*Acta Math. Hung.*
**48 (1—2)** (1986), 147—154.

# ON DIVISORS OF SUMS OF INTEGERS. I

A. SÁRKÖZY (Budapest) and C. L. STEWART* (Waterloo)

§ 1. **Introduction.** Let $N$ be a positive integer and let $A_1, \ldots, A_k$ be non-empty subsets of $\{1, \ldots, N\}$. Let $|A_i|$ denote the cardinality of $A_i$. For any integer $n$ larger than one let $P(n)$ denote the greatest prime factor of $n$. In [1], Balog and Sárközy proved, by means of the large sieve inequality, that if $|A_1||A_2| > 100N(\log N)^2$ and $N$ is sufficiently large then there exist $a_1 \in A_1$ and $a_2 \in A_2$ such that

$$P(a_1+a_2) > \frac{1}{16}\frac{(|A_1||A_2|)^{1/2}}{\log N}.$$

In the same article they obtained a slightly weaker result by means of the Hardy—Littlewood method. We propose to employ the Hardy—Littlewood method in connection with this problem in a sequel to this article. However, the purpose of this note is to estimate $P(a_1+\ldots+a_k)$ where $a_1, \ldots, a_k$ are chosen from the $k$ sets $A_1, \ldots, A_k$ respectively. Put

$$T = \Big(\prod_{j=1}^{k} |A_i|\Big)^{1/k}.$$

THEOREM. *Let $A_1, \ldots, A_k$ be non-empty subsets of $\{1, \ldots, N\}$ with $|A_1| = \min_i |A_i|$ and $k > 1$, and let $\varepsilon$ be a positive real number. If*

$$\sum_{i=1}^{k} |A_i| > (1+\varepsilon)N,$$

*then for any prime $p$ with $N < p < \left(1+\frac{\varepsilon}{2}\right)N$, there exist $a_i \in A_i$, for $i = 1, \ldots, k$, such that*

$$P(a_1+\ldots+a_k) = p, \tag{1}$$

*whenever $N > N_0(\varepsilon, k)$. If $T > 8N^{1/2}\log N$, then there exist $a_i \in A_i$, for $i = 1, \ldots, k$, such that*

$$P(a_1+\ldots+a_k) > \frac{kT}{14\log T}, \tag{2}$$

* This research was supported in part by Grant A3528 from the Natural Sciences and Engineering Research Council of Canada.

*for* $N > N_1(k)$. *Further, there exist* $a_i \in A_i$, *for* $i=1, \ldots, k$, *such that*

$$P(a_1+\ldots+a_k) > \frac{|A_1|}{N^{1/k+\varepsilon}}, \tag{3}$$

*for* $N > N_2(\varepsilon, k)$. *Here* $N_0(\varepsilon, k)$, $N_1(k)$ *and* $N_2(\varepsilon, k)$ *are numbers which are effectively computable in terms of* $\varepsilon$ *and* $k$, $k$, *and* $\varepsilon$ *and* $k$ *respectively.*

To prove (1) we appeal to the Cauchy—Davenport Lemma. Note that we are able to specify the greatest prime factor of $a_1+\ldots+a_k$ in this case. For the proof of (2) we use the large sieve inequality in conjunction with the Cauchy—Davenport Lemma. If $k=2$ then (2) yields the result of Balog and Sárközy referred to above. Finally, (3) is obtained using the Cauchy—Davenport Lemma and Gallagher's larger sieve.

In the following result, which is an immediate consequence of our theorem, we require that all the summands be taken from a single set.

COROLLARY. *Let* $A$ *be a non-empty subset of* $\{1, \ldots, N\}$, *let* $\varepsilon$ *be a positive real number and let* $k$ *be an integer larger than one. If* $|A| > (1+\varepsilon)N/k$ *and* $p$ *is any prime number with* $N < p < \left(1+\frac{\varepsilon}{2}\right)N$ *then there exist* $a_1, \ldots, a_k$ *in* $A$ *such that*

$$P(a_1+\ldots+a_k) = p,$$

*for* $N$ *sufficiently large in terms of* $\varepsilon$ *and* $k$. *Further, if* $|A| > 8N^{1/2}\log N$ *then there exist* $a_1, \ldots, a_k$ *in* $A$ *such that*

$$P(a_1+\ldots+a_k) > \frac{k|A|}{14\log|A|},$$

*for* $N$ *sufficiently large in terms of* $k$. *Furthermore, there exist* $a_1, \ldots, a_k$ *in* $A$ *such that*

$$P(a_1+\ldots+a_k) > \frac{|A|}{N^{1/k+\varepsilon}},$$

*for* $N$ *sufficiently large in terms of* $\varepsilon$ *and* $k$.

