TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 302, Number 1, July 1987

ON THE DISTRIBUTION OF THE NUMBER OF PRIME FACTORS OF SUMS $a+b$

P. ERDÖS, H. MAIER AND A. SÁRKÖZY

ABSTRACT. We continue a series of investigations by A. Balog and two of the authors (P. Erdös and A. Sárközy) on the arithmetic properties of the elements $a+b$, where $a \in \mathbf{A}$, $b \in \mathbf{B}$, $\mathbf{A}$ and $\mathbf{B}$ "dense sequences."

The present paper transfers the famous Erdös-Kac theorem on the normal distribution of the number of distinct prime factors of integers to such "sum sequences."

1. Throughout this paper we use the following notations: For any real number x let $[x]$ denote the greatest integer less than or equal to x, and let $\|x\|$ denote the distance from x to the nearest integer: $\|x\| = \min(x-[x], 1+[x]-x)$. We write $e^{2\pi i x} = e(x)$. The cardinality of the finite set X is denoted by $|X|$. $\nu(n)$ denotes the number of distinct prime factors of n, while $\Omega(n)$ denotes the total number of prime factors of n counted with multiplicity. We denote the distribution function of the normal distribution by $\phi(x)$:

$$\phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-u^2/2}\, du.$$

2. Recently in several papers Balog, Erdös, Sárközy and Stewart have studied problems of the following type: If $\mathbf{A}, \mathbf{B}$ are "dense" sequences of positive integers then what can be told about the arithmetic properties of the sums $a+b$ with $a \in \mathbf{A}$, $b \in \mathbf{B}$? (See, e.g., [**1**, **4**, **7** and **8**].) These results show that from many aspects the behavior of these sums is similar to the behavior of the sequence of the consecutive integers. This fact led us to the following question: is it true that if $\mathbf{A}, \mathbf{B}$ are "dense" sequences, then the sums $a+b$ with $a \in \mathbf{A}$, $b \in \mathbf{B}$ must satisfy an Erdös-Kac type theorem (see [**3**])? In other words, is it true that if $\mathbf{A}, \mathbf{B}$ are sets of integers not exceeding x and $|\mathbf{A}|, |\mathbf{B}|$ are "large" in terms of x, then the distribution of the numbers

$$(\nu(a+b) - \log\log x)/(\log\log x)^{1/2}$$

can be approximated well by the normal distribution?

First we are going to show that if we count any integer n that can be represented in the form $a+b=n$ only once (independently of the number of solutions of $a+b=n$) then the answer to the question above is negative. To see this, define the set $\mathbf{A}$ in the following way: $\mathbf{A} = 2 \cup \{n \colon n \le x,\, n \equiv 1 \pmod 2,\, \nu(n+2) > \log\log x\}$,

Received by the editors June 27, 1986.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11N60; Secondary 11B75.
Research supported by an NSF grant.

and put $\mathbf{B} = \mathbf{A}$. By the Erdös-Kac theorem, we have $|\mathbf{A}| = |\mathbf{B}| \sim x/4$. It can be shown (by using, e.g., the Chinese remainder theorem) that almost all the even integers $n = 2k$ with $n \leq 2x$ can be represented in the form $a + a'$, $a \in \mathbf{A}$, $a' \in \mathbf{A}$ so that the number of the integers that can be represented in the form $a + a'$ is altogether about $x + x/4 = 5x/4$. On the other hand, again by the Erdös-Kac theorem, about half of the even integers $n = 2k \leq 2x$ satisfy

$$\nu(n) > \log\log x \tag{1}$$

(and they almost all are of the form $a+a'$), and also all the odd integers n satisfying $3 \leq n \leq x+2$ and (1) are of the form $a + a'$, $a \in \mathbf{A}$, $a' \in \mathbf{A}$ (with $a = 2$) so that there are about $x/2 + x/4 = 3x/4$ integers satisfying (1) and of the form $a+a'$. This is much more than their expected number $\sim 1/2 \cdot 5x/4 = 5x/8$ (since the expected number would be half of the number of all the integers of the form $a+a'$), and thus no Erdös-Kac type theorem can hold.

On the other hand, we will show that if we count the sums $a+b$ *with multiplicity*, then the following Erdös-Kac type theorem holds.

THEOREM. *There exist absolute constants x_0, C_1 such that if $x > x_0$, $\mathbf{A}, \mathbf{B}$ are sets of positive integers not exceeding x and l is an arbitrary positive integer, then we have*

$$\left| |\{(a,b)\colon a \in \mathbf{A},\ b \in \mathbf{B},\ \nu(a+b) \leq l\}| - \phi\left(\frac{l - \log\log x}{(\log\log x)^{1/2}}\right) |\mathbf{A}||\mathbf{B}| \right| < C_1 x (|\mathbf{A}||\mathbf{B}|)^{1/2} (\log\log x)^{-1/4}. \tag{2}$$

We note that one can prove the analogous assertion with $\Omega(a+b)$ in place of $\nu(a+b)$ in near the same way. However, some technical details are slightly simpler in case of the ν function, and this is the reason that we have preferred to study this case.

