JOURNAL OF NUMBER THEORY 44, 93–104 (1993)

On Elements of Sumsets with Many Prime Factors

P. ERDŐS

Mathematical Institute, Hungarian Academy of Sciences, H-1053 Budapest, Hungary

C. POMERANCE

Department of Mathematics, University of Georgia, Athens, Georgia 30602

A. SÁRKÖZY

Mathematical Institute, Hungarian Academy of Sciences, H-1053 Budapest, Hungary

AND

C. L. STEWART

Department of Pure Mathematics, University of Waterloo, Waterloo, Ontario N2L 3Gl, Canada

Communicated by Alan C. Woods

Received July 15, 1991

In this paper we show that if $\mathscr{A}$ and $\mathscr{B}$ are "dense" sets of positive integers, then there is some member of the sumset $\mathscr{A}+\mathscr{B}$ with many prime factors.

1. INTRODUCTION

Let $\nu(n)$ denote the number of distinct prime factors of the positive integer n. Further, let $\Omega(n)$ denote the number of prime factors of n counted with multiplicity. In this paper we are concerned with showing that ν and Ω are forced to be quite large on some elements of the sumset $\mathscr{A}+\mathscr{B}$ if $\mathscr{A}$ and $\mathscr{B}$ are "dense" sets of positive integers.

Recently in several papers, Balog, Elliott, Maier, Tenenbaum, and the authors have studied problems of the following type: if $\mathscr{A}$ and $\mathscr{B}$ are

0022-314X/93 $5.00

"dense" sets of integers, then what can be said of the arithmetical properties of the elements of $\mathscr{A}+\mathscr{B}$? In particular, Balog and Sárközy [1] studied the problem of finding suitable conditions on $\mathscr{A}$ and $\mathscr{B}$ that would force $\mathscr{A}+\mathscr{B}$ to have an element with only small prime factors, while in [6] and in several papers referenced therein, the problem of when $\mathscr{A}+\mathscr{B}$ contains an element with a very large prime factor is considered. In [2, 3] the problem of when there is an Erdős–Kac type theorem for the distribution of the numbers $v(a+b)$ for $a\in\mathscr{A}, b\in\mathscr{B}$ is studied. See [5] for further problems and references.

Let $m(N)$ denote the largest integer m for which $p_1 p_2 \cdots p_m \leqslant N$, where p_i denotes the ith prime. Thus

$$m(N)=\max\{v(k): k\leqslant N\}.$$

We show below that for each $\varepsilon>0$ there are numbers $c(\varepsilon)$, $N(\varepsilon)$ such that whenever $N\geqslant N(\varepsilon)$ and $\mathscr{A}$ and $\mathscr{B}$ are sets of integers in $[1, N/2]$ with $|\mathscr{A}|\,|\mathscr{B}|>\varepsilon N^2$, we have some $a\in\mathscr{A}, b\in\mathscr{B}$ with

$$v(a+b)>m(N)-c(\varepsilon)\sqrt{m(N)}.$$

We further show that this result is nearly best possible.

In addition, we show some similar results for the function Ω.

Throughout the paper, all latin letters except c will represent positive integers. Further, if a, k are integers with $k>0$, then let $r(a, k)$ denotes the integer in $[-k/2, k/2)$ that is congruent to $a \bmod k$.

2. Large Values of v on $\mathscr{A}+\mathscr{A}$

We first consider the case $\mathscr{A}=\mathscr{B}$, so that we wish to show that if $\mathscr{A}$ is "dense," then there are $a, a'\in\mathscr{A}$ with $v(a+a')$ large.

