^{\mu}\leq m\leq x^{\mu+\kappa}}} a_m b_n + O(Y). \tag{21}$$

Let c_r be a sequence of complex numbers such that

$$|c_r| \leq 1, \quad \textit{and if} \;\; c_r \neq 0, \quad \textit{then} \;\; p|r \Rightarrow p > x^{\epsilon}, \tag{22}$$

for some $\epsilon > 0$. Then, if $x^{\mu} < M$, $2R < \min(x^{1-\mu}, M)$, and $M > x^{1-\mu}$ if $2R > x^{\mu+\kappa}$, we have

$$\sum_{r\sim R} c_r S(\mathcal{A}_r, x^{\kappa}) = \lambda \sum_{r\sim R} c_r S(\mathcal{B}_r, x^{\kappa}) + O(Y\log^3 x). \tag{23}$$

Now by (13) and (16) we can apply Lemma 5 with $\mu = \frac{4}{9}, \kappa = \frac{1}{9}, M = x^{\frac{5}{9}}, \lambda = \alpha$. Let $z = x^{1/9}, z' = x^{5/36}, z^*(P) = x^{5/9}/P$ for $x^{\frac{1}{3}} \leq P < x^{4/9}$. We can thus obtain formulae for

$$\sum_{r\sim R} a_r S(\mathcal{A}_r, z), \;\; R < x^{\frac{5}{9}}, \quad \text{and for} \;\; S(\mathcal{A}, z),$$

immediately from Lemma 5. For example, we have

$$S(\mathcal{A}, z) = \alpha S(\mathcal{B}, z)(1 + o(1)) + O\left(y(\log x)^{-7}\right).$$

We can extend this to give a formula for $S(\mathcal{A}, z')$ using Buchstab's identity applied four times as this will lead to sums

$$S(\mathcal{A}, z), \quad \sum_{p_1} S(\mathcal{A}_{p_1}, z), \quad \ldots, \sum_{p_1,p_2,p_3,p_4} S(\mathcal{A}_{p_1p_2p_3p_4}, p_4).$$

The first terms can all be evaluated by Lemma 5, while the final sum can be estimated immediately from (16) since

$$x^{4/9} \le p_1 p_2 p_3 p_4 \le z'^4 = x^{5/9}.$$

We can also evaluate, for $x^{\frac{1}{3}} \le P \le x^{\frac{4}{9}}$,

$$\sum_{p \sim P} S(\mathcal{A}_p, z^*(P)) = \sum_{p \sim P} S(\mathcal{A}, z) - \sum_{\substack{p \sim P \\ z \le q < z^*(P)}} S(\mathcal{A}_{pq}, q),$$

as the final sum has $x^{\frac{4}{9}} \le pq \le x^{\frac{5}{9}}$. Similarly we can deal with

$$\sum_{p \sim P} S(\mathcal{A}_p, z^*(P)) \quad \text{where} \quad z^*(P) = \begin{cases} (x^{5/9}/P)^{\frac{1}{2}} \text{ if } x^{2/9} < P < x^{1/3} \\ (x^{5/9}/P)^{\frac{1}{3}} \text{ if } x^{1/9} < P < x^{2/9}. \end{cases}$$

Henceforth we take $P = 2^k, k \in \mathbb{N}$ so that in an expression involving p we define P by $P \le p < 2P$. We also put $z^*(P) = P$ for $P > x^{\frac{4}{9}}$. We can thus apply Buchstab's identity twice to obtain

$$S(\mathcal{A}, X) = S(\mathcal{A}, z') - \sum_{z' \le p < X} S(\mathcal{A}_p, z^*(P)) + \sum_{\substack{z' \le p < x^{\frac{4}{9}} \\ z^*(P) \le q \le \min(p, (x/p)^{\frac{1}{2}})}} S(\mathcal{A}_{pq}, q)\,. \quad (24)$$

In the final sum above we can decompose twice more if $pq^2 < x^{5/9}$. Also we can give a formula immediately from (16) when $x^{4/9} \le pq \le x^{5/9}$.

