ACTA ARITHMETICA
LXX.1 (1995)

On addition of two distinct sets of integers

by

Vsevolod F. Lev (Tel-Aviv) and
Pavel Y. Smeliansky (Wollongong, N.S.W.)

What is the structure of a pair of finite integers sets $A, B \subset \mathbb{Z}$ with the small value of $|A+B|$? We answer this question for addition coefficient 3. The obtained theorem sharpens the corresponding results of G. Freiman.

1. Introduction and historical comments. Let $A = \{a_1, \ldots, a_k\}$, $B = \{b_1, \ldots, b_l\}$ be two sets of integers, so that $k = |A|$, $l = |B|$, and suppose $0 = a_1 < \ldots < a_k$, $0 = b_1 < \ldots < b_l$. As usual, we write $A + B$ for the set $\{a_i + b_j \mid 1 \le i \le k,\ 1 \le j \le l\}$, and put $2A = A + A$. By $(a_1, \ldots, a_k)$ we denote the greatest common divisor of $a_1, \ldots, a_k$, and by $(a_1, \ldots, b_l)$ the greatest common divisor of $a_1, \ldots, a_k, b_1, \ldots, b_l$.

In [1] G. Freiman proved the following:

Theorem 1. (i) *Let* $a_k \le 2k - 3$. *Then* $|2A| \ge a_k + k$.

(ii) *Let* $a_k \ge 2k - 2$ *and* $(a_1, \ldots, a_k) = 1$. *Then* $|2A| \ge 3k - 3$.

The present paper is devoted to the generalization of this theorem to the case of summation of two distinct sets A and B. Without loss of generality, we may assume $a_k \ge b_l$, and put

$$\delta = \begin{cases} 1 & \text{if } b_l = a_k, \\ 0 & \text{if } b_l < a_k. \end{cases}$$

Our main result is:

Theorem 2. (i) *Let* $a_k \le k + l - 2 - \delta$. *Then* $|A + B| \ge a_k + l$.

(ii) *Let* $a_k \ge k + l - 1 - \delta$ *and* $(a_1, \ldots, a_k) = 1$. *Then* $|A + B| \ge k + 2l - 2 - \delta$.

We would like to note at this point that Theorem 2 will be deduced in the next section from the following lemma, which in turn will be proved in Section 3:

Lemma 1. *Let* $(a_1, \ldots, a_k) = 1$. *Then* $|A + B| \ge \min\{a_k, k + l - 2 - \delta\} + l$.

The question which so far remains unanswered is: how can one estimate $|A+B|$ in the case of $a_k \geq k+l-1-\delta$ and $(a_1,\ldots,a_k) = d > 1$? We may, of course, assume that $(a_1,\ldots,b_l) = 1$ (else both A and B should be reduced by the greatest common divisor) and then the following technique may be used.

Break B into the union of pairwise disjoint sets, $B = B_1 \cup \ldots \cup B_s$, where s is the number of residue classes modulo d having non-empty intersection with B, and B_i $(i = 1,\ldots,s)$ are those intersections. Then obviously $A+B_i$ also are pairwise disjoint, hence

$$|A+B| = \sum_{i=1}^{s} |A+B_i|.$$

Using the well-known estimate $|A+B_i| \geq |A| + |B_i| - 1$ and observing that $s \geq 2$ (in view of $(a_1,\ldots,b_l) = 1$) we immediately obtain:

LEMMA 2. *Let* $(a_1,\ldots,a_k) > 1$ *and* $(a_1,\ldots,b_l) = 1$. *Then* $|A+B| \geq 2k+l-2$.

The more accurate approach is to estimate $|A+B_i|$ using Lemma 1 (which firstly requires the application of a suitable linear transformation to both A and B_i). This readily gives

$$|A+B| \geq l + \sum_{i=1}^{s} \min\{a_k/d,\ k+l_i-2-\delta_i\}, \tag{1}$$

where we set $l_i = |B_i|$ (so that $l_1 + \ldots + l_s = l$) and

$$\delta_i = \begin{cases} 1 & \text{if } 0 \in B_i \text{ and } \delta = 1, \\ 0 & \text{if } 0 \notin B_i \text{ or } \delta = 0. \end{cases}$$

The sum on the right-hand side of (1) should now be estimated on the basis of specific features of a particular problem. Actually, we will use this approach later on in this paper to deduce Theorem 2 from Lemma 1.

