SYMMETRIC KNESER'S THEOREM WITH TRIOS AND 3-TRANSFORM

DAVID J. GRYNKIEWICZ AND VSEVOLOD F. LEV

Abstract. We give a new equivalent restatement and a new proof in terms of trios to the classical Kneser's theorem. In the finite case, our restatement takes the following, particularly symmetric shape: if A, B, and C are subsets of a finite abelian group G such that $A+B+C \neq G$, then, denoting by H the period of the sumset $A+B+C$, we have

$$|A|+|B|+|C| \le |G|+|H|.$$

The proof is based on an extension of the familiar Dyson transform onto set systems containing three (or more) sets.

1. Introduction: Kneser's Theorem and Trios

For a subset S of an abelian group, let $\pi(S)$ denote the *period (stabilizer)* of S; that is, $\pi(S)$ is the subgroup consisting of all those group elements g with $S+g=S$.

One of the most basic yet robust results in additive combinatorics, Kneser's theorem, is standardly formulated as follows.

Theorem 1 (Kneser [Kn53, Kn55])**.** *If A and B are finite subsets of an abelian group, then*

$$|A+B| \ge |A|+|B|-|\pi(A+B)|.$$

The goal of our paper is to restate Kneser's theorem in a "symmetric" form, and give our restatement an independent proof in terms of trios.

Following [BDM15], by a *trio* in an abelian group G, we mean a triple (A,B,C) of non-empty subsets of G such that $A+B+C \neq G$ and each of A, B, and C is either finite or co-finite in G. Since the sum of two co-finite subsets of an infinite group is the whole group, every trio can have at most one infinite component.

The *deficiency* of a trio (A,B,C), denoted $\delta(A,B,C)$, is defined as follows. If A,B, and C are all finite, while the underlying group G is infinite, then we let $\delta(A,B,C)=-\infty$; this case is in fact of no importance as we are interested in "large" trios. Otherwise, if G is infinite, then exactly one of the sets A, B, and C is co-finite, while the other two are finite, and if G is finite, then all these three sets are both finite and co-finite. In either case, we can rename the sets involved so that A

is co-finite while B and C are finite, and with this assumption, we let $\delta(A,B,C) = -|\overline{A}| + |B| + |C|$, where for a set $S \subseteq G$, we write $\overline{S} := G \setminus S$. Notice that, if G is finite, then

$$-|\overline{A}| + |B| + |C| = |A| - |\overline{B}| + |C| = |A| + |B| - |\overline{C}| = |A| + |B| + |C| - |G|,$$

showing that deficiency is well-defined in this case.[1]

We can now present our restatement of Kneser's theorem (cf. [BDM15, Theorem 3.6]).

Theorem 2. *For any trio (A,B,C), we have $\delta(A,B,C) \le |\pi(A+B+C)|$.*

Observe that, in the finite case, Theorem 2 can be given a particularly simple shape.

Theorem 2′. *If A, B, and C are subsets of a finite abelian group G such that $A+B+C \neq G$, then*

$$|A| + |B| + |C| \le |G| + |\pi(A+B+C)|.$$

We keep using the convention that, for a set S, the complement of S in the underlying group is denoted by $\overline{S}$. The equivalence of Theorems 1 and 2 is easy to establish using the following simple fact.

Claim 1. *For any subset S of an abelian group, we have $\pi(S-\overline{S}) = \pi(S)$. Moreover, if S is either finite or co-finite, then indeed*

$$S - \overline{S} = \overline{S} - S = \overline{\pi(S)}.$$

Proof. For a group element g, we have $g \notin \pi(S)$ if and only if either $S+g \not\subseteq S$, or $S-g \not\subseteq S$. The former relation is equivalent to $g \in \overline{S}-S$, and the latter to $g \in S-\overline{S}$. We thus conclude that $\overline{\pi(S)} = (S-\overline{S}) \cup (\overline{S}-S)$, whence, in view of $S-\overline{S} = -(\overline{S}-S)$,

$$\pi(S) = \pi(\overline{\pi(S)}) = \pi\big((S-\overline{S}) \cup (\overline{S}-S)\big) \ge \pi(S-\overline{S}) \ge \pi(S),$$

implying the first assertion.

Furthermore, if S is either finite or co-finite, then $g \notin \pi(S)$ if and only if $S+g \not\subseteq S$, which yields $\overline{\pi(S)} = \overline{S} - S$. Switching the roles of S and $\overline{S}$ and observing that $\pi(S) = \pi(\overline{S})$, we get $\overline{\pi(S)} = S - \overline{S}$, and the second assertion follows. □

To derive Theorem 2 from Theorem 1, assume that (A,B,C) is a trio with B and C finite and A co-finite, and fix a group element $g \notin A+B+C$; we then have $B+C \subseteq g-\overline{A}$, whence, by Theorem 1,

$$\delta(A,B,C) \le |B| + |C| - |B+C| \le |\pi(B+C)| \le |\pi(A+B+C)|.$$

[1]The last expression suggests that *redundancy* might be a more intuitive term than *deficiency*; however, we stick with the terminology of [BDM15].

Conversely, assuming Theorem 2, and given finite, non-empty subsets A and B of an abelian group G, let $C := -\overline{A+B}$. If $C = \varnothing$, then $A + B = G$ and $|A + B| \geq |A| + |B| - |\pi(A+B)|$ is immediate. If $C \neq \varnothing$, then $0 \notin A+B-\overline{A+B} = A+B+C$, showing that (A, B, C) is a trio, and from Theorem 2 and Claim 1 (applied with $S = A + B$), it follows that

$$|A| + |B| - |A + B| = \delta(A, B, C) \leq |\pi(A + B + C)| = |\pi(A + B)|.$$

We have shown that Theorems 1 and 2 are equivalent in the sense that each of them follows easily from the other one. Our goal is to give Theorem 2 an independent, "symmetric" proof. As preparation steps, in the next section we collect some background facts about trios, and in Section 3 we develop a multiple-set transform, the basic tool employed in our proof. The proof itself is then presented in Sections 4 and 5. Concluding remarks are gathered in Section 6.

2. Trios

In this section we provide the background about trios needed for the proof of Theorem 2. Most of the material here is contained, in this or another form, in [BDM15].

Refining the definition from the previous section, for an abelian group G and an element $g \in G$, we say that a triple (A, B, C) of non-empty subsets of G is a *g-trio* if $g \notin A + B + C$ and each of A, B, and C is either finite or co-finite in G.

The *period* of the trio (A, B, C) is the period of the sumset $A + B + C$. Since this sumset is either a finite or a co-finite subset of the underlying group, the period of a trio is always finite. The trio (A, B, C) is *aperiodic* if its period is the trivial subgroup, and *periodic* otherwise. It is easily verified that if (A, B, C) is a trio in an abelian group G, and $H = \pi(A + B + C)$, then the images of A, B, and C under the canonical homomorphism $G \to G/H$ form an aperiodic trio in the quotient group G/H.

Clearly, if (A, B, C) is a g-trio in an abelian group G, then $(A - a, B - b, C - c)$ is a $(g - a - b - c)$-trio in G for any $a, b, c \in G$, and both trios share the same period.

As a direct consequence of Claim 1, we have the following corollary.

Corollary 1. *Suppose that (A, B, C) is a g-trio in an abelian group G. If $C = g - \overline{A+B}$, then, letting $H := \pi(A + B + C)$, we have $A + B + C = G \setminus (g + H)$ and $\pi(C) = H$.*

Furthermore, we have the following lemma.

Lemma 1. *Suppose that (A, B, C) is a g-trio and let $C' := g - \overline{A+B}$. Then (A, B, C') is also a g-trio, $C \subseteq C'$, and $\pi(A + B + C') \leq \pi(A + B + C)$.*

Proof. The definition of C' readily implies that $g \notin A + B + C'$, and that C' is either finite or co-finite (the latter follows from finiteness or co-finiteness of A and B); consequently, (A, B, C') is a g-trio. Since (A, B, C) is a g-trio, we have $g \notin A+B+C$, whence $C \subseteq g - \overline{A+B} = C'$. Finally, by Corollary 1,

$$\pi(A + B + C') = \pi(C') = \pi(A + B) \leq \pi(A + B + C).$$

□

The trio (A, B, C) is *contained* in the trio (A', B', C') if $A \subseteq A'$, $B \subseteq B'$, and $C \subseteq C'$; in this case, the former trio is also said to be a *subtrio* of the latter, and the latter a *supertrio* of the former.

We say that (A, B, C) is a *maximal* g-trio if (in addition to being a g-trio) it is not properly contained in any other g-trio; that is, for any g-trio (A', B', C') with $A \subseteq A'$, $B \subseteq B'$, and $C \subseteq C'$, we actually have $A = A'$, $B = B'$, and $C = C'$. By Lemma 1, for (A, B, C) to be a maximal g-trio, it is necessary and sufficient that $A = g-\overline{B+C}$, $B = g-\overline{C+A}$, and $C = g-\overline{A+B}$. Hence, by Corollary 1, if (A, B, C) is a maximal g-trio, then, letting $H := \pi(A + B + C)$, we have $\pi(A) = \pi(B) = \pi(C) = H$ and $A + B + C = G \setminus (g + H)$. In particular, if (A, B, C) is a maximal *aperiodic* g-trio, then $A + B + C = G \setminus \{g\}$.