Now we must pause and take stock at this point for we have only given a formula for sums involving $\mathcal{A}$ in terms of sums involving $\mathcal{B}$. Since $y < x^{7/12}$ we may not have an asymptotic formula for all the sums involving $\mathcal{B}$. Certainly we do not have such a formula for $S(\mathcal{B}, X)$. However, we can use the detailed analysis given in [3] to produce asymptotic formulae for most of the resulting sums and we now show that this suffices. For example, we immediately have by the method expounded there that

$$S(\mathcal{B}, z) = y(x\delta)^{-1} S(\mathcal{C}, z)(1 + o(1))\,.$$

Rather than going through the working in [3] *de nouveau*, we shall take it as it stands (it is valid for $y > x^{11/20}$ and we have $y > x^{5/9}$). To illustrate what happens, we have so far shown that

$$\sum_{z' \le p < X} S(\mathcal{A}_p, z^*(P)) = \alpha(1 + o(1)) \sum_{z' \le p < X} S(\mathcal{B}_p, z^*(P)) + O\left(y(\log x)^{-7}\right)\,.$$

We can write, with $z_0 = x^{\frac{1}{10}}$,

$$\sum_{z' \le p < X} S(\mathcal{B}_p, z^*(P)) = \sum_{z' \le p < X} S(\mathcal{B}_p, z_0) - \sum_{\substack{z' \le p < X \\ z_0 \le q < \min(z^*(P),(x/p)^{\frac{1}{2}})}} S(\mathcal{B}_{pq}, q)$$
$$= \mathcal{S}_1 - \mathcal{S}_2 \ \text{ say.}$$

We then have

$$\mathcal{S}_1 = y(x\delta)^{-1}(1 + o(1)) \sum_{z' \le p < X} S(\mathcal{C}_p, z_0)\,.$$

We can then read off [3, Diagram 1] those values of p, q for which we are unable to give the required asymptotic formula for the corresponding parts of $\mathcal{S}_2$. We can do the same immediately for the final sum on the right hand side of (24). We thus arrive at

$$S(\mathcal{A}, X) \ge \alpha y(x\delta)^{-1} \left(S(\mathcal{C}, X)(1 + o(1)) - Ex\delta(\log x)^{-1} \right),$$

where E corresponds to the expected size of the sums we have had to discard because we either could not obtain the required formula for $\mathcal{A}$ in terms of $\mathcal{B}$ or of $\mathcal{B}$ in terms of $\mathcal{C}$. By standard methods we can turn the sums into integrals involving Buchtab's function $\omega(u)$ (see [8, pp 15-16]). We give a couple of examples of the components of E by way of illustration before stating the contributions from each region. The numerical calculations performed here were done twice: once using BASIC on a modern 64-bit personal computer and once using Mathematica on a Raspberry Pi. The calculations agreed to more decimal places than we shall quote here.

We consider the part of the double sum with $pq^2 < x^{5/9}$ where it is possible to decompose twice more. This contributes to E

$$\int_w \int_x \int_y \int_z \frac{\omega((1 - w - x - y - z)/z)}{wxyz^2} \, dz \, dy \, dx \, dw\,.$$

Here the region of integration is defined by

$$\frac{5}{36} \le w \le \frac{2}{9}, \ \ \frac{1}{3}\left(\frac{5}{9} - w\right) \le x \le \min\left(w, \frac{1}{2}\left(\frac{5}{9} - w\right)\right), \ \ \frac{1}{9} \le z \le y \le x\,,$$

excluding the region with

$$w + 4x \ge 0.82, \quad w + x \le 0.36$$

(for this corresponds to the region Δ_2 on [3, Diagram 1]) and it is assumed that no combination of the variables lies between $\frac{4}{9}$ and $\frac{5}{9}$. Thus this is quite a small region of integration resulting in a contribution to E not exceeding 0.0019. Here and henceforth we replace $\omega(u)$ by an upper bound approximation given by

$$\omega(u) \begin{cases} = \ 1/u & \text{if } 1 \le u \le 2, \\ = \ (1+\log(u-1))/u & \text{if } 2 \le u \le 3, \\ \le \ \frac{1}{3}(1+\log 2) & \text{if } u \ge 3. \end{cases} \tag{25}$$

Now consider the contribution from the part with $x^{\frac{2}{9}} \le p \le x^{\frac{1}{3}}, pq < x^{\frac{4}{9}}$ or $pq > x^{\frac{5}{9}}$. For the part of the sum with $pq^3 < x$ we may use the Buchstab identity to write

$$\sum_{p,q} S(\mathcal{A}_{pq}, q) = \sum_{p,q} S(\mathcal{A}_{pq}, (x/pq)^{\frac{1}{2}}) + \sum_{\substack{p,q \\ q \le r < (x/pq)^{\frac{1}{2}}}} S(\mathcal{A}_{pqr}, r)\,.$$

Although we cannot evaluate the first sum on the right hand side above, we are be able to deal with part of the second sum. In fact this leads to a contribution < 0.249 in place of < 0.34.