And now a brief historical reference. The first generalization of Theorem 1 to the case of two distinct summands was done by G. Freiman in [2]. The results obtained may be formulated as follows:

THEOREM 3. (i) *Let* $a_k \leq k+l-3$. *Then* $|A+B| \geq a_k + l$.

(ii) *Let* $a_k \geq k+l-2$ *and* $(a_1,\ldots,b_l) = 1$. *Then* $|A+B| \geq k+l+\min\{k,l\}-3$.

Later, J. Steinig gave in [5] a somewhat simplified proof of Theorem 3.

Note that this theorem follows easily from Theorem 2 and Lemma 2 according to the scheme below:

1) If $a_k \le k+l-3$, we apply Theorem 2(i);
2) If $a_k \ge k+l-2$ and $(a_1,\ldots,b_l)=1$:
   2.1) If $(a_1,\ldots,a_k)>1$, we apply Lemma 2;
   2.2) If $(a_1,\ldots,a_k)=1$:
      2.2.1) If $a_k \ge k+l-1-\delta$, we apply Theorem 2(ii);
      2.2.2) If $a_k \le k+l-2-\delta$, then $\delta=0$, $a_k=k+l-2$ and we apply Theorem 2(i).

2. Deduction of the main theorem from Lemma 1. We assume $(a_1,\ldots,a_k)=d>1$ and

$$a_k \le k+l-2-\delta \tag{2}$$

(else Theorem 2 follows from Lemma 1 automatically) and make use of (1). First observe that B is situated in s of the d available residue classes modulo d. Therefore

$$l \le s\frac{a_k}{d}+\delta,$$

which together with (2) gives

$$a_k \le k+s\frac{a_k}{d}-2, \qquad a_k(d-s)\le (k-2)d$$

and then, in view of $a_k \ge (k-1)d$, we obtain $s=d$. Hence, the result will follow from (1) as soon as we show that for each $i=1,\ldots,d$,

$$a_k/d \le k+l_i-2-\delta_i. \tag{3}$$

Using (2) once again we obtain

$$\begin{aligned}\#\{0\le c<a_k \mid c\notin B\} &= a_k-l+\delta \le k-2,\\ l_i &= \#\{0\le c<a_k \mid c\in B_i\}+\delta_i\\ &\ge a_k/d-\#\{0\le c<a_k \mid c\notin B\}+\delta_i\\ &\ge a_k/d-(k-2)+\delta_i,\end{aligned}$$

which proves (3) and therefore the whole theorem. ∎

3. Proof of Lemma 1. Let G be an abelian group, and let $\overline{C}\subseteq G$ be a finite subset of G. By $H(\overline{C})$ we will denote the *period* of $\overline{C}$, that is, the subgroup of all those elements $h\in G$ which satisfy $\overline{C}+h=\overline{C}$. Obviously, $H(\overline{C})$ is always finite. If $|H(\overline{C})|>1$, the set $\overline{C}$ is called *periodic*.

We will need the following result, due to M. Kneser ([3], [4]):

Theorem 4. *Let $\overline{A},\overline{B}\subseteq G$ be finite non-empty subsets of G satisfying*

$$|\overline{A}+\overline{B}|\le|\overline{A}|+|\overline{B}|-1.$$

Then $H=H(\overline{A}+\overline{B})$ satisfies

$$|\overline{A}+\overline{B}|+|H|=|\overline{A}+H|+|\overline{B}+H|.$$

Hence, $\overline{A}+\overline{B}$ *is periodic if*

$$|\overline{A}+\overline{B}| \le |\overline{A}|+|\overline{B}|-2.$$

Proof of Lemma 1. Suppose

$$|A+B| \le k+2l-3-\delta \tag{4}$$

and prove that

$$|A+B| \ge a_k+l.$$

Set $q=a_k$, $\overline{A}=\varphi A$, $\overline{B}=\varphi B$, where $\varphi:\mathbb{Z}\to\mathbb{Z}_q$ is the canonical homomorphism of $\mathbb{Z}$ onto $\mathbb{Z}_q$. Then $|\overline{A}|=k-1$, $|\overline{B}|=l-\delta$, and obviously, $\overline{A}+\overline{B}=\varphi A+\varphi B=\varphi(A+B)$. As the first step, we show that