It is easy to show that the function $m=m(N)$ defined above satisfies

$$m=(1+o(1))\frac{\log N}{\log\log N} \qquad \text{as} \quad N\to\infty. \tag{1}$$

For $N\geqslant e^4$, let $x=x(N)$ denote the largest integer with $p_x\leqslant\sqrt{\log N}$. Define $n=n(N)$ to be the largest integer with

$$p_{x+1}p_{x+2}\cdots p_{x+n}\leqslant N.$$

Clearly we have

$$n\leqslant m\leqslant x+n.$$

Thus from (1), we have as $N \to \infty$,

$$0 \leqslant m - n \leqslant x = \pi(\sqrt{\log N}) = o(\sqrt{m}). \tag{2}$$

THEOREM 1. *There exist effectively computable positive constants c_0 and N_0 such that if N is an integer with $N > N_0$, $\mathscr{A} \subset \{1, 2, ..., N\}$, L is a positive integer with $L < n/2$, and*

$$\frac{|\mathscr{A}|}{N} > c_0 2^{-n} \sum_{l=0}^{L} \binom{n}{l}, \tag{3}$$

then there exist integers $a, a' \in \mathscr{A}$ with

$$v(a + a') > 6L - 2n. \tag{4}$$

COROLLARY 1. *For each $\varepsilon > 0$ there exist effectively computable positive numbers $c(\varepsilon)$, $N(\varepsilon)$ such that if N is an integer with $N > N_0(\varepsilon)$ and $\mathscr{A}$ is a set of integers in $[1, N]$ with $|\mathscr{A}| > \varepsilon N$, then there exist integers $a, a' \in \mathscr{A}$ with*

$$v(a + a') > m - c(\varepsilon)\sqrt{m}. \tag{5}$$

Proof of Corollary 1. It is easy to see that the hypothesis implies there are effectively computable positive numbers $c'(\varepsilon)$ and $N'(\varepsilon)$ such that if $N > N'(\varepsilon)$ then (3) holds with $L > 0.5n - c'(\varepsilon)\sqrt{n}$. Thus the Corollary follows from (2) (with an effective estimate for $\pi(\sqrt{\log N})$) and (4).

Note that we clearly have

$$\max_{a, a' \in \mathscr{A}} v(a + a') \leqslant \max_{i \leqslant 2N} v(i) = m(2N) \leqslant m + 1,$$

so that (5) is best possible apart from an $O(\sqrt{m})$ term.

It is clear that Theorem 1 is only interesting in the case that $L > n/3$. Suppose that $|\mathscr{A}| > N \exp(-c' \log N/\log\log N)$ where c' is a small positive constant. Then for sufficiently large N, (3) holds with $L = [0.34n] + 1$. Hence, in view of (2), Theorem 1 implies for large N that there are $a, a' \in \mathscr{A}$ with

$$v(a + a') > 6L - 2n > \frac{n}{25} > \frac{m}{26} > \frac{\log N}{27 \log\log N}.$$

We thus have the following result.

COROLLARY 2. *There exist effectively computable positive constants c_1, N_1 such that if N is an integer with $N > N_1$ and $\mathscr{A}$ is a set of integers in $[1, N]$ with*

$$|\mathscr{A}| > N \exp\left(-c_1 \frac{\log N}{\log\log N}\right),$$

then there exist integers $a, a' \in \mathcal{A}$ *with*

$$v(a+a') > \frac{\log N}{27 \log\log N}.$$

3. A Combinatorial Lemma

The proof of Theorem 1 will be based on a combinatorial lemma (Lemma 2 below) which can be derived from the following result of Katona [4].

Lemma 1. *Let* k *and* n *be integers of the same parity with* $0 < k < n$. *Let* $\mathcal{S}$ *be any set of cardinality* n *and let* $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_t$ *be distinct subsets of* $\mathcal{S}$ *with*

$$t > \sum_{l=0}^{(n-k)/2} \binom{n}{l}. \tag{6}$$

Then there exist subsets $\mathcal{S}_i, \mathcal{S}_j$ *with* $i \neq j$ *and*

$$|\mathcal{S}_i \cap \mathcal{S}_j| < k. \tag{7}$$

(Note that the lower bound in (6) is the best possible. In fact, taking all the subsets $\mathcal{T}$ of $\mathcal{S}$ with $|\mathcal{T}| \geqslant (n+k)/2$, there are $\sum_{l=0}^{(n-k)/2} \binom{n}{l}$ of them and for any pair $\mathcal{T}, \mathcal{T}'$ of them, $|\mathcal{T} \cap \mathcal{T}'| \geqslant k$.)

