Sumsets, Zero-Sums and Extremal Combinatorics

Thesis by

David J. Grynkiewicz

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2006

(Defended August 10, 2005)

Acknowledgements

I first thank my advisor, Rick Wilson, for his continual support throughout my stay at Caltech. I would also like to thank all of my coauthors, Arie Bialostocki, Paul Dierker, Mark Lotspeich, Rasheed Sabar and Andy Shultz, for the efforts they contributed, and I also thank Arie Bialostocki for his problem suggestions and counsel. I express my gratitude to the NSF and the University of Idaho, for, respectively, funding and hosting the Research Experiences for Undergraduates (REU) that precipitated much of my interest in this area; to Caltech, where I have spent the past four years developing the results now contained in this thesis; and to Bates College, where my study in mathematics principally began. Lastly, I thank all my friends and family for keeping me sane throughout the whole process.

Abstract

This thesis develops and applies a method of tackling zero-sum additive questions—especially those related to the Erdős-Ginzburg-Ziv Theorem (EGZ)—through the use of partitioning sequences into sets, i.e., set partitions. Much of the research can alternatively be found in the literature spread across nine separate articles, but is here collected into one cohesive work augmented by additional exposition. Highlights include a new combinatorial proof of Kneser's Theorem (not currently located elsewhere); a proof of Caro's conjectured weighted Erdős-Ginzburg-Ziv Theorem; a partition analog of the Cauchy-Davenport Theorem that encompasses several results of Mann, Olson, Bollobás and Leader, and Hamidoune; a refinement of EGZ showing that an essentially dichromatic sequence of $2m-1$ terms from an abelian group of order m contains a mostly monochromatic m-term zero-sum subsequence; an interpretation of Kemperman's Structure Theorem (KST) for critical pairs (i.e., those finite subsets A and B of an abelian group with $|A+B| < |A|+|B|$) through quasi-periodic decompositions, which establishes certain canonical aspects of KST and facilitates its use in practice; a draining theorem for set partitions showing that a set partition of large cardinality sumset can have several elements removed from its terms and still have the sumset remain of large cardinality; a proof of a subsequence sum conjecture of Hamidoune; the determination of the $g(m, k)$ function introduced by Bialostocki and Lotspeich (defined as the least n so that a sequence of terms from $\mathbb{Z}/m\mathbb{Z}$ of length n with at least k distinct terms

must contain an m-term zero-sum subsequence) for m large with respect to k; the determination of $g(m, 5)$ for $m \geq 5$, including the details to the abbreviated proof found in the literature; various zero-sum results concerning modifications to the nondecreasing diameter problem of Bialostocki, Erdős, and Lefmann; an extension of EGZ to a class of hypergraphs; and a lower bound on the number of zero-sum m-term subsequences in a sequence of n terms from an abelian group of order m that establishes Bialostocki's conjectured value for small $n \leq 6\frac{1}{3}m$.

Contents

List of Definitions and Notation

$[a, b]$: The integers between a and b inclusive, where a and b are both integers with $a \leq b$. If $b = a - 1$, then by convention $[a, b] = [a, a - 1]$ is the empty set. (Part II (Interlude))

$[G : H]$: The index of the subgroup H in the group G.

ϕ_a: The natural homomorphism $\phi_a : G \to G/H_a$. See also H_a. (Sec 1.1)

$\eta_b(A, B)$: The number of $c \in A + b$ such that $\nu_c(A, B) = 1$. (Sec 5.2)

$\nu_g(A, B)$: The number of ways g can be represented as a sum $a + b = g$ with $a \in A$ and $b \in B$. (Sec 1.2)

$\langle A \rangle$: The subgroup generated by the subset or element A.

$A - B$: $A + (-B)$. (Sec 1.1)

$\overline{A}$: The complement of the set A. (Sec 1.1)

$\overline{a}$: The least positive integer representative of the element $a \in \mathbb{Z}/m\mathbb{Z}$. (Sec 4.2)

$-A$: The set of inverses to the elements of A, where A is a subset of an abelian group, i.e., $\{-a \mid a \in A\}$. (Sec 1.1).

$A + B$: The sumset of A and B, namely $\{a + b \mid a \in A, b \in B\}$, where A and B are subsets of an abelian group, see also sumset. (Sec 1.1)

$A \setminus b$: $A \setminus \{b\}$.

(c.x): A labelled comment from Chapter 5.

$f(L^t_m, r)$: See Part II (Interlude).

$f_{zs}(L_m^t, r)$: See Part II (Interlude).

first$_k(Z)$: The set $\{z_1, \ldots, z_{\min\{k,\,m\}}\}$, where Z is a set of integers $z_1 < z_2 < \ldots < z_m$. (Sec 7.2)

$H < K$: H is a proper subgroup of K.

$H(A)$: The stabilizer of the subset A of an abelian group G, namely the subgroup given by $H(A) = \{g \in G \mid g + A = A\}$. (Sec 1.1)

H**-hole**: An H-hole of a subset A of an abelian group is an element $h \in (H + A) \setminus A$, where H is a subgroup. (Sec 1.1)

H_a: Often used to denote a subgroup of an abelian group, whose index, if finite, is associated to a. (Sec 1.1)

H_a**-doubled**: An element $y \in \phi_a(A_j)$ is H_a-doubled if $|\phi_a^{-1}(y) \cap A_j| \geq 2$. (Sec 3.1)

H_a**-exception**: An element $y \in G/H_a$ that is not an H_a-nonexception. See H_a-nonexception. (Sec 3.1)

H_a**-nonexception**: An element $y \in G/H_a$, such that $y \in \phi_a(A_i)$ for all i, where $A_1, \ldots, A_n$ is an n-set partition whose sumset is H_a-periodic. (Sec 3.1)

H_a**-periodic part**: See quasi-periodic decomposition. (Sec 5.2)

H**-periodic**: A subset A of an abelian group is H-periodic (with period H), where H is a subgroup of G, if A is a union of H-cosets. (Sec 1.1)

int$_i(Z)$: The element z_i, where Z is a set of integers $z_1 < z_2 < \ldots < z_m$ and $i \leq m$. (Sec 7.2)

last$_k(Z)$: The set $\{z_{\max\{1,\,m-k+1\}}, \ldots, z_m\}$, where Z is a set of integers $z_1 < z_2 < \ldots < z_m$. (Sec 7.2)

m**-set**: A set of cardinality m. (Sec 7.2)

m**-uniform**: An m-uniform hypergraph is a hypergraph in which every edge has cardinality m. (Sec. 11.1)

m-zsf: A sequence of terms from $\mathbb{Z}/m\mathbb{Z}$ is m-zsf if it contains no m-term zero-sum subsequence. (Sec 9.1)

n-set partition: An *n-set partition* of a sequence S is a partition of the sequence S into n nonempty subsequences, $A_1, \ldots, A_n$, such that the terms in each subsequence A_i are all distinct, and thus the A_i can be regarded as sets. (Sec 2.2)

$\binom{n}{m}$: The number $\frac{n(n-1)\cdots(n-m+1)}{m(m-1)\cdots 1}$, where $m \in \mathbb{Z}$.

$n \wedge S$: For a sequence S, $n \wedge S$ is the set of elements that can be represented as a sum of terms from some n-term subsequence of S. (Sec 8.1)

r-coloring: An r-coloring of a set X is a function $\Delta : X \to C$, where C is a cardinality r set considered as the set of colors. (Part II (Interlude))

$|S|$: (1) The length of the sequence S. (2) The cardinality of the set S. (Sec 2.2)

$S \setminus S'$: (1) If S and S' are sets, then $S \setminus S' = \{s \in S \mid s \notin S'\}$. (Sec 1.1) (2) if S is a sequence and S' is a subsequence of S, then $S \setminus S'$ is the sequence obtained by removing all terms from S that are in S'. (Sec 2.2)

$S \cup S'$: (1) If S and S' are sets, then $S \cup S'$ is their union. (2) If S and S' are sequences, then $S \cup S'$ denotes the concatenation of S' with S. (Sec 10.2)

$X \cap S$: For a sequence S and set X, $X \cap S$ denotes the subsequence of S consisting of terms from X. (Sec 4.1)

$X \prec Y$: The notation $X \prec Y$ indicates that $\max X < \min Y$, where X and Y are subsets of the integers. (Sec 7.2)

$x + A$: $\{x\} + A$, where A is a subset of an abelian group G, and $x \in G$. See also sumset. (Sec 1.1)

(w, m): The greatest common divisor of w and m.

wA: $\{wa \mid a \in A\}$, where A is a subset of an abelian group and w is an integer. Note that $2A$ is NOT in general equal to $A + A$. (Sec 2.3)

$\mathbb{Z}^+$: The positive integers.

Affine transformation: An affine transformation, as used in this thesis, is a map from $\mathbb{Z}/m\mathbb{Z}$ to $\mathbb{Z}/m\mathbb{Z}$ of the form $x \mapsto ax + b$, with a, $b \in \mathbb{Z}/m\mathbb{Z}$ and $(a, m) = 1$. (Sec 9.1)

Aperiodic: A subset A is aperiodic if it is not periodic. Equivalently A is maximally H-periodic with H the trivial subgroup. (Sec 1.1)

Aperiodic part: See quasi-periodic decomposition. (Sec 5.2)

Arithmetic progression: An arithmetic progression in an abelian group G is a set of the form $\{\alpha + id \mid i = 1, 2, \ldots, l\}$ with $\alpha \in G$, with difference $d \in G$, and with length l a positive integer. (Sec 1.1)

Cauchy: A subset B of an abelian group G is Cauchy if B is finite and nonempty, and $|A + B| \geq \min\{|G|, |A| + |B| - 1\}$ for every finite, nonempty subset $A \subseteq G$. (Sec 5.1)

Coloring: A coloring of a set or sequence X by a set of colors C is a mapping $\Delta : X \to C$ that assigns a color from C to each element or term of X. See also r-coloring. (Part II (Interlude))

Complete m-uniform hypergraph: The m-uniform hypergraph on a vertex set V that has every possible cardinality m subset of V as an edge. (Sec 11.1)

Critical pair: A pair of finite subsets A and B of an abelian group such that $|A+B| \leq |A| + |B| - 1$. (Sec 5.1)

Exponent: The exponent of an abelian group G is the minimal integer k such $kg = 0$ for all $g \in G$. (Sec 2.3)

Hole: See H-hole. (Sec 1.1)

Hypergraph: A set system, namely, a collection of subsets, referred to as edges, of a set V, whose elements are referred to as the vertices of the hypergraph. If all edges have cardinality two, then the hypergraph is just a graph. (Sec 11.1)

Maximally H-periodic: A subset A of an abelian group G is maximally H-periodic if H is the maximal subgroup for which A is H-periodic. Equivalently $H = H(A)$ is the stabilizer of A. (Sec 1.1)

Monochromatic: (1) All having the same color (sometimes used to mean that all terms of a sequence are equal, in which case, each term is implicitly assumed to color itself). (2) Every edge is monochromatic as above, in reference to a uniform hypergraph having a vertex coloring. See also Part II (Interlude).

Monovalent: A vertex contained in precisely one edge. See hypergraph. (Sec 11.1)

Natural numbers: The non-negative integers.

Periodic: A subset is periodic if it is H-periodic for some nontrivial subgroup. Equivalently, $H(A)$ is nontrivial. (Sec 1.1)

Punctured periodic set: A punctured periodic set is a set A for which there exists $\alpha \in G \setminus A$ such that $A \cup \{\alpha\}$ is periodic. (Sec 5.2)

Quasi-periodic: A set A of an abelian group is quasi-periodic if A has a quasi-periodic decomposition $A = A_1 \cup A_0$ with A_1 nonempty. (Sec 5.2)

Quasi-periodic decomposition: For a subset A of an abelian group, and a nontrivial subgroup H_a, a quasi-periodic decomposition of A with quasi-period H_a is a partition $A = A_1 \cup A_0$ of A into two disjoint (each possibly empty) subsets such that A_1 is H_a-periodic or empty and A_0 is a subset of an H_a-coset. The set A_0 is the aperiodic part of the decomposition, and the set A_1 is the H_a-periodic part of the decomposition. (Sec 5.2)

Quasi-period: See quasi-periodic decomposition. (Sec 5.2)

Rearranged subsequence: A rearranged subsequence S' of a sequence S is sequence that under some permutation of its terms is a subsequence of S. (Sec 2.1)

Reduced quasi-periodic decomposition: A quasi-periodic decomposition $A_1 \cup A_0$ is reduced if A_0 is not quasi-periodic. (Sec 5.2)

Set partition: See n-set partition. (Sec 2.2)

Singleton Set: A set of cardinality one.

Sumset: The set of all pairwise sums of elements from two sets of an abelian group, see also $A + B$. (Sec 1.1)

Stabilizer: See $H(A)$. (Sec 1.1)

w.l.o.g.: Without loss of generality.

Zero-sum: (1) Either having the sum or the sum of colors being zero (the constituent terms, or colors, coming from an abelian group). (2) Every edge is zero-sum as above, in reference to a uniform hypergraph having a vertex coloring with colors from an abelian group. See also Part II (Interlude).

Zero-sum generalization: See Part II (Interlude).

Preface and Introduction

This thesis contains most of the work in zero-sum additive theory done during my stay at Caltech, much of which appears in print in various journals (see the end of this Preface and Introduction for specifics), but which is here collected into one body of work.

Zero-sum additive theory is an area of mathematics whose oldest roots trace back to Cauchy, but which has only recently begun experiencing rapid growth and development. Given two subsets A and B of an abelian group G, their sumset $A+B$ is the set of all pairwise sums, i.e., $A+B=\{a+b \mid a\in A,\, b\in B\}$. Given a sequence S of terms from the abelian group G, a subsequence S' of S is called zero-sum if the sum of terms from S' is zero. There are many natural questions that arise concerning the structures just described. If the cardinality of $A+B$ (or a particular restricted subset of $A+B$) is small in comparison to $|A|$ and $|B|$, then what can we say about the structure of A, B and $A+B$ (or that particular restricted subset of $A+B$)? If S contains no zero-sum subsequence (of possibly fixed length), or if the number of elements of G that can be represented as a sum of some subsequence of S (of possibly fixed length) is small, then what can we say about the structure of S?

One cornerstone result from the area concerning the second question is due to Erdős, Ginzburg and Ziv, and is quite simple in its statement: if a sequence S of terms from an abelian group of order m contains no m-term zero-sum, then the length of S must be at most $2m-2$. With a little thought, one soon realizes that this Erdős-Ginzburg-Ziv Theorem (EGZ for short) is an algebraic generalization of the pigeonhole principle (just consider sequences that consist only of 0's and 1's from a cyclic group). It having sometimes been said that Ramsey Theory (and maybe some related areas of extremal combinatorics) is an extension of the pigeonhole principle, then it took little time before the question was raised of whether other questions from generalized Ramsey Theory might also zero-sum generalize in a manner analogous to that of the pigeonhole principle by EGZ (namely if one replaced a 2-coloring by a coloring with the elements from a cyclic group, would the size of the structure needed to guarantee a sought-after substructure, having the sum of its colors equaling zero, be the same as the size needed in the 2-coloring case to guarantee a sought-after monochromatic substructure?).

Lacking very many tools for dealing with such problems, progress progressed with some difficulty, though things went smoother when m was prime (since this assumption made the cyclic group also a field). One method (the one in fact used originally to handle the prime cases in the proof of EGZ) involved making use of answers to the first question mentioned about zero-sum additive theory to help find answers to the second question. For instance, if $A_1,\ldots,A_n$ were nonempty subsets of $\mathbb{Z}/m\mathbb{Z}$ having a large cardinality sumset, say $|\sum_{i=1}^{n} A_i|\geq m$, then it would be possible to select an element a_i from each A_i so that the resulting sum of the a_i was whatever element of $\mathbb{Z}/m\mathbb{Z}$ one might desire, including zero. Hence if all the elements of the A_i were also terms in a sequence S, then theorems used to derive information about the structure of the A_i when $|\sum_{i=1}^{n} A_i|$ was small could be used to derive information about the structure of the sequence S when S contained no n-term zero-sum. Of course there might be many different ways to so associate n sets $A_1,\ldots,A_n$ to the sequence S, and if any one of these had a large enough cardinality then we could deduce the existence of an n-term zero-sum. One might naturally think that forcing all of these set partitions $A_1, \ldots, A_n$ to have small cardinality sumset would impose yet stronger structure restrictions on the sequence S than any gained from knowing that just a particular one of these set partitions $A_1, \ldots, A_n$ had a small cardinality sumset. Several years later, and after much growth and development, this simple idea yields its fruit in this thesis.

I have attempted in part to make this thesis into a small primer on the implementation of the set partition methods and techniques (developed during my course of study at Caltech, and building upon the research originating from my time at Bates College) for solving various zero-sum-related questions. The thesis is for the most part self contained, though several lengthy proofs of results from additive number theory have been omitted with citations given instead in such instances. Otherwise, only a basic understanding of the fundamentals of mathematics, combinatorics, group theory and calculus are prerequisite. I have divided the thesis into two main parts—the first developing tools and machinery, the second containing zero-sum applications of that developed machinery. Chapter 5 also contains a few additional non-zero-sum applications separate from the remainder of the thesis. Notation and definitions are introduced as needed, and for ease of reference also appear in the preceding section accompanied by the section in which they were introduced. Brief specifics of each chapter's topics are given below.

Chapter 1 is devoted to the fundamental result of Kneser, concerning the structure of sets with very small sumset, and several of its immediate consequences (such as the Cauchy-Davenport Theorem). A new proof of Kneser's Theorem (more geometric in nature) has been included, and the results of this chapter will be heavily used in subsequent chapters.

Chapter 2 begins by introducing the notion of set partitions and their most basic properties. It then continues with a weighted version of the Erdős-Ginzburg-Ziv Theorem that establishes a conjecture of Y. Caro [9], as well as many cases in a related conjecture of Y. O. Hamidoune [33]. The proof employs the same techniques to be used for the main result of Chapter 3, but in a relatively simpler form. As such, the material from Chapter 2 provides a more gradual introduction and cushioning to the apparent technical complexity of Chapter 3. The material for this chapter now appears in [28], with small portions from [7].

Chapter 3 contains the foundational result for using set partitions to solve zero-sum problems. The main result generalizes and unifies existing results of Mann [48], Olson [50], Bollobás and Leader [8], and Hamidoune [38]. The chapter is a hybrid of work done both at Bates College and Caltech, and can be pieced together from various articles by the author in the literature [24] [26] [28] [30].

Chapter 4 contains a refinement of the Erdős-Ginzburg-Ziv Theorem that shows that a sequence of $2m-1$ terms from an abelian group of order m consisting 'mostly' of two distinct terms (with only a few exceptions) contains a zero-sum m-term subsequence consisting 'mostly' of one distinct term. While the statement of the theorem may at first seem slightly esoteric, such conditions arise naturally for a sequence that does not contain a sufficiently compressed set partition. A common theme that emerges when using set partitions for solving zero-sum problems is that one often finds the majority of cases falling rather easily to a straightforward implementation of the results from Chapter 3, nonetheless leaving behind several very stubborn and difficult special cases consisting of highly structured sequences, which often must be handled by more ingenious tactics (Chapter 12 gives an extreme example of this issue). The existence of a sufficiently compressed set partition is the input needed to set in motion the machinery of Chapter 3, and thus results like those of this chapter can sometimes help provide the necessary priming for the more heavy machinery.

The results of Chapter 4 are put to use later in Chapter 11 and can be found in [27].

Chapter 5 introduces the notion of quasi-periodic decompositions, and then uses this notion to give a simpler interpretation of Kemperman's involved and recursive description of the structure of a pair of finite subsets A and B of an abelian group satisfying $|A+B| < |A|+|B|$ (i.e., those pairs of subsets for which the triangle inequality fails). The material will be used at the end of the chapter to extend and simplify several non-zero-sum results that could previously only be handled by the more involved isoperimetric method introduced by Hamidoune [37] [35]. The material will also be used in Chapter 6, and all results from the chapter appear in [25].

Chapter 6 contains, with great output of technical effort, a draining result for set partitions that can be used in certain cases to boost the effectiveness of the main result from Chapter 3. In essence, the result of Chapter 6 allows several terms of the sequence to be used twice, thus doubling the effectiveness of what can be done with this certain small portion of the sequence. Like the result of Chapter 4, the added benefits obtained from Chapter 6 are of greatest use for more involved and intricate zero-sum questions. The results of Chapter 6 will be needed for Chapters 11 and 12, and occur in [27], with small portions from [25].

Chapter 7 marks a break in the thesis between the first half, which deals with the development of necessary machinery, and the second half, which deals with the application of this machinery to zero-sum questions. This chapter contains a simple application to a zero-sum generalizing problem from generalized Ramsey Theory. The majority of work done in this chapter is due to my co-author A. Schultz, with the major contribution of myself being the interface between the work Schultz provided and the general method developed in prior chapters of this thesis. Like the relevant portion of Chapter 9 to follow, Chapter 7 requires little additional set partition machinery other than the simplest version of the core result of Chapter 3. The research from this chapter is currently contained in the submitted preprint [29].

Chapter 8 derives the structure of sequences, of terms from an abelian group of order m, that represent very few elements as a sum of some m-term subsequence, affirming a conjecture of Hamidoune [38]. The result is similar in flavor to Kneser's Theorem (from Chapter 1) as well as a result of Olson [50]. The material from this chapter now appears in [26] and will be used in Chapter 9 as well.

Chapter 9 deals with looking for zero-sums in sequences with a fixed number of distinct terms. Another theme that emerges when studying zero-sum questions is the curious observation that a larger number of distinct terms usually makes finding a zero-sum subsequence easier instead of harder. To capture this idea, Bialostocki and Lotspeich introduced the function $g(m, k)$ defined as the minimal length of a sequence of terms, from a cyclic group of order m containing at least k distinct terms, that guarantees an m-term zero-sum subsequence. This chapter derives the value of $g(m, k)$ for $k \leq 5$ (the cases $k \leq 4$ were known; the case $k = 5$ was previously still open), as well as the exact value of $g(m, k)$ for sufficiently large m with respect to k. The material appears in [7], which was coauthored with A. Bialostocki, P. Dierker and M. Lotspeich. The initial portion of the chapter contains several results of Gao [20] [18], as well as the statement of the now affirmed Erdős-Heilbronn conjecture [12] [2], whose proofs are omitted but which will be used in the remainder of the chapter. The results related to $g(m, k)$ for $k \leq 4$ will also be used in Chapter 10.

Chapter 10 contains the (chronologically) first application of Chapter 3 that began to tap below the surface of Chapter 3, and was the application that inspired one of the major improvements now incorporated into Chapter 3. The chapter begins by providing upper and lower bounds for a question from generalized Ramsey-Theory (a modification of a problem considered by Bialostocki, Erdős and Lefmann [6]), and then obtains further partial and exact results for the corresponding zero-sum version. The material can be found in the submitted preprint [30], jointly co-authored with R. Sabar.

Chapter 11 extends the Erdős-Ginzburg-Ziv Theorem to a class of hypergraphs. Many of the initial zero-sum generalizing problems, such as those of Chapters 7 and 10, deal with looking for several simultaneous disjoint zero-sums. It is natural to wonder about similar problems when the intersection structures are non-disjoint. However, these problems turn out to be more difficult. There is not always a nice zero-sum generalization, similar to that of EGZ and the pigeon-hole principle, for any given intersection structure. The results of this chapter, however, show that if the intersection structure is very weak, then such a nice zero-sum generalization still occurs. The results can be found in [27].

Chapter 12 concludes the thesis with an application to the multiplicity of m-term zero-sums in a sequence of terms from an abelian group of order m. A lower bound on how many such subsequences must exist (as a function of the length of the sequence and m) is given that affirms the bound conjectured by Bialostocki [3] [4] for sequences of small length. The material currently appears in the submitted preprint [31].

Part I

METHODS AND TOOLS

Chapter 1

Kneser's Theorem

1.1 Discussion

In this chapter, we give a new proof for one of the most foundational results of inverse additive number—Kneser's Theorem—which will be used extensively in this thesis. This chapter also serves as an introduction to set addition and related topics.

Given two subsets A and B of an abelian group G, their *sumset* is the set of all pairwise sums, denoted $A + B = \{a + b \mid a \in A,\, b \in B\}$. From the basic properties of addition, we have that $A + B = B + A$, and that the sumset of more than two sets, denoted $\sum\limits_{i=1}^{n} A_i = \{\sum\limits_{i=1}^{n} a_i \mid a_i \in A_i\}$, is well defined. We often use the convention that $\sum\limits_{i\in\emptyset} A_i = \{0\}$. For sumsets with a single element set, we abbreviate $\{x\} + A$ by $x + A$. Substraction of sets is handled similarly; for instance, $-A = \{-a \mid a \in A\}$ and $A - B = A + (-B)$.

Arithmetic progressions in an arbitrary abelian group G (with length l and difference d) are sets of the form $\{\alpha + id \mid i = 1, 2, \ldots, l\}$, with $\alpha,\, d \in G$ and $l \in \mathbb{Z}^+$, and are closely related to the prototypical cases that arise when studying sumsets with small cardinality. The complement of a subset A of an abelian group G, denoted $\overline{A}$, also proves useful when G is finite.

Next we introduce some notation and definitions related to the structural description of the sumset $A+B$ provided by Kneser's Theorem. Given a subgroup H of G, and a subset A of G, the set A is said to be *H-periodic* (with period H) if A is a union of H-cosets. Note that any subset A is H-periodic with H the trivial subgroup. Since subgroups are closed under addition (i.e., $H+H=H$), it follows from the definition that if A is H-periodic, then $H+A=A$. On the other hand, if H is a subgroup for which $H+A=A$, then A must be H-periodic (since otherwise there would be an element $a \in A$ such that $H+a \nsubseteq A$, contradicting that $H+a \subseteq H+A=A$). Note that if a subset A is H-periodic, then $A+B$ is also H-periodic, since $H+(A+B)=(H+A)+B=A+B$.

The maximal subgroup (with respect to inclusion) for which A is H-periodic is called the *stabilizer* of A and is denoted $H(A)$. In light of previous discussion, it is unique and equal to $H(A)=\{g \in G \mid g+A=A\}$ (which is easily checked to be a subgroup). We will often say that A is *maximally H-periodic*, meaning that $H=H(A)$ is the stabilizer of A. Since the subgroup $H(A)$ for which A is maximally periodic is unique, it follows that any other subgroup H for which A is H-periodic is also a subgroup of $H(A)$.

A subset that is maximally H-periodic, with H the trivial group, is *aperiodic*. A subset that is H-periodic, with H nontrivial, is *periodic*. An H-hole of a set A is an element $h \in (H+A)\setminus A$. Often the subgroup H will be implicity understood without confusion and in such cases will be dropped from the notation.

Finally, we need some useful notation for the homomorphisms that will repeatedly show up. Given a subgroup H_a of an abelian group G, we use $\phi_a : G \to G/H_a$ to denote the natural homomorphism. If H_a has finite index in G, then we will associate a with the index, namely $[G : H_a]=a$. Note for noncyclic groups that $a=b$ does NOT imply that $H_a=H_b$, nor that $\phi_a=\phi_b$. It is an important observation that if A is maximally H_a-periodic, then $\phi_a(A)$ is aperiodic.

We can now state Kneser's Theorem. The original proof appears in [44] with English translations in [49] [41], a density version appears in [43] with English translation in [32], and a vector-space version (which provides an alternative proof to the original theorem) appears in [39].

Kneser's Theorem. *Let G be an abelian group, and let $A_1, A_2, \ldots, A_n$ be a collection of finite, nonempty subsets of G. If $\sum_{i=1}^{n} A_i$ is maximally H_a-periodic, then*

$$\left|\sum_{i=1}^{n} \phi_a(A_i)\right| \geq \sum_{i=1}^{n} |\phi_a(A_i)| - n + 1.$$

There are many alternative formulations of Kneser's Theorem, stated below in Theorem 1.1, which reflect varying aspects of Kneser's Theorem. Statement (vii) is inspired by the prime case in [53].

Theorem 1.1. *Let G be an abelian group, let $A_1, A_2, \ldots, A_n$ be a collection of finite, nonempty subsets of G, and suppose that $\sum_{i=1}^{n} A_i$ is maximally H_a-periodic. Then the following statements are all equivalent:*

(i) Kneser's Theorem: $|\sum_{i=1}^{n} \phi_a(A_i)| \geq \sum_{i=1}^{n} |\phi_a(A_i)| - n + 1$,

(ii) $|\sum_{i=1}^{n} A_i| \geq \sum_{i=1}^{n} |H_a + A_i| - (n-1)|H_a|$,

(iii) $|\sum_{i=1}^{n} A_i| \geq \sum_{i=1}^{n} |A_i| - (n-1)|H_a|$,

(iv) either $|\sum_{i=1}^{n} A_i| \geq \sum_{i=1}^{n} |A_i| - n + 1$ *or* $\sum_{i=1}^{n} A_i$ *is periodic,*

(v) either $|\sum_{i=1}^{n} A_i| \geq \sum_{i=1}^{n} |H_a + A_i| - (n-2)|H_a|$ *or* $|\sum_{i=1}^{n} A_i| = \sum_{i=1}^{n} |H_a + A_i| - (n-1)|H_a|$,

(vi) any one of the above five statements only in the case $n = 2$.

Additionally, in the case where G is finite, the following statement is also equivalent to any of the above statements:

(vii) if $n = 3$ and $\sum_{i=1}^{3} |A_i| \geq |G| + |H_a| + 1$, *then* $\sum_{i=1}^{3} A_i = G$.

Before giving a proof of Theorem 1.1, we first provide relevant commentary on the meaning and implications of Kneser's Theorem. In the special case that G is of prime order, then G contains no proper, nontrivial subgroups. Hence in this case, Kneser's Theorem reduces to what is known as the Cauchy-Davenport Theorem (CDT), which was originally proven by Cauchy [10], and later independently re-derived by Davenport [11].

Cauchy-Davenport Theorem (CDT). *If $A_1, A_2, \ldots, A_n$ are a collection of nonempty subsets of $\mathbb{Z}/p\mathbb{Z}$ with p prime, then*

$$|\sum_{i=1}^{n} A_i| \geq \min\{p, \sum_{i=1}^{n} |A_i| - n + 1\}.$$

Observe that the $n > 2$ case of CDT follows from its $n = 2$ case by a simple inductive argument. One way to view CDT is as follows. No matter how you choose a selection of elements $a_j \in A_j$, $j \geq 2$, then the addition of a_j to the sumset $\sum_{i=1}^{j-1} A_i$ transfers (by translation) all previous elements of $\sum_{i=1}^{j-1} A_i$ into the new sumset $\sum_{i=1}^{j} A_i$, while each further element of A_j guarantees one additional element in the sumset $\sum_{i=1}^{j} A_i$ beyond those of $\sum_{i=1}^{j-1} A_i + a_j$ (unless, of course, the sumset $\sum_{i=1}^{j} A_i$ is so large as to contain all p possible elements from $\mathbb{Z}/p\mathbb{Z}$). Observe that the Cauchy-Davenport bound is achieved when the A_i are all arithmetic progressions with common difference.

If $|\sum_{i=1}^{j} A_i| \geq |\sum_{i=1}^{j-1} A_i| + |A_j| - 1$ holds for every $j = 2, \ldots, n$, then a simple inductive argument shows that $|\sum_{i=1}^{n} A_i| \geq \sum_{i=1}^{n} |A_i| - n + 1$ (i.e., if the Cauchy-Davenport bound holds locally for each pairwise sumset of $\sum_{i=1}^{j-1} A_i$ and A_j, $j = 2, \ldots, n$, then the Cauchy-Davenport bound holds globally for the n-fold sumset of the A_i)—this is an important observation that will be used at several points in this thesis.

Part of the utility of Kneser's Theorem comes from the fact that it either gives the

Cauchy-Davenport bound on the cardinality of the sumset of the A_i, or else reduces the question from one about subsets A_i of G, to one involving subsets $\phi_a(A_i)$ of the simpler abelian group G/H_a.

