ACTA ARITHMETICA
LXXV.2 (1996)

Problems and results on $\alpha p - \beta q$

by

K. Ramachandra, A. Sankaranarayanan
and K. Srinivas (Bombay)

*Dedicated to Professor K. Chandrasekharan
on his seventy-fifth birthday*

1. Introduction. This paper is a continuation of the work of K. Ramachandra [10]. It is in fact a development of the methods adopted there. We prove three theorems of which Theorem 1 (below) is a remark on Goldbach numbers (i.e. numbers which can be represented as a sum of two odd primes). We state and prove Theorems 2 and 3 in Sections 2 and Section 3 respectively. In Section 4 we make some concluding remarks.

THEOREM 1. *Let θ $(3/55 < \theta \le 1)$ be any constant and let x exceed a certain large positive constant. Then the number of Goldbach numbers in $(x, x + x^\theta)$ exceeds a positive constant times x^θ. In particular, if g_n denotes the nth Goldbach number then*

$$g_{n+1} - g_n \ll g_n^\theta.$$

Remark 1. This theorem and its proof has its genesis in Section 9 of the paper [8] of H. L. Montgomery and R. C. Vaughan.

Remark 2. Let α be any positive constant. By considering (in our proof of Theorem 1) the expression

$$\frac{1}{Y}\sum_{y=Y}^{2Y}(\vartheta(x+h-y)-\vartheta(x-y))\left(\vartheta\left(\frac{y+h}{\alpha}\right)-\vartheta\left(\frac{y}{\alpha}\right)\right)$$

$$\left(\text{resp. } \frac{1}{Y}\sum_{y=Y}^{2Y}(\vartheta(x+h+y)-\vartheta(x+y))\left(\vartheta\left(\frac{y+h}{\alpha}\right)-\vartheta\left(\frac{y}{\alpha}\right)\right)\right)$$

we can prove that every interval $(x, x+x^\theta)$ contains a number of the form

1991 *Mathematics Subject Classification*: 11xx, 11Nxx, 11N05.
Key words and phrases: primes, inequalities, Harman's result, Selberg's sieve, Brun's sieve.

$p + q\alpha$ (resp. $p - q\alpha$) where p and q are primes, provided x exceeds a large positive constant.

Since the proof of Theorem 1 follows nearly the corresponding result of [10], with $7/72 < \theta \le 1$, we will prove this theorem in the introduction itself. In both Section 2 and Section 3 and also in Section 4 we deal with the problem: How small can $\alpha p - \beta q$ (> 0) be made if α and β are positive constants and p and q primes? We also consider similar questions about $|\alpha p - \beta q|$.

Proof of Theorem 1. Let ε $(0 < \varepsilon < 1/100)$ be any constant. Let x be any integer which exceeds a large positive constant depending on ε. Put $Y = [x^{6/11+\varepsilon}]$, $h = [Y^{1/10+\varepsilon}]$. As in [10] consider the sum

$$J \equiv Y^{-1} \sum_{y=Y}^{2Y} (\vartheta(x+h-y) - \vartheta(x-y))(\vartheta(y+h) - \vartheta(y)).$$

According to a theorem of G. Harman (see [3]) the number of integers y with $\vartheta(y+h) - \vartheta(y) \le \eta h$ is $o(Y)$ provided η (> 0) is a certain small constant. Since the terms in the sum defining J are non-negative we have

$$J \ge J_1 \equiv \eta h Y^{-1} {\sum_{y=Y}^{2Y}}' (\vartheta(x+h-y) - \vartheta(x-y)),$$

where the accent indicates the restriction of the sum to integers y with $\vartheta(y+h) - \vartheta(y) > \eta h$. (Note that the omitted values are $o(Y)$ in number.) We now include these omitted values of y, and by applying a well-known theorem of V. Brun (see [11]) these contribute $o(h^2)$ to J_1. Thus

$$J \ge J_1 = \eta h Y^{-1} J_2 + o(h^2),$$

where

$$J_2 = \sum_{y=Y}^{2Y} (\vartheta(x+h-y) - \vartheta(x-y)).$$

Now

$$J_2 = \sum_{x+h-2Y \le n \le x+h-Y} \vartheta(n) - \sum_{x-2Y \le n \le x-Y} \vartheta(n).$$

Here the first sum is over $x+h-2Y \le n \le x-Y$ and $x-Y < n \le x+h-Y$, and the second is over $x-2Y \le n < x+h-2Y$ and $x+h-2Y \le n \le x-Y$. Hence

$$J_2 = \sum_{x-Y < n \le x+h-Y} (\vartheta(n) - \vartheta(n-Y-1)).$$

Now by applying the results of S. Lou and Q. Yao (see [7]), we see that each term of the last sum is $\gg Y$ and so $J_1 \gg h^2$ and thus $J \gg h^2$. From this we deduce Theorem 1 as in [10] by applying Corollary 5.8.3 on page 179 of [2].