Lemma 2. *Let* k *and* n *be integers of the same parity with* $0 < k < n$ *and put* $L = (n-k)/2$. *Let* $\mathcal{R}$ *be a set of cardinality* n *and let* $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_z$ *be distinct subsets of* $\mathcal{R}$ *with*

$$z > 2 \sum_{l=0}^{L} \binom{n}{l}. \tag{8}$$

Then there exist subsets $\mathcal{R}_i, \mathcal{R}_j$ *with* $i \neq j$ *such that the symmetric difference* $\mathcal{R}_i \triangle \mathcal{R}_j$ *satisfies*

$$|\mathcal{R}_i \triangle \mathcal{R}_j| = |\mathcal{R}_i - \mathcal{R}_j| + |\mathcal{R}_j - \mathcal{R}_i| > 6L - 2n. \tag{9}$$

Proof of Lemma 2. Clearly,

$$|\mathcal{R}_i| > L \tag{10}$$

holds for all but $\sum_{l=0}^{L} \binom{n}{l}$ subsets $\mathcal{R}_i$. Thus by (8), the number t of subsets $\mathcal{R}_i$ satisfying (10) is such that $t > \sum_{l=0}^{L} \binom{n}{l}$, so that t satisfies (6). Thus

Lemma 1 can be applied with these subsets, so that there are subsets $\mathscr{R}_i$, $\mathscr{R}_j$ satisfying (10) with $i \neq j$ and

$$|\mathscr{R}_i \cap \mathscr{R}_j| < k. \tag{11}$$

Then by (11) we have

$$\begin{aligned} |\mathscr{R}_i \triangle \mathscr{R}_j| &= (|\mathscr{R}_i| - |\mathscr{R}_i \cap \mathscr{R}_j|) + (|\mathscr{R}_j| - |\mathscr{R}_i \cap \mathscr{R}_j|) \\ &> (L-k) + (L-k) = 6L - 2n, \end{aligned}$$

which proves (9) and so completes the proof of Lemma 2.

4. The Proof of Theorem 1

Recall the definitions of $x = x(N), n = n(N)$ from Section 2. Let $Q = p_{x+1} p_{x+2} \cdots p_{x+n}$. For $i = 1, 2, \ldots, [N/Q] + 1$ put $\mathscr{A}_i = ((i-1)Q, iQ] \cap \mathscr{A}$. By the pigeon hole principle, there is an integer j with $1 \leqslant j \leqslant [N/Q] + 1$ and

$$|\mathscr{A}_j| \geqslant \frac{|\mathscr{A}|}{[N/Q]+1} \geqslant \frac{|\mathscr{A}|}{2N} Q, \tag{12}$$

where we use $Q \leqslant N$.

Recall the definition of $r(a, k)$ from Section 1. To any $a \in \mathscr{A}_j$, we assign the n-tuple $\mathbf{u}(a) = (r(a, p_{x+1}), \ldots, r(a, p_{x+n}))$. Note that by the Chinese remainder theorem, the mapping that sends $a \in \mathscr{A}_j$ to the vector $\mathbf{u}(a)$ is injective, so that if $\mathscr{U}$ is defined as the set of $\mathbf{u}(a)$ for $a \in \mathscr{A}_j$, then by (12)

$$|\mathscr{U}| = |\mathscr{A}_j| \geqslant \frac{|\mathscr{A}|}{2N} Q. \tag{13}$$