Lemma 2. *If (A, B, C) is a maximal g-trio, then it is in fact a maximal f-trio for each $f \notin A + B + C$.*

Proof. Clearly (A, B, C) is an f-trio, whence

$$A \subseteq f - \overline{B+C},\ B \subseteq f - \overline{C+A},\ \text{and } C \subseteq f - \overline{A+B}. \tag{1}$$

On the other hand, since (A, B, C) is a maximal g-trio,

$$A = g - \overline{B+C},\ B = g - \overline{C+A},\ \text{and } C = g - \overline{A+B}. \tag{2}$$

We now claim that none of the inclusions in (1) is strict; for if we had, for instance, $C \subsetneq f - \overline{A+B}$, then in view of (2) this would imply $C \subsetneq (f - g) + C$, which is impossible since C is either finite or co-finite. This shows that (A, B, C) is a maximal f-trio. □

With Lemma 2 in mind, we can speak about maximal trios without indicating the specific value of g. In addition, Lemma 2 shows that (A, B, C) is a maximal g-trio for some group element g if and only if it is a maximal trio; that is, not properly contained in any other trio.

Lemma 3. *For a g-trio (A, B, C), define $A' := g - \overline{B+C}$, and then subsequently $B' := g - \overline{A'+C}$ and $C' := g - \overline{A'+B'}$. Then*

i) $A \subseteq A'$, $B \subseteq B'$, *and* $C \subseteq C'$;
ii) (A', B', C') *is a maximal* g*-trio;*
iii) $\pi(A' + B' + C') \leq \pi(A + B + C)$.

Proof. The fact that (A', B', C') is a g-trio containing (A, B, C), and also the relation $\pi(A'+B'+C') \leq \pi(A+B+C)$, follow readily by repeated application of Lemma 1. To see why (A', B', C') is maximal, notice that, if it is contained in a g-trio (A'', B'', C''), then

$$A' \subseteq A'' \subseteq g - \overline{B'' + C''} \subseteq g - \overline{B' + C'} \subseteq g - \overline{B + C} = A'$$

implying $A'' = A'$, and similarly $B'' = B'$ and $C'' = C'$. □

As it follows from Lemma 3, every aperiodic trio can be embedded into a maximal aperiodic trio.

A trio is said to be *deficient* if its deficiency is positive. To conclude this section, we record the following corollary of Theorem 2.

Corollary 2. *If* (A, B, C) *is a maximal, deficient trio, then we have* $\delta(A, B, C) = |\pi(A + B + C)|$.

Proof. Let $H := \pi(A + B + C)$ and assume for definiteness that B and C are finite. By maximality, we have $\pi(A) = \pi(B) = \pi(C) = H$, so that $|B|$, $|C|$ and $|\overline{A}|$ are all divisible by $|H|$; hence also $\delta(A, B, C)$ is divisible by $|H|$. On the other hand, deficiency gives $\delta(A, B, C) > 0$, and the conclusion now follows from Theorem 2. □

3. The n-transform

In this section we introduce a version of the classical Dyson transform for a set system potentially involving more than two sets. We call it the *n-transform*, where n can be substituted with the actual number of sets; thus, the Dyson transform is the 2-transform, and what we ultimately need for the proof of Theorem 2 is the 3-transform.

Let $\mathcal{A} = (A_\nu)_{\nu \in \mathcal{N}}$ be a system of subsets of some ground set G. For an integer $i \geq 1$, denote by $\tau_i(\mathcal{A})$ the set of all those elements of G belonging to at least i sets from $\mathcal{A}$, and let $\tau(\mathcal{A}) = (\tau_i(\mathcal{A}))_{i \geq 1}$. If $\mathcal{A}$ is finite with, say, $|\mathcal{N}| = n$, then the sets $\tau_i(\mathcal{A})$ are empty for $i > n$, and we then identify $\tau(\mathcal{A})$ with the finite sequence $(\tau_i(\mathcal{A}))_{1 \leq i \leq n}$; notice that, in this case, $\tau_n(\mathcal{A}) = \cap_{\nu \in \mathcal{N}} A_\nu$, and that we always have $\tau_1(\mathcal{A}) = \cup_{\nu \in \mathcal{N}} A_\nu$.

Although we are interested in the situation where $\mathcal{A}$ is a finite sequence of subsets of an *abelian group*, we start with two general set-theoretic properties of the n-transform.

Lemma 4. *If $A_1, \dots, A_n$ are finite sets, then letting $(A_1^*, \dots, A_n^*) = \tau(A_1, \dots, A_n)$, we have*

$$|A_1^*| + \cdots + |A_k^*| \geq |A_1| + \cdots + |A_k| \quad \text{for } k \in [1, n] \tag{3}$$

with equality for $k = n$. If, indeed, equality holds in (3) for each $k \in [1, n]$, *then $A_1 \supseteq \cdots \supseteq A_n$ (whence $A_k^* = A_k$ for all $k \in [1, n]$).*

Proof. The equality $|A_1^*| + \cdots + |A_n^*| = |A_1| + \cdots + |A_n|$ follows by observing that for every element g of the ground set, the number of the sets A_i that contain g is equal to the number of the sets A_i^* containing g. For $1 \leq k < n$, replacing each of the sets $A_{k+1}, \dots, A_n$ with the empty set (which does not affect the sum $|A_1| + \cdots + |A_k|$, and can only make the sum $|A_1^*| + \cdots + |A_k^*|$ smaller), we reduce the situation to the case $k = n$ just considered.

To prove the second assertion, we first notice that if equality holds in (3) for all $k \in [1, n]$, then $|A_1^*| = |A_1|, \dots, |A_n^*| = |A_n|$, and then use induction by n. The case $n = 1$ is immediate, and we assume therefore that $n \geq 2$. Since $A_1^* = A_1 \cup \cdots \cup A_n$, from $|A_1^*| = |A_1|$ we derive that, in fact, $A_1 = A_1^* = A_1 \cup \cdots \cup A_n$, whence $A_k \subseteq A_1$ for all $k \in [2, n]$. Hence, for each $k \in [2, n]$, the set A_k^* consists of all those elements contained in at least $k - 1$ of the sets $A_2, \dots, A_n$; that is, $(A_2^*, \dots, A_n^*) = \tau(A_2, \dots, A_n)$. By the induction hypothesis, we have then $A_2 \supseteq \cdots \supseteq A_n$ and the assertion follows. □

Lemma 5. *For a sequence of sets $(A_1, A_2, \dots)$ to be stable under the n-transform, it is necessary and sufficient that $A_1 \supseteq A_2 \supseteq \dots$.*

Proof. Clearly, the condition is sufficient: if $A_1 \supseteq A_2 \supseteq \dots$, then $\tau_k(A_1, A_2, \dots) = A_k$ for each $k \geq 1$. It is also necessary for, in general, if $(B_1, B_2, \dots)$ is an image of some sequence under the n-transform, then $B_1 \supseteq B_2 \supseteq \cdots$. □

We now turn to the properties of the n-transform specific to subsets of abelian groups.

For integers $a_1, \dots, a_n, b_1, \dots, b_n$, we write $(a_1, \dots, a_n) \prec (b_1, \dots, b_n)$ if $(b_1, \dots, b_n)$ majorizes $(a_1, \dots, a_n)$; that is, if $a_1 + \cdots + a_k \leq b_1 + \cdots + b_k$ for each $k \in [1, n]$, with equality for $k = n$ and strict inequality for at least one $k \in [1, n-1]$. Notice that, if $(a_1, \dots, a_n) \prec (b_1, \dots, b_n)$, then $(a_1, \dots, a_n)$ precedes $(b_1, \dots, b_n)$ also in the lexicographic order $\prec_\ell$.

Lemma 6. *For any finite subsets $A_1, \dots, A_n$ of an abelian group, one of the following holds:*

i) *There exist elements $a_k \in A_k$ $(k \in [1, n])$ such that, letting $(A_1^*, \dots, A_n^*) := \tau(A_1 - a_1, \dots, A_n - a_n)$, we have*

$$(|A_1|, \dots, |A_n|) \prec (|A_1^*|, \dots, |A_n^*|).$$

ii) *We have $A_k - A_k \subseteq \pi(A_{k-1})$ for all $k \in [2,n]$; that is, for each $k \in [2,n]$, the set A_k is contained in a coset of the period of the set A_{k-1}.*

Proof. By Lemma 4, if $(|A_1|,\dots,|A_n|) \prec (|A_1^*|,\dots,|A_n^*|)$ does *not* hold for some specific choice of the elements $a_k \in A_k$, then for each $k \in [2,n]$, we have $A_k - a_k \subseteq A_{k-1} - a_{k-1}$, and therefore $A_k - a_k + a_{k-1} \subseteq A_{k-1}$. If now $a_k \in A_k$ and $a_{k-1} \in A_{k-1}$ can be chosen arbitrarily, this leads to $A_k - A_k + A_{k-1} \subseteq A_{k-1}$, whence $A_k - A_k \subseteq \pi(A_{k-1})$. □

Lemma 7. *If $A_1,\dots,A_n$ are subsets of an abelian group, then letting $(A_1^*,\dots,A_n^*) = \tau(A_1,\dots,A_n)$ we have*

$$A_1^* + \cdots + A_n^* \subseteq A_1 + \cdots + A_n.$$

Proof. The assertion follows by fixing, for each $g \in A_1^* + \cdots + A_n^*$, a representation $g = b_1 + \cdots + b_n$ with $b_i \in A_i^*$ for each $i \in [1,n]$, and then recursively choosing indices $i_1,\dots,i_n \in [1,n]$ so that, having $i_1,\dots,i_{k-1}$ found, the next index i_k is chosen to satisfy $i_k \notin \{i_1,\dots,i_{k-1}\}$ and $b_k \in A_{i_k}$. The details are straightforward. □

For a trio (A,B,C), let $(A^*,B^*,C^*) := \tau(A,B,C)$. As a corollary of Lemma 7, if $C^* \neq \varnothing$, then (A^*,B^*,C^*) is a trio, too.