Note that if $A+B$ is maximally H_a-periodic and $\rho = |A+H_a| - |A| + |B+H_a| - |B|$ is the number of holes in A and B, then Kneser's Theorem (via Theorem 1.1(ii)) implies $|A+B| \geq |A| + |B| - |H_a| + \rho$. Consequently, if either A or B contains a unique element from some H_a-coset, then $|A+B| \geq |A|+|B|-1$. More generally, if $\rho = \sum\limits_{i=1}^{n} |H_a + A_i| - |A_i|$ is the total number of holes in the A_i, then $|\sum\limits_{i=1}^{n} A_i| \geq \sum\limits_{i=1}^{n} |A_i| - (n-1)|H_a| + \rho$. Hence, if $|\sum\limits_{i=1}^{n} A_i| < \sum\limits_{i=1}^{n} |A_i| - n + 1$ (i.e, the Cauchy-Davenport bound does not hold), then

$$\rho < (n-1)(|H_a| - 1).$$

If the inequality $|\sum\limits_{i=1}^{n} \phi_a(A_i)| \geq \sum\limits_{i=1}^{n} |\phi_a(A_i)| - \frac{n-2}{|H_a|}$ were to hold, then $|H_a| \cdot |\sum\limits_{i=1}^{n} \phi_a(A_i)| \geq |H_a| \cdot \sum\limits_{i=1}^{n} |\phi_a(A_i)| - n + 2$ would follow. Thus, since $\sum\limits_{i=1}^{n} A_i$ is H_a-periodic, and since $\rho \geq 0$, it would follow that

$$|\sum_{i=1}^{n} A_i| \geq \sum_{i=1}^{n} |H_a + A_i| - n + 2 = \sum_{i=1}^{n} |A_i| + \rho - n + 2 \geq \sum_{i=1}^{n} |A_i| - n + 2.$$

Consequently, if $|\sum\limits_{i=1}^{n} A_i| \leq \sum\limits_{i=1}^{n} |A_i| - n + 1$, then $|\sum\limits_{i=1}^{n} \phi_a(A_i)| < \sum\limits_{i=1}^{n} |\phi_a(A_i)| - \frac{n-2}{|H_a|}$ must hold; and in particular, if $n = 2$ as well, then equality must hold in Theorem 1.1(i).

Summarizing, Kneser's Theorem says that the Cauchy-Davenport bound holds modulo the stabilizer H_a of $\sum\limits_{i=1}^{n} A_i$. Consequently, if the Cauchy-Davenport bound does not hold for the sumset $\sum\limits_{i=1}^{n} A_i$, then $\sum\limits_{i=1}^{n} A_i$ must be periodic. Additionally, as noted above, there cannot be very many H_a-holes contained among the sets A_i. Indeed, the sets A_i must on average be essentially H_a-periodic sets themselves, i.e., they must be very close to being periodic sets, with the average difference $\frac{1}{n}\sum_{i=1}^{n}|(H_a+A_i)\setminus A_i|$ at most $\frac{n-1}{n}|H_a|$ (via Theorem 1.1(ii)).

We conclude this section with the proof of Theorem 1.1.

Proof. Multiplying both sides of the inequality from (i) by $|H_a|$ yields the inequality from (ii). Hence (i) implies (ii). The implications, (ii) implies (iii) and (iii) implies (iv), are immediate. Since H_a is maximal, it follows (as noted earlier) that $\phi_a\left(\sum_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\phi_a(A_i)$ is aperiodic, whence (iv) immediately implies (i). Thus we see that the first four statements are equivalent. That (v) implies (ii) is also immediate. Since $\sum_{i=1}^{n}A_i$ and each H_a+A_i are all H_a-periodic, it follows that every term in the inequality from (ii) is divisible by $|H_a|$. Hence if the inequality in (ii) is strict, then $|\sum_{i=1}^{n}A_i|\geq\sum_{i=1}^{n}|H_a+A_i|-(n-1)|H_a|+|H_a|$. Consequently, it follows that (ii) implies (v), showing that (v) is also equivalent to the first four statements.

Since the arguments for the above implications work equally well when restricted only to the $n=2$ cases, then to show the equivalence of (vi) with the first five statements, it suffices to show that the $n=2$ case of (iv) implies the general case of (iv). We proceed to do so.

As noted earlier, if $|\sum_{i=1}^{j}A_i|\geq|\sum_{i=1}^{j-1}A_i|+|A_j|-1$ holds for every $j=2,\ldots,n$, then $|\sum_{i=1}^{n}A_i|\geq\sum_{i=1}^{n}|A_i|-n+1$. Hence, if $|\sum_{i=1}^{n}A_i|<\sum_{i=1}^{n}|A_i|-n+1$, then there exists an index $j\geq 2$ such that $|\sum_{i=1}^{j}A_i|<|\sum_{i=1}^{j-1}A_i|+|A_j|-1$. Thus applying statement (iv) in the case $n=2$ to the pair of sets $\sum_{i=1}^{j-1}A_i$ and A_j, it follows that $\sum_{i=1}^{j}A_i$ is periodic, implying that $\sum_{i=1}^{n}A_i$ is periodic. Thus the $n=2$ case of (iv) does imply the general case of (iv).

It remains to show the equivalence of (vii). Applying (iii) to the sets A_i, $i=1,2,3$, and using the hypotheses of (vii), it follows that $|\sum_{i=1}^{3}A_i|\geq\sum_{i=1}^{3}|A_i|-2|H_a|\geq|G|-|H_a|+1$. However, since $\sum_{i=1}^{3}A_i$ is H_a-periodic, and hence $|\sum_{i=1}^{3}A_i|$ is divisible by $|H_a|$, then the inequality from the previous sentence implies that $|\sum_{i=1}^{3} A_i| \geq |G|$, whence $\sum_{i=1}^{3} A_i = G$. Thus we see that (iii) implies (vii).

We proceed to complete the proof by showing that (vii) implies the $n = 2$ case of (iii). Suppose that (iii) does not hold for the pair A_1 and A_2. Hence $|A_1+A_2| \leq |A_1|+|A_2|-|H_a|-1$. Thus $|-\overline{A_1 + A_2}| \geq |G| - |A_1| - |A_2| + |H_a| + 1$, whence the sets A_1, A_2, and $-\overline{A_1 + A_2}$ satisfy the hypothesis of (vii). Applying (vii), it follows that $A_1 + A_2 - \overline{A_1 + A_2} = G$. Thus $0 \in A_1 + A_2 - \overline{A_1 + A_2}$, so that $0 = a_1 + a_2 - c$ for some $a_1 \in A_1$, $a_2 \in A_2$ and $c \in \overline{A_1 + A_2}$. As a result, $c = a_1 + a_2$, whence $c \in A_1 + A_2$, which contradicts that $c \in \overline{A_1 + A_2}$. So the $n = 2$ case of (iii) does follow from (vii), completing the proof. □

1.2 A New Proof

The original proof of Kneser's Theorem made use of the Theorem 1.1(ii) formulation (in the case $n = 2$) and an extremal argument involving what is known as the Dyson e-transform—the Dyson e-transform of the pair of sets A and B is the pair of sets $A(e) = (e + A) \cup (-e + B)$ and $B(e) = (e + A) \cap (-e + B)$. Note that $A(e) + B(e) \subseteq A + B$, and from inclusion-exclusion that $|A| + |B| = |e + A| + |-e + B| = |A(e)| + |B(e)|$ (these two properties, for instance, allow one to conclude $|A+B| \geq |A|+|B|-1$ provided $|A(e)+B(e)| \geq |A(e)| + |B(e)| - 1$ is known). The original proof also made use of the following lemma, which, though it will not be needed for our proof of Kneser's Theorem, will be needed later in Chapter 6. [41]

Kneser Lemma. *Let C_0 be a finite subset of an abelian group. If $C_0 = C_1 \cup C_2$ with $C_i \neq C_0$ $(i = 1, 2)$, then $\min_{i=1,2}\{|C_i| + |H_{k_i}|\} \leq |C_0| + |H_{k_0}|$, where H_{k_i} is the maximal group for which C_i is H_{k_i}-periodic $(i = 0, 1, 2)$.*

The proof we present below will instead make use of the Theorem 1.1(iv) formulation (in the case $n = 2$), a new extremal argument involving a simplified version of the Dyson e-transform, and the following two basic but important propositions. However, before proceeding, we need to introduce the notation of $\nu_g(A, B)$ for the number of ways that an element g can be written as a sum of the form $a + b$ with $a \in A$ and $b \in B$. Hence those $g \in G$ with $\nu_g(A, B) > 0$ are the elements of $A + B$. Also note that $\nu_g(A, B) = |(g - B) \cap A| = |(g - A) \cap B|$.

Proposition 1.2. *Let A and B be nonempty subsets of a finite abelian group G. If $|A| + |B| \geq |G| + 1$, then $A + B = G$.*

Proof. Since $|A|+|B| \geq |G|+1$, it follows from the pigeonhole principle that the intersection $(g-B)\cap A$ is nonempty for every $g \in G$. Hence $\nu_g(A, B) = |(g-B)\cap A| \geq 1$ for all $g \in G$. $\square$

Proposition 1.3 below is quite useful since it shows that if the pair A and B does not satisfy the Cauchy-Davenport bound, then elements can be removed from A and B without affecting the sumset—in any possible way—until the cardinalities of A and B are reduced sufficiently so that the Cauchy-Davenport bound is achieved in the resulting sets. Alternatively, Theorem 1.3 implies that if $A+B$ contains a unique expression element, then A and B must satisfy the Cauchy-Davenport bound. We also remark that it is a consequence of results of Kemperman and Liapin [41] [40] [13] that the following proposition holds in a nonabelian group setting as well. The proof below follows Scherk [52].

Proposition 1.3. *Let A and B be nonempty, finite subsets of an abelian group G, and let r be an integer. If $|A + B| < |A| + |B| - r$, then $\nu_g(A, B) > r$ for every $g \in A + B$.*

Proof. Note that the conclusion is equivalent to saying that $A+B' = A+B$ for every subset $B' \subseteq B$ with $|B'| \geq |B| - r$ (which is how the proposition will often be used). The cases

$r \leq 0$ are trivial, and the general $r > 0$ case follows from r applications of the $r = 1$ case. So we may assume $r = 1$. We prove the contrapositive.

By hypothesis there is a unique expression element $a + b$, with $a \in A$ and $b \in B$. By translation, we may w.l.o.g. assume $a = b = 0$. If $|B| = 1$, then $|A + B| \geq |A| + |B| - 1$ holds trivially. We proceed by induction on $|B|$. Choose nonzero $b \in B$. Since 0 is a unique expression element, it follows that $0 \notin A + b$, whence $A + b \neq A$. Hence by cardinality considerations, it follows that there is an element $a_0 \in A$ with $a_0 + b \notin A$. Let $B_1 = \{b_i \in B \mid a_0 + b_i \notin A\}$, and let $A_1 = \{a_0 + b_i \mid b_i \in B_1\}$. Note that $0 \notin B_1$ and that $b \in B_1$. Hence, letting $A_2 = A \cup A_1$ and $B_2 = B \setminus B_1$, it follows that A_1 and B_2 are both nonempty, finite subsets, and that $|B_2| < |B|$. Let $a_2 \in A_2$ and $b_2 \in B_2$. If $a_2 \in A$, then $a_2 + b_2 \in A + B$ follows from $B_2 \subseteq B$. Otherwise, $a_2 = a_0 + b_i$ for some $b_i \in B_1$. Thus $a_2 + b_2 = a_0 + b_i + b_2 = (a_0 + b_2) + b_i$. Since $b_2 \in B_2$, then $b_2 \notin B_1$, whence it follows from the definition of B_1 that $a_0 + b_2 \in A$. Hence $a_2 + b_2 = (a_0 + b_2) + b_i \in A + B$ in this case as well. Consequently, $A_2 + B_2 \subseteq A + B$. Also note that $|A_2| + |B_2| = |A| + |B|$. Thus if we knew $A_2 + B_2$ had a unique expression element, then the proposition would follow by applying the induction hypothesis to the pair A_2 and B_2. We proceed to show that 0 is such an element.

Note $0 \in A_2$ (since $0 \in A$ and $A \subseteq A_2$) and $0 \in B_2$ (since $0 \notin B_1$). Suppose $a_2 + b_2 = 0$ with $a_2 \in A_2$ and $b_2 \in B_2$. If $a_2 \in A$, then since 0 is a unique expression element in $A + B$, it follows that $a_2 = 0$ and $b_2 = 0$. Otherwise, $a_2 = a_0 + b_j$, for some $b_j \in B_1$. Then $0 = a_2 + b_2 = (a_0 + b_2) + b_j$ with $a_0 + b_2 \in A$ (since $b_2 \notin B_1$). Hence, since 0 is a unique expression element in $A + B$, it follows that $0 = b_j \in B_1$. However, this contradicts that $0 \notin B_1$. Thus 0 is a unique expression element in $A_2 + B_2$, completing the proof as remarked earlier. $\square$

We now proceed with the proof of Kneser's Theorem.

Proof. Let A and B be finite, nonempty subsets of an abelian group G. We need to show that either $|A+B| \geq |A|+|B|-1$ or else $A+B$ is periodic. We proceed by induction on the lexigraphic order of the unordered pair $|A|$ and $|B|$. Note that if either $|A|=1$ or $|B|=1$, then $|A+B| \geq |A|+|B|-1$ follows trivially. Inductively assume the theorem true for any pair of subsets A_0 and B_0 satisfying either $|A_0|+|B_0| < |A|+|B|$ or else $|A_0|+|B_0| = |A|+|B|$ and $\min\{|A_0|, |B_0|\} < \min\{|A|, |B|\}$.

Let A' and B' be a maximal (with respect to inclusion) pair of subsets such that $A \subseteq A'$, $B \subseteq B'$ and $A'+B' = A+B$. Since $|A|$ and $|B|$ are finite, it follows that $|A+B|$ is finite. Hence A' and B' exist, and both have finite cardinality at most $|A+B|$. Let $r = |A'|+|B'|-|A|-|B|$. We may w.l.o.g. assume $|A'| \geq |B'|$, and we may also assume

$$|A'+B'| \leq |A'|+|B'|-r-2, \tag{1.1}$$

since otherwise $|A+B| \geq |A|+|B|-1$ follows, completing the proof.

Suppose first that there is a finite, nontrivial subgroup H_a such that there are at most $|H_a|-1$ H_a-holes contained among the sets A' and B' (i.e., that the pair A' and B' is very close to being a pair of periodic sets). Hence, given any $\alpha \in A'$ and $\beta \in B'$, it follows that $|(\alpha+H_a) \cap A'| + |(\beta+H_a) \cap B'| \geq |H_a|+1$. Note that we can translate the sumset $((\alpha+H_a) \cap A') + ((\beta+H_a) \cap B')$ by adding $-\alpha-\beta$. Hence

$$|((\alpha+H_a) \cap A') + ((\beta+H_a) \cap B')| =$$

$$|-\alpha-\beta+((\alpha+H_a) \cap A') + ((\beta+H_a) \cap B')| = |(H_a \cap (-\alpha+A')) + (H_a \cap (-\beta+B'))|. \tag{1.2}$$

Applying Proposition 1.2 to the pair $H_a \cap (-\alpha + A')$ and $H_a \cap (-\beta + B')$ of subsets of the abelian group H_a, it follows that their sumset is all of H_a. Thus from (1.2) we conclude that the entire H_a-coset containing $\alpha + \beta$ is contained in $A' + B' = A + B$. Since $\alpha \in A'$ and $\beta \in B'$ were both arbitrary, this implies that $A + B$ is H_a-periodic with nontrivial period H_a, completing the proof. So we may assume that there are at least $|H_a|$ H_a-holes contained among the sets A' and B' for any finite, nontrivial subgroup H_a.

For any $e \in A' - B'$, we have that $A' \cap (e + B') \neq \emptyset$. However, if $e + B' \subseteq A'$ for every $e \in A' - B'$, then $A' - B' + B' = A'$ (since $0 \in B' - B'$), implying that $B' - B' \subseteq H(A')$. Thus, since the case $|B'| = 1$ (implying $|B| = 1$) has already been handled in the base of the induction, it follows that A' is periodic, whence $A' + B' = A + B$ is also periodic, completing the proof. So we can choose $e \in A' - B'$ such that $|A' \cap (e + B')|$ is maximized subject to $e + B' \nsubseteq A'$.

Let $A'(e) = A' \cup (e + B')$ and let $B'(e) = A' \cap (e + B')$. Since $e \in A' - B'$, then both $A'(e)$ and $B'(e)$ are nonempty. Note $|A'(e)| + |B'(e)| = |A'| + |B'|$ and $A'(e) + B'(e) \subseteq e + A + B$. Hence $|A'(e) + B'(e)| < |A'(e)| + |B'(e)| - r - 1 < |A'(e)| + |B'(e)| - r$ follows from (1.1). Thus applying Proposition 1.3, it follows that any subset $B''(e) \subseteq B'(e)$, with $|B''(e)| = |B'(e)| - r$, satisfies $A'(e) + B''(e) = A'(e) + B'(e)$. Note $|A'(e)| + |B''(e)| = |A| + |B|$. However, since $e + B' \nsubseteq A'$, and since $|A'| \geq |B'|$, it follows that $|B'(e)| < \min\{|A'|, |B'|\} \leq \min\{|A|, |B|\} + r$. Hence $|B''(e)| < \min\{|A|, |B|\}$, allowing us to apply the induction hypothesis to the pair $A'(e)$ and $B''(e)$. Since $|A'(e)| + |B''(e)| = |A| + |B|$, and since $A'(e) + B'(e) \subseteq e + A' + B' = e + A + B$, then if $|A'(e) + B''(e)| \geq |A'(e)| + |B''(e)| - 1 = |A| + |B| - 1$, it follows that $|A + B| \geq |A| + |B| - 1$, completing the proof. So after applying the induction hypothesis it follows that $A'(e) + B''(e) = A'(e) + B'(e)$ is maximally H_a-periodic with nontrivial period.

Next suppose that $B'(e)$ is not H_a-periodic. Then $B'(e)$ must have an H_a-hole x. Hence, since $A'(e)+B'(e)$ is H_a-periodic, it follows that

$$(A'\cup(e+B'))+(B'(e)\cup\{x\})=A'(e)+(B'(e)\cup\{x\})=A'(e)+B'(e)\subseteq e+A+B.$$

Consequently, $x-e+A'\subseteq A+B$ and $x+B'\subseteq A+B$. Since x is an H_a-hole in $B'(e)$, then (via the definition of $B'(e)$) either $x\notin A'$ or $x-e\notin B'$. Hence, in view of the last two sentences, it follows that we can contradict the maximality of the pair A' and B' by either adding x to A' (if $x\notin A'$) or else by adding $x-e$ to B' (if $x-e\notin B'$). So we may assume that $B'(e)$ is H_a-periodic.

Since $|A'(e)+B'(e)|\leq|A'+B'|<|A'|+|B'|-r-1=|A'(e)|+|B'(e)|-r-1$ (in view of (1.1)), then from Proposition 1.3 it follows that we can remove any combination of $r+1$ elements from the sets $A'(e)$ and $B'(e)$, yielding subsets A'' and B'', such that $A''+B''=A'(e)+B'(e)$ and $|A''|+|B''|=|A|+|B|-1$. Choose such subsets A'' and B'' that maximize $|\phi_a(A'')|+|\phi_a(B'')|$. From the maximality, it follows that if $|\phi_a(A'')|+|\phi_a(B'')|<|\phi_a(A'(e))|+|\phi_a(B'(e))|$, then $|\phi_a(A'')|=|A''|$ and $|\phi_a(B'')|=|B''|$. Hence, since $A'(e)+B'(e)$ is H_a-periodic with nontrivial period, it follows that

$$\begin{aligned}&|A+B|=|A'+B'|\geq|A'(e)+B'(e)|=|A''+B''|\geq\\&\quad|H_a|\cdot\max\{|\phi_a(B'')|,\,|\phi_a(A'')|\}\geq 2\cdot\max\{|B''|,\,|A''|\}\geq|A''|+|B''|=|A|+|B|-1,\end{aligned}$$

completing the proof. So we may instead assume that

$$|\phi_a(A'(e))|+|\phi_a(B'(e))|=|\phi_a(A'')|+|\phi_a(B'')|<|A|+|B|.$$

Consequently, we see that we can apply the induction hypothesis to the subsets $\phi_a(B'(e))$

and $\phi_a(A'(e))$ of the abelian group G/H_a. Since $A'(e)+B'(e)$ is maximally H_a-periodic with nontrivial period, it follows that $\phi_a(A'(e)) + \phi_a(B'(e))$ is aperiodic, whence the induction hypothesis implies that

$$|\phi_a(A'(e)) + \phi_a(B'(e))| \geq |\phi_a(A'(e))| + |\phi_a(B'(e))| - 1. \tag{1.3}$$

Let ρ be the number of H_a-holes contained in the pair $A'(e)$ and $B'(e)$, and let ρ' be the number of H_a-holes contained in the pair A' and B'. Partition the set A' into the disjoint sets $A' \cap (e + B')$, A'_1 and A'_2, where A'_1 consists of those elements of A' which modulo H_a are contained in $\phi_a(A') \cap \phi_a(e+B')$ but which are not in $A' \cap (e+B')$, and where A'_2 are the remaining elements of A' not contained modulo H_a in $\phi_a(A') \cap \phi_a(e+B')$. Likewise partition the set $e+B' = (A' \cap (e+B')) \cup B'_1 \cup B'_2$. Let ρ'' be the number of H_a-holes contained among A'_2 and B'_2. Since $A' \cap (e + B') = B'(e)$ is H_a-periodic, it follows that $\phi_a(A'_1) = \phi_a(B'_1)$. Hence, since $A'_1 \cap B'_1$ is empty, then it follows that $|A'_1| + |B'_1| = |A_1 \cup B_1| \leq |H_a||\phi_a(A'_1)|$. Also, observe (since $A' \cap (e + B')$ is H_a-periodic) that

$$\rho = \rho'' + |H_a| \cdot |\phi_a(A'_1)| - |A'_1| - |B'_1|. \tag{1.4}$$

Multiplying both sides of (1.3) by $|H_a|$, it follows that

$$|A'(e) + B'(e)| \geq |A'| + |B'| - |H_a| + \rho \geq |A| + |B| - |H_a| + \rho. \tag{1.5}$$

Suppose $|\phi_a(A'_1)| = 0$. Hence, since $A' \cap (e + B')$ is H_a-periodic, it follows from (1.4) that $\rho = \rho'' = \rho'$. Since the pair A' and B' contains at least $|H_a|$ H_a-holes, it follows that $\rho' \geq |H_a|$, whence (1.5) implies $|A+B| \geq |A'(e)+B'(e)| \geq |A|+|B|$, completing the proof.

So we may assume $\phi_a(A_1')$ is nonempty.

Let $\alpha_1, \ldots, \alpha_n \in G$ be a set of distinct modulo H_a representatives for $\phi_a(A_1')$, let $C_i = (\alpha_i + H_a) \cap A'$, and let $D_i = (\alpha_i + H_a) \cap (e + B')$. Note that $e' \in C_i - D_i \subseteq H_a$ are exactly those elements such that $(e' + D_i) \cap C_i$ is nonempty. Additionally, since $A' \cap (e + B')$ is H_a-periodic, and since $e' \in H_a$, it follows that $A' \cap (e + B') \subseteq A' \cap (e' + e + B')$. Thus, in view of the previous two sentences, unless $e' + e + B' \subseteq A'$, then the element $e' + e$ will contradict the maximality of e. Hence in order to avoid this contradiction we must have: (a) B_2' empty (else w.l.o.g. there will be an H_a-coset $\beta + H_a$ that intersects $e' + e + B'$ but not A'), and (b) $e' + D_i \subseteq C_i$ for each $e' \in C_i - D_i$ (else there will be an element from the coset $\alpha_i + H_a$ contained in $e' + e + B'$ but not in A'), and (c) $C_i - D_i = C_j - D_j$ for all i and j (else there will be an element $e' \in C_i - D_i$ but $e' \notin C_j - D_j$, whence the elements from the coset $e' + \alpha_j + H_a$ contained in $e' + e + B'$ will not be contained in A', but some element from the coset $e' + \alpha_i + H_a$ contained in $e' + e + B'$ will also be contained in A').

Since $e' + D_i \subseteq C_i$ for each $e' \in C_i - D_i$, it follows that $C_i - D_i + D_i = C_i$, implying that $D_i - D_i \subseteq H(C_i)$. Hence $C_i - D_i = -d_i + C_i$ for any $d_i \in D_i$. Thus, since $C_i - D_i = C_j - D_j$ for all i and j, it follows that $C_i = C_j + (d_i - d_j)$. Consequently, the C_i are all just translates of one another, implying that $H(C_i) = H(C_j) = H_{ka} \leq H_a$ and that $|\phi_{ka}(D_i)| = 1$ (whence $|D_i| \leq |H_{ka}|$), for all i and j. Note H_{ka} must be a proper subgroup of H_a, else $C_i \cap D_i$ would be nonempty, a contradiction.

Since B_2' is empty, and since $A'(e) + B'(e)$ is H_a-periodic, then by the maximality of A', it follows that each α_i must have another $\alpha_{\sigma(i)}$ (for some mapping $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$), such that $\alpha_i + \alpha_{\sigma(i)} + H_a \nsubseteq A'(e) + B'(e)$. Hence $C_i + D_{\sigma(i)}$ is disjoint from $A'(e) + B'(e)$. Let h_i be the number of holes contained in $C_i \cup (d_i - d_{\sigma(i)} + D_{\sigma(i)})$. Since all the C_i are H_{ka}-periodic, and since $|\phi_{ka}(D_i)| = 1$, it follows that $\rho - \rho'' \geq h_i$. On the other hand,

$$|C_i + D_{\sigma(i)}| = |C_i| \geq |H_a| - |H_{ka}| - h_i. \tag{1.6}$$

Since $\rho \geq \rho - \rho'' \geq h_i$, and since $C_i + D_{\sigma(i)}$ is disjoint from $A'(e) + B'(e)$, then H_{ka} is nontrivial, else $|A + B| \geq |A| + |B| - 1$ follows from (1.5) and (1.6). Hence A'_2 must be nonempty, and there must be elements $\beta \in A'_2$ and $\beta' \in B'_1$ such that $\beta + \beta' \notin ((A' \setminus A'_2) + (e + B')) \cup (A'(e) + B'(e))$, else $A' + B' = ((A' \setminus A'_2) + (e + B')) \cup (A'(e) + B'(e))$ will be H_{ka}-periodic with nontrivial period H_{ka}, completing the proof.

However, we must have $|(\beta + H_a) \cap A'| = |H_a| - |H_{ka}| + s$ with $s > 0$, else $\rho'' \geq |H_{ka}|$, whence (1.5) and (1.6) and $\rho \geq \rho'' + h_i$ will imply $|A + B| \geq |A| + |B| - 1$. Consequently, $|\phi_{ka}((\beta + H_a) \cap A')| = |\phi_{ka}(H_a)|$. Note $\rho'' \geq |H_{ka}| - s$. Also note, since $H_{ka} < H_a$, that $|H_a| - |H_{ka}| \geq |H_{ka}|$, and that $\phi_a(\beta+\beta') \notin A'(e)+B'(e)$ follows from $A'(e)+B'(e)$ being H_a-periodic with $\beta+\beta' \notin A'(e)+B'(e)$. Hence if $\phi_a(\beta+\beta') \notin \phi_a((A' \setminus A'_2)+e+B'))$, then together with (1.5), (1.6) and $\rho \geq h_i$, it follows that $|A + B| \geq |A| + |B| - 1$, completing the proof. Thus we may w.l.o.g. (by an appropriate choice for σ) assume $\phi_a(\beta + \beta') = \phi_a(\alpha_i + \alpha_{\sigma_i(i)})$ for some i. Hence, since $|\phi_{ka}((\beta + H_a) \cap A')| = |\phi_{ka}(H_a)|$, then it follows, in view of (1.6), that $|(\alpha_i + \alpha_{\sigma(i)} + H_a) \cap (e + A + B)| \geq |C_i| + s \geq |H_a| - |H_{ka}| - h_i + s$, which combined with (1.5), with $\rho \geq \rho'' + h_i$, and with $\rho'' \geq |H_{ka}| - s$, shows that $|A + B| \geq |A| + |B| - 1$, completing the proof. $\square$

Chapter 2

Weighted EGZ

2.1 Discussion

In 1961, Erdős, Ginzburg and Ziv proved the following theorem, which has become one of the foundational theorems of zero-sum additive theory [15].

Erdős-Ginzburg-Ziv Theorem (EGZ). *If S is a sequence of $2m-1$ elements from an abelian group G of order m, then S contains an m-term zero-sum subsequence.*

Note that if $G = \mathbb{Z}/m\mathbb{Z}$ and if S consists of only 0's and 1's, then the m-term zero-sum subsequences of S correspond precisely with the m-term monochromatic subsequences. Hence EGZ can be viewed as an algebraic generalization of the pigeonhole principle. Their theorem spurred the growth of the now-developing field of zero-sum Ramsey Theory (which will be elaborated on in Part II) and has been the subject of numerous and varied generalizations. One such generalization, given below, was proposed by Y. Caro in the early 1990s and was immediately affirmed in the case $n = m$ with m prime by N. Alon. The statement requires that we define a *rearranged subsequence* S' of a sequence S, which is just a sequence that under some permutation of its terms is a subsequence of S. Such a definition is useful, since though many theorems from zero-sum additive theory are phrased in the language of sequences, it is only the allowance for terms with multiplicity, and not the ordering of terms, that is important.

Conjecture 2.1. *Let* $W = \{w_i\}_{i=1}^n$ *be a sequence of integers such that* $\sum_{i=1}^n w_i \equiv 0 \mod m$. *Let* $S = \{b_i\}_{i=1}^{m+n-1}$ *be another sequence of integers. Then there exists a rearranged subsequence* $\{b_{j_i}\}_{i=1}^n$ *of* S *such that* $\sum_{i=1}^n w_i b_{j_i} \equiv 0 \mod m$.

After communicating with A. Bialostocki and Y. Caro, the proof was soon extended to arbitrary n and m prime, where the status of the problem remained. The conjecture was included a few years later in a survey of Y. Caro on problems in zero-sum combinatorics [9], where a reference was made to the (unpublished) proof of the prime case [1]. Soon after, Y. O. Hamidoune published a pair of papers where, in the first he proved that an equivalent form of Conjecture 2.1 holds (in a more general abelian group setting) provided each w_i for $i \in \{1, \ldots, n-1\}$ was relatively prime to m [34], and in the second he introduced the following conjecture, which he verified for $n = m$ [33].

Conjecture 2.2. *Let* S *be a sequence of* $m+n-1$ *elements from a nontrivial abelian group* G *of order* m*, and let* $W = \{w_i\}_{i=1}^n$ *be a sequence of integers whose sum is zero modulo* m. *If the multiplicity of every term of* S *is at most* n*, and if each* w_i *for* $i \le n-1$ *is relatively prime to* m*, then there is a nontrivial subgroup* H *of* G *such that for every* $h \in H$ *there is a rearranged subsequence* $\{b_{h_i}\}_{i=1}^n$ *of* S *with* $\sum_{i=1}^n w_i b_{h_i} = h$.

No further progress was made on either conjecture until 2003, when W. Gao and X. Jin established Conjecture 2.1 in the case of $m = p^2$, for p prime [22].

In this chapter we will completely affirm the conjecture of Caro, as well as many cases in the related conjecture of Hamidoune. The proof technique is a simplified variation of the involved extremal argument that will be used in Chapter 3, and so an understanding of the proof from Section 2.3 will aid greatly in understanding Chapter 3.

2.2 Set Partitions

Before stating the main result explicitly, we first introduce and develop the basic properties of the important concept of set partitions, which will be used throughout this thesis. Let S be a sequence. The length of S is denoted $|S|$. An *n-set partition* of S is a partition of the sequence S into n nonempty subsequences, $A_1, \ldots, A_n$, such that the terms in each subsequence A_i are all distinct, allowing the A_i to be regarded as sets. Set partitions are extremely useful tools in zero-sum additive theory and were first implemented in the original proof of EGZ. They provide an important linking connection between the inverse additive results like those described in Chapter 1 and the zero-sum additive results like EGZ.

