Acta Arith. 87(1998), no. 2, 159–169.

ON SUMS OF DISTINCT REPRESENTATIVES

Hui-Qin Cao and Zhi-Wei Sun

1. Introduction

In combinatorics, for a finite sequence

$$\{A_i\}_{i=1}^n \tag{1}$$

of sets, a sequence

$$\{a_i\}_{i=1}^n \tag{2}$$

of elements is called a *system of distinct representatives* (abbreviated to SDR) of (1) if $a_1 \in A_1, \cdots, a_n \in A_n$ and $a_i \neq a_j$ for all $1 \leqslant i < j \leqslant n$. A celebrated theorem of P. Hall [H] says that (1) has an SDR if and only if

$$\Big|\bigcup_{i\in I} A_i\Big| \geqslant |I| \qquad \text{for all } I \subseteq \{1, \cdots, n\}. \tag{3}$$

Clearly (1) has an SDR provided that $|A_i| \geqslant i$ for all $i = 1, \cdots, n$, in particular an SDR of (1) exists if $|A_1| = \cdots = |A_n| \geqslant n$ or $0 < |A_1| < \cdots < |A_n|$.

Let G be an additive abelian group and $A_1, \cdots, A_n$ its subsets. We associate any SDR (2) of (1) with the sum $\sum_{i=1}^n a_i$ and set

$$S(\{A_i\}_{i=1}^n) = S(A_1, \cdots, A_n) = \{a_1 + \cdots + a_n : \ \{a_i\}_{i=1}^n \text{ forms an SDR of } \{A_i\}_{i=1}^n\}. \tag{4}$$

Of course, $S(A_1, \cdots, A_n) \neq \emptyset$ if and only if (3) holds. A fascinating and challenging problem is to give a sharp lower bound for $|S(\{A_i\}_{i=1}^n)|$ and determine when the bound can be reached.

Let p be a prime. In 1964 P. Erdös and H. Heilbronn (cf. [EH] and [G]) conjectured that for each nonempty subset A of $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ there are at least

---

1991 *Mathematics Subject Classification.* Primary 11B75; Secondary 05A05.

The research is supported by the Return-from-abroad Foundations of the Chinese Educational Committee and Nanjing City.

$\min\{p, 2|A|-3\}$ elements of $\mathbb{Z}_p$ that can be written as the sum of two distinct elements of A. With the help of Grassmann spaces this was confirmed by J. A. Dias da Silva and Y. O. Hamidoune [DH] in 1994, in fact they proved the following generalization for n-fold sums: If $A \subseteq \mathbb{Z}_p$ then

$$|n^\wedge A| \geqslant \min\left\{p,\ n|A|-n^2+1\right\} \tag{5}$$

where $n^\wedge A$ denotes the set of sums of n distinct elements of A, i.e. $n^\wedge A = S(A,\cdots,A)$ with A repeated n times on the right hand side. In 1995 and 1996, N. Alon, M. B. Nathanson and I. Z. Ruzsa [ANR1,ANR2] introduced an ingenious polynomial method and obtained the following result by contradiction: Let F be any field of characteristic p and $A_1,\cdots,A_n$ its subsets with $0<|A_1|<\cdots<|A_n|<\infty$. Then

$$|S(A_1,\cdots,A_n)| \geqslant \min\left\{p,\ \sum_{i=1}^n |A_i| - \frac{n(n+1)}{2}+1\right\} \tag{6}$$

providing $M<p$ where $M=\sum_{i=1}^n|A_i|-\frac{n(n+1)}{2}$. We mention that (6) also holds in the case $M \geqslant p$. In fact, let s be the smallest positive integer with $\sum_{i=1}^s(|A_i|-i) > M-p$ and choose $A_1' \subseteq A_1,\cdots,A_n'\subseteq A_n$ so that $|A_i'|=i$ for $i<s$, $A_i'=A_i$ for $i>s$, and

$$|A_s'| = s-1+\sum_{i=1}^s(|A_i|-i)-(M-p) \leqslant |A_s|-1,$$

then $|A_1'|<\cdots<|A_n'|$, $M'=\sum_{i=1}^n|A_i'|-\frac{n(n+1)}{2}=p-1$ and so

$$|S(A_1,\cdots,A_n)| \geqslant |S(A_1',\cdots,A_n')| \geqslant M'+1 = p = \min\{p,M\}.$$

Alon, Nathanson and Ruzsa [ANR2] posed the question when the lower bound in (6) can be reached and considered it as an interesting one.

In view of the fundamental theorem of finitely generated abelian groups (cf. [J]), if a finite addition theorem holds in $\mathbb{Z}$ then it holds in any torsion-free abelian groups. So, without any loss of generality, we may work within $\mathbb{Z}$.

