ILLINOIS JOURNAL OF MATHEMATICS
Volume 39, Number 1, Spring 1995

# SOME SINGULAR SERIES AVERAGES AND THE DISTRIBUTION OF GOLDBACH NUMBERS IN SHORT INTERVALS

J.B. FRIEDLANDER[1] AND D.A. GOLDSTON[2]

## 1. Introduction

The Goldbach conjecture asserts that every even integer exceeding two can be written as the sum of two primes. As this has still not been substantiated, there is reason to distinguish those even integers which can be written as the sum of two primes; we call such an integer a Goldbach number. There are many results that have been proven about Goldbach numbers (see, for example, the introduction in [Go$_2$]).

In this paper we shall be concerned with the question of the existence of Goldbach numbers in short intervals and the asymptotic formula for the number of representations of the even integers in a short interval as the sum of two primes.

It was proven by Montgomery and Vaughan [MV$_2$] that every interval $(N-K, N]$ contains Goldbach numbers provided that $K > N^{7/72+\varepsilon}$ and $N > N_0(\varepsilon)$. More recently Perelli and Pintz [PP] have proven that almost every even integer in the interval $(N-K, N]$ is a Goldbach number if $K > N^{7/36+\varepsilon}$.

In the case that one admits conditional results it is possible to treat significantly shorter intervals and there has been a history of results based on certain unproved hypotheses. The first such result (which preceded the unconditional results) was due to Linnik [L$_1$] who proved, under the assumption of the Riemann Hypothesis, that one could find Goldbach numbers in every interval $(N-K, N]$ with $K > (\log N)^{3+\varepsilon}$ and $N > N_0(\varepsilon)$. Linnik's result was sharpened by Kátai [K] and later but independently by Montgomery and Vaughan [MV$_2$] so as to replace $(\log N)^{3+\varepsilon}$ by $C \log^2 N$ for a suitable absolute constant $C$, again under the assumption of the Riemann Hypothesis.

The next step was taken by Goldston [Go$_2$]. To describe this we need to define the integral

$$J(N,h) = \int_1^N (\psi(x+h) - \psi(x) - h)^2 \, dx, \tag{1.1}$$

Received November 3, 1992.
1991 Mathematics Subject Classification. Primary 11P32; Secondary 11N37, 11N13.
[1]Research supported in part by grants from the NSERC and the National Science Foundation.
[2]Research supported in part by a grant from the National Science Foundation.

where $\psi(x) = \sum_{n \le x} \Lambda(n)$, and $\Lambda$ is the von Mangoldt function, defined to be $\log p$ if $n = p^m$, $m \ge 1$, and zero otherwise. This integral, first studied by Selberg [Se], was used in the work of Montgomery and Vaughan, and, in essence, in the work of Kátai. It is expected that $J(N, h)$ satisfies the asymptotic formula

$$J(N,h) \sim hN \log(N/h), \qquad 1 \le h \le N^{1-\varepsilon}. \tag{1.2}$$

It has been proven in [GM] that, assuming the Riemann Hypothesis, this relation is equivalent to a form of Montgomery's pair correlation conjecture [$M_2$]. We shall use this formula in a shorter range but in the stronger form

$$J(N,h) = hN \log(N/h) + O(hN), \qquad 1 \le h \le \log N. \tag{1.3}$$

Much more than this is expected to be true, see [$Go_2$, (36)]. We actually do not need to assume as strong an error term as in (1.3), as will be clear from the proof.

As consequences of the main theorem of [$Go_2$] it follows that, subject to (1.2), there exists a Goldbach number in every interval $(N - K, N]$ with $K > (2 + \varepsilon)\log N$, $N > N(\varepsilon)$ and moreover, if one takes $K/\log N \to \infty$ then the number of representations

$$r(n) = \sum_{p+p'=n} 1$$

satisfies the expected asymptotic formula

$$\sum_{N-K<n\le N} r(n) \sim K\frac{N}{(\log N)^2}. \tag{1.4}$$

The method in [$Go_2$] provides an upper bound to the error term in (a weighted version of) (1.4) in terms of the integral $J(N, K)$. As was indicated in [$Go_2$] a known lower bound for $J(N, K)$ proved in [$Go_1$] limits the method of [$Go_2$] to intervals no shorter than $\log N$.

In this paper we consider a different method which allows us to study Goldbach numbers in much shorter intervals. This is accomplished by relating the problem to the distribution of primes in arithmetic progressions and thereby in particular to the well-known Elliott-Halberstam conjecture [EH].

For integers $a$ and $q$ with $q \ge 1$ let

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

and $E'(x; q, a) = \psi(x; q, a) - x\chi_0(a)/\phi(q)$ where $\chi_0$ is the principal character mod $q$. We consider estimates of the form

$$\max_a \sum_{q \le Q} |E'(x; q, a)| \ll \frac{x}{(\log x)^A}, \tag{1.5}$$

which hold for arbitrary $A > 0$ with an implied constant that depends on $A$. The Bombieri-Vinogradov theorem [B] implies that (1.5) holds for $Q = x^{1/2}(\log x)^{-B}$ for a suitable $B = B(A)$, and the Elliott-Halberstam conjecture predicts that it holds with much larger $Q$ (not too much smaller than $x$). Upper limits have been found in [FG] and [FGHM] for the size of $Q$ such that (1.5) can hold and, in particular, it follows from [FGHM] that for each $A > 0$, (1.5) will not hold in general for any

$$Q > x/L^{(1-\varepsilon)A}, \qquad \text{where } L = L(x) = \exp\big((\log\log x)^2/\log\log\log x\big). \tag{1.6}$$

This level of $Q$ represents the essential limit of the method in [FGHM].

We shall also need a modified version of (1.5). We define the truncated von Mangoldt function for $n \ge 2$ by

$$\Lambda_Q(n) = \sum_{\substack{d|n \\ d \le Q}} \mu(d) \log(Q/d) \tag{1.7}$$

and equal to zero otherwise; for $Q \ge n$ this is just $\Lambda(n)$. We define

$$\psi_Q(x; q, a) = \sum_{\substack{n \le x,\, n \ne a \\ n \equiv a (\mathrm{mod}\ q)}} \Lambda_Q(n)$$

(it turns out to be convenient to delete the term $n = a$; in the case of $\Lambda$ this is unimportant), and we let $E'_Q(x; q, a) = \psi_Q(x; q, a) - x\chi_0(a)/\phi(q)$. We shall require the upper bound

$$\max_a \sum_{q \le Q} \big|E'_Q(x; q, a)\big| \ll \frac{x}{(\log x)^A}, \tag{1.8}$$

for arbitrary $A > 0$; in fact, as will be evident in what follows, we require such bounds as this and (1.5) only maximized over some range of $a$ and only for some $A$ not very large.