The connection comes in part from the observation that if an element g lies in the sumset $\sum_{i=1}^{n} A_i$ of some n-set partition $A_1, \ldots, A_n$ of a sequence S, then an appropriate selection of a term from each A_i yields an n-term subsequence of S whose terms sum to g. In particular, if $|\sum_{i=1}^{n} A_i| \geq |G|$, then every element, including 0, can be represented as the sum of some n-term subsequence. As a consequence, if one is looking for a zero-sum n-term subsequence, then by contradiction every n-set partition of S must have small cardinality sumset, whence inverse structure theorems, like Kneser's Theorem or Cauchy-Davenport Theorem (Chapter 1), can be used to derive structural information about the sets A_i and, consequently, about the sequence S that the A_i partition as well. Chapter 3 will provide an even stronger structure theorem for a sequence S assuming all its n-set partitions have small cardinality sumset.

However, before the machinery of any such structure theorems can be invoked, one first needs to know that S actually has an n-set partition. Fortunately, this is not a very stringent restriction, as the following proposition indicates.

Proposition 2.3. *Let n' and n be positive integers with $n \leq n'$. A sequence S of terms from G has an n'-set partition $A = A_1, \ldots, A_{n'}$ with $|A_i| = 1$ for $i > n$ if and only if $|S| \geq n'$, and for every nonempty subset $X \subseteq G$ with $|X| \leq \frac{|S|-n'-1}{n} + 1$ there are at most $n' + (|X| - 1)n$ terms of S from X. In particular, S has an n-set partition if and only if $|S| \geq n$ and the multiplicity of every term of S is at most n.*

Proposition 2.3 tells us when S has a sufficiently compressed set partition $A_1, \ldots, A_{n'}$, i.e., one which has sufficiently many cardinality one sets $A_{n+1}, \ldots, A_{n'}$. The existence of a sufficiently compressed set partition can be quite useful. For instance, S may not have an n-set partition (if some term has too great a multiplicity), but it is often even more useful to know that S has an n'-set partition (for some $n' \geq n$), $A_1, \ldots, A_{n'}$, sufficiently compressed so that all but at most n of the sets have cardinality one. One reason for this is that it allows us to conclude that $|\sum_{i=1}^{n'} A_i| = |\sum_{i=1}^{n} A_i|$. In essence, a sequence S having a sufficiently compressed n'-set partition, with all but at most n sets having cardinality one, means that a large subsequence of S has an n-set partition, even if S does not. Another reason, is that there may be more terms in the sequence S than need to be included at any one time in the n-set partition (for the purposes of applying a structure theorem), and so the weaker conditions needed to guarantee a compressed n'-set partition versus a non-compressed n-set partition can provide a useful advantage. Additionally, many results and problems from zero-sum additive theory involve restrictions that can often via Proposition 2.3 be immediately translated into the existence of a compressed set partition.

If S does not have a sufficiently compressed set partition, then Proposition 2.3 implies that S is 'essentially' r-chromatic, where $1 \leq r \leq \frac{|S|-n'-1}{n} + 1$, by which we mean that the terms of S are all equal to one of the r distinct terms from X with at most a 'small' number (no more than $|S| - n' - (r-1)n - 1$) of exceptions. In zero-sum applications, these essentially r-chromatic cases often prove the most troublesome. The chapters from the second part will exhibit a large variety of tactical methods used to handle such cases, but unfortunately no overriding strategic method has yet emerged.

Next, we introduce the notation of $S \setminus S'$, when S' is a subsequence of a sequence S, to denote the sequence that results by deleting all terms from S contained in S', and proceed with the proof of Proposition 2.3.

Proof. Suppose S has an n'-set partition $A_1, \ldots, A_{n'}$ with $|A_i| = 1$ for $i > n$. Since none of the n' sets in the set partition can be empty, it trivially follows that $|S| \geq n'$. Additionally, given any subset X, since all terms in the A_i are distinct, and since $|A_i| = 1$ for $i > n$, then it follows there can be at most $n|X|$ terms of S from X partitioned by $A_1, \ldots, A_n$, and at most $n' - n$ terms of S from X partitioned by $A_{n+1}, \ldots A_{n'}$, for a total of at most $(|X| - 1)n + n'$ terms of S from X.

Next suppose that $|S| \geq n'$ and that there are at most $(|X|-1)n+n'$ terms of S from any subset X with $|X| \leq \frac{|S|-n'-1}{n}+1$. Note that there are trivially at most $(|X|-1)n+n'$ terms of S from X from a subset X with $|X| > \frac{|S|-n'-1}{n}+1$ (since the inequality $|X| > \frac{|S|-n'-1}{n}+1$ implies $|S| \leq (|X| - 1)n + n')$. Thus, regardless of the cardinality of X, we may assume that there are at most $(|X| - 1)n + n'$ terms of S from X.

Let $s_1, \ldots, s_u$ be the distinct terms of S arranged in nondecreasing order of multiplicity. Let $X = \{s_1, \ldots, s_{u'}\}$ (possibly empty) be the distinct terms of S with multiplicity at least n, and let m_i, for $i = 1, \ldots, u'$, denote their respective multiplicities. Remove $m_i - n$ terms from S equal to s_i, for $i = 1, \ldots, u'$. Note that if the total number of terms removed in this way exceeded $n'-n$, then this would imply that there were at least $|X|n+n'-n+1$ terms of S from X, contradicting the assumption to the contrary. Hence we can remove an additional

$n-n'-\sum_{i=1}^{u'}(m_i-n)$ terms from S. Let $S''=y_1,\ldots,y_{n'-n}$ denote the subsequence of terms of S removed in this manner, let $S'=S\backslash S''=x_1,\ldots,x_{kn+r}$, where $|S|-(n'-n)=|S'|=kn+r$ with $0\le r<n$. We may w.l.o.g. (by reordering the sequence S) assume the terms x_i of S' have been ordered so that all terms equal to s_1 come first, followed by all terms equal to s_2, and so forth, terminating with the terms equal to s_u. To complete the proof, it suffices to show that S' has an n-set partition—since appending the remaining $n'-n$ terms y_i as singleton sets will then give the n'-set partition of S with the desired properties.

Consider the following sequence A of n subsequences of S' written vertically.

$$A=\begin{pmatrix} x_1 \\ x_{n+1} \\ \vdots \\ x_{(k-1)n+1} \\ x_{kn+1}\end{pmatrix},\ldots,\begin{pmatrix} x_r \\ x_{n+r} \\ \vdots \\ x_{(k-1)n+r} \\ x_{kn+r}\end{pmatrix},\begin{pmatrix} x_{r+1} \\ x_{n+r+1} \\ \vdots \\ x_{(k-1)n+r+1} \\ .\end{pmatrix},\ldots,\begin{pmatrix} x_n \\ x_{2n} \\ \vdots \\ x_{kn} \\ .\end{pmatrix}$$

We will show that A is an n-set partition of S'. Indeed, since $|S'|\ge|S|-(n'-n)\ge n$ (since $|S|\ge n'$), it follows that none of the sets in A are empty. Furthermore, in view of the definition of S' and the fact that the maximum multiplicity of a term in S' does not exceed n, it follows that $x_{j_1n+i}\neq x_{j_2n+i}$, for every i and every $j_1\neq j_2$. Thus A is an n-set partition of S, and the proposition is established.

Note in the case $n=n'$, that the condition that there be at most $(|X|-1)n+n'=n|X|$ terms from S from a subset X is equivalent (in view of the Pigeonhole Principle) to every term of S having multiplicity at most n, i.e., the case when $|X|=1$, whence the latter remark of the proposition follows. $\square$

In a similar spirit to Proposition 1.3, we have the following basic draining result for set partitions for which the Cauchy-Davenport bound does not hold. Proposition 2.4 is often used in the case $B = \{0\}$ by first applying (i) to reduce the number of sets in the set partition to a desired smaller number, and then applying (ii) with $r' = r$ to reduce the number of elements partitioned by the set partition sufficiently so that the Cauchy-Davenport bound holds in the resulting set partition.

Proposition 2.4. *Let S be a finite sequence of elements from an abelian group G, let B be a finite, nonempty subset of G, and let $A = A_1, \ldots, A_n$ be an n-set partition of S, where $|B + \sum_{i=1}^{n} A_i| - |B| + 1 = r$, and $\max_i\{|B + A_i| - |B| + 1\} = s$. Furthermore, let $b_1, \ldots, b_n$ be a subsequence of S such that $b_i \in A_i$ for $i = 1, \ldots, n$, and let r' be an integer with $1 \le r' \le r$.*

(i) There exists a subsequence S' of S and an n'-set partition $A' = A_{j_1}, \ldots, A_{j_{n'}}$ of S', which is a subsequence of the n-set partition $A = A_1, \ldots, A_n$, such that $n' \le r - s + 1$ and $|B + \sum_{i=1}^{n'} A_{j_i}| = |B + \sum_{i=1}^{n} A_i|$.

(ii) There exists a subsequence S' of S of length at most $n + r' - 1$, and an n-set partition $A' = A'_1, \ldots, A'_n$ of S', where $A'_i \subseteq A_i$ for $i = 1, \ldots, n$, such that $|B + \sum_{i=1}^{n} A'_i| \ge |B| - 1 + r'$. Furthermore, $b_i \in A'_i$ for $i = 1, \ldots, n$.

Proof. We first prove (i). Assume w.l.o.g. that $|B + A_1| = |B| - 1 + s$. We will construct the n'-set partition A' in n steps as follows, and S' will be implied implicitly. Denote by $A^{(k)} = A'_1, \ldots, A'_{a_k}$ the constructed sequence after k steps, and hence $A' = A^{(n)}$ and $n' = a_n$. Let $A^{(1)} = A_1$, and for $k = 1, 2, \ldots, n-1$, let

$$A^{(k+1)} = \begin{cases} A^{(k)} & \text{if } |B + \sum_{i=1}^{a_k} A'_i + A_{k+1}| = |B + \sum_{i=1}^{a_k} A'_i| \\ A^{(k)}, A_{k+1} & \text{if } |B + \sum_{i=1}^{a_k} A'_i + A_{k+1}| > |B + \sum_{i=1}^{a_k} A'_i|. \end{cases}$$

It is easily seen by the above algorithm that $|B+\sum_{i=1}^{a_n} A_i'| = |B+\sum_{i=1}^{n} A_i| = r+|B|-1$. Furthermore, since each kept term increases the cardinality of the sumset of the previous terms of A' by at least one, and since $|B+A_1| = |B|-1+s$, it follows that at most $r-s$ terms, excluding A_1, were kept, and thus $a_n = n' \leq 1+r-s$.

The proof of (ii) is similar to that of (i). First, for $i=1,\ldots,n$, let the elements of A_i be $\{a_1^{(i)},\ldots,a_{|A_i|}^{(i)}\}$, where $a_1^{(i)} = b_i$. We will construct the n-set partition A' in a two-loop algorithm. The outer loop has n steps, where at the ith step the set A_i' is constructed using the inner loop. In turn, the inner loop, at the ith step of the outer loop, constructs A_i' in $|A_i|$ steps. For a given i, where $1 \leq i \leq n$, let $A_i^{(k)}$ denote the set constructed after k steps of the inner loop at the ith step of the outer loop, and hence $A' = A_1^{(|A_1|)},\ldots,A_n^{(|A_n|)}$ with S' implied implicitly. For a given j, where $1 \leq j \leq n$, let $A_j^{(1)} = \{b_j\}$, and for $k=1,\ldots,|A_j|-1$, let

$$A_j^{(k+1)} = \begin{cases} A_j^{(k)} & \text{if } |B+\sum_{i=1}^{j-1} A_i^{(|A_i|)} + A_j^{(k)}| = |B+\sum_{i=1}^{j-1} A_i^{(|A_i|)} + (A_j^{(k)} \cup \{a_{k+1}^{(j)}\})|, \\ & \text{or if } |B+\sum_{i=1}^{j-1} A_i^{(|A_i|)} + A_j^{(k)}| \geq |B|-1+r', \\ A_j^{(k)} \cup \{a_{k+1}^{(j)}\} & \text{otherwise.} \end{cases}$$

It is easily seen by the above algorithm that $A_i^{(|A_i|)} \subseteq A_i$ and $b_i \in A_i^{(|A_i|)}$ for $i=1,\ldots,n$, and that $|B+\sum_{i=1}^{n} A_i^{(|A_i|)}| \geq |B|-1+r'$. Furthermore, since each kept element $a_k^{(j)}$ with $k>1$ increases the cardinality of the sumset by at least one, it follows that at most $r'-1$ terms, excluding the b_i's, were kept, and hence $|S'| \leq n+r'-1$. $\square$

2.3 A Weighted Version of EGZ

In this section we present the main result (Theorem 2.5) from the chapter. However, we first give some additional definitions. The *exponent* of an abelian group G is the minimal integer k such that $kg = 0$ for all $g \in G$. If G is finite of order m, then such a k will always exist, and $k|m$ follows as well. We regard the abelian group G as a $\mathbb{Z}$-module. For $w \in \mathbb{Z}$ and $A \subseteq G$, we let $wA = \{wa_i \mid a_i \in A\}$. Thus $2A$ is NOT equal to $A + A$, as is sometimes customary in the literature.

Theorem 2.5. *If S is a sequence of $m + n - 1$ elements from a nontrivial abelian group G of order m and exponent k, and if $W = \{w_i\}_{i=1}^n$ is a sequence of integers whose sum is zero modulo k, then there exists a rearranged subsequence $\{b_i\}_{i=1}^n$ of S such that $\sum\limits_{i=1}^{n} w_i b_i = 0$.*

Furthermore, if S has an n-set partition $A = A_1, \ldots, A_n$ such that $|w_i A_i| = |A_i|$ for all i, then there exists a nontrivial subgroup H of G and an n-set partition $A' = A'_1, \ldots, A'_n$ of S with $H \subseteq \sum\limits_{i=1}^{n} w_i A'_i$ and $|w_i A'_i| = |A'_i|$ for all i.

We note that the example $W = (\underbrace{1, \ldots, 1}_{m-2}, 0, 2)$ and $S = (-1, \underbrace{0, \ldots, 0}_{m-1}, \underbrace{1, \ldots, 1}_{m-1})$ with $G = \mathbb{Z}/m\mathbb{Z}$ and $n = m$ shows that in the above theorem we cannot require $\{b_i\}_{i=1}^n$ to be an actual (including order) subsequence of S. Also, since $|w_i A_i| = |A_i|$ for w_i relatively prime to k (and since both conditions (b) and (c) to be stated at the end of the sentence imply, in view of Proposition 2.3, there exists an n-set partition of S with at least one set A_i of cardinality one), it follows that Theorem 2.5 implies Conjecture 2.2, provided any one of the following conditions also holds: (a) w_n is also relatively prime to m, or (b) $n \geq m$, or (c) every term of S has multiplicity at most $n - 1$.

We proceed with the proof of Theorem 2.5. In the proof, we will essentially be considering an n-set partition $A = A_1, \ldots, A_n$ of S that iteratively maximizes $\sum\limits_{i=1}^{n} |w_i A_i|$, then

$|\sum_{i=1}^{n} w_iA_i|$, and finally $\sum_{i=1}^{n}|\phi_a(w_iA_i)|$, where $\sum_{i=1}^{n} w_iA_i$ is maximally H_a-periodic. With the help of Kneser's Theorem, we will be able to show that we can remove some term b of S from the set partition A leaving the third maximized quantity unaffected. If the second maximized quantity is also preserved, then this will allow us to place the term b back into the n-set partition in such a way as to preserve the first quantity and increase one of the later two quantities, a contradiction, unless the term $\phi_a(b)$ is contained in every set w_iA_i, in which case $H_a = \sum_{i=1}^{n} w_ib + H_a \subseteq \sum_{i=1}^{n} w_iA_i$ will follow from Kneser's Theorem, completing the proof. On the other hand, if removing the term b from its set w_jA_j would destroy the second maximized quantity, then we will use Proposition 1.3 to show that the set w_jA_j locally adds lots of elements to the sumset $\sum_{i=1}^{n} w_iA_i$. An extremal argument will then be used to show that either there must be a term of S that can be removed from A while preserving both the later two maximized quantities, or else there will be very many sets w_iA_i that locally add lots of elements to $\sum_{i=1}^{n} w_iA_i$, enough so that we can conclude that the sumset $\sum_{i=1}^{n} w_iA_i$ has large enough cardinality globally to represent every element of G.

Proof. If there is a term x of S whose multiplicity is at least $n+1$, then S cannot have an n-set partition and Theorem 2.5 follows by choosing $b_i = x$ for all i. Hence we may assume each term of S has multiplicity at most n, whence it follows in view of Proposition 2.3 that there exists an n-set partition $A = A_1, \ldots, A_n$ of S. Choose A such that $\sum_{i=1}^{n}|w_iA_i|$ is maximized.

Suppose $|w_jA_j| < |A_j|$ for some index j (so that the conditions from the furthermore part of Theorem 2.5 do not hold), and let $b,\, b' \in A_j$ with $w_jb = w_jb'$ and $b \neq b'$. If there exists an index r such that $w_rb \notin w_rA_r$, then the n-set partition $A'_1, \ldots, A'_n$ defined by $A'_j = A_j \setminus \{b\}$, $A'_r = A_r \cup \{b\}$ and $A'_i = A_i$ for $i \neq j, r$, contradicts the maximality of

$\sum_{i=1}^{n}|w_iA_i|$. Thus we may assume $w_ib \in w_iA_i$ for all i. Hence, since $\sum_{i=1}^{n} w_i \equiv 0 \mod k$, it follows that $0 = \sum_{i=1}^{n} w_ib \in \sum_{i=1}^{n} w_iA_i$, and the proof is complete by an appropriate selection of a term from each A_i. So we may assume $|w_iA_i| = |A_i|$ for all i. Furthermore, assume A is chosen such that $|\sum_{i=1}^{n} w_iA_i|$ is maximized subject to $|w_iA_i| = |A_i|$ for all i.

If $|\sum_{i=1}^{n} w_iA_i| \geq m$, then the proof is complete with $H = G$. Hence, since $|S| = m + n - 1$ and since $|w_iA_i| = |A_i|$, it follows that we may assume that

$$|\sum_{i=1}^{n} w_iA_i| < \sum_{i=1}^{n}|w_iA_i| - n + 1, \tag{2.1}$$

whence from Kneser's Theorem it follows that $\sum_{i=1}^{n} w_iA_i \stackrel{def}{=} X$ is maximally H_a-periodic for some proper, nontrivial subgroup H_a of G. Assume that A was chosen, from among all n-set partitions $A' = A'_1, \ldots, A'_n$ of S that satisfy $|w_iA'_i| = |A'_i|$ and $\sum_{i=1}^{n} w_iA'_i = X$, such that $\sum_{i=1}^{n}|\phi_a(w_iA_i)|$ is maximized.

If every set w_iA_i with $i \geq 2$ contains an element that is the unique element from its H_a-coset in w_iA_i, then there are at least $(n-1)(|H_a|-1)$ holes among the sets w_iA_i, whence Kneser's Theorem implies that $|\sum_{i=1}^{n} w_iA_i| \geq \sum_{i=1}^{n}|w_iA_i| - (n-1)|H_a| + (n-1)(|H_a|-1) = \sum_{i=1}^{n}|A_i| - n + 1$, contradicting (2.1). Therefore we may assume $|\phi_a(w_jA_j)| < |w_jA_j|$ for some index $j \geq 2$.

Let $j \geq 2$ be an index such that $|\phi_a(w_jA_j)| < |w_jA_j|$. Suppose that

$$\sum_{i=1}^{j} w_iA_i = \sum_{i=1}^{j-1} w_iA_i + w_j(A_j \setminus \{b\}), \tag{2.2}$$

for some $b \in A_j$ such that $\phi_a(w_jA_j) = \phi_a(w_j(A_j \setminus \{b\}))$. Hence, if there exists an index r such that $\phi_a(w_rb) \notin \phi_a(w_rA_r)$, then the n-set partition defined by $A'_j = A_j \setminus \{b\}$, $A'_r = A_r \cup \{b\}$

and $A'_i = A_i$ for $i \neq j, r$, contradicts the maximality of either $|\sum_{i=1}^{n} w_iA_i|$ or $\sum_{i=1}^{n}|\phi_a(w_iA_i)|$. Therefore we may assume $\phi_a(w_ib) \in \phi_a(w_iA_i)$ for all i. Hence, since $\sum_{i=1}^{n} w_i \equiv 0 \mod k$, it follows that $0 = \sum_{i=1}^{n}\phi_a(w_ib) \in \sum_{i=1}^{n}\phi_a(w_iA_i)$. Thus, since $\sum_{i=1}^{n} w_iA_i$ is H_a-periodic, it follows that $H_a \subseteq \sum_{i=1}^{n} w_iA_i$, and the proof is complete with $H = H_a$. So we may assume that (2.2) does not hold, whence in view of Proposition 1.3 it follows that

$$|\sum_{i=1}^{j} w_iA_i| \geq |\sum_{i=1}^{j-1} w_iA_i| + |w_jA_j| - 1. \tag{2.3}$$

Let l, where $2 \leq l \leq n$, be the minimal integer, allowing re-indexing of the w_iA_i, such that

$$|\sum_{i=1}^{j} w_iA_i| \geq |\sum_{i=1}^{j-1} w_iA_i| + |w_jA_j| - 1, \tag{2.4}$$

for all $j \geq l$. From the conclusions of the last two paragraphs, and since by re-indexing we may assume $j = n$ in the previous paragraph, it follows that l exists. Observe that

$$|\sum_{i=1}^{l-1} w_iA_i| < \sum_{i=1}^{l-1}|w_iA_i| - (l-1) + 1, \tag{2.5}$$

since otherwise applying (2.4) iteratively yields $|\sum_{i=1}^{n} w_iA_i| \geq \sum_{i=1}^{n}|w_iA_i| - n - 1$, contradicting (2.1). Hence from Kneser's Theorem and the maximality of H_a, it follows that $\sum_{i=1}^{l-1} w_iA_i$ is maximally H_{ka}-periodic for some nontrivial subgroup $H_{ka} \leq H_a$. Note that (2.5) can only hold provided $l - 1 \geq 2$. Furthermore, if every set w_iA_i with $2 \leq i \leq l-1$ contains an element that is the unique element from its H_{ka}-coset in w_iA_i, then there will be at least $(l-2)(|H_{ka}| - 1)$ holes among the sets w_iA_i with $i \leq l-1$, whence Kneser's Theorem implies that $|\sum_{i=1}^{l-1} w_iA_i| \geq \sum_{i=1}^{l-1}|w_iA_i| - (l-2)|H_{ka}| + (l-2)(|H_{ka}|-1) = \sum_{i=1}^{l-1}|A_i| - (l-1) + 1$, contradicting (2.5). Therefore there must exist a set A_j with $2 \leq j \leq l-1$, such that w_jA_j

does not contain an element that is the unique element from its H_{ka}-coset in $w_j A_j$. Hence, since $H_{ka} \leq H_a$, it follows that $|\phi_a(w_j A_j)| < |w_j A_j|$ for some index j with $2 \leq j \leq l-1$. Thus, since by re-indexing we may assume $j = l-1$, it follows, in view of the paragraph before (2.3), that (2.3) holds with $j = l-1$, which in view of (2.4) contradicts the minimality of l. $\square$

Chapter 3

A Composite Analog of the Cauchy-Davenport Theorem

3.1 Discussion

Let $A = A_1, \ldots, A_n$ be an n-set partition of a sequence S of elements from an abelian group G whose sumset is H_a-periodic. Let $y \in G/H_a$. If $y \in \phi_a(A_i)$ for all i, then y is an H_a-*nonexception*, and otherwise y is an H_a-*exception*. If $|\phi_a^{-1}(y) \cap A_j| \geq 2$, then y is an H_a-*doubled* element of $\phi_a(A_j)$. The number of $y \in G/H_a$ that are H_a-nonexceptions of A is denoted by $N(A, H_a)$. The number of terms x of S such that $\phi_a(x)$ is an H_a-exception of A is denoted by $E(A, H_a)$. Note that

$$N(A, H_a) = \frac{1}{|H_a|}|\bigcap_{i=1}^{n}(A_i + H_a)|$$

and that

$$E(A, H_a) = \sum_{j=1}^{n}(|A_j| - |A_j \cap \bigcap_{i=1}^{n}(A_i + H_a)|).$$

The main result of this chapter is Theorem 3.1. We also give a particular corollary to Theorem 3.1, namely Theorem 3.2, that is often convenient to use in practice, since it incorporates several routine consequences whose arguments are not entirely succinct.

Theorem 3.1. *Let S' be a subsequence of a finite sequence S of terms from an abelian group G, let $W = w_1, \ldots, w_n$ be a sequence of integers such that $w_i g \neq 0$ for all i and all nonzero $g \in G$, let $A = A_1, \ldots, A_n$ be an n-set partition of S', and let $a_i \in A_i$ for $i = 1, \ldots, n$. Then there exists an n-set partition $A' = A'_1, \ldots, A'_n$ of a subsequence S'' of S such that $\sum_{i=1}^{n} w_i A'_i$ is H_a-periodic, $|S'| = |S''|$, $\sum_{i=1}^{n} w_i A_i \subseteq \sum_{i=1}^{n} w_i A'_i$, $a_i \in A'_i$, and*

$$\left|\sum_{i=1}^{n} w_i A'_i\right| \geq \big(E(A', H_a) + (N(A', H_a) - 1)n + 1\big)\,|H_a|.$$

Furthermore, if H_a is nontrivial, then $\phi_a(x) \in \phi_a(A'_i)$ for every $i = 1, \ldots, n$ and $x \in S \setminus S''$.

Theorem 3.2. *Let S' be a subsequence of a finite sequence S of terms from an abelian group G of order m and exponent k, let $W = w_1, \ldots, w_n$ be a sequence of integers such that $(w_i, k) = 1$ for all i, let $P = P_1, \ldots, P_n$ be an n-set partition of S', let $a_i \in P_i$ for $i = 1, \ldots, n$, and let p be the smallest prime divisor of m. If $n \geq \min\{\frac{m}{p} - 1, \frac{|S'|-n+1}{p} - 1\}$, then either:*

(i) there is an n-set partition $A = A_1, \ldots, A_n$ of a subsequence S'' of S with $|S'| = |S''|$, $\sum_{i=1}^{n} w_i P_i \subseteq \sum_{i=1}^{n} w_i A_i$, $a_i \in A_i$, and

$$\left|\sum_{i=1}^{n} w_i A_i\right| \geq \min\{m,\, |S'| - n + 1\},$$

(ii) there is a proper, nontrivial subgroup H_a of index a, a coset $\alpha + H_a$ such that all but e terms of S are from $\alpha + H_a$, where $e \leq \min\{a - 2, \left\lfloor \frac{|S'|-n}{|H_a|} \right\rfloor - 1\}$, an n-set partition $A = A_1, \ldots, A_n$ of subsequence S'' of S with $|S''| = |S'|$, $\sum_{i=1}^{n} w_i P_i \subseteq \sum_{i=1}^{n} w_i A_i$, $a_i \in A_i$, and $\left|\sum_{i=1}^{n} w_i A_i\right| \geq (e+1)|H_a|$, and an n-set partition $B = B_1, \ldots, B_n$ of a subsequence S''_0 of S, with all terms of S''_0 from $\alpha + H_a$ and $|S''_0| \leq n + |H_a| - 1$, such that $\sum_{i=1}^{n} w_i B_i = \alpha \sum_{i=1}^{n} w_i + H_a$.

For a sequence S, let $n \wedge S$ be the set of elements that can be represented as a sum of terms from some n-term subsequence of S. In 1967, Mann gave an easy extension of EGZ, by showing that if $m = |G|$ is prime, $|S| = m + n - 1$, and every term of S has multiplicity at most n, then $n \wedge S = G$ [48]. In 1977, Olson generalized this result in the case $n = m$ to an arbitrary abelian group of order m, by showing that if $|S| = 2m - 1$, and if every term of S has multiplicity at most m, then either $m \wedge S = G$, or there exists a proper, nontrivial subgroup H_a of index a such that $H_a \subseteq m \wedge S$, and all but at most $a - 2$ terms of S are from the same H_a-coset [50]. Unfortunately, while the conclusion of Olson's Theorem was quite strong, including a structure restriction on the sequence S, it failed to cover sequences with length smaller than $2m - 1$. In an effort to alleviate this restriction, Bollobás and Leader obtained a weaker version of Olson's result that was valid for sequences of any length; they showed that if $0 \notin m \wedge S$, then $|m \wedge S| \geq |S| - m + 1$ [8]. Hamidoune improved upon this result—extending, as in Mann's result, from m-sums to arbitrary n-sums—by showing that either $|n \wedge S| \geq |S| - n + 1$ or else there exists a term x of S with $nx \in n \wedge S$ [38].

Theorem 3.1 accomplishes the task of fully generalizing the previous results of Mann, Olson, Bollobás and Leader, and Hamidoune. For non-weighted applications, the sequence W may always be taken to be an n-term sequence consisting entirely of 1's, which is how it will be used in this thesis. However, the weighted versions of Theorems 3.1 and 3.2 are obtained with equal ease from the arguments used in the non-weighted case and provide much more potent tools (as compared with the results of Chapter 2) for zero-sum applications involving weights w_i that are all relatively prime to the exponent of G.

Since a set partition $A_1, \ldots, A_n$ of S partitions the terms of S, it follows that $\sum_{i=1}^{n} |A_i| = |S|$, and hence having the Cauchy-Davenport bound hold for an n-set partition $A_1, \ldots, A_n$ means that $|\sum_{i=1}^{n} A_i| \geq |S| - n + 1$. Thus, unless $N(A', H_a) > 0$ and H_a is a proper, nontrivial subgroup, then Theorem 3.1 implies that the Cauchy-Davenport bound holds for the weighted sumset, namely that $|\sum_{i=1}^{n} w_i A'_i| \geq \min\{|G|, |S'| - n + 1\}$.

In loose terms, Theorem 3.2 gives the existence of an n-set partition satisfying the Cauchy-Davenport bound, except when S is essentially (i.e., with very few exceptions, an immediate upper bound for which is $a-2$) a sequence of terms from some smaller nontrivial subgroup translate $\alpha + H_a$ of G with the existence result then holding modulo H_a. But under these restrictive conditions it follows from a applications of EGZ, with the appropriate terms translated to be considered elements of the corresponding subgroup H_a of index a, that any subsequence of S with length $m + \frac{m}{a} - 1 + (a-2)$ must contain an m-term zero-sum subsequence. Since $m + \frac{m}{a} + a - 3 \leq \lfloor \frac{3}{2} m \rfloor - 1$ (some basic calculus shows that $\frac{m}{a} + a$ is maximized, as a function of a, for the boundary values of a), this is often a sufficiently significant improvement over EGZ.

The assertion of Theorem 3.1 is more natural than it might at first seem. From Kneser's Theorem we know that if a given n-set partition $A = A_1, \ldots, A_n$ fails to satisfy the Cauchy-Davenport bound, then its sumset must be H_a-periodic with nontrivial period. If H_a is maximal, then modulo H_a the sumset of A is aperiodic. Thus if in some set A_i of A there are two elements from the same H_a-coset, and if there is some set A_j of A that does not contain an element from this coset, then we know that the bound given by Kneser's Theorem on the cardinality of the sumset of A modulo H_a will increase when we move one of the two elements from A_i to A_j. It is natural to think the sumset (not modulo H_a) will likewise increase, and thus repeating this moving procedure we should be able to attain the Cauchy-Davenport bound unless a small number of H_a-cosets contain most of the terms of S. Theorem 3.2 asserts that this is essentially true. The Cauchy-Davenport bound asserts that each term of S partitioned by the n-set partition A—minus one term per A_j with $j \geq 2$

that instead transfers all elements in the sumset $\sum_{i=1}^{j-1} A_i$ to the sumset $\sum_{i=1}^{j} A_i$—contributes at least one element to the sumset $\sum_{i=1}^{n} A_i$. Theorem 1 says that there exists an n-set partition A of S with sumset H_a-periodic such that, if we equate all terms of S that both belong to the same H_a-nonexception and are also contained in the same set A_i, then each resulting term of S, minus one resulting term per A_j with $j \geq 2$, contributes at least one H_a-coset to the sumset $\sum_{i=1}^{n} w_i A_i$. Hence only terms of S that belong to an H_a-nonexception will contribute to any deficit between the Cauchy-Davenport bound and the actual cardinality of $\sum_{i=1}^{n} w_i A_i$.