For a finite subset A of $\mathbb{Z}$, in 1995 Nathanson [N] showed the inequlity $|n^\wedge A| \geqslant n|A|-n^2+1$ and proved that if the equality holds then A must be an AP providing $2 \leqslant n < |A|-2$. The same result was independently obtained by Y. Bilu [B].

Let $A_1,\cdots,A_n$ be finite subsets of $\mathbb{Z}$ with $0<|A_1|<\cdots<|A_n|$. Take a sufficiently large prime p such that it is greater than $\sum_{i=1}^n|A_i|-n(n+1)/2$ and the largest element of $S(A_1, \cdots, A_n)$. Applying the Alon-Nathanson-Ruzsa result stated above, we have the inequality

$$|S(A_1, \cdots, A_n)| \geqslant \sum_{i=1}^{n} |A_i| - \frac{n(n+1)}{2} + 1 = \sum_{i=1}^{n} (|A_i| - i) + 1. \tag{7}$$

In this paper we will make a new approach to sums of distinct representatives. The method allows us to give a somewhat constructive proof of (7) provided that $A_1, \cdots, A_n$ are finite nonempty subsets of $\mathbb{Z}$ with distinct cardinalities. Furthermore we are able to make key progress in the equality case.

Let's first look at two examples.

Example 1. Let $a \in \mathbb{Z}$ and $d \in \mathbb{Z} \setminus \{0\}$. Let $k \geqslant n \geqslant 1$, $A = \{a + jd : \ j = 0, 1, \cdots, k-1\}$, and $A_1, \cdots, A_n$ be subsets of A with $|A_i| = k - n + i$ for every $i = 1, \cdots, n$. Obviously $S(A_1, \cdots, A_n) \subseteq n^\wedge A$. If $S \subseteq A$ and $|S| = n$, then for each $i = 1, \cdots, n$ at least i elements of S lie in A_i since $|S \setminus A_i| \leqslant |A \setminus A_i| = n - i$, therefore we can write S in the form $\{a_1, \cdots, a_n\}$ where $a_1 \in A_1, \cdots, a_n \in A_n$. So $n^\wedge A \subseteq S(A_1, \cdots, A_n)$. Let $X = \{j_1 + \cdots + j_n : \ 0 \leqslant j_1 < \cdots < j_n < k\}$. For each $j = 0, 1, \cdots, n(k-n)$, there exist $0 \leqslant u < n$ and $0 \leqslant v \leqslant k - n$ such that $j = u(k-n) + v$, hence

$$\begin{aligned} \frac{n(n-1)}{2} + j &= \sum_{i=1}^{n} (i-1) + u(k-n) + v \\ &= \sum_{0<i<n-u} (i-1) + (n-u-1+v) + \sum_{n-u<i\leqslant n} (k-n+i-1) \end{aligned}$$

belongs to X. Thus $\{\frac{n(n-1)}{2} + j : \ 0 \leqslant j \leqslant n(k-n)\} \subseteq X$. Apparently the least and the largest elements of X are $0+1+\cdots+(n-1) = \frac{n(n-1)}{2}$ and $(k-n)+\cdots+(k-1) = \frac{n(n-1)}{2} + n(k-n)$ respectively. So, by the above

$$\begin{aligned} S(A_1, \cdots, A_n) = n^\wedge A &= \left\{ \sum_{i=1}^{n} (a + j_i d) : \ 0 \leqslant j_1 < \cdots < j_n < k \right\} = \{na + xd : \ x \in X\} \\ &= \left\{ na + \left(\frac{n(n-1)}{2} + j \right) d : \ 0 \leqslant j \leqslant n(k-n) \right\} \end{aligned}$$

and hence

$$|S(A_1, \cdots, A_n)| = |n^\wedge A| = n(|A| - n) + 1 = \sum_{i=1}^{n} |A_i| - \frac{n(n+1)}{2} + 1.$$

Example 2 (cf. [N]). Let $a_0, a_1, a_2, a_3 \in \mathbb{Z}$, $a_0 < a_1 < a_2 < a_3$ and $a_3 - a_2 = a_1 - a_0$ (but $a_2 - a_1$ may be different from $a_1 - a_0$). Let $A_1 = \{a_0, a_1, a_2\}$, $A_2 = \{a_0, a_1, a_2, a_3\}$. Then

$$S(A_1, A_2) = \{a_0 + a_1,\ a_0 + a_2,\ a_0 + a_3 = a_1 + a_2,\ a_1 + a_3,\ a_2 + a_3\}.$$

Note that in this example $|A_1| = 3 < |A_2| = 4 < |S(A_1, A_2)| = 5 = |A_1| + |A_2| - \frac{2(2+1)}{2} + 1$.