The function $\Lambda_Q$ has been introduced by Goldston in [Go$_3$] where it was used to study primes in short intervals and to give a new proof of a theorem of Bombieri and Davenport [BD] which avoids the circle method and offers a number of advantages. The estimate (1.8) can be shown to hold in the range

$Q < x^{1/2}(\log x)^{-B}$ and may be expected to hold in essentially the same range as (1.5). Indeed, the Bombieri-Vinogradov theorem has been found capable of very wide generalization. Actually, in some respects, the behaviour of $\Lambda_Q$ in arithmetic progressions is better than that of $\Lambda$ since one can prove, using ideas similar to those in Proposition 3 below, an asymptotic formula for $\psi_Q(x; q, a)$ with a much greater uniformity in $q$ than is known to hold in the corresponding formula for $\psi(x; q, a)$. In the case of a closely related function this has been done already by Heath-Brown in [H-B].

Nevertheless, partly because it is not as immediately clear for $\Lambda_Q$ as for $\Lambda$ that its support is almost completely concentrated in the reduced residue classes, it seems worthwhile to also give the results in terms of a modified version of these mean value statements (which in the case of (1.5) is easily seen to be equivalent). Because this formulation is a natural one which seems to be useful for other functions we state it more generally.

We left $f$ denote a general arithmetic function and let $a$ and $q$ be given such that their greatest common divisor is $(a, q) = \Delta$. We consider the error term given by

$$E_f(x; q, a) = \sum_{\substack{n \le x,\, n \ne a \\ n \equiv a (\mathrm{mod}\, q)}} f(n) - \frac{1}{\phi(q/\Delta)} \sum_{\substack{n \le x \\ (n, q) = \Delta}} f(n)$$

and we consider the estimate

$$\max_a \sum_{q \le Q} \left| E_f(x; q, a) \right| \ll \frac{x}{(\log x)^A}. \tag{1.9}$$

In the particular cases $f = \Lambda, \Lambda_Q$ we denote $E_f$ by $E, E_Q$.

Returning to the Goldbach problem, we let

$$R(n) = \sum_{j \le n} \Lambda(j) \Lambda(n - j).$$

The expected asymptotic formula for $R(n)$ is

$$R(n) = n\mathfrak{S}(n) + o(n) \tag{1.10}$$

where $\mathfrak{S}(n)$ is defined for all integers $n \ne 0$ by

$$\mathfrak{S}(n) = \begin{cases} 0 & \text{if } n \text{ is odd;} \\ \mathfrak{S} \prod_{\substack{p \mid n \\ p > 2}} \left( \frac{p-1}{p-2} \right) & \text{if } n \text{ is even, } n \ne 0; \end{cases} \tag{1.11}$$

and

$$\mathfrak{S} = 2\prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right).$$

Our main result is:

THEOREM. *Let $\varepsilon > 0$, $A > 0$ be fixed. Then, for some $c = c(A)$ and for all $N, Q, K$ with $N \geq 3$, $N^{\varepsilon} \leq Q \leq N(L(N))^{-c(A)}$, $2 \leq K \leq Q^{1/4}$, we have*

$$\begin{aligned}\left|\sum_{k=-K}^{K}(K - |k|)(R(N+k) - (N+k)\mathfrak{S}(N+k))\right| &\leq J(N,K) - KN\log(Q/K) + O(KN)\\ &\quad + O\big(K^2N(\log N)^{-A}\big)\\ &\quad + O\Big(K^2\log N\max_k\max_a\sum_{d\leq Q}\big(|E(N;d,a)|\\ &\qquad + |E_Q(N+k;d,a)|\big)\Big), \qquad (1.12)\end{aligned}$$

*where the maxima are taken over $|k| \leq K$, $0 < |a| \leq 2N$. The same result holds with $E'$, $E'_Q$ in place of $E$, $E_Q$.*

Our first corollary is concerned with the expected asymptotic formula (1.4).

COROLLARY 1. *Assume that* (1.3) *holds for $J(N,K)$ and that we have* (1.5) *and* (1.8) (*or alternatively that we have* (1.9) *for both $f = \Lambda$ and $f = \Lambda_Q$*) *for $|x - N| \leq K$, over $0 < |a| \leq 2N$, and $Q = N\exp(-(\log N)^{2/3})$. Then, for $K = (\log N)^{2/3+\varepsilon}$,*

$$\sum_{N-K<n\leq N} r(n) \sim \frac{1}{(\log N)^2}\sum_{N-K<n\leq N} n\mathfrak{S}(n) \sim K\frac{N}{(\log N)^2}.$$

The limit 2/3 of the method comes from an appeal to Vinogradov's estimates for exponential sums. By making an assumption that such exponential sums satisfy heuristically expected estimates one could reduce it further. It is more interesting to note however that, if one asks not for the asymptotic formula but only for the existence of Goldbach numbers in such intervals, then one can, without such an additional assumption, treat substantially shorter intervals. This is the content of our second corollary.

COROLLARY 2. *Assume that* (1.3) *holds for* $J(N, K)$ *and that* (1.5) *and* (1.8) (*or* (1.9) *for* $f = \Lambda$ *and* $f = \Lambda_Q$) *both hold for* $|x - N| \leq K$, *over* $0 < |a| \leq 2N$, *and a given level* $Q \leq N(L(N))^{-c(A)}$. *Then there exist Goldbach numbers in the interval* $(N - K, N]$ *provided that* $K > c\log(N/Q)$ *for a suitable positive absolute constant* $c$.

In particular, if we assume that these mean value estimates hold for $Q$ just below the limit $N(L(N))^{-c}$ then one has Goldbach numbers for all $K > C(\log\log N)^2/\log\log\log N$, for some $C$, and, even assuming a much safer level, one can get such numbers in all intervals of length $K > (\log\log N)^B$ for some $B$.

One of the advantages of the method of [$\mathrm{Go}_3$] over the earlier method of [BD] for studying primes in short intervals is the potential it offers, given strong mean value statements, to treat shorter intervals. Thus, as remarked in [$\mathrm{Go}_3$], the assumption of (1.5) and (1.8) for $Q \leq x^{1-\varepsilon}$ for all $\varepsilon > 0$ implies that $\liminf(p_{n+1} - p_n)/\log p_n = 0$. As a by-product of the present work we get the following analogue of Corollary 2.

COROLLARY 2′. *Assume that* (1.5) *and* (1.8) (*or* (1.9) *for* $f = \Lambda$ *and* $f = \Lambda_Q$) *both hold for* $x = N$, *over* $0 < |a| \leq K$, *and a given level* $Q \leq N(\log N)^{-B(A)}$. *Then, for some* $n$ *with* $N(\log N)^{-2} \leq p_n \leq N$, *we have* $p_{n+1} - p_n < K$ *provided that* $K > c\log(N/Q)$ *for a suitable positive absolute constant* $c$.

Actually, with a little more effort one can choose, say, $N/2 \leq p_n \leq N$. In the case where one assumes (1.5) and (1.8) rather than (1.9) the above result is more or less implicit in [$\mathrm{Go}_3$]. It is worthwhile to note that in this case, as opposed to the statements about the Goldbach problem, we do not require $|a|$ to be very large, certainly we may take it bounded by $\log x$. For $a$ of this size it may possibly be that bounds such as (1.9) hold even with $Q$ as large as $x(\log x)^{-B}$; the methods devised so far cannot rule this out. In this most optimistic case one deduces that $\liminf(p_{n+1} - p_n)/\log\log p_n$ is bounded. One should note however, that for this problem in contrast to the Goldbach problem, the assumption of (1.3) gives a great deal in an almost trivial fashion. Indeed it follows from (1.3) that $\liminf(p_{n+1} - p_n)$ is actually bounded.