It would be interesting if one could obtain results of comparable strength to the above for subsets of $\{1, \ldots, N\}$ of cardinality less than $N^{1/k}$. The only result of which we are aware in this connection is due to Erdős and Turán [4]. They showed, in 1934, by means of an elementary argument that for any finite set of positive integers $A$ there exist integers $a_1$ and $a_2$ from $A$ such that

$$P(a_1+a_2) > c\log|A|,$$

for a positive constant $c$.

**§2. Preliminary lemmas.** Let $\mathbf{Z}$ denote the set of integers.

LEMMA 1 (Cauchy—Davenport [2], [3]). *Let* $p$ *be a prime number and let* $A$ *and* $B$ *be a subsets of* $\mathbf{Z}/p\mathbf{Z}$. *If* $|A|=m$ *and* $|B|=n$ *then* $|A+B| \geqq \min\{m+n-1, p\}$; *here* $A+B=\{a+b \mid a\in A, b\in B\}$.

LEMMA 2 (large sieve). *Let $\mathcal{N}$ be a set of integers in the interval $[M+1, M+N]$. For each prime p let $\nu(p)$ denote the number of residue classes modulo p that contain an element of $\mathcal{N}$. Then for any positive integer Q we have*

$$|\mathcal{N}| \leqq \frac{N+Q^2}{L}, \quad for \quad L = \sum_{q\leqq Q}' \prod_{p|q} \frac{p-\nu(p)}{\nu(p)},$$

*where the summation is taken over square-free positive integers q.*

PROOF. See Theorem 7.1 of [8].

LEMMA 3 (Gallagher [5]). *Let $\mathcal{N}$ be a set of integers in the interval $[M+1, M+N]$. For each prime p let $\nu(p)$ denote the number of residue classes modulo p that contain an element of $\mathcal{N}$. Then for any finite set of primes S we have*

$$|\mathcal{N}| \leqq \frac{\sum_{p\in S} \log p - \log N}{\sum_{p\in S} \frac{\log p}{\nu(p)} - \log N},$$

*provided that the denominator is positive.*

We shall also require the following result.

LEMMA 4. *Let p and k be integers with $k \geqq 2$ and $p-1 \geqq (k-1)^k$. Let $D = \left\{(x_1, \dots, x_k) \in \mathbf{R}^k \,|\, x_1+\dots+x_k \leqq 1+\frac{k-2}{p} \text{ and } \frac{1}{p} \leqq x_i \leqq \frac{p-1}{p} \text{ for } i=1, \dots, k\right\}$. Then*

$$\min_D \prod_{i=1}^k \left(\frac{1}{x_i}-1\right) = \left(\frac{k}{1+\frac{k-2}{p}}-1\right)^k \tag{4}$$

*and*

$$\min_D \sum_{i=1}^k \frac{1}{x_i} = \frac{k^2}{1+\frac{k-2}{p}}. \tag{5}$$

PROOF. First we shall establish (4) by induction on $k$. It is readily checked that (4) holds for $k=2$ and so we may assume that $k>2$. Our inductive hypothesis is that (4) holds with $k-1$ in place of $k$. We observe that the minimum of $\prod_{i=1}^k \left(\frac{1}{x_i}-1\right)$ in $D$, occurs in $D_0$ where $D_0 = \left\{(x_1, \dots, x_k) \in D \,|\, x_1+\dots+x_k = 1+\frac{k-2}{p}\right\}$. Note also that $\prod_{i=1}^k \left(\frac{1}{x_i}-1\right)$ and $\sum_{i=1}^k \log\left(\frac{1}{x_i}-1\right)$ achieve their minimum value in $D_0$ at the same points. Applying the method of Lagrange multipliers we conclude that if $\sum_{i=1}^k \log\left(\frac{1}{x_i}-1\right)$ has a local minimum at $(x_1, \dots, x_k)$ in the interior of $D_0$ then for all integers $i$ and $j$ with $1 \leqq i, j \leqq k$ either $x_i = x_j$ or $x_i = 1 - x_j$. If $x_i = x_j$ for all $i$ and $j$ then

$$\prod_{i=1}^{k} \left(\frac{1}{x_i} - 1\right) = \left(\frac{k}{1 + \frac{k-2}{p}} - 1\right)^k.$$