Let us call two vectors $\mathbf{u}(a), \mathbf{u}(a')$ in $\mathscr{U}$ *equivalent* if $r(a, p_{x+i}) = \pm r(a', p_{x+i})$ for each $i = 1, 2, \ldots, n$. This is clearly an equivalence relation on $\mathscr{U}$ and the number T of equivalence classes satisfies

$$T \leqslant \prod_{i=1}^{n} \frac{p_{x+i}+1}{2} = 2^{-n} Q \prod_{i=1}^{n} \left(1 + \frac{1}{p_{x+i}}\right) \leqslant 2^{-n} Q \exp\left(\sum_{i=1}^{n} \frac{1}{p_{x+i}}\right) \ll 2^{-n} Q.$$

Thus for some absolute constant c_2 we have

$$T < c_2 2^{-n} Q. \tag{14}$$

Let us write $\mathscr{U}_i$ for the vectors in $\mathscr{U}$ in the ith equivalence class, $i = 1, 2, ..., T$. By (13) and (14), there is a class $\mathscr{U}_v$ with

$$|\mathscr{U}_v| \geqslant \frac{|\mathscr{U}|}{T} > \frac{1}{c_2}\frac{|\mathscr{A}|}{2N}2^n. \tag{15}$$

Let $\mathscr{A}_{j,v}$ denote the set of $a \in \mathscr{A}_j$ with $\mathbf{u}(a) \in \mathscr{U}_v$ and for $a \in \mathscr{A}_{j,v}$, let

$$\mathscr{R}(a) = \{i: 1 \leqslant i \leqslant n, r_i(a) > 0\}.$$

The mapping that sends $a \in \mathscr{A}_{i,v}$ to $\mathscr{R}(a) \subset \{1, 2, ..., n\}$ is clearly injective, so that the number of such sets $\mathscr{R}(a)$ is $|\mathscr{A}_{j,v}| = |\mathscr{U}_v|$, which is estimated in (15).

Let c_0 in Theorem 1 be $4c_2$. Then by (3) and (15), the number of sets $\mathscr{R}(a)$ for $a \in \mathscr{A}_{j,v}$ exceeds

$$\frac{1}{c_2}\frac{|\mathscr{A}|}{2N}2^n > 2\sum_{l=0}^{L}\binom{n}{l}.$$

We thus may apply Lemma 2 with $\mathscr{R} = \{1, 2, ..., n\}$, $k = n - 2L$, and with the subsets $\mathscr{R}(a)$. We obtain that there are $a, a' \in \mathscr{A}_{j,v}$ with

$$|\mathscr{R}(a) \triangle \mathscr{R}(a')| > 6L - 2n. \tag{16}$$

But if $i \in \mathscr{R}(a) \triangle \mathscr{R}(a')$, then $r(a, p_{x+i}) + r(a', p_{x+i}) = 0$; that is, $p_{x+i} / a + a'$. Thus

$$v(a + a') \geqslant |\mathscr{R}(a) \triangle \mathscr{R}(a')| > 6L - 2n,$$

which proves (4) and thus completes the proof of Theorem 1.

5. The Case $\mathscr{A} \neq \mathscr{B}$

We now generalize the statement in Theorem 1 by considering the case of two sets of integers $\mathscr{A}, \mathscr{B}$ that are not necessarily equal. This can be done by combining Theorem 1 with Lemma 3 below.

For a set of integers $\mathscr{A}$, we write $2 \times \mathscr{A}$ for the set of numbers $2a$ with $a \in \mathscr{A}$.

Lemma 3. *Let N be a positive integer and let $\mathscr{A}, \mathscr{B}$ be nonempty subsets of $\{1, 2, ..., N\}$. Then there is a set $\mathscr{D}$ such that*

(i) $\mathscr{D} \subset \{1, 2, ..., 2N\}$,

(ii) $\mathscr{D} + \mathscr{D} \subset 2 \times (\mathscr{A} + \mathscr{B})$, *and*

(iii) $|\mathscr{D}| > |\mathscr{A}|\,|\mathscr{B}|/2N$.