Remark 1. There is an improvement of log and log log factors in density results in our previous paper [12]. This gives

$$\frac{1}{X}\int_X^{2X}(\vartheta(x+H)-\vartheta(x)-H)^2\,dx = o(H^2(\log X)^{-1})$$

provided $H = X^{1/6}(\log X)^{137/12}(\log\log X)^{47/12}f(X)$, where $f(X)$ is any function of X which tends to infinity as $X\to\infty$. But this has no advantage over the results of G. Harman (see [3]) which in the direction of Theorem 1 is more powerful than all the results known so far.

Remark 2. If we assume a certain obvious hypothesis we can improve Theorem 1. However, assuming a stronger hypothesis like the Riemann hypothesis (R.H.) Theorem 1 can be improved much further (see Section 4 for a remark in this direction).

After Theorem 1 we turn to a different question. We ask whether for every ε (>0), the inequality $0 < p - 2q < p^{\varepsilon}$ has infinitely many solutions in primes p, q. (In Section 2 we prove a more general result which shows that this is so if $\varepsilon > 1/10$.) But if $\varepsilon \le 1/10$ we do not know whether even $|p-2q| < p^{\varepsilon}$ has infinitely many solutions. A milder question is this: Do there exist positive integer constants a and b with $\mu(ab) = -1$ for which $0 < ap - bq < p^{\varepsilon}$ has infinitely many solutions in primes p, q (for every $\varepsilon > 0$)? We do not know whether even $\mu(n(n+1)) = -1$ has infinitely many solutions in positive integers n.

Notation. We use standard notation. The symbol $\equiv$ denotes a definition. The symbols $\ll$ (resp. $\gg$) mean less than (resp. greater than) a positive constant multiple of. Sometimes we specify the constants on which these constants depend by indicating them below these signs. For any positive function g of $X, o(g)$ means a term which when divided by g tends to zero as $X\to\infty$. The symbol $O(g)$ means a quantity which when divided by g remains bounded as $X\to\infty$. The function $\vartheta(x)$ is as usual $\sum \log p$ summed up over all primes $p \le x$. But in Section 3, $\vartheta(x)$ may have a slightly different meaning which will be explained at relevant places. We put $\mu(1) = 1, \mu(n) = 0$ if p^2 divides n for some prime p, and otherwise $\mu(n) = \pm 1$ according as the number of prime factors of n is even or odd. Also if $n > 1$ we define $\Omega(n)$ to be the number of prime factors of n counted with multiplicity. For any real x, $[x]$ will mean the greatest integer $\le x$. We use x (or X) to represent the only independent variable and unless otherwise stated they will be supposed to exceed a large positive constant. Any other notation will be explained at relevant places.

2. Statement and proof of Theorem 2. In the statement of Theorem 2 positive functions $F(x)$ of x with a certain property play an important role. The property is this: For all but $o(X)$ integers n with $X \le n \le 2X$, the intervals $(n, n+F(X))$ should contain $\ge \eta F(X)(\log X)^{-1}$ primes, where $\eta > 0$ is a constant. Owing to the work of G. Harman (see [3]) we can take $F(X)$ to be $X^{1/10+\varepsilon}$, where $\varepsilon > 0$ is any constant and η depends on ε. This work of his has already been referred to in Section 1. (If, however, we assume (R.H.) then there is the following result due to A. Selberg (see [14] and for references to the related material see page 349 of [5] and §9 of [8]):

$$\frac{1}{X}\int_X^{2X} (\vartheta(x+h)-\vartheta(x)-h)^2\,dx \ll (h^{1/2}\log X)^2 \qquad \text{valid for } h \ge 1,\ X \ge 2.$$

Here the left hand side is equal to

$$\frac{1}{X}\sum_{X \le n \le 2X} (\vartheta(n+h)-\vartheta(n)-h)^2 + O(\log X)^2$$

and so we can choose $F(X)$ to be $(\log X)^2$ times any function which tends to ∞ as $X \to \infty$, and η to be any constant < 1.) *It should be stressed that our Theorem 2 is a self contained statement and does not depend on any external unproved hypothesis.* We are now in a position to state Theorem 2.