On the other hand if $x_i = 1 - x_j$ for some $i$ and $j$ then $x_i + x_j = 1$ and by the definition of $D_0$ we have $x_l = \frac{1}{p}$ for some integer $l$. Similarly if $(x_1, \dots, x_k)$ is on the boundary of $D_0$ then we again have $x_l = \frac{1}{p}$ for some integer $l$. However, if $x_l = \frac{1}{p}$ and $(x_1, \dots, x_k)$ is in $D_0$ then

$$\prod_{i=1}^{k} \left(\frac{1}{x_i} - 1\right) = (p-1) \prod_{\substack{i=1 \\ i \neq l}}^{k} \left(\frac{1}{x_i} - 1\right) \geqq (p-1) \min_{D'} \prod_{\substack{i=1 \\ i \neq l}}^{k} \left(\frac{1}{x_i} - 1\right),$$

where

$$D' = \{(x_1, \dots, x_{l-1}, x_{l+1}, \dots, x_k) \in \mathbf{R}^{k-1} \mid x_1 + \dots + x_{l-1} + x_{l+1} + \dots + x_k \leqq 1 + \frac{k-3}{p}$$

$$\text{and} \quad \frac{1}{p} \leqq x_i \leqq \frac{p-1}{p} \quad \text{for} \quad i = 1, \dots, l-1, l+1, \dots, k\}.$$

By our inductive hypothesis the minimum over $D'$ of $\prod_{\substack{i=1 \\ i \neq l}}^{k} \left(\frac{1}{x_i} - 1\right)$ is $\left(\frac{k-1}{1 + \frac{k-3}{p}} - 1\right)^{k-1}$, which is at least 1 since $k > 2$. Therefore if $(x_1, \dots, x_k)$ is a point in $D_0$ with $x_l = \frac{1}{p}$ then

$$\prod_{i=1}^{k} \left(\frac{1}{x_i} - 1\right) \geqq p - 1 \geqq (k-1)^k > \left(\frac{k}{1 + \frac{k-2}{p}} - 1\right)^k,$$

consequently the minimum of $\prod_{i=1}^{k} \left(\frac{1}{x_i} - 1\right)$ on $D$ occurs with

$$x_1 = \dots = x_k = \frac{1 + \frac{k-2}{p}}{k}.$$

Thus (4) holds and this completes the induction.

To establish (5) requires only a routine application of the method of Lagrange multipliers. Alternatively (5) can be deduced from the arithmetic-harmonic mean inequality.

§3. **Proof of the main theorem.** We shall first prove (1). Let $p$ be a prime with $N<p<\left(1+\frac{\varepsilon}{2}\right)N$. Assume that $N>\max\left\{\frac{2(k-1)}{\varepsilon}, k\right\}$ and put $A_i(p)=\{a+p\mathbf{Z}|a\in A_i\}$ for $i=1, \ldots, k$. By repeated application of Lemma 1 we find that

$$|A_1(p)+\ldots+A_k(p)| \geqq \min\left\{\sum_{i=1}^{k}|A_i(p)|-(k-1), p\right\}. \tag{6}$$

Since $A_i\subseteqq\{1, \ldots, N\}$ and $p>N$, $|A_i(p)|=|A_i|$. Therefore

$$\sum_{i=1}^{k}|A_i(p)| > (1+\varepsilon)N$$

and, since $\frac{\varepsilon}{2}N>k-1$, the minimum on the right hand side of (6) is $p$. Accordingly, $|A_1(p)+\ldots+A_k(p)|=p$, hence $A_1(p)+\ldots+A_k(p)=\mathbf{Z}/p\mathbf{Z}$. Therefore there exist $a_i\in A_i$, for $i=1, \ldots, k$, with $p|a_1+\ldots+a_k$. Since $a_1+\ldots+a_k\leqq kN$, $k<N$ and $p>N$, $P(a_1+\ldots+a_k)=p$ as required.