$$|\overline{A}+\overline{B}| \le |A+B|-l. \tag{5}$$

To this purpose, in the case of $b_l<a_k$ it is sufficient to observe that

$$\varphi(a_1+b_i)=\varphi(a_k+b_i) \quad (i=1,\ldots,l)$$

while all the sums $a_1+b_i,\ a_k+b_i$ $(i=1,\ldots,l)$ are pairwise distinct:

$$a_1+b_1<\ldots<a_1+b_l<a_k+b_1<\ldots<a_k+b_l.$$

And in the case of $b_l=a_k$, here we have

$$\varphi(a_1+b_i)=\varphi(a_k+b_i) \quad (i=2,\ldots,l-1),$$
$$\varphi(a_1+b_1)=\varphi(a_1+b_l)=\varphi(a_k+b_l)$$

while all the sums above are pairwise distinct:

$$a_1+b_1<a_1+b_2<\ldots<a_1+b_{l-1}<a_1+b_l$$
$$<a_k+b_2<\ldots<a_k+b_{l-1}<a_k+b_l.$$

In either case, (5) holds, and thus (4) implies

$$|\overline{A}+\overline{B}| \le (k+2l-3-\delta)-l=|\overline{A}|+|\overline{B}|-2,$$

which in view of Kneser's theorem shows that $\overline{A}+\overline{B}$ is periodic. Put $H=H(\overline{A}+\overline{B})\subseteq\mathbb{Z}_q$ so that $H=d\mathbb{Z}_q$ for some $d\,|\,q,\ d>0$ (the requirement $d>0$ effectively means $|H|>1$). Moreover, if $d=1$, then $H=\mathbb{Z}_q$, that is, $\overline{A}+\overline{B}=\mathbb{Z}_q$, and hence

$$|A+B| \ge |\overline{A}+\overline{B}|+l=q+l=a_k+l,$$

which was to be proved. We now assume $d>1$ and show that $d\,|\,(a_1,\ldots,a_k)$, in contradiction with the assumptions of the lemma.

Denote by σ the canonical homomorphism $\sigma:\mathbb{Z}_q\to\mathbb{Z}_q/H$, and let $\widetilde{A}=\sigma\overline{A},\ \widetilde{B}=\sigma\overline{B}$. Since

$$|\overline{A}+H|=|\widetilde{A}||H|, \quad |\overline{B}+H|=|\widetilde{B}||H|, \quad |\overline{A}+\overline{B}|=|\widetilde{A}+\widetilde{B}||H|,$$

Kneser's theorem gives

$$(6) \qquad |\widetilde{A}+\widetilde{B}| = |\widetilde{A}| + |\widetilde{B}| - 1,$$

$$(7) \qquad |(\overline{A}+H)\setminus\overline{A}| + |(\overline{B}+H)\setminus\overline{B}| = |H| - (|\overline{A}| + |\overline{B}| - |\overline{A}+\overline{B}|).$$

Each element $c \in A+B$ satisfies either $\sigma\varphi c \in \widetilde{B}$, or $\sigma\varphi c \in (\widetilde{A}+\widetilde{B})\setminus\widetilde{B}$. We will now separately count the number of elements c of both types:

1. Since $\overline{B}+H \subseteq \overline{A}+\overline{B}$, we have

$$(8) \qquad \begin{aligned}\#\{c \in A+B \mid \sigma\varphi c \in \widetilde{B}\} &= \#\{c \in A+B \mid \varphi c \in \overline{B}+H\} \\ &\ge l + \#\{\overline{c} \in \overline{A}+\overline{B} \mid \overline{c} \in \overline{B}+H\} \\ &= l + |\overline{B}+H| = l + |\widetilde{B}||H|.\end{aligned}$$