We now consider the situation where one of the sets involved can be infinite. Lemmas 5 and 7 do not in fact assume finiteness, while Lemmas 4 and 6 extend onto the infinite case as follows.

Lemma 4′. *Suppose that $A_1,\dots,A_n$ are subsets of some ground set G such that $G \setminus A_1$, $A_2,\dots,A_n$ are all finite, and let $(A_1^*,\dots,A_n^*) := \tau(A_1,\dots,A_n)$. Then also $G \setminus A_1^*$, $A_2^*,\dots,A_n^*$ are all finite, and*

$$-|G \setminus A_1^*| + |A_2^*| + \cdots + |A_k^*| \geq -|G \setminus A_1| + |A_2| + \cdots + |A_k| \quad \text{for } k \in [1,n] \tag{4}$$

with equality for $k = n$. If, indeed, equality holds in (4) for each $k \in [1,n]$, then $A_1 \supseteq \cdots \supseteq A_n$ (whence $A_k^* = A_k$ for all $k \in [1,n]$).

Proof. The finiteness of $G \setminus A_1^*$, $A_2^*,\dots,A_n^*$ is immediate. The remaining assertions follow from Lemma 4 by considering the finite sets $U := A_2 \cup \cdots \cup A_n$ and $A_1' := A_1 \cap U$, and observing that

$$\tau(A_1', A_2, \dots, A_n) = (U, A_2^*, \dots, A_n^*),$$

that

$$-|G \setminus A_1^*| + |G \setminus A_1| = |A_1^* \setminus A_1| = |U \setminus A_1'| = |U| - |A_1'|,$$

and that $A_1' \supseteq A_2$ implies $A_1 \supseteq A_2$. □

Lemma 6′. *For any subsets $A_1, \dots, A_n$ of an abelian group G such that $G \setminus A_1$, $A_2, \dots, A_n$ are all finite, one of the following holds:*

i) *There exist elements $a_k \in A_k$ ($k \in [1,n]$) such that, letting $(A_1^*, \dots, A_n^*) := \tau(A_1 - a_1, \dots, A_n - a_n)$, we have*

$$(-|G \setminus A_1|, |A_2|, \dots, |A_n|) \prec (-|G \setminus A_1^*|, |A_2^*|, \dots, |A_n^*|).$$

ii) *We have $A_k - A_k \subseteq \pi(A_{k-1})$ for all $k \in [2,n]$; that is, for each $k \in [2,n]$, the set A_k is contained in a coset of the period of the set A_{k-1}.*

Proof. The proof is almost identical to that of Lemma 6, except that we now apply Lemma 4′ instead of Lemma 4.

By Lemma 4′, if $(-|G \setminus A_1|, \dots, |G \setminus A_n|) \prec (-|G \setminus A_1^*|, \dots, |A_n^*|)$ does *not* hold for some specific choice of $a_k \in A_k$, then for each $k \in [2,n]$, we have $A_k - a_k + a_{k-1} \subseteq A_{k-1}$. If now $a_k \in A_k$ and $a_{k-1} \in A_{k-1}$ can be chosen arbitrarily, this leads to $A_k - A_k + A_{k-1} \subseteq A_{k-1}$, meaning that $A_k - A_k \subseteq \pi(A_{k-1})$ as $A_k - A_k$ is symmetric. □

Since the order $\prec$ implies the lexicographic order $\prec_\ell$, from Lemma 6′ we conclude that either there exist elements $a_k \in A_k$ such that $(|G \setminus A_1^*|, -|A_2^*|, \dots, -|A_n^*|) \prec_\ell (|G \setminus A_1|, -|A_2|, \dots, -|A_n|)$, where $(A_1^*, \dots, A_n^*) = \tau(A_1 - a_1, \dots, A_n - a_n)$, or $A_k - A_k \subseteq \pi(A_{k-1})$ holds for each $k \in [2,n]$.

4. The Main Lemma and Overview of the Proof.

The following result is the central ingredient of the proof of Theorem 2.

Main Lemma. *Let (A,B,C) be an aperiodic, maximal, deficient trio in an abelian group G such that A is co-finite. For a triple $(a,b,c) \in G^3$, let $(A^*, B^*, C^*) := \tau(A - a, B - b, C - c)$ and suppose that there exists $(a,b,c) \in G^3$ with $C^* \neq \varnothing$ and $(A^*, B^*, C^*) \neq (A - a, B - b, C - c)$. With these assumptions, choose $(a,b,c) \in G^3$, satisfying the conditions just mentioned, for which $|G \setminus (A^* + B^* + C^*)|$ is smallest possible, and let $H := \pi(A^* + B^* + C^*)$. Then*

$$|(A^* + H) \setminus A^*| + |(B^* + H) \setminus B^*| + |(C^* + H) \setminus C^*| \geq |H| - 1.$$

We actually prove the Main Lemma and Theorem 2 simultaneously, using induction, as we now proceed to describe.

Clearly, it suffices to prove Theorem 2 only for those trios (A,B,C) with A co-finite. To every such trio, we associate the quadruple

$$\sigma(A,B,C) := (|G|, |G \setminus A|, -|B|, -|C|),$$

and we denote by $\mathfrak{S}$ the set of all quadruples that can arise this way, ordered lexicographically. The proof of the Main Lemma and Theorem 2 goes by induction on

$\sigma(A,B,C)$. The induction is well-founded as $\mathfrak{S}$ does not contain infinite descending chains (with respect to the lexicographic order). This follows by observing that, for $|G\setminus A|$ fixed, the set of possible values of $|B|$ and $|C|$ is finite, because $A+B+C\neq G$ implies $\max\{|B|,|C|\}\le |B+C|\le |G\setminus A|$.

As a part of our inductive argument, we now show that, loosely speaking, if the Main Lemma is true for the trio (A,B,C) with A co-finite, and Theorem 2 is true for all trios (A',B',C') with A' co-finite and $\sigma(A',B',C')<\sigma(A,B,C)$, then Theorem 2 is also true for the trio (A,B,C).

Proposition 1. *Let (A,B,C) be a trio with A co-finite and suppose that*

i) *either the assumptions of the Main Lemma fail for (A,B,C) or the assertion of the Main Lemma holds for (A,B,C);*
ii) *the estimate $\delta(A',B',C')\le |\pi(A'+B'+C')|$ holds for all trios (A',B',C') with A' co-finite and $\sigma(A',B',C')<\sigma(A,B,C)$.*

Then $\delta(A,B,C)\le |\pi(A+B+C)|$.

Proof. If (A,B,C) is not deficient, then the required estimate $\delta(A,B,C)\le |\pi(A+B+C)|$ is immediate; suppose therefore that (A,B,C) is deficient.

If (A,B,C) is not maximal, then we consider a maximal trio (A',B',C') containing (A,B,C) and satisfying $\pi(A'+B'+C')\le \pi(A+B+C)$, cf. Lemma 3. Since (A',B',C') strictly contains (A,B,C), it follows that $\sigma(A',B',C')<\sigma(A,B,C)$ and $\delta(A,B,C)<\delta(A',B',C')$, and in view of the assumption ii), we then get

$$\delta(A,B,C)<\delta(A',B',C')\le |\pi(A'+B'+C')|\le |\pi(A+B+C)|.$$

Suppose thus that (A,B,C) is maximal.

Let G denote the underlying group, and for a subgroup $K\le G$ denote by φ_K the canonical homomorphism from G onto the quotient group G/K.

If (A,B,C) is periodic, then, denoting its period by K, we get

$$\delta(A,B,C)\le \delta(A+K,B+K,C+K)=|K|\,\delta(\varphi_K(A),\varphi_K(B),\varphi_K(C))\le |K|,$$

with the last inequality following from the assumption ii) in view of $|G/K|\le |G|$ and

$$|\varphi_K(G)\setminus\varphi_K(A)|=|G\setminus(A+K)|/|K|\le |G\setminus A|/|K|<|G\setminus A|.$$

We therefore suppose that (A,B,C) is aperiodic.