Finally, observe that if Theorem 3.1 does not hold with H_a trivial (i.e., the Cauchy-Davenport bound does not hold), then $(e + (N-1)n + 1)|H_a| \leq |S'| - n$ follows, where $N = N(A', H_a)$ and $e = E(A', H_a)$, implying $Nn|H_a| - |S'| \leq n(|H_a| - 1) - |H_a| - e|H_a|$. Hence

$$\rho < (n - 1 - e)(|H_a| - 1) \leq (n - 1)(|H_a| - 1),$$

where $\rho = Nn|H_a| - |S'| + e$ is the number of H_a-holes contained among the sets $A'_j \cap \bigcap_{i=1}^{n} (A'_i + H_a)$, $j = 1, \ldots, n$. This mirrors a similar bound on the number of holes obtained from Kneser's Theorem discussed earlier in Chapter 1.

3.2 Composite Cauchy-Davenport

The proof of Theorem 3.1 is somewhat constructive in nature and will be presented as a series of lemmas. In what follows, n is a fixed positive integer, S' is a subsequence of a finite sequence S of elements from an abelian group G, and $A = A_1, \ldots, A_n$ is an n-set partition of S' that by contradiction does not satisfy Theorem 3.1. Note that the conditions on the w_i imply that $|w_i A_i| = |A_i|$, that $w_i x \in w_i A_i$ if and only if $x \in A_i$, that $\phi_a(w_i A_i) = w_i \phi_a(A_i)$, and that $\phi_a(x) \in \phi_a(A_i)$ is doubled if and only if $\phi_a(w_i x) \in \phi_a(w_i A_i)$ is doubled, where

$A_i \subseteq G$ and $H_a \leq G$—all of which will be used implicitly in the proof. The proof makes use of an n-set partition that satisfies a list of iterated extremal conditions that are rigorously described by the following two lengthy definitions.

Definition 1. For a fixed integer $r \leq n$, an *r-maximal partition set* of S, denoted by Λ_r, is the set consisting of those ordered n-set partitions, of a subsequence of S with length $|S'|$, that can be constructed recursively by the method described below. For the sake of clarity, in addition to Λ_i, we introduce four associated entities denoted by $\mathcal{F}_i$, Υ_i, $\mathcal{G}_i$ and $H_{k_{i+1}}$, for $i = 0, \ldots, r-1$.

Λ_0 consists of all ordered n-set partitions, $(Z_1, \ldots, Z_n)$, of a subsequence of S with length $|S'|$, such that $\sum\limits_{i=1}^{n} w_i A_i \subseteq \sum\limits_{i=1}^{n} w_i Z_i$ and $a_i \in Z_i$ for $i \leq n$.

$\mathcal{F}_0 = (A_1^0, \ldots, A_n^0)$ is a fixed element of Λ_0.

Υ_0 is the subset of Λ_0 consisting of all ordered n-set partitions, $(Z_1, \ldots, Z_n)$, for which $|\sum\limits_{i=1}^{n} w_i Z_i|$ is maximized.

$\mathcal{G}_0 = B_1^0, \ldots, B_n^0$ is a fixed element of Υ_0. Different choices of $\mathcal{G}_0$ may result in different Λ_r's.

H_{k_1} is the maximal subgroup for which the sumset $\sum\limits_{i=1}^{n} w_i B_i^0$ is periodic.

Suppose Λ_{j-1}, $\mathcal{F}_{j-1} = (A_1^{j-1}, \ldots, A_n^{j-1})$, Υ_{j-1}, $\mathcal{G}_{j-1} = (B_1^{j-1}, \ldots, B_n^{j-1})$, and H_{k_j} have been constructed; then we proceed as follows:

$$\Lambda_j = \left\{ (Z_1, \ldots, Z_n) \in \Upsilon_{j-1} \mid \sum_{i=1}^{n} |\phi_{k_j}(w_i Z_i)| \text{ is maximum subject to } \sum_{i=j}^{n} w_i Z_i = \sum_{i=j}^{n} w_i B_i^{j-1} \right\}.$$

$\mathcal{F}_j = (A_1^j, \ldots, A_n^j)$ is a fixed element of Λ_j. Different choices of $\mathcal{F}_j$ may result in different Λ_r's.

$$\Upsilon_j = \Big\{ (Z_1, \ldots, Z_n) \in \Lambda_j \mid \left| \sum_{i=j+1}^{n} w_i Z_i \right|$$
$$\text{is maximum subject to } \phi_{k_j}(w_i A_i^j) \subseteq \phi_{k_j}(w_i Z_i) \text{ for all } i \Big\}.$$

$\mathcal{G}_j = (B_1^j, \ldots, B_n^j)$ is a fixed element of Υ_j. Different choices of $\mathcal{G}_j$ may result in different Λ_r's.

$H_{k_{j+1}}$ is the maximal subgroup for which the sumset $\sum_{i=j+1}^{n} w_i B_i^j$ is periodic.

Definition 2. For a fixed integer ρ, where $0 \leq \rho \leq n-2$, a *ρ-factor form* of S is an ordered n-set partition of a subsequence of S with length $|S'|$, say $F_\rho = (Z_1, \ldots, Z_n) = (X_1, \ldots, X_\rho, Y_{\rho+1}, \ldots, Y_n)$, which satisfies:

(I) if $1 \leq j \leq \rho + 1$, then $\sum_{i=j}^{n} w_i Z_i$ is maximally H_{k_j}-periodic with H_{k_j} a proper, nontrivial subgroup—for simplicity we will sometimes denote $k_{\rho+1}$ by k;

(II) $|\sum_{i=1}^{n} \phi_k(w_i Z_i)| \geq |\sum_{i=\rho+1}^{n} \phi_k(w_i Y_i)| + \sum_{i=1}^{\rho} |\phi_k(w_i X_i)| - (\rho + 1) + 1;$

(III) $|\sum_{i=\rho+1}^{n} w_i Y_i| < \sum_{i=\rho+1}^{n} |Y_i| - (n-\rho) + 1$;

(IV) each term X_i, for all $i \leq \rho$, contains an element mapped to an H_{k_i}-exception;

(V) there exists a $(\rho+1)$-maximal partition set $\Lambda_{\rho+1}$ of S, such that $F_\rho \in \Lambda_{\rho+1}$.

If F_ρ is an ordered n-set partition, of a subsequence of S with length $|S'|$, that satisfies (I), (IV) and (V), then it is called a *weak ρ-factor form*. It should be noted that (I) easily implies that $H_{k_{j+1}} \leq H_{k_j}$ for $j = 1, \ldots, \rho$. Also, due to the maximality of Λ_j, it follows that the definition of Υ_j is unchanged by changing the inclusion $\phi_{k_j}(w_i A_i^j) \subseteq \phi_{k_j}(w_i Z_i)$ to an equality. A similar statement concerning equality versus inclusion, due to the maximality of Υ_{j-1}, holds concerning the equality $\sum_{i=j}^{n} w_i Z_j = \sum_{i=j}^{n} w_i B_i^{j-1}$ in the definition of Λ_j. Consequently, the mod H_{k_j} structure, with $j \leq r$, of an element of Υ_r is fixed, while the structure $\sum_{i=j}^{n} w_i Z_i$, with $j \leq r$, is fixed for an element $(Z_1, \ldots, Z_n) \in \Lambda_r$.

Lemma 1. *If S has a weak ρ-factor form, $F_\rho = (Z_1, \ldots, Z_n) = (X_1, \ldots, X_\rho, Y_{\rho+1}, \ldots, Y_n)$ such that for some index q there exists $x \in Z_q$, where $\phi_k(x)$ is an H_k-doubled H_k-exception, then $q \geq \rho + 1$ and $\sum_{\substack{i=\rho+1 \\ i \neq q}}^{n} w_i Z_i + w_q(Z_q \setminus \{x\})$ is not H_k-periodic.*

Proof. Since $\phi_k(x)$ is doubled, it follows that there are at least two elements of Z_q mapped by ϕ_k to $\phi_k(x)$. Hence w.l.o.g. we may assume $x \neq a_q$. Let $l = \min\{\rho + 1, q\}$. From the definition of an exception, it follows that there must exist a term D of F_ρ such that $\phi_k(x) \notin \phi_k(D)$. Suppose $\sum_{\substack{i=l \\ i \neq q}}^{n} w_i Z_i + w_q(Z_q \setminus \{x\})$ is still H_k-periodic. Then by (I) and the definition of a doubled element, it follows that $\sum_{i=l}^{n} w_i Z_i = \sum_{\substack{i=l \\ i \neq q}}^{n} w_i Z_i + w_q(Z_q \setminus \{x\})$. Hence, since $x \neq a_q$, it follows that if we remove x from Z_q and place it in D, we obtain a new ordered n-set partition, say $F'_\rho = (Z'_1, \ldots, Z'_n)$, such that

$$\sum_{i=j}^{n} w_i Z_i \subseteq \sum_{i=j}^{n} w_i Z'_i \text{ for every } j \leq \rho + 1. \tag{3.1}$$

Since $\phi_k(x)$ is doubled and since $\phi_k(x) \notin \phi_k(D)$, it follows that

$$\sum_{i=1}^{n} |\phi_k(w_i Z'_i)| > \sum_{i=1}^{n} |\phi_k(w_i Z_i)|. \tag{3.2}$$

By (V) and the definition of an r-maximal partition set, it follows that $F_\rho \in \Lambda_{\rho+1} \subseteq \Upsilon_l$, for every $l \leq \rho$. Hence in view of (3.1), since $\phi_k(x)$ is a doubled exception in F_ρ, since $H_k \leq H_{k_i}$ for all i, and since $F_\rho \in \Lambda_{\rho+1}$, it follows by a simple inductive argument passing from j to $j+1$, where $j = 0, \ldots, \rho$, that $F'_\rho \in \Upsilon_j$ and $F'_\rho \in \Lambda_{j+1}$ (simply note that the moving procedure can only increase any of the maximized quantities and can only further increase (by inclusion) any of the fixed ‘subject to’ conditions, which due to their maximality (see the remarks before Lemma 1) must then still remain fixed). Consequently, $F'_\rho \in \Lambda_{\rho+1}$. Since $F_\rho \in \Lambda_{\rho+1}$ and since $F'_\rho \in \Lambda_{\rho+1}$, from the definition of an r-maximal partition set it follows that $\sum_{i=1}^{n} |\phi_k(w_i Z'_i)| = \sum_{i=1}^{n} |\phi_k(w_i Z_i)|$, contradicting (3.2). So $\sum_{\substack{i=l \\ i \neq q}}^{n} w_i Z_i + w_q(Z_q \setminus \{x\})$ is not H_k-periodic. Hence it follows from (I) that $q \geq \rho + 1$, whence $l = \rho + 1$, completing the proof of Lemma 1. $\square$

Lemma 2. *If S has a weak ρ-factor form $F_\rho = (Z_1, \ldots, Z_n) = (X_1, \ldots, X_\rho, Y_{\rho+1}, \ldots, Y_n)$ that satisfies (III), and for which for some index q there exists $x \in Z_q$, where $\phi_k(x)$ is an H_k-doubled H_k-exception, then $|\sum_{i \neq q} w_i Y_i| < \sum_{i \neq q} |Y_i| - (n - \rho - 1) + 1$.*

Proof. From Lemma 1, Proposition 1.3 and (I), it follows that $q \geq \rho + 1$ and

$$|\sum_{i \neq q} w_i Y_i + w_q Y_q| \geq |\sum_{i \neq q} w_i Y_i| + |Y_q| - 1. \tag{3.3}$$

If the conclusion of the lemma is false, then (3.3) implies $|\sum_{i=\rho+1}^{n} w_i Y_i| \geq \sum_{i=\rho+1}^{n} |Y_i| - (n-\rho) + 1$, contradicting (III). □

Lemma 3. *If S has a weak ρ-factor form $F_\rho = (Z_1, \ldots, Z_n) = (X_1, \ldots, X_\rho, Y_{\rho+1}, \ldots, Y_n)$, which satisfies (II), then F_ρ is a ρ-factor form.*

Proof. Note that we need only show that (III) holds. From Lemma 1 it follows that there cannot exist a term X_r of F_ρ, where $r \leq \rho$, such that $\phi_k(X_r)$ contains an H_k-doubled H_k-exception. Hence, since $H_k \leq H_{k_j}$ for all j, then from (IV) it follows that each term X_r, and hence each term $w_r X_r$, with $r \leq \rho$, contains a unique element from some H_k-coset. Thus, if (III) does not hold, then (by counting holes—the argument is the same as the one used with Kneser's Theorem to show the CDT bound holds provided the number of holes ρ is at least $(n-1)(|H_a| - 1)$) it follows from (I) and (II) that Theorem 3.1 holds with the trivial group, contrary to assumption. □

Lemma 4. *If S has a weak ρ-factor form $F_\rho = (Z_1, \ldots, Z_n) = (X_1, \ldots, X_\rho, Y_{\rho+1}, \ldots, Y_n)$, then F_ρ is a ρ-factor form.*

Proof. From Lemma 3 we see it suffices to show that (II) holds. To this end, note that it suffices to show

$$|\sum_{i=j+1}^{n} \phi_k(w_i Z_i) + \phi_k(w_j X_j)| \geq |\sum_{i=j+1}^{n} \phi_k(w_i Z_i)| + |\phi_k(w_j X_j)| - 1, \text{ for all } j \leq \rho. \tag{3.4}$$

Let $j \leq \rho$ be arbitrary. From (IV) it follows that there exists $x \in X_j$ such that $\phi_{k_j}(x)$ is an H_{k_j}-exception. Suppose $\phi_{k_j}(x)$ is H_{k_j}-doubled. Then w.l.o.g. $x \neq a_j$. If $\sum_{i=j+1}^{n} \phi_k(w_i Z_i) + \phi_k(w_j X_j) \neq \sum_{i=j+1}^{n} \phi_k(w_i Z_i) + (\phi_k(w_j X_j) \setminus \{\phi_k(w_j x)\})$, then (3.4) follows from Proposition 1.3. Otherwise, it follows that we can remove x from X_j and place it in some term D with $\phi_{k_j}(x) \notin D$, yielding a contradiction by the arguments used in the proof of Lemma 1. So we may assume $\phi_{k_j}(x)$ is not H_{k_j}-doubled. Hence it follows that $\phi_k(w_j x)$ is the only element from its H_{k_j}/H_k-coset in $\phi_k(w_j X_j)$. From (I) it follows that $\sum_{i=j+1}^{n} \phi_k(w_i Z_i) + \phi_k(w_j X_j)$ is maximally H_{k_j}/H_k-periodic. Hence (3.4) follows from Kneser's Theorem and the conclusions of the last two sentences. □

Lemma 5. *If S has a ρ-factor form $F_\rho = (Z_1, \ldots, Z_n) = (X_1, \ldots, X_\rho, Y_{\rho+1}, \ldots, Y_n)$, then S has a $(\rho+1)$-factor form.*

Proof. From Lemma 4 it suffices to show that S has a weak $(\rho+1)$-factor form. Suppose there does not exist an H_k-doubled H_k-exception. Hence from (II), (I) and Kneser's Theorem it follows, since Theorem 3.1 does not hold with H_k, that there exists $x \in S \setminus S''$ and a term D of F_ρ such that $\phi_k(x) \notin \phi_k(D)$, where S'' is the subsequence of S that F_ρ partitions. In view of (III) it follows that there exists an index j, with $\rho + 1 \leq j < n$, such that $|\sum_{i=j}^{n} w_i Z_i| < |\sum_{i=j+1}^{n} w_i Z_i| + |Z_j| - 1$. Hence from Kneser's theorem it follows that $\sum_{i=j}^{n} w_j Z_j$ is maximally H-periodic with nontrivial period, and that there cannot be an element in $w_j Z_j$ that is the unique element from its H-coset. Consequently, since $H \leq H_k$ follows from (I), it follows that there cannot be an element in $w_j Z_j$ that is the unique element from its H_k-coset. Hence, since there are no H_k-doubled H_k-exceptions, it follows that all elements of $\phi_k(Z_j)$ are H_k-nonexceptions and that $|\phi_k^{-1}(\beta) \cap w_j Z_j| \geq 2$ for each H_k-nonexception $\beta \in G/H_k$. Since $|\sum_{i=j}^{n} w_i Z_i| < |\sum_{i=j+1}^{n} w_i Z_i| + |Z_j| - 1$, it follows in view of Proposition 1.3

that $\sum_{i=j}^{n} w_i Z_i = \sum_{i=j+1}^{n} w_i Z_i + w_j(Z_j \setminus \{y\})$ for $y \in Z_j$. Hence, since $|\phi_k^{-1}(\beta) \cap w_j Z_j| \geq 2$ for each $\beta \in \phi_a(w_j Z_j)$, it follows that we can choose $y \in Z_j$ such that $a_j \neq y$, such that $|\phi_k(w_j A_j)| = |\phi_k(w_j(A_j \setminus \{y\}))|$, and such that $\sum_{i=j}^{n} w_i Z_i = \sum_{i=j+1}^{n} w_i Z_i + w_j(Z_j \setminus \{y\})$. Hence it follows that we can remove y from the set partition F_ρ and place x in D to obtain a new ordered n-set partition $F'_\rho = (Z'_1, \ldots, Z'_n)$ of the sequence $S''' = (S'' \setminus \{y\}) \cup \{x\}$, yielding a contradiction to the maximality of $\sum_{i=1}^{n} |\phi_k(w_j Z_j)|$ for F_ρ by the arguments used in the proof of Lemma 1. So we may assume there exists an H_k-doubled H_k-exception.

However, by Lemma 1 it follows that no term Z_i with $i \leq \rho$ can contain an element mapped to an H_k-doubled H_k-exception. Hence, there exists a term Y_q, such that $\phi_k(Y_q)$ contains an H_k-doubled H_k-exception. Since the order of terms Y_i for $i > \rho$ is inconsequential, we may assume w.l.o.g. that $q = \rho + 1$. Define $\Upsilon_{\rho+1}$ to be

$$\Upsilon_{\rho+1} \stackrel{def}{=} \Big\{(Z'_1, \ldots, Z'_n) \in \Lambda_{\rho+1} \mid \left|\sum_{i=\rho+2}^{n} w_i Z'_i\right| \text{ is maximum subject to } \phi_k(w_i Z_i) \subseteq \phi_k(w_i Z'_i) \text{ for all } i\Big\},$$

and let

$$F'_\rho = (Z'_1, \ldots, Z'_n) = (X'_1, \ldots, X'_\rho, Y'_{\rho+1}, \ldots, Y'_n),$$

be an arbitrarily chosen element of $\Upsilon_{\rho+1}$. Note since (V) implies $F_\rho \in \Lambda_{\rho+1}$, and since $F'_\rho \in \Lambda_{\rho+1}$, then it follows, in view of the remarks before Lemma 1, that (I), (IV) and (V) hold for F'_ρ. Hence by Lemma 4 it follows that F'_ρ is a ρ-factor form.

Next we will show the inequality

$$|\sum_{i=\rho+2}^{n} w_i Y'_i| < \sum_{i=\rho+2}^{n} |Y'_i| - (n - \rho - 1) + 1. \tag{3.5}$$

From the definitions of $\Lambda_{\rho+1}$ and $\Upsilon_{\rho+1}$ it follows (as remarked before Lemma 1) that $\phi_k(w_iZ_i) = \phi_k(w_iZ_i')$ for all i. Hence, since $\phi_k(Z_{\rho+1})$ contained an H_k-exception, it follows that $\phi_k(Z'_{\rho+1})$ still contains an H_k-exception, say $\phi_k(x)$, where $x \in Z'_{\rho+1}$. If $\phi_k(x)$ is H_k-doubled, then Lemma 2 implies (3.5). Hence we may assume that $w_{\rho+1}x$ is the unique element from its H_k-coset in $w_{\rho+1}Z'_{\rho+1}$. Thus from (I) and Kneser's Theorem, it follows that $|\sum_{i=\rho+1}^{n} w_iY_i'| \geq |\sum_{i=\rho+2}^{n} w_iY_i'| + |Y'_{\rho+1}| - 1$, whence in view of (III) it follows that (3.5) must hold in this case as well. So (3.5) does hold as desired. Consequently, $\rho+1 \leq n-2$. Furthermore, by Kneser's Theorem it follows from (3.5) that $\sum_{i=\rho+2}^{n} w_iY_i'$ is maximally $H_{k_{\rho+2}}$-periodic, with $H_{k_{\rho+2}}$ a proper, nontrivial subgroup. We can further assume w.l.o.g. that we chose F'_ρ so that $\sum_{i=1}^{n}|\phi_{k_{\rho+2}}(w_iZ_i')|$ is maximum with respect to all $(Z_1'', \ldots, Z_n'') \in \Upsilon_{\rho+1}$ with $\sum_{i=\rho+2}^{n} w_iZ_i'' = \sum_{i=\rho+2}^{n} w_iZ_i'$. Thus the n-set partition $(X_1', \ldots, X_\rho', Z'_{\rho+1}, Y'_{\rho+2}, \ldots, Y_n')$ satisfies all conditions for a weak $(\rho+1)$-factor form with $Z'_{\rho+1} = X'_{\rho+1}$. □

We can now complete the proof of Theorem 3.1.

Proof. Let $A' = (A_1', \ldots, A_n')$ be an n-set partition of a subsequence of S with length $|S'|$ whose weighted sumset $\sum_{i=1}^{n} w_iA_i'$ has maximal cardinality subject to $\sum_{i=1}^{n} w_iA_i \subseteq \sum_{i=1}^{n} w_iA_i'$ and $a_i \in A_i'$ for $i \leq n$. Since we have assumed Theorem 3.1 does not hold for A with the trivial group (i.e., the CDT bound does not hold), it follows that (III) holds with $\rho = 0$. Hence from Kneser's Theorem it follows that $\sum_{i=1}^{n} w_iA_i'$ is maximally H_{k_1}-periodic with H_{k_1} a nontrivial subgroup. Since Theorem 3.1 does not hold with the group G, it follows that H_{k_1} must also be proper. Thus the set partition A' satisfies (I), (II), (III) and (IV) for $\rho = 0$, and we may assume that A' has been chosen such that $\sum_{i=1}^{n}|\phi_{k_1}(w_iA_i')|$ is maximum over all n-set partitions $(Z_1, \ldots, Z_n)$ of S with $\sum_{i=1}^{n} w_iZ_i = \sum_{i=1}^{n} w_iA_i'$ and $a_i \in A_i'$ for all i. Thus the sequence S has a 0-factor form given by the n-set partition A'. Let γ be the maximum integer for which S has a γ-factor form; it follows that γ exists, since ρ is bounded from above by $n-2$ from the definition of a ρ-factor form. However, it follows from Lemma 5 that S has a $(\gamma+1)$-factor form, contradicting the maximality of γ. $\square$

We conclude with the proof of Theorem 3.2.

Proof. We use induction on $|S|$ with n fixed. Note that (i) holds trivially with $A = P$ for the base case $|S| = n$. Apply Theorem 3.1 to the subsequence S' of S with n-set partition P, and let $A = A_1, \ldots, A_n$ be the resulting set partition and H_a the corresponding subgroup. Since $n \geq \min\{\frac{m}{p} - 1, \frac{|S'|-n+1}{p} - 1\}$, then from Theorem 3.1 we may assume that H_a is a proper, nontrivial subgroup, that $N(A, H_a) = 1$, that $|\sum_{i=1}^{n} w_i A_i| \geq (e+1)|H_a|$, and that

$$e \leq \min\{a-2, \left\lfloor \frac{|S'|-n}{|H_a|} \right\rfloor - 1\}, \tag{3.6}$$

where $e = E(A, H_a)$, since otherwise (i) follows. Thus all but $e \leq \min\{a-2, \left\lfloor \frac{|S'|-n}{|H_a|} \right\rfloor - 1\}$ terms of S are from the same H_a-coset, say $\alpha + H_a$, where $\phi_a(\alpha)$ is the H_a-nonexception, and $|\sum_{i=1}^{n} w_i A_i| \geq (e+1)|H_a|$. Hence we may assume $e > 0$, since otherwise in view of Proposition 2.4 applied to A it follows that (ii) holds with $e = 0$.

Let S_0 be the subsequence of S consisting of all terms from $\alpha + H_a$, let $A' = A'_1, \ldots, A'_n$ where $A'_i = A_i \cap (\alpha + H_a)$, and let S'_0 be the subsequence of S_0 that A' partitions. Note since $N(A, H_a) = 1$, that $|A'_i| > 0$ for all i, and thus A' is an n-set partition of S'_0. Since $(e+1)|H_a| \leq |\sum_{i=1}^{n} w_i A_i| < |S'| - n + 1$, it follows that $|S'| \geq n + (e+1)|H_a|$. Hence, since $N(A, H_a) = 1$, then it follows that $|S'_0| \geq n + (e+1)|H_a| - e \geq n + |H_a|$. Since $e > 0$, it follows that $|S_0| < |S|$. We may also w.l.o.g. assume $\alpha = 0$. Hence we can apply the induction hypothesis to the subsequence S'_0 of S_0 with set partition A' and with $G = H_a$.

If (i) holds for S_0, then since $|S_0'| \geq n + |H_a|$, it follows, in view of $|\sum_{i=1}^{n} w_i A_i| \geq (e+1)|H_a|$, (3.6), and Proposition 2.4, that (ii) holds for S with subgroup H_a. So assume (ii) holds for S_0 with subgroup $H_{ka} \leq H_a$ of index $k = [H_a : H_{ka}]$, with coset $\beta + H_{ka}$, and with n-set partition $B = B_1, \ldots, B_n$ satisfying $\sum_{i=1}^{n} w_i B_i = \beta \sum_{i=1}^{n} w_i + H_{ka}$. In this case, since at most $k-2$ terms of S_0 are not from the coset $\beta + H_{ka}$ (follows by induction hypothesis), and since $|S'| \geq |S_0'| \geq n + |H_a| = n + \frac{m}{a}$, it follows in view of (3.6) that there are at most

$$k - 2 + \min\{a-2, \frac{|S'|-n}{|H_a|} - 1\} = \min\{k-2+a-2, k-2+\frac{a(|S'|-n)}{m} - 1\} \leq$$
$$\min\left\{ka - 4, \frac{ka(|S'|-n)}{m} - 1 + \left(k - 2 - (k-1)\frac{a(|S'|-n)}{m}\right)\right\} <$$
$$\min\{ka-2, \frac{(|S'|-n)}{|H_{ka}|} - 1\},$$

terms of S not from the coset $\beta + H_{ka}$. Also,

$$|\sum_{i=1}^{n} w_i A_i| \geq (e+1)|H_a| = k(e+1)|H_{ka}| \geq (k-1+e)|H_{ka}| \geq (e'+1)|H_{ka}|,$$

where e' is the number of terms of S not from $\beta + H_{ka}$. Hence (ii) holds for S with subgroup H_{ka}, coset $\beta + H_{ka}$, and set partitions $A = A_1, \ldots, A_n$ and $B = B_1, \ldots, B_n$. $\square$

Chapter 4

Mostly Monochromatic Zero-Sums

4.1 Discussion

In this chapter, we present a result that shows that in a mostly two-color sequence of terms from an abelian group of order m with length $2m-1$, there is a mostly monochromatic m-term zero-sum subsequence. More precisely, if we introduce the notation of $X \cap S$, where X is a set and S a sequence, to denote the subsequence of S consisting of terms equal to an element from X, then the main result of this chapter is the following.

Theorem 4.1. *Let S be a sequence of elements from a finite abelian group G of order m, and suppose there exist a, $b \in G$ such that $|(G \setminus \{a, b\}) \cap S| \leq \lfloor \frac{m}{2} \rfloor$. If $|S| \geq 2m-1$, then there exists an m-term zero-sum subsequence S' of S with $|\{a\} \cap S'| \geq \lfloor \frac{m}{2} \rfloor$ or $|\{b\} \cap S'| \geq \lfloor \frac{m}{2} \rfloor$.*

The sequence $S = (\underbrace{0, \ldots, 0}_{m-1}, \underbrace{1, \ldots, 1}_{m-1}, \lceil \frac{m}{2} \rceil)$ with $G = \mathbb{Z}/m\mathbb{Z}$ shows that the lower bound $\lfloor \frac{m}{2} \rfloor$ in Theorem 4.1 is also tight, although the theorem likely remains true under a less restrictive condition than $|(G \setminus \{a, b\}) \cap S| \leq \lfloor \frac{m}{2} \rfloor$.

Theorem 4.1 can sometimes be used to handle the stubborn cases when a sufficiently compressed n-set partition does not exist. The results of Section 2.2 show that when such a set partition does not exist, then the majority of terms are equal to one of a small selection of elements. In one of the most basic cases, the majority of terms will be equal to one of two fixed elements, in which case the hypotheses of Theorem 4.1 are readily attained. The conclusion of Theorem 4.1 then not only gives an m-term zero-sum, but also one with a considerable amount of structure, which may be exploitable.

4.2 A Refinement of EGZ

Before beginning the proof, we first introduce the notation of $\overline{\alpha}$, for $\alpha \in \mathbb{Z}/m\mathbb{Z}$, to denote the least positive integer representative of α. In what follows, we implicitly use the fact that $\overline{\alpha + \beta}$ either equals $\overline{\alpha}+\overline{\beta}$ or $\overline{\alpha}+\overline{\beta}-m$. The proof of Theorem 4.1, which we begin below, follows a method introduced by Gao and Hamidoune [21].

Proof. Let $|\{a\} \cap S| = n_0$, let $|\{b\} \cap S| = n_1$, and let $t = |S| - n_0 - n_1$. We may w.l.o.g. assume $|S| = 2m - 1$, $n_1 \leq n_0 \leq m - 1$, and $a = 0$. Hence, since by hypothesis

$$t \leq \left\lfloor \frac{m}{2} \right\rfloor, \tag{4.1}$$

and since there can be no monochromatic m-term zero-sum, else the proof is complete, then it follows that

$$\left\lceil \frac{m}{2} \right\rceil \leq m - t \leq n_1 \leq n_0 \leq m - 1, \tag{4.2}$$

and, in view of the pigeonhole principle, that

$$m - \left\lfloor \frac{t+1}{2} \right\rfloor \leq n_0. \tag{4.3}$$

Let c be the order of b. Suppose first that $c < m$. Let l be the least integer such that $\lfloor \frac{t+1}{2} \rfloor \leq l$ and $c|l$. Observe $l \leq \lfloor \frac{t+1}{2} \rfloor + c - 1$. Hence, if $c < \frac{m}{3}$, then in view of (4.1) it follows that $l \leq \lfloor \frac{m+2}{4} \rfloor + \frac{m}{4} - 1 \leq \lceil \frac{m}{2} \rceil$. On the other hand, if $c \geq \frac{m}{3}$, then from (4.1) it follows that $\lfloor \frac{t+1}{2} \rfloor \le c$, whence $l = c \le \lceil \frac{m}{2} \rceil$. Hence, in view of (4.2) and (4.3), it follows in both cases that the proof is complete by selecting l terms equal to b and $m - l$ terms equal to 0. So we may assume that $c = m$, whence G is cyclic and w.l.o.g. $b = 1$.

Let $W = w_1, w_2, \ldots, w_l$ be a subsequence of the terms of S not equal to 0 or 1, and let $\sum_{i=1}^{l} w_i = w$. Observe that the m-term sequence

$$(\underbrace{0, \ldots, 0}_{\overline{w}-l}, \underbrace{1, \ldots, 1}_{m-\overline{w}}, w_1, \ldots, w_l)$$

is zero-sum provided $\overline{w} \ge l$. Hence, in view of (4.2), it follows that if $\overline{w} \ge \lfloor \frac{m}{2} \rfloor + l$, then

$$\overline{w} \ge n_0 + l + 1, \tag{4.4}$$

and if $l \le \overline{w} \le \lceil \frac{m}{2} \rceil$, then

$$\overline{w} \le m - n_1 - 1, \tag{4.5}$$

else the proof is complete.

