Journal de Théorie des Nombres
de Bordeaux **17** (2005), 181–193

Restricted set addition in Abelian groups: results and conjectures

par Vsevolod F. LEV

Résumé. Nous présentons un ensemble de conjectures imbriquées qui peuvent être considérées comme des analogues pour l'addition restreinte des théorèmes classiques dûs à Kneser, Kemperman et Scherk. Les connections avec le théorème de Cauchy-Davenport, la conjecture d'Erdős-Heilbronn et la méthode polynomiale d'Alon-Nathanson-Ruzsa sont étudiées.

Cet article ne suppose pas d'expertise de la part du lecteur et peut servir d'introduction au sujet.

Abstract. We present a system of interrelated conjectures which can be considered as restricted addition counterparts of classical theorems due to Kneser, Kemperman, and Scherk. Connections with the theorem of Cauchy-Davenport, conjecture of Erdős-Heilbronn, and polynomial method of Alon-Nathanson-Ruzsa are discussed.

The paper assumes no expertise from the reader and can serve as an introduction to the subject.

0. Quick start.

Is it true that for any sets A and B of residues modulo a positive integer, such that $A \cap (-B) = \{0\}$, there are at least $|A| + |B| - 3$ residues representable as $a + b$ with $a \in A$, $b \in B$, and $a \neq b$? We believe that the answer is "yes", and there is not much exaggeration in saying that the goal of this paper is to explain why this conjecture is so exciting, where it came from, and what it is related to.

In the next section we review three remarkable theorems on set addition due to Cauchy, Davenport, Kneser, Kemperman, and Scherk; these theorems are of utmost importance in what follows. In Section 2 we briefly discuss the conjecture of Erdős-Heilbronn (presently the theorem of Dias da Silva-Hamidoune) which is the restricted addition analog of the theorem of Cauchy-Davenport. In Sections 3 and 4 we consider conjectural restricted addition counterparts of the theorems of Kneser and Kemperman-Scherk, respectively.

1. Historical background, I. Unrestricted set addition.

Additive combinatorial number theory is a branch of mathematics which can be traced down to its very roots. It is generally believed that the first result in this area is a well-known theorem due to Cauchy and Davenport.

Let A and B be non-empty sets of residues modulo a prime p, and denote by $A+B$ the set of all residues representable as a sum $a+b$ with $a \in A$ and $b \in B$. Given the number of elements in A and B (and no other information on the two sets), how small can $A+B$ be? If A and B are arithmetic progressions with the same common difference, then it is immediately seen that either $A+B$ contains all p residues, or otherwise $|A+B| = |A|+|B|-1$. The theorem of Cauchy-Davenport shows that this is the extremal case.

Theorem 1.1 (Cauchy-Davenport)**.** *For any non-empty sets A and B of residues modulo a prime p we have*

$$|A+B| \ge \min\{|A|+|B|-1,\, p\}.$$

This theorem was first established in 1813 by Cauchy (see [C13]) who used it to give an alternative proof of the result of Lagrange that the congruence $ax^2+by^2+c \equiv 0 \pmod p$ is solvable for any non-zero residues a, b, and c. (Lagrange's result is actually a lemma in his proof of the famous four-squares theorem.) In 1935 Theorem 1.1 was independently re-discovered by Davenport (see [D35, D47]) as a "modulo p analog" of a conjecture of Landau and Schnirelmann concerning the density of the sum of integer sequences.

Among numerous generalizations of the theorem of Cauchy-Davenport the deepest and most powerful is, undoubtedly, an extension onto arbitrary abelian groups due to Kneser. Loosely speaking, it says that if A and B are finite, non-empty subsets of an abelian group, then "normally" there are at least $|A|+|B|-1$ group elements representable as $a+b$ with $a \in A$ and $b \in B$. In other words, letting

$$A+B := \{a+b\colon a \in A,\ b \in B\} \tag{1.1}$$

we have $|A+B| \ge |A|+|B|-1$, unless there are some "special reasons" for this to fail. To state Kneser's theorem precisely we have to introduce some notation.

Throughout the paper, for subsets A and B of an abelian group G we define the sum set $A+B$ by (1.1). (For an element $b \in G$ we write $A+b$ instead of $A+\{b\}$.) Notice, that if H is a subgroup of G and if for $C \subseteq G$ we denote by $\overline{C}$ the image of C under the canonical homomorphism $G \to G/H$, then $\overline{A+B} = \overline{A}+\overline{B}$ for any $A, B \subseteq G$. The *period* (or *stabilizer*) of a set $C \subseteq G$ is defined by $H(C) := \{g \in G\colon C+g = C\}$, and C is said to be *periodic* or *aperiodic* according to whether $H(C) \ne \{0\}$ or $H(C) = \{0\}$.

