ACTA ARITHMETICA
LXXXVII.2 (1998)

On sums of distinct representatives

by

Hui-Qin Cao and Zhi-Wei Sun (Nanjing)

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 $a_1 \in A_1, \ldots, a_n \in A_n$ and $a_i \neq a_j$ for all $1 \leq i < j \leq n$. A celebrated theorem of P. Hall [H] says that (1) has an SDR if and only if

$$\Big|\bigcup_{i\in I} A_i\Big| \geq |I| \quad \text{for all } I \subseteq \{1, \ldots, n\}. \tag{3}$$

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

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

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

Of course, $S(A_1, \ldots, A_n) \neq \emptyset$ 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. Erdős 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 $\min\{p, 2|A| - 3\}$ 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

---

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.

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

$$(5) \qquad |n^\wedge A| \geq \min\{p, n|A| - n^2 + 1\}$$

where $n^\wedge A$ denotes the set of sums of n distinct elements of A, i.e. $n^\wedge A = S(A, \ldots, A)$ 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 $A_1, \ldots, A_n$ its subsets with $0 < |A_1| < \ldots < |A_n| < \infty$. Then

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

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

$$|A_s'| = s - 1 + \sum_{i=1}^{s}(|A_i| - i) - (N - p) \leq |A_s| - 1.$$

Then $|A_1'| < \ldots < |A_n'|$, $N' = \sum_{i=1}^{n} |A_i'| - n(n+1)/2 = p - 1$ and so

$$|S(A_1, \ldots, A_n)| \geq |S(A_1', \ldots, A_n')| \geq N' + 1 = p = \min\{p, N\}.$$

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

In view of the fundamental theorem on 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 inequality $|n^\wedge A| \geq n|A| - n^2 + 1$ and proved that if equality holds then A must be an arithmetic progression providing $2 \leq n < |A| - 2$. The same result was independently obtained by Y. Bilu [B].

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

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

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 $A_1, \ldots, A_n$ are finite nonempty subsets of $\mathbb{Z}$ with distinct cardinalities. Furthermore we are able to make key progress in the equality case.

Let us first look at two examples.

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

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

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

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

and hence

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

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 $A_1 = \{a_0, a_1, a_2\}$ and $A_2 = \{a_0, a_1, a_2, a_3\}$. 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 $|A_1| = 3 < |A_2| = 4 < |S(A_1, A_2)| = 5 = |A_1| + |A_2| - 2(2+1)/2 + 1$.

Now we introduce some notations to be used 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

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

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

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

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

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

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

When $|n^\wedge A| = n|A| - n^2 + 1$, we have

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

hence by the Theorem if $2 \le n \le |A| - 3$ (i.e., $n \ge 2$ and $|A_1| \ge 4$) then $A = A_n$ is an AP.

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

2. Auxiliary results

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

(i) *There exists a $J \subseteq \{1, \ldots, n\}$ containing r such that if $J \subseteq I \subseteq \{1, \ldots, n\}$ then*

$$S_r^{(I)}\Big(\bigcup_{j \in J} A_j\Big) \subseteq S(\{A_i\}_{i \in I}), \qquad \Big\{i \in I \setminus \{r\} : a_i \in \bigcup_{j \in J} A_j\Big\} = J \setminus \{r\}$$

and hence

$$\Big|S_r^{(I)}\Big(\bigcup_{j \in J} A_j\Big)\Big| = \Big|\bigcup_{j \in J} A_j\Big| - |J| + 1,$$

where for any subset A of G we let

$$S_r^{(I)}(A) = \Big\{\sum_{i \in I} a_i : a_r \in A \setminus \{a_i : i \in I \setminus \{r\}\}\Big\}. \tag{8}$$

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

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

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

Proof. (i) Let $\mathcal{J}$ be the class of those $J \subseteq \{1, \ldots, n\}$ containing r such that if $J \subseteq I \subseteq \{1, \ldots, n\}$ then for each $j \in J$ there exists a one-to-one mapping $\sigma_{I,j} : I \setminus \{r\} \to I \setminus \{j\}$ 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 $J \subseteq I \subseteq \{1, \ldots, n\}$.