We guess that (2) can be improved by replacing the exponent $1/4$ on the right-hand side by $1/2$ but unfortunately we have not been able to prove this. (On the other hand, certainly (2) does not hold with an exponent larger than $1/2$ in place of $1/4$.)

By the theorem above,

$$(|\mathbf{A}||\mathbf{B}|)^{-1} |\{(a,b)\colon a \in \mathbf{A},\ b \in \mathbf{B},\ \nu(a+b) \leq l\}|$$

can be approximated by

$$\phi\left(\frac{l - \log\log x}{(\log\log x)^{1/2}}\right)$$

provided that

$$x^2 (|\mathbf{A}||\mathbf{B}|)^{-1} = o((\log\log x)^{1/2}).$$

Probably the upper bound on the right-hand side can be improved. On the other hand, it cannot be replaced by $O(\exp(\varepsilon\sqrt{\log\log x}\log\log\log x))$. This can be shown by the following construction: let P denote the product of the primes not exceeding $(\varepsilon/3)\sqrt{\log\log x}\log\log\log x$, and let $\mathbf{A} = \mathbf{B}$ be the set of the integers of the form kP where $1 \leq k \leq x/P$. It is easy to see that in this case,

$$x^2 (|\mathbf{A}||\mathbf{B}|)^{-1} > \exp(\varepsilon\sqrt{\log\log x}\log\log\log x),$$

and, on the other hand, the sums $a+b$ have "too many" prime factors, so that, e.g., for $l=[\log\log x]$ we have

$$(|\mathbf{A}||\mathbf{B}|)^{-1}|\{(a,b)\colon a\in\mathbf{A},\ b\in\mathbf{B},\ \nu(a+b)\le l\}|<\frac{1}{2}-2\delta$$
$$<\phi\left(\frac{l-\log\log x}{(\log\log x)^{1/2}}\right)-\delta$$

for some $\delta=\delta(\varepsilon)>0$.

We will prove the theorem above by using the Hardy-Littlewood method. Another alternative approach could be to use the moment method but it would give much weaker estimates. On the other hand, the moment method has the advantage that it is more flexible so that one can use it to derive some nontrivial results also in some cases when the more analytical approach does not work, e.g., when the cardinalities of the given sequences are smaller. We hope to return to these questions in a subsequent paper.

3. Our theorem follows relatively easily from Lemma 1 below.

For $l=0,1,2,\dots$, put

$$S(x,l,\alpha)=\sum_{\substack{n\le x\\ \nu(n)\le l}} e(n\alpha)$$

and

$$E(x,l)=\phi\left(\frac{l-\log\log x}{(\log\log x)^{1/2}}\right).$$

LEMMA 1. *There exist absolute constants x_1, C_2 such that for $x>x_1$, $l=0,1,2,\dots$, and any real number α we have*

$$\left|S(x,l,\alpha)-E(x,l)\sum_{n=1}^{[x]}e(n\alpha)\right|<C_2x(\log\log x)^{-1/4}.$$

In this section we are going to derive our theorem from Lemma 1, while §§3–8 will be devoted to the proof of Lemma 1.

Put

$$F(\alpha)=\sum_{a\in\mathbf{A}}e(a\alpha),\qquad G(\alpha)=\sum_{b\in\mathbf{B}}e(b\alpha).$$

Clearly, we have

$$\begin{aligned}&|\{(a,b)\colon a\in\mathbf{A},\ b\in\mathbf{B},\ \nu(a+b)\le l\}|\\
&=\sum_{\substack{a\in\mathbf{A},b\in\mathbf{B}\\ n\le 2x,\nu(n)\le l\\ a+b-n=0}}1=\sum_{\substack{a\in\mathbf{A},b\in\mathbf{B}\\ n\le 2x,\nu(n)\le l}}\int_0^1 e((a+b-n)\alpha)\,d\alpha\\
&=\int_0^1\left(\sum_{a\in\mathbf{A}}e(a\alpha)\right)\left(\sum_{b\in\mathbf{B}}e(b\alpha)\right)\left(\sum_{\substack{n\le 2x\\ \nu(n)\le l}}e(-n\alpha)\right)d\alpha\\
&=\int_0^1 F(\alpha)G(\alpha)S(2x,l,-\alpha)\,d\alpha.\end{aligned}$$