The following properties of the period are almost immediate and yet worth to be recorded explicitly:

(i) $H(C)$ is a subgroup of G;
(ii) $H(C) = G$ if and only if $C = \varnothing$ or $C = G$;
(iii) C is a union of cosets of $H(C)$. Indeed, for a subgroup $H \leq G$ the set C is a union of H-cosets if and only if $H \leq H(C)$;
(iv) $C/H(C)$ (the image of C under the canonical homomorphism $G \to G/H(C)$) is an aperiodic subset of the quotient group $G/H(C)$;
(v) if C is finite, then so is $H(C)$ and $|C| = |C/H(C)|\,|H(C)|$;
(vi) for any $B \subseteq G$ we have $H(C) \leq H(B + C)$.

Theorem 1.2 (Kneser, [Kn53, Kn55]; see also [Ma65])**.** *Suppose that A and B are finite non-empty subsets of an abelian group satisfying*

$$|A + B| \leq |A| + |B| - 1. \tag{1.2}$$

Then letting $H := H(A + B)$ we have

$$|A + B| = |A + H| + |B + H| - |H|.$$

Corollary 1.1. *Let A and B be finite non-empty subsets of an abelian group. If $|A + B| < |A| + |B| - 1$, then $A + B$ is periodic.*

Notice that Kneser's theorem implies readily the theorem of Cauchy-Davenport: for the only non-empty periodic subset of a group of residues modulo a prime is the group itself.

Theorem 1.2 shows that any pair of subsets A, B of an abelian group G, satisfying (1.2), can be obtained by "lifting" subsets $\overline{A}, \overline{B}$ of a quotient group such that

$$|\overline{A} + \overline{B}| = |\overline{A}| + |\overline{B}| - 1. \tag{1.3}$$

Indeed, let $H := H(A+B)$ and denote by $\overline{A}$ and $\overline{B}$ the images of A and B, respectively, under the canonical homomorphism $G \to G/H$. Then (1.3) holds by Theorem 1.2, and furthermore,

$$(|A + H| - |A|) + (|B + H| - |B|) = |H| + (|A + B| - |A| - |B|) < |H|.$$

Thus A and B are obtained from $A + H$ and $B + H$ (which are the full inverse images of $\overline{A}$ and $\overline{B}$ in G) by removing less than $|H|$ elements totally.

It is not difficult to see that Theorem 1.2 is equivalent to Corollary 1.1 in the sense that the former can be easily deduced from the latter. We leave the deduction as an exercise to the interested reader.

The third result we discuss in this section will be referred to as the theorem of Kemperman-Scherk. It relates the size of the sum set $A + B$ to the number of representations of group elements as $a + b$ with $a \in A$ and $b \in B$.

Assuming that finite subsets A and B of an abelian group G are fixed and implicit from the context, for every $c \in G$ we write

$$\nu(c) := |\{(a,b)\colon c = a+b,\ a \in A,\ b \in B\}|,$$

the number of representations of c as a sum of an element of A and an element of B. (A more precise but heavier notation would be $\nu_{A,B}(c)$.) Evidently, $\nu(c) > 0$ if and only if $c \in A+B$. The theorem of Kemperman-Scherk shows that if $A+B$ is small, then any $c \in A+B$ has many representations in the above indicated form.

Theorem 1.3 (Kemperman-Scherk)**.** *Let A and B be finite non-empty subsets of an abelian group. Then $\nu(c) \ge |A| + |B| - |A+B|$ for any $c \in A+B$. In other words,*

$$|A+B| \ge |A| + |B| - \min_{c \in A+B} \nu(c).$$

The history of Theorem 1.3 dates back to 1951 when Moser proposed in the American Mathematical Monthly ([Mo51]) a problem which can be formulated as follows:

> *For finite subsets A and B of the torus group $\mathbb{R}/\mathbb{Z}$ such that $0 \in A \cap B$ and $\nu(0) = 1$, prove that $|A+B| \ge |A| + |B| - 1$.*

In 1955 (see [S55]) Scherk published a solution which actually assumed A and B to be subsets of an arbitrary abelian group (not necessarily the torus group). The argument of Scherk requires very minor modifications to yield the proof of Theorem 1.3, but it was not until Kemperman's papers [Ke56, Ke60] that the full version of the theorem has been stated. Indeed, in the second of the two papers Kemperman ascribes the theorem to Scherk; on the other hand, in his solution of Moser's problem Scherk indicates that the argument follows the ideas of his earlier joint paper with Kemperman. With this in mind, we believe that the theorem should bear both names.

Here is a short proof of the theorem of Kemperman-Scherk based on Kneser's theorem.