If there do not exist $a,b,c\in G$ such that, letting $(A^*,B^*,C^*):=\tau(A-b,B-b,C-c)$, we have $C^*\neq\varnothing$ and $(A^*,B^*,C^*)\neq(A-a,B-b,C-c)$, then by the remark at the end of Section 3, the set C is contained in a coset of $\pi(B)$, and the set B is contained in a coset of $\pi(A)$. Since (A,B,C) is aperiodic, this yields $|B|=|C|=1$,

whence

$$\delta(A,B,C) = -|G \setminus A| + |B| + |C| = 2 - |G \setminus A| \le 1 \le |\pi(A+B+C)|.$$

We thus assume that the trio (A,B,C) satisfies all the assumptions of the Main Lemma: namely, it is maximal, aperiodic, and deficient, and there exists $(a,b,c) \in G^3$ such that, letting $(A^*,B^*,C^*) := \tau(A-a, B-b, C-c)$, we have $C^* \neq \varnothing$ and $(A^*,B^*,C^*) \neq (A-a,B-b,C-c)$. Moreover, we assume that the triple (a,b,c) is chosen to minimize $|G \setminus (A^*+B^*+C^*)|$. Notice that the condition $(A^*,B^*,C^*) \neq (A-a,B-b,C-c)$ implies $\sigma(A^*,B^*,C^*) < \sigma(A-a,B-b,C-c) = \sigma(A,B,C)$ by Lemma 4′.

Let $H := \pi(A^*+B^*+C^*)$. If $H = \{0\}$, then, by Lemma 4′ and assumption ii), we have

$$\delta(A,B,C) = \delta(A^*,B^*,C^*) \le |H| = 1 \le |\pi(A+B+C)|.$$

Suppose therefore that $H \neq \{0\}$. In view of $|G/H| \le |G|$ and

$$\begin{aligned} |\varphi_H(G) \setminus \varphi_H(A^*)| &\le |\varphi_H(G) \setminus \varphi_H(A)| \\ &= |G \setminus (A+H)|/|H| \le |G \setminus A|/|H| < |G \setminus A|, \end{aligned} \tag{5}$$

we can apply assumption ii) to the aperiodic trio $(\varphi_H(A^*), \varphi_H(B^*), \varphi_H(C^*))$ to obtain

$$\delta(A^*+H, B^*+H, C^*+H) = |H|\,\delta(\varphi_H(A^*), \varphi_H(B^*), \varphi_H(C^*)) \le |H|. \tag{6}$$

On the other hand, from the Main Lemma,

$$\delta(A^*+H, B^*+H, C^*+H) - \delta(A^*,B^*,C^*) \ge |H| - 1. \tag{7}$$

Comparing (6) and (7) and using Lemma 4′, we obtain

$$\delta(A,B,C) = \delta(A^*,B^*,C^*) \le 1 \le |\pi(A+B+C)|.$$

□

5. Proof of the Main Lemma and Theorem 2

5.1. **The set-up and initial observations.** As follows from Proposition 1, to establish the Main Lemma and Theorem 2, it suffices to prove the former assuming, as an induction hypothesis, that the latter is true for all "smaller" trios. Having the components of the trio under consideration appropriately translated, we thus have the following set of assumptions:

i) (A,B,C) is an aperiodic, maximal, deficient trio in an abelian group G, with A co-finite.

ii) the triple $(A^*, B^*, C^*) := \tau(A, B, C)$ satisfies $C^* \neq \varnothing$ and $(A^*, B^*, C^*) \neq (A, B, C)$; thus, (A^*, B^*, C^*) is a trio with $\delta(A^*, B^*, C^*) = \delta(A, B, C)$ and $\sigma(A^*, B^*, C^*) < \sigma(A, B, C)$ (by Lemmas 4′ and 6′).

iii) for any $a, b, c \in G$, letting $(U^*, V^*, W^*) := \tau(A-a, B-b, C-c)$, we have either $W^* = \varnothing$ or $(U^*, V^*, W^*) = (A - a, B - b, C - c)$ or $|G \setminus (U^* + V^* + W^*)| \geq |G \setminus (A^* + B^* + C^*)|$;

iv) for any trio (A', B', C') with A' co-finite and $\sigma(A', B', C') < \sigma(A, B, C)$, we have $\delta(A', B', C') \leq |\pi(A' + B' + C')|$.

We let $H := \pi(A^* + B^* + C^*)$, and we want to prove that

$$|(A^* + H) \setminus A^*| + |(B^* + H) \setminus B^*| + |(C^* + H) \setminus C^*| \geq |H| - 1. \tag{8}$$

Denote the left-hand side of (8) by ρ and, for a contradiction, assume that

$$\rho \leq |H| - 2; \tag{9}$$

notice that this implies $|H| \geq 2$.

For an element $x \in G$ and a set $S \subseteq G$, let $S_x := S \cap (x + H)$ be the *x-slice of S*; thus, if $x \equiv y \pmod H$, then $S_x = S_y$. From now on, we will write $\overline{S} := (S+H)\setminus S$ for the H-complement of S. (Although this is inconsistent with the notation of Section 2, no confusion should arise as the "old notation" will not be used anymore.) Thus, for instance, $\overline{S_x}$ is the complement of S in $x + H$, except that, if S does not have any elements in this coset, then $\overline{S_x}$ is empty:

$$\overline{S_x} = \begin{cases} (x + H) \setminus S & \text{if } S_x \neq \varnothing, \\ \varnothing & \text{if } S_x = \varnothing; \end{cases}$$

as a result, $|\overline{S_x}| = t|H| - |S_x|$, where $t = 1$ if $S_x \neq \varnothing$, and $t = 0$ otherwise.

Observing that $(A^*_x, B^*_x, C^*_x) = \tau(A_x, B_x, C_x)$, we get $|A^*_x| + |B^*_x| + |C^*_x| = |A_x| + |B_x| + |C_x|$ by Lemma 4. Consequently, we have

$$\rho = \sum_x \rho_x,$$

where x runs over the representatives of all cosets of H, and

$$\begin{aligned} \rho_x &= |(A^*_x + H) \setminus A^*| + |(B^*_x + H) \setminus B^*| + |(C^*_x + H) \setminus C^*| \\ &= |\overline{A^*_x}| + |\overline{B^*_x}| + |\overline{C^*_x}| \\ &= t_x|H| - |A^*_x| - |B^*_x| - |C^*_x| \\ &= t_x|H| - |A_x| - |B_x| - |C_x|, \end{aligned} \tag{10}$$

$t_x \in [0,3]$ being the number of non-empty slices among A_x^*, B_x^*, and C_x^*. In particular, if $A_x^* \neq \varnothing$ (meaning that least one A_x, B_x, and C_x is non-empty), then

$$\rho_x \geq |\overline{A_x^*}| \geq |H| - |A_x| - |B_x| - |C_x|, \tag{11}$$

and if C_x^* is non-empty (so that also A_x^* and B_x^* are non-empty), then

$$\rho_x \geq 3|H| - |A_x| - |B_x| - |C_x| = |\overline{A_x}| + |\overline{B_x}| + |\overline{C_x}|. \tag{12}$$

We say that a subset S of an H-coset is *partial* if $0 < |S| < |H|$ and is *full* if $|S| = |H|$.

Since (A, B, C) is maximal and aperiodic, there is a unique element of G not lying in $A+B+C$ (see a remark before the statement of Lemma 2); we denote this element by g_0, so that $A + B + C = G \setminus \{g_0\}$. Notice that, for every $x \in G$ with A_x partial, there exist $y, z \in G$ with both B_y and C_z partial and $x + y + z \equiv g_0 \pmod H$. Indeed, otherwise, for any y and z with $x + y + z \equiv g_0 \pmod H$, we would have either $B_y = \varnothing$ or $C_z = \varnothing$; this would lead to $g_0 \notin (x + H) + B + C$ and consequently $(A \cup (x + H), B, C)$ would also be a trio, contradicting the maximality of (A, B, C). Similar remarks apply to the situation where B_y or C_z is partial for some $y, z \in G$. This observation will be used repeatedly in the proof.

Recall that, for a subgroup $K \leq G$, by φ_K we denote the canonical homomorphism from G onto G/K.

Lemma 8.

i) *If $A_x^* = \varnothing$ for some $x \in G$, then there exist slices B_y^*, $C_z^* \neq \varnothing$ with $x+y+z \equiv g_0 \pmod H$.*
ii) *If $B_y^* = \varnothing$ for some $y \in G$, then there exist slices A_x^*, $C_z^* \neq \varnothing$ with $x+y+z \equiv g_0 \pmod H$.*
iii) *If $C_z^* = \varnothing$ for some $z \in G$, then there exist slices A_x^*, $B_y^* \neq \varnothing$ with $x+y+z \equiv g_0 \pmod H$.*

Proof. We prove the first assertion only; the other two follow in an identical way.