Now we introduce some notations throughout the paper. For a subset A of $\mathbb{Z}$, $-A$ refers to $\{-a : a \in A\}$, $\min A$ and $\max A$ denote the least and the largest elements of A respectively. If there exist $a \in \mathbb{Z}$, $d \in \mathbb{Z} \setminus \{0\}$ and a positive integer k such that

$$A = \{a + jd : 0 \leqslant j < k\},$$

then we call A an *arithmetic progression* (in short, AP).

In this paper, by a novel method we obtain the following

Theorem. *Let $A_1, \cdots, A_n$ be subsets of $\mathbb{Z}$ with $0 < |A_1| < \cdots < |A_n| < \infty$. Then inequality (7) holds. Moreover, in the equality case we have $\bigcup_{i=1}^m A_i = A_m$ for every $m \in M = \{1 \leqslant j < n : |A_{j+1}| > |A_j| + 1\} \cup \{n\}$, and A_n forms an AP unless $n = 1$ or $|A_1| \leqslant 3$.*

The result of Nathanson and Bilu stated above actually follows from Theorem. For $i = 1, \cdots, n$ let $A_i \subseteq A$ and $|A_i| = |A| - (n - i)$. Obviously $0 < |A_1| < \cdots < |A_n| < \infty$. It follows from Example 1 and Theorem that

$$|n^\wedge A| = |S(A_1, \cdots, A_n)| \geqslant \sum_{i=1}^n |A_i| - \frac{n(n+1)}{2} + 1 = n|A| - n^2 + 1.$$

When $|n^\wedge A| = n|A| - n^2 + 1$, we have $|S(A_1, \cdots, A_n)| = \sum_{i=1}^n |A_i| - \frac{n(n+1)}{2} + 1$, hence by Theorem if $2 \leqslant n \leqslant |A| - 3$ (i.e., $n \geqslant 2$ and $|A_1| \geqslant 4$) then $A = A_n$ is an AP.

In the next section we shall provide two lemmas. A proof of Theorem will be given in Section 3.

2. Auxiliary Results

Lemma 1. *Let G be an additive abelian group, and $A_1, \cdots, A_n$ its finite subsets. Let $r \in \{1, \cdots, n\}$ and suppose that $\{a_i\}_{i \neq r}$ forms an SDR of $\{A_i\}_{i \neq r}$. Then*

(i) *There exists a* $J \subseteq \{1, \cdots, n\}$ *containing* r *such that if* $J \subseteq I \subseteq \{1, \cdots, n\}$ *then* $S_r^{(I)}(\bigcup_{j\in J} A_j) \subseteq S(\{A_i\}_{i\in I})$, $\{i \in I\backslash\{r\} : a_i \in \bigcup_{j\in J} A_j\} = J\backslash\{r\}$ *and hence* $|S_r^{(I)}(\bigcup_{j\in J} A_j)| = |\bigcup_{j\in J} A_j| - |J| + 1$, *where for any subset* A *of* G *we let*

$$S_r^{(I)}(A) = \left\{ \sum_{i\in I} a_i : \ a_r \in A\backslash\{a_i : i \in I\backslash\{r\}\} \right\}.$$

(ii) *Let* $k_r = |A_r| < \cdots < k_n = |A_n|$. *For any* J *described in (i) and and* $I \subseteq \{1, \cdots, n\}$ *containing* J, *we have the inequality*

$$\left| S_r^{(I)}\left(\bigcup_{j\in J} A_j \right) \right| \geqslant k_r - r + 1, \tag{9}$$

and moreover the equality holds if and only if there exists an $l \in \{r, \cdots, n\}$ *for which* $J = \{1, \cdots, l\}$, $\bigcup_{j=1}^{l} A_j = A_l$ *and* $k_l = k_r + (l - r)$.

Proof. i) Let $\mathcal{J}$ be the class of those $J \subseteq \{1, \cdots, n\}$ containing r such that if $J \subseteq I \subseteq \{1, \cdots, n\}$ then for each $j \in J$ there exists a one-to-one mapping $\sigma_{I,j} : I\backslash\{r\} \to I\backslash\{j\}$ for which $a_i \in A_{\sigma_{I,j}(i)}$ for all $i \in I \setminus \{r\}$. Obviously $\mathcal{J}$ is nonempty (for, $\{r\}$ belongs to $\mathcal{J}$) and finite. Let J be any maximal set in $\mathcal{J}$ with respect to the semiorder $\subseteq$, and let $J \subseteq I \subseteq \{1, \cdots, n\}$.