Proof of Theorem 1.3. If $|A+B| \ge |A| + |B| - 1$, the assertion is trivial; suppose that $|A+B| = |A| + |B| - k$ with some $k \ge 2$ and show that $\nu(c) \ge k$ for any $c \in A+B$. Let $H := H(A+B)$ and fix a representation $c = a_0 + b_0$ with $a_0 \in A$ and $b_0 \in B$. By Kneser's theorem we have

$$\begin{aligned}
|(A-a_0) \cap H| + |(b_0 - B) \cap H| &= |A \cap (a_0+H)| + |B \cap (b_0+H)| \\
&= 2|H| - |(a_0+H) \setminus A| - |(b_0+H) \setminus B| \\
&\ge 2|H| - |(A+H) \setminus A| - |(B+H) \setminus B| \\
&= |H| + (|A| + |B| - |A+B|) \\
&= |H| + k
\end{aligned}$$

and by the boxing principle, $(A - a_0) \cap H$ and $(b_0 - B) \cap H$ share at least k common elements. However, each pair (a, b) such that $a \in A$, $b \in B$, and $a - a_0 = b_0 - b$ yields a solution to $a + b = c$, whence $\nu(c) \geq k$ as claimed. □

2. Historical background, II. Restricted set addition.

Return for a moment to the theorem of Cauchy-Davenport, but instead of counting *all* sums $a + b$ consider only those with $a \neq b$. How many are there? This question was asked by Erdős and Heilbronn in the mid sixties (of the twentieth century).

For subsets A and B of an abelian group G write

$$A\dot{+}B := \{a + b\colon a \neq b,\ a \in A,\ b \in B\}.$$

In contrast with the regular sum set $A+B$, the sum set $A\dot{+}B$ is often called "restricted". Suppose that G is the group of residues modulo a prime p. If A and B coincide and form an arithmetic progression modulo p with at least two terms, then either $A\dot{+}B$ contains all residues (this holds if $|A| = |B| \geq (p + 3)/2$), or $|A\dot{+}B| = |A| + |B| - 3$. Erdős and Heilbronn conjectured that in fact, for any non-empty sets A and B of residues modulo a prime p one has

$$|A\dot{+}B| \geq \min\{|A| + |B| - 3,\ p\}.$$

This conjecture was frequently mentioned by Erdős in his talks, but it seems that it was first published in [EG80] only. It took about 30 years until Dias da Silva and Hamidoune, using advanced linear algebra tools, proved the conjecture; see [DH94]. Analyzing their proof, Alon, Nathanson, and Ruzsa proposed in [ANR95, ANR96] an elegant and powerful "polynomial method" which has simplified the original argument drastically. Moreover, their method, further developed by Alon in [A99], allows one to handle a variety of unrelated combinatorial problems. The basic idea of the method, as applied to the problem of Edős-Heilbronn, is to consider the polynomial $P(x, y) := (x-y)\prod_{c\in A\dot{+}B}(x+y-c)$ over the field $\mathrm{GF}(p)$. We have $P(a, b) = 0$ for any pair (a, b) such that $a \in A$ and $b \in B$, and since the polynomial has "many roots" its degree (which is $|A\dot{+}B| + 1$) is to be large.

Another approach is pursued in [FLP99, L00a] where combinatorial considerations are combined with the character sum technique. For instance, in [L00a, Theorem 2] we were able to establish the structure of sets A of residues modulo a prime p such that $200 \leq |A| \leq p/50$ and $|A\dot{+}A| \leq 2.18|A| - 6$.

3. Restricted set addition and Kneser's theorem

Once we have a restricted set addition analog of the Cauchy-Davenport theorem, it is most natural to ask for such an analog for Kneser's theorem.

Let G be an abelian group. For two finite non-empty subsets $A, B \subseteq G$, how large can the restricted sum set $A\dot{+}B$ be given that it is aperiodic? Set $G_0 := \{g \in G\colon 2g = 0\}$, the subgroup of G consisting of zero and all elements of order two. For brevity we write $L_G := |G_0|$. In connection with restricted addition problems this invariant of the group G was introduced in [L00a, L00b, L01]; its importance stems from the fact that for any $c \in A+B$ such that $\nu(c) > L_G$ we have $c \in A\dot{+}B$. Fix an integer $n \geq 2$ and an element $d \in G$ of order at least $2n-1$ and let $A = B := \{0, d, \dots, (n-1)d\} + G_0$. Then $A+B = \{0, d, \dots, 2(n-1)d\}+G_0$ and $A\dot{+}B = (A+B)\setminus\{0, 2(n-1)d\}$ so that $|A\dot{+}B| = (2n-1)|G_0|-2 = |A|+|B|-(L_G+2)$; on the other hand, it is easily verified that $A\dot{+}B$ is aperiodic. Based on numerical evidence and several particular cases which we can settle, we conjectured in [L00a] that if $|A\dot{+}B| < |A| + |B| - (L_G + 2)$ holds for finite non-empty subsets $A, B \subseteq G$, then $A\dot{+}B$ is periodic and indeed, we have $A\dot{+}B = A + B$.