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

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

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

This proves part (i).

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

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

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

$$|A| - |J| \ge |A_{\max J}| - |J| \ge k_r + \max J - r - |J| \ge k_r - r,$$

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.,

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

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

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

Proof. Let $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_l\}$ where $a_1 < \ldots < a_k$ and $b_1 < \ldots < b_l$. Put $C = \{a_1 + b_2, \ldots, a_1 + b_{l-1}, a_1 + b_l, \ldots, a_k + b_l\}$. 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 $a_k \ne b_l$, for $i = 2, \ldots, k$ we may suppose that $a_i = b_{f(i)}$ where $i \le f(i) < l$. Because

$$\begin{aligned} S(A, B) \supseteq \{a_1 + b_j : 2 \le j < f(i)\} \cup \{a_i + b_j : j \ne f(i)\} \\ \cup \{a_j + b_l : i < j \le k\}, \end{aligned}$$

we have

$$k + l - 2 = |S(A, B)| \ge (f(i) - 2) + (l - 1) + (k - i) = k + l + f(i) - i - 3,$$

i.e. $a_i \in \{b_i, b_{i+1}\}$. Observe that $a_3 < a_k \le b_{l-1}$. 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 $a_3 + b_{l-1} = a_2 + b_l$, it follows that

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

If $a_2 \ne b_2$, 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 $a_2 + b_{l-2} = a_1 + b_{l-1}, \ldots, a_2 + b_3 = a_1 + b_4$. Thus

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

If $a_3 \ne b_3$, 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. The case $n = 1$ is trivial. Below we let $n \ge 2$ and assume the statement holds for smaller values of n.

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

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

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

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

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

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

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 $J \subseteq \{1, \ldots, n\}$ containing r for which

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

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 $S(A_1, \ldots, A_n) \supseteq S(A'_1, \ldots, A'_n) \cup S_r(A)$ and

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

we have

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

From now on we assume that

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

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 $l \in \{r, \ldots, n\}$ for which $J = \{1, \ldots, l\}$, $\bigcup_{j=1}^{l} A_j = A_l$, $k_l = k_r + (l - r)$ and

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

(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 $|S(-A_1, \ldots, -A_n)| = |S(A_1, \ldots, A_n)| = \sum_{i=1}^{n} |-A_i| - n(n+1)/2 + 1$, similarly we have $-b \in -A_n$. So $b \in A'_n = A_n \setminus \{a\}$. Choose the smallest $s \leq n$ such that $b \in A_s$.

Let $m \in M$. We now 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 $\bigcup_{i=1}^{m} A_i = \bigcup_{i=1, i \neq r}^{m} A_i = A_m$.

Assume $r < n$. 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 $j \leq l$. Since $b = a'_j \in A'_j \subseteq A_j$, we have $j \geq s$. If $l = s$, then $b \in A_l$, $s \leq j \leq l = s$, $s = j \neq r$.

Now suppose that $r < m \leq n$. If $m < t = n$, then $m \in I$ by (12), and $k'_{m+1} - k'_m \geq (k_{m+1} - 1) - (k_m - 1) > 1$. By (11), (12) and the induction hypothesis $\bigcup_{i=1, i \neq r}^{m} A'_i = A'_m$. If $1 \leq i < r$, then $A_i = A'_i \subseteq A'_m \subseteq A_m$; if $r < i \leq m$ and $i \in I$, then $A_i = A'_i \cup \{a\} \subseteq A'_m \cup \{a\} = A_m$; if $r < i \leq m$ but $i \notin I$, then $|A_i| \geq k_r + 1 \geq 2$ 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 find 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 $l \leq m$ 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 us check that A_n is an AP except the case $k_1 \leq 3$.