Set $A = \bigcup_{j\in J} A_j$. Apparently $J' = \{r\} \bigcup \{i \in I\backslash\{r\} : a_i \in A\}$ contains J. Let $J' \subseteq I' \subseteq \{1, \cdots, n\}$. Since $J \in \mathcal{J}$ and $J \subseteq I'$, for $j \in J$ there is a one-to-one mapping $\sigma_{I',j} : I'\backslash\{r\} \to I'\backslash\{j\}$ such that $a_i \in A_{\sigma_{I',j}(i)}$ for all $i \in I'\backslash\{r\}$. For $j' \in J'\backslash J$, there is a $j \in J$ with $a_{j'} \in A_j$. Since $J \in \mathcal{J}$ and $I'' = I'\backslash\{j'\} \supseteq J$, there also exists a one-to-one mapping $\sigma_{I'',j} : I''\backslash\{r\} \to I''\backslash\{j\}$ such that $a_i \in A_{\sigma_{I'',j}(i)}$ for $i \in I''\backslash\{r\}$. Obviously by letting $j' \in I'\backslash\{r\}$ correspond to $j \in I'\backslash\{j'\}$ we can extend $\sigma_{I'',j}$ to a one-to-one mapping $\sigma_{I',j'} : I'\backslash\{r\} \to I'\backslash\{j'\}$ for which $a_i \in A_{\sigma_{I',j'}(i)}$ for all $i \in I'\backslash\{r\}$. Thus $J' \in \mathcal{J}$. As $J \subseteq J'$ and J is a maximal set in $\mathcal{J}$, we must have $J' = J$, i.e., $\{i \in I\backslash\{r\} : \ a_i \in \bigcup_{j\in J} A_j\} = J\backslash\{r\}$.

If $j \in J$ and $x_j \in A_j\backslash\{a_i : i \in I\backslash\{r\}\}$, then $x_j + \sum_{i\in I\backslash\{r\}} a_i \in S(\{A_i\}_{i\in I})$ because $a_i \in A_{\sigma_{I,j}(i)}$ for $i \in I \setminus \{r\}$. So $S_r^{(I)}(A) \subseteq S\left(\{A_i\}_{i\in I}\right)$. Note that

$$\begin{aligned} |S_r^{(I)}(A)| &= |A\backslash\{a_i : i \in I\backslash\{r\}\}| = |A| - |\{i \in I\backslash\{r\} : a_i \in A\}| \\ &= |A| - |J\backslash\{r\}| = |A| - |J| + 1. \end{aligned}$$

This proves part (i).

ii) Let J be as described in (i), $A = \bigcup_{j\in J} A_j$ and $J \subseteq I \subseteq \{1, \cdots, n\}$. If $|J| < r$, then

$$|A| - |J| \geqslant |A_r| - |J| > k_r - r.$$

When $|J| \geqslant r$, clearly $\max J \geqslant r$ and $k_{\max J} - k_r = \sum_{r<i\leqslant \max J}(k_i - k_{i-1}) \geqslant \max J - r$, therefore

$$|A| - |J| \geqslant |A_{\max J}| - |J| \geqslant k_r + (\max J - r) - |J| \geqslant k_r - r,$$

and $|A| - |J| = k_r - r$ if and only if

$$A = A_{\max J}, \quad k_{\max J} = k_r + (\max J - r), \quad \max J = |J|,$$

i.e.,

$$J = \{1, \cdots, |J|\}, \quad A = A_{|J|}, \quad k_{|J|} = k_r + (|J| - r).$$

This together with the equality $|S_r^{(I)}(A)| = |A| - |J| + 1$ yields the second part. We are done.

Lemma 2. *Let A and B be finite subsets of $\mathbb{Z}$ with $4 \leqslant k = |A| < l = |B|$, $A \subseteq B$, $\min A = \min B$, $\max A \neq \max B$ and $|S(A,B)| = k + l - 2$. Then B is an AP.*

Proof. Let $A = \{a_1, \cdots, a_k\}$ and $B = \{b_1, \cdots, b_l\}$ where $a_1 < \cdots < a_k$ and $b_1 < \cdots < b_l$. Put $C = \{a_1 + b_2, \cdots, a_1 + b_{l-1}, a_1 + b_l, \cdots, a_k + b_l\}$. Clearly $C \subseteq S(A,B)$ and $|C| = k + l - 2$. As $|S(A,B)| = k + l - 2$, $S(A,B)$ coincides with C. Since $A \subseteq B$ and $a_k \neq b_l$, for $i = 2, \cdots, k$ we may suppose that $a_i = b_{f(i)}$ where $i \leqslant f(i) < l$. Because

$$S(A,B) \supseteq \{a_1 + b_j :\ 2 \leqslant j < f(i)\} \cup \{a_i + b_j :\ j \neq f(i)\} \cup \{a_j + b_l :\ i < j \leqslant k\},$$