In the course of proving the asymptotic formula in Corollary 1 we are led to the problem of estimating the sum $\sum_{n \leq x}\mathfrak{S}(n)$ with as small an error term as possible. Improving earlier results of Hardy-Littlewood [HL] and of Bombieri-Davenport [BD], Montgomery [$\mathrm{M}_1$] proved the formula

$$\frac{\mathfrak{S}}{2}\sum_{n\leq x}\prod_{\substack{p|n\\(p,2r)=1}}\left(\frac{p-1}{p-2}\right) = (x + O(\log x))\prod_{\substack{p|r\\p>2}}\left(1 - \frac{1}{(p-1)^2}\right)$$

where $r$ is a positive integer and the implied constant is absolute. There appears to have been some doubt as to whether this estimate could be improved although in [$\mathrm{MV}_1$, p 210] it is suggested that "possibly ... one might replace the error by $c \log x + o(\log x)$". We shall prove this in the following stronger form:

PROPOSITION 1. *Let $r$ be a positive integer. We have for $x \geq r$,*

$$\frac{\mathfrak{S}}{2} \sum_{n \leq x} \prod_{\substack{p \mid n \\ (p,2r)=1}} \left(\frac{p-1}{p-2}\right) = \left(x - \frac{\phi(2r)}{4r} \log x + O\left(\frac{r}{\phi(r)} (\log x)^{2/3}\right)\right)$$

$$\times \prod_{\substack{p \mid r \\ p>2}} \left(1 - \frac{1}{(p-1)^2}\right),$$

*where the implied constant is absolute.*

The restriction $r \leq x$ can be dropped on replacing $r/\phi(r)$ by $\log\log 3r$ in the above $O$-term (and perhaps even without doing so). In our case we need only $r = 1$ and, in the notation of (1.11), this is

$$\sum_{n \leq x} \mathfrak{S}(n) = x - \frac{1}{2} \log x + O\left((\log x)^{2/3}\right). \tag{1.13}$$

The paper is arranged as follows. In §2 we prove Proposition 1 and give an analogous formula for a similar weighted sum. In §3 we prove that the two forms of the Theorem, that with $E$, $E_Q$, and that with $E'$, $E'_Q$ are equivalent. In §4 and §5 we recall from [$\mathrm{Go}_3$], and in some cases refine, some basic facts about the truncated von Mangoldt function and its use in studying primes in short intervals. In §6 we relate the latter problem to the Goldbach problem and complete the proofs.

## 2. Some singular series averages

In addition to Proposition 1 we shall require for later use the following result of [$\mathrm{Go}_2$] which refines [$\mathrm{M}_1$, Lemma 17.5] and [$\mathrm{Go}_1$, Lemma 3].

PROPOSITION 2. *We have*

$$\sum_{\substack{k=-K \\ k \neq 0}}^{K} (K - |k|)\mathfrak{S}(k) = K^2 - K \log K$$

$$+ K(1 - \gamma - \log 2\pi) + O_\varepsilon(K^{1/2+\varepsilon}). \tag{2.1}$$

This result follows from a standard contour integration argument; see the remarks accompanying Lemma 2.1 below.

*Proof of Proposition* 1. We let $r$ be a positive integer and define

$$H(n) = H_r(n) = \prod_{\substack{p\mid n\\(p,2r)=1}} \left(\frac{p-1}{p-2}\right).$$

Thus we must show that, for $x \geq r$,

$$\frac{\mathfrak{S}}{2} \prod_{\substack{p\mid r\\p>2}} \left(1 - \frac{1}{(p-1)^2}\right)^{-1} \sum_{n\leq x} H(n) = x - \frac{\phi(2r)}{4r}\log x + O\left(\frac{r}{\phi(r)}(\log x)^{2/3}\right) \quad (2.2)$$

with an absolute implied constant. Equation (1.13) follows at once from the case $r = 1$, since in that case $\mathfrak{S}\sum_{n\leq x} H(n) = \sum_{n\leq 2x}\mathfrak{S}(n)$.

We have

$$H(n) = \prod_{\substack{p\mid n\\(p,2r)=1}} \left(1 + \frac{1}{p-2}\right) = \sum_{\substack{d\mid n\\(d,2r)=1}} \frac{\mu^2(d)}{\phi_2(d)}$$

where $\phi_2(p) = p - 2$ and $\phi_2$ is extended to square-free $d$ by multiplicativity. Now,

$$\begin{aligned}\sum_{n\leq x} H(n) &= \sum_{\substack{d\leq x\\(d,2r)=1}} \frac{\mu^2(d)}{\phi_2(d)} \sum_{\substack{n\leq x\\n\equiv 0(d)}} 1\\ &= x \sum_{\substack{d\leq x\\(d,2r)=1}} \frac{\mu^2(d)}{d\phi_2(d)} - \frac{1}{2}\sum_{\substack{d\leq x\\(d,2r)=1}} \frac{\mu^2(d)}{\phi_2(d)} - \sum_{\substack{d\leq x\\(d,2r)=1}} \frac{\mu^2(d)}{\phi_2(d)} P\left(\frac{x}{d}\right),\end{aligned} \quad (2.3)$$

where $P(t) = \{t\} - 1/2$. For the first sum we have

$$\begin{aligned}\sum_{\substack{d\leq x\\(d,2r)=1}} \frac{\mu^2(d)}{d\phi_2(d)} &= \sum_{\substack{d=1\\(d,2r)=1}}^{\infty} \frac{\mu^2(d)}{d\phi_2(d)} - \sum_{\substack{d>x\\(d,2r)=1}} \frac{\mu^2(d)}{d\phi_2(d)}\\ &= \prod_{p\nmid 2r}\left(1 + \frac{1}{p(p-2)}\right) + O\left(\frac{1}{x}\right).\end{aligned} \quad (2.4)$$

For the second sum in (2.3) we use:

LEMMA 2.1. *We have, for $x \geq r$,*

$$\sum_{\substack{d \leq x \\ (d,2r)=1}} \frac{\mu^2(d)}{\phi_2(d)} = \frac{\phi(2r)}{2r} \prod_{p \nmid 2r} \left(1 + \frac{1}{p(p-2)}\right) \log x + O(1), \quad (2.5)$$

*where the implied constant is absolute.*

For the third sum we shall prove:

LEMMA 2.2. *We have*

$$\sum_{\substack{d \leq x \\ (d,2r)=1}} \frac{\mu^2(d)}{\phi_2(d)} P\left(\frac{x}{d}\right) \ll \frac{r}{\phi(r)} (\log x)^{2/3}. \quad (2.6)$$

Inserting (2.4), (2.5) and (2.6) we at once have (2.2) and so Proposition 1 is proven, subject to Lemmata 2.1 and 2.2.