Conjecture 3.1. *Let G be an abelian group, and let A and B be finite non-empty subsets of G satisfying*

$$|A\dot{+}B| < |A| + |B| - (L_G + 2). \tag{3.1}$$

Then $A\dot{+}B = A + B$.

For some partial results towards the proof of Conjecture 3.1 the reader is referred to [L00a, L00b, L01]. In particular, [L00b, Theorem 4] implies that for any finite non-empty subsets $A, B \subseteq G$ such that $A\dot{+}B \neq A+B$ one has $|A\dot{+}B| > (1-\delta)(|A|+|B|)-(L_G+2)$, where $\delta = |A||B|/(|A|+|B|)^2 \leq 0.25$. Also, we have verified computationally Conjecture 3.1 for all cyclic groups $G = \mathbb{Z}_l$ with $l \leq 25$, and in the case $A = B$ with $l \leq 36$. The reader will check easily that the conjecture is valid for all cyclic groups of prime order (when it is equivalent to the Erdős-Heilbronn conjecture); for the infinite cyclic group (then $|A\dot{+}B| \geq |A| + |B| - 3$ for any finite, non-empty $A, B \subseteq G$) and as one can deduce from it, for all torsion-free abelian groups; and for elementary 2-groups (in which case the assertion is trivial).

In fact, Conjecture 3.1 is not an analog, but rather a counterpart of Kneser's theorem. It shows that either the restricted sum set $A\dot{+}B$ is "large", or it coincides with the regular sum set $A + B$ and then the well-established machinery of Kneser's theorem can be used to study it.

Remark. The conditions

(i) $A\dot{+}B = A + B$

and

(ii) $\nu(c) \geq 2$ for any $c \in A + B$

are not to be confused with each other. It is not difficult to see that if $A = B$, then (i) implies (ii); on the other hand, if G has no elements of even order, then (ii) implies (i). However, in general none of the two

conditions follows from another one: consider, for instance, the subsets $A = \{0, 1, 2\}$, $B = \{10, 11, 12\}$ of the additive group of integers (where (i) holds but (ii) does not), and the subsets $A = B = \{1, 2, 4, 5\}$ of the group of residues modulo 6 (here (ii) holds but (i) fails).

In conjunction with the theorems of Kneser and Kemperman-Scherk, Conjecture 3.1 readily implies

Corollary 3.1. *Let G be an abelian group, and let A and B be finite nonempty subsets of G satisfying* (3.1). *Then*

(i) $\min_{c\in A+B} \nu(c) \geq 3 + L_G$;
(ii) $|H(A\dot{+}B)| \geq 3 + L_G$;
(iii) $|A\dot{+}B| = |A + H| + |B + H| - |H|$, *where* $H := H(A\dot{+}B)$.

Proof. (i) If there were $c \in A + B$ such that $\nu(c) \leq 2 + L_G$, then by the theorem of Kemperman-Scherk we had $|A + B| \geq |A| + |B| - (2 + L_G)$. However, $|A + B| = |A\dot{+}B| < |A| + |B| - (2 + L_G)$ by Conjecture 3.1.

(ii) Follows from (iii) and the trivial estimates $|A+H| \geq |A|$, $|B+H| \geq |B|$.

(iii) Immediate from Kneser's theorem and since $|A + B| = |A\dot{+}B| < |A| + |B| - 1$. □

Each one of conditions (i)–(iii) of the last corollary, along with assumption (3.1), implies that $A\dot{+}B$ is periodic. For (ii) and (iii) this is obvious; as to (i), it gives $A\dot{+}B = A + B$, and periodicity of this sum follows from Kneser's theorem. This allows us to state yet another corollary of Conjecture 3.1, which is parallel to Corollary 1.1.

Corollary 3.2. *Let G be an abelian group, and let A and B be finite nonempty subsets of G satisfying* (3.1). *Then $A\dot{+}B$ is periodic.*

Just as Corollary 1.1 is equivalent to Kneser's theorem, Corollary 3.2 is equivalent to Conjecture 3.1.

Proposition 3.1. *Corollary 3.2 implies Conjecture 3.1.*

Proof. Suppose that Corollary 3.2 holds true and let A and B be finite nonempty subsets of an abelian group G, satisfying (3.1). We use induction by $|A\dot{+}B|$ to prove that $A\dot{+}B = A + B$. The case $|A\dot{+}B| = 1$ is easy in view of $|A\dot{+}B| \geq \max\{|A|, |B|\} - 1$, and we assume that $|A\dot{+}B| \geq 2$.