THEOREM 2. *Let $F(x)$ (with $(\log x)^2 \le F(x) \le x(\log x)^{-2}$) be a function of x such that for all but $o(X)$ integers n in the interval $X \le n \le 2X$, the interval $(n, n+F(X))$ contains at least $\eta F(X)(\log X)^{-1}$ primes, where $\eta > 0$ is some constant. Let $\alpha_1, \ldots, \alpha_N$ $(N \ge 1)$ be any fixed positive constants. Let $e_1, \ldots, e_N$ denote integers which are 0 or 1. Then there exist infinitely many $(N+1)$-tuples of primes $p, q_1, \ldots, q_N$ such that*

$$0 < (-1)^{e_k}(p - \alpha_k q_k) < F(p)f(p) \qquad (k = 1, \ldots, N), \tag{2.1}$$

where $f(x)$ is any function of x which tends to ∞ as $x \to \infty$.

For the purposes of the proof we introduce some notation. First we may assume $f(x) = O(\log\log x)$. n^+ will denote the smallest prime $\ge n$, and n_- the largest prime $< n$, while $\|p\|^+$ will denote the minimum of $q - p$ taken over all primes $q > p$. For any constant δ $(0 < \delta < 1/100)$ we define $\vartheta_\delta(x)$ to be the sum $\sum \log p$ taken over all primes p subject to $p \le x$ and $\|p\|^+ \ge \delta \log p$. We put $h = h(X) = F(X)\sqrt{f(X)}$. C will denote a large positive constant.

We now start with the auxiliary function

$$Q \equiv \sum_{X \le n \le CX} \prod_{k=1}^{N} A_k^{(\delta)}(n^+), \tag{2.2}$$

where for any integer n we have written

$$A_k^{(\delta)}(n) \equiv \left\{\left(\vartheta_\delta\left(\frac{n}{\alpha_k}\right) - \vartheta_\delta\left(\frac{n-[h\alpha_k]}{\alpha_k}\right)\right)^{1-e_k}\left(\vartheta_\delta\left(\frac{n+[h\alpha_k]}{\alpha_k}\right) - \vartheta_\delta\left(\frac{n}{\alpha_k}\right)\right)^{e_k}\right\}.$$

Our aim is to prove that $Q \neq 0$. (In fact, we prove that $Q \gg h^N X$.) This would clearly prove Theorem 2. We write

$$Q_1 \equiv \sum_{X \le n \le CX} \prod_{k=1}^{N} A_k^{(\delta)}(n).$$

Our first step consists in proving that

$$Q \ge Q_1 + o(h^N X) \tag{2.3}$$

for a suitable choice of δ and C. We begin by proving that if p_1 and p_2 are two unequal primes with $\|p_1\|^+ \ge \delta \log p_1$ and $\|p_2\|^+ \ge \delta \log p_2$ then $|p_1 - p_2| \ge \min(\delta \log p_1, \delta \log p_2)$. To see this let $p_1 > p_2$. Then $p_1 - p_2 = p_1 - p_2^* + p_2^* - p_2 \ge p_2^* - p_2 = \|p_2\|^+$, where p_2^* denotes the prime next to p_2. Similarly if $p_2 > p_1$ then $p_2 - p_1 \ge \|p_1\|^+$. This proves our assertion. From this it follows that $A_k^{(\delta)}(n)$ and $A_k^{(\delta)}(n^+)$ are both $O(h\delta^{-1})$ and also their difference is $O(((n^+ - n_-)(\delta \log X)^{-1} + 1)\log X)$. Thus

$$\sum_{n_- \le n \le n^+} |A_k^{(\delta)}(n) - A_k^{(\delta)}(n^+)| = O(\min(h(n^+ - n_-)\delta^{-1}, ((n^+ - n_-)^2(\delta \log X)^{-1} + (n^+ - n_-))\log X)).$$

Hence if in (2.2) we replace $A_N^{(\delta)}(n^+)$ by $A_N^{(\delta)}(n)$ the total error is

$$O\left(\left(\frac{h}{\delta}\right)^{N-1} \sum_{X \le n^+ \le CX} \min\left(\frac{h}{\delta}(n^+ - n_-), \frac{(n^+ - n_-)^2}{\delta} + (n^+ - n_-)\log X\right)\right).$$

Put $H = F(X)$. Then the sum

$$\sum_{X \le n^+ \le CX,\, n^+ - n_- \ge H} (n^+ - n_-)$$

is easily seen to be $\le 2\sum_n 1$, where the sum is over all n for which $(n, n + F(X))$ does not contain any prime. This can be seen by considering the intervals $[n_-, n^+]$ and $[n^+, n^+ + H]$. Hence by hypothesis this sum is $o(X)$. Again

$$\sum_{\substack{X \le n^+ \le CX \\ n^+ - n_- \le H}} \left(\frac{(n^+ - n_-)^2}{\delta} + (n^+ - n_-)\log X\right) = O\left(\frac{HX}{\delta}\right) = o\left(\frac{hX}{\delta}\right),$$

since $\sum_{X\le n^+\le CX}(n^+-n_-)=O(X)$ and $\sum_{X\le n^+\le CX}1=O(X/\log X)$. Hence the total error obtained by replacing in (2.2) the numbers $A_N^{(\delta)}(n^+)$ by $A_N^{(\delta)}(n)$ is $o(h^NX)$. By repeating this process we can replace all the numbers $A_k^{(\delta)}(n^+)$ by $A_k^{(\delta)}(n)$ $(k=1,2,\ldots,N)$ with a total error which is $o(h^NX)$. This proves (2.3). (Note that we have retained δ in the estimates to give a rough idea of how we obtained them. But it is not important.)