Lemma 2.1, whose proof we do not give, follows from familiar contour integration arguments (indeed, in sharper form than stated) since the generating Dirichlet series

$$\prod_{p \nmid 2r} \left(1 + \frac{1}{(p-2)p^s}\right)$$

may be written as $\zeta(s+1)G(s)$ where $G(0)$ is the multiple of $\log x$ in (2.5), $G(s)$ is analytic for $\sigma > -1/2$ and satisfies, say for $\sigma > -1/4$, a bound $G(s) \ll \tau(r)$, (in order to drop the restriction $x \geq r$ a more careful bound is required). In proving Proposition 2 we use a similar argument starting from the formula

$$\sum_{n \leq x} (x-n)H(n) = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \zeta(s)\zeta(s+1)G(s) \frac{x^{s+1}}{s(s+1)}\, ds.$$

*Proof of Lemma* 2.2. We take as our point of departure an argument of Sitaramachandrarao [Si, Lemma 2.2] wherein the estimate (2.6) is deduced in the case $r = 1$ for the corresponding sum with $\phi_2(d)$ replaced by $\phi(d)$. As there, the key ingredient is the estimate of Walfisz [W]

$$\sum_{n \leq x} \frac{1}{n} P\left(\frac{x}{n}\right) \ll (\log x)^{2/3} \quad (2.7)$$

which is in turn a consequence of Vinogradov's estimates for exponential sums. For $n$ odd we have

$$\frac{\phi(n)}{\phi_2(n)} = \prod_{p|n}\left(1 + \frac{1}{p-2}\right) = \sum_{d\delta=n} \frac{\mu^2(d)}{\phi_2(d)} \tag{2.8}$$

so that

$$\begin{aligned}\sum_{\substack{n\le x\\(n,2r)=1}} \frac{\mu^2(n)}{\phi_2(n)} P\left(\frac{x}{n}\right) &= \sum_{\substack{n\le x\\(n,2r)=1}} \frac{\mu^2(n)}{\phi(n)} P\left(\frac{x}{n}\right) \sum_{d\delta=n} \frac{\mu^2(d)}{\phi_2(d)}\\ &= \sum_{\substack{d\delta\le x\\(d,\delta)=1\\(d\delta,2r)=1}} \frac{\mu^2(d)}{\phi(d)\,\phi_2(d)} \frac{\mu^2(\delta)}{\phi(\delta)} P\left(\frac{x}{d\delta}\right)\\ &= \sum_{\substack{d\le x\\(d,2r)=1}} \frac{\mu^2(d)}{\phi(d)\,\phi_2(d)} \sum_{\substack{\delta\le x/d\\(\delta,2dr)=1}} \frac{\mu^2(\delta)}{\phi(\delta)} P\left(\frac{x/d}{\delta}\right).\end{aligned} \tag{2.9}$$

The inner sum is essentially the same as the original one but with $\phi_2$ replaced by $\phi$. We iterate the previous argument. Similarly to (2.8) we have,

$$\frac{n}{\phi(n)} = \prod_{p|n}\left(1 + \frac{1}{p-1}\right) = \sum_{d\delta=n} \frac{\mu^2(d)}{\phi(d)}, \tag{2.10}$$

and so

$$\begin{aligned}\sum_{\substack{n\le x\\(n,2k)=1}} \frac{\mu^2(n)}{\phi(n)} P\left(\frac{x}{n}\right) &= \sum_{\substack{n\le x\\(n,2k)=1}} \frac{\mu^2(n)}{n} P\left(\frac{x}{n}\right) \sum_{d\delta=n} \frac{\mu^2(d)}{\phi(d)}\\ &= \sum_{\substack{d\le x\\(d,2k)=1}} \frac{\mu^2(d)}{d\phi(d)} \sum_{\substack{\delta\le x/d\\(\delta,2dk)=1}} \frac{\mu^2(\delta)}{\delta} P\left(\frac{x/d}{\delta}\right).\end{aligned} \tag{2.11}$$

Here the inner sum is, apart from the coprimality condition, one estimated in [Si] but for completeness we give the argument. Since $\mu^2(n) = \sum_{d^2|n}\mu(d)$ we have

$$\sum_{\substack{n \le x \\ (n,K)=1}} \frac{\mu^2(n)}{n} P\left(\frac{x}{n}\right) = \sum_{\substack{d^2\delta \le x \\ (d\delta, K)=1}} \frac{\mu(d)}{d^2\delta} P\left(\frac{x}{d^2\delta}\right)$$

$$= \sum_{\substack{d \le x^{1/2} \\ (d,K)=1}} \frac{\mu(d)}{d^2} \sum_{\substack{\delta \le x/d^2 \\ (\delta, K)=1}} \frac{1}{\delta} P\left(\frac{x/d^2}{\delta}\right). \tag{2.12}$$

Finally, for the inner sum here we use

$$\sum_{\substack{n \le x \\ (n,K)=1}} \frac{1}{n} P\left(\frac{x}{n}\right) = \sum_{n \le x} \frac{1}{n} P\left(\frac{x}{n}\right) \sum_{\substack{d \mid n \\ d \mid K}} \mu(d)$$

$$= \sum_{d \mid K} \frac{\mu(d)}{d} \sum_{\delta \le x/d} \frac{1}{\delta} P\left(\frac{x/d}{\delta}\right). \tag{2.13}$$

Since

$$\sum_{d \mid j} \frac{\mu^2(d)}{d} = \prod_{p \mid j} \left(1 + \frac{1}{p}\right) \le \frac{j}{\phi(j)},$$

inserting the Walfisz estimate (2.7) in (2.13) we have

$$\sum_{\substack{n \le x \\ (n,K)=1}} \frac{1}{n} P\left(\frac{x}{n}\right) \ll \frac{K}{\phi(K)} (\log x)^{2/3}.$$

Using this estimate in (2.12) yields

$$\sum_{\substack{n \le x \\ (n,K)=1}} \frac{\mu^2(n)}{n} P\left(\frac{x}{n}\right) \ll \frac{K}{\phi(K)} (\log x)^{2/3}.$$

Inserting this in (2.11) we have

$$\sum_{\substack{n \le x \\ (n,2k)=1}} \frac{\mu^2(n)}{\phi(n)} P\left(\frac{x}{n}\right) \ll \frac{k}{\phi(k)} (\log x)^{2/3}, \tag{2.14}$$

from the multiplicativity of $\phi$ and the convergence of the series

$$\sum_{d=1}^{\infty} \frac{\mu^2(d)}{\phi^2(d)}.$$

Finally, inserting (2.14) in (2.9) the estimate (2.6), that is Lemma 2.2, follows from the convergence of the series

$$\sum_{\substack{d=1 \\ (d,2)=1}}^{\infty} \frac{\mu^2(d)d}{\phi^2(d)\,\phi_2(d)}.$$

*Remark.* It is a simple matter to extend the Walfisz estimate to the bound

$$\sum_{n \le x} \frac{\tau_j(n)}{n} P\left(\frac{x}{n}\right) \ll_j (\log x)^{j-1/3}$$

for the divisor functions $\tau_j$, $j = 2, 3, \dots$. Using this and iterations of the above argument one can treat sums $\sum_{n \le x} H_{k,l}(n)$ where $l \ge 0$ and

$$H_{k,l}(n) = \prod_{p|n} \left(\frac{p-k+l}{p-k}\right).$$

## 3. Two mean value statements

As indicated in the introduction we have chosen to state our results in terms of the error terms $E$, $E_Q$ given in (1.9), which seem to us the more natural, whereas the proofs lend themselves to the error terms $E'$, $E'_Q$ given in (1.5) and (1.8). In this section we prove a result, Proposition 3, that allows us to reduce the first problem to the second.