The key observation is that the aperiodic trio $(\varphi_H(A^*), \varphi_H(B^*), \varphi_H(C^*))$ is maximal: otherwise by Lemma 3 it would be properly contained in an aperiodic maximal trio (U, V, W) to which the induction hypothesis applies in view of $|\varphi_H(G) \setminus U| \leq |\varphi_H(G) \setminus \varphi_H(A^*)|$ and (5). This would lead to

$$\delta(\varphi_H(A^*), \varphi_H(B^*), \varphi_H(C^*)) < \delta(U, V, W) \leq |\pi(U + V + W)| = 1,$$

and thus to

$$\delta(A, B, C) = \delta(A^*, B^*, C^*) = H\delta(\varphi_H(A^*), \varphi_H(B^*), \varphi_H(C^*)) \leq 0,$$

contrary to the deficiency assumption. Now, the maximality of $(\varphi_H(A^*), \varphi_H(B^*), \varphi_H(C^*))$ shows that $(A^* \cup (x+H), B^*, C^*)$ is *not* a trio, which readily implies the assertion. □

Lemma 9. *Let (A_x, B_y, C_z) be a triple of slices with $x + y + z \equiv g_0 \pmod H$.*

i) *If $A_x \neq \varnothing$, then $|B_y| + |C_z| \leq |H|$;*
ii) *if $B_y \neq \varnothing$, then $|C_z| + |A_x| \leq |H|$;*
iii) *if $C_z \neq \varnothing$, then $|A_x| + |B_y| \leq |H|$.*

Proof. By the pigeonhole principle, from $|B_y| + |C_z| > |H|$ we would get $B_y + C_z = y+z+H$. If $A_x \neq \varnothing$, then this implies $g_0 \in x+y+z+H = A_x+B_y+C_z \subseteq A+B+C$, contrary to the choice of g_0. This proves i), and in the same way one obtains ii) and iii). □

5.2. **Recovering the structure.** We prove the Main Lemma in a series of claims sharing all the assumptions and notation of Section 5.1.

Claim A. *Let (U, V, W) be a permutation of (A, B, C) such that (U_x, V_y, W_z) has $C_x^* \neq \emptyset$ and V_y, $W_z \neq \emptyset$, where x, y, $z \in G$ satisfy $x + y + z \equiv g_0 \pmod H$. Then $V_z, W_y \neq \varnothing$ and $U_y = U_z = B_y^* = B_z^* = \varnothing$. In addition,*

$$|\overline{A_y^*}| + |\overline{A_z^*}| + 2|\overline{U_x}| \geq |H| \quad \text{and} \quad |\overline{A_y^*}| + |\overline{A_z^*}| + 4|\overline{U_x}| \geq 2|H|.$$

Proof. If we had $x \equiv y \pmod H$, then (12) would give

$$\rho \geq \rho_x \geq 3|H| - (|U_x| + |V_x| + |W_x|),$$

while $|U_x| + |V_x| = |U_x| + |V_y| \leq |H|$ by Lemma 9 (as $W_z \neq \varnothing$). It would then follow that $\rho \geq |H|$, contradicting (9).

Switching the roles of y and z in this argument, we similarly rule out the situation where $x \equiv z \pmod H$. Thus, we actually have $x \not\equiv y \pmod H$ and $x \not\equiv z \pmod H$.

If we had $U_z = V_z = \varnothing$, then from (12) and (10) we would obtain

$$\rho_x \geq 3|H| - |U_x| - |V_x| - |W_x| \geq |H| - |U_x|$$

and

$$\rho_z \geq |H| - |U_z| - |V_z| - |W_z| = |H| - |W_z|,$$

while $|U_x| + |W_z| \leq |H|$ by Lemma 9; consequently, $\rho \geq \rho_x + \rho_z \geq |H|$, contradicting (9). Thus, at least one of U_z and V_z is non-empty.

If *both* U_z and V_z were non-empty, then we would get a contradiction from

$$|U_x| + |V_y| \leq |H|, \quad |V_x| + |W_y| \leq |H|, \quad |W_x| + |U_y| \leq |H|$$

(by Lemma 9) and

$$\rho \geq \rho_x + \rho_y \geq (3|H| - |U_x| - |V_x| - |W_x|) + (|H| - |U_y| - |V_y| - |W_y|)$$

(by (12) and (10)). It follows that *exactly one* of U_z and V_z is empty. Switching the roles of y and z and of V and W, in the very same way we conclude that exactly one of U_y and W_y is empty.

We now claim that, indeed, V_z and W_y are non-empty, while U_y and U_z are empty, for if, say, we had $V_z = \varnothing$, then from

$$|U_x| + |W_z| \le |H|, \quad |U_z| + |W_x| \le |H|$$

(Lemma 9) we would get

$$\rho \ge \rho_x + \rho_z \ge (3|H| - |U_x| - |V_x| - |W_x|) + (|H| - |U_z| - |W_z|) \ge 2|H| - |V_x| \ge |H|,$$

and in a similar way we get a contradiction with (9) assuming that $W_y = \varnothing$.

We have thus shown that $V_z, W_y \ne \varnothing$ and $U_y = U_z = \varnothing$. Now, if we had $B^*_y \ne \varnothing$, then in view of $U_y = \varnothing$ this would result in

$$\rho \ge \rho_x + \rho_y \ge (3|H| - |U_x| - |V_x| - |W_x|) + (2|H| - |V_y| - |W_y|)$$

which, in conjunction with $|U_x| + |W_y| \le |H|$ and $|V_x| + |W_x| + |V_y| \le 3|H|$, contradicts (9). In the same way we obtain a contradiction assuming $B^*_z \ne \varnothing$. Thus, $B^*_y = B^*_z = \varnothing$.

Finally, since

$$|U_x| + |W_z| \le |H|, \quad |U_x| + |V_y| \le |H|, \quad |W_y| + |V_z| \le |H|$$

by Lemma 9, it follows in view of (11) that

$$\begin{aligned} &|\overline{A^*_y}| + |\overline{A^*_z}| + 2|\overline{U_x}| \\ &\qquad \ge (|H| - |V_y| - |W_y|) + (|H| - |V_z| - |W_z|) + 2(|H| - |U_x|) \ge |H| \end{aligned}$$

and similarly, from

$$|U_x| + |V_y| \le |H|, \quad |U_x| + |W_y| \le |H|, \quad |U_x| + |V_z| \le |H|, \quad |U_x| + |W_z| \le |H|,$$

we get

$$\begin{aligned} &|\overline{A^*_y}| + |\overline{A^*_z}| + 4|\overline{U_x}| \\ &\qquad \ge (|H| - |V_y| - |W_y|) + (|H| - |V_z| - |W_z|) + 4(|H| - |U_x|) \ge 2|H|. \end{aligned}$$

□

Claim B. *There is at most one coset $x + H$ such that C^*_x is partial. Moreover, if C^*_x is partial, then exactly one of A_x, B_x and C_x is partial while the other two are full.*

Proof. Assume by contradiction that C^*_x and C^*_ξ are both partial with $x \not\equiv \xi \pmod H$. Since C^*_x is partial, all three slices A_x, B_x and C_x are nonempty with at least one of them partial. Let (U, V, W) be a permutation of (A, B, C) such that U_x is partial.

Likewise, all three slices A_ξ, B_ξ and C_ξ are nonempty with at least one partial. Let (U', V', W') be a permutation of (A, B, C) such that U'_ξ is partial.

Recalling the observation above Lemma 8, let (U_x, V_y, W_z) and $(U'_\xi, V'_\eta, W'_\zeta)$ be triples of partial (in particular, nonempty) slices with

$$x + y + z \equiv \xi + \eta + \zeta \equiv g_0 \pmod{H}. \tag{13}$$

Without loss of generality, we assume that $|\overline{U'_\xi}| \geq |\overline{U_x}|$.

We have $B^*_x \supseteq C^*_x \neq \varnothing$ and $B^*_\xi \supseteq C^*_\xi \neq \varnothing$ while $B^*_y = B^*_z = B^*_\eta = B^*_\zeta = \varnothing$ by Claim A, and it follows that x and ξ are distinct modulo H from each of y, z, η, ζ. Consequently, by (12), (11), and the second inequality in Claim A,

$$\begin{aligned}\rho \geq \rho_x + \rho_\xi + \max\{\rho_y, \rho_z\} &\geq |\overline{U_x}| + |\overline{U'_\xi}| + \max\{|\overline{A^*_y}|, |\overline{A^*_z}|\} \\ &\geq 2|\overline{U_x}| + \frac{1}{2}\big(|\overline{A^*_y}| + |\overline{A^*_z}|\big) \geq |H|.\end{aligned}$$

This contradicts (9), showing that there is at most one coset $x + H$ such that C^*_x is partial.