We now prove the following lemma.

LEMMA 2.1. *We have, for* $x\ge 2$,

$$0\le \vartheta(x)-\vartheta_\delta(x)\le 4x\delta_1$$

where $\delta_1=\delta_1(x)\to\delta$ *as* $x\to\infty$.

Remark. We postpone the proof of this lemma to the end of the proof of Lemma 2.4.

LEMMA 2.2. *Let*

$$A_k(n)=\left\{\left(\vartheta\left(\frac{n}{\alpha_k}\right)-\vartheta\left(\frac{n-[h\alpha_k]}{\alpha_k}\right)\right)^{1-e_k}\left(\vartheta\left(\frac{n+[h\alpha_k]}{\alpha_k}\right)-\vartheta\left(\frac{n}{\alpha_k}\right)\right)^{e_k}\right\}.$$

Then $A_k(n)\ge A_k^{(\delta)}(n)$ *and*

$$\sum_{X\le n\le CX}(A_k(n)-A_k^{(\delta)}(n))\le 4Xh\delta_1C$$

for $k=1,\ldots,N$ *with* $\delta_1=\delta_1(X)\to\delta$ *as* $X\to\infty$.

Proof. Put $\psi_\delta(x)=\vartheta(x)-\vartheta_\delta(x)$. Then

$$A_k(n)-A_k^{(\delta)}(n)=\psi_\delta\left(\frac{n+[h\alpha_k]}{\alpha_k}\right)-\psi_\delta\left(\frac{n}{\alpha_k}\right)$$

provided $e_k=1$ (the case $e_k=0$ can be treated similarly). Now by treating the sum

$$\sum_{X\le n\le CX}\left(\psi_\delta\left(\frac{n+[h\alpha_k]}{\alpha_k}\right)-\psi_\delta\left(\frac{n}{\alpha_k}\right)\right)$$

just as we treated J_2 in the proof of Theorem 1 of Section 1 we are led to the lemma in view of Lemma 2.1.

LEMMA 2.3. *We have*

$$A_k(n)-A_k^{(\delta)}(n)\le h\sqrt{\delta}\qquad (X\le n\le CX)$$

except for $O(X\delta^{1/2})$ *integers* n. *Also (by hypothesis)* $A_k(n)\ge\eta h$ *for some constant* $\eta>0$ *and for all integers* n $(X\le n\le CX)$ *with the exception of* $o(X)$ *integers* n.

Proof. The proof follows from Lemma 2.2.

LEMMA 2.4. *We have*

$$A_k^{(\delta)}(n) \gg h \qquad (k = 1, 2, \ldots, n)$$

for $\gg X$ integers n in $X \le n \le CX$.

Proof. Follows from Lemma 2.3 by choosing a small δ. Lemmas 2.1 to 2.4 prove that Q_1 and therefore Q are $\gg h^N X$, for a suitable choice of X and δ, provided we prove Lemma 2.1.

Proof of Lemma 2.1. We have to prove that

$$\sum_{p \le x,\, \|p\|^+ \le \delta \log p} \log p \le 4x\delta_1.$$

The proof of this is based on the following lemma.

Note. In Lemma 2.5 and its proof we use h, H. This should not be confused with the earlier ones.

LEMMA 2.5. *Let h be any non-zero integer. Then the number of primes $p \le x$ for which $p + h$ $(= p')$ is again prime is*

$$\le \left\{ 8 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{2<p\mid h} \left(\frac{p-1}{p-2}\right) \right\} \left(1 + O\left(\frac{\log\log x}{\log x}\right)\right) \frac{x}{(\log x)^2},$$

uniformly in h provided h is even. If h is odd the number in question is ≤ 1.

Remark. For the proof of this lemma we refer to Theorem 3.11 on page 117 of [2]. This result is due to E. Bombieri and H. Davenport (see [1]) and independently also to L. F. Kondakova and N. I. Klimov (see [6]). Both these discoveries use the Selberg sieve.