We require the following lemma which strengthens (and corrects a minor obscurity in) Lemma 1.1.7 of [$L_2$]; the ideas in the proof are the same.

LEMMA 3.1. *Let $A > 0$. There exists $c' = c'(A)$ such that for all $n \ge 3$ we have*

$$\sum_{\substack{\delta | n \\ \delta > \Delta_0}} \frac{\tau(\delta)}{\delta} \ll (\log n)^{-A}, \tag{3.1}$$

*where $\Delta_0 = (L(n))^{c'}$.*

*Proof.* We split the interval $[\Delta_0, n]$ into $\ll \log n$ subintervals of the type $[N', N]$ where $N' \geq N/2$ and seek to bound the contribution to the sum from each subinterval. This contribution is majorized by

$$N^{-1} \sum_{\substack{N' < \delta \leq N \\ \delta | n}} \tau(\delta) \ll N^{-1} \Big( \sum_{\delta \leq N} \tau^2(\delta) \Big)^{1/2} \Big( \sum_{\substack{\delta \leq N \\ \delta | n}} 1 \Big)^{1/2}, \tag{3.2}$$

by Cauchy's inequality. The first sum on the right is bounded by $N(\log N)^3$. Elementary considerations show that the second sum on the right can only be increased if the prime divisors of $n$, say $\nu$ in number, are replaced by the first $\nu$ primes and in that case (since, for large $n$, we have $\nu \leq 2 \log n$) that sum is bounded by the well known function $\psi(N, 2\log n)$. Since, for suitable $c'$ we have

$$\log n \leq \Delta_0^{c' \log\log\log n / \log\log n} \leq N^{c' \log\log\log n / \log\log n} = N^{1/u}$$

the lemma follows from well known bounds for $\psi$; cf. [HT]. Indeed, it suffices to use the rather crude bound

$$N^{-1} \psi(N, N^{1/u}) \ll \exp(-(1 - \varepsilon) u \log u),$$

valid for all $N$ in $(\Delta_0, n)$ for this choice of $u$.

*Remarks.* We shall actually use Lemma 3.1 in a slightly different form namely: If $n$ is an integer satisfying $0 < n \leq x$ then (3.1) still holds if $L(n)$ is replaced by $L(x)$ and $\log n$ is replaced by $\log x$. To see this version we note that the new sum on the left in (3.1) can only be increased on replacing $n$ by a multiple of $n$ and so, by multiplying if necessary by a power of 2, we are able to assume that $x/2 < n \leq x$ so that $\log n = \log x + O(1)$. The proof now follows as before.

The function $L$ in (1.6) which provided the limitation to the range of $Q$ in (1.5) in the method of [FGHM] also provides the limitation to the above lemma and thereby to the range in the statement of the theorem. In fact, by exerting a little more care in the above lemma and using Hölder's inequality in place of Cauchy's, we can take $c'(A)$ in the lemma (and for $A > A_0(\varepsilon)$, $c(A)$ in the theorem) to be $(1 + \varepsilon)A$ in comparison to the $(1 - \varepsilon)A$ in (1.6). It's a little difficult to believe that this is a coincidence.

PROPOSITION 3. (a) *Let $A > 0$ be fixed. There exists a constant $c = c(A)$ such that, uniformly for all $a, Q$ with $0 < |a| \le x$, $Q \le x(L(x))^{-c}$ we have*

$$\sum_{\Delta | a} \sum_{\substack{q \le Q \\ q \equiv 0(\Delta)}} \frac{1}{\phi(q/\Delta)} \left| \sum_{\substack{m \le x \\ (m,q)=\Delta}} \Lambda_Q(m) - x \sum_{\delta | \Delta} \mu(\delta) \right| \ll x(\log x)^{-A}, \quad (3.3)$$

*and the same result holds with $\Lambda$ in place of $\Lambda_Q$.*

(b) *If the range of $a$ is restricted to $0 < |a| \le (\log x)^A$ then the conclusions hold in the wider range $Q \le x(\log x)^{-2A-1}$.*

*Proof.* The fact that (3.3) holds (and indeed, in much stronger form than stated) when $\Lambda_Q$ is replaced by $\Lambda$ is an easy consequence of the prime number theorem with error term and the fact that the support of $\Lambda$ is very small on the non-reduced residue classes.

We consider $\Lambda_Q$. Let $\Delta_0 = (L(x))^c$ with $c$ taken to exceed $c'(A+3)$ from the previous lemma and write the sum in (3.3) as $\Sigma_{\Delta \le \Delta_0} S_\Delta + \Sigma_{\Delta > \Delta_0} S_\Delta$. Since $|\Lambda_Q(m)| \le \tau(m) \log Q$ we have

$$\left| \sum_{\substack{m \le x \\ (m,q)=\Delta}} \Lambda_Q(m) \right| \le \sum_{\substack{m \le x \\ m \equiv 0(\Delta)}} \tau(m) \log Q \ll x(\log x)^2 \Delta^{-1} \tau(\Delta),$$

and thus, by Lemma 3.1 and the remark following its proof,

$$\sum_{\Delta > \Delta_0} S_\Delta \ll x(\log x)^3 \sum_{\substack{\Delta | a \\ \Delta > \Delta_0}} \frac{\tau(\Delta)}{\Delta} \ll x(\log x)^{-A}. \quad (3.4)$$

We now consider $\Sigma_{\Delta \le \Delta_0}$. By the definition of $\Lambda_Q$,

$$\sum_{\substack{m \le x \\ (m,q) \equiv \Delta}} \Lambda_Q(m) = \sum_{d \le Q} \mu(d) \log(Q/d) \sum_{\substack{m \le x \\ (m,q)=\Delta \\ m \equiv 0(d)}} 1.$$

We write $m = rd\Delta/(d, \Delta)$ and the above sum becomes

$$\sum_{\substack{d \le Q \\ (d/(d,\Delta), q/\Delta)=1}} \mu(d) \log(Q/d) \sum_{\substack{r \le x(d,\Delta)/d\Delta \\ (r, q/\Delta)=1}} 1$$

which, by the well-known estimate

$$\sum_{j\le y,(j,m)=1} 1 = \frac{\phi(m)}{m}y + O(\tau(m)),$$

may be written as

$$xq^{-1}\phi(q/\Delta)A_q + O(B_q), \tag{3.5}$$

say, where $B_q = Q\tau(q/\Delta)$ and

$$A_q = \sum_{\substack{d\le Q\\(d/(d,\Delta),Q/\Delta)=1}} \frac{\mu(d)}{d}(d,\Delta)\log(Q/d). \tag{3.6}$$

We sum the contribution from the error terms $B_q$, first over $q$ and then over $\Delta$. Thus

$$\sum_{\substack{q\le Q\\q\equiv 0(\Delta)}} \frac{B_q}{\phi(q/\Delta)} \ll Q\sum_{j\le Q/\Delta}\frac{\tau(j)}{\phi(j)} \ll Q(\log Q)^2,$$

and this makes a contribution to $\Sigma_{\Delta\le\Delta_0}$ which is

$$\ll \Delta_0 Q\log Q \ll x(\log x)^{-A}, \tag{3.7}$$

since $c$ has been chosen to exceed the constant $c'$ from Lemma 3.1.