Proof. For each integer y, let $f(y)$ denote the number of pairs $(a, b) \in \mathscr{A} \times \mathscr{B}$ with $a - b = y$. Then clearly we have

$$\sum_{y=-N+1}^{N-1} f(y) = |\mathscr{A} \times \mathscr{B}| = |\mathscr{A}|\,|\mathscr{B}|,$$

so that there is an integer y_0 with $-N < y_0 < N$ for which

$$f(y_0) \geqslant \frac{1}{2N-1}\,|\mathscr{A}|\,|\mathscr{B}|. \tag{17}$$

Let $\mathscr{D}$ denote the set of integers d that can be written in the form

$$d = 2a - y_0 = 2b + y_0 = a + b \qquad (a \in \mathscr{A},\ b \in \mathscr{B}). \tag{18}$$

Thus (i) holds trivially, while (iii) holds by (17). Finally, let $d, d' \in \mathscr{D}$, so that there are $a \in \mathscr{A}$, $b' \in \mathscr{B}$ with $d = 2a - y_0$, $d' = 2b' + y_0$. Thus

$$d + d' = 2(a + b') \in 2 \times (\mathscr{A} + \mathscr{B}),$$

which proves that also (ii) holds and thus completes the proof of Lemma 3.

From Lemma 3 and Theorem 1, we easily deduce the following.

COROLLARY 3. *The constants c_0, N_0 of Theorem 1 and the function $n = n(N)$ of Section 2 have the following property. Suppose N is an integer with $N > N_0$, $\mathscr{A}$ and $\mathscr{B}$ are subsets of $\{1, 2, \ldots, [N/2]\}$, L is an integer with $L < n/2$ and*

$$\frac{|\mathscr{A}|\,|\mathscr{B}|}{N^2} > c_0 2^{-n} \sum_{l=0}^{L} \binom{n}{l}.$$

Then there exist integers $a \in \mathscr{A}$, $b \in \mathscr{B}$ with $v(a + b) > 6L - 2n - 1$.

6. AN EXAMPLE

We now give an example to show that Corollary 1 and thus also Theorem 1 are best possible apart from a factor $\log m$ in the secondary term. We are not sure which result is closer to the "truth." The first two of us vote for Corollary 1, while the latter two vote for Theorem 2 below. So even if such matters were decided by votes, things would still be inconclusive.

THEOREM 2. *There are effectively computable positive constants c_3, c_4, N_2 such that for each integer $N > N_2$ there is a set $\mathscr{A} \subset \{1, 2, ..., N\}$ with*

$$|\mathscr{A}| > c_3 N \tag{19}$$

and for all $a, a' \in \mathscr{A}$ we have

$$v(a + a') < m - c_4 \frac{\sqrt{m}}{\log m}. \tag{20}$$

Proof. Put $t = [m/3]$ and let $Q = p_{t+1} p_{t+2} \cdots p_{2t}$. Let $\mathscr{K}(a)$ denote the set of i for which $r(a, p_{t+i}) > 0$. Note that for any set $\mathscr{S} \subset \{t+1, t+2, ..., 2t\}$, the number of integers a with $0 < a \leqslant N$ and $\mathscr{K}(a) = \mathscr{S}$ is

$$N \prod_{i \in \mathscr{S}} \frac{p_i - 1}{2p_i} \prod_{i \in [t+1, 2t] - \mathscr{S}} \frac{p_i + 1}{2p_i} + \theta(\mathscr{S})\, Q,$$

where $\theta(\mathscr{S})$ is some number of absolute value at most 1. Thus the number of these integers a is

$$2^{-t} N(1 + O(1/\log m)) + \theta(\mathscr{S})\, Q = 2^{-t} N(1 + O(1/\log m))$$

uniformly for all large N and all subsets $\mathscr{S} \subset \{t+1, ..., 2t\}$.