We continue the proof of Lemma 2.1. Put $H = [\delta \log x(1 + o(1))]$. Then it suffices to prove that

$$8\Big(\sum_{1 \le h \le H/2} \prod_{2<p\mid h} \left(\frac{p-1}{p-2}\right)\Big) \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \le 4\delta_1 \log x.$$

For $s = \sigma + it$, $\sigma > 1$, we have

$$\begin{aligned} G(s) &\equiv \sum_{h=1}^{\infty} \Big(\prod_{2<p\mid h} \left(\frac{p-1}{p-2}\right)\Big) h^{-s} \\ &= \sum_{h \text{ odd}} \Big(\sum_{p\mid h} \left(\frac{p-1}{p-2}\right)\Big) h^{-s} + 2^{-s} \sum_{2 \le h \equiv 2 \pmod 4} \prod_{p\mid h/2} (\ldots) \left(\frac{h}{2}\right)^{-s} \\ &\quad + 4^{-s} \sum_{4 \le h \equiv 4 \pmod 8} \prod_{p\mid h/4} (\ldots) \left(\frac{h}{4}\right)^{-s} + \ldots \end{aligned}$$

$$= \Big(\prod_{p>2}\Big\{1+p^{-s}\Big(\frac{p-1}{p-2}\Big)+p^{-2s}\Big(\frac{p-1}{p-2}\Big)+\ldots\Big\}\Big)\sum_{n=0}^{\infty}2^{-ns}$$
$$= \zeta(s)\prod_{p>2}\Big\{\Big(1-\frac{1}{p^s}\Big)\Big(1+\frac{p^{-s}}{1-p^{-s}}\Big(\frac{p-1}{p-2}\Big)\Big)\Big\}$$
$$= \zeta(s)\prod_{p>2}\Big\{1+\Big(\frac{p-1}{p-2}-1\Big)p^{-s}\Big\} = \zeta(s)\prod_{p>2}\Big(1+\frac{p^{-s}}{p-2}\Big).$$

Hence by standard methods (using Perron's formula etc.) we deduce that the quantity in question is

$$\leq 8\Big(\prod_{p>2}\Big\{\Big(1+\frac{1}{p(p-2)}\Big)\Big(1-\frac{1}{(p-1)^2}\Big)\Big\}\Big)\Big(\frac{1}{2}H+o(H)\Big) = 4H(1+o(1)).$$

Thus Lemma 2.1, and hence Theorem 2, is completely proved.

Before leaving this section we record a remark.

Remark. Let S be a fixed infinite set of primes with the property that

$$\Pi(S,x) \equiv \sum_{p\leq x,\, p\in S} 1$$

satisfies $\Pi(S,x) \geq (d+o(1))x(\log x)^{-1}$, where d $(0 < d \leq 1)$ is a constant. We define

$$\vartheta_\delta(S,x) \equiv \sum_{p\leq x,\, p\in S,\ \|p\|^+\geq\delta\log p} \log p \quad \text{and} \quad \vartheta(S,x) \equiv \sum_{p\leq x,\, p\in S} \log p.$$

Then trivially (with any constant $\alpha > 0$) the arguments in the previous paragraphs lead to

$$\sum_{X\leq n\leq CX}\Big(\vartheta_\delta\Big(S,\frac{n+[h\alpha]}{\alpha}\Big)-\vartheta_\delta\Big(S,\frac{n}{\alpha}\Big)\Big)$$
$$\geq \sum_{CX<n\leq CX+[h\alpha]}\Big(\vartheta\Big(S,\frac{n}{\alpha}\Big)-\frac{4n\delta_1}{\alpha}\Big) - \sum_{X\leq n<X+[h\alpha]}\vartheta\Big(\frac{n}{\alpha}\Big)$$
$$\geq \Big(\frac{d}{\alpha}+o(1)-\frac{4\delta_1}{\alpha}\Big)CX[h\alpha]-(1+o(1))\Big(\frac{X+[h\alpha]}{\alpha}\Big)[h\alpha]$$

(by dividing the sum $\vartheta(S,n/\alpha)$ into primes $\leq X(\log X)^{-2}$ and the rest)

$$\geq (C(d-4\delta)-1+o(1))hX.$$

3. Statement and proof of Theorem 3. We begin by stating an earlier result due to K. Ramachandra (see [10]) obtained by applying Selberg's sieve.

Result. *Let ε be any positive constant < 1 and let N be any natural number $> 2\varepsilon^{-1}$. Let $\alpha_1, \ldots, \alpha_N$ be any distinct positive constants. Then there exist two of these constants, say β and γ, such that the inequality*

$$|\beta p - \gamma q| < p^{\varepsilon} \tag{3.1}$$

holds for infinitely many prime pairs (p, q).

By a careful use of Brun's sieve, S. Srinivasan [15] has shown that (3.1) holds with p^{ε} replaced by a certain constant times $\log p$, provided N exceeds a certain large constant. He has also shown that the set of primes can be replaced by a slightly thinner set such as our set S. The purpose of this section is to develop the method of proof of Theorem 2 and prove the following theorem.