we have $k + l - 2 = |S(A,B)| \geqslant (f(i) - 2) + (l - 1) + (k - i) = k + l + f(i) - i - 3$, i.e. $a_i \in \{b_i, b_{i+1}\}$. Observe that $a_3 < a_k \leqslant b_{l-1}$. Since

$$a_1 + b_{l-1} < a_2 + b_{l-1} < a_3 + b_{l-1} < a_3 + b_l,$$

$a_2 + b_{l-1} = a_1 + b_l$ and $a_3 + b_{l-1} = a_2 + b_l$, therefore

$$a_2 - a_1 = b_l - b_{l-1} = a_3 - a_2.$$

If $a_2 \neq b_2$, then $a_2 = b_3$, $a_3 = b_4$, $b_2 - b_1 < a_2 - b_1 = a_3 - a_2$, $a_1 + b_3 < b_2 + a_2 < a_3 + b_1 = a_1 + b_4$, this contradicts the fact $a_2 + b_2 \in S(A,B) = C$. So $a_2 = b_2$. As $a_2 + b_{l-1} \in C$ we must have $a_2 + b_{l-1} = a_1 + b_l$, similarly $a_2 + b_{l-2} = a_1 + b_{l-1}, \cdots, a_2 + b_3 = a_1 + b_4$. Thus

$$b_l - b_{l-1} = \cdots = b_4 - b_3 = a_2 - a_1 = b_2 - b_1 = a_3 - b_2.$$

If $a_3 \neq b_3$, then $a_3 = b_4$ and hence $b_2 = b_2 - a_3 + b_4 = b_3$ which is impossible. So $a_3 = b_3$ and B forms an AP.

3. Proof of Theorem

Proof of Theorem: The case $n = 1$ is trivial. Below we let $n \geqslant 2$ and assume the theorem for smaller values of n.

Put $k_i = |A_i|$ for $i = 1, \cdots, n$. Set $a = \min \bigcup_{i=1}^n A_i$, $I = \{1 \leqslant i \leqslant n : a \in A_i\}$, $r = \min I$ and $t = \max I$. For $i \in I$ let

$$A'_i = \begin{cases} A_i \setminus \{a\} & \text{if } i \neq r, \\ \{a\} & \text{if } i = r; \end{cases}$$

and for $i \in \bar{I} = \{1, \cdots, n\} \setminus I$ put

$$A'_i = \begin{cases} A_i \setminus \{a_i\} & \text{if } r \leqslant i \leqslant t \text{ and } i \notin M, \\ A_i & \text{otherwise}, \end{cases}$$

where a_i is an arbitrary element of A_i. Apparently all the A'_i are finite, nonempty and contained in $\mathbb{Z}$, also $S(A'_1, \cdots, A'_n) = S(\{A'_i\}_{i \neq r})$. Let $k'_i = |A'_i|$ for $i = 1, \cdots, n$. Observe that $k'_i < k'_j$ if $1 \leqslant i < j \leqslant n$ and $i, j \neq r$. By the induction hypothesis,

$$|S(\{A'_i\}_{i \neq r})| \geqslant \sum_{i \neq r} k'_i - \frac{(n-1)(n-1+1)}{2} + 1 > 0.$$

Suppose that $\max S(\{A'_i\}_{i \neq r}) = \sum_{i \neq r} a'_i$ where $\{a'_i\}_{i \neq r}$ is an SDR of $\{A'_i\}_{i \neq r}$. By Lemma 1 there exists a $J \subseteq \{1, \cdots, n\}$ containing r for which

$$J \setminus \{r\} = \{i \neq r :\ a'_i \in A\},\ S_r(A) \subseteq S(\{A_i\}_{i=1}^n) \text{ and } |S_r(A)| \geqslant k_r - r + 1$$

where $A = \bigcup_{j \in J} A_j$ and $S_r(A) = \{\sum_{i=1}^n a'_i : a'_r \in A \setminus \{a'_i : i \neq r\}\}$. As $S(A_1, \cdots, A_n)$ contains $S(A'_1, \cdots, A'{}_n) \cup S_r(A)$ and

$$\max S(A'_1, \cdots, A'_n) = a + \sum_{i \neq r} a'_i = \min S_r(A),$$

we have

$$\begin{aligned} |S(A_1, \cdots, A_n)| &\geqslant |S(A'_1, \cdots, A'_n)| + |S_r(A)| - 1 \\ &\geqslant \sum_{i \neq r} k'_i - \frac{n(n-1)}{2} + 1 + (k_r - r + 1) - 1 \\ &\geqslant \sum_{i \neq r} k_i - (t - r) - \frac{n(n-1)}{2} + 1 + k_r - r \\ &= \sum_{i=1}^n k_i - \frac{n(n+1)}{2} + n - t + 1 \\ &\geqslant \sum_{i=1}^n k_i - \frac{n(n+1)}{2} + 1. \end{aligned}$$