Let $Y = y_1, \ldots, y_{r_y}$ be the subsequence of S consisting of terms y_i such that $1 < \overline{y_i} \le \frac{m}{2}$, and let $Z = z_1, \ldots, z_{r_z}$ be the subsequence of S consisting of terms z_i such that $\frac{m}{2} < \overline{z_i} \le m - 1$. Applying (4.4) with $W = \{z_i\}$, it follows that $\overline{z_i} \ge n_0 + 2$ for all i. Hence, since $\frac{m}{2} < \overline{z_i} \le m - 1$, then in view of (4.1), (4.4) applied to $W = z_1, \ldots, z_{l-1}$, and (4.3), it follows from an easy inductive argument passing from $l - 1$ to l that $\lfloor \frac{m}{2} \rfloor + l \le \overline{\sum_{i=1}^{l} z_i}$ for all $l \in \{1, \ldots, r_z\}$. Hence, since $\frac{m}{2} < \overline{z_i} \le m - 1$, it follows that $\overline{\sum_{i=1}^{l} z_i} \le m - l$. Consequently from (4.4) applied with $W = Z$, it follows that

$$r_z \le \frac{m - n_0 - 1}{2}. \tag{4.6}$$

Let $Y' = y'_1, \ldots, y'_l$ be a subsequence of Y with length l. We next show by induction on l, passing from $l-1$ to l, that

$$\overline{\sum_{i=1}^{l} y'_i} \leq \left\lfloor \frac{m}{2} \right\rfloor + l - 1, \tag{4.7}$$

for all $l \in \{1, \ldots, r_y\}$. The case $l = 1$ follows from the definition of Y. Since $2m - 1 = n_0 + n_1 + t$, then applying (4.5) with $W = \{y_i\}$, it follows that $\overline{y'_i} \leq t - m + n_0$ for all i. Hence by induction hypothesis it follows that

$$n_0 - \left\lceil \frac{m}{2} \right\rceil + l - 2 + t \geq \overline{\sum_{i=1}^{l} y'_i}. \tag{4.8}$$

If (4.7) does not hold, then applying (4.4) with $W = Y'$, it follows that $\overline{\sum_{i=1}^{l} y'_i} \geq n_0 + l + 1$. Hence from (4.8) it follows that $t \geq \lceil \frac{m}{2} \rceil + 3$, contradicting (4.1). So we may assume that (4.7) holds.

We proceed to show that

$$\overline{\sum_{i=1}^{l} y'_i} = \sum_{i=1}^{l} \overline{y'_i}. \tag{4.9}$$

Since $\overline{y'_i} \leq \frac{m}{2}$, it follows that (4.9) holds for $l = 1$ and $l = 2$. Assume inductively that (4.9) holds up to $(l-1)$, where $l \geq 3$. Letting $j, j' \in \{1, \ldots, l\}$ be arbitrary distinct indices, it follows in view of (4.7) and the induction hypothesis that $\sum_{\substack{i=1 \\ i \neq j}}^{l} \overline{y'_i} = \overline{\sum_{\substack{i=1 \\ i \neq j}}^{l} y'_i} \leq \lfloor \frac{m}{2} \rfloor + l - 2$. Hence, using the estimate $\overline{y'_i} \geq 2$ for $i \neq j'$, it follows that

$$\overline{y'_{j'}} \leq \left\lfloor \frac{m}{2} \right\rfloor - l + 2, \tag{4.10}$$

for all $j' \in \{1, \ldots, l\}$. But then from (4.10), induction hypothesis and (4.7), it follows that

$$\sum_{i=1}^{l} \overline{y'_i} = \overline{y'_l} + \sum_{i=1}^{l-1} \overline{y'_i} = \overline{y'_l} + \overline{\sum_{i=1}^{l-1} y'_i} \leq \left\lfloor \frac{m}{2} \right\rfloor - l + 2 + \left\lfloor \frac{m}{2} \right\rfloor + l - 2 = 2 \left\lfloor \frac{m}{2} \right\rfloor \leq m,$$

from which (4.9) immediately follows.

In view of (4.6) and (4.3), it follows that

$$r_y \geq \frac{3t+1}{4}. \tag{4.11}$$

Let l be the maximal integer for which there exists a subsequence $Y' = y'_1, \ldots, y'_l$ of Y satisfying $\sum_{i=1}^{l} \overline{y'_i} \leq \lceil \frac{m}{2} \rceil$. Note, since (4.11) implies $r_y > 0$, and since $\overline{y_i} \leq \frac{m}{2}$, that such a subsequence exists. Hence, since $2m - 1 = n_0 + n_1 + t$, and since $\overline{y_i} \geq 2$, it follows, in view of (4.5) and (4.9), that

$$2l \leq \sum_{i=1}^{l} \overline{y'_i} \leq n_0 + t - m. \tag{4.12}$$

Hence, since $m - n_0 \geq 1$, it follows that $l \leq \frac{t-1}{2}$. Hence from (4.11) it follows that there are at least $\lceil \frac{t+3}{4} \rceil$ terms of Y not in the maximal subsequence Y'. Furthermore, since $l \geq 1$, it follows that $t \geq 3$. Let $A = a_1, \ldots, a_{\lceil (t+3)/4 \rceil}$ be a subsequence of $Y \setminus Y'$. Define α by $\sum_{i=1}^{l} \overline{y'_i} = n_0 + t - m - \alpha$. From (4.12) it follows that $\alpha \geq 0$. Hence, in view of the maximality of Y', it follows that $\overline{y} \geq \lceil \frac{m}{2} \rceil + m - n_0 - t + 1 + \alpha$ for each $y \in Y \setminus Y'$. Hence by considering lower and upper bounds for $\sum_{a \in A} \overline{a} + \sum_{y' \in Y} \overline{y'}$, it follows, in view of (4.7) and (4.9), that

$$\left\lceil \frac{t+3}{4} \right\rceil \left(\left\lceil \frac{m}{2} \right\rceil + m - n_0 - t + 1 + \alpha \right) + (n_0 + t - m - \alpha) \leq \left\lfloor \frac{m}{2} \right\rfloor + l + \left\lceil \frac{t+3}{4} \right\rceil - 1.$$

Hence, since $\alpha \geq 0$, and since $\lceil \frac{t+3}{2} \rceil - 1 > 0$ follows from $t \geq 3$, then it follows that

$$\left\lceil \frac{t+3}{4} \right\rceil \left(\left\lceil \frac{m}{2} \right\rceil + (m - n_0) - t \right) - \left(\left\lfloor \frac{m}{2} \right\rfloor + (m - n_0) - t \right) \leq l - 1,$$

whence

$$\left(\left\lceil \frac{t+3}{4} \right\rceil - 1 \right) \left(\left\lceil \frac{m}{2} \right\rceil + (m - n_0) - t \right) \leq l - 1.$$

Thus, since $m - n_0 \geq 1$, then it follows in view of (4.1) that if m is odd, or $m - n_0 \geq 2$, or $t < \lfloor \frac{m}{2} \rfloor$, then the above inequality implies $l \geq \frac{t+1}{2}$, a contradiction to $l \leq \frac{t-1}{2}$. Hence, in view of (4.1), we may assume m is even, $t = \frac{m}{2}$, and $n_0 = m - 1$. Hence from (4.6) it follows that $r_y = \frac{m}{2}$. Thus from (4.9) it follows that $y_i = 2$ for all i, whence in view of (4.2) the proof is complete by selecting $\frac{m}{2}$ terms equal to 0 and $\frac{m}{2}$ terms equal to 2. $\square$

Chapter 5

Quasi-Periodic Decompositions and the Kemperman Structure Theorem

5.1 Discussion

We begin by defining a subset B of an abelian group G to be *Cauchy* if B is finite and nonempty, and $|A+B| \geq \min\{|G|, |A|+|B|-1\}$ for every finite, nonempty subset $A \subseteq G$. We proceed with the discussion.

The problems of describing the structure of sets A and B (of an abelian group G) for which $A+B$ is small and of estimating the size of $A+B$ are important in many applications ranging from analysis to zero-sum Ramsey Theory. Finite sets such that $|A+B| \leq |A| + |B| - 1$ are called *critical pairs* and, despite some confusion to the contrary, a complete recursive description of their structure was first given by Kemperman [41] (we refrain from stating the theorem until we have developed further notation). However, the description is somewhat complicated and seemingly unwieldy to use (the full recursive description was spread across two separate theorems, Theorems 3.4 and 5.1, and some remarks at the end of Section 5 [41]). Owing to this fact, several attempts were made to obtain more readily usable theorems related to the Kemperman Structure Theorem (KST) [45] [37] [35]. In

[45], Lev gave a weaker but simpler necessary condition for a pair (A, B) to be critical. In [37] [35], Hamidoune used his isoperimetric method—a sophisticated method, applicable to a wide range of additive problems, that uses global properties to infer results about local structure—to (a) determine the structure of those finite, nonempty subsets $B \subseteq G$ for which $|A + B| \geq \min\{|G| - 1, |A| + |B|\}$ holds for every finite subset $A \subseteq G$ with $|A| \geq 2$, and to (b) give for a fixed Cauchy subset $B \subseteq G$ a recursive description of the structure of those finite, nonempty subsets $A \subseteq G$ such that $|A + B| = |A| + |B| - 1$.

The aim of this chapter is to introduce the geometrically intuitive concept of quasi-periodic decompositions and develop their basic properties in relation to KST. This yields a fuller understanding of KST and gives a way to more effectively use KST. As one consequence, we will give a centralized and (relatively) compact statement of the full recursive version of KST.

To illustrate how, for questions involving critical pairs, these results can often be used as an alternative to the isoperimetric method, we will subsequently in Section 5.3 simplify and generalize the previously mentioned results of Hamidoune [37] [35]. Specifically, we will (a) give a new and simple proof of the description of the structure of those finite, nonempty subsets $B \subseteq G$ for which $|A + B| \geq \min\{|G| - 1, |A| + |B|\}$ holds for every finite subset $A \subseteq G$ with $|A| \geq 2$, and will (b) give for a Cauchy subset $B \subseteq G$ a *nonrecursive* description of the structure of those finite, nonempty subsets $A \subseteq G$ such that $|A + B| = |A| + |B| - 1$. We will accomplish (b) by showing that the recursive description given by Kemperman terminates after one or two iterations, provided one of the two subsets is Cauchy.

5.2 Quasi-Periodic Decompositions and KST

This section contains many comments and observations concerning quasi-periodic decompositions and KST, which while important are also straightforward to verify. Thus we will generally state the simpler observations, attaching to the ends of the corresponding sentences labels of the form (c.x), with $x \in \mathbb{Z}$, for ease of future reference, and will provide proofs and explanations for the more involved statements.

Let G be an abelian group, and let H_a be a nontrivial subgroup. We use $\eta_b(A, B)$ to denote the number of $c \in A + b$ such that $\nu_c(A, B) = 1$ (recall that $\nu_c(A, B)$ is the number of representations of $c = a + b$ with $a \in A$ and $b \in B$). If $A \subseteq G$, then a *quasi-periodic decomposition* of A with *quasi-period* H_a is a partition $A = A_1 \cup A_0$ of A into two disjoint (each possibly empty) subsets such that A_1 is H_a-periodic or empty and A_0 is a subset of an H_a-coset. A set $A \subseteq G$ is *quasi-periodic* if A has a quasi-periodic decomposition $A = A_1 \cup A_0$ with A_1 nonempty. We remark that this definition of quasi-periodic differs from that used in Kemperman's original proof, though his definition inspired the one used here. Given a quasi-periodic decomposition $A_1 \cup A_0$ with quasi-period H_a, we refer to A_1 as the *H_a-periodic part* and refer to A_0 as the *aperiodic part* (although it may be periodic if A is periodic). Such a decomposition is *reduced* if A_0 is not quasi-periodic. Note that if A is finite and has a quasi-periodic decomposition $A_1 \cup A_0$ with quasi-period H, then A has a reduced quasi-periodic decomposition $A'_1 \cup A'_0$ with quasi-period $H' \leq H$ and $A'_0 \subseteq A_0$ (c.1). Additionally, a pair of quasi-periodic decompositions $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ with common quasi-period H_a induce a quasi-periodic decomposition of $A + B = C$ with quasi-period H_a given either by $(C \setminus (A_0 + B_0)) \cup (A_0 + B_0)$, if $\phi_a(A_0 + B_0)$ is a unique expression element in $\phi_a(A) + \phi_a(B)$, and otherwise by $(C \setminus (A_0 + B_0 + H_a)) \cup (A_0 + B_0 + H_a)$ (c.2).

Useful examples of non-quasi-periodic sets include arithmetic progressions with difference d and at most $|\langle d\rangle|-2$ terms (c.3). A *punctured periodic set*, i.e., a set A for which there exists $\alpha \in G\setminus A$ such that $A\cup\{\alpha\}$ is maximally H_a-periodic with nontrivial period, has a reduced quasi-periodic decomposition for each prime order subgroup of H_a (c.4). However, as the following proposition shows, reduced quasi-periodic decompositions are otherwise canonical.

Proposition 5.1. *If $A_1\cup A_0$ and $A_1'\cup A_0'$ are both reduced quasi-periodic decompositions of a subset A of an abelian group G, with A_1 maximally H-periodic and A_1' maximally L-periodic, then either (i) $A_1=A_1'$ and $A_0=A_0'$ or (ii) $H\cap L$ is trivial, $A_0\cap A_0'=\emptyset$, $|H|$ and $|L|$ are prime, and there exists $\alpha\in G\setminus A$ such that $A_0\cup\{\alpha\}$ is an H-coset, $A_0'\cup\{\alpha\}$ is an L-coset, and $A\cup\{\alpha\}$ is $(H+L)$-periodic.*

Proof. To show (i) it suffices to show $A_1=A_1'$. We may assume A_1 and A_1' are nonempty, since if w.l.o.g. $A_1=\emptyset$ and $A_1'\neq\emptyset$, then $A_0=A=A_1'\cup A_0'$ is quasi-periodic, contradicting that $A_1\cup A_0$ is reduced. Note that $H\cap L$ is trivial, since otherwise $(A_0'\cap A_1)\cup(A_0'\cap A_0)=A_0'$ and $(A_0\cap A_1')\cup(A_0\cap A_0')=A_0$ imply either $A_1=A_1'$, completing the proof, or that one of A_0' or A_0 is quasi-periodic with quasi-period $H\cap L$, a contradiction.

Suppose $A_1'\subseteq A_1$. Then each L-coset of A_1' is contained in an $(H+L)$-coset contained in A_1. Hence, since $H\cap L$ is trivial, it follows that there must be an entire L-coset, say $\gamma+L$, contained in one of these $(H+L)$-cosets contained in A_1, such that $\gamma+L$ is not in A_1', since otherwise A_1' will be $(H+L)$-periodic, contradicting the maximality of L. But then A_0' must contain $\gamma+L$, implying $A_0'=\gamma+L$, which contradicts that A_0' is not quasi-periodic. So $A_1'\cap A_0\neq\emptyset$.

By repeating the above argument for A_1, it follows that $A_1\cap A_0'\neq\emptyset$ as well. Now A_0' is contained in an $(H+L)$-coset, and this $(H+L)$-coset decomposes as a union of H-cosets.

Since $A_1 \cap A_0' \neq \emptyset$, one of these H-cosets, say $\gamma + H$, is contained in A_1. Hence, since $H \cap L$ is trivial, it follows that part of $\gamma + H$ is contained in A_1'. Let $\beta + L$ be an L-coset in A_1' that intersects $\gamma + H$. If every H-coset that meets $\beta + L$ is in A_1, then this implies that the entire $(H + L)$-coset, which contains the L-coset in which A_0' is contained, is in A_1. Hence A_0' is periodic, contradicting that A_0' is not quasi-periodic. So there exists an H-coset, say $\gamma' + H$, that meets $\beta + L$, and which is not contained in A_1. Then $\gamma' + H$ must be the H-coset containing A_0, and hence also the unique H-coset that meets $\beta + L$ not in A_1. Thus the entire $(H + L)$-coset containing A_0' is contained in A_1 except for (possibly) elements in $\gamma' + H$. Hence, if $\beta' + L$ is the L-coset containing A_0', then the only elements that can be missing from $\beta' + L$ in A are those in $(\beta' + L) \cap (\gamma' + H)$. Hence, since $H \cap L$ is trivial, and since A_0' is not periodic, it follows that A_0' is obtained from $\beta' + L$ by deleting the single element α in $(\beta' + L) \cap (\gamma' + H)$. The same is true of A_0, and (ii) immediately follows. □

It follows from Proposition 5.1 that the complement of a punctured periodic set, i.e., a set A such that $A \setminus \{\beta\}$ is maximally H_a-periodic with nontrivial period for some $\beta \in A$, is aperiodic, whence it follows that a punctured periodic set is also aperiodic (c.5). Concerning punctured periodic sets, we also have the following proposition.

Proposition 5.2. *If A is a punctured H-periodic subset of an abelian group G with $|H| > 2$, then there is a unique $\alpha \in G$ such that $A \cup \{\alpha\}$ is periodic.*

Proof. Assume to the contrary that there are distinct $\alpha, \beta \in G$ such that $A \cup \{\alpha\}$ is maximally H-periodic, $|H| > 2$, while $A \cup \{\beta\}$ is maximally L-periodic with nontrivial period. In view of (c.4), we obtain (for each prime order subgroup H_a of H) a reduced quasi-periodic decomposition $A = A_1 \cup A_0$ with quasi-period H_a. Likewise for each prime order subgroup H_b of L we obtain a quasi-periodic decomposition $A = B_1 \cup B_0$ with quasi-period H_b. Note α is the missing element from the coset $\alpha + H_a = A_0 \cup \{\alpha\}$ in A, while β is the missing element from the coset $\beta + H_b = B_0 \cup \{\beta\}$ in A. If $B_0 \neq A_0$, then Proposition 5.1 implies that $\alpha = \beta$, a contradiction. So we may assume $A_0 = B_0$.

Suppose $|H_a| > 2$. Then the punctured H_a-coset A_0 contains two elements, the difference of which generates the prime order group H_a. Since $A_0 = B_0$, it follows that these elements are also contained in the same H_b-coset, whence their difference generates H_b as well. Consequently, $H_a = H_b$ follows, implying that $\alpha = \beta$, a contradiction. So we may assume $|H_a| = 2$. The same argument also shows that $|H_b| = 2$.

Let $K = H_a + H_b$. From the previous paragraph, it follows that K is isomorphic to the Klein four group. Since $A \cup \{\alpha\}$ is H_a-periodic with $\beta \notin A \cup \{\alpha\}$, it follows that the other element, say γ, from the same H_a-coset as β, is also not contained in A. Hence the single element from $A_0 = B_0$ is the only element from its K-coset contained in A (as the other elements, α, β and γ, are not). Thus $A \setminus A_0$ is K-periodic. Consequently, since $H_a < K$, it follows that $\phi_a(A \setminus A_0)$ is periodic. Hence, in view of (c.5), it follows that $\phi_a(A \setminus A_0) \cup \phi_a(A_0) = \phi_a(A)$ is aperiodic. However, $\phi_a(A) = \phi_a(A \cup \{\alpha\})$. Thus, since $\phi_a(A) = \phi_a(A \cup \{\alpha\})$ is aperiodic, it follows that $A \cup \{\alpha\}$ is maximally H_a-periodic. Hence, since $A \cup \{\alpha\}$ is maximally H-periodic, it follows that $H_a = H$, which since $|H_a| = 2 < |H|$, is impossible. $\square$

In view of Propositions 5.1 and 5.2, it follows that a set A, which is the complement of a punctured periodic set, either has a unique $\beta \in A$ such that $A \setminus \{\beta\}$ is periodic (or empty), or else there is a unique $\alpha \notin A$ such that $A \cup \{\alpha\}$ is K-periodic, where K is isomorphic to the Klein four group (c.6).

We can now state the structure theorem for critical pairs proved by Kemperman.

Kemperman Structure Theorem I (KST). *Let A and B be finite, nonempty subsets of an abelian group G. Then $|A+B| = |A|+|B|-1$, with either $A+B$ aperiodic or else $\nu_c(A,B) = 1$ for some c, if and only if there exist quasi-periodic decompositions $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ with nonempty aperiodic parts and common quasi-period H_a, such that:*

(i) $\nu_c(\phi_a(A), \phi_a(B)) = 1$, where $c = \phi_a(A_0) + \phi_a(B_0)$,

(ii) $|\phi_a(A) + \phi_a(B)| = |\phi_a(A)| + |\phi_a(B)| - 1$, and

(iii) the pair (A_0, B_0) is of one of the following types (all of which imply $|A_0 + B_0| = |A_0| + |B_0| - 1$):

(I) $|A_0| = 1$ or $|B_0| = 1$;

(II) A_0 and B_0 are arithmetic progressions with common difference d, where the order of d is at least $|A_0| + |B_0| - 1$, and $|A_0| \geq 2$, $|B_0| \geq 2$ (hence, $A_0 + B_0$ is an arithmetic progression with difference d, while $\nu_c(A_0, B_0) = 1$ for exactly two $c \in A_0 + B_0$);

(III) $|A_0| + |B_0| = |H_a| + 1$, and precisely one element g_0 satisfies $\nu_{g_0}(A_0, B_0) = 1$ (hence, B_0 has the form $B_0 = (g_0 - \overline{A_0} \cap (g_1 + H_a)) \cup \{g_0 - g_1\}$, where $g_1 \in A_0$);

(IV) A_0 is aperiodic, B_0 is of the form $B_0 = g_0 - \overline{A_0} \cap (g_1 + H_a)$, with $g_1 \in A_0$ (hence, $A_0 + B_0 = (g_0 + H_a) \setminus \{g_0\}$), and $\nu_c(A_0, B_0) \neq 1$ for all c.

Note that KST(i) and KST(ii) insure that we can apply KST modulo H_a (c.7). Next observe that (II) implies that $|\{c \in A+B \mid \nu_c(A,B) = 1\}| = 2$, that (III) implies $A+B$ is periodic and $|\{c \in A+B \mid \nu_c(A,B) = 1\}| = 1$, and that (IV) implies $|\{c \in A+B \mid \nu_c(A,B) = 1\}| = 0$ (c.8). Hence if $|\{c \in A+B \mid \nu_c(A,B) = 1\}| > 2$, then (A,B) must be of type (I) (c.9). Also if $\nu_c(A,B) = 1$ for $c = a+b$ with $a \in A$ and $b \in B_1$, or if $\eta_b(B,A) \geq 2$ for some $b \in A$, then (A,B) must have type (I) with $|A_0| = 1$ (c.10).

In view of Proposition 5.1, (c.1), (c.3), the characterization of type (IV) given in KST(iii), and a simple counting argument, it follows that the subsets A_0 and B_0 from KST can always be taken to be the respective aperiodic parts of (some) reduced quasi-periodic decompositions of A and B, provided $A + B$ is aperiodic, and, furthermore, assuming A_0 and B_0 have been chosen such, then $A_0 + B_0$ will be non-quasi-periodic, provided $A + B$ is not a punctured periodic set (c.11).

Note that the union of an arithmetic progression, having at least two terms, having difference d, and having at most $|\langle d\rangle| - 2$ terms, and a disjoint nonempty $\langle d\rangle$-periodic set cannot satisfy KST(iii), as in view of Proposition 5.1 and (c.3) it is no longer an arithmetic progression and hence not of type (II), nor is it a set with a single element and hence not of type (I), nor since $|\{c \in A + B \mid \nu_c(A, B) = 1\}| = 2 \notin \{0, 1\}$ of type (III) or (IV). Hence in view of Proposition 5.1 and (c.1) it follows that if (A, B) has type (II), then $A_1 \cup A_0$ and $B_1 \cup B_0$ must be taken to be the unique reduced quasi-periodic decompositions of A and B—unless one of A or B, say A, is a punctured $\langle d\rangle$-periodic set, with $|\langle d\rangle| \geq 3$, in which case $B_1 \cup B_0$ is reduced with $|B_0| = 2$ and $A + B$ is periodic; but note in this case that $B_0 = \{b \in B \mid \eta_b(A, B) > 0\}$ with the difference d of elements in B_0 determining A_0 (c.12). Hence in view of (c.8), and since (A, B) of type (I) implies $|\{c \in A+B \mid \nu_c(A, B) = 1\}| > 0$ so that (A, B) cannot be type (IV), and since (A, B) of type (I) with $A + B$ H_a-periodic with nontrivial period implies that $|\{c \in A + B \mid \nu_c(A, B) = 1\}| \geq |H_a| \geq 2$ and that one of A or B is periodic, so that (A, B) cannot have type (III), nor type (II) with $A+B$ periodic, it follows that the type of a pair (A, B) is unique and depends only on (A, B) and not the choice of quasi-periodic decompositions that satisfy KST (c.13).

If $A + B$ is maximally H_a-periodic with nontrivial period, then from Kneser's Theorem it follows that KST(ii) holds with H_a, and that there are exactly $|H_a| - 1$ holes in A and B. If there does not exist a pair of subsets $A_0 \subseteq A$ and $B_0 \subseteq B$, each contained in an H_a-coset, such that all $|H_a| - 1$ holes in A and B are contained in $(A_0 + H_a) \setminus A_0$ and $(B_0 + H_a) \setminus B_0$, then from Propositions 1.2 and 1.3 it follows that there will not be a unique expression element in $A+B$. Hence if (A,B) has type (III) with $A+B$ maximally H_a-periodic, then it follows from the previous two sentences that there will be quasi-periodic decompositions of A and B that satisfy KST with quasi-period H_a (c.14). The following proposition gives a canonical decomposition for (A,B) of type (I).

Proposition 5.3. *Let A and B be finite subsets of an abelian group G such that $|A+B| = |A|+|B|-1$, and let $A_0 = \{b \in A \mid \eta_b(B,A) > 0\}$, $A_1 = \{b \in A \mid \eta_b(B,A) = 0\}$, $B_0 = \{b \in B \mid \eta_b(A,B) > 0\}$, and $B_1 = \{b \in B \mid \eta_b(A,B) = 0\}$. If (A,B) has type (I), then $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ are a pair of quasi-periodic decompositions that satisfy KST.*

Proof. Since A and B are finite, we may w.l.o.g. assume G is finitely generated. Let $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ be quasi-periodic decompositions that satisfy KST with quasi-period H_a maximal. Since (A,B) has type (I), then w.l.o.g. $|A_0| = 1$. If $|A| = 1$ or $|B| = 1$, then the proof is trivial. So we may assume $|A| > 1$ and $|B| > 1$. If $\eta_b(A,B) = 0$ for all $b \in B_1$, and $\eta_b(B,A) = 0$ for all $b \in A_1$, then the proof is complete. Hence in view of (c.10) we may w.l.o.g. assume $\eta_{b'}(A,B) > 0$ for some $b' \in B_1'$. In view of (c.7), apply KST modulo H_a, and let $\phi_a(A) = \phi_a(A_1') \cup \phi_a(A_0')$ and $\phi_a(B) = \phi_a(B_1') \cup \phi_a(B_0')$, with $A = A_1' \cup A_0'$ and $B = B_1' \cup B_0'$, be corresponding quasi-periodic decompositions that satisfy KST with quasi-period $H_{a'}/H_a$ maximal. Note that $\eta_b(A,B) > 0$ for $b \in B$ implies $\eta_{\phi_a(b)}(\phi_a(A), \phi_a(B)) > 0$. Hence, in view of KST(i), and since $\eta_{b'}(A,B) > 0$ for some $b' \in B_1$, it follows that $\eta_{\phi_a(a_0)}(\phi_a(B), \phi_a(A)) \geq 2$, where $A_0 = \{a_0\}$. Hence from (c.10) it follows that $\phi_a(A)$ must have type (I) with $A_0' = A_0$, implying that $A_1' = A_1$. Thus since $|A| > 1$, it follows that A_1' is $H_{a'}$-periodic and nonempty.

Suppose that $\eta_{\phi_a(b)}(\phi_a(A), \phi_a(B)) = 0$ for all $b \in B_1'$. Hence from KST(i) it follows that $B_0 \subseteq B_0'$. Hence B_1' is $H_{a'}$-periodic. Thus, since $A_0' = A_0 = \{a_0\}$, it follows that

$A = A_1 \cup A_0$ and $B = B_1' \cup B_0'$ are a pair of quasi-periodic decompositions that satisfies KST with quasi-period $H_{a'}$, contradicting the maximality of H_a. So we may assume that $\eta_{\phi_a(b'')}(\phi_a(A), \phi_a(B)) > 0$ for some $b'' \in B_1'$. Hence we can iterate the above arguments indefinitely, yielding an infinite chain of strictly increasing subgroups $H_a < H_{a'} < \ldots$, which is impossible in a finitely generated abelian group. $\square$

We will refer to the pair of quasi-periodic decompositions that satisfy KST with quasi-period H_a maximized as the *Kemperman decompositions* of A and B. Note in view of (c.2) that the decompositions mentioned in Proposition 5.3, (c.12) and (c.14) are those that satisfy KST with H_a maximal, for types (I), (II) and (III), respectively, and that they are each unique (c.15). We proceed to show the following proposition that in view of (c.2) will characterize the Kemperman decomposition for (A, B) of type (IV). Note if (A, B) has type (IV), then KST implies that $A + B$ is a punctured maximally H_a-periodic set, with $|H_a| > 2$, whence Proposition 5.2 shows that H_a is unique for $A + B$.

Proposition 5.4. *Let A and B be finite subsets of an abelian group. If (A, B) has type (IV), $A + B$ is a punctured maximally H_a-periodic set, and $|A + B| = |A| + |B| - 1$, then there exist quasi-periodic decompositions of A and B that satisfy KST with quasi-period H_a.*

Proof. From KST(iii) and Proposition 1.2 it follows that there exists an element $b \notin A$, from the coset containing the aperiodic part of the Kemperman decomposition of A, such that $|A \cup \{b\} + B| = |A \cup \{b\}| + |B| - 1$. Hence, since the inclusion of b increased the cardinality of the sumset by one, it follows that $\eta_b(B, A \cup \{b\}) = 1$, and hence, since (A, B) has (IV), that $(A \cup \{b\}, B)$ has type (III). Hence, let $A \cup \{b\} = A_1 \cup A_0$ and $B = B_1 \cup B_0$ be the Kemperman decompositions with quasi-period H_a. Since $\eta_b(B, A \cup \{b\}) = 1$, and since $(A \cup \{b\}, B)$ has type (III), it follows that $b \in A_0$. Hence, since $A + B$ is aperiodic from (c.5), it follows that $|A_0| > 1$. Thus from the characterizations of sets satisfying (III) and (IV) found in KST(iii), it follows that $A_0 \setminus \{b\}$ and B_0 satisfy type (IV) and hence $A = A_1 \cup (A_0 \setminus \{b\})$ and $B = B_1 \cup B_0$ are a pair of quasi-periodic decompositions that satisfy KST with quasi-period H_a, completing the proof. $\square$

In view of Propositions 5.3 and 5.4, (c.11) and (c.15), it follows, for (A, B) of type (I) or (IV) with $A + B$ aperiodic, that there are two main choices for the quasi-periodic decompositions that satisfy KST. The first being to take reduced quasi-period decompositions of A and B, which from Proposition 5.1 will be unique provided $A + B$ is not a punctured periodic set, and the second being to take the Kemperman decompositions.

In view of Proposition 5.3, (c.10) and KST(iii), it follows that either $\eta_b(A, B) \leq 1$ for all $b \in B$ or $\eta_b(B, A) \leq 1$ for all $b \in A$ (c.16). Hence, if $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ are quasi-periodic decompositions that satisfy KST with quasi-period H_a, and if $A + B = C_1 \cup C_0$ is the corresponding induced quasi-periodic decomposition, then applying (c.16) modulo H_a, it follows from KST(i) that either $A_1 + B = C_1$ or $A + B_1 = C_1$ (c.17). Note too that if $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ are the Kemperman decompositions, then $\eta_b(A, B) = 0$ for all $b \in B_1$ and $\eta_b(B, A) = 0$ for all $b \in A_1$ (c.18).