Returning to $A_q$ defined in (3.6), we write that sum in the form $\Sigma_{\delta|\Delta}\theta_\delta$, where $\theta_\delta$ collects the contribution to $A_q$ of those $d$ for which $(d,\Delta)=\delta$. Letting $r = d/\delta$ one finds, after a little computation,

$$\theta_\delta = \mu(\delta)\sum_{\substack{r\le Q/\delta\\(r,q)=1}} \frac{\mu(r)}{r}\log(Q/\delta r).$$

The estimate

$$\sum_{\substack{r\le y\\(r,q)=1}} \frac{\mu(r)}{r}\log(y/r) = \frac{q}{\phi(q)} + O\big((\log y)^{-A'}\big),$$

which holds for $y > q^\varepsilon$, follows by a contour integration argument similar to that mentioned in §2, but using also the standard zero-free region for $\zeta(s)$.

Inserting this we have (recall $\delta \le \Delta \le \Delta_0$),

$$\theta_\delta = \mu(\delta)\frac{q}{\phi(q)} + O\big((\log x)^{1-A'}\big).$$

Thus, for $\Delta = 1$ we have

$$x \sum_{\substack{q \le Q \\ q \equiv 0(\Delta)}} \frac{A_q}{q} = x \sum_{q \le Q} \frac{1}{\phi(q)} + O\big(x(\log x)^{1-A'}\big), \tag{3.8}$$

and if $1 < \Delta \le \Delta_0$, we have

$$x \sum_{\substack{q \le Q \\ q \equiv 0(\Delta)}} \frac{A_q}{q} \ll x(\log x)^{-A} \sum_{\substack{q \le Q \\ q \equiv 0(\Delta)}} \frac{1}{q} \ll x\Delta^{-1}(\log x)^{1-A'}. \tag{3.9}$$

Choosing $A' = A + 2$, summing these over $\Delta \le \Delta_0$, and combining the result with (3.7) we have $\Sigma_{\Delta \le \Delta_0} S_\Delta \ll x(\log x)^{-A}$. Combining this with (3.4) we complete the proof of (a).

In the case of (b) we choose $\Delta_0 = (\log x)^A$. The sum estimated by (3.4) is now empty and, with this smaller value of $\Delta_0$, the estimate (3.7) holds even for the larger $Q$. Combining these with (3.8) and (3.9) we get the result.

## 4. A truncated von Mangoldt function

We recall from (1.7) the formula

$$\Lambda_Q(n) = \sum_{\substack{d|n \\ d \le Q}} \mu(d)\log(Q/d).$$

For $Q \ge n$ this is just the von Mangoldt function $\Lambda(n)$. We shall use it rather for $Q \le n$ as an approximation to $\Lambda(n)$ in estimating sums over primes. The remaining sum

$$\tilde{\Lambda}_Q(n) = \sum_{\substack{d|n \\ d > Q}} \mu(d)\log(Q/d) \tag{4.1}$$

is regarded as an error term so that, obviously, the larger the level $Q$ that we can choose, the sharper our results. The level $Q$ will be determined by the quality of the Elliott-Halberstam type estimates (1.5) and (1.8) (or (1.9)) that we are able to use.

The starting point is the evident formula, valid for $n \geq 2$,

$$\Lambda(n) = \Lambda_Q(n) + \tilde{\Lambda}_Q(n). \tag{4.2}$$

When applied to $R(n) = \sum_{j \leq n} \Lambda(j)\Lambda(n-j)$ this gives

$$\begin{aligned} R(n) = &\sum_{j \leq n} \big(\Lambda_Q(j)\Lambda(n-j) + \Lambda(j)\Lambda_Q(n-j)\big) \\ &- \sum_{j \leq n} \Lambda_Q(j)\Lambda_Q(n-j) + \sum_{j \leq n} \tilde{\Lambda}_Q(j)\tilde{\Lambda}_Q(n-j). \end{aligned} \tag{4.3}$$

Note that the first sum on the right is just $2\sum_{j \leq n} \Lambda_Q(j)\Lambda(n-j)$. In Lemma 6.1 below we shall give asymptotic formulae for each of the above sums other than the last one. We are unable to do this for the last sum for individual $n$ but are able, also in the final section, to bound an average of this sum for $n$ in a short interval by relating it to the similar problem for primes in short intervals.

We take this opportunity to stress, in connection with the sums occurring in (4.3) as well as several occurring in the next two sections, that we have defined both $\Lambda(n)$ and $\Lambda_Q(n)$ to be zero for $n = 1, 0$, or a negative integer.

## 5. Small gaps between primes

In [Go$_3$] the function $\Lambda_Q$ was used to study the problem of primes in short intervals. There one is concerned with the counting function

$$Z(N;k) = \sum_{n \leq N} \Lambda(n)\Lambda(n+k)$$

and, in analogy to (4.3), one has

$$\begin{aligned} Z(N;k) = &\sum_{n \leq N} \big(\Lambda_Q(n)\Lambda(n+k) + \Lambda(n)\Lambda_Q(n+k)\big) \\ &- \sum_{n \leq N} \Lambda_Q(n)\Lambda_Q(n+k) + \sum_{n \leq N} \tilde{\Lambda}_Q(n)\tilde{\Lambda}_Q(n+k) \end{aligned} \tag{5.1}$$

The sums on the right of (5.1) apart from the last one are evaluated by means of the following lemma.

LEMMA 5.1. *Let $\varepsilon > 0$, $A > 0$. For $N^\varepsilon \leq Q \leq N$, we have*

$$\sum_{n \leq N} \Lambda_Q(n)\Lambda(n) = N \log Q + O(Q) + O\big(N(\log N)^{-A}\big), \tag{5.2}$$

$$\sum_{n \leq N} \Lambda_Q^2(n) = N \log Q + O(N), \tag{5.3}$$

*and for some* $c(A)$ *and for all* $Q, k$ *with* $N^\varepsilon \le Q \le N(L(N))^{-c}$, $0 < |k| \le Q^{1/4}$,

$$\sum_{n\le N} \Lambda_Q(n)\Lambda(n+k) = \mathfrak{S}(k)N + O\Big(\log N \sum_{d\le Q} |E(N; d, k)|\Big) + O\big(N(\log N)^{-A}\big), \tag{5.4}$$

$$\sum_{n\le N} \Lambda_Q(n)\Lambda_Q(n+k) = \mathfrak{S}(k)N + O\Big(\log N \sum_{d\le Q} |E_Q(N; d, k)|\Big) + O\big(N(\log N)^{-A}\big). \tag{5.5}$$

*The equations* (5.4) *and* (5.5) *also hold with* $E'$, $E'_Q$ *in place of* $E$, $E_Q$ *and, if* $k$ *is restricted to satisfy* $0 < |k| \le (\log N)^A$, *they hold in the wider range* $N^\varepsilon \le Q \le N(\log N)^{-2A-1}$.