From now on we assume that

$$|S(A_1,\cdots,A_n)| = \sum_{i=1}^{n} k_i - \frac{n(n+1)}{2} + 1. \tag{10}$$

The above deduction yields

$$|S(\{A'_i\}_{i\neq r})| = \sum_{i\neq r} k_i - (t-r) - \frac{n(n-1)}{2} + 1, \tag{11}$$

$$r < i < t \Rightarrow i \notin \bar{I} \cap M, \tag{12}$$

$$|S_r(A)| = k_r - r + 1, \tag{13}$$

$$t = n. \tag{14}$$

By (13) and Lemma 1 there is an $l \in \{r,\cdots,n\}$ for which $J = \{1,\cdots,l\}$, $\bigcup_{j=1}^{l} A_j = A_l$, $k_l = k_r + (l-r)$ and

$$\{1,\cdots,l\}\setminus\{r\} = J\setminus\{r\} = \{i \neq r :\ a'_i \in A = A_l\}. \tag{15}$$

(14) indicates that $a \in A_n$. Let $b = \max\bigcup_{i=1}^{n} A_i$. Clearly $a \neq b$, for, otherwise each A_i contains exactly one element, which contradicts the inequality $k_1 < k_n$. As $-b = \min\bigcup_{i=1}^{n} -A_i$ and $|S(-A_1,\cdots,-A_n)| = |S(A_1,\cdots,A_n)| = \sum_{i=1}^{n} |-A_i| - \frac{n(n+1)}{2} + 1$, similarly we have $-b \in -A_n$. So $b \in A'_n = A_n\setminus\{a\}$. Choose the smallest $s \leqslant n$ such that $b \in A_s$.

Let $m \in M$. We come to show that $\bigcup_{i=1}^{m} A_i = A_m$, i.e. A_m contains both $\bigcup_{i=1,\ i\neq r}^{m} A_i$ and A_r.

If $m = r$, then $r \in M$, hence $l = r = m$ and $A_m = \bigcup_{i=1}^{m} A_i \supseteq \bigcup_{i=1,\ i\neq r}^{m} A_i$.

Since $A'_i = A_i$ for all $i < r$, by (11),(12) and the induction hypothesis, if $m < r$ then $\bigcup_{i=1}^{m} A_i = \bigcup_{i=1,\ i\neq r}^{m} A_i = A_m$.

Clearly $b \in \{a'_i : i \neq r\}$ (otherwise $\sum_{i\neq r,n} a'_i + b \in S(\{A'_i\}_{i\neq r})$ would be greater than $\sum_{i\neq r} a'_i = \max S(\{A'_i\}_{i\neq r})$). Suppose that $b = a'_j$ where $j \neq r$. In view of (15), $b \in A_l$ if and only if $j \leqslant l$. Since $b = a'_j \in A'_j \subseteq A_j$, we have $j \geqslant s$. If $l = s$, then $b \in A_l$, $s \leqslant j \leqslant l = s$, $s = j \neq r$.

Now suppose that $r < m \leqslant n$. If $m < t = n$, then $m \in I$ by (12), and $k'_{m+1} - k'_m = (k_{m+1}+1) - (k_m+1) > 1$. By (11), (12) and the induction hypothesis $\bigcup_{i=1,\ i\neq r}^{m} A'_i = A'_m$. If $1 \leqslant i < r$, then $A_i = A'_i \subseteq A'_m \subseteq A_m$; if $r < i \leqslant m$ and $i \in I$, then $A_i = A'_i \cup \{a\} \subseteq A'_m \cup \{a\} = A_m$; if $r < i \leqslant m$ but $i \notin I$, then $|A_i| \geqslant k_r + 1 \geqslant 2$ and hence for any given $x_i \in A_i$ by taking $a_i \in A_i$ different from x_i at the beginning we obtain that $x_i \in A_i\setminus\{a_i\} = A'_i \subseteq A'_m \subseteq A_m$. So $\bigcup_{i=1,\ i\neq r}^{m} A_i \subseteq A_m$.

Since s is the smallest index such that $-A_s$ contains $\min\bigcup_{i=1}^n -A_i = -b$, by analogy $\bigcup_{i=1,\ i\neq s}^m -A_i \subseteq -A_m$. Thus, if $r \neq s$ then $-A_r \subseteq -A_m$, i.e. $A_r \subseteq A_m$. If $r = s$, then by the above $l \neq s = r$, also $l \leqslant m$ since $k_l = k_r + l - r$, therefore $A_r \subseteq \bigcup_{j=1}^l A_j = A_l \subseteq A_m$. So $\bigcup_{i=1}^m A_i = A_m$.

Now let's check that A_n is an AP except the case $k_1 \leqslant 3$.