Theorem 3. *Let S be a fixed infinite set of primes satisfying*

$$\sum_{p \in S,\, p \le x} 1 \ge (d + o(1))x(\log x)^{-1} \qquad (x \ge 2) \tag{3.2}$$

where d $(0 < d \le 1)$ is any constant. Let δ $(0 < \delta < d/4)$ be any constant and $r \ge 2$ any integer constant. Put

$$N = \left[\frac{r!}{2\delta(d - 4\delta)}\right] + 1. \tag{3.3}$$

Let $\alpha_1, \ldots, \alpha_N$ be any given distinct positive constants. Then there exist r of these constants, say $\beta_1, \ldots, \beta_r$ such that the $\frac{1}{2}r(r-1)$ inequalities

$$|\beta_i p_i - \beta_j p_j| \le \delta(\beta_i + \beta_j)L \qquad (i, j = 1, \ldots, r;\ i < j), \tag{3.4}$$

where $L = \min(\log p_1, \ldots, \log p_r)$, hold for an infinite set of r-tuples of primes $p_1, \ldots, p_r$ all belonging to S.

By taking $d = 1$, $r = 2$, $\delta = 1/8$, $N = 17$ we have the following corollary.

Corollary 1. *Given any 17 distinct positive constants $\alpha_1, \ldots, \alpha_{17}$, there exist two of them, say β and γ, such that the inequality*

$$|\beta p - \gamma q| \le \tfrac{1}{8}(\beta + \gamma) \log p \tag{3.5}$$

has infinitely many solutions in prime pairs (p, q).

Letting δ to be arbitrary and choosing $\alpha_1, \ldots, \alpha_N$ to be suitable distinct rational constants close to 1, we have the following corollary.

Corollary 2. *Given any $\delta > 0$ there exist infinitely many rational constants $\beta > 0$ $(\beta \ne 1)$ with*

$$\lim_{p \to \infty} \inf\{(\min_q |p - \beta q|)(\log p)^{-1}\} \le \delta. \tag{3.6}$$

Here p and q denote primes.

For the proof of Theorem 3 we adopt the notation explained in the last remark of Section 2. For simplicity of notation we do not mention the dependence of $\vartheta_\delta(\ldots)$ etc. on the set S. We implicitly involve S in the notation of this section. We need a few lemmas.

Lemma 3.1. *Let C be a large positive constant and α any positive constant. Then for any function $h = h(X)$ (which is $O(\log X)$ but tends to ∞ as $X \to \infty$) and any constant δ $(0 < \delta < d/4)$, we have*

$$\sum_{X \le n \le CX} \left(\vartheta_\delta\left(\frac{n+[h\alpha]}{\alpha}\right) - \vartheta_\delta\left(\frac{n}{\alpha}\right)\right) \ge (C(d-4\delta) - 1 + o(1))hX. \tag{3.7}$$

Proof. See the last remark in Section 2.

Lemma 3.2. *There exists an integer n (with $X \le n \le CX$) such that*

$$\sum_{j=1}^{N} \left(\vartheta_\delta\left(\frac{n+[h\alpha_j]}{\alpha_j}\right) - \vartheta\left(\frac{n}{\alpha_j}\right)\right) \ge \frac{hN(C(d-4\delta) - 1 + o(1))}{C-1}. \tag{3.8}$$

Proof. The proof follows from Lemma 3.1.

From now on n will be fixed as the one given by Lemma 3.2.

Lemma 3.3. *Put*

$$A_j = \vartheta_\delta\left(\frac{n+[h\alpha_j]}{\alpha_j}\right) - \vartheta_\delta\left(\frac{n}{\alpha_j}\right) \quad \text{and} \quad A = \sum_{j=1}^{N} A_j.$$

Then for any integer $r \ge 2$, we have

$$A^r \le \tfrac{1}{2}(r!\Delta)A^{r-1} + r!Q_2, \tag{3.9}$$

where Q_2 is the sum of the square-free products of $A_1, \ldots, A_N$ occurring in A^r and $\Delta = \max(A_1, \ldots, A_N)$.

Proof. Clearly

$$A^r - r!Q_2 \le \frac{r!}{2}(A_1^2 A^{r-2} + A_2^2 A^{r-2} + \ldots + A_N^2 A^{r-2})$$

and $A_1^2 + A_2^2 + \ldots + A_N^2 \le \Delta A$. Thus the lemma is proved.

Lemma 3.4. *We have, with $R = \log\log X$,*

$$\Delta \le (1 + [h(\delta \log X - 2\delta R)^{-1}])(\log X + R). \tag{3.10}$$

Proof. By our remarks immediately following (2.3) we see that the distance between any two primes occurring in

$$A_j = \sum_{n\alpha_j^{-1} < p < (n+[h\alpha_j])\alpha_j^{-1},\ \|p\|^+ \ge \delta \log p} \log p$$

is $\ge \delta(\log X - 2R)$. This proves the lemma.