Write $D := \{d \in A \cap B\colon 2d \notin A\dot{+}B\}$ so that $A + B = (A\dot{+}B) \cup \{2d\colon d \in D\}$, the two sets at the right-hand side being disjoint. Our goal is to show that $D = \varnothing$. Assuming the opposite, we set $n := |\{2d\colon d \in D\}|$ and $k := \min_{d\in D} \nu(2d)$, so that $1 \leq k \leq L_G$ and $|D| \geq kn$. Notice, that by the theorem of Kemperman-Scherk we have $|A + B| \geq |A| + |B| - k$, whence $n = |A + B| - |A\dot{+}B| \geq L_G + 3 - k$.

Let $H := H(A\dot{+}B)$, so that $|H| > 1$ by Corollary 3.2. We claim that $(d + H) \cap A = (d + H) \cap B = \{d\}$ for any $d \in D$. Indeed, assuming, say,

$a \in d + H$, $a \in A \setminus \{d\}$, we get $d \in a + H$, whence $2d \in a + d + H \subseteq A\dot{+}B$ (in view of $a + d \in A\dot{+}B$), a contradiction. This shows that for all $d \in D$ the sets $(d + H) \setminus \{d\}$ are disjoint with A. Moreover, they are disjoint with each other: for if $(d' + H) \cap (d'' + H) \neq \varnothing$, then $d'' \in (d' + H) \cap A$, hence $d'' = d'$. Thus

$$|A + H| \geq |A| + (|H| - 1)|D| \geq |A| + (|H| - 1)kn,$$

and similarly $|B + H| \geq |B| + (|H| - 1)kn$. Therefore, if $\overline{A}, \overline{B}$, and $\overline{A\dot{+}B}$ denote the images of A, B, and $A\dot{+}B$, respectively, under the canonical homomorphism $G \to G/H$, then

$$\begin{aligned}
|\overline{A}\dot{+}\overline{B}||H| &\leq |\overline{A\dot{+}B}||H| \\
&= |A\dot{+}B| \\
&\leq |A| + |B| - (L_G + 3) \\
&\leq |A + H| + |B + H| - 2kn(|H| - 1) - (L_G + 3) \\
&= (|\overline{A}| + |\overline{B}| - (L_G + 3))|H| - (2kn - L_G - 3)(|H| - 1) \\
&< (|\overline{A}| + |\overline{B}| - (L_G + 3))|H|,
\end{aligned}$$

the last inequality following from

$$2kn \geq 2k(L_G + 3 - k) \geq 2(L_G + 2) > L_G + 3.$$

As $L_G \geq L_{G/H}$, it follows that $|\overline{A}\dot{+}\overline{B}| < |\overline{A}| + |\overline{B}| - (L_{G/H} + 3)$, hence $\overline{A}\dot{+}\overline{B} = \overline{A} + \overline{B}$ by the induction hypothesis. However, for any $d \in D$ its image $\overline{d} \in G/H$ satisfies $2\overline{d} \in (\overline{A} + \overline{B}) \setminus (\overline{A}\dot{+}\overline{B})$, a contradiction. □

In fact, it suffices to prove Conjecture 3.1 (or Corollary 3.2 to which it is equivalent) in the particular case $B \subseteq A$ only.

Conjecture 3.1′. *Let G be an abelian group, let A and B be finite non-empty subsets of G, satisfying* (3.1)*, and suppose that $B \subseteq A$. Then $A\dot{+}B = A + B$.*

Corollary 3.2′. *Let G be an abelian group, let A and B be finite non-empty subsets of G, satisfying* (3.1)*, and suppose that $B \subseteq A$. Then $A\dot{+}B$ is periodic.*

Proposition 3.2. *Corollary 3.2′ implies Conjecture 3.1′, which implies Conjecture 3.1, which implies Corollary 3.2. Thus, all four results are equivalent in the sense that each of them implies the other three.*

Proof. Exactly as in the proof of Proposition 3.1, one can see that Conjecture 3.1′ follows from Corollary 3.2′. We show that Conjecture 3.1 follows from Conjecture 3.1′.

To this end, suppose that Conjecture 3.1′ is true, and let A and B be finite non-empty subsets of an abelian group G, satisfying (3.1). We want

to show that $A\dot{+}B = A + B$. Set $A^* := A \cup B$ and $B^* := A \cap B$; we can assume that $B^* \neq \varnothing$, for otherwise the assertion is immediate. Plainly, we have $|A^*| + |B^*| = |A| + |B|$ and $A^*\dot{+}B^* \subseteq A\dot{+}B$, whence $|A^*\dot{+}B^*| < |A^*| + |B^*| - (L_G + 2)$. From Conjecture 3.1′ we derive $A^*\dot{+}B^* = A^* + B^*$, and this shows that for any $d \in B^*$ there exist $a^* \in A^*$, $b^* \in B^*$ such that $a^* \neq b^*$ and $2d = a^* + b^*$. Since $a^* + b^* \in A^*\dot{+}B^* \subseteq A\dot{+}B$, it follows that $A\dot{+}B = A + B$, as required. □