Thus by using Lemma 1 (with $2x$ in place of x), Cauchy's inequality and the Parseval formula, we obtain that

$$\begin{aligned}
&||\{(a,b)\colon a\in\mathbf{A},\ b\in\mathbf{B},\ \nu(a+b)\le l\}|-E(x,l)|\mathbf{A}||\mathbf{B}||\\
&\quad\le ||\{(a,b)\colon a\in\mathbf{A},\ b\in\mathbf{B},\ \nu(a+b)\le l\}|-E(2x,l)|\mathbf{A}||\mathbf{B}||\\
&\qquad+|E(2x,l)-E(x,l)||\mathbf{A}||\mathbf{B}|\\
&\quad=\left|\int_0^1 F(\alpha)G(\alpha)S(2x,l,-\alpha)\,d\alpha-E(2x,l)\sum_{a\in\mathbf{A},b\in\mathbf{B}}1\right|\\
&\qquad+o\left((\log\log x)^{-1/2}|\mathbf{A}||\mathbf{B}|\right)\\
&\quad=\left|\int_0^1 F(\alpha)G(\alpha)S(2x,l,-\alpha)\,d\alpha-E(2x,l)\sum_{\substack{a\in\mathbf{A},b\in\mathbf{B}\\ n\le 2x,a+b-n=0}}1\right|\\
&\qquad+o\left((\log\log x)^{-1/2}|\mathbf{A}||\mathbf{B}|\right)\\
&\quad=\left|\int_0^1 F(\alpha)G(\alpha)S(2x,l,-\alpha)\,da\right.\\
&\qquad\qquad\left.-E(2x,l)\int_0^1 F(\alpha)G(\alpha)\sum_{n=1}^{[2x]}e(-n\alpha)\,d\alpha\right|\\
&\qquad+o\left((\log\log x)^{-1/2}|\mathbf{A}||\mathbf{B}|\right)\\
&\quad\le\int_0^1|F(\alpha)||G(\alpha)|\left|S(2x,l,-\alpha)-E(2x,l)\sum_{n=1}^{[2x]}e(-n\alpha)\right|\,d\alpha\\
&\qquad+o\left((\log\log x)^{-1/2}|\mathbf{A}||\mathbf{B}|\right)\\
&\quad=O\left(x(\log\log x)^{-1/4}\int_0^1|F(\alpha)||G(\alpha)|\,d\alpha\right)+o\left((\log\log x)^{-1/2}|\mathbf{A}||\mathbf{B}|\right)\\
&\quad=O\left(x(\log\log x)^{-1/4}\left(\int_0^1|F(\alpha)|^2\,d\alpha\int_0^1|G(\alpha)|^2\,d\alpha\right)^{1/2}\right)\\
&\qquad+o\left((\log\log x)^{-1/2}|\mathbf{A}||\mathbf{B}|\right)\\
&\quad=O\left(x(\log\log x)^{-1/4}(|\mathbf{A}||\mathbf{B}|)^{1/2}\right)+O\left((\log\log x)^{-1/2}|\mathbf{A}||\mathbf{B}|\right)\\
&\quad=O\left(x(|\mathbf{A}||\mathbf{B}|)^{1/2}(\log\log x)^{-1/4}\right)
\end{aligned}$$

(since clearly $|E(2x,l)-E(x,l)|=o((\log\log x)^{-1/2})$ uniformly in l) and this completes the proof of our theorem (assuming Lemma 1).

4. We start with an outline for the proof of Lemma 1. The method bears resemblance with that of Vinogradov applied in the proof of his Three Primes Theorem [**9**]. We decompose the interval $I=[0,1)$. Let

$$\mathbf{M}_0=\{\alpha\in I\colon \|\alpha\|\le x^{-1}(\log\log x)^{1/4}\}.$$

The asymptotic equation

$$S(x,l,\alpha)\sim E(x,l)\sum_{n=1}^{[x]} e(n\alpha)$$

is shown only for $\alpha\in\mathbf{M}_0$. It follows easily from the Erdös-Kac theorem with error term [**2**] by partial summation. For $\alpha\notin\mathbf{M}_0$ we merely show that $S(x,l,a)$ and $E(x,l)\sum_{n=1}^{[x]} e(n\alpha)$ are both small.