*Proof.* Formula (5.2) is contained in Lemma 1 of [Go$_3$]. Formula (5.3) is due to S. Graham [Gr]. The statement of (5.4) and (5.5) with $E'$, $E'_Q$ in place of $E$, $E_Q$ sharpens Lemmata 1 and 3 of [Go$_3$] and is achieved by replacing (19) of that paper by the estimate

$$\sum_{\substack{d\le Q\\(d,k)=1}} \frac{\mu(d)}{\phi(d)} \log(Q/d) = \mathfrak{S}(k) + O\big((\log Q)^{-A}\big),$$

which is valid for $Q > k^\varepsilon$, again by a contour integration argument. The fact that (5.4) and (5.5) follow from the corresponding statements for $E'$, $E'_Q$ is an immediate consequence of Proposition 3 and the triangle inequality. This completes the proof of Lemma 5.1.

We substitute equations (5.2)–(5.5) into (5.1) and obtain

$$Z(N; 0) = N\log Q + \sum_{n\le N} \tilde{\Lambda}_Q^2(n) + O(N), \tag{5.6}$$

and, in the case $k \ne 0$,

$$Z(N; k) = \mathfrak{S}(k)N + \sum_{n\le N} \tilde{\Lambda}_Q(n)\tilde{\Lambda}_Q(n+k) + O\big(N(\log N)^{-A}\big) + O\Big(\log N \sum_{d\le Q} \big(|E(N; d, k)| + |E(N; d, -k)| + |E_Q(N; d, k)|\big)\Big). \tag{5.7}$$

Let $t(k) = \max(K - |k|, 0)$, $T(\alpha) = \sum_k t(k)e(-k\alpha)$, and $\tilde{S}(\alpha) = \sum_{m \le M} \tilde{\Lambda}_Q(m)e(m\alpha)$, where $M = N + K$. We consider

$$\int_0^1 |\tilde{S}(\alpha)|^2 T(\alpha)\, d\alpha = \sum_k t(k) \sum_{\substack{n, m \le M \\ n-m=k}} \tilde{\Lambda}_Q(n)\tilde{\Lambda}_Q(m)$$
$$= \sum_k t(k) \sum_{n \le N} \tilde{\Lambda}_Q(n)\tilde{\Lambda}_Q(n+k)$$
$$+ O\big(K^3\tau^2(n)(\log Q)^2\big) \qquad (5.8)$$

(using $|\tilde{\Lambda}_Q(n)| \le \tau(n)\log Q$). By (5.6), (5.7), and the prime number theorem, this is equal to

$$KN\log(N/Q) + O(KN)$$
$$+ \sum_{k \ne 0} t(k)\{Z(N;k) - \mathfrak{S}(k)N\} + O\big(K^2N(\log N)^{-A}\big)$$
$$+ O\Big(K^2 \log N \max_a \sum_{d \le Q} \big(|E(N;d,a)| + |E_Q(N;d,a)|\big)\Big). \qquad (5.9)$$

At this point we are able to prove Corollary 2′. Since, as is well-known, $T(\alpha) \ge 0$ for all $\alpha$ the quantity in (5.9) is non-negative. From the bound $\mathfrak{S}(2k) \ge \mathfrak{S} \ge 1$ and the assumptions of Corollary 2′, we deduce that for suitable $c$, if $c\log(N/Q) \le K \le \log N$, then we have $\sum_{k \ne 0} t(k)Z(N;k) \gg K^2N$. The contribution to this sum from pairs where one of the integers is a power of a prime rather than a prime itself may be disregarded since it is trivially $\ll K^2N^{1/2}\log N$. Similarly the contribution from those pairs less than say $N(\log N)^{-2}$ may be neglected. This gives Corollary 2′.

We now return to (5.9). By Proposition 2 and by the following formula, a slight modification of [Go$_2$, (25)],

$$J(N,K) = KN\log N + O(KN) + \sum_{k \ne 0} t(k)Z(N,k)$$
$$- K^2N + O\big(K^2N(\log N)^{-A}\big),$$

we have for some $c(A)$, and for all $N^\varepsilon \le Q \le N(L(N))^{-c}$, $K \le Q^{1/4}$,

$$\int_0^1 |\tilde{S}(\alpha)|^2 T(\alpha)\, d\alpha$$
$$= J(N,K) - KN\log(Q/K) + O(KN) + O\big(K^2N(\log N)^{-A}\big)$$
$$+ O\Big(K^2 \log N \max_a \sum_{d \le Q} \big(|E(N;d,a)| + |E_Q(N;d,a)|\big)\Big). \qquad (5.10)$$

## 6. Goldbach numbers in short intervals

The reduction of our problem on Goldbach numbers to the above problem on primes in short intervals comes from the trivial inequality

$$\left|\int_0^1 \left(\tilde{S}(\alpha)\right)^2 e(-N\alpha) T(\alpha)\, d\alpha\right| \le \int_0^1 \left|\tilde{S}(\alpha)\right|^2 T(\alpha)\, d\alpha, \tag{6.1}$$

which follows since $T(\alpha) \ge 0$ for all $\alpha$. The integral on the left hand side is equal to

$$\sum_k t(k) \sum_{j \le N+k} \tilde{\Lambda}_Q(j) \tilde{\Lambda}_Q(N+k-j),$$

which by (4.3) is equal to

$$\sum_k t(k) \left\{ R(N+k) - 2 \sum_{j \le N+k} \Lambda_Q(j) \Lambda(N+k-j) \right.$$
$$\left. + \sum_{j \le N+k} \Lambda_Q(j) \Lambda_Q(N+k-j) \right\}.$$

The two inner sums are evaluated by the following lemma which corresponds to Lemma 5.1 but for the Goldbach sum and is proven in the same fashion. It is perhaps worthwhile to remark that it is really only for (6.3) that we require Proposition 3(a); the much simpler (b) version suffices for Lemma 5.1 in the range where the latter is really required.