To prove (2) we assume that $T\geqq 8N^{1/2}\log N$ and we put $Q=\frac{kT}{7\log T}$. Further, we shall suppose that $N$ is chosen sufficiently large for the subsequent argument; in particular, large enough that $\frac{Q}{2}>(k-1)^k$, $T^{1/7}>k$ and $N^{1/2}>8\log N$. We shall now show that the assumption that $P(a_1+\ldots+a_k)<\frac{Q}{2}$ whenever $a_i\in A_i$, $i=1, \ldots, k$, leads to a contradiction and this will establish (2).

Applying Lemma 2 with $M=0$, we find that

$$|A_i| < \frac{N+Q^2}{\sum\limits_{Q/2<p<Q}\frac{p-\nu_i(p)}{\nu_i(p)}},$$

where the summation in the denominator is taken over primes $p$ and where $\nu_i(p)$ is the number of residue classes modulo $p$ that contain an element of $A_i$. Thus

$$T < \frac{N+Q^2}{H}, \tag{7}$$

where

$$H = \left(\prod_{i=1}^{k}\sum_{Q/2<p<Q}\frac{p-\nu_i(p)}{\nu_i(p)}\right)^{1/k}.$$

By a generalization of the Cauchy—Schwarz inequality (see 81.3, page 68 of [7]),

$$H \geqq \sum_{Q/2<p<Q}\left(\prod_{i=1}^{k}\frac{p-\nu_i(p)}{\nu_i(p)}\right)^{1/k}. \tag{8}$$

Define $A_i(p)$ as above and notice that, by Lemma 1, we again obtain (6). However, for each prime $p$ with $\frac{Q}{2}<p<Q$, $A_1(p)+\ldots+A_k(p)$ does not contain the zero

residue class hence $|A_1(p)+\ldots+A_k(p)|\leqq p-1$. Further, $\nu_i(p)=|A_i(p)|$ and therefore

(9) $$\nu_1(p)+\ldots+\nu_k(p)\leqq p+k-2.$$

Certainly $1\leqq \nu_i(p)\leqq p-1$ and thus putting $\frac{\nu_i(p)}{p}=x_i$ and applying (4) of Lemma 4 we find, since $p>\frac{Q}{2}>(k-1)^k$, that

(10) $$\left(\prod_{i=1}^{k}\left(\frac{p}{\nu_i(p)}-1\right)\right)^{1/k}\geqq \frac{k}{1+\frac{k-2}{p}}-1\geqq \frac{k}{2}.$$

By the prime number theorem,

(11) $$\sum_{Q/2<p<Q}\frac{k}{2}>\frac{kQ}{5\log Q},$$

for $N$ sufficiently large. Combining (7), (8) (10) and (11) we obtain

$$T<\frac{N+Q^2}{\frac{kQ}{5\log Q}}.$$

By assumption $N^{1/2}>8\log N$ and so $N<\frac{1}{5}Q^2$. Thus

$$\frac{kT}{6}<Q\log Q\leqq \frac{kT}{7\log T}\log kT,$$

and, since $T^{1/7}>k$,

$$\frac{kT}{7\log T}\log kT<\frac{kT}{6}.$$

This gives the required contradiction.

Finally, we shall prove (3). We may assume without loss of generality that $\varepsilon$ is less than one. Put

$$Q=\frac{|A_1|}{N^{1/k+\varepsilon/2}}\quad\text{and}\quad Q_1=|A_1|.$$

We shall assume that $N$ is sufficiently large in terms of $\varepsilon$ and $k$ for the validity of the argument to follow. Further we shall assume that $Q>N^{\varepsilon/2}$ and that $P(a_1+\ldots+a_k)\leqq Q$, whenever $a_i\in A_i$, $i=1,\ldots,k$. Let $\nu_i(p)$ denote the number of residue classes modulo $p$ that contain an element of $A_i$. By Lemma 3,