Let $C_3>0$ be a fixed large constant. The set $\mathbf{M}_1$ is defined by $\mathbf{M}_1=\{\alpha\colon \|\alpha\|\le x^{-1}\log^{C_3}x\}-\mathbf{M}_0$. For each rational number a/q with $1<q\le\log^{C_3}x$, $0<a<q$, $(a,q)=1$, we form the neighborhood

$$\mathbf{M}_{a,q}=\{\alpha\colon |\alpha-a/q|\le x^{-1}\log^{C_3}x\}.$$

We denote the union of these "major arcs" by $\mathbf{M}_2$:

$$\mathbf{M}_2=\bigcup_{\substack{1<q\le\log^{C_3}x\\(a,q)=1}}\mathbf{M}_{a,q}.$$

To estimate $S(x,l,\alpha)=\sum_{n\le x;v(n)\le l} e(n\alpha)$ for $\alpha\in\mathbf{M}_2$ we factor $n=mP_1(n)$. (Here and in the sequel $P_i(n)$ will denote the ith largest prime-factor of n.) We apply the prime number theorem of Page-Siegel-Walfisz to deal with the sums over the $P_1(n)$'s. The estimate for $\alpha\in\mathbf{M}_1$ is much simpler but depends on the same basic ideas.

Let $\mathbf{M}_3=I-(\mathbf{M}_0\cup\mathbf{M}_1\cup\mathbf{M}_2)$. Here we again use the factorization $n=mP_1(n)$. However the sums over $P_1(n)$ are estimated by a mean-value argument.

5. Before estimating $S(x,l,\alpha)$ on the "major arcs" we prove some preliminary lemmas.

LEMMA 2.

$$\sum_{n\le x} e(n\alpha)=\int_1^x e(u\alpha)\,du+O(\|\alpha\|x).$$

(*Here and in the sequel the constants implied in the O- and $\ll$-symbols are absolute unless indicated otherwise.*)

PROOF. This follows immediately by Euler's summation formula.

LEMMA 3. *Let $z\ge 3$, $A\ge 1$, $0<|\beta|\le z^{-1/2}(\log z)^{-1}$. Then*

$$\int_2^z \frac{e(\beta u)}{\log u}\,du\ll\beta^{-1}(\log z)^{-1}\qquad(z\to\infty).$$

PROOF. We have

$$\int_2^z \frac{e(\beta u)}{\log u}\,du=\int_2^{z^{1/2}}\frac{e(\beta u)}{\log u}\,du+\left[\beta^{-1}e(\beta u)(\log u)^{-1}\right]_{u=z^{1/2}}^{u=z}+\beta^{-1}\int_{z^{1/2}}^z\frac{e(\beta u)}{u(\log u)^2}\,du.$$

Each of the three terms is $\ll|\beta|^{-1}(\log z)^{-1}$.

LEMMA 4. *Let* $1 < \tilde{q} \le (\log z)^{3C_3}$; $(c, \tilde{q}) = 1$. *Then we have*

$$\sum_{\substack{p \le z \\ p \equiv c \bmod \tilde{q}}} e(\beta p) = \frac{1}{\phi(\tilde{q})} \int_2^z \frac{e(\beta u)}{\log u}\, du + O_{C_3}\left((1 + |\beta| z) z \exp(-C_4 (\log z)^{1/2})\right),$$

where $C_4 > 0$ *is an absolute constant.*

PROOF. We apply the prime number theorem of Page-Siegel-Walfisz [**5**, p. 144, Satz 8.3],

$$\pi(u; \tilde{q}, c) = \frac{1}{\phi(\tilde{q})} \operatorname{li} u + O_{C_3}\left(u \exp(-C_4 (\log u)^{1/2})\right)$$

for $z^{1/2} < u \le z$. Lemma 4 follows immediately by partial summation.

LEMMA 5. *Let* $\psi(z, y)$ *denote the number of* $n \le z$ *that have only prime factors* $\le y$. *Then we have for* $y \le z$,

$$\psi(z, y) < z \exp\left\{ -\frac{\log\log\log y}{\log y} \log x + \log\log y + O\left(\frac{\log\log y}{\log\log\log y}\right)\right\}$$

for $y = y(z) \to \infty$ *as* $z \to \infty$.

PROOF. This is a result of Rankin [**6**].

In the sequel we will use repeatedly conditions of the form

$$P_1(m) \ge \exp\left(\frac{\log x}{(\log\log x)^2}\right).$$

Therefore we introduce

$$Y = Y(x) = \exp\left(\frac{\log x}{(\log\log x)^2}\right). \tag{5.1}$$

LEMMA 6.

$$\sum_{m : x/(\log x)^{3C_3} < m P_1(m) \le x} \frac{1}{m} \ll_{C_3} (\log\log x)^3.$$

PROOF. We set $m = k \cdot P$, $P = P_1(m)$, $B = (\log x)^{3C_3}$. Then the summation condition for $mP_1(m)$ implies

$$(x/Bk)^{1/2} < P_1(m) \le (x/k)^{1/2}.$$

By Lemma 5 we have

$$\sum_{m : x/B < m P_1(m) \le x} \frac{1}{m} = \sum_{k \le x/Y} \sum_{\substack{(x/Bk)^{1/2} < P \le (x/k)^{1/2} \\ P \text{ prime}}} \frac{1}{P} + O((\log x)^{-1}).$$