The region with $x^{5/27} \le p \le x^{2/9}, pq^2 > x^{5/9}$ contributes < 0.077. The region with $x^{1/3} \le p \le x^{4/9}$ contributes < 0.522. The regions for which we cannot evaluate terms like $S(\mathcal{B}_{pq}, q)$ contribute < 0.01 from the working in [3]. We thus conclude that $E < 0.86$ which more than suffices to establish (3) in this case.

5 The case $q < (\log x)^{40}$

As we remarked in [2] it is the possibility that T' could be as large as T^2 that causes problems in using Lemma 4. If $T' \ll T$ then we could use the same type of working that leads to $y = x^{21/40}$ in [4]. Instead we must adapt the working in [2, 3] to intervals of length $[x, x + x^{5/9+\epsilon}]$ to obtain the required arithmetical information. We must assume that we have either a

$$\begin{aligned} &\text{Type I sum:} && b(k) = \sum_{\substack{mn=k \\ m\sim M}} a_m \\ &\text{Type I/II sum:} && b(k) = \sum_{\substack{mnr=k \\ m\sim M, n\sim N}} a_m b_n\,, \\ \text{or}\quad &\text{Type II* sum:} && b(k) = \sum_{\substack{mnr=k \\ m\sim M, n\sim N}} a_m b_n c_r\,. \end{aligned}$$

In the final sum the coefficient c_r is a convolution of the characteristic functions of primes on certain intervals. In all the sums we assume that $a_n, b_m \ll 1$.

Lemma 6. *We obtain* (17) *for a Type I sum with* $M \le x^{5/9}$.

Proof. This corresponds to [2, Lemma 3.6]. □

Lemma 7. *We obtain* (17) *for a Type I/II sum if*

$$M \le x^{13/27}, \qquad N \le x^{7/27}, \qquad MN \le x^{19/27}. \tag{26}$$

Proof. There is no corresponding result in [2]. Instead we must adapt [8, Lemma 7.4] or equivalently [4, Lemma 10]. Now we cannot use the reflection principle or the fourth power moment of an L-function as there. Instead we must use Jutila's result [12] that, for any $\epsilon > 0$,

$$\sum_{\chi \bmod q} \int_{T'}^{T'+T} \left| L\left(\tfrac{1}{2} + it, \chi\right)\right|^4 dt \ll \left(qT + qT'^{\frac{2}{3}}\right)(qT')^{\epsilon}\,.$$

Using this, Hölder's inequality, and the mean-value theorem for Dirichlet polynomials we quickly obtain that the left hand side of (17) is

$$\ll (M+T)^{\frac{1}{2}}(N^2+T)^{\frac{1}{4}}T^{\frac{1}{3}+\frac{1}{4}\epsilon} \ll Tyx^{-\frac{1}{2}}(\log x)^{-A}$$

as required from (26). □

Lemma 8. *We obtain* (17) *for a Type II* sum if*

$$x^{-\frac{1}{9}} \le M/N \le x^{\frac{1}{9}}, \qquad MN \ge x^{\frac{7}{9}}. \tag{27}$$

Proof. This corresponds to [2, Lemma 3.1] and [8, Lemma 7.3] with $g = 2$. □

Lemma 9. *We obtain* (17) *for a Type II* sum if*

$$\max(M, N) \le x^{\frac{7}{15}}, \qquad MN \ge x^{\frac{16}{21}}, \qquad \min(M, N) \ge x^{\frac{8}{27}}. \tag{28}$$

Proof. This corresponds to [2, Lemma 3.2] after replacing $T = x^{9/20}$ by $x^{4/9}$ in [3, Lemma 4]. □

We note that the above lemmas supply us with the required bounds on exponential sums when their hypotheses are satisfied via Lemma 4 and (19). We now use the above information to provide useful asymptotic formulae.

Lemma 10. *Suppose that either* (27) *or* (28) *holds. Then we have*

$$\sum_{\substack{m\sim M\\ n\sim N}} a_m b_n S(\mathcal{A}_{mn}, f(m,n)) = \alpha \frac{y(1+o(1))}{x\delta} \sum_{\substack{m\sim M\\ n\sim N}} a_m b_n S(\mathcal{C}_{mn}, f(m,n)) .$$

Here $f(m,n)$ is a smooth function (possibly constant) of m and/or n such as $f(m,n) = n$ or $f(m,n) = (x/mn)^{\frac{1}{2}}$.