If $r = 1$ then $\min\{k'_i : i \neq r\} = k'_2 \geq k_2 - 1 \geq k_1$, and if $r > 1$ then $\min\{k'_i : i \neq r\} = k'_1 = k_1$. So $\min\{k'_i : i \neq r\} \geq k_1$. Below we assume that $k_1 \geq 4$.

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

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 $a = b_1 < \ldots < b_{k_2} = b$ be all the elements of A_2. If $1 \le i < j \le k_2$, 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

$$\begin{aligned} S(A_1, A_2) &\supseteq S(A_2 \setminus \{b\}, A_2) \\ &\supseteq C = \{b_1 + b_2, \ldots, b_1 + b_{k_2}, b_2 + b_{k_2}, \ldots, b_{k_2-1} + b_{k_2}\}. \end{aligned}$$

As $|S(A_1, A_2)| = k_1 + k_2 - 2 = 2k_2 - 3 = |C|$, we have $S(A_1, A_2) = C = S(A_2 \setminus \{b\}, A_2)$. Clearly $|A_2 \setminus \{b\}| = k_2 - 1 = k_1 \ge 4$, $\min(A_2 \setminus \{b\}) = \min A_2$ and $\max(A_2 \setminus \{b\}) \neq \max A_2$. 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

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

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 $|A_n^- \cap A_n^+| \ge k_{n-1} \ge k_1 \ge 4$. Among $A_1, \ldots, A_{n-1}, A_n^-$, 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 $A_1, \ldots, A_{n-1}, A_n^+$, 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 $|A_n^-| > k_{n-1} > \ldots > k_1$. According to the previous reasoning,

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

Observe that

$$\begin{aligned} \max S(A_1, \ldots, A_{n-1}, A_n^-) &\le \max S(A'_1, \ldots, A'_{n-1}) + \max A_{n-1} \\ &< \max S(\{a'_1\}, \ldots, \{a'_{n-1}\}, A_n \setminus A_n^-), \end{aligned}$$

and that $|S(\{a'_1\}, \ldots, \{a'_{n-1}\}, A_n \setminus A_n^-)| = |A_n \setminus A_n^-|$. So

$$\begin{aligned} |S(A_1, \ldots, A_n)| &\ge |S(A_1, \ldots, A_{n-1}, A_n^-)| + |S(\{a'_1\}, \ldots, \{a'_{n-1}\}, A_n \setminus A_n^-)| \\ &\ge \sum_{i=1}^{n-1} k_i + |A_n^-| - \frac{n(n+1)}{2} + 1 + (k_n - |A_n^-|) \\ &= \sum_{i=1}^{n} k_i - \frac{n(n+1)}{2} + 1. \end{aligned}$$

Since (10) holds, we must have

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

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

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

Assume that $s<r=n$. Then both r^+ and $s^+=s$ are less than n. If $A_n^+\ne A_{n-1}$, 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 $A_n\setminus\{b\}$ by the above, therefore A_n forms an AP. If $n=2$, then $s=1$, $\min(-A_1)=\min(-A_2)$, $\max(-A_1)\ne\max(-A_2)$ (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 $-A_1,\dots,-A_n$ instead of $A_1,\dots,A_n$, we see 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 $A_n^+\ne A_{n-1}$, and A_n^- forms an AP if $A_n^-\ne A_{n-1}$. Providing $n>2$, both A_n^+ and A_n^- are APs, therefore A_n forms an AP. When $n=2$, if $A_n=A_2$ is not an AP, then A_2^- or A_2^+ coincides with A_1, hence $\min A_1=\min A_2^-=\min A_2$ and $\max A_1\ne\max A_2$ (since $s=2$), or $\min(-A_1)=\min(-A_2^+)=\min(-A_2)$ and $\max(-A_1)\ne\max(-A_2)$ (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 the Theorem is finished.

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] —, —, —, *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*, ibid. 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. Erdős 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, New York, 1994, 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

Received on 1.3.1998
and in revised form on 9.8.1998 (3346)