The P-sum is

$$\ll \frac{\log B}{\log(x/k)} \ll_{C_3} \frac{(\log\log x)^3}{\log x}.$$

6. We now turn to the "major arcs" estimate.

LEMMA 7. (i) *For* $\alpha \in \mathbf{M}_2$ *we have*

$$S(x,l,\alpha) \ll_{C_3} x(\log\log x)^{-1/2}.$$

(ii) *For* $\alpha \in \mathbf{M}_1$ *we have*

$$S(x,l,\alpha) \ll_{C_3} \|\alpha\|^{-1} + x(\log\log x)^6/(\log x).$$

PROOF. We first prove (i) and then indicate the modifications necessary for the proof of (ii).

Given (a,q) with $q \le (\log x)^{C_3}$; $(a,q) = 1$; $|\alpha - a/q| \le x^{-1}(\log x)^{C_3}$. Let $q = p_1^{\gamma_1}\cdots p_r^{\gamma_r}, p_1 < \cdots < p_r, \gamma_i > 0$, be the prime factorization of q. By application of the Chinese remainder theorem we get a decomposition

$$\frac{a}{q} = \frac{b_1}{p_1^{\gamma_1}} + \frac{b_2}{p_2^{\gamma_2}} + \cdots + \frac{b_r}{p_r^{\gamma_r}} \quad \text{with } (b_i, q_i) = 1.$$

We factor $n = p_1^k mP$, where $(m,p_1) = 1$, $k \ge 0$, $P = P_1(n)$. Then we replace $S(x,l,\alpha)$ by a slightly modified sum $S^{(1)}(\alpha)$. (For the sake of simplicity we omit the dependence on x and l in the notations during the proof of Lemma 7.)

$$S(x,l,\alpha) = S^{(1)}(\alpha) + O(x/p_1(\log\log x)^{1/2}),$$

where

$$\begin{aligned} S^{(1)}(\alpha) &= \sum_{\substack{n\le x;(n,p_1)=1\\ \nu(n)\le l}} e(n\alpha) + \sum_{\substack{n\le x; n\equiv 0 \bmod p_1\\ \nu(n)\le l+1}} e(n\alpha)\\ &= \sum_{k\ge 0} S_{p_1,k}(\alpha), \end{aligned}$$

with

$$S_{p_1,k}(\alpha) = \sum_{\substack{m\le x\\ \nu(m)\le l-1;(m,p_1)=1\\ P_1(m)\le x/mp_1^k}} \sum_{\substack{P\in I(p_1,k,m)\\ P \text{ prime}}} e(p_1^k mP\alpha) + O\left(\frac{x}{p_1^k \log x}\right).$$

Here $I(p_1,k,m) = [\max(P_1(m),p_1); x/mp_1^k]$. The error-term $O(x/(p_1^k\log x))$ contains the contributions from n with double prime factor $P_1(n)^2$. We now determine the positive integer $L = L(p_1)$ by

$$p_1^L \le (\log x)^{2C_3} < p_1^{L+1}.$$

Furthermore we define

$$\begin{aligned} \mathbf{M} &= \mathbf{M}(x,l,p_1)\\ &= \{m \le xY^{-1} : (m,p_1) = 1; mP_1(m) \le x/(\log x)^{3C_3}, v(m) \le l-1\}. \end{aligned}$$

Then we replace $S^{(1)}(\alpha)$ by $S^{(2)}(\alpha)$:

$$S^{(1)}(\alpha) = S^{(2)}(\alpha) + R, \tag{6.1}$$

where

$$S^{(2)}(\alpha) = \sum_{0\le k\le L} \sum_{m\in\mathbf{M}} \sum_{\substack{P\le x/mp_1^k\\ P \text{ prime}}} e(p_1^k mP\alpha).$$

For the estimate of the error R we have to consider the effects of three modifications: the restriction of the k-sum to $k \le L$, the additional condition imposed on m, $m \in \mathbf{M}$, and finally the change in the range of summation from $P \in I(p_1, k, m)$ to $P \le x/mp_1^k$. We have

$$\sum_{k>L} S_{p_1,k}(\alpha) \ll \frac{x}{(\log x)^{2C_3}}. \tag{6.2}$$

In the second step we apply Lemma 6,

$$\sum_{0\le k\le L} \sum_{\substack{m\notin \mathbf{M} \\ m\le xY^{-1}}} \sum_{\substack{P\in I(p_1,k,m) \\ P \text{ prime}}} 1 \ll \frac{x}{\log x}(\log\log x)^6. \tag{6.3}$$