To complete the proof, we now show that, if C^*_x is partial, then exactly one of A_x, B_x, and C_x is partial; since $C^*_x \neq \varnothing$ ensures that all three slices A_x, B_x and C_x are nonempty, this will also show that the other two slices are full. For a contradiction, suppose that (U, V, W) is a permutation of (A, B, C) such that U_x and V_x are both partial and find then y, z, η, ζ satisfying

$$x + y + z \equiv x + \eta + \zeta \equiv g_0 \pmod{H}$$

so that all the components of (U_x, V_y, W_z) and (V_x, W_η, U_ζ) are partial. As above, from Claim A we derive that x is distinct modulo H from each of y, z, η, ζ. Furthermore, by Claim A, the unique empty slice in (U_y, V_y, W_y) is U_y, the unique empty slice in (U_z, V_z, W_z) is U_z, the unique empty slice in (V_η, W_η, U_η) is V_η, and the unique empty slice in $(V_\zeta, W_\zeta, U_\zeta)$ is V_ζ; it follows that y is distinct modulo H from each of η and ζ, and similarly z is distinct modulo H from each of η and ζ. Also, from (10) and Claim A,

$$|\overline{U_x}| + \max\{\rho_y, \rho_z\} \geq |\overline{U_x}| + \frac{1}{2}\big(|\overline{A^*_y}| + |\overline{A^*_z}|\big) \geq \frac{1}{2}|H|$$

and

$$|\overline{V_x}| + \max\{\rho_\eta, \rho_\zeta\} \geq |\overline{V_x}| + \frac{1}{2}\big(|\overline{A^*_\eta}| + |\overline{A^*_\zeta}|\big) \geq \frac{1}{2}|H|.$$

Since $\rho_x \geq |\overline{U_x}| + |\overline{V_x}|$ by (12), we derive that

$$\rho \geq \rho_x + \max\{\rho_y, \rho_z\} + \max\{\rho_\eta, \rho_\zeta\} \geq |H|,$$

contradicting (9). □

Claim C. *If $U = A + a$, $V = B + b$, and $W = C + c$, with $a, b, c, \in H$, then letting $(U^*, V^*, W^*) := \tau(U, V, W)$, we have*

$$A^* \subseteq U^* + H, \quad B^* \subseteq V^* + H \quad \text{and} \quad C^* \subseteq W^* + H.$$

Also,

$$A^* + B^* + C^* \subseteq U^* + V^* + W^*.$$

Proof. The first assertion can be equivalently restated as follows: if, for a group element x, some of the slices A^*_x, B^*_x, and C^*_x are non-empty, then the corresponding slices from among U^*_x, V^*_x, and W^*_x are non-empty, too. Let t_x be the number of slices from among A^*_x, B^*_x and C^*_x that are non-empty. Then (9) and (10) give

$$|H| > \rho \geq t_x|H| - \big(|A_x| + |B_x| + |C_x|\big),$$

which further leads to $|U_x|+|V_x|+|W_x| = |A_x|+|B_x|+|C_x| > (t_x-1)|H|$; consequently, the pigeonhole principle ensures that at least t_x slices from among U^*_x, V^*_x and W^*_x are nonempty, and since $W^*_x \subseteq V^*_x \subseteq U^*_x$ and $C^*_x \subseteq B^*_x \subseteq A^*_x$ by definition of τ, the claimed result follows.

We proceed to prove the inclusion $A^* + B^* + C^* \subseteq U^* + V^* + W^*$. Assuming for a contradiction that it fails to hold, there exists a coset g_1+H contained in $A^*+B^*+C^*$ but not in $U^*+V^*+W^*$. Find group elements x, y, and z with $x+y+z \equiv g_1 \pmod H$ such that (A^*_x, B^*_y, C^*_z), and therefore also (U^*_x, V^*_y, W^*_z), has all its components non-empty. Since $U^*_x + V^*_y + W^*_z \subsetneq g_1 + H$ and $U^*_x \neq \varnothing$, the pigeonhole principle gives $|V^*_y| + |W^*_z| \leq |H|$, and hence

$$|\overline{V^*_y}| + |\overline{W^*_z}| \geq |H|. \tag{14}$$

If C^*_z were full, then all of A_z, B_z, C_z, and consequently W^*_z, would be full, contradicting $U^*_x + V^*_y + W^*_z \subsetneq g_1 + H$. Thus C^*_z is partial, and by Claim B, two of the slices A_z, B_z, and C_z are full. As a result, using (12) we obtain

$$\rho_z \geq |\overline{C^*_z}| = |\overline{W^*_z}|, \tag{15}$$

and we also conclude that B^*_z and V^*_z both are full. Consequently, if we had $W^*_y \neq \varnothing$, this would result in

$$g_1 + H = U^*_x + V^*_z + W^*_y \subseteq U^* + V^* + W^*,$$

a contradiction; thus, $W^*_y = \varnothing$, and comparing this to $W^*_z \neq \varnothing$, we obtain $y \not\equiv z \pmod H$. Since $B^*_y \neq \varnothing$ and $W^*_y = \varnothing$, from (10) we now get

$$\begin{aligned}\rho_y \geq 2|H| - (|A_y| + |B_y| + |C_y|) &= 2|H| - (|U_y| + |V_y| + |W_y|) \\ &= 2|H| - (|U^*_y| + |V^*_y| + |W^*_y|) \geq |H| - |V^*_y| \geq |\overline{V^*_y}|.\end{aligned}$$

In view of (15) and (14), this yields

$$\rho \geq \rho_y + \rho_z \geq |\overline{V_y^*}| + |\overline{W_z^*}| \geq |H|,$$

contradicting (9). □

For a set $S \subseteq G$, by $\langle S \rangle$ we denote the subgroup of G generated by S. Thus, $\langle S - S \rangle$ is the smallest subgroup $H \leq G$ such that S lies in an H-coset.

Claim D. *We have $H \leq \pi(C^*)$; that is, C^* is a union of H-cosets.*

Proof. If the assertion is wrong, then, by Claim B, there is a unique coset $z+H$ such that C_z^* is partial; moreover, of the three slices A_z, B_z, and C_z, one is partial while the other two are full. To begin with, we show that C_z partial, whereas B_z and C_z are full.

Aiming at a contradiction, assume that, for instance, B_z is partial, and therefore there exist $x, y \in G$ with $x + y + z \equiv g_0 \pmod{H}$ such that (A_x, B_z, C_y) has all its components non-empty. Observing that $A_y \neq \varnothing$ by Claim A, fix arbitrarily an element $a \in A_y - C_y \subseteq H$. Letting $(U^*, V^*, W^*) := \tau(A - a, B, C)$, we have then $U_x^* \neq \varnothing$ (as $A_x \neq \varnothing$), $V_y^* \neq \varnothing$ (as $(A_y - a) \cap C_y \neq \varnothing$), and $W_z^* \neq \varnothing$ (by Claim C). Hence,

$$(U^* + V^* + W^*) \cap (g_0 + H) \neq \varnothing, \tag{16}$$

whereas we know that

$$(A^* + B^* + C^*) \cap (g_0 + H) = \varnothing. \tag{17}$$

Since

$$A^* + B^* + C^* \subseteq U^* + V^* + W^* \tag{18}$$

by Claim C, this contradicts minimality of $|G \setminus (A^* + B^* + C^*)|$, unless $(U^*, V^*, W^*) = (A - a, B, C)$; that is, unless $C \subseteq B \subseteq A - a$. This, however, is inconsistent with the assumption that B_z is partial and C_z is full.

We have shown that B_z cannot be partial, and a similar argument shows that neither can A_z. Consequently, C_z is partial while both A_z and B_z are full, and we now re-use the argument above in these new settings.

Since C_z is partial, there exist $x, y \in G$ with $x + y + z \equiv g_0 \pmod{H}$ such that (A_x, B_y, C_z) has all its components non-empty. Let X be a set of representatives modulo H for all possible such choices of x and likewise let Y be a set of representatives modulo H for all such choices of y. By Claim A, for every pair $(x, y) \in X \times Y$ with $x + y + z \equiv g_0 \pmod{H}$, we have $B_x, A_y \neq \varnothing$; hence, X and Y coincide modulo H, and we can assume that, indeed, $X = Y$ holds.

Fix $(x,y) \in X \times Y$ with $x+y+z \equiv g_0 \pmod H$, and suppose that $b \in B_y - A_y \subseteq H$. Letting $(U^*, V^*, W^*) := \tau(A, B-b, C)$, we have then $U^*_x \neq \varnothing$ (as $A_x \neq \varnothing$), $V^*_y \neq \varnothing$ (as $(B_y - b) \cap A_y \neq \varnothing$), and $W^*_z \neq \varnothing$ (by Claim C). Hence, (16) holds true, and comparing it with (17) and (18), we get $(U^*, V^*, W^*) = (A, B-b, C)$, implying $C \subseteq B - b \subseteq A$, for otherwise the minimality of $|G \setminus (A^* + B^* + C^*)|$ would be contradicted. Likewise, for $a \in A_y - B_y \subseteq H$, letting $(U^*, V^*, W^*) := \tau(A-a, B, C)$, we have then $U^*_x \neq \varnothing$ (as $A_x \neq \varnothing$), $V^*_y \neq \varnothing$ (as $(A_y - a) \cap B_y \neq \varnothing$), and $W^*_z \neq \varnothing$ (by Claim C). Hence, (16) holds true, and comparing it with (17) and (18), we get $(U^*, V^*, W^*) = \tau(A-a, B, C)$, implying $C \subseteq B \subseteq A - a$. To summarize, for each $y \in Y$ and each $b \in B_y - A_y$ and $a \in A_y - B_y$, we have

$$C \subseteq B - b \subseteq A \quad \text{and} \quad C \subseteq B \subseteq A - a. \tag{19}$$

As a corollary, $B - B_y + A_y \subseteq A$, implying $A_y + B_\eta - B_y \subseteq A_\eta$ for all $y, \eta \in Y$. Switching the roles of y and η, we also get $A_\eta + B_y - B_\eta \subseteq A_y$, and as a result,

$$A_y + B_\eta - B_\eta + B_y - B_y \subseteq A_y.$$

Letting $K := \sum_{\eta \in Y} \langle B_\eta - B_\eta \rangle$, we conclude in view of $Y = X$ that $K \leq \pi(A_x)$ for each $x \in X$. From (19), we also see that

$$C \subseteq A \cap B.$$

Thus $C = A \cap B \cap C = C^*$. By Claim B, the set C has then exactly one partial slice; namely, C_z. It follows that all non-trivial triples (A_ξ, B_η, C_ζ) with $\xi + \eta + \zeta \equiv g_0 \pmod H$ have $\zeta \equiv z \pmod H$, and therefore have $\xi \in X$. Since the above-defined subgroup K lies below the period of each set A_ξ with $\xi \in X$, it must also lie below the period of $(A+B+C) \cap (g_0 + H) = (g_0 + H) \setminus \{g_0\}$. This, however, is only possible if $K = \{0\}$, forcing $|B_\eta| = 1$ for each $\eta \in Y$.