Let $\mathscr{A}$ denote the set of all integers a with $0 < a \leqslant N$ and

$$|\mathscr{K}(a)| > \frac{t}{2} + \sqrt{t}. \tag{21}$$

Since the number of subsets $\mathscr{S}$ of $\{t+1, ..., 2t\}$ with $|\mathscr{S}| > t/2 + \sqrt{t}$ is $\gg 2^t$, it follows from the above that $|\mathscr{A}| \gg N$; that is, (19) holds.

Assume now that $a, a' \in \mathscr{A}$. From (21) it is clear we have

$$|\mathscr{K}(a) \cap \mathscr{K}(a')| > 2\sqrt{t}. \tag{22}$$

Further, if $i \in \mathscr{K}(a) \cap \mathscr{K}(a')$, then $0 < r(a, p_{t+i}) + r(a', p_{t+i}) < p_{t+i}$, so that

$$a + a' \equiv r(a, p_{t+i}) + r(a', p_{t+i}) \not\equiv 0 \bmod p_{t+i}.$$

Thus by (22) we have

$$u = u(a + a') := |\{j \leqslant 2t: p_j \nmid a + a'\}| > 2\sqrt{t}. \tag{23}$$

If $a, a' \in \mathscr{A}$ and if $v = v(a + a')$, then

$$2N \geqslant a + a' \geqslant p_{2t}^{-u} \prod_{j=1}^{v+u} p_j \geqslant p_{2t}^{-u} p_{m+1}^{v+u-m-1} \prod_{j=1}^{m+1} p_j$$

$$> N \left(\frac{p_{m+1}}{p_{2t}} \right)^u p_{m+1}^{v-m-1},$$

so that

$$p_{m+1}^{m+1-v} > \frac{1}{2} \left(\frac{p_{m+1}}{p_{2t}} \right)^u. \tag{24}$$

For all large N we have $p_{m+1}/p_{2t} > 3/2$, so that with (23) and (24) we have

$$m + 1 - v > \frac{-\log 2 + u \log(3/2)}{\log p_{m+1}} > \frac{-\log 2 + 2\sqrt{[m/3]} \log(3/2)}{\log m}.$$

Fix some positive number $c_4 < (2/\sqrt{3}) \log(3/2)$. Then for all large N and all $a, a' \in \mathscr{A}$, we have (20). This completes the proof of Theorem 2.

7. The Case of Ω

In this section we show the following result. The proof is very much like that of Theorem 1, so that we only sketch it.

THEOREM 3. *There exist effectively computable positive constants c_5, c_6, N_3 such that if $N > N_3$ is an integer and $\mathscr{A} \subset \{1, 2, ..., N\}$ with*

$$|\mathscr{A}| > N \exp\left(-c_5 \frac{\log N}{\log \log N} \right), \tag{25}$$

then there are integers $a, a' \in \mathscr{A}$ with

$$\Omega(a + a') > c_6 \frac{\log N}{\log \log(2 + N/|\mathscr{A}|)}. \tag{26}$$

Proof. Write $s = [c_7 \log(2N/|\mathscr{A}|)]$, where c_7 is an effectively computable positive constant to be specified later. If N_3 is sufficiently large and $c_5 = 1/(3c_7)$, then (25) implies that $\prod_{i=1}^{s} p_i \leqslant \sqrt{N}$. Let

$$\alpha = \left[\frac{\log N}{\log(p_1 p_2 \cdots p_s)} \right], \qquad Q = \left(\prod_{i=1}^{s} p_i \right)^{\alpha},$$

so that $\alpha \geqslant 2$ and $Q \leqslant N$. As in the proof of Theorem 1, there is an integer j such that if $\mathscr{A}_j = ((j-1)Q, jQ] \cap \mathscr{A}$, then (12) holds.