2. We have

$$\#\{c \in A+B \mid \sigma\varphi c \in (\widetilde{A}+\widetilde{B})\setminus\widetilde{B}\} = \sum_{\widetilde{c}\in(\widetilde{A}+\widetilde{B})\setminus\widetilde{B}} \#\{c \in A+B \mid \sigma\varphi c = \widetilde{c}\}.$$

For each $\widetilde{c} \in (\widetilde{A}+\widetilde{B})\setminus\widetilde{B}$ fix $\widetilde{a} \in \widetilde{A}$, $\widetilde{b} \in \widetilde{B}$ in such a way that $\widetilde{c} = \widetilde{a}+\widetilde{b}$. Then

$$\begin{aligned}\#\{c \in A+B \mid \sigma\varphi c = \widetilde{c}\} &\ge |\varphi^{-1}\sigma^{-1}\widetilde{a}\cap A + \varphi^{-1}\sigma^{-1}\widetilde{b}\cap B| \\ &\ge |\varphi^{-1}\sigma^{-1}\widetilde{a}\cap A| + |\varphi^{-1}\sigma^{-1}\widetilde{b}\cap B| - 1 \\ &\ge |\sigma^{-1}\widetilde{a}\cap \overline{A}| + |\sigma^{-1}\widetilde{b}\cap \overline{B}| - 1 \\ &\ge 2|H| - 1 - |(\overline{A}+H)\setminus\overline{A}| - |(\overline{B}+H)\setminus\overline{B}| \\ &= |H| - 1 + |\overline{A}| + |\overline{B}| - |\overline{A}+\overline{B}|\end{aligned}$$

(we used here (7)). Therefore, in view of (6),

$$(9) \qquad \begin{aligned}&\#\{c \in A+B \mid \sigma\varphi c \in (\widetilde{A}+\widetilde{B})\setminus\widetilde{B}\} \\ &\quad\ge |(\widetilde{A}+\widetilde{B})\setminus\widetilde{B}|(|H| - 1 + |\overline{A}| + |\overline{B}| - |\overline{A}+\overline{B}|) \\ &\quad= (|\widetilde{A}| - 1)(|H| - 1 + |\overline{A}| + |\overline{B}| - |\overline{A}+\overline{B}|).\end{aligned}$$

Summing up (8) and (9) and taking into account (6), we obtain

$$\begin{aligned}|A+B| &\ge l + |\widetilde{B}||H| + (|\widetilde{A}| - 1)(|H| - 1 + |\overline{A}| + |\overline{B}| - |\overline{A}+\overline{B}|) \\ &= l + (|\widetilde{A}| + |\widetilde{B}| - 1)|H| + (|\widetilde{A}| - 1)(|\overline{A}| + |\overline{B}| - |\overline{A}+\overline{B}| - 1) \\ &= l + |\overline{A}+\overline{B}| + (|\widetilde{A}| - 1)(|\overline{A}| + |\overline{B}| - |\overline{A}+\overline{B}| - 1).\end{aligned}$$

Now (4) gives $|A+B| \le (k-1) + (l-\delta) + l - 2 = |\overline{A}| + |\overline{B}| + l - 2$, hence

$$|\overline{A}| + |\overline{B}| + l - 2 \ge l + |\overline{A}+\overline{B}| + (|\widetilde{A}| - 1)(|\overline{A}| + |\overline{B}| - |\overline{A}+\overline{B}| - 1),$$

that is,

$$1 + (|\widetilde{A}| - 2)(|\overline{A}| + |\overline{B}| - |\overline{A}+\overline{B}| - 1) \le 0,$$

and the obtained inequality shows that $|\widetilde{A}| = 1$. But in view of $0 \in \widetilde{A}$ this means $d \mid (a_1, \ldots, a_k)$, a contradiction. ∎

4. Consequences. Two situations permanently arise in applications and are worth mentioning here.

The first is when B is a subset of A. This is an additional information, and we use it to reject in Theorem 2 the restriction concerning the greatest common divisor of elements of A. This also allows us to put the conclusion of the theorem in a more compact form, like that of Lemma 1.