A recursive description, for all (A, B) with $A+B$ aperiodic or $A+B$ containing a unique expression element, is obtained from KST by repeatedly applying KST modulo the quasi-period H_a. In view of KST(i), it follows that type (IV) can never occur in one of the recursive iterations other than in the initial pair of quasi-periodic decompositions (c.19). If $A + B$ is maximally H_a-periodic, then in view of Kneser's it follows that $\phi_a(A + B)$ is aperiodic and that $|\phi_a(A) + \phi_a(B)| = |\phi_a(A)| + |\phi_a(B)| - 1$. Hence the recursive description given by KST can be used to describe the mod H_a skeletons of A and B. From Kneser's Theorem it follows that A and B must satisfy $|A| + |B| = |A + H_a| + |B + H_a| - |H_a| + 1$, while in view of

Proposition 1.2 and Kneser's Theorem it follows that any pair of subsets $A' \subset A + H_a$ and $B' \subset B + H_a$ with $|A'| + |B'| = |A + H_a| + |B + H_a| - |H_a| + 1$ satisfies $A' + B' = A + B$ and $|A' + B'| = |A'| + |B'| - 1$. Combining the last two sentences we obtain a complete recursive characterization for sets A and B with $A + B$ periodic and $|A + B| = |A| + |B| - 1$. As noted by Kemperman [40], to describe A and B for which $|A + B| = |A| + |B| - 1 - \rho$ with $\rho \geq 1$, we simply use Kneser's Theorem to conclude that $A + B$ is maximally H_a-periodic and that $|\phi_a(A) + \phi_a(B)| = |\phi_a(A)| + |\phi_a(B)| - 1$, and then we use the recursive description given by KST for $A + B$ aperiodic or containing a unique expression element. This gives us the mod H_a skeletons for A and B. To complete the description we simply take $A + H_a$ and $B + H_a$ (well defined since both of these sets depend only on the H_a skeleton) and delete any $|H_a| - 1 - \rho$ total elements from $A + H_a$ and $B + H_a$ collectively. In view of KST(i) and (c.18), it follows that by choosing the Kemperman decompositions at each step of the recursion we are assured that if $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ are the Kemperman decompositions with quasi-period H_a, and if $\phi_a(A) = \phi_a(A_1') \cup \phi_a(A_0')$ and $\phi_a(B) = \phi_a(B_1') \cup \phi_a(B_0')$ with $A = A_1' \cup A_0'$ and $B = B_1' \cup B_0'$ are the Kemperman decompositions modulo H_a, then $A_0 \subseteq A_0'$ and $B_0 \subseteq B_0'$. To put all this in more rigorous summary, we restate the Kemperman Structure Theorem with the described recursive aspects included.

Kemperman Structure Theorem II (with Recursion). *Let A and B be nonempty and finite subsets of an abelian group G. Then $|A + B| = |A| + |B| - 1$, with either $A + B$ aperiodic or else $\nu_c(A, B) = 1$ for some c, if and only if there exist an integer $r \geq 1$, partitions $A = A_r \cup \ldots \cup A_1 \cup A_0$ and $B = B_r \cup \ldots \cup B_1 \cup B_0$ of A and B into disjoint (possibly empty) subsets, and a sequence of subgroups $H_{a_r} > \ldots > H_{a_1} > H_{a_0} = 0$, such that A_0 and B_0 are nonempty, $A_r = B_r = \emptyset$, and for each $l \in \{1, \ldots, r\}$:*

(i) the pair $\phi_{a_{l-1}}(A) = \phi_{a_{l-1}}(A_r \cup \ldots \cup A_l) \cup \phi_{a_{l-1}}(A_{l-1} \cup \ldots \cup A_0)$ and $\phi_{a_{l-1}}(B) =$

$\phi_{a_{l-1}}(B_r\cup\ldots\cup B_l)\cup\phi_{a_{l-1}}(B_{l-1}\cup\ldots\cup B_0)$ are the Kemperman decompositions with common quasi-period $H_{a_l}/H_{a_{l-1}}$,

(ii) $\nu_{c_l}(\phi_{a_l}(A),\phi_{a_l}(B))=1$, where $c_l=\phi_{a_l}(A_{l-1}\cup\ldots\cup A_0)+\phi_{a_l}(B_{l-1}\cup\ldots\cup B_0)$,

(iii) $|\phi_{a_l}(A)+\phi_{a_l}(B)|=|\phi_{a_l}(A)|+|\phi_{a_l}(B)|-1$,

(iv) $\eta_b(\phi_{a_{l-1}}(A),\phi_{a_{l-1}}(B))=\eta_{b'}(\phi_{a_{l-1}}(B),\phi_{a_{l-1}}(A))=0$ for all $b\in\phi_{a_{l-1}}(B_r\cup\ldots\cup B_l)$ and $b'\in\phi_{a_{l-1}}(A_r\cup\ldots\cup A_l)$,

(v) the pair (A_l',B_l'), where $A_l'=\phi_{a_{l-1}}(A_{l-1}\cup\ldots\cup A_0)$ and $B_l'=\phi_{a_{l-1}}(B_{l-1}\cup\ldots\cup B_0)$, is of one of the below types, with type (IV) possible only for $l=1$:

(I) $|A_l'|=1$ or $|B_l'|=1$;

(II) A_l' and B_l' are arithmetic progressions with common difference d, where the order of d is at least $|A_l'|+|B_l'|-1$, and $|A_l'|\geq 2$, $|B_l'|\geq 2$ (hence, $A_l'+B_l'$ is an arithmetic progression with difference d, while $\nu_c(A_l',B_l')=1$ for exactly two $c\in A_l'+B_l'$);

(III) $|A_l'|+|B_l'|=|H_{a_l}/H_{a_{l-1}}|+1$, and precisely one element g_0 satisfies $\nu_{g_0}(A_l',B_l')=1$ (hence, B_l' has the form $B_l'=(g_0-\overline{A_l'}\cap(g_1+(H_{a_l}/H_{a_{l-1}})))\cup\{g_0-g_1\}$, where $g_1\in A_l'$);

(IV) A_l' is aperiodic, B_l' is of the form $B_l'=g_0-\overline{A_l'}\cap(g_1+(H_{a_l}/H_{a_{l-1}}))$, with $g_1\in A_l'$ (hence, $A_l'+B_l'=(g_0+(H_{a_l}/H_{a_{l-1}}))\setminus\{g_0\}$), and $\nu_c(A_l',B_l')\neq 1$ for all c.

Furthermore, $|A+B|<|A|+|B|-1$ or $|A+B|=|A|+|B|-1$ with $A+B$ periodic, if and only if $A+B$ is maximally H_a-periodic with nontrivial period, the pair $(\phi_a(A),\phi_a(B))$ satisfies the conditions from the above paragraph, and $|A+H_a|+|B+H_a|=|A+B|+|H_a|$.

However, in many applications it suffices to deal only with single level quasi-periodic decompositions and use KST without the above recursive aspects included. The following proposition, like Proposition 5.4, gives conditions when a quasi-periodic decomposition of $A+B$ can be realized as the induced decomposition of a pair of decompositions that satisfies the conditions of KST, and hence can be used to pull back a quasi-periodic decomposition from sum to components, an ability that can sometimes be quite useful.

Proposition 5.5. *Let A, B, C be finite subsets of an abelian group G, such that $A+B=C$ and $|A+B|=|A|+|B|-1$. Suppose C is neither periodic nor a punctured periodic set, and let $C=C_1\cup C_0$ be the reduced quasi-periodic decomposition. If C_1 is maximally H_a-periodic, then there exist quasi-periodic decompositions $A=A_1\cup A_0$ and $B=B_1\cup B_0$ that satisfy KST with quasi-period H_a such that $A_0+B_0=C_0$.*

Proof. From Proposition 5.1, (c.2) and (c.11), it follows that there exist reduced quasi-periodic decompositions $A=A_1\cup A_0$ and $B=B_1\cup B_0$ that satisfy KST with quasi-period $H_{a'}\leq H_a$ such that $A_0+B_0=C_0$. Hence C_0 is contained in an $H_{a'}$-coset, and the proof is complete unless C_1 is nonempty. Let A_0' be the maximal subset of A containing A_0 that is contained in an H_a-coset. Define B_0' likewise. Since C_0 is contained in an $H_{a'}$-coset, since C_1 is maximally H_a-periodic, and since $H_{a'}\leq H_a$, it follows that $A_0'+B_0'=A_0+B_0=C_0$. Hence $A_0'=A_0$ and $B_0'=B_0$, since otherwise $|\phi_{a'}(C_0)|=|\phi_{a'}(A_0'+B_0')|>1$, contradicting that C_0 is contained in an $H_{a'}$-coset. Since in view of KST(i) $A_0+B_0=C_0$ is a unique expression element modulo $H_{a'}$, since $H_{a'}\leq H_a$, and since C_1 is H_a-periodic, it follows that $A_0+B_0=C_0$ is a unique expression element modulo H_a. Hence it remains to show that KST(ii) holds with H_a and that A_1 and B_1 are H_a-periodic.

Suppose that $|\phi_a(A)+\phi_a(B)|>|\phi_a(A)|+|\phi_a(B)|-1$. Hence, since C_1 is H_a-periodic, it follows that $|\phi_{a'}(C)|\geq(|\phi_a(A)|+|\phi_a(B)|-1)|H_a/H_{a'}|+1$. However, since $A_0=A_0'$ and $B_0=B_0'$ are each a subset of an $H_{a'}$-coset, it follows from KST(ii) that $|\phi_{a'}(C)|\leq((|\phi_a(A)|-1)|H_a/H_{a'}|+1)+((|\phi_a(B)|-1)|H_a/H_{a'}|+1)-1$, contradicting the bound from the previous sentence. So we may assume that $|\phi_a(A)+\phi_a(B)|\leq|\phi_a(A)|+|\phi_a(B)|-1$.

Suppose that $|\phi_a(A)+\phi_a(B)|<|\phi_a(A)|+|\phi_a(B)|-1$. Hence, since $A_0'=A_0$, $B_0'=B_0$ and $A_0+B_0=C_0$ are each a subset of an $H_{a'}$-coset, since C_1 is H_a-periodic, and in view of KST(i) with $H_{a'}$, it follows that $|\phi_a(A_1)+\phi_a(B)| < |\phi_a(A_1)|+|\phi_a(B)|-1$ and $|\phi_a(A)+\phi_a(B_1)| < |\phi_a(A)|+|\phi_a(B_1)|-1$. Since $A_0+B_0=C_0$, it follows from (c.17) that w.l.o.g. $A_1+B=C_1$. Thus from the conclusions of the previous two sentences, and in view of Kneser's Theorem, it follows that $\phi_a(C_1)$ is periodic, contradicting that C_1 is maximally H_a-periodic. So in view of the previous paragraph we may assume that KST(ii) holds with H_a. Hence, since C_1 is H_a-periodic, it then follows from a simple counting argument that A_1 and B_1 are H_a-periodic, completing the proof. □

5.3 Some Illustrative Examples

In this section we give some examples relating the results from the previous section with similar results obtained using the isoperimetric method. However, we first note that it is a result of Mann, or an easily derived consequence of Kneser's Theorem, that a finite, nonempty subset $B \subseteq G$ being Cauchy is equivalent to there not existing a finite subgroup H of G such that $|H+B| < \min\{|G|, |H|+|B|-1\}$ (c.20) [46] [47], i.e., B cannot have too few H-holes for any subgroup H such that $H+B \neq G$.

The following is a simple proof of Theorem 4.6 from [35].

Theorem 5.6. *Let G be an abelian group, let $B \subseteq G$ be a Cauchy subset, and let $B = B_1 \cup B_0$ be a reduced quasi-periodic decomposition of B. Then a necessary and sufficient condition for there to exist a finite, nonempty subset $A \subseteq G$ such that $|A+B| \leq \min\{|G|-2, |A|+|B|-1\}$ and $|A| \geq 2$, is that $|B| < |G|-2$ and one of the following conditions holds:*

(i) B_0 is an arithmetic progression with at least two terms and either B is not quasi-periodic (so that $B = B_0$ and the difference of the progression either has infinite order or else is of order $|G|$) or $\overline{B}$ is an arithmetic progression of finite length (so that G is finite),

(ii) $|B_0| = 1$,

(iii) for any $b \in B$, *there exists a finite subgroup* H *generated by* $(B-b)\cap H$ *such that* $|H+B| = |H|+|B|-1 < |G|$ *and* $|H| \geq 3$.

Proof. To show sufficiency, in case (i) let $A = \{0, d\}$, where d is the difference of the arithmetic progression B_0, in case (ii) let $A = \{0, h\}$, where h is any nonzero element of a quasi-period of $B = B_1 \cup B_0$, and in case (iii) let $A = H$. We next show necessariness.

If $|B| \geq |G|-1$, then $|A+B| \geq |G|-1$. Furthermore, if $|B| = |G|-2$, then since B is Cauchy, it follows that $|A+B| \geq \min\{|G|, |A|+|B|-1\} \geq |B|+1 = |G|-1$ for any finite subset $A \subseteq G$ with $|A| \geq 2$. Thus it follows that $|B| < |G|-2$. If B does not have a unique reduced quasi-periodic decomposition, then, since B is Cauchy, it follows in view of Proposition 5.1 and (c.20) that $B = G \setminus \{g\}$ for some $g \in G$, contradicting that $|B| < |G|-2$. Thus we may assume B has a unique reduced quasi-periodic decomposition.

Since B is Cauchy, it follows from hypothesis that $|A+B| = |A|+|B|-1 < |G|$. Suppose that $A+B$ is maximally H_a-periodic with nontrivial period. Hence $A' = A + H_a$ satisfies $|A'+B| \leq |A|+|B|-1 < |G|$, whence $A' = A$, since otherwise $|A'+B| < |A'|+|B|-1$, contradicting that B is Cauchy. Thus, since $|A+B| = |A|+|B|-1$, then from Kneser's Theorem it follows that

$$|H_a + B| - |B| = |H_a| - 1. \tag{5.1}$$

Let $b \in B$ and let H be the subgroup generated by $H_a \cap (B-b)$.

First suppose that $|H| = 1$. Hence $|H_a \cap (B-b)| = 1$, whence (ii) follows in view of (5.1) and the uniqueness of the reduced quasi-periodic decomposition for B. Next suppose that $|H| = 2$. Hence $|H_a \cap (B-b)| = 2$, and from (5.1) it then follows that B has a reduced quasi-periodic decomposition with quasi-period H and with its aperiodic part having cardinality one. Thus, as in the previous sentence, it follows that (ii) holds. So we may assume that $|H| \geq 3$.

Since $H+B \subset H_a+B \subset (A-a_0)+B \neq G$, where $a_0 \in A$, it follows that $|H+B| < |G|$. In view of (5.1) and the definition of H, it follows by counting holes that

$$|H+B| - |B| \leq |H_a+B| - |B| - (l-1)|H| =$$
$$|H_a| - 1 - (l-1)|H| = l|H| - 1 - (l-1)|H| = |H| - 1,$$

where $l = [H_a, H]$. Since B is Cauchy, and since $|H+B| < |G|$, then in view of (c.20) it follows that we must have equality in the above inequality, and (iii) follows. So we may assume that there does not exist a subset A satisfying the hypothesis of the theorem with the additional property that $A+B$ is periodic.

Since $A+B$ is aperiodic, apply KST to the pair (A, B) and let $A = A_1' \cup A_0'$ and $B = B_1' \cup B_0'$ be the Kemperman decomposition of (A, B) with quasi-period H_a. Since $A+B$ is aperiodic, it follows in view of (c.8) that (A, B) cannot have type (III). If (A, B) has type (IV), then from the characterization of type (IV) it follows that we can find an element $a_0 \in G \setminus A$ such that $(A \cup \{a_0\}, B)$ will be a type (III) pair. Furthermore, since $|A+B| < |G|-1$, and since $|A+B|$ is congruent to -1 modulo $|H_a|$ for type (IV), it follows that $|A+B| \leq |G| - |H_a| - 1 \leq |G| - 3$, implying $|(A \cup \{a_0\}) + B| = |A+B| + 1 \leq |G| - 2$. Hence this reduces to the previously handled case. If (A, B) has type (I) with $|B_0'| = 1$, then (ii) follows by the uniqueness of a reduced quasi-periodic decomposition for B.

Suppose (A, B) has type (I) with $|B_0'| \geq 2$ and $|A_0'| = 1$. Since B is Cauchy, it follows that if $B = B_1'' \cup B_0''$ is a quasi-periodic decomposition with quasi-period H, then $H+B = G$, or $|B_0''| = 1$, or H is infinite and $B = B_0$. Hence, since $|B_0'| \geq 2$, and since H_a must be finite else $|A| = |A_0| = 1$, a contradiction, it follows that $\phi_a(B) = G/H_a$, implying from KST(ii) that $\phi_a(A) = 1$, whence $A = A_0'$. However, since $|A_0'| = 1$ and since $|A| \geq 2$, this is a contradiction. So we may assume that (A, B) has type (II), implying that B_0' is an arithmetic progression with $|B_0'| \geq 2$ and also, by the characterization of the type (II) Kemperman decomposition, that $B = B_1' \cup B_0'$ is reduced (as $A+B$ is aperiodic). Since B is Cauchy, then the remainder of conclusion (i) follows easily from (c.20) and the uniqueness of the reduced quasi-periodic decomposition for B. □

The following is the (corrected) Theorem 6.6 from [36], which we will derive as a basic corollary to Theorem 5.6 (there is a typo in the original statement of Theorem 6.6; namely the inequality in Theorem 6.6(iii) should not be strict, as is easily seen by the example $G = \mathbb{Z}/6\mathbb{Z}$, $B = \{0, 3, 1\}$).

Theorem 5.7. *Let B be a finite, nonempty subset of an abelian group G. If $|B| \leq |G|/2$, then one of the following conditions holds:*

(i) $|A + B| \geq \min\{|G| - 1, |A| + |B|\}$, for all finite subsets $A \subseteq G$ with $|A| \geq 2$,

(ii) B is an arithmetic progression,

(iii) there is a finite, nontrivial subgroup H, such that $|H + B| \leq \min\{|G| - 1, |H| + |B| - 1\}$.

Proof. We may assume B is Cauchy, else (iii) follows in view of (c.20). We may also assume that the hypothesis of Theorem 5.6 holds for B, else (i) follows. Apply Theorem 5.6 to B. If Theorem 5.6(iii) holds, then (iii) follows. We may assume $|B| > 1$, else (ii) follows. Hence, if Theorem 5.6(ii) holds, then we may assume $B = B_1 \cup B_0$ is quasi-periodic with quasi-period H. Hence (iii) follows unless $H + B = G$, in which case $|B| > |G|/2$, a contradiction. So we may assume that Theorem 5.6(i) holds. However, $|B| \leq |G|/2$ and B being Cauchy prevent B from being quasi-periodic, whence $B = B_0$ is an arithmetic progression, and (ii) follows. $\square$

The following theorem gives a nonrecursive description of those finite, nonempty subsets A for which $|A + B| = |A| + |B| - 1$, where B is a fixed Cauchy subset. This shows that additionally assuming one of the sets from a critical pair is Cauchy allows for a significant simplification of the structure of the pair.

Theorem 5.8. *Let G be an abelian group, let A, $B \subseteq G$ be finite, nonempty subsets, and let $B = B_1 \cup B_0$ be a reduced quasi-periodic decomposition of B. Suppose that B is Cauchy. Then $|A+B| = |A| + |B| - 1$, with either $A+B$ aperiodic or else $\nu_c(A, B) = 1$ for some c, if and only if one of the following conditions holds:*

(i) A is aperiodic and $A = g_0 - \overline{B}$, for some $g_0 \in G$ (in which case $A + B = G \setminus \{g_0\}$),

(ii) $A = (g_0 - \overline{B}) \cup \{g_1\}$, for some $g_0 \in G$ and $g_1 \notin g_0 - \overline{B}$ (in which case $A + B = G$),

(iii) $|A| = 1$ or $|B| = 1$,

(iv) A and B_0 are arithmetic progressions with common difference d, where the order of d is at least $|A| + |B_0| - 1$, and either B is not quasi-periodic (in which case $B = B_0$) or $\overline{B}$ is an arithmetic progression with difference d and finite length,

(v) $|B_0| = 1$, and there exists a quasi-period H_a of $B = B_1 \cup B_0$ (namely the maximal quasi-period of the type (I) Kemperman decomposition of (A, B)) such that A has a quasi-periodic decomposition $A = A'_1 \cup A'_0$ with quasi-period H_a and $A'_0 \neq \emptyset$, such that $\nu_c(\phi_a(A), \phi_a(B)) = 1$, where $c = \phi_a(A'_0) + \phi_a(B_0)$, such that $\phi_a(B)$ is Cauchy, and such that the pair $(\phi_a(A), \phi_a(B))$ satisfies one of (ii), (iii) or (iv) with $G = G/H_a$.

Furthermore, $|A + B| = |A| + |B| - 1 < |G|$ with $A + B$ maximally H_k-periodic with nontrivial period, if and only if A is maximally H_k-periodic with nontrivial period, $A+B \neq$

G, $|B + H_k| - |B| = |H_k| - 1$, and the pair $(\phi_k(A), \phi_k(B))$ satisfies the hypotheses from the above paragraph with $G = G/H_k$; and $|A + B| = |A| + |B| - 1 = |G|$ if and only if $|A| = |G| - |B| + 1$.

Proof. We first show that the furthermore statement of the theorem follows from the first part of the theorem. Note that the last part of the furthermore statement is a consequence of Proposition 1.2.

Suppose A is maximally H_k-periodic with nontrivial period, $A + B \neq G$, B is Cauchy, $|B + H_k| - |B| = |H_k| - 1$, and the pair $(\phi_k(A), \phi_k(B))$ satisfies the hypotheses from the first part of the theorem with $G = G/H_k$. Then by the first part of the theorem it follows that $|\phi_k(A) + \phi_k(B)| = |\phi_k(A)| + |\phi_k(B)| - 1$. Hence, since A is H_k-periodic, and since $|B + H_k| - |B| = |H_k| - 1$, it follows that $|A + B| = |A| + |B| - 1$ with $A + B$ being H_k-periodic. Furthermore, since $|A+B| = |A|+|B|-1$, since A is maximally H_k-periodic, and since $A+B \neq G$, it follows that $A+B$ is maximally H_k-periodic, since otherwise $A+H$ will contradict that B is Cauchy, where $A + B$ is maximally H-periodic.

Next suppose that B is Cauchy and that $|A + B| = |A| + |B| - 1 < |G|$ with $A + B$ maximally H_k-periodic with nontrivial period. Hence, by the reasoning from the previous paragraph, it follows that A must be maximally H_k-periodic (else we contradict that B is cauchy) and $A+B \neq G$. Thus, since $|A+B| = |A|+|B|-1$, then in view of Kneser's Theorem it follows that $|B+H_k|-|B| = |H_k|-1$, and that $|\phi_k(A)+\phi_k(B)| = |\phi_k(A)|+|\phi_k(B)|-1$. Also, by the maximality of H_k it follows that $\phi_k(A) + \phi_k(B)$ is aperiodic. Finally, since B is Cauchy and since $|B + H_k| - |B| = |H_k| - 1$, then in view of (c.20) it follows by counting holes that $\phi_k(B)$ is Cauchy. Thus the pair $(\phi_k(A), \phi_k(B))$ satisfies the hypotheses of the first part of the theorem with $G = G/H_k$, and the proof of the furthermore statement of the theorem is complete.

Sufficiency of the first part of the theorem follows directly from KST-II. Thus it remains to show necessariness. Assume B is Cauchy, $|A+B| = |A|+|B|-1$, and, moreover, if $A+B$ is periodic, then $\nu_c(A,B) = 1$ for some c. Apply KST to the pair (A,B) and let $A = A_1' \cup A_0'$ and $B = B_1' \cup B_0'$ be the corresponding Kemperman decompositions with maximal quasi-period H_a.

Since B is Cauchy, it follows that if $B = B_1'' \cup B_0''$ is a quasi-periodic decomposition with quasi-period H, then $H+B = G$, or $|B_0''| = 1$, or H is infinite and $B = B_0''$ (c.21). Hence from KST it follows that (i) or (ii) holds provided (A,B) has type (IV) or (III), respectively.

Suppose B does not have a unique reduced quasi-periodic decomposition. Hence, since B is Cauchy, it follows in view of Proposition 5.1 and (c.20) that $B = G \setminus \{g\}$ for some $g \in G$. Thus $|A+B| = |A|+|B|-1$ implies $|A| \leq 2$, and it is easily seen that (ii) or (iii) holds. So we may assume B has a unique reduced quasi-periodic decomposition (c.22).

Suppose (A,B) has type (II). Hence in view of KST, the characterization of the Kemperman decomposition for type (II), and (c.22), it follows that $B_0 = B_0'$ (since the case where B is a punctured H-periodic set with $|H| > 2$ reduces to the argument of the previous paragraph) and A_0' are both arithmetic progressions with common difference d, where the order of d is at least $|A_0'| + |B_0| - 1$, that $|A_0'| \geq 2$ and that $|B_0| \geq 2$. Hence, in view of (c.21) it follows that either $B = B_0$ and $A = A_0$ (since an infinite quasi-period is possible only if both periodic parts are empty), yielding (iv), or that $\phi_a(B) = G/H_a$. Thus from KST(ii) it follows that $\phi_a(A) = 1$, implying $A = A_0'$. Furthermore, since B_0 is an arithmetic progression with difference d, then it follows from (c.21) that if B is quasi-periodic, then $\overline{B}$ is a finite arithmetic progression with difference d. Thus (iv) follows. So we may assume (A,B) has type (I).

Suppose $|B_0'| > 1$. Hence from (c.21) it follows that $|\phi_a(A)| = 1$, whence $A = A_0'$.

Hence, since (A, B) has type (I) with $|B_0'| > 1$, it follows that $|A| = |A_0'| = 1$ whence (iii) holds. So we may assume $|B_0'| = 1$. Thus in view of (c.22) it follows that $B_0' = B_0$, whence we may assume $|B_1| \neq 0$, else (iii) follows with $|B| = 1$.

Since $|B_0| = 1$, then in view of KST and the above work, it follows that (v) will hold, provided we can additionally show that $\phi_a(B)$ is Cauchy, and also that $(\phi_a(A), \phi_a(B))$ does not have type (I) with $|B''| = 1$, where B'' is the aperiodic part of the corresponding Kemperman decomposition of $\phi_a(B)$.

Suppose $\phi_a(B)$ is not Cauchy. Hence by (c.20) it follows that there exists a finite subgroup H of G such that $|H/H_a + \phi_a(B)| < |G/H_a|$ and $|H/H_a + \phi_a(B)| < |H/H_a| + |\phi_a(B)| - 1$. Hence, since B has exactly $|H_a| - 1$ H_a-holes, it follows by multiplying the previous inequality by H_a that $|H + B| < |H| + (|B| + |H_a| - 1) - |H_a| = |H| + |B| - 1$. Also, $|H/H_a + \phi_a(B)| < |G/H_a|$ implies that $|H + B| < |G|$, whence in view of (c.20) and the last sentence it follows that B is not Cauchy, a contradiction. So we may assume $\phi_a(B)$ is Cauchy.

Let $\phi_a(B) = \phi_a(B_1'') \cup \phi_a(B_0'')$ and $\phi_a(A) = \phi_a(A_1'') \cup \phi_a(A_0'')$, with $B = B_1'' \cup B_0''$ and $A = A_1'' \cup A_0''$, be the corresponding modulo H_a Kemperman decompositions with maximal quasi-period H/H_a. Suppose $(\phi_a(A), \phi_a(B))$ has type (I) with $|\phi_a(B_0'')| = 1$. Hence, $B_0'' = B_0$ and $B_1'' = B_1$. Thus, since $|B_1| > 0$, it follows that H/H_a is nontrivial. Since $B_0'' = B_0$, it follows for $a_0 \in A$, in view of the characterization of the type (I) Kemperman decomposition and KST, that $\eta_{a_0}(B, A) > 0$ if and only if $\eta_{\phi_a(a_0)}(\phi_a(B), \phi_a(A)) > 0$, whence the characterization of the type (I) Kemperman decomposition implies that $A_0'' = A_0'$ and $A_1'' = A_1'$. Thus $A = A_1' \cup A_0'$ and $B = B_1 \cup B_0$ are a quasi-periodic decomposition that satisfies KST with quasi-period H, whence by the maximality of H_a it follows that $H = H_a$. Hence H/H_a is trivial, a final contradiction. $\square$

Chapter 6

Vampirish Set Partition Draining Results

6.1 Discussion

In this chapter, we present a useful draining result for set partitions. Given an n-set partition satisfying the Cauchy-Davenport bound, the main result will, modulo some restrictions, allow us to drain $n-1$ terms from the set partition in such a way that the resulting set partition still satisfies the original bound. In applications, this allows the drained terms to then be put to other ends, essentially allowing them to be used twice, which gives a small but noticeable boost in the effectiveness of the methods developed in the first part of this thesis. The utility of the theorem is most notable when $|S| \geq 3n-1$, and so a sufficiently compressed set partition is required before implementing it. The statement of the result is as follows.

Theorem 6.1. *Let S be a finite sequence of elements from an abelian group G. If S has an n-set partition, $A = A_1, \ldots, A_n$, such that*

$$|\sum_{i=1}^{n} A_i| \geq \sum_{i=1}^{n} |A_i| - n + 1, \tag{6.1}$$

then there exists a subsequence S' of S, with length $|S'| \le \max\{|S| - n + 1, 2n\}$, and with an n-set partition, $A' = A'_1, \ldots, A'_n$, such that $|\sum_{i=1}^{n} A'_i| \ge \sum_{i=1}^{n} |A_i| - n + 1$. Furthermore, if $||A_i| - |A_j|| \le 1$ for all i and j, or if $|A_i| \ge 3$ for all i, then $A'_i \subseteq A_i$.

Note that the sequence $S = (\underbrace{0, \ldots, 0}_{n}, \underbrace{1, \ldots, 1}_{n}, \underbrace{2, \ldots, 2}_{n'})$, where $n' \le n$ and $G = \mathbb{Z}/m\mathbb{Z}$, shows that the bound on $|S'|$ in Theorem 6.1 is tight for $|S| \le 3n$. However, it is not immediately evident from the proof what other set partitions critically satisfy the bound on the number of possible removed terms, though such a description might prove useful. It is also not evident whether the conditions $|A_i| \ge 3$ or $||A_i| - |A_j|| \le 1$ are needed to obtain the conditions $A'_i \subseteq A_i$, or if $|A_i| \ge 2$ would suffice.

6.2 Draining Results

We begin first with the case when $n = 2$, where we have the following stronger version.

Theorem 6.2. *Let G be an abelian group, and let $A, B \subseteq G$ be finite subsets such that $|A| \ge 2$ and $|B| \ge 3$. If $|A + B| \ge |A| + |B| - 1$, then either:*

(i) there exists $b \in B$ such that $|A + (B \setminus \{b\})| \ge |A| + |B| - 1$, or

(ii) (a) $|A + B| = |A| + |B| - 1$, (b) there exists $a \in A$ such that $A \setminus \{a\}$ is H_a-periodic, and (c) there exists $\alpha \in G$ such that $B \subseteq \alpha + H_a$.

Proof. Suppose (i) does not hold. Hence $\eta_b(A, B) \ge 1$ for all $b \in B$. Furthermore, if $|A + B| > |A| + |B| - 1$, then $\eta_b(A, B) \ge 2$ for all $b \in B$, whence $|A + (B \setminus \{b\})| \ge |A| + 2(|B| - 2) \ge |A| + |B| - 1$ for any $b \in B$. So we may assume $|A + B| = |A| + |B| - 1$. Hence apply KST to (A, B) and let $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ be the Kemperman decompositions with quasi-period H_a. Since $|B| \ge 3$, and since $\eta_b(A, B) \ge 1$ for each $b \in B$, it follows from (c.9) that (A, B) has type (I) with $|A_0| = 1$, whence the remainder of the theorem follows from the characterization of the Kemperman decomposition for type (I) given in Proposition 5.3. □

The next lemma gives us conditions that allow us to remove a term while preserving a large cardinality sumset both locally and globally. Theorem 6.2 and Lemma 6 will both be used in the proof of Theorem 6.1

Lemma 6. *Let G be an abelian group, and let A, B, $C_1, \ldots, C_r \subseteq G$ be finite subsets with $|B| \geq 3$. If $|A+B| > |A|+|B|-1$, $|A+B+\sum_{i=1}^{r} C_i| \geq |A|+|B|+\sum_{i=1}^{r}|C_i|-(r+2)+1$, and $|A+\sum_{i=1}^{r} C_i| \geq |A|+\sum_{i=1}^{r}|C_i|-(r+1)+1$, then there exists $b \in B$ such that $|A+(B\setminus\{b\})| \geq |A|+|B|-1$ and $|A+(B\setminus\{b\})+\sum_{i=1}^{r} C_i| \geq |A|+|B|+\sum_{i=1}^{r}|C_i|-(r+2)+1$.*

Proof. Let $b_1, \ldots, b_k$ be those $b_i \in B$ such that $|A+(B\setminus\{b\})| \geq |A|+|B|-1$, and let $b_{k+1}, \ldots, b_n$ be the remaining elements of B. Note $\eta_{b_i}(A,B) \geq 1$ for all i, else the proof is complete with $b = b_i$. Since $|A+B| > |A|+|B|-1$, then for each b_j with $j > k$, it follows that $\eta_{b_j}(A,B) \geq 2$. Thus, if $k \leq n-2$, then for $j > k$ it follows, in view of $\eta_{b_i}(A,B) \geq 1$ for all i, that $|A+(B\setminus\{b_j\})| \geq \min\{|A|+2(n-k-1)+k-1,\ |A|+2n-4\} \geq |A|+|B|-1$, contradicting that $j > k$. So $k \geq n-1$.