Proof. This follows on combining Lemmas 1, 4, (19) and whichever of Lemmas 8 or 9 is appropriate. □

Lemma 11. *Suppose $M \le x^{\frac{1}{2}}$. Then*

$$\sum_{m\sim M} S(\mathcal{A}_m, z) = \alpha \frac{y(1+o(1))}{x\delta} \sum_{m\sim M} S(\mathcal{C}_m, z) .$$

Proof. This can be established by converting the problem to exponential sums as in the last section then working in the same way as [2, Lemma 3.8]. The reader should note that we do not need to go via $S(\mathcal{B}_m, z)$ (as in the last section) because the arithmetical information required to establish asymptotic formulae for $S(\mathcal{B}_m, z)$ includes that for obtaining $S(\mathcal{A}_m, z)$ (as we only need $T = T'$ for the $\mathcal{B}$ case). □

Lemma 12. *Suppose that $M \le x^{13/27}, x^{\frac{1}{9}} \le R \le x^{\frac{2}{9}}$. Then*

$$\sum_{\substack{m\sim M\\ r\sim R}} a_m c_r S(\mathcal{A}_{mr}, z) = \alpha \frac{y(1+o(1))}{x\delta} \sum_{\substack{m\sim M\\ r\sim r}} a_m c_r S(\mathcal{C}_{mr}, z) .$$

Proof. We work in a similar way to Lemma 11. First suppose that $M \geq x^{\frac{4}{9}}$. We can then write, with $w = \exp((\log x)^{\frac{9}{10}})$,

$$\sum_{\substack{m\sim M\\ r\sim R}} a_m c_r S(\mathcal{A}_{mr}, z) = \sum_{\substack{m\sim M\\ r\sim R}} a_m c_r S(\mathcal{A}_{mr}, w) - \sum_{\substack{m\sim M, r\sim R\\ w\leq p<z}} a_m c_r S(\mathcal{A}_{mrp}, p)\,. \quad (29)$$

The first sum on the right hand side above can be estimated using Lemma 7 using [8, Theorem 4.3] as for the similar term in [8, Lemma 7.5]. The second sum can be handled via Lemma 8 (with p here taking the role of r in that result).

Now define ρ by $R = x^{\rho}$. To apply Lemma 8, with r here now having the same meaning as in that lemma, we need a variable $m' = x^a$ with a in the range

$$\frac{1-\rho-\frac{1}{9}}{2} \leq a \leq \frac{1-\rho+\frac{1}{9}}{2}\,.$$

This is an interval of length $\frac{1}{9}$ whose left hand end-point is between $\frac{1}{3}$ and $\frac{7}{18}$. We can proceed using the basic idea of the author's sieve method to take out the prime factors of the implicit variable counted in the $S(\cdot,\cdot)$ notation by repeated use of Buchstab's identity; that is, we iterate the step (29). We can keep iterating so long as Lemma 7 is applicable to estimate $\sum S(\mathcal{A}_{mrs}, w)$ and we can stop when combining m with the new prime variables gives the variable m' required. Since each new prime variable is $\leq x^{\frac{1}{9}}$ we must eventually succeed (compare [2, Lemma 3.8]). $\square$

We are now in a position to perform our decomposition of $S(\mathcal{A}, X)$ into sums which we can either evaluate or discard. We write

$$S(\mathcal{A}, X) = S(\mathcal{A}, z) - \sum_{z\leq p<X} S(\mathcal{A}_p, z) + \sum_{\substack{z\leq q\leq \min(p,(x/p)^{\frac{1}{2}})\\ z\leq p<X}} S(\mathcal{A}_{pq}, q)\,.$$

We can evaluate the first two terms on the right hand side above using Lemma 11. Some parts of the final term above can be evaluated immediately from Lemma 10. If $pq^2 \leq x^{\frac{1}{2}}$ we can apply Buchstab twice more using Lemma 11. In the remaining sum we can decompose twice more using Lemma 12 if $pq \leq x^{13/27}, q \leq x^{\frac{2}{9}}$. Large sections of the resulting quadruple sums can be evaluated using Lemma 10 since there are so many possible combinations of variables. Even for the parts of the final double sum for which we have not been able to apply Buchstab again, we can use the device from the previous section and consider which almost-primes counted are amenable to detection by Lemma 10. Some calculations then give the following results for the proportion of the expected value we must discard:

Double sum discarded	< 0.78,
Quadruple sum discarded	< 0.005.

We thus conclude that

$$S(\mathcal{A}, X) > \alpha \frac{y(1+o(1))}{5x\delta} S(\mathcal{C}, X)\,,$$

which more than suffices to complete the proof.

6 Proof of Theorem 3

We note the following result proved by Zhan as [17, Theorem 2].