The contribution of $m \ge xY^{-1}$ is $\ll x/(\log x)$ by Lemma 5. In the last estimate we can assume that $m \in \mathbf{M}$. We observe that $\max(P_1(m), p_1) \ll (x/mp_1^k)(\log x)^{-C_3}$ and thus

$$\sum_{\substack{P\in I(p_1,k,m) \\ P \text{ prime}}} e(p_1^k mP\alpha) = \sum_{\substack{P\le x/mp_1^k \\ P \text{ prime}}} e(p_1^k mP\alpha) + O\left(\frac{x}{mp_1^k(\log x)^{C_3}}\right)$$

which gives

$$\begin{aligned} &\sum_{0\le k\le L} \sum_{m\in\mathbf{M}} \sum_{\substack{P\in I(p_1,k,m) \\ P \text{ prime}}} e(p_1^k mP\alpha) \\ &= \sum_{0\le k\le L} \sum_{m\in\mathbf{M}} \sum_{\substack{P\le x/mp_1^k \\ P \text{ prime}}} e(p_1^k mP\alpha) + O\left(\frac{x}{\log x}\right). \end{aligned} \tag{6.4}$$

From (6.1), (6.2), (6.3) we get

$$R \ll x(\log\log x)^6/(\log x). \tag{6.5}$$

We now set $\tilde q = q/p_1^{\gamma_1}$ and decompose the sum $S^{(2)}(\alpha)$ according to the residue class mod $\tilde q$ of $n = p_1^k mP$. We set

$$S^{(2)}(\alpha) = \sum_{c \bmod \tilde q} S_c(\alpha), \tag{6.6}$$

where

$$S_c(\alpha) = \sum_{0\le k\le L} \sum_{m\in\mathbf{M}} \sum_{P\le x/mp_1^k} e(p_1^k mP\alpha).$$

We denote the residue classes mod $p_1^{\gamma_1}$, by sp_1^k, where $(s, p_1) = 1$ and decompose $S_c(\alpha)$ according to residue classes of $p_1^k mP \bmod p_1^{\gamma_1}$.

We set $\alpha - a/q = \theta$.

$$S_c(\alpha) = \vartheta \sum_{m\in\mathbf{M}} \left[\sum_{0\le k<\gamma_1} \sum_{1\le s\le p_1-1} e\left(\frac{p_1^k s b_1}{p_1^{\gamma_1}}\right) \otimes \sum_{\substack{P\le x/mp_1^k \\ p_1^k mP\equiv c \bmod \tilde{q} \\ P\equiv m^{-1}s \bmod p_1^{\gamma_1-k}}} e(\theta p_1^k mP) + \sum_{\gamma_1\le k\le L} \sum_{\substack{P\le x/mp_1^k \\ p_1^k mP\equiv c \bmod \tilde{q}}} e(\theta p_1^k mP) \right],$$

where $\vartheta = e(c(b_2/p_2^{\gamma_2}+\cdots+b_r/p_r^{\gamma_r}))$, c the common residue class of $n = p_1^k mP \bmod \tilde{q}$. Let

$$\sum_{(k,p_1,m,\theta)} := \sum_{\substack{P\le x/mp_1^k, P \text{ prime} \\ mp_1^k\equiv c \bmod \tilde{q} \\ P\equiv m^{-1}s \bmod p_1^{\gamma_1-k}}} e(\theta p_1^k mP) \quad \text{for } k<\gamma_1,$$

and

$$\sum_{(k,p_1,m,\theta)} := \sum_{\substack{P\le x/mp_1^k; P \text{ prime} \\ mp_1^k P\equiv c \bmod \tilde{q}}} e(\theta p_1^k mP) \quad \text{for } k\ge\gamma_1.$$

By Lemma 4 we have now with $\beta = \theta p_1^k m$, $z = x/mp_1^k$:

$$\sum_{(k,p_1,m,\theta)} := \begin{cases} \dfrac{1}{\phi(qp_1^{-k})} \displaystyle\int_2^{x/mp_1^k} \frac{e(\theta p_1^k mu)}{\log u}\, du \\ \qquad + O\left(\dfrac{x}{m} \exp(-C_4(\log x)^{1/3}\right) & \text{for } k<\gamma_1, \\ \dfrac{1}{\phi(qp_1^{-\gamma_1})} \displaystyle\int_2^{x/mp_1^k} \frac{e(\theta p_1^k mu)}{\log u}\, du \\ \qquad + O\left(\dfrac{x}{m} \exp(-C_4(\log x)^{1/3})\right) & \text{for } k\ge\gamma_1. \end{cases}$$