Let (A_ξ, B_η, C_z) be a nontrivial triple with $(\xi, \eta) \in X \times Y$ and $\xi + \eta + z \equiv g_0 \pmod H$. Since B_z is full while $|B_\xi| = 1$, we have $z \not\equiv \xi \pmod H$, whence

$$\rho \geq \rho_z + \rho_\xi \geq (|H| - |C_z|) + (|H| - |A_\xi| - |B_\xi| - |C_\xi|)$$

by (12) and (10). Since $|B_\xi| = 1$, and $C^*_z \neq \varnothing$ yields $C_\xi = \varnothing$ by Claim A, we conclude that

$$\rho \geq 2|H| - 1 - (|C_z| + |A_\xi|),$$

and to obtain a contradiction with (9) and complete the proof it remains to notice that $|C_z| + |A_\xi| \leq |H|$ by Lemma 9. □

Claim E. *If, for some $y \in G$, at least two among the slices A_y, B_y, and C_y are non-empty, then also B^*_y is non-empty.*

Proof. Suppose for a contradiction that $B_y^* = \varnothing$. By Lemma 8, there exist $x, z \in G$ with $x+y+z \equiv g_0 \pmod H$ such that A_x^* and C_z^* are non-empty. As a result, at least one of A_x, B_x, and C_x is non-empty (as $A_x^* \neq \varnothing$), at least two of A_y, B_y, and C_y are non-empty (by the assumption of the claim), and all three slices A_z, B_z, and C_z are non-empty (as follows from $C_z^* \neq \varnothing$). Consequently, there is a permutation (U, V, W) of the original trio (A, B, C) such that U_x, V_y, and W_z are all non-empty. Moreover, by Claim D, from $C_z^* \neq \varnothing$ it follows that A_z, B_z, and C_z are full. In particular, W_z is full, and so $g_0 + H = U_x + V_y + W_z \subseteq A + B + C$, a contradiction. □

Claim F. *Let Z be a set of representatives of all those cosets $z + H$ such that $A_z, B_z, C_z \neq \varnothing$ but $C_z^* = \varnothing$. Assuming that $Z \neq \varnothing$, let $K_B := \sum_{z\in Z}\langle B_z - B_z\rangle$ and $K_C := \sum_{z\in Z}\langle C_z - C_z\rangle$. Then for each $z \in Z$, we have $K_B \leq \pi(A_z)$ and $K_C \leq \pi(B_z) \cap \pi(A_z)$.*

Proof. Fix $z \in Z$. By Lemma 8, there exist $x, y \in G$ with $x + y + z \equiv g_0 \pmod H$ and $A_x^*, B_y^* \neq \varnothing$. Furthermore, for each $b \in B_z - A_z \subseteq H$, we have $(B_z - b) \cap A_z \neq \varnothing$, and we can find $c \in H$ so that, indeed, $A_z \cap (B_z - b) \cap (C_z - c) \neq \varnothing$. Letting $(U^*, V^*, W^*) := \tau(A, B - b, C - c)$, we thus have $W_z^* \neq \varnothing$ and, by Claim C, we have $U_x^*, V_y^* \neq \varnothing$. This shows that

$$(U^* + V^* + W^*) \cap (g_0 + H) \;\neq\; (A^* + B^* + C^*) \cap (g_0 + H) \;=\; \varnothing,$$

and since $A^* + B^* + C^* \subseteq U^* + V^* + W^*$ by Claim C, the minimality of the quantity $|G\setminus(A^*+B^*+C^*)|$ implies $(U^*, V^*, W^*) = (A, B-b, C-c)$; that is, $C-c \subseteq B-b \subseteq A$. Recalling that b was chosen to be an arbitrary element of $B_z - A_z$, we conclude that $B + A_z - B_z \subseteq A$, and in particular, $A_z + B_\zeta - B_z \subseteq A_\zeta$ for any $\zeta \in Z$. Switching the roles of z and ζ, we also get $A_\zeta + B_z - B_\zeta \subseteq A_z$, and combining these inclusions, we obtain $A_z + (B_z - B_z) + (B_\zeta - B_\zeta) \subseteq A_z$. As a result, $K_B \leq \pi(A_z)$, as required.

The second assertion follows in a similar way: for each $c \in C_z - B_z$, there exists $a \in H$ with $(C_z-c)\cap B_z \cap (A_z-a) \neq \varnothing$, and then the minimality of $|G\setminus(A^*+B^*+C^*)|$ gives $C - c \subseteq B \subseteq A - a$; this shows that $C_\zeta + B_z - C_z \subseteq B_\zeta$ for all $z, \zeta \in Z$, and combining this with $C_z + B_\zeta - C_\zeta \subseteq B_z$ yields $K_C \leq \pi(B_z)$.

The final portion of the claim also follows in a similar way. For each $c \in C_z - A_z$, there exists $b \in H$ with $(C_z - c) \cap (B_z - b) \cap A_z \neq \emptyset$, and then the minimality of $|G \setminus (A^* + B^* + C^*)|$ gives $C - c \subseteq B - b \subseteq A$. This shows that $C_\zeta + A_z - C_z \subseteq A_\zeta$ for all $z, \zeta \in Z$, and combining this with $C_z + A_\zeta - C_\zeta \subseteq A_z$ yields $K_C \subseteq \pi(A_z)$. □

Corollary 3. *If, for some $z \in G$, we have $A_z, B_z, C_z \neq \varnothing$ while $C_z^* = \varnothing$, then each of B_z and C_z is contained in a coset of $\pi(A_z)$.*

Proof. Let $K_B, K_C \le H$ be as in Claim F. Then $\langle B_z - B_z\rangle \le K_B \le \pi(A_z)$ shows that B_z is contained in a coset of $\pi(A_z)$, and then from $\langle C_z - C_z\rangle \le K_C \le \pi(A_z)$ we derive that C_z is contained in a coset of $\pi(A_z)$. □

Claim G. *If, for some $y \in G$, the slice B_y is partial, then A_y and C_y are both non-empty while C_y^* is empty.*

Proof. Since B_y is partial, there exist $x, z \in G$ such that $x + y + z \equiv g_0 \pmod H$ and A_x, C_z are both partial. We notice that none of the slices C_x^*, C_y^*, and C_z^* is full; hence by Claim D, all of them are actually empty.

Suppose that $w \in \{x, y, z\}$. If exactly one of the slices A_w, B_w, and C_w is non-empty, then denoting by U the corresponding set from among A, B, and C, we have

$$\rho_w = |\overline{A_w^*}| = |\overline{U_w}|.$$

If exactly two of A_w, B_w, and C_w are non-empty, then $B_w^* \ne \varnothing$ by Claim E, and denoting by U and V the sets from among A, B, and C corresponding to the non-empty slices, we have

$$\rho_w = |\overline{A_w^*}| + |\overline{B_w^*}| = |\overline{U_w}| + |\overline{V_w}|.$$

If A_w, B_w, and C_w are all non-empty with A_w partial, then $B_w^* \ne \varnothing$ by Claim E, and by Corollary 3, each of B_w and C_w lies in a coset of $\pi(A_w)$ (recall that $C_w^* = \emptyset$ as noted at the beginning of the proof); consequently,

$$\rho_w = 2|H| - |A_w| - |B_w| - |C_w| \ge 2|H| - 2|\pi(A_w)| - (|H| - |\pi(A_w)|) \ge \frac{1}{2}\,|H|. \quad (20)$$

With these preliminary observations, we can now prove that A_y and C_y are non-empty.

If $x \equiv y \equiv z \pmod H$, then the assertion is immediate as we have chosen x and z so that $A_x, C_z \ne \varnothing$. Assume now that x, y, and z all are different modulo H, and, for a contradiction, that there are at most two non-empty slices among A_y, B_y, C_y. As we have shown above, the latter assumption implies $\rho_y \ge |\overline{B_y}|$ (as B_y is nonempty by hypothesis). Also, if there were at most two non-empty slices among A_x, B_x and C_x, then we would have $\rho_x \ge |\overline{A_x}|$, whence Lemma 9 yields

$$\rho \ge \rho_x + \rho_y \ge |\overline{A_x}| + |\overline{B_y}| = 2|H| - (|A_x| + |B_y|) \ge |H|,$$

which contradicts (9). Thus, A_x, B_x, and C_x are all non-empty, and in a similar way, A_z, B_z, and C_z are all non-empty. Now $A_x + B_y + C_z \ne g_0 + H$ shows that A_x is partial, and $A_z + B_y + C_x \ne g_0 + H$ shows that A_z is partial. Hence, (20) yields

$$\rho \ge \rho_x + \rho_z \ge |H|,$$

contradicting (9).