If $r = 1$ then $\min\{k_i' : i \neq r\} = k_2' = k_2 - 1 \geqslant k_1$, if $r > 1$ then $\min\{k_i' : i \neq r\} = k_1' = k_1$. So $\min\{k_i' : i \neq r\} \geqslant k_1$. Below we assume that $k_1 \geqslant 4$.

Suppose $n > 2$. By (11), (12) and the induction hypothesis, if $r < n$ then $A_n' = A_n\backslash\{a\}$ is an AP. Similarly, if $s < n$ then $-A_n\backslash\{-b\}$ is an AP and hence so is $A_n\backslash\{b\}$. Thus, if $r < n$ and $s < n$ then $A_n = \{a\} \cup (A_n\backslash\{a\}) = (A_n\backslash\{b\}) \cup \{b\}$ forms an AP. (Note that $|A_n| = k_n > k_1 \geqslant 4$.)

Now consider the case $r = s = 1 < n = 2$. By the above $l = 2$ (since $l \neq s = 1$) and $k_2 = k_1 + 1$. Let $a = b_1 < \cdots < b_{k_2} = b$ be all the elements of A_2. If $1 \leqslant i < j \leqslant k_2$, then either b_i or b_j belongs to A_1 because $|A_2 \setminus A_1| = 1$, therefore $b_i + b_j \in S(A_1, A_2)$. So

$$S(A_1, A_2) \supseteq S(A_2\backslash\{b\}, A_2) \supseteq C = \{b_1 + b_2, \cdots, b_1 + b_{k_2}, b_2 + b_{k_2}, \cdots, b_{k_2-1} + b_{k_2}\}.$$

As $|S(A_1, A_2)| = k_1 + k_2 - 2 = 2k_2 - 3 = |C|$, we have $S(A_1, A_2) = C = S(A_2\backslash\{b\}, A_2)$. Clearly $|A_2\backslash\{b\}| = k_2 - 1 = k_1 \geqslant 4$, $\min(A_2\backslash\{b\}) = \min A_2$ and $\max(A_2\backslash\{b\}) \neq \max A_2$. Applying Lemma 2 we find that $A_n = A_2$ forms an AP.

With respect to the case $r = n$ we make the following remarks:

i) Since $A_n = \bigcup_{j=1}^n A_j$, we have $A_{n-1} \subset A_n$. If $n > 2$, then by (11), (12) and the induction hypothesis $A_{n-1}' = \bigcup_{i=1}^{n-1} A_i'$ forms an AP, i.e. $\bigcup_{i=1}^{n-1} A_i = A_{n-1}$ is an AP. Set

$$A_n^- = \{x \in A_n :\ x \leqslant \max A_{n-1}\} \ \text{ and } \ A_n^+ = \{x \in A_n :\ \min A_{n-1} \leqslant x\}.$$

Whether $n = 2$ or $n > 2$ we always have $\bigcup_{i=1}^{n-1} A_i = A_{n-1} \subseteq A_n^- \cap A_n^+$ and hence $|A_n^- \cap A_n^+| \geqslant k_{n-1} \geqslant k_1 \geqslant 4$. Among $A_1, \cdots, A_{n-1}, A_n^-$, the index r^- of the first one containing $\min(\bigcup_{i=1}^{n-1} A_i \cup A_n^-) = a$ is identical with r while the index s^- of the first one containing $\max(\bigcup_{i=1}^{n-1} A_i \cup A_n^-) = \max A_{n-1}$ is less than n. Similarly, among $A_1, \cdots, A_{n-1}, A_n^+$, the index r^+ of the first one containing $\min(\bigcup_{i=1}^{n-1} A_i \cup A_n^+) = \min A_{n-1}$ is less than n while the index s^+ of the first one containing $\max(\bigcup_{i=1}^{n-1} A_i \cup A_n^+) = b$ is equal to s.

ii) Suppose that $A_n^- \neq A_{n-1}$. Then $|A_n^-| > k_{n-1} > \cdots > k_1$. According to the previous reasoning,

$$|S(A_1, \cdots, A_{n-1}, A_n^-)| \geqslant \sum_{i=1}^{n-1} k_i + |A_n^-| - \frac{n(n+1)}{2} + 1 > 0.$$

Observe that

$$\max S(A_1,\cdots,A_{n-1},A_n^-)\leqslant \max S(A_1',\cdots,A_{n-1}')\ +\max A_{n-1}$$
$$<\min S(\{a_1'\},\cdots,\{a_{n-1}'\},A_n\backslash A_n^-),$$

and that $|S(\{a_1'\},\cdots,\{a_{n-1}'\},A_n\backslash A_n^-)|=|A_n\backslash A_n^-|$. So

$$\begin{aligned}|S(A_1,\cdots,A_n)|&\geqslant |S(A_1,\cdots,A_{n-1},A_n^-)|+|S(\{a_1'\},\cdots,\{a_{n-1}'\},A_n\backslash A_n^-)|\\&\geqslant \sum_{i=1}^{n-1}k_i+|A_n^-|-\frac{n(n+1)}{2}+1+(k_n-|A_n^-|)\\&=\sum_{i=1}^{n}k_i-\frac{n(n+1)}{2}+1.\end{aligned}$$