Let $C = \sum_{i=1}^{r} C_i$. If $|A+C+B| < |A+C|+|B|-1$, then it follows from Proposition 1.3 that the proof is complete with $b = b_1$. Thus $|A+C+B| \geq |A+C|+|B|-1$. Suppose $|A+C+B| > |A+C|+|B|-1$. Hence, since $|A+\sum_{i=1}^{r} C_i| \geq |A|+\sum_{i=1}^{r}|C_i|-(r+1)+1$, it follows that $\eta_{b_j}(A+C,B) \geq 2$ for all $j \leq k$, else the proof is complete with $b = b_j$. Hence, since $k \geq n-1$, it follows that $|A+C+(B\setminus\{b_1\})| \geq |A+C|+2n-4 \geq |A+C|+|B|-1$. Thus in view of $|A+\sum_{i=1}^{r} C_i| \geq |A|+\sum_{i=1}^{r}|C_i|-(r+1)+1$ it follows that the proof is complete with $b = b_1$. So we may assume

$$|A + C + B| = |A + C| + |B| - 1. \tag{6.2}$$

Since $|A + \sum_{i=1}^{r} C_i| \geq |A| + \sum_{i=1}^{r} |C_i| - (r + 1) + 1$, it follows that the proof will be complete with $b = b_j$, $j \leq k$, unless

$$|A + C + (B \setminus \{b_j\})| \leq |A + C| + |B| - 2. \tag{6.3}$$

However, since $k \geq 2$, if the inequality in (6.3) is sharp for some $j \leq k$, then from (6.2) and Proposition 1.3, it follows for $j' \leq k$, $j' \neq j$ that $|A + C + (B \setminus \{b_{j'}\})| \geq |A + C| + |B| - 1$, contradicting (6.3). Hence, for $j \leq k$, it follows that

$$|A + C + (B \setminus \{b_j\})| = |A + C| + |B| - 2. \tag{6.4}$$

If $A + C$ is periodic, then $|A + C + B| - |A + C + (B \setminus \{b_j\})|$ must be a multiple of the nontrivial period's cardinality, contradicting (6.4) and (6.2). So we may assume $A + C$ is aperiodic. Hence C is aperiodic, whence from Kneser's Theorem it follows that $|C| \geq \sum_{i=1}^{r} |C_i| - r + 1$. Hence if $|A + C| > |A| + |C| - 1$, then $|A + C| > |A| + \sum_{i=1}^{r} |C_i| - (r + 1) + 1$, whence in view of (6.4) the proof is complete with $b = b_1$. So

$$|A + C| = |A| + |C| - 1. \tag{6.5}$$

Note that $\eta_{b_j}(A + C, B) \geq 1$ for b_j with $j \leq k$ else the proof is complete. Suppose $\eta_{b_n}(A + C, B) \geq 1$. Hence, since $k \geq n - 1$ and since $\eta_{b_j}(A + C, B) \geq 1$ for b_j with $j \leq k$, then from Theorem 6.2 and (6.2) it follows that $A+C$ has a quasi-periodic decomposition $C_1 \cup C_0$, where $C_0 = \{c_0\}$ and C_1 is maximally H_a-periodic, and that B is a subset of an H_a-coset. Since B is a subset of an H_a-coset, and since $|B| \geq 3$, it follows that $|H_a| \geq 3$. Thus from Proposition 5.1 it follows that $A+C$ is not a punctured periodic set. Hence, from (6.5) and Proposition 5.5 applied to $A+C$, it follows that A has a quasi-periodic decomposition $A_1 \cup A_0$ where A_1 is H_a-periodic and $|A_0| = 1$. Hence, since B is a subset of an H_a-coset, it follows that $|A+B| = |A| + |B| - 1$, a contradiction. So we may assume that $\eta_{b_n}(A+C, B) = 0$ and, since $\eta_{b_j}(A+C, B) \geq 1$ for b_j with $j \leq k$, that $k = n-1$.

Since $k \geq 2$, it follows from (6.4) and (6.2) that we can apply KST to $(A+C, B)$. Hence, let $A+C = C_1 \cup C_0$ and $B = B_1 \cup B_0$ be the Kemperman decompositions with quasi-period H_a. Since b_n is the unique $b \in B$ with $\eta_b(A+C, B) = 0$, it follows in view of (c.18) that $|B_1| \leq 1$ and hence, since B_1 is periodic, that $|B_1| = 0$. Hence B_0 is a subset of an H_a-coset, and since $\eta_{b_n}(A+C, B) = 0$, it follows in view of Proposition 5.3 that $(A+C, B)$ cannot have type (I) with $|C_0| = 1$. Hence, in view of (c.9) and since $\eta_{b_j}(A+C, B) \geq 1$ for b_j with $j \leq k = n-1$, it follows that we may assume $n = 3$; furthermore $|\{c \in A+B \mid \nu_c(A, B) > 0\}| = 2$, implying that $(A+C, B)$ has type (II) with (b_1, b_3, b_2) an arithmetic progression with difference $d = b_1 - b_3 = b_3 - b_2$, that $C_1 \cup C_0$ is the unique reduced quasi-periodic decomposition of $A+C$ (in view of (c.12) since $|B_0| > 2$), that $2 \leq |C_0| \leq |\langle d \rangle| - 2$, and that C_0 is an arithmetic progression with difference d. Hence $A+C$ is not a punctured periodic set, and from (6.5) and Proposition 5.5 it follows that A has a quasi-periodic decomposition $A_1 \cup A_0$ with quasi-period H_a, and that C has a quasi-periodic decomposition $C = C_1' \cup C_0'$, such that $A_0 + C_0' = C_0$. If $|A_0| = 1$, then since B is a subset of an H_a-coset it follows that $|A+B| = |A| + |B| - 1$, a contradiction. So we may assume $|A_0| \geq 2$. Thus, since $A_0 + C_0' = C_0$, and since $|C_0| \leq |\langle d \rangle| - 2$, it follows that $2 \leq |A_0| \leq |\langle d\rangle| - 2$. Hence $|\langle d\rangle| \geq 4$. Since C_0 is an arithmetic progression with difference d, and since $|A_0| \geq 2$, then if (A, C) has type (I) it follows that A_0 is an arithmetic progression with difference d. Otherwise, since $A + C$ is aperiodic and not a punctured periodic set, it follows that (A, C) has type (II). Thus A_0 and C_0' are arithmetic progressions with $A_0 + C_0' = C_0$ an arithmetic progression with difference d and at most $|\langle d\rangle| - 2$ terms. Since C_0 has at most $|\langle d\rangle| - 2$ terms, it follows that the difference of the arithmetic progression C_0 is unique up to sign. Hence A_0 must be an arithmetic progression with difference d in this case as well. Thus A_0 is an arithmetic progression with difference d regardless of the type of (A, C). Hence, since $2 \leq |A_0| \leq |\langle d\rangle| - 2$, then it follows from Proposition 5.1 that $A_1 \cup A_0$ is the unique reduced quasi-periodic decomposition of A.

Since $k = n - 1 = 2$, it follows from the definition of k that $|A + \{b_1, b_2\}| \leq |A| + |\{b_1, b_2\}| - 1$; furthermore, in view of Proposition 1.3 it follows that the proof is complete with $b = b_2$, unless $|A + \{b_1, b_2\}| = |A| + |\{b_1, b_2\}| - 1$. Hence, since $\eta_{b_1}(A, B) \geq 1$, it follows that we can apply KST to the pair $(A, \{b_1, b_2\})$. Let $A = A_1' \cup A_0'$ and $\{b_1, b_2\} = B_1' \cup B_0'$ be the Kemperman decompositions. Since B_0' is nonempty, and since B_1' periodic implies $|B_1'| \geq 2$ or $|B_1'| = 0$, it follows that $B_0' = \{b_1, b_2\}$. Hence, since $\eta_{b_i}(A, B) \geq 1$ for $i = 1, 2$, then it follows from KST(iii) that A_0' is an arithmetic progression with difference $b_1 - b_2$, and that $(A, \{b_1, b_2\})$ has type (I) or (II). However from the conclusion of the last paragraph it follows that $A = A_1 \cup A_0$ is the unique reduced quasi-periodic decomposition of A, and that $|A_0| \geq 2$. Hence $(A, \{b_1, b_2\})$ must be of type (II), whence from (c.12) it follows that $A_1' = A_1$ and $A_0' = A_0$. Thus $A_0 = A_0'$ is an arithmetic progression with difference $d = b_1 - b_3 = b_3 - b_2$ as well as an arithmetic progression with difference $b_1 - b_2$. Hence, since $2 \leq |A_0| \leq |\langle d\rangle| - 2$ so that the difference of A_0 is unique up to sign, it follows that $\pm(b_2 - b_1) = b_1 - b_3 = b_3 - b_2$, contradicting that the b_i are distinct or that $|\langle d\rangle| \geq 4$. $\square$

We note that conclusion (ii) of Theorem 6.2 implies both that $|A+(B\backslash\{b\})| \geq |A|+|B|-2$ for all $b \in B$ and that $|A| > |B|$, so that by interchanging the roles of A and B we can be assured that (i) will hold. We can now begin the proof of Theorem 6.1.

Proof. We may assume $|S| \geq 2n+1$ and $n \geq 2$, else the theorem is trivial. Let $|S| = sn+r$, where $s \geq 2$ and $0 \leq r < n$. If neither of the conditions of the furthermore part of Theorem 6.1 hold, then we may w.l.o.g. assume that A was chosen from all n-set partitions of S that satisfy (6.1) so that the cardinality s' of the minimal cardinality set A_i in A is maximal, and such that, subject to prior conditions, the number of terms A_i in A with cardinality s' is minimal. Re-index so that the cardinalities of the A_i are nondecreasing, and assume that $|A_i| \geq s+2$ for $i > k_2$, and that $|A_i| \leq \min\{2,\, s-1\}$ for $i < k_1$.

The remainder of the proof is divided into two cases. The first handles the case when either all sets A_i are of cardinality at least three or all are of cardinality equal to two or three. Under these conditions, we show in Case 1b that we can inductively remove terms from the sets A_i one by one, unless highly restrictive conditions occur. Under these restrictive conditions, we show in Case 1a that we can complete the removal of the remaining terms in one swipe. We note that the complexity of the induction statement in Case 1b arises from the exceptional case in Theorem 6.2, and that without this problem the induction would go through quite smoothly. Finally, Case 2 handles the case when the set-partition A can't be reduced to one satisfying the conditions of Case 1. In this case, a similar argument to that of Case 1a works quite simply provided the Cauchy-Davenport bound does not hold for every subsequence of A. Thus the majority of Case 2 is spent showing that it is quite difficult for a set-partition A to satisfy Cauchy-Davenport everywhere and not be reducible to a set partition either with a larger minimal cardinality set or with a fewer number of minimal cardinality sets.

Case 1a: Suppose that $k_1 = 1$, and if $s = 2$ that $k_2 = n$ (i.e., one of the hypotheses of the furthermore part of Theorem 6.1 holds, whence in view of Theorem 6.1 it follows that we may assume $n \geq 3$). Further suppose that, allowing re-indexing, there exist an n-set partition, $A' = A'_1, \ldots, A'_n$, of a subsequence S' of S, and an integer l with $2 \leq l \leq n$, such that

$$|\sum_{i=1}^{n} A'_i| \geq \sum_{i=1}^{n} |A_i| - n + 1, \tag{6.6}$$

$A'_i \subseteq A_i$, $\sum_{i=1}^{l} |A'_i| = |A_1| + \sum_{i=2}^{l} \max\{2,\ |A_i| - 1\}$, $\sum_{i=1}^{l} A'_i$ is maximally H_a-periodic with nontrivial period, $|A_1| = \min_i \{|A_i|\}$, $A'_i = A_i$ for $i > l$, $|A'_l| \geq \max\{2,\ |A_l| - 1\}$,

$$|\sum_{i=1}^{l-1} A'_i| \geq \sum_{i=1}^{l-1} |A_i| - (l-1) + 1, \tag{6.7}$$

and

$$|\sum_{i=1}^{l} A'_i| < \sum_{i=1}^{l} |A_i| - l + 1. \tag{6.8}$$

Let b be the integer such that

$$b|H_a| < \sum_{i=1}^{n} |A_i| - n + 1 \leq (b+1)|H_a|, \tag{6.9}$$

let ρ be the integer such that

$$|\sum_{i=1}^{l} A'_i| = |\sum_{i=1}^{l-1} A'_i| + |A'_l| - 1 - \rho, \tag{6.10}$$

let $s_2 = \sum_{i=l+1}^{n} |A_i|$, let $s_1 = \sum_{i=1}^{l} |A_i|$, and let $s'_1 = \sum_{i=1}^{l} |A'_i|$.

Since $|A'_l| \geq |A_l| - 1$ and since $A'_l \subseteq A_l$, then in view of (6.7), (6.8) and (6.10), it follows that $0 \leq \rho \leq |A_l| - 1$. Furthermore, in view of Proposition 1.3, it follows that there exists a proper subset $T \subseteq A'_l$ of cardinality ρ such that $\sum_{i=1}^{l-1} A'_i + (A'_l \setminus T) = \sum_{i=1}^{l} A'_i$.

Let S'' be a minimal length subsequence of the terms of S' partitioned by the $A'_i = A_i$ where $i \geq l+1$, with an $(n-l)$-set partition, $B' = B_1, \ldots, B_{n-l}$, such that $|\sum_{i=1}^{l} \phi_a(A'_i) + \sum_{i=1}^{n-l} \phi_a(B_i)| \geq b+1$ and $B_i \subseteq A_{i+l}$ (since $\sum_{i=1}^{l} A'_i$ is H_a-periodic, such a subsequence exists by (6.6) and (6.9)). Since $\sum_{i=1}^{l} |A'_i| = |A_1| + \sum_{i=2}^{l} \max\{2, |A_i| - 1\}$, since $|A_1| = \min_i\{|A_i|\}$, since $A'_i \subseteq A_i$, since $k_1 = 1$, since $k_2 = n$ if $s = 2$, and since $\sum_{i=1}^{l} A'_i$ is H_a-periodic, it follows in view of (6.9) and the conclusion of the last paragraph that the proof will be complete unless

$$s_2 - s'_2 \leq n - l - 1 - \rho, \tag{6.11}$$

where $s'_2 = |S''|$. Hence $l < n$. From the minimality of S'' it follows that $|B_j| = |\phi_a(B_j)|$, and furthermore, for $x \in B_j$ with $|B_j| \geq 2$, that

$$\eta_{\phi_a(x)} \left(\sum_{i=1}^{l} \phi_a(A'_i) + \sum_{i=1}^{j-1} \phi_a(B_i),\, \phi_a(B_j) \right) \geq 1. \tag{6.12}$$

Hence, since $A'_i \subseteq A_i$, since $\sum_{i=1}^{l} A'_i$ is H_a-periodic, and since $|A'_l| \geq |A_l| - 1$, it follows, in view of (6.12), (6.7), (6.10) and (6.9), that we can remove an element y from S'' contained in the set B_j with greatest index such that $|B_j| \geq 2$ (since $k_1 = 1$ and $A'_i \subseteq A_i$, such a set exists in view of (6.11)) and contradict the minimality of S'' unless

$$(s'_2 - (n-l) - 1)|H_a| \leq \left| \left(\sum_{i=1}^{l} A'_i + \sum_{\substack{i=1 \\ i \neq j}}^{n-l} B_i + (B_j \setminus \{y\}) \right) \setminus \sum_{i=1}^{l} A'_i \right| \leq s_2 - (n-l) + \rho. \tag{6.13}$$

Using the estimate $|H_a| \geq 2$, it follows from (6.13) that

$$s_2' \leq (s_2 - s_2') + \rho + (n-l) + 2. \tag{6.14}$$

However, (6.14) and (6.11) imply that

$$s_2' \leq 2(n-l) + 1. \tag{6.15}$$

Hence the proof is complete unless $\rho = 0$ and equality holds in (6.15), which can only occur if $|H_a| = 2$.

From Proposition 5.1 and (c.1), it follows that a finite, nonempty set A is periodic if and only if any quasi-periodic decomposition of A has its aperiodic part being periodic or empty. Hence, if $|A_l'| \geq 3$, then since $\rho = 0$, and since $\sum_{i=1}^{l} A_i'$ is periodic with maximal period H_a, it follows from (6.10), Proposition 5.1 and Theorem 6.2, that either we can remove an additional element from A_l' leaving the sumset unchanged, whence the proof is complete, or else A_l' is maximally $H_{a'}$-periodic with nontrivial $H_{a'} \leq H_a$, whence since $|H_a| = 2$ it follows that A_l' is maximally H_a-periodic. If $|A_l'| = 2$, then since $\rho = 0$, and since $\sum_{i=1}^{l} A_i'$ is maximally H_a-periodic, it follows from (6.10) and Kneser's Theorem that $|\phi_a(A_l')| = 1$, whence since $|H_a| = 2$ it follows that A_l' is H_a-periodic. Thus regardless of the cardinality of A_l' we may assume A_l' is H_a-periodic. Hence it follows that there does not exist a set A_j' with $j < l$ and $|\phi_a(A_j')| < |A_j'|$, since otherwise we can remove an additional element from A_j' leaving the sumset unchanged and completing the proof. Hence, since $\sum_{i=1}^{l} A_i'$ is maximally H_a-periodic, and since $|H_a| = 2$, it follows in view of Kneser's Theorem and (6.8) that $s_1 - l \geq |\sum_{i=1}^{l} A_i'| \geq 2(s_1' - l + 1 - |A_l'|) + |A_l'|$. Since $A_i' \subseteq A_i$, since $k_1 = 1$, and since $s'_1 = \sum_{i=1}^{l} |A'_i| = |A_1| + \sum_{i=2}^{l} \max\{2, |A_i| - 1\}$, it follows that

$$s_1 \leq s'_1 + l - 1. \tag{6.16}$$

Hence, since $s_1 - l \geq 2(s'_1 - l + 1 - |A'_l|) + |A'_l|$, it follows that $s'_1 \leq 2l - 3 + |A'_l|$. Hence, if $|A'_l| = 2$, then in view of (6.15) it follows that the proof is complete. So we may assume $|A'_l| > 2$. Thus, since A'_l is H_a-periodic, and since $|H_a| = 2$, it follows that $|A'_l| \geq 4$. Hence, since $k_2 = n$ if $s = 2$, and since $A'_l \subseteq A_l$, it follows that $s \geq 3$. Since $s'_1 \leq 2l - 3 + |A'_l|$, it follows that $\sum_{i=1}^{l-1} |A'_i| \leq 2(l-1) - 1$. Consequently, since $s \geq 3$, since $k_1 = 1$, and since $A'_i \subseteq A_i$, it follows that $s_1 \geq s'_1 + l$, a contradiction to (6.16).

Case 1b: Suppose that $k_1 = 1$, and if $s = 2$ that $k_2 = n$. We proceed by induction on a parameter l, with $1 \leq l \leq n$, as follows. Inductively assume, passing from $l - 1$ to l, that (allowing re-indexing) we can remove elements from the sets A_i with $i \leq l - 1$, yielding new, nonempty sets A'_i, such that $\sum_{i=1}^{l-1} |A'_i| = |A_1| + \sum_{i=2}^{l-1} \max\{2, |A_i| - 1\}$, such that $|A_1| = \min_i\{|A_i|\}$, such that (6.6) and (6.7) hold with $A'_i = A_i$ for $i > l - 1$, and such that $|A'_{l-1}| \geq \max\{2, |A_{l-1}| - 1\}$; furthermore, if $l - 1 > 1$, if equality holds in (6.7), if

$$\sum_{i=1}^{l-1} A'_i = H \cup \{b\}, \tag{6.17}$$

where H is maximally H_a-periodic and $b \notin H$, and if $|H_a| > 2$, then

$$|\sum_{i=1}^{(l-1)-1} A'_i| \geq \sum_{i=1}^{(l-1)-1} |A_i| - ((l-1) - 1) + \epsilon, \tag{6.18}$$

where $\epsilon = 0$ if $|A'_{l-1}| > 3$ and $|A'_{l-1}| = |A_{l-1}|$, and $\epsilon = 1$ if $|A'_{l-1}| \leq 3$ or $|A'_{l-1}| = |A_{l-1}| - 1$. The case $l = 1$ is trivial. Note also that the $l = n$ case completes the proof, so that Case 1 will be complete once the induction is completed. Further note that (6.7) with parameter $l-1$ implies (6.18) with parameter l (in place of $(l-1)$).

Suppose there exists a set A_r with $r > l-1$ such that $|\sum_{i=1}^{l-1} A'_i + A_r| < \sum_{i=1}^{l-1}|A_i| + |A_r| - l + 1$. Hence from (6.7) it follows that $|\sum_{i=1}^{l-1} A'_i + A_r| < |\sum_{i=1}^{l-1} A'_i| + |A_r| - 1$, whence from Kneser's Theorem it follows that $\sum_{i=1}^{l-1} A'_i + A_r$ is maximally H_a-periodic with nontrivial period, and from Proposition 1.3 it follows (for $|A_r| \geq 3$) that we can remove some element x from A_r to yield a new set A'_r, such that $\sum_{i=1}^{l-1} A'_i + A_r = \sum_{i=1}^{l-1} A'_i + A'_r$. Hence, after re-indexing, the conditions of Case 1a are met, and so we may assume $|\sum_{i=1}^{l-1} A'_i + A_r| \geq \sum_{i=1}^{l-1}|A_i| + |A_r| - l + 1$. Consequently, we may assume $|A_r| > 2$ for $r > l-1$, else the induction is complete.

Suppose there exists a set A_r with $r > l-1$ such that $|\sum_{i=1}^{l-1} A'_i + A_r| < |\sum_{i=1}^{l-1} A'_i| + |A_r| - 1$. Then from Proposition 1.3 it follows that we can remove some element x from A_r to yield a new set A'_r such that $\sum_{i=1}^{l-1} A'_i + A_r = \sum_{i=1}^{l-1} A'_i + A'_r$. If $|\sum_{i=1}^{l-1} A'_i + A_r| \geq \sum_{i=1}^{l-1}|A_i| + |A_r| - l + 1$, then the induction is complete, and otherwise we reduce to the conditions of the previous paragraph. So we may assume that $|\sum_{i=1}^{l-1} A'_i + A_r| \geq |\sum_{i=1}^{l-1} A'_i| + |A_r| - 1$ for all $r > l-1$.

Suppose that the inequality in (6.7) is strict. Suppose further that $|\sum_{i=1}^{l-1} A'_i + \sum_{i=l+1}^{n} A_i| < |\sum_{i=1}^{l-1} A'_i| + \sum_{i=l+1}^{n} |A_i| - (n-l+1) + 1$. Hence in view of Proposition 1.3 it follows that there exists a set A_r with $r \geq l+1$ such that $\sum_{i=1}^{l-1} A'_i + \sum_{i=l+1}^{n} A_i = \sum_{i=1}^{l-1} A'_i + \sum_{\substack{i=l+1 \\ i \neq r}}^{n} A_i + (A_r \setminus \{x\})$ for all $x \in A_r$. In view of Theorem 6.2 and the conclusion of the last paragraph, it follows that there exists $x \in A_r$ such that $|\sum_{i=1}^{l-1} A'_i + (A_r \setminus \{x\})| \geq |\sum_{i=1}^{l-1} A'_i| + |A_r| - 2$. Hence since the inequality in (6.7) is strict, and since $\sum_{i=1}^{l-1} A'_i + \sum_{i=l+1}^{n} A_i = \sum_{i=1}^{l-1} A'_i + \sum_{\substack{i=l+1 \\ i \neq r}}^{n} A_i + (A_r \setminus \{x\})$, it follows that the induction is complete letting $A'_l = A_r \setminus \{x\}$. So we may assume $|\sum_{i=1}^{l-1} A'_i + \sum_{i=l+1}^{n} A_i| \geq |\sum_{i=1}^{l-1} A'_i| + \sum_{i=l+1}^{n} |A_i| - (n-l+1) + 1$.

Since the inequality in (6.7) is strict, and in view of the conclusion of the third paragraph of Case 1b (with $r = l$), then it follows from Proposition 1.3 that $|\sum_{i=1}^{l-1} A'_i + (A_l \setminus \{x\})| \geq \sum_{i=1}^{l} |A_i| - l + 1$, for all but at most one (say x_0) $x \in A_l$ (since $|\sum_{i=1}^{l} A'_i + (A_l \setminus \{x_0\})| < |\sum_{i=1}^{l} A'_i| + |(A_l \setminus \{x_0\})| - 1$ for such x_0). Hence the induction is complete letting $A'_l = A_l \setminus \{x\}$, with $x \in A_l$ and $x \neq x_0$, unless $|\sum_{i=1}^{l-1} A'_i + (A_l \setminus \{x\}) + \sum_{i=l+1}^{n} A_i| < \sum_{i=1}^{n} |A_i| - n + 1$. Hence, in view of strict inequality in (6.7) and the conclusion of the last paragraph, it follows that $|\sum_{i=1}^{l-1} A'_i + (A_l \setminus \{x\}) + \sum_{i=l+1}^{n} A_i| < |\sum_{i=1}^{l-1} A'_i + \sum_{i=l+1}^{n} A_i| + |(A_l \setminus \{x\})| - 1$, whence in view of Proposition 1.3 it follows that the induction is complete by letting $A'_l = A_l \setminus \{x'\}$ for any $x' \in A_l \setminus \{x, x_0\}$. So (since $|A_l| \geq 3$) we may assume that equality holds in (6.7).

Suppose there exists a set A_r with $r > l - 1$ such that $|\sum_{i=1}^{l-1} A'_i + A_r| = |\sum_{i=1}^{l-1} A'_i| + |A_r| - 1$. Hence, since $|A'_1| \leq |A_1| \leq |A_r|$, and since $|A_r| \geq 3$, then from Theorem 6.2 it follows that either the induction is complete or else (6.17) holds with $|H_a| > 2$, $A_r \subseteq \alpha + H_a$ for some $\alpha \in G$, and $l > 2$. Hence, since equality holds in (6.7), it follows by inductive assumption that (6.18) holds. Hence, since equality holds in (6.7), and since $|A'_{l-1}| \geq |A_{l-1}| - 1$, it follows that there exists a subset $H' \subset H \cup \{b\}$ with cardinality at most $|A'_{l-1}| + 1 - \epsilon$, such that $\sum_{i=1}^{l-2} A'_i = \beta + (H \cup \{b\}) \setminus H'$, for some $\beta \in G$.

Suppose $|H_a| > |A'_{l-1}| + 2 - \epsilon$. Hence, since H is H_a-periodic, and since $|H'| \leq |A'_{l-1}| + 1 - \epsilon$, it follows that if an H_a-coset $\gamma + H_a$ contains at least two elements of $\sum_{i=1}^{l-1} A'_i = H \cup \{b\}$, then the H_a-coset $(\beta + \gamma) + H_a$ will contain at least two elements of $\sum_{i=1}^{l-2} A'_i$. Hence, since $|A'_{l-1}| \geq 2$, it follows from (6.17) that $|\phi_a(A'_{l-1})| > 1$ and that $b \notin H'$, since if the contrary holds in either case, then $H \cup \{b\}$ will contain at least two elements from every H_a-coset that intersects $H \cup \{b\}$ (since the minimum number of elements from an H_a-coset in $\sum_{i=1}^{l-1} A'_i + A'_l$ is at least the minimum number of elements from an H_a-coset in $\sum_{i=1}^{l-1} A'_i$), a contradiction. Hence from the conclusions of the last two sentences it follows that $\phi_a(\sum_{i=1}^{l-2} A'_i) = \phi_a(\sum_{i=1}^{l-1} A'_i)$, whence since $|\phi_a(A'_{l-1})| > 1$, it follows in view of Proposition 1.3 applied modulo H_a that $\nu_{\phi_a(b)}(\sum_{i=1}^{l-2}\phi_a(A'_i),\, \phi_a(A'_{l-1})) \geq 2$. Hence there are two elements, c, $d \in \sum_{i=1}^{l-2} A'_i$ say, that are distinct modulo H_a, and each of which can be summed with some element of A'_{l-1} to give us an element from the coset $b + H_a$. Consequently, if the coset class represented by c has at least x elements contained in $\sum_{i=1}^{l-2} A'_i$, then the coset class of b must also contain at least x elements in $\sum_{i=1}^{l-1} A'_i$. Likewise for d. However, by (6.17) we know that b is the unique element from its H_a-coset in $\sum_{i=1}^{l-1} A'_i$, and thus by the previous two sentences both c and d must be the unique element from their coset class in $\sum_{i=1}^{l-2} A'_i$. However, it follows from the second sentence of this paragraph that if a coset class contained at least two elements in $\sum_{i=1}^{l-1} A'_i$, then the corresponding (up to translation) coset class of $\sum_{i=1}^{l-2} A'_i$ must also contain at least two elements. Since this is not the case for the two distinct coset classes c and d, it follows that there must be two distinct coset classes with a unique element in $\sum_{i=1}^{l-1} A'_i$, which contradicts (6.17). So we may assume $|H_a| \leq |A'_{l-1}| + 2 - \epsilon$.

Hence, since $|A_r| \geq 3$ and since A_r is a subset of an H_a-coset, it follows that

$$3 \leq |A_r| \leq |H_a| \leq |A'_{l-1}| + 2 - \epsilon. \tag{6.19}$$

Let $x \in A'_{l-1}$. If $\sum_{i=1}^{l-2} A'_i + (A'_{l-1} \setminus \{x\}) = \sum_{i=1}^{l-2} A'_i + A'_{l-1}$, then the induction will be complete by letting $A'_{l-1} = A'_{l-1} \setminus \{x\}$ and letting $A'_l = A_r$. Hence $\eta_x(\sum_{i=1}^{l-2} A'_i, A'_{l-1}) \geq 1$ for all $x \in A'_{l-1}$.

Suppose $\eta_{x_i}(\sum_{i=1}^{l-2} A'_i, A'_{l-1}) = 1$ holds for at least two distinct x_1, $x_2 \in A'_{l-1}$. Hence for one of these x_i, say x_1, it follows from (6.17) that

$$\left|\phi_a\left(\sum_{i=1}^{l-2} A'_i + (A'_{l-1} \setminus \{x_1\})\right)\right| = \left|\phi_a\left(\sum_{i=1}^{l-1} A'_i\right)\right|, \tag{6.20}$$

whence, since $|A_r| \geq 3$, since A_r is a subset of an H_a-coset, and since $\eta_{x_1}(\sum_{i=1}^{l-2} A_i', A_{l-1}') = 1$, it follows from (6.17) and from Proposition 1.2 that $\sum_{i=1}^{l-2} A_i' + (A_{l-1}' \setminus \{x_1\}) + A_r = \sum_{i=1}^{l-1} A_i' + A_r$, whence the induction is complete for $|A_r| > 3$ by letting $A_{l-1}' = A_{l-1}' \setminus \{x_1\}$ and letting $A_l' = A_r$. So assume $|A_r| = 3$. Hence, since A_r is a subset of an H_a-coset, it follows in view of (6.17) and (6.20) that $\left(\sum_{i=1}^{l-2} A_i' + (A_{l-1}' \setminus \{x_1\})\right) + A_r$ has a quasi-periodic decomposition $B_1 \cup B_0$ with $|B_0| = 3$. Hence, in view of Proposition 5.1, it follows that $\left(\sum_{i=1}^{l-2} A_i' + (A_{l-1}' \setminus \{x_1\})\right) + A_r$ cannot have a reduced quasi-periodic decomposition $B_1' \cup B_0'$ where $|B_0'| = 1$ and B_1' is maximally $H_{a'}$-periodic with $|H_{a'}| > 2$, since if that were the case, then it would follow from (c.1) and the uniqueness of $B_1' \cup B_0'$ that $B_0' \subseteq B_0$ and that $B_0 \setminus B_0'$ was $H_{a'}$-periodic, contradicting that $|B_0 \setminus B_0'| = 2 < |H_{a'}|$. Hence (6.17) cannot hold for $\left(\sum_{i=1}^{l-2} A_i' + (A_{l-1}' \setminus \{x_1\})\right) + A_r$ with $|H_a| > 2$. Thus, since $\sum_{i=1}^{l-2} A_i' + (A_{l-1}' \setminus \{x_1\}) + A_r = \sum_{i=1}^{l-1} A_i' + A_r$, it follows that the induction will be complete by letting $A_{l-1}' = A_{l-1}' \setminus \{x_1\}$ and letting $A_l' = A_r$. So we may assume that $\eta_x(\sum_{i=1}^{l-2} A_i', A_{l-1}') \geq 2$ for all but at most one $x \in A_{l-1}'$.