For $a \in \mathscr{A}_j$, let $\mathbf{u}(a) = (r(a, p_1^\alpha), \ldots, r(a, p_s^\alpha))$. As before, we say two such vectors are equivalent if the corresponding coordinates are equal in absolute value. The number T of equivalence classes satisfies

$$T \leqslant (2^{\alpha-1}+1)\prod_{i=2}^{s}\frac{p_i^\alpha+1}{2} = 2^{-s}Q\left(1+\frac{2}{2^\alpha}\right)\prod_{i=2}^{s}\left(1+\frac{1}{p_i^\alpha}\right) \ll 2^{-s}Q,$$

since $\alpha \geqslant 2$. Thus the set $\mathscr{A}_{j,y}$ of elements in $\mathscr{A}_j$ which map to the largest equivalence class has cardinality

$$|\mathscr{A}_{j,y}| \geqslant \frac{|\mathscr{A}_j|}{T} \gg \frac{|\mathscr{A}_j|\,2^s}{Q} \gg \frac{|\mathscr{A}|\,2^s}{N}, \tag{27}$$

where we use (12) for the last inequality. From the definition of s above, we have $2^{0.02s} > N/|\mathscr{A}|$ if c_7 is chosen large enough. Thus from (27), $|\mathscr{A}_{j,y}| \gg 2^{0.98s}$, so that if c_7 is chosen sufficiently large, we have

$$|\mathscr{A}_{j,y}| > 2\sum_{l=0}^{L}\binom{s}{l},$$

where $L = [0.4s] + 1$.

Thus as in the proof of Theorem 1, we may use Lemma 2 to show there are $a, a' \in \mathscr{A}_{j,y}$ such that $r(a, p_i^\alpha) + r(a', p_i^\alpha) = 0$ holds for at least $6L - 2s > 0.4s$ distinct values of i. For each such i we have $p_i^\alpha \mid a + a'$, so that

$$\Omega(a+a') > 0.4s\alpha \gg s\frac{\log N}{\log(p_1 p_2 \cdots p_s)} \gg \frac{\log N}{\log s} \gg \frac{\log N}{\log\log(2+N/|\mathscr{A}|)}.$$

This completes the proof of Theorem 3.

We remark that by the use of Lemma 3, it is a simple matter to prove a version of Theorem 3 that holds for two sets of integers $\mathscr{A}, \mathscr{B}$ as in Corollary 3.

In addition, we remark that it is easy to see that apart from the constant, (26) is best possible. Indeed, if $S < 0.9 \log N$ and $\mathscr{A}$ is defined as the set of integers $a \in \{1, \ldots, N\}$ with $r(a, 4) > 0$ and $r(a, p) > 0$ for each odd prime $p \leqslant S$, then $|\mathscr{A}| \gg N/2^S$. Further, if $a, a' \in \mathscr{A}$ then $4 \nmid a + a'$ and $p \mid a + a'$ for each odd prime $p \leqslant S$, so that $\Omega(a+a') \leqslant 1 + \log N/\log S$. Thus we see that (26) is best possible, apart from the choice of c_6.

The maximal order of Ω on the interval $[1, 2N]$ is easy to compute; it is $\log(2N)/\log 2$. One may ask how dense must $\mathscr{A}$ be for there to be $a, a' \in \mathscr{A}$ with $\Omega(a+a') \sim \log N/\log 2$. Certainly density $1/4$ is not

sufficient, since if we take for $\mathscr{A}$ the integers up to N that are 1 mod 4, then $\Omega(a+a') \leqslant 1 + \log N/\log 3$ for all $a, a' \in \mathscr{A}$. It is interesting that for some values of N, density 1/2 is sufficient and for other values it is not. For example, if N is a power of 2, say $N = 2^k$, and if $\mathscr{A} \subset \{1, ..., N\}$ with $|\mathscr{A}| = N/2$, then either $N \in \mathscr{A}$ or there are $a, a' \in \mathscr{A}$ with $a + a' = N$. In either case, we get at least $\log N/\log 2$ for the maximal value of Ω on $\mathscr{A} + \mathscr{A}$.