Since (10) holds, we must have

$$|S(A_1,\cdots,A_{n-1},A_n^-)|=\sum_{i=1}^{n-1}k_i+|A_n^-|-\frac{n(n+1)}{2}+1. \tag{16}$$

iii) By analogy, when $-A_n^+=\{-x\in -A_n:-x\leqslant \max -A_{n-1}\}\neq -A_{n-1}$ (i.e. $A_n^+\neq A_{n-1}$), we have

$$|S(A_1,\cdots,A_{n-1},A_n^+)|=|S(-A_1,\cdots,-A_{n-1},-A_n^+)|=\sum_{i=1}^{n-1}k_i+|A_n^+|-\frac{n(n+1)}{2}+1. \tag{17}$$

Assume that $s<r=n$. Then both r^+ and $s^+=s$ are less than n. If $A_n^+\neq A_{n-1}$, then (17) holds and hence A_n^+ forms an AP by previous arguments. If $n>2$, then A_n^+ is an AP anyway and so is $A_n\backslash\{b\}$ by the above, therefore A_n forms an AP. If $n=2$, then $s=1$, $\min -A_1=\min -A_2$, $\max -A_1\neq \max -A_2$ (since $r=2$), and $|S(-A_1,-A_2)|=|S(A_1,A_2)|=k_1+k_2-2$, hence $-A_2$ is an AP by Lemma 2, thus $A_n=A_2$ forms an AP.

In the case $r<s=n$, by applying the above result to the subsets $-A_1,\cdots,-A_n$ instead of $A_1,\cdots,A_n$, we obtain that $-A_n$ forms an AP, i.e., A_n is an AP.

Finally we handle the remaining case $r=s=n$. Since $r^+<s^+=s=n$ and $s^-<r^-=r=n$, by the above A_n^+ forms an AP if $A_n^+\neq A_{n-1}$, and A_n^- forms an AP if $A_n^-\neq A_{n-1}$. Providing $n>2$, both A_n^+ and A_n^- are APs, therefore A_n forms an AP. When $n=2$, if $A_n=A_2$ isn't an AP, then A_2^- or A_2^+ coincides with A_1, hence $\min A_1=\min A_2^-=\min A_2$ and $\max A_1\neq \max A_2$ (since $s=2$), or $\min -A_1=\min -A_2^+=\min -A_2$ and $\max -A_1\neq \max -A_2$ (since $r=2$), thus A_2 forms an AP by Lemma 2, which leads a contradiction. So, whether $n>2$ or $n=2$, A_n always forms an AP.

The induction step is now completed and the proof of Theorem is ended.

References

[ANR1] N. Alon, M. B. Nathanson and I. Z. Ruzsa, *Adding distinct congruence classes modulo a prime*, Amer. Math. Monthly **102** (1995), 250–255.

[ANR2] N. Alon, M. B. Nathanson and I. Z. Ruzsa, *The polynomial method and restricted sums of congruence classes*, J. Number Theory **56** (1996), 404–417.

[B] Y. Bilu, *Addition of sets of integers of positive density*, J. Number Theory **64** (1997), 233–275.

[DH] J. A. Dias da Silva and Y. O. Hamidoune, *Cyclic space for Grassmann derivatives and additive theory*, Bull. London Math. Soc. **26** (1994), 140–146.

[EH] P. Erdös and H. Heilbronn, *On the addition of residue classes* mod p, Acta Arith. **9** (1964), 149–159.

[G] R. K. Guy, *Unsolved Problems in Number Theory* (2nd ed.), Springer-Verlag, New York, 1994, pp. 129–131.

[H] P. Hall, *On representatives of subsets*, J. London Math. Soc. **10** (1935), 26–30.

[J] N. Jacobson, *Basic Algebra I* (2nd ed.), W.H. Freeman and Company, New York, 1985.

[N] M. B. Nathanson, *Inverse theorems for subset sums*, Trans. Amer. Math. Soc. **347** (1995), 1409–1418.

Hui-Qin Cao, the Fundamental Division, Nanjing Auditing College, Nanjing 210029, the People's Republic of China

Zhi-Wei Sun, Department of Mathematics, Nanjing University, Nanjing 210093, the People's Republic of China. *E-mail*: `zwsun@netra.nju.edu.cn`