The substitution $v = p_1^k u$ gives

$$\sum_{(k,p_1,m,\theta)} = \begin{cases} \dfrac{1}{\phi(q)} \displaystyle\int_2^{x/m} \frac{e(\theta mv)}{\log v}\, dv \\ \qquad + O\left(\dfrac{x(\log\log x)^3}{\phi(q)m(\log x)^2}\right) & \text{for } k<\gamma_1, \\ \dfrac{1}{\phi\left(qp_1^{-\gamma_1}\right) P_1^k} \displaystyle\int_2^{x/m} \frac{e(\theta mv)}{\log v}\, dv \\ \qquad + O\left(\dfrac{x(\log\log x)^3}{\phi\left(qp_1^{-\gamma_1}\right) p_1^k m(\log x)^2}\right) & \text{for } k\ge\gamma_1. \end{cases}$$

Thus

$$S_c(\alpha) = \vartheta \sum_{m\in\mathbf{M}} \frac{1}{\phi(q)} \sum_{s \bmod p_1^{\gamma_1}} e\left(\frac{s}{p_1^{\gamma_1}}\right) \int_2^{x/m} \frac{e(\theta m v)}{\log v}\, dv$$

$$\cdot \left(1 + O\left[\frac{(\log\log x)^3}{\log x}\right]\right) dv.$$

Since the inner sum without the error term is zero we get

$$S^{(2)}(\alpha) = \sum_{c \bmod \tilde{q}} S_c(\alpha) \ll \frac{(\log\log x)^4}{\log x}. \tag{6.7}$$

Part (i) of Lemma 7 now follows from (6.4) and (6.8). The proof of part (ii) is much simpler since we need not consider residue classes.

In analogy to (6.1) we obtain

$$S(x,l,\alpha) = S^{(2)}(\alpha) + O\left(x(\log\log x)^6/\log x\right) \tag{6.8}$$

with

$$S^{(2)}(\alpha) = \sum_{m\in\mathbf{M}} \sum_{\substack{P\le x/m \\ P \text{ prime}}} e(mP\alpha).$$

We now follow the estimate of $\sum_{(k,p_1,m,\theta)}$ but omit the factor p_1^k and the congruence conditions. By Lemma 3 we get $S^{(2)}(\alpha) \ll \|\alpha\|^{-1}$ which together with (6.9) proves (ii) of Lemma 7.

7. We now estimate $S(x,l,\alpha)$ on the "minor arcs". As a preparation we need

LEMMA 8.

$$|\{n \le x\colon P_1(n) \le 2P_2(n)\}| \ll \frac{x(\log\log x)^2}{\log x}.$$

PROOF. We factor $n = mP_2(n)P_1(n)$ and get

$$\sum_{P_2\le x} \sum_{P_2\le P_1\le 2P_2} \frac{x}{P_2P_1} \ll \sum_{Y/2\le P_2\le x} \frac{x}{P_2\log P_2} + \sum_{\substack{n\le x \\ P_1(n)\le Y}} 1$$

$$\ll \frac{x(\log\log x)^2}{\log x},$$

LEMMA 9. *If f,a,q',q are integers, U,α are real numbers with $0 < q' \le q$, $U \ge 1$, $|\alpha - a/q| < 1/q^2$ then we have*

$$\sum_{y=f}^{f+q'} \min\left(U, \frac{1}{2\|y\alpha\|}\right) < 4U + q\log q.$$

This is a special case of Lemma 8a in [**9**, p. 24].

LEMMA 10. *Let $1 < q < x$, $(a,q) = 1$, $\alpha = a/q + \theta/q^2$, $|\theta| \le 1$. Then we have*

$$S(x,l,\alpha) \ll_{C_3} x\frac{(\log\log x)^2}{\log x}$$

$$+ x(\log x)^{5/2}\left(\frac{1}{q} + \frac{q}{x} + (\log x)^{-C_3} + \frac{(\log x)^{C_3}}{x}\right)^{1/2}.$$

PROOF. In $S(x,l,\alpha)$ we can restrict the summation to n with

$$P_1(n) \le x/(\log x)^{C_3}.$$

The number of exceptional n is $\ll_{C_3} x/(\log x)(\log\log x)$. This follows from

$$\sum_{\substack{x/(\log x)^{C_3}<P\le x\\ P \text{ prime}}} \frac{1}{P} \ll_{C_3} \frac{\log\log x}{\log x}.$$

We again decompose $n = mP$, where $P = P_1(n)$. We get

$$S(x,l,\alpha) = \sum_{U,V} S_{U,V}(\alpha) + O\left(\frac{x\log\log x}{\log x}\right),$$

where

$$S_{U,V}(\alpha) = \sum_{\substack{U<m\le U'\\ V<P_1(m)\le 2V\\ (U'\le 2U)}} \quad \sum_{\substack{2V<P\le x/m\\ P \text{ prime}}} e(mP\alpha).$$