Remark. It is easily seen that, without loss of generality, one can assume $0 \in A \cap B$, $0 \notin A\dot{+}B$ in Conjectures 3.1 and 3.1′ and Corollaries 3.2 and 3.2′. In fact, it is the last corollary with this extra assumption that has been verified computationally for cyclic groups of small order. Some further reformulations are possible. For instance, Conjecture 3.1 is equivalent to the assertion that if $A, B \subseteq G$ satisfy $0 \in A \cap B$ and $0 \notin A\dot{+}B$, then $|A\dot{+}B| \geq |A| + |B| - (L_G + 2)$.

4. Restricted set addition and the theorem of Kemperman-Scherk

What is the restricted addition analog of the theorem of Kemperman-Scherk? That is, how small can $\min_{c \in A+B} \nu(c)$ be for $|A|, |B|$, and $|A\dot{+}B|$ given? Consider an example similar to that constructed in Section 3. Fix a subgroup $H \leq G$ such that $2h = 0$ for any $h \in H$, an integer $n \geq 2$, and an element $d \in G$ of order at least $2n - 1$, and let $A = B := \{0, d, \ldots, (n-1)d\} + H$. Then $|A\dot{+}B| = |A| + |B| - (|H| + 2)$ whence $\nu(0) = |H| = |A| + |B| - 2 - |A\dot{+}B|$.

Conjecture 4.1. *Let A and B be finite non-empty subsets of an abelian group. Then $\nu(c) \geq |A| + |B| - 2 - |A\dot{+}B|$ for any $c \in A + B$. In other words,*

$$|A\dot{+}B| \geq |A| + |B| - 2 - \min_{c \in A+B} \nu(c).$$

As it is the case with Conjecture 3.1, we have verified Conjecture 4.1 computationally for all cyclic groups $\mathbb{Z}_l$ with $l \leq 25$, and in the case $A = B$ with $l \leq 36$. Also, Conjecture 4.1 is true for torsion-free abelian groups (then $|A\dot{+}B| \geq |A| + |B| - 3$ for any finite non-empty A and B), for groups of residues modulo a prime (when it is equivalent to the conjecture of Erdős-Heilbronn), and for elementary abelian 2-groups (follows from the theorem of Kemperman-Scherk and the observation that $A\dot{+}B = (A + B) \setminus \{0\}$ for these groups).

Corollary 4.1. *Let A and B be finite non-empty subsets of an abelian group. If there exists an element $c \in A + B$ with a unique representation as $c = a + b$ ($a \in A$, $b \in B$), then*

$$|A\dot{+}B| \geq |A| + |B| - 3.$$

Proposition 4.1. *Corollary 4.1 implies Conjecture 4.1.*

Proof. Suppose that Corollary 4.1 holds true, and let A and B be finite non-empty subsets of an abelian group. We fix arbitrarily $c \in A+B$ and show that $|A\dot{+}B| \ge |A|+|B|-2-\nu(c)$.

Write $k := \nu(c)$ and let $c = a_1+b_1 = \cdots = a_k+b_k$, where $a_i \in A$, $b_i \in B$ $(i = 1,\dots,k)$, be the k representations of c. Define $B' := B \setminus \{b_1,\dots,b_{k-1}\}$. Then c has a unique representation as $c = a+b'$ with $a \in A$ and $b' \in B'$ (namely, $c = a_k+b_k$), whence by Corollary 4.1

$$|A\dot{+}B| \ge |A\dot{+}B'| \ge |A|+|B'|-3 = |A|+|B|-2-k,$$

as required. □

The following immediate corollary of Conjecture 4.1, as we will see shortly, is equivalent to Conjecture 3.1. The former conjecture, therefore, is a strengthening of the latter.

Corollary 4.2. *Let A and B be finite non-empty subsets of an abelian group G. If there exists an element $c \in A+B$ such that $\nu(c) \le L_G$, then*

$$|A\dot{+}B| \ge |A|+|B|-(L_G+2).$$

Proposition 4.2. *Corollary 4.2 is equivalent to Conjecture 3.1.*

Proof. It is clear that Corollary 3.1 (i), and therefore Conjecture 3.1, implies Corollary 4.2. On the other hand, if Corollary 4.2 holds true and if A and B are finite non-empty subsets of an abelian group G such that $|A\dot{+}B| < |A|+|B|-(L_G+2)$, then $\nu(c) > L_G$ for any $c \in A+B$, whence $A\dot{+}B = A+B$. □

Remark. It is not clear whether Corollary 4.2 implies Conjecture 4.1; in other words, whether Conjectures 3.1 and 4.1 are equivalent.