LEMMA 3.5. *Put $h = \delta(\log X - R^2)$, where as before $R = \log\log X$. Then*

$$Q_2 \neq 0. \tag{3.11}$$

Proof. Otherwise we have (by Lemma 3.3) $A \le \frac{1}{2}(r!\Delta)$. Note that $\Delta \le \log X + R$ and so

$$\frac{2hN(C(d-4\delta)-1+o(1))}{C-1} \le r!(\log X + R).$$

Here dividing by $\log X$ and letting $X \to \infty$, we have

$$2N\delta(C(d-4\delta)-1)(C-1)^{-1} \le r!.$$

Since C can be chosen to be a large constant it follows that this is possible only if $2N\delta(d-4\delta) \le r!$, which is a contradiction since $N > r!(2\delta(d-4\delta))^{-1}$.

Remark. In fact, we can obtain a contradiction even if δ is replaced by a slightly smaller number.

Proof of Theorem 3. By Lemma 3.5, the hypothesis of the theorem implies that $Q_2 \neq 0$. Thus there exist constants $\beta_1, \ldots, \beta_r$ out of $\alpha_1, \ldots, \alpha_N$ for which (for a suitable n with $X \le n \le CX$) we have

$$\vartheta_\delta\left(\frac{n+[h\beta_j]}{\beta_j}\right) - \vartheta_\delta\left(\frac{n}{\beta_j}\right) > 0 \quad (j = 1, \ldots, r).$$

Hence there exist $p_1, \ldots, p_r$ (all in S) with $X\beta_j^{-1} < p_j < (CX + [h\beta_j])\beta_j^{-1}$ satisfying the inequalities

$$\frac{n+[h\beta_j]}{\beta_j} \ge p_j > \frac{n}{\beta_j} \quad (j = 1, \ldots, r),$$

i.e.

$$|\beta_i p_i - \beta_j p_j| < h(\beta_i + \beta_j) \quad (i, j = 1, \ldots, r; i < j). \tag{3.12}$$

Now $h = \delta(\log X - R^2)$ and $\log p_i = \log X + O(1)$ $(i = 1, \ldots, r)$. By choosing C large and using the remark at the end of the proof of Lemma 3.5, we see that $O(1)$ can be omitted. Next $\beta_1, \ldots, \beta_r$ depend on X. But there are at most $N!(r!(N-r)!)^{-1}$ choices for them and the set of r-tuples of primes which figure in (3.12) is infinite. Thus we have (3.4) for an infinite set of r-tuples of primes for a suitable set of r constants (chosen from $\alpha_1, \ldots, \alpha_N$) which we again denote by $\beta_1, \ldots, \beta_r$ for simplicity of notation. This proves Theorem 3 completely.

4. Concluding remarks. (1) Our proof of Theorem 1 actually gives the following corollary. Suppose θ_1 and θ_2 are two positive constants such that $\pi(x+Y) - \pi(x) \gg Y(\log Y)^{-1}$ for $Y = [x^{\theta_1+\varepsilon}]$ and with $h = [Y^{\theta_2+\varepsilon}]$ we have $\pi(y+h) - \pi(y) \gg h(\log h)^{-1}$ for all integers y in $[Y, 2Y]$ with $o(Y)$ exceptions ($\varepsilon > 0$ being arbitrary). Then Theorem 1 holds with $\theta_1\theta_2 < \theta \le 1$. In proving Theorem 1 we have taken $\theta_1 = 6/11$ and $\theta_2 = 1/10$, which are known results giving the above hypothesis for these values of θ_1 and θ_2. However, we have to use the method of [10]. We will deal with consequences of R.H. and Montgomery's pair correlation conjecture in another paper. According to A. Perelli the last two hypotheses together seem to imply that there are $\gg \log x$ Goldbach numbers in the interval $(x, x + D\log x)$, where $D > 0$ is a certain large constant.

(2) G. Harman has shown in [4] that almost all intervals $(n, n+(\log n)^{7+\delta})$ (where $\delta > 0$ is any constant) contain a number of the type p_1p_2, where p_1 and p_2 are odd, p_1 "big" and p_2 "small". From this it follows by our method that there exist infinitely many pairs (n_1, n_2) of positive integers with $\Omega(2n_1n_2) \le 5, \mu(2n_1n_2) = -1$ and $|n_1 - 2n_2| \le (\log n_1)^{7+\delta}$. (See also D. Wolke [16] for a bigger constant in place of 7 in Harman's result. Mention has to be made of Y. Motohashi's result [9]. Using his result we can prove that each of $0 < p - 2p_1p_2 < p^{\varepsilon}$ and $0 < 2p_1p_2 - p < p^{\varepsilon}$ (taken separately) has infinitely many solutions in odd primes p, p_1 and p_2 with p_1 "big" and p_2 "small".)