The second situation is when we cannot decide in advance which one of the two sets A and B is longer. We have to pay for this uncertainty by relaxing the estimates in Theorem 2.

In this section, we do not assume that the minimal elements of A and B are 0, so the definition of δ should be changed, to say, as follows:

$$\delta = \begin{cases} 1 & \text{if } A \text{ and } B \text{ are of the same length,} \\ 0 & \text{otherwise.} \end{cases}$$

Here by the length of a set we mean the difference between its maximal and minimal elements.

We need also the notion of *reduced length*. For $A = \{a_1, \ldots, a_k\}$ put $a_i' = a_i - a_1$ $(i = 1, \ldots, k)$, and denote by d the greatest common divisor of the elements of the set $A' = \{a_1', \ldots, a_k'\}$. Then the reduced length of A is defined by $a = a_k'/d$.

THEOREM 5. *Let A be a finite set of integers of reduced length a, and $B \subseteq A$. Then*

$$|A + B| \ge \min\{a, k + l - 2 - \delta\} + l.$$

Proof. We define A' and d as above and put

$$\begin{aligned} a_i'' &= a_i'/d \qquad && (i = 1, \ldots, k), \quad A'' = \{a_1'', \ldots, a_k''\}, \\ b_i'' &= (b_i - b_1)/d \qquad && (i = 1, \ldots, l), \quad B'' = \{b_1'', \ldots, b_l''\}, \end{aligned}$$

so that $a_k'' = a$ is the reduced length of A. Then our theorem follows immediately from Lemma 1 as applied to the sets A'' and B''. ∎

The second situation of the two discussed above is covered by

THEOREM 6. *Define $d = (a_1', \ldots, a_k', b_1', \ldots, b_l')$, $a = a_k'/d$, $b = b_l'/d$ (where $a_i' = a_i - a_1$, $b_i' = b_i - b_1$) and put $c = \max\{a, b\}$. Then*

$$|A + B| \ge \min\{c, k + l - 2 - \delta\} + \min\{k, l\}.$$

Proof. We may assume $d = 1$, $a_1 = b_1 = 0$ and also (due to the symmetry between A and B) $c = a_k \ge b_l$. Then in the case of $(a_1, \ldots, a_k) = 1$ we apply Lemma 1, and otherwise, Lemma 2. ∎

It should be pointed out that theorems of this type are usually utilized to estimate the length for a given value of $|A+B|$, like the following:

COROLLARY 1. *Let* A, B, a *be as in Theorem* 5, *and assume that* $T = |A+B| < k + 2l - 2 - \delta$. *Then* $a \le T - l$.

COROLLARY 2. *Let* A, B, c *be as in Theorem* 6, *and assume that* $T = |A+B| < k + l + \min\{k, l\} - 2 - \delta$. *Then* $c \le T - \min\{k, l\}$.

References

[1] G. Freiman, *On addition of finite sets*, *I*, Izv. Vyssh. Uchebn. Zaved. Mat. 1959 (6), 202–213.

[2] —, *Inverse problems of additive number theory*, *VI. On addition of finite sets*, *III*, ibid. 1962 (3), 151–157.

[3] M. Kneser, *Abschätzung der asymptotischen Dichte von Summenmengen*, Math. Z. 58 (1953), 459–484.

[4] —, *Ein Satz über abelschen Gruppen mit Anwendungen auf die Geometrie der Zahlen*, ibid. 61 (1955), 429–434.

[5] J. Steinig, *On Freiman's theorems concerning the sum of two finite sets of integers*, in: Preprints of the conference on Structure Theory of Set Addition, CIRM, Marseille, 1993, 173–186.

SCHOOL OF MATHEMATICAL SCIENCES
SACKLER FACULTY OF EXACT SCIENCES
TEL-AVIV UNIVERSITY
69978 TEL-AVIV, ISRAEL
E-mail: SEVA@MATH.TAU.AC.IL

2/56 SMITH STR.
WOLLONGONG, NEW SOUTH WALES 2000
AUSTRALIA
E-mail: PAVEL@SBDEV.OZ.AU

Received on 6.5.1994
and in revised form on 12.7.1994 (2610)