But now say $N = 2^{2k} + 2^{k+1}$ for some integer $k \geqslant 2$ and $\mathscr{A}$ is the set of integers $a \in [1, N]$ with $r(a, 2^{k+2}) > 0$. Then $|\mathscr{A}| = (2^{k-2} + 1)(2^{k+1} - 1) > N/2$. But for all $a, a' \in \mathscr{A}$ we have $a + a' \not\equiv 0 \bmod 2^{k+2}$. Thus $\Omega(a+a') \leqslant k + 1 + \log(N/2^k)/\log 3$. It is not hard to show that something close to this upper bound actually does occur for this example, so that as $N \to \infty$ through this sequence of numbers and $\mathscr{A}$ is as just constructed, then the maximal order of Ω on $\mathscr{A} + \mathscr{A}$ is $(c + o(1)) \log N$ where

$$c = \frac{1}{2\log 2} + \frac{1}{2\log 3} < \frac{1}{\log 2}.$$

Generalizing these thoughts, we have the following two results.

THEOREM 4. *Suppose $\varepsilon > 0$, N is a positive integer, and $\mathscr{A} \subset \{1, ..., N\}$ with $|\mathscr{A}| \geqslant (N + N^{1-\varepsilon})/2$. Then there are $a, a' \in \mathscr{A}$ with $\Omega(a+a') > (1-\varepsilon)\log N/\log 2$.*

THEOREM 5. *Let $0 < \varepsilon \leqslant 1/2$ be arbitrary, but fixed. There are infinitely many positive integers N for which there is a set $\mathscr{A} \subset \{1, ..., N\}$ with $|\mathscr{A}| > (N + N^{1-\varepsilon})/2$ and*

$$\max_{a, a' \in \mathscr{A}} \Omega(a+a') \leqslant \left(\frac{1-\varepsilon}{\log 2} + \frac{\varepsilon}{\log 3} + o(1)\right) \log N,$$

for N running through the infinite set asserted to exist.

As with the gap between Corollary 1 and Theorem 2, we are not sure which of Theorems 4 and 5 is closer to the truth. It is perhaps interesting to note the following corollary of Theorem 4: For each $\delta > 0$, there is a number c_δ, such that if N is a positive integer and $\mathscr{A} \subset \{1, ..., N\}$ with $|\mathscr{A}| > (1+\delta) N/2$, then there are $a, a' \in \mathscr{A}$ with $v(a+a') > -c_\delta + \log N/\log 2$. In fact, $c_\delta = |\log \delta|/\log 2$.

ACKNOWLEDGMENTS

The first author would like to acknowledge the hospitality of the University of Georgia where some of the work on this paper was done. The second author was supported in part

by an NSF grant. In addition he would like to acknowledge the hospitality of the University of Waterloo and Macquarie University where some of the work on this paper was done. The third author would like to acknowledge the hospitality of the University of Waterloo where some of the work on this paper was done. The work of the fourth author was supported in part by a Killam Research Fellowship and by Grant A3528 from the Natural Sciences and Engineering Council of Canada.

REFERENCES

1. A. BALOG AND A. SÁRKÖZY, On sums of sequences of integers, I, *Acta Arith.* **44** (1984), 73–86.
2. P. D. T. A. ELLIOTT AND A. SÁRKÖZY, The distribution of the number of prime factors of sums $a+b$, *J. Number Theory* **29** (1988), 94–99.
3. P. ERDŐS, H. MAIER, AND A. SÁRKÖZY, On the distribution of the number of prime factors of sums $a+b$, *Trans. Amer. Math. Soc.* **302** (1987), 269–280.
4. GY. KATONA, Intersection theorems for systems of finite sets, *Acta Math. Acad. Sci. Hung.* **15** (1964), 329–337.
5. C. POMERANCE, A. SÁRKÖZY, AND C. L. STEWART, On divisors of sums of integers, III, *Pacific J. Math.* **133** (1988), 363–379.
6. A. SÁRKÖZY AND C. L. STEWART, On divisors of sums of integers, II, *J. Reine Angew. Math.* **365** (1986), 171–191.