Corollary 4.2 can be reduced to the case $B \subseteq A$.

Corollary 4.2′. *Let A and B be finite non-empty subsets of an abelian group G, satisfying $B \subseteq A$. If there exists an element $c \in A+B$ such that $\nu(c) \le L_G$, then*

$$|A\dot{+}B| \ge |A|+|B|-(L_G+2).$$

Proposition 4.3. *Corollary 4.2′ is equivalent to Corollary 4.2.*

First proof. Consider the implications

$$\text{(Corollary 4.2}') \Rightarrow \text{(Conjecture 3.1}') \Rightarrow \text{(Conjecture 3.1)} \Rightarrow \text{(Corollary 4.2)}.$$

□

Second proof. We follow the proof of Proposition 3.2. Suppose that Corollary 4.2′ is true and let A and B be finite non-empty subsets of an abelian group G. Assuming that $\min_{c\in A+B}\nu(c) \le L_G$ and $|A\dot{+}B| < |A|+|B|-(L_G+2)$ we will obtain a contradiction.

We observe that $|A+B| \ge |A|+|B|-L_G$ by the theorem of Kemperman-Scherk, whence $A\dot{+}B \ne A+B$. Set $A^* := A\cup B$ and $B^* := A\cap B$. We have $B^* \ne \varnothing$, for otherwise $A\dot{+}B = A+B$. In view of $A^*\dot{+}B^* \subseteq A\dot{+}B$ we obtain

$$|A^*\dot{+}B^*| \le |A\dot{+}B| < |A|+|B|-(L_G+2) = |A^*|+|B^*|-(L_G+2),$$

hence by Corollary 4.2′ every $c\in A^*+B^*$ has at least L_G+1 representations as $c=a^*+b^*$ with $a^*\in A^*$, $b^*\in B^*$. Thus $A^*\dot{+}B^* = A^*+B^*$ and (as in the proof of Proposition 3.2) we conclude that $A\dot{+}B = A+B$, the sought contradiction. □

The reader is urged to compare our next corollary with the question we started our paper with.

Corollary 4.3. *Suppose that A and B are finite subsets of an abelian group satisfying $A\cap(-B)=\{0\}$. Then*

$$|A\dot{+}B| \ge |A|+|B|-3.$$

Proof. From $A\cap(-B)=\{0\}$ it follows that $\nu(0)=1$ (the only representation of 0 being $0=0+0$). □

Remark. It is not difficult to see that if $L_G=1$, then Conjectures 3.1 and 4.1 are equivalent to each other and to Corollary 4.3. Indeed, in this case Corollaries 4.1 and 4.2 coincide; however, we saw that the former of them is equivalent to Conjecture 4.1 and the latter is equivalent to Conjecture 3.1. To verify that Corollary 4.3 implies Conjecture 3.1 for groups G satisfying $L_G=1$, observe that if $A\dot{+}B \ne A+B$ then there exists $c\in A\cap B$ such that $2c\notin A\dot{+}B$; letting $A' := A-c$ and $B' := B-c$ we get $A'\cap(-B')=\{0\}$ whence $|A\dot{+}B| = |A'\dot{+}B'| \ge |A'|+|B'|-3 = |A|+|B|-3$ by Corollary 4.3.

In conclusion, we note that for restricted set addition it may be natural to replace $\nu(c)$ by the "restricted representation function"

$$\dot\nu(c) := |\{(a,b)\colon c=a+b,\ a\ne b,\ a\in A,\ b\in B\}|.$$

From Conjecture 4.1 it is not difficult to derive

Corollary 4.4. *Let A and B be finite non-empty subsets of an abelian group G. Then*

$$|A\dot{+}B| \ge |A|+|B|-L_G-1-\min_{c\in A\dot{+}B}\dot\nu(c).$$

Proof. If $\min_{c\in A+B}\nu(c)\le L_G$ then, assuming Conjecture 4.1, we get

$$|A\dot{+}B|\ge |A|+|B|-2-L_G\ge |A|+|B|-1-L_G-\min_{c\in A\dot{+}B}\dot{\nu}(c).$$

Otherwise we have $A\dot{+}B=A+B$ whence by the theorem of Kemperman-Scherk

$$|A\dot{+}B|=|A+B|\ge |A|+|B|-\min_{c\in A+B}\nu(c).$$

The result now follows from

$$\min_{c\in A+B}\nu(c)\le \min_{c\in A\dot{+}B}\nu(c)\le \min_{c\in A\dot{+}B}(\dot{\nu}(c)+L_G).$$

□

5. A unifying conjecture

When this paper was essentially completed we were able to state a rather general conjecture which is somewhat simpler, yet stronger than Conjectures 3.1 and 4.1.