Lemma 13. *Assume that*

$$\alpha = \frac{a}{q} + \lambda, \quad (a, q) = 1, \quad |\lambda| \leq \frac{1}{q\tau}, \quad 1 \leq q \leq \tau.$$

Write $L = \log x$. *Then, for any* $B > 0$, *there exists* $c_j > 0, 1 \leq j \leq 3$ *such that if*

$$\tau = A^2 N^{-1} L^{-c_1}, \quad q \geq L^{c_2}, \quad \textit{and} \quad N^{\frac{3}{5}} L^{c_3} \leq A \leq N,$$

then we have

$$\sum_{N-A<n\leq N} \Lambda(n) e(n\alpha) \ll AL^{-B}\,.$$

We apply the above with $B = 20, A = y, N = x + y$. It quickly follows from this that for all large x we have

$$\sum_{\ell \leq L^{10}} \left| \sum_{x \leq p < x+y} e(\ell p \alpha) \right| \ll yL^{-10},$$

which establishes Theorem 3 by Lemma 1. To see this, we use Dirichlet's theorem in Diophantine Approximation to find a q_ℓ for each $1 \leq \ell \leq L^{10}$ with an associated a_ℓ such that

$$q_\ell \leq y^2 x^{-1} L^{-c_1}, \quad \ell\alpha = \frac{a_\ell}{q_\ell} + \lambda_\ell, \quad |\lambda_\ell| \leq \frac{xL^{c_1}}{q_\ell y^2}\,.$$

Since α is of finite type we have, for some $A \geq 1$,

$$\frac{xL^{c_1}}{y^2} \geq q_\ell |\lambda_\ell| = ||\ell q_\ell \alpha|| \gg (\ell q_\ell)^{-A},$$

and so

$$\ell q_\ell \gg \left(\frac{y^2}{xL^{c_1}} \right)^{\frac{1}{A}} \gg L^{c_2+10}$$

as required to give $q_\ell > L^{c_2}$.

References

[1] R. C. Baker, *Diophantine Inequalities*, LMS Monographs (New Series), vol. 1, Clarendon Press, Oxford, 1986.

[2] R. C. Baker and G. Harman, The three primes theorem with almost equal summands. *Phil. Trans. R. Soc. Lond.* A 356 (1998), 763–780.

[3] R. C. Baker, G. Harman, and J. Pintz, The exceptional set for Goldbach's problem in short intervals. In *Sieve Methods, Exponential Sums and Their Applications in Number Theory*, Cambridge University Press, Cambridge, 1997, pp. 11–54.

[4] R. C. Baker, G. Harman, and J. Pintz, The difference between consecutive primes II. *Proc. London Math. Soc.* (3) **83** (2001), 532–562.

[5] R. C. Baker and L. Zhao, Gaps between primes in Beatty sequences. http://arxiv.org/pdf/1411.2989.pdf

[6] W. D. Banks and I. Shparlinski, Prime numbers in Beatty sequences. *Colloq. Math.* 115 (2009), 147–157.

[7] J. Brüdern and K. Kawada, The localisation of primes in arithmetic progressions of irrational modulus. *Colloq. Math.* 123 (2011), 53–61.

[8] G. Harman, *Prime-detecting Sieves*, London Mathematical Society Monographs Series, **33**, Princeton University Press, Princeton, NJ, 2007.

[9] G. Harman, Primes in intersections of Beatty sequences. *Journal of Integer Sequences* **18** (2015), Article 15.7.3.

[10] D. R. Heath-Brown and H. Iwaniec, On the difference between consecutive primes. *Invent. Math.* **55** (1979), 49–69.

[11] M. N. Huxley and H. Iwaniec, Bombieri's theorem in short intervals. *Mathematika* **22** (1975), 188–194.

[12] M. Jutila, Mean value estimates for exponential sums with applications to L-functions. *Acta Arith.* **57** (1991), 93-114.

[13] G. Lü and H. Sun, The ternary Goldbach-Vinogradov theorem with almost equal primes from the Beatty sequence. *Ramanujan J.* **30** (2013), 153–161.

[14] J. Maynard, Small gaps between primes. *Ann. of Math.* (2) **181** (2015), no. 1, 383–413.

[15] R. C. Vaughan, *The Hardy-Littlewood Method*, Cambridge Tracts in Mathematics **80**, Cambridge University Press, Cambridge, second edition 1997.

[16] R. C. Vaughan, The general Goldbach problem with Beatty primes. *Ramanujan J.* **34** (2014) 347–359.

[17] Zhan Tao, On the representation of large odd integer as a sum of three almost equal primes. *Acta Math. Sinica* **7** (1991), 259–272.

Glyn Harman
Royal Holloway, University of London,
Egham, Surrey TW20 0EX, UK
E-mail: G.Harman@rhul.ac.uk