(3) In Sections 2 and 3 we get economical constants since we have used the results based on Selberg's sieve. If, however, we use results based on Brun's sieve (see [11] for an exposition of Brun's sieve) we obtain bad constants; but the analogous results are still true. We take this opportunity to point out some numerical corrections in [11]: page 90, 2_-, $4a\log 2 - 3 \to 4^a a\log 2 - 2$; page $91, 4^+, 4a\log 2 \to 4^a a\log 2; 3_-, 4a\log 2 < D + 3 \to 4^a a\log 2 < D + 2$; page $92, 9^+, 4a\log 2 \to 4^a a\log 2; 10^+, 4a\log 2 < D + 3 \to 4^a a\log 2 < D + 2$.

(4) From the results of S. Lou and Q. Yao (see [7]) it follows that given any constant $\alpha > 0$ and any prime p there exists a prime q such that $0 < \alpha p - q \ll_{\varepsilon} p^{6/11+\varepsilon}$, where ε (> 0) is any arbitrary constant.

Acknowledgements. We are very much indebted to the work [2] and its continuation [13] for beautiful expositions of Brun's sieve and Selberg's sieve. We should also mention that the exposition of Brun's sieve in [11] was also useful. The first author wishes to acknowledge the help given to him by H. Halberstam during his stay for three months in Nottingham, U.K., during 1978.

References

[1] E. Bombieri and H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. London Ser. A 293 (1966), 1–18.
[2] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.
[3] G. Harman, *Primes in short intervals*, Math. Z. 180 (1982), 335–348.
[4] —, *Almost primes in short intervals*, Math. Ann. 258 (1981), 107–112.

[5] A. Ivić, *The Riemann Zeta-function*, Wiley-Interscience, New York, 1985.
[6] L. F. Kondakova and N. I. Klimov, *Certain additive problems*, Volzh. Mat. Sb. 7 (1969), 41–44 (in Russian).
[7] S. Lou and Q. Yao, *A Chebychev's type of prime number theorem in a short interval II*, Hardy-Ramanujan J. 15 (1992), 1–33.
[8] H. L. Montgomery and R. C. Vaughan, *The exceptional set of Goldbach's problem*, Acta Arith. 27 (1975), 353–370.
[9] Y. Motohashi, *A note on almost primes in short intervals*, Proc. Japan Acad. Ser. A Math. Sci. 55 (1979), 225–226.
[10] K. Ramachandra, *Two remarks in prime number theory*, Bull. Soc. Math. France 105 (1977), 433–437.
[11] —, *Viggo Brun (13-10-1885 to 15-8-1978)*, Math. Student 49 (1981), 97–95.
[12] K. Ramachandra, A. Sankaranarayanan and K. Srinivas, *Ramanujan's lattice point problem, prime number theory and other remarks*, Hardy-Ramanujan J. 19 (1996), 2–56.
[13] H.-E. Richert, *Sieve Methods*, Lecture Notes 55, Tata Institute of Fundamental Research, Bombay, 1976.
[14] A. Selberg, *On the normal density of primes in small intervals, and the difference between consecutive primes*, Arch. Math. Naturvid. 47 (1943), 87–105.
[15] S. Srinivasan, *A note on* $|\alpha p - q|$, Acta Arith. 41 (1982), 15–18.
[16] D. Wolke, *Fast-Primzahlen in kurzen Intervallen*, Math. Ann. 244 (1979), 233–242.

Post-script (November 1995)**.** Regarding Remark (1) of Section 4 the following results have been proved recently (information from Professor A. Perelli): $\theta_1 = 535/1000$ due to R. C. Baker and G. Harman (to appear in Proc. London Math. Soc.), $\theta_2 = 1/14$ due to Nigel Watt (*Short intervals almost all containing primes*, Acta Arith. 72 (1995), 131–167). The result of Nigel Watt has been improved by K. C. Wong (a student of Glyn Harman). His exponent is 1/18. This is in the course of publication. More recently we came to know from the Editors that the exponent has been improved to 1/20 by Jia Chaohua (*Almost all short intervals containing prime numbers*, to appear in Acta Arith.). Thus in Theorem 1 we can replace 3/55 by 535/20000. This seems to be the best known result of this kind.

School of Mathematics
Tata Institute of Fundamental Research
Homi Bhabha Road
Colaba, Bombay 400 005, India
E-mail: kram@tifrvax.tifr.res.in
sank@tifrvax.tifr.res.in
srini@tifrvax.tifr.res.in

Received on 29.5.1995 (2800)