LEMMA 6.1. *Let $\varepsilon > 0$, $A > 0$. We have for some $c(A)$, and for all $N, Q, n$ satisfying $N^{\varepsilon} \le Q \le N(L(N))^{-c}$, $N/2 \le n \le 2N$,*

$$\sum_{j \le n} \Lambda_Q(j) \Lambda(n-j) = \mathfrak{S}(n) n + O\left(\log N \sum_{d \le Q} |E(n; d, n)|\right)$$
$$+ O\left(N (\log N)^{-A}\right), \tag{6.2}$$

*and*

$$\sum_{j \le n} \Lambda_Q(j) \Lambda_Q(n-j) = \mathfrak{S}(n) n + O\left(\log N \sum_{d \le Q} |E_Q(n; d, n)|\right)$$
$$+ O\left(N (\log N)^{-A}\right). \tag{6.3}$$

On substituting (6.2) and (6.3) we find that the left hand side of (6.1) is equal to

$$\sum_k t(k)(R(N+k) - \mathfrak{S}(N+k)(N+k)) + O\big(K^2N(\log N)^{-A}\big)$$
$$+ O\bigg(K^2 \log N \max_{|k|\le K} \max_a \sum_{d\le Q} \big(|E(N;d,a)| + |E_Q(N+k;d,a)|\big)\bigg).$$

Here we have used the fact that

$$|E(N+H;d,a) - E(N;d,a)| \ll (|H|/d + 1)\log|NH| + |H|/\phi(d).$$

It seems not so easy to prove a similar estimate for $E_Q$ and this is the reason for the lack of symmetry between the two in the inequality (1.12). The Theorem now follows from (5.10) and (6.1).

To prove Corollary 1, we follow the corresponding argument for Corollary 1 of [Go$_2$]. By Propositions 1 and 2 and partial summation we obtain, for $(\log N)^{2/3} \le K \le \log N$,

$$\sum_{N-K<n\le N} n\mathfrak{S}(n) = KN + O\big(N(\log N)^{2/3}\big),$$

and

$$\sum_{k=-K}^{K} (K - |k|)(N+k)\mathfrak{S}(N+k) = K^2N + O\big(KN(\log N)^{2/3}\big).$$

From (1.3), (1.12), and the choice of $Q$ we have

$$\sum_{k=-K}^{K} (K - |k|)R(N+k) = K^2N + O\big(KN(\log N)^{2/3}\big). \tag{6.4}$$

Let $K = (\log N)^{2/3+\varepsilon}$ and $K_1 = K(1 - (\log N)^{-\varepsilon/2})$. We apply (6.4) with $K_1$ and with $K$ and then subtract the first from the second. Used with the well-known sieve upper bound for $R(n)$, cf [HR, p 117], this gives

$$\sum_{N-K<n\le N} R(n) = KN\big(1 + O(\log N)^{-\varepsilon/3}\big). \tag{6.5}$$

(Actually, one can derive (6.5) without appealing to the sieve by using the positivity of $R(n)$ as in [Go$_2$].) The transition from $R(n)$ to $r(n)$ is achieved by noting first that the contribution to (6.5) from higher powers of primes is trivially bounded by $KN^{1/2}\log N$ and second that, for $N - K < n \le N$, we have $\log n = \log N + O(K/N)$.

Corollary 2 follows from the lower bound $\mathfrak{S}(2n) \geq \mathfrak{S} \geq 1$ which implies

$$\sum_{k=-K}^{K} (K - |k|)(N + k)\mathfrak{S}(N + k) \gg K^2 N,$$

which will dominate the right hand side of (1.12) under the conditions of Corollary 2. The contribution from the higher powers of primes can be removed as before.

## References


[B] E. Bombieri, *On the large sieve*, Mathematika **12** (1965), 201–225.

[BD] E. Bombieri and H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. Ser. A **293** (1966), 1–18.

[EH] P.D.T.A. Elliott and H. Halberstam, *A conjecture in prime number theory*, Symp. Math. IV, Rome, 1968–69, 59–72.

[FG] J. Friedlander and A. Granville, *Limitations to the equi-distribution of primes* I, Ann. of Math. **129** (1989), 363–382.

[FGHM] J. Friedlander, A. Granville, A. Hildebrand and H. Maier, *Oscillation theorems for primes in arithmetic progressions and for sifting functions*, J. Amer. Math. Soc. **4** (1991), 25–86.

[Go$_1$] D.A. Goldston, *The second moment for prime numbers*, Quart. J. Math. Oxford (2) **35** (1984), 153–163.

[Go$_2$] ______, *Linniks' theorem on Goldbach numbers in short intervals*, Glasgow Math. J. **32** (1990), 285–297.

[Go$_3$] ______, *On Bombieri and Davenport's theorem concerning small gaps between primes*, Mathematika **39** (1992), 10–17.

[GM] D.A. Goldston and H.L. Montgomery, "Pair correlation of zeros and primes in short intervals" in *Analytic number theory and diophantine problems*, Birkhaüser, Boston 1987, 183–203.

[Gr] S. Graham, *An asymptotic estimate related to Selberg's sieve*, J. Number Theory **10** (1978), 83–94.

[HR] H. Halberstam and H.-E. Richert, *Sieve methods*, LMS Monographs **4**, Academic Press, London, 1974.

[HL] G.H. Hardy and J.E. Littlewood, *Note on Messrs. Shah and Wilson's paper entitled: 'On an empirical formula connected with Goldbach's Theorem'*, Proc. Cambridge Phil. Soc. **19** (1919), 245–254.

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

[HT] A. Hildebrand and G. Tenenbaum, *On integers free of large prime factors*, Trans. Amer. Math. Soc. **296** (1986), 265–290.

[K] I. Kátai, *A remark on a paper of Ju. V. Linnik*, Magyar Tud. Akad. Mat. Fiz. Oszt. Közl. **17** (1967), 99–100.

[L$_1$] Yu.V. Linnik, *Some conditional theorems concerning the binary Goldbach problem*, Izv. Akad. Nauk SSSR **16** (1952), 503–520.

[L$_2$] ______, *The dispersion method in binary additive problems*, Amer. Math. Soc., Providence, R.I., 1963.

[M$_1$] H.L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Mathematics, no. 227, Springer-Verlag, New York, 1971.

[M$_2$] ______, *The pair correlation of zeros of the zeta function*, Proc. Sympos. Pure Math., no. 24, Amer. Math. Soc., Providence, R.I., 1973, 181–193.

[$MV_1$] H.L. MONTGOMERY and R.C. VAUGHAN, *Error terms in additive prime number theory*, Quart. J. Math. Oxford (2) **24** (1973), 207–216.

[$MV_2$] ______, *The exceptional set in Goldbach's problem*, Acta Arith. **27** (1975), 353–370.

[PP] A. PERELLI and J. PINTZ, *On the exceptional set for Goldbach's problem in short intervals*, J. London Math. Soc. (2) **47** (1993), 41–49.

[Se] A. SELBERG, *On the normal density of primes in small intervals, and the difference between consecutive primes*, Arch. Math. Naturvid. **47** (1943), 87–105.

[Si] R. SITARAMACHANDRARAO, *On an error term of Landau*, Indian J. Pure Appl. Math. **13** (1982), 882–885.

[W] A. WALFISZ, *Weylsche Expontentialsummen in der Neueren Zahlentheorie*, VEB Deutscher Verlag, Berlin, 1963.

MATHEMATICAL SCIENCES RESEARCH INSTITUTE
BERKELEY, CALIFORNIA

UNIVERSITY OF TORONTO
TORONTO, ONTARIO, CANADA

SAN JOSE STATE UNIVERSITY
SAN JOSE, CALIFORNIA