(12) $$|A_i|\leqq\frac{\sum_{Q<p<Q_1}\log p-\log N}{\sum_{Q<p<Q_1}\frac{\log p}{\nu_i(p)}-\log N},$$

for $i=1, \ldots, k$, whenever the denominator is positive; here the summations are taken over all primes $p$ between $Q$ and $Q_1$. We shall show that for at least one integer $i$ the denominator in (12) is at least $\frac{\varepsilon}{2}\log N$. As before we find that (9) holds for each prime $p$ with $Q<p<Q_1$. Since $1\leqq v_i(p)\leqq p-1$, on putting $\frac{v_i(p)}{p}=x_i$ and applying (5) of Lemma 4 we find that

$$\frac{1}{k}\sum_{i=1}^{k}\frac{\log p}{v_i(p)}\geqq\frac{\log p}{p}\frac{k}{1+\frac{k-2}{p}}.$$

Since $p>Q\geqq N^{\varepsilon/2}$,

$$\frac{k}{1+\frac{k-2}{p}}>k-\frac{\varepsilon}{8}$$

for $N$ sufficiently large. Thus

$$\frac{1}{k}\sum_{i=1}^{k}\left(\sum_{Q<p<Q_1}\frac{\log p}{v_i(p)}-\log N\right)=\sum_{Q<p<Q_1}\left(\frac{1}{k}\sum_{i=1}^{k}\frac{\log p}{v_i(p)}\right)-\log N\geqq$$

$$\geqq\sum_{Q<p<Q_1}\left(k-\frac{\varepsilon}{8}\right)\frac{\log p}{p}-\log N.$$

By Theorem 425 of [6] there is a constant $C$ such that the right hand side of the above inequality is

$$\geqq\left(k-\frac{\varepsilon}{8}\right)(\log Q_1-\log Q-C)-\log N\geqq\left(k-\frac{\varepsilon}{4}\right)\log(Q_1/Q)-\log N.$$

This in turn is $>\frac{\varepsilon}{2}\log N$, since $k\geqq 2$, $\varepsilon<1$ and $Q_1/Q=N^{1/k+\varepsilon/2}$. Since the average of the denominators in (12) is at least $\varepsilon/2\log N$, for at least one integer $i$,

$$|A_i|\leqq\frac{\sum_{Q<p<Q_1}\log p-\log N}{\frac{\varepsilon}{2}\log N}.$$

Hence, by the prime number theorem,

$$|A_i|\leqq\frac{4Q_1}{\varepsilon\log N}=\frac{4|A_1|}{\varepsilon\log N},$$

for $N$ sufficiently large. But $|A_1|\leqq|A_i|$ and so we have a contradiction for $N$ sufficiently large. Therefore either $P(a_1+\ldots+a_k)>Q$ for some $a_1\in A_i$, $i=1, \ldots, k$ or $Q<N^{\varepsilon/2}$. Consequently, for some $a_i\in A_i$, $i=1, \ldots, k$,

$$P(a_1+\ldots+a_k)>\frac{Q}{N^{\varepsilon/2}}=\frac{|A_i|}{N^{1/k+\varepsilon}},$$

as required.

## References

[1] A. Balog and A. Sárközy, *On sums of sequences of integers*, II, to appear.
[2] H. Davenport, On the addition of residue classes, *J. London Math. Soc.*, **10** (1935), 30—32.
[3] H. Davenport, A historical note, *J. London Math. Soc.*, **22** (1947), 100—101.
[4] P. Erdős and P. Turán, On a problem in the elementary theory of numbers, *American Math. Monthly*, **41** (1934), 608—611.
[5] P. X. Gallagher, A larger sieve, *Acta Arith.*, **18** (1971), 77—81.
[6] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed. (Oxford, 1979).
[7] G. Pólya and G. Szegő, *Problems and theorems in analysis*, Vol. I, Springer-Verlag (Berlin, 1972).
[8] H. E. Richert, *Lectures on sieve methods*, Tata Institute of Fundamental Research (Bombay, 1976).

*(Received August 2, 1984)*

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
1053 BUDAPEST, REÁLTANODA U. 13—15
HUNGARY

DEPARTMENT OF PURE MATHEMATICS
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO
CANADA
N2L 3G1