We have thus shown that exactly two of x, y, and z coincide modulo H. If $x \equiv y \not\equiv z \pmod{H}$, then $A_y = A_x \neq \varnothing$ whence, assuming $C_y = \varnothing$, we would get

$$\rho \geq \rho_y = |\overline{A_y}| + |\overline{B_y}| = 2|H| - |A_x| - |B_y| \geq |H|$$

by Lemma 9. In a similar way we obtain a contradiction if $x \not\equiv y \equiv z \pmod{H}$, and it remains to consider the case where $x \equiv z \not\equiv y \pmod{H}$. If in this case $B_x = B_z = \varnothing$, then we obtain a contradiction from

$$\rho \geq \rho_x + \rho_y \geq |\overline{A_x}| + |\overline{B_y}| = 2|H| - |A_x| - |B_y| \geq |H|,$$

the last estimate following by Lemma 9.

Assume therefore that $B_x = B_z \neq \varnothing$. In this case A_x, B_x, and $C_x = C_z$ are all non-empty while $C_x^* = \varnothing$, whence

$$|B_x| \leq |\pi(A_x)| \tag{21}$$

by Corollary 3. On the other hand,

$$|A_x| \leq |H| - |\pi(A_x)| \tag{22}$$

since A_x is partial (see the beginning of the proof). Furthermore, $A_x, B_x, C_x \neq \varnothing$ yields $B_x^* \neq \varnothing$ by Claim E, implying

$$\rho_x \geq 2|H| - |A_x| - |B_x| - |C_x| \tag{23}$$

in view of (10), while

$$\rho_y \geq |\overline{B_y}| = |H| - |B_y| \tag{24}$$

as follows from an observation at the beginning of the proof. Finally,

$$|B_y| + |C_x| = |B_y| + |C_z| \leq |H| \tag{25}$$

by Lemma 9. Combining (21)–(25), we get

$$\begin{aligned}\rho \geq \rho_x + \rho_y &\geq (2|H| - |A_x| - |B_x| - |C_x|) + (|H| - |B_y|) \\ &= 3|H| - (|B_y| + |C_x|) - (|A_x| + |B_x|) \geq |H|,\end{aligned}$$

contradicting (9). This shows that A_y and C_y are both nonempty, and now $C_y^* = \emptyset$ follows from Claim D, else B_y would be full, contrary to hypothesis. □

5.3. **Conclusion of the proof.** We are ready to complete the proof of the Main Lemma.

Let $y \in G$ be an arbitrary element such that the slice B_y is partial. (Notice that such elements exist since otherwise B would be H-periodic, while we assume that (A, B, C) is an aperiodic trio.) By Claim G, we have $A_y, C_y \neq C_y^* = \varnothing$, and keeping the notation Z, K_B, and K_C of Claim F, we then conclude that $y \in Z$ and $K_C \leq \pi(B_y)$. Thus, every partial slice of B is K_C-periodic, and it follows that B itself

is K_C-periodic, implying $K_C = \{0\}$; that is, $|C_z| = 1$ for each $z \in Z$. In particular, $|C_y| = 1$.

If now A_z were not full for some $z \in Z$, then we would have $|A_z| \le |H| - |K_B|$ by Claim F, and since B_z^* is non-empty by Claim E, using (10) we would obtain

$$\rho \ge \rho_y \ge 2|H| - |A_z| - |B_z| - |C_z| \ge 2|H| - (|H| - |K_B|) - |K_B| - 1 = |H| - 1,$$

contradicting (9). Thus A_z is full for every $z \in Z$. In particular, A_y is full and

$$\rho_y \ge 2|H| - |A_y| - |B_y| - |C_y| = |H| - |B_y| - 1.$$

Since B_y is partial, we can find $x, z \in G$ with $x + y + z \equiv g_0 \pmod H$ and $A_x, C_z \ne \varnothing$. By Claim D, for each $w \in \{x, y, z\}$ we have $C_w^* = \varnothing$: for otherwise A_w, B_w, and C_w all would be full, leading to $A_x + B_y + C_z = g_0 + H$. Consequently, there is at least one empty slice among A_x, B_x, and C_x: else $x \in Z$ and (as we have just shown) A_x would then be full, whence $A_x + B_y + C_z = g_0 + H$. Since, in contrast, A_y, B_y and C_y are all non-empty, we have $x \not\equiv y \pmod H$. Furthermore, arguing as at the beginning of the proof of Claim G, we get

$$\rho_x \ge |\overline{A_x}|.$$

We have shown that that $x \not\equiv y \pmod H$, whence combining the above inequalities yields

$$\rho \ge \rho_x + \rho_y \ge |\overline{A_x}| + (|H| - |B_y| - 1) = 2|H| - 1 - (|B_y| + |A_x|) \ge |H| - 1$$

by Lemma 9, a proof concluding contradiction.

6. Concluding remarks.

In hindsight, the following stronger (and simpler) version of the Main Lemma follows easily from Theorem 2.

Lemma 10. *Suppose that (A, B, C) is an aperiodic, maximal, deficient trio, and let $(A^*, B^*, C^*) := \tau(A, B, C)$ and $H := \pi(A^* + B^* + C^*)$. If $C^* \ne \varnothing$, then*

$$|(A^* + H) \setminus A^*| + |(B^* + H) \setminus B^*| + |(C^* + H) \setminus C^*| \ge |H| - 1.$$

Lemma 10 is the "ideal-world main lemma". To derive it from Theorem 2, notice that by Lemma 4′ and Corollary 2,

$$\delta(A^* + H, B^* + H, C^* + H) \ge \delta(A^*, B^*, C^*) = \delta(A, B, C) = 1,$$

and that

$$\delta(A^* + H, B^* + H, C^* + H) = |H|\, \delta(\varphi_H(A^*), \varphi_H(B^*), \varphi_H(C^*)).$$

Thus, we have in fact

$$\delta(A^* + H, B^* + H, C^* + H) \geq |H|,$$

which implies

$$\delta(A^* + H, B^* + H, C^* + H) - \delta(A^*, B^*, C^*) \geq |H| - 1.$$

This is equivalent to the inequality of Lemma 10.

It is a major challenge to give Lemma 10 a simple, independent proof.

Interestingly, Theorem 2 is equivalent to the following statement:

$$\text{For any } \textit{maximal} \text{ trio } (A, B, C), \text{ we have } \delta(A, B, C) \leq |\pi(A + B + C)|. \tag{26}$$

To derive Theorem 2 from (26), given a trio (A, B, C), construct (A', B', C') as in Lemma 3. Since (A', B', C') is maximal, applying (26) to it we get

$$\delta(A, B, C) \leq \delta(A', B', C') \leq |\pi(A' + B' + C')| \leq |\pi(A + B + C)|.$$

An easy consequence of Theorem 2 is a characterization of deficient trios as those which can be obtained be removing few elements from a *maximal* deficient trio.

Claim 2. *A trio (A, B, C) is deficient if and only if there exists a* maximal *deficient trio (A', B', C') such that $A \subseteq A'$, $B \subseteq B'$, $C \subseteq C'$, and*

$$|A' \setminus A| + |B' \setminus B| + |C' \setminus C| < |\pi(A' + B' + C')|. \tag{27}$$

Proof. If (A', B', C') is maximal and (A, B, C) satisfies (27), then by Corollary 2,

$$\delta(A, B, C) = \delta(A', B', C') - (|A' \setminus A| + |B' \setminus B| + |C' \setminus C|) > 0.$$

Conversely, given a deficient trio (A, B, C), for the supertrio (A', B', C') of Lemma 3 we have

$$\begin{aligned} |A' \setminus A| + |B' \setminus B| + |C' \setminus C| &= \delta(A', B', C') - \delta(A, B, C) \\ &< \delta(A', B', C') = |\pi(A' + B' + C')| \end{aligned}$$

(the last equality uses Corollary 2 again). □

Finally, we note that Lemma 7 can be extended to take into account the number of representations of group elements.

Lemma 11. *For any finite subsets $A_1, \dots, A_n$ of an abelian group, letting $(A_1^*, \dots, A_n^*) := \tau(A_1, \dots, A_n)$, the number of representations of any group element as $a_1^* + \dots + a_n^*$ (with $a_i^* \in A_i^*$ for each $i \in [1, n]$) does not exceed the number of its representations as $a_1 + \dots + a_n$ (with $a_i \in A_i$ for each $i \in [1, n]$).*

We omit the proof.

References

[BDM15] T. Boothby, M. DeVos, and A. Montejano, A new proof of Kemperman's theorem, *INTEGERS* **15** (2015), #A15.

[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.

E-mail address: diambri@hotmail.com

Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA

E-mail address: seva@math.haifa.ac.il

Department of Mathematics, The University of Haifa at Oranim, Tivon 36006, Israel