Conjecture 5.1. *For any finite non-empty subsets A and B of an abelian group we have*

$$|A\dot{+}B|\ge \min\{|A+B|,\,|A|+|B|-1\}-2.$$

In other words, if $|(A+B)\setminus(A\dot{+}B)|\ge 3$, then $|A\dot{+}B|\ge |A|+|B|-3$.

We make several remarks:

(i) It is easily seen that Conjecture 5.1 implies Conjectures 3.1 and 4.1. Moreover, if $L_G=1$ then all three conjectures are *equivalent* to each other.

(ii) The general case reduces easily to the case $0\in B\subseteq A$, $0\notin A\dot{+}B$.

(iii) We have verified Conjecture 5.1 (in fact its special case described in (ii)) for all cyclic groups $\mathbb{Z}_l$ of order $l\le 25$, and assuming equal set summands ($A=B$) of order $l\le 36$.

(iv) Conjecture 5.1 is valid for torsion-free abelian groups (in this case $|A\dot{+}B|\ge |A|+|B|-3$ holds for *any* finite non-empty subsets A and B).

(v) It is valid also for groups of residues modulo a prime (when it is equivalent to Erdős-Heilbronn).

(vi) It is valid also for elementary abelian 2-groups (in these groups we have $A\dot{+}B=(A+B)\setminus\{0\}$ for any A and B).

(vii) Perhaps, using Kemperman's results it is feasible to prove Conjecture 5.1 in the case $|A+B|\le |A|+|B|-1$. This would show that the conjecture is equivalent to the assertion that if $|A\dot{+}B|<|A|+B|-3$, then $|A+B|<|A|+|B|-1$.

References

[A99] N. Alon, *Combinatorial Nullstellensatz. Recent trends in combinatorics (Mťrahźa, 1995)*. Combin. Probab. Comput. **8** (1–2) (1999), 7–29.

[ANR95] N. Alon, M.B. Nathanson, I.Z. Ruzsa, *Adding distinct congruence classes modulo a prime*. American Math. Monthly **102** (1995), 250–255.

[ANR96] N. Alon, M.B. Nathanson, I.Z. Ruzsa, *The polynomial method and resricted sums of congruence classes*. J. Number theory **56** (1996), 404–417.

[C13] A. Cauchy, *Recherches sur les nombres*. Jour. Ecole polytechn. **9** (1813), 99–116.

[D35] H. Davenport, *On the addition of residue classes*. J. London Math. Soc. **10** (1935), 30–32.

[D47] ———, *A historical note*. J. London Math. Soc. **22** (1947), 100–101.

[DH94] J.A. Dias da Silva, Y.O. Hamidoune, *Cyclic spaces for Grassmann derivatives and additive theory*. Bull. London Math. Soc. **26** (1994), 140–146.

[EG80] P. Erdős, R. Graham, *Old and new problems and results in combinatorial number theory*. L'Enseignement Mathématique, Geneva (1980).

[FLP99] G. Freiman, L. Low, J. Pitman, *Sumsets with distinct summands and the conjecture of Erdős-Heilbronn on sums of residues*. Astérisque **258** (1999), 163–172.

[Ke56] J.H.B. Kemperman, *On complexes in a semiroup*. Indag. Math. **18** (1956), 247–254.

[Ke60] ———, *On small sumsets in an abelian group*. Acta Math. **103** (1960), 63–88.

[Kn53] M. Kneser, *Abschätzung der asymptotischen Dichte von Summenmengen*. Math. Z. **58** (1953), 459–484.

[Kn55] ———, *Ein Satz über abelsche Gruppen mit Anwendungen auf die Geometrie der Zahlen*. Math. Z. **61** (1955), 429–434.

[L00a] V.F. Lev, *Restricted set addition in groups, I. The classical setting*. J.London Math. Soc. (2) **62** (2000), 27–40.

[L00b] ———, *Restricted set addition in groups, II. A generalization of the Erdős-Heilbronn conjecture*. Electron. J. Combin. **7** (1) (2000), Research Paper 4, 10 pp. (electronic).

[L01] ———, *Restricted set addition in groups, III. Integer sumsets with generic restrictions*. Periodica Math. Hungarica **42** (2001), 89–98.

[Ma65] H. B. Mann, *Addition Theorems: The Addition Theorems of Group Theory and Number Theory*. Interscience Publishers, a division of John Wiley and Sons, New York, 1965.

[Mo51] L. Moser, *Problem 4466*. American Math. Monthly **58** (10) (1951), 703.

[S55] P. Scherk, *Distinct elements in a set of sums (solution to Problem 4466)*. American Math. Monthly **62** (1) (1955), 46–47.

Vsevolod F. Lev
Department of Mathematics
The University of Haifa at Oranim
Tivon 36006, Israel
E-mail : seva@math.haifa.ac.il
URL: http://math.haifa.ac.il/~seva/