The U,V-sum runs over $\ll (\log x)^2$ pairs (U,V), where $U \ge (\log x)^{C_3}$. The number of n not counted in the sums $S_{U,V}(\alpha)$ is, by Lemma 8, $\ll x(\log\log x)^2/(\log x)$. We now estimate $S_{U,V}(\alpha)$:

$$\begin{aligned}
S^2_{U,V}(\alpha) &\ll U \sum_{U<m\le U'} \left| \sum_{\substack{2V<P\le x/m\\ P \text{ prime}}} e(mP\alpha)\right|^2 \\
&\ll U \sum_{P_1,P_2} \sum_{U<m\le \min(x/P_1,x/P_2,U')} e(m(P_1-P_2)\alpha).
\end{aligned}$$

Thus

$$\begin{aligned}
S^2_{U,V} &\ll U \sum_{P_1,P_2\le x/U} \min\left(U, \frac{1}{\|\alpha(P_1-P_2)\|}\right) \\
&= U \sum_{-x/U<n<x/U} \left(\sum_{\substack{n_1,n_2\le x/U\\ n_1-n_2=n}} 1\right) \min\left(U, \frac{1}{\|\alpha n\|}\right) \\
&\ll U\cdot\frac{x}{U} \sum_{j=-[x/Uq]}^{[x/Uq]} \sum_{n=jq+1}^{(j+1)q} \min\left(U,\frac{1}{\|\alpha n\|}\right) \\
&\ll x \sum_{j=-[x/Uq]-1}^{[x/Uq]} (U+q\log q) \ll x\left(\frac{1}{Uq}+1\right)(U+q\log q) \\
&\ll x\left(\frac{x}{Uq}+1\right)(U+q)\log x = x^2\log x\left(\frac{1}{q}+\frac{q}{x}+\frac{U}{x}+\frac{1}{U}\right) \\
&\ll x^2 \log x\left(\frac{1}{q}+\frac{q}{x}+(\log x)^{-C_3}+\frac{(\log x)^{C_3}}{x}\right).
\end{aligned}$$

This proves Lemma 10.

8. We now conclude the proof of Lemma 1 and thus of our theorem. The Erdös-Kac theorem with remainder [**2**] gives

$$\sum_{\substack{n\le u\\ v(n)\le l}} 1 = E(x,l)u + O(u(\log\log u)^{-1/2}) \quad \text{for } u \ge x^{1/2}.$$

From this it follows by partial summation that

$$S(x,l,\alpha) = E(x,l)\int_1^x e(u\alpha)\,du + O((1+\|\alpha\|x)x(\log\log x)^{-1/2}).$$

By Lemma 2 the claim of Lemma 1 now follows for $\alpha \in \mathbf{M}_0$. By Lemma 2 the claim of Lemma 1 now follows from lemmas 2 and 7 and from the fact that $\int_1^x e(u\alpha)\,du \ll \|\alpha\|^{-1}$. For $\alpha \in \mathbf{M}_3$ by Dirichlet's approximation theorem there exists a, q with $(a,q) = 1$, $\log^{C_3} x < q = x(\log x)^{-C_3}$, and $|\alpha - a/q| \le \theta/q^2$, $|\theta| \le 1$. Lemma 1 then follows from Lemma 10. That concludes the proof.

REFERENCES

1. A. Balog and A. Sárközy, *On sums of sequences of integers.* III, Acta Math. Acad. Sci. Hungar. **44** (1984), 339–349.
2. P. D. T. A. Elliott, *Probabilistic number theory.* II, Springer-Verlag, 1980.
3. J. Erdös and M. Kac, *The Gaussian law of errors in the theory of additive number-theoretic functions*, Amer. J. Math. **62** (1940), 738–742.
4. P. Erdös and A. Sárközy, *On divisibility properties of integers of the form* $a + a'$, Acta. Math. Acad. Sci. Hungar. (to appear).
5. K. Prachar, *Primzahlverteilung*, Springer-Verlag, 1957.
6. R. A. Rankin, *The difference between consecutive prime numbers*, J. London Math. Soc. **13** (1938), 242–247.
7. A. Sárközy, *On the number of prime factors of integers of the form* $a_i + b_j$, Studia Sci. Math. Hungar. (to appear).
8. A. Sárközy and C. L. Stewart, *On divisors of sums of integers.* II, J. Reine Angew. Math. **365** (1986), 171–191.
9. I. M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, Interscience, New York, 1954.

MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES, 1053 BUDAPEST, REALTANODA U. 13–15, HUNGARY (Current address of P. Erdös and A. Sárközy)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF GEORGIA, ATHENS, GEORGIA 30602 (Current address of H. Maier)