Hence from (6.18) it follows that

$$|\sum_{i=1}^{l-1} A_i'| \geq \sum_{i=1}^{l-2} |A_i| - (l-2) + \epsilon + 2(|A_{l-1}'| - 1), \tag{6.21}$$

which, from the definition of ϵ, and since $|A_{l-1}'| \geq \max\{2, |A_{l-1}| - 1\}$, contradicts that equality holds in (6.7) unless $|A_{l-1}'| = 2$ and equality holds in (6.21), whence it follows that $\eta_{x_i}(\sum_{i=1}^{l-2} A_i', A_{l-1}') \leq 2$ for both $x_1, x_2 \in A_{l-1}'$. Since $|A_{l-1}'| = 2$, implying $\epsilon = 1$ by induction hypothesis, it follows in view of (6.19) that $|H_a| = 3$. Hence, since $|A_{l-1}'| = 2$, since $\eta_{x_i}(\sum_{i=1}^{l-2} A_i', A_{l-1}') \leq 2$, and in view of (6.17), it follows for at least one of x_1 and x_2, say x_1, that (6.20) holds. Hence, since A_r is a subset of an H_a-coset, since $|A_r| \geq 3$, and since

$|H_a| = 3$, it follows that A_r is an H_a-coset, that $\sum_{i=1}^{l-1} A'_i + A_r = \sum_{i=1}^{l-2} A'_i + (A'_{l-1} \setminus \{x_1\}) + A_r$, and that $\sum_{i=1}^{l-2} A'_i + (A'_{l-1} \setminus \{x_1\}) + A_r$ is H_a-periodic. Hence, since in view of (c.5) the complement of punctured periodic set is aperiodic, it follows that (6.17) cannot hold for $\sum_{i=1}^{l-2} A'_i + (A'_{l-1} \setminus \{x_1\}) + A_r$, whence the induction is complete by letting $A'_{l-1} = A'_{l-1} \setminus \{x_1\}$ and letting $A'_l = A_r$. So we may assume that $|\sum_{i=1}^{l-1} A'_i + A_r| \neq |\sum_{i=1}^{l-1} A'_i| + |A_r| - 1$ for all $r > l-1$.

Hence, in view of the conclusion of the third paragraph of Case 1b, it follows that every set A_r with $r > l-1$ satisfies

$$|\sum_{i=1}^{l-1} A'_i + A_r| > |\sum_{i=1}^{l-1} A'_i| + |A_r| - 1. \tag{6.22}$$

Let $B_1, \ldots, B_{l'}$ be a nonempty subsequence of $A_l, \ldots, A_n$. If

$$|\sum_{i=1}^{l-1} A'_i + \sum_{i=1}^{l'} B_i| \leq |\sum_{i=1}^{l-1} A'_i| + \sum_{i=1}^{l'} |B_i| - (l'+1) + 1, \tag{6.23}$$

then, in view of (6.22) and Proposition 1.3, it follows that there exists a set B_w such that $\sum_{i=1}^{l-1} A'_i + \sum_{\substack{i=1 \\ i \neq w}}^{l'} B_i + (B_w \setminus \{x\}) = \sum_{i=1}^{l-1} A'_i + \sum_{i=l}^{l'} B_i$, for every $x \in B_w$. Hence from (6.22) and Theorem 6.2 it follows that an $x \in B_w$ can be found so that the induction is complete by letting $A'_l = B_w \setminus \{x\}$. So we may assume for any l' that (6.23) does not hold. Hence, since $|A_l| \geq 3$, then in view of (6.22) it follows that the induction is complete by applying Lemma 6 with $A = \sum_{i=l}^{l-1} A'_i$, $B = A_l$, and $C_i = A_{l+i}$.

Case 2: If $s \neq 2$, then suppose $k_1 \neq 1$, and if $s = 2$, then suppose $k_1 \neq 1$ or $k_2 \neq n$. Let s' be the minimal cardinality of a set A_i. Note from the assumptions of the case that $s' \leq 2$. Let $k \leq n$ be the index such that $|A_i| \geq s' + 2$ for $i \geq k$. Let $A_{j'}$ be a subset with $|A_{j'}| = s'$. Note, for $j \geq k$ and for every $t \in A_j \setminus A_{j'}$, that we can remove t from A_j and place t in $A_{j'}$ to form a new set $A'_{j'}$ with $|A'_{j'}| > |A_{j'}|$. Hence

$$\eta_t(\sum_{i=1}^{l} A_{b_i}, A_j) \geq 1, \tag{6.24}$$

where $A' = (A_{b_1}, \ldots, A_{b_l})$ is any nonempty subsequence of $A = (A_1, \ldots, A_n)$ that does not include the term A_j, since otherwise

$$|\sum_{\substack{i=1 \\ i \neq j,j'}}^{n} A_i + (A_{j'} \cup \{t\}) + (A_j \setminus \{t\})| \geq \sum_{i=1}^{n} |A_i| - n + 1, \tag{6.25}$$

contradicting the extremal assumptions originally assumed for A. From (6.24) and Proposition 1.3 it follows that

$$|\sum_{i=1}^{l} A_{b_i} + (A_j \setminus A'_j)| \geq |\sum_{i=1}^{l} A_{b_i}| + |(A_j \setminus A'_j)| - 1, \tag{6.26}$$

where $A' = (A_{b_1}, \ldots, A_{b_l})$ is any nonempty subsequence of $A = (A_1, \ldots, A_n)$ that does not include the term A_j, and A'_j is a proper subset of $A_j \setminus A_{j'}$.

Suppose that

$$|\sum_{i=1}^{l} A_{b_i}| \geq \sum_{i=1}^{l} |A_{b_i}| - l + 1, \tag{6.27}$$

for every nonempty subsequence $A' = (A_{b_1}, \ldots, A_{b_l})$ of $A = (A_1, \ldots, A_n)$. Since $|A_j| - |A_{j'}| \geq 2$, then in view of (6.27) and (6.26) with $A'_j = \{t\}$ and $A' = A \setminus (A_j)$, it follows that (6.25) holds, a contradiction to the extremal assumptions originally assumed for A, unless equality holds in (6.27) and (6.26) with $A'_j = \{t\}$ and $A' = A \setminus (A_j)$, and

$$|\sum_{\substack{i=1 \\ i \neq j}}^{n} A_i + (A_j \setminus \{t\})| = \sum_{i=1}^{n} |A_i| - n, \tag{6.28}$$

for each $t \in A_j \setminus A_{j'}$. However, since (6.25) cannot hold, then in view of Kneser's Theorem and (6.28), it follows that $\sum_{\substack{i=1\\ i\neq j}}^{n} A_i + (A_j \setminus \{t\}) = \sum_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t\}) + (A_{j'} \cup \{t\})$ is maximally H_{a_t}-periodic with nontrivial period. Hence, in view of (6.24) with $A' = A \setminus (A_j)$ it follows that each $t \in A_j \setminus A_{j'}$ is the only element from its H_{a_t}-coset in A_j.

Suppose $A_{j'}$ does not contain an element from the same H_{a_t}-coset as t. Thus t is the unique element from its H_{a_t}-coset in $A_{j'} \cup \{t\}$. Hence, since $\sum_{\substack{i=1\\ i\neq j}}^{n} A_i + (A_j \setminus \{t\}) = \sum_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t\}) + (A_{j'} \cup \{t\})$ is maximally H_{a_t}-periodic, and in view of Kneser's Theorem, it follows that $|\sum_{\substack{i=1\\ i\neq j',j}}^{n} A_i + (A_j \setminus \{t\}) + (A_{j'} \cup \{t\})| \geq |\sum_{\substack{i=1\\ i\neq j',j}}^{n} A_i + (A_j \setminus \{t\})| + |(A_{j'} \cup \{t\})| - 1$. Hence from (6.27) and (6.26) with $A'_j = \{t\}$ and $A' = A \setminus (A_{j'}, A_j)$, it follows that (6.25) holds, a contradiction. So we may assume $\phi_{a_t}(t) \in \phi_{a_t}(A_{j'})$. Thus, since each $t \in A_j \setminus A_{j'}$ is the only element from its H_{a_t}-coset in A_j (from second paragraph of Case 2), it follows that $A_{j'} \nsubseteq A_j$. Hence $|A_j \setminus A_{j'}| \geq 3$.

Hence in view of (6.28), (6.24), (6.27) and (6.26) with $A'_j = \{t_1, t_2\}$ and $A' = A \setminus (A_j)$, it follows that

$$|\sum_{\substack{i=1\\ i\neq j}}^{n} A_i + (A_j \setminus \{t_1, t_2\})| = \sum_{i=1}^{n} |A_i| - n - 1, \tag{6.29}$$

for any pair of distinct $t_1, t_2 \in A_j \setminus A_{j'}$. Hence, in view of (6.28) and (6.24) with $A' = A \setminus (A_j)$, it follows that $\eta_t(\sum_{\substack{i=1\\ i\neq j}}^{n} A_i, A_j) = 1$ for each $t \in A_j \setminus A_{j'}$.

Since $\sum_{\substack{i=1\\ i\neq j}}^{n} A_i + (A_j \setminus \{t\})$ is periodic, it follows that $\sum_{i=1}^{n} A_i$ is the disjoint union of that periodic set, say T, and all those elements of $\sum_{i=1}^{n} A_i$ that have precisely one representation in the sumset $\sum_{\substack{i=1\\ i\neq j}}^{n} A_i + A_j$ and with that one representation using the term t. Since $\eta_t(\sum_{\substack{i=1\\ i\neq j}}^{n} A_i, A_j) = 1$, it follows that there is precisely one such element of $\sum_{i=1}^{n} A_i$, say x, that has precisely one representation in the sumset $\sum_{\substack{i=1\\i\neq j}}^{n} A_i + A_j$ and with that one representation using the term t. Hence $\sum_{i=1}^{n} A_i = T \cup \{x\}$ is a reduced quasi-periodic decomposition of $\sum_{i=1}^{n} A_i$. Thus from (c.5) it follows that $\sum_{i=1}^{n} A_i$ must be aperiodic.

Next apply the Kneser Lemma with $C_0 = \sum_{i=1}^{n} A_i$, $C_1 = \sum_{\substack{i=1\\i\neq j}}^{n} A_i + (A_j \setminus \{t_1\})$ and $C_2 = \sum_{\substack{i=1\\i\neq j}}^{n} A_i + (A_j \setminus \{t_2\})$, where t_1 and t_2 are an arbitrary pair of distinct elements from $A_j \setminus A_{j'}$. Since $C_0 = \sum_{i=1}^{n} A_i$ is aperiodic (from the previous paragraph), it follows that $|H_{k_0}| = 1$ in the Lemma. Also note by their definitions that $H_{a_{t_1}} = H_{k_1}$ and $H_{a_{t_2}} = H_{k_2}$, in the notation of the lemma. Since $\eta_t(\sum_{\substack{i=1\\i\neq j}}^{n} A_i, A_j) = 1$ for each $t \in A_j \setminus A_{j'}$, including t_1 and t_2, then it follows that $|C_1| = |C_2| = |C_0| - 1$. Hence the inequality given by the Kneser Lemma implies that either $|H_{k_1}| \leq 2$ or $|H_{k_2}| \leq 2$. Hence, since both H_{k_1} and H_{k_2} are nontrivial by their definition, it follows that either $|H_{k_1}| = 2$ or $|H_{k_2}| = 2$. If there were two distinct elements t_1 and t_2 from $A_j \setminus A_{j'}$ both with $|H_{k_1}| \neq 2$ and $|H_{k_2}| \neq 2$, then applying the above argument with these two t_i would yield a contradiction. Thus we can assume that $|H_{a_t}| = 2$ for all but at most one (say t_0) $t \in A_j \setminus A_{j'}$.

Let $t \in A_j \setminus A_{j'}$ with $t \neq t_0$. Since $\sum_{i=1}^{n} A_i$ is aperiodic, it follows that every set A_i is aperiodic. Since $|H_{a_t}| = 2$, and since $\sum_{\substack{i=1\\i\neq j}}^{n} A_i + A_j \setminus \{t\}$ is maximally H_{a_t}-periodic, then from Kneser's Theorem it follows that

$$|\sum_{\substack{i=1\\i\neq j}}^{n} A_i + (A_j \setminus \{t\})| = \sum_{\substack{i=1\\i\neq j}}^{n} |A_i| + |A_j \setminus \{t\}| - (n-1)|H_{a_t}| + \rho = \sum_{i=1}^{n} |A_i| - 2n + 1 + \rho,$$

where ρ is the number of H_{a_t}-holes contained collectively from the sets A_i, $i \neq j$, and from $A_j \setminus \{t\}$. Since each set A_i is aperiodic, it follows that each set A_i, $i \neq j$, contains at least one H_{a_t}-hole, and thus $\rho \geq n-1$, implying $|\sum_{\substack{i=1\\i\neq j}}^{n} A_i + A_j \setminus \{t\}| \geq \sum_{i=1}^{n}|A_i| - 2n + 1 + (n - 1) = \sum_{i=1}^{n}|A_i| - n$. However, by (6.28) we know that equality holds in this inequality, and consequently it follows that each set A_i, $i \neq j$, must contain exactly one H_{a_t}-hole, and that $A_j \setminus \{t\}$ must contain no H_{a_t}-holes. Hence each set A_i is a union of an H_{a_t}-periodic set and a disjoint element, say x. However, since $|H_{a_t}| = 2$, then adding the other element (besides x) from the H_{a_t}-coset that contains x to the set A_i will complete the coset and make the resulting set H_{a_t}-periodic. Thus each A_i is a punctured H_{a_t}-periodic set. Hence, since $\phi_{a_t}(t) \in \phi_{a_t}(A_{j'})$ (from third paragraph of Case 2), and since $t \notin A_{j'}$, it follows that $A_{j'} \cup \{t\}$ is H_{a_t}-periodic, and that if $t' \in A_j \setminus A_{j'}$, $t' \neq t$, then $\phi_{a_t}(t') \notin \phi_{a_t}(A_{j'})$.

Since every set A_i is a punctured H_{a_t}-periodic set, and since $|H_{a_t}| = 2$, it follows that $|A_i|$ is odd for every $i \leq n$. Hence, since $s' \leq 2$, it follows that $s' = 1$, and that there is no set A_i with $|A_i| = s' + 1 = 2$.

Suppose

$$|\sum_{\substack{i=1\\i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\})| \leq \sum_{i=1}^{n}|A_i| - n, \tag{6.30}$$

for distinct $t, t' \in A_j \setminus A_{j'}$, $t \neq t_0$. Hence from Kneser's Theorem, it follows that $\sum_{\substack{i=1\\i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\})$ is maximally $H_{a'}$-periodic with nontrivial period.

Suppose the inequality in (6.30) is strict. Hence, since

$$\sum_{\substack{i=1\\i\neq j}}^{n} A_i + (A_j \setminus \{t,t'\}) \subseteq \sum_{\substack{i=1\\i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\}),$$

it follows in view of (6.29) that

$$\sum_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\}) = \sum_{\substack{i=1\\ i\neq j}}^{n} A_i + (A_j \setminus \{t,t'\}).$$

Hence, in view of (6.29) and (6.28), it follows that $\sum\limits_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\})$ is a punctured H_{a_t}-periodic set. Thus from (c.5) it follows that $\sum\limits_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\})$ cannot be periodic, contradicting that $\sum\limits_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\})$ is $H_{a'}$-periodic with nontrivial period. So we may assume that (6.30) can only hold with equality.

Since $\sum\limits_{\substack{i=1\\ i\neq j}}^{n} A_i + (A_j \setminus \{t,t'\}) \subseteq \sum\limits_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t\})$, then in view of (6.29) it follows that $|\sum\limits_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t\})| \geq \sum\limits_{i=1}^{n} |A_i| - n - 1$. Suppose

$$|\sum_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t\})| > \sum_{i=1}^{n} |A_i| - n - 1.$$

Hence, since $\sum\limits_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t\}) \subseteq \sum\limits_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t\}) + (A_{j'} \cup \{t\})$, it follows in view of (6.28) that

$$\sum_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t\}) = \sum_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t\}) + (A_{j'} \cup \{t\}) = \sum_{\substack{i=1\\ i\neq j}}^{n} A_i + (A_j \setminus \{t\}).$$

Hence in view of (6.30) it follows that

$$\sum_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t\}) = \sum_{\substack{i=1\\ i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\})$$

is maximally H_{a_t}-periodic. Hence, since $\phi_{a_t}(t') \notin \phi_{a_t}(A_{j'})$ (from seventh paragraph of Case 2), since t is the only element from its H_{a_t}-coset in A_j (from second paragraph of Case 2), since $|H_{a_t}| = 2$, and since each A_i is a punctured H_{a_t}-coset (from seventh paragraph of Case 2), it follows from Kneser's Theorem (by counting holes) that $|\sum_{\substack{i=1\\i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\})| \geq \sum_{i=1}^{n}|A_i| - n + 2$, contradicting (6.30). So we may assume that $|\sum_{\substack{i=1\\i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t\})| = \sum_{i=1}^{n}|A_i| - n - 1$.

Hence, since equality holds in (6.30), it follows that $\sum_{\substack{i=1\\i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t\})$ is punctured from the $H_{a'}$-periodic set $\sum_{\substack{i=1\\i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\})$, and thus is aperiodic by (c.5). However, since $A_{j'} \cup \{t\}$ is H_{a_t}-periodic (from seventh paragraph of Case 2), it follows that $\sum_{\substack{i=1\\i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t\})$ is periodic, a contradiction. So we may assume (6.30) does not hold, i.e., that

$$|\sum_{\substack{i=1\\i\neq j,j'}}^{n} A_i + (A_j \setminus \{t,t'\}) + (A_{j'} \cup \{t,t'\})| \geq \sum_{i=1}^{n}|A_i| - n + 1, \tag{6.31}$$

for distinct $t, t' \in A_j \setminus A_{j'}$, $t \neq t_0$.

If $|A_j| - |A_{j'}| > 2$, then in view of (6.31) it follows that the set partition obtained by moving t and t' from A_j to $A_{j'}$ satisfies (6.1) and contains one less set of cardinality s', contradicting the extremal conditions originally assumed for A. Thus we may assume $|A_j| - |A_{j'}| = 2$. Hence $|A_j| = s' + 2 = 3$. Consequently, since A_j and $A_{j'}$ with $|A_j| \geq s' + 2$ and $|A_{j'}| = s'$ were arbitrary, and since there are no sets A_i with $|A_i| = s' + 1$ (from eighth paragraph of Case 2), it follows that $|A_i| = 1$ for $i < k$ and that $|A_i| = 3$ for $i \geq k \geq 2$. Thus $s = 2$, and hence applying Case 1 to the $(n - k + 1)$-set partition $A_k, A_{k+1}, \ldots, A_n$ completes the proof. So we may assume (6.27) does not hold.

Since (6.27) does not hold, then let l be the minimal integer such that, allowing re-indexing,

$$|\sum_{i=1}^{l} A_i| < \sum_{i=1}^{l} |A_i| - l + 1. \tag{6.32}$$

Hence from Kneser's Theorem it follows that $\sum_{i=1}^{l} A_i$ is maximally H_a-periodic with nontrivial period. Since $s' \leq 2$, then in view of (6.24), Proposition 1.3 and the minimality of l, it follows that $|A_i| \leq s' + 1 \leq 3$ for $i \leq l$. Hence, in view of Kneser's Theorem and the minimality of l, it follows (by counting holes) that each A_i with $i \leq l$ is contained in an H_a-coset. Thus, since $\sum_{i=1}^{l} A_i$ is H_a-periodic, it follows that $\sum_{i=1}^{l} A_i$ is an H_a-coset. Let b, s_1 and s_2 be as defined in Case 1a. Since $\sum_{i=1}^{l} A_i$ is an H_a-coset, then in view of Proposition 2.4(ii), it follows that we can remove elements from the sets in A_i with $i \leq l$, yielding new, nonempty sets A'_i, such that $s'_1 \stackrel{def}{=} \sum_{i=1}^{l} |A'_1| \leq |H_a| + l - 1$ and $\sum_{i=1}^{l} A'_1 = \sum_{i=1}^{l} A_1$.

Let S' be a minimal length subsequence of the terms of S partitioned by the A_i where $i \geq l+1$, with an $(n-l)$-set partition, $B' = B_1, \ldots, B_{n-l}$, such that $|\sum_{i=1}^{n-l} \phi_a(B_i)| \geq b+1$ (since $\sum_{i=1}^{l} A'_i$ is an H_a-coset, such a subsequence exists by (6.1) and (6.9)). In view of Proposition 2.4(ii) it follows that $|S'| \leq (n-l) + b$.

Letting $s'_2 = |S'|$, letting $r' = r$ for $s \geq 3$, and letting $r' = n - 1$ for $s = 2$, observe that the proof will be complete unless

$$s'_2 + s'_1 \geq (s-1)n + r' + 2. \tag{6.33}$$

Hence from the conclusions of the last two paragraphs, it follows that

$$(s-1)n + r' + 2 \leq |H_a| + l - 1 + (n - l) + b,$$

implying $(s-1)n \leq \frac{s-1}{s-2}(|H_a|+b-r'-3) \leq 2(|H_a|+b-r'-3)$ for $s \geq 3$, and that $n \leq |H_a|+b-2$ for $s=2$. Hence in view of (6.9), it follows that $b|H_a| \leq 2|H_a|+2b-5$, implying $(b-2)|H_a| \leq 2b-5$, whence $b \leq 1$. Since $|A_i| \leq s'+1 \leq 3$ for $i \leq l$, it follows from the minimality of l that $|A_i|=2$ or $|A_i|=3$ for all $i \leq l$. Hence, in view of (6.32), it follows that applying Proposition 2.4(ii) to the A_i with $i \leq l$ yields sets $A_i' \subseteq A_i$ such that $\sum_{i=1}^{l} A_i' = \sum_{i=1}^{l} A_i$, such that $|\sum_{i=1}^{l} A_i'| = \sum_{i=1}^{l} |A_i'| - l + 1$, such that $|A_r'| \leq 2$ for some r, and such that the conditions of Case 1 hold for the subsequence of the A_i' consisting of those A_i' with $|A_i'| > 1$. Hence, since $|A_r'| \leq 2$ for some r, then applying Case 1 it follows that we may assume that $s_1' \leq 2l$. Hence, since $b \leq 1$, and since $s_2' \leq (n-l)+b$, it follows that $s_1'+s_2' \leq n+l+1$. Thus from (6.33) it follows that $n+l+1 \geq 2n+1$, whence $n \leq l$ contradicting (6.1) or (6.32), and completing the proof. □

Part II

ZERO-SUM APPLICATIONS

Interlude: Zero-Sum Generalizations

Part II, beginning with Chapter 7, initiates the material on zero-sum applications. We remarked in Chapter 2 that the Erdős-Ginzburg-Ziv Theorem can be viewed as a generalization of the pigeonhole principle. Having discovered a zero-sum generalization for the pigeonhole principle, one of the simplest Ramsey-type extremal problems, it was natural to wonder if a similar generalization might also occur for more complex extremal questions.

One particular incarnation of this idea can be described as follows. Let L_m^t be a fixed system of inequalities (often linear) in tm variables $x_1^1, \ldots, x_m^1, x_1^2, \ldots, x_m^2, \ldots, x_1^t, \ldots, x_m^t$, let $f(L_m^t, r)$ denote the minimal integer N such that no matter how the first N integers, denoted $[1, N]$, are r-colored, say by $\Delta : [1, N] \rightarrow \{0, 1, \ldots, r-1\}$, there will always be an integer solution to L_m^t, given by $x_i^j = t_i^j$, with $\Delta(t_1^j) = \Delta(t_2^j) = \ldots = \Delta(t_m^j)$, for each $j = 1, \ldots, t$ (we call such a solution *monochromatic*). Likewise, let $f_{zs}(L_m^t, 2)$ denote the minimal integer N such that no matter how the first N integers are colored using the elements from $\mathbb{Z}/m\mathbb{Z}$, say by $\Delta : [1, N] \rightarrow \mathbb{Z}/m\mathbb{Z}$, there will always be an integer solution to L_m^t, given by $x_i^j = t_i^j$, such that $\sum_{i=1}^{m} \Delta(t_i^j) = 0$, for each $j = 1, \ldots, t$ (we call such a solution *zero-sum*). In short, the previous extremal functions involve looking for a collection of m-uniform (in number of elements) solutions to a (usually linear) system of inequalities, each individually monochromatic or zero-sum, respectively, with additional inequality relations required amongst the t solutions of size m. Note that EGZ says that $f(P_m^1, 2) = f_{zs}(P_m^1, 2) = 2m - 1$, where P_m^1 is the system $x_1 < x_2 < \ldots < x_m$ in m variables (i.e., no restriction on the variables except that they be distinct in value).

Since we are allowed to use only 0's and 1's in the coloring for $f_{zs}(L_m^t, 2)$ (in which case the m-term zero-sum solutions would be in exact correspondence with the m-term monochromatic solutions), it is easily seen that a lower bound construction for $f(L_m^t, 2)$ yields a lower bound construction for $f_{zs}(L_m^t, 2)$. Hence $f(L_m^t, 2) \leq f_{zs}(L_m^t, 2)$. On the other hand, since a monochromatic solution is always zero-sum, we have the inequality $f_{zs}(L_m^t, 2) \leq f(L_m^t, m)$. If the first inequality is an equality, i.e., $f(L_m^t, 2) = f_{zs}(L_m^t, 2)$, as it is for instance for EGZ, then we say the system L_m^t zero-sum generalizes. In essence, the system zero-sum generalizing means that best way to avoid zero-sum solutions is to avoid monochromatic solutions. One might at first think this a very unusual occurrence, particularly since there is such additional freedom when coloring with $\mathbb{Z}/m\mathbb{Z}$ versus $\{0, 1\}$; however many examples attaining equality have been found. Though no formal proof or theorem is known, it is generally believed (by at least some) that as long as the restrictions on the variables are 'sufficiently nonrestrictive,' such a zero-sum generalization will occur.

One might wonder if there is a natural way to obtain a zero-sum generalization for $f(L_m^t, r)$ when $r > 2$. The easiest and most straightforward way is to simply replace pairs of colors from $\{0, 1, \ldots, r-1\}$ with disjoint copies of $\mathbb{Z}/m\mathbb{Z}$, leaving intact an odd-person-out color (if r is odd). Formally, let $f_{zs}(L_m^t, r)$ denote the minimal integer N such that no matter how the first N integers are colored using the elements from $\lfloor \frac{r}{2} \rfloor$ disjoint copies of $\mathbb{Z}/m\mathbb{Z}$ and (if r is odd) an additional disjoint color class, say by

$$\Delta : [1, N] \to (\mathbb{Z}/m\mathbb{Z})^{(1)} \bigsqcup (\mathbb{Z}/m\mathbb{Z})^{(2)} \bigsqcup \ldots \bigsqcup (\mathbb{Z}/m\mathbb{Z})^{(\lfloor \frac{r}{2} \rfloor)},$$

if r is even, or by

$$\Delta : [1, N] \to (\mathbb{Z}/m\mathbb{Z})^{(1)} \bigsqcup (\mathbb{Z}/m\mathbb{Z})^{(2)} \bigsqcup \ldots \bigsqcup (\mathbb{Z}/m\mathbb{Z})^{(\lfloor \frac{r}{2} \rfloor)} \sqcup \{\infty\}$$

if r is odd, then there will always be an integer solution to L_m^t, given by $x_i^j = t_i^j$, such that, for each $j = 1, \ldots, t$, we have both that $\widetilde{\Delta}(t_1^j) = \widetilde{\Delta}(t_2^j) = \ldots = \widetilde{\Delta}(t_m^j)$, and that either $\Delta(t_i^j) = \infty$ for all i or else $\sum_{i=1}^{m} \Delta(t_i^j) = 0$, where $\widetilde{\Delta} : [1, N] \to \{1, \ldots, \lceil \frac{r}{2} \rceil\}$ is the coloring given by $\widetilde{\Delta}(t) = s$ if $\Delta(t) \in (\mathbb{Z}/m\mathbb{Z})^{(s)}$, and otherwise $\widetilde{\Delta}(t) = \lceil \frac{r}{2} \rceil$. Then we once more have the inequalities

$$f(L_m^t, r) \leq f_{zs}(L_m^t, r) \leq f(L_m^t, (m-1)\left\lfloor \frac{r}{2} \right\rfloor + \left\lceil \frac{r}{2} \right\rceil),$$

and there is an r-color zero-sum generalization whenever $f(L_m^t, r) = f_{zs}(L_m^t, r)$.

Very few examples of r-color zero-sum generalizations with $r > 2$ are known—due (perhaps?) to the added difficulty of such problems—but there have been a handful of examples. One might lament that this definition for an r-color zero-sum generalization is somewhat unnatural, particularly in the odd case, and is thus not entirely satisfactory. On the positive side, this 'weak' notion of zero-sum generalization is defined for every $r \geq 2$ and requires no machinery to show it is well defined. There is (sometimes) an alternative 'strong' notion of zero-sum generalization, that, though not dealt with in this thesis, is worth mentioning.

Let $\tau(m, s)$ be the maximal integer τ such that there exists a cardinality τ subset X of $\underbrace{\mathbb{Z}/m\mathbb{Z} \times \ldots \times \mathbb{Z}/m\mathbb{Z}}_{s}$ with the property that every m-term zero-sum subsequence with its terms from X must be monochromatic. Let $\kappa(m, s)$ be the minimal integer κ such that every sequence of terms from $\underbrace{\mathbb{Z}/m\mathbb{Z} \times \ldots \times \mathbb{Z}/m\mathbb{Z}}_{s}$ with length κ contains an m-term zero-sum subsequence. Observe that taking $m-1$ terms equal to each of the elements from the set X from $\tau(m, s)$ gives a lower bound for $\kappa(m, s)$. Hence $\tau(m, s)(m-1) + 1 \leq \kappa(m, s)$. That equality holds for $s = 1$ follows from EGZ, while it is a recent result of C. Reheir [51], affirming the long-standing Kemnitz Conjecture, that equality holds for $s = 2$ as well. The determination of $\kappa(m,s)$ for $s>3$ seems extremely difficult, and even the determination of $\tau(m,s)$ is quite nontrivial. However, whenever $\tau(m,s)(m-1)+1=\kappa(m,s)$, one can define $f_{sz}(L^t_m,\tau(m,s))$ to be the minimal integer N such that no matter how the first N integers are colored using the elements from $\underbrace{\mathbb{Z}/m\mathbb{Z}\times\ldots\times\mathbb{Z}/m\mathbb{Z}}_{s}$, say by $\Delta:[1,N]\rightarrow \underbrace{\mathbb{Z}/m\mathbb{Z}\times\ldots\times\mathbb{Z}/m\mathbb{Z}}_{s}$, then there will always be an integer solution to L^t_m, given by $x^j_i=t^j_i$, such that $\sum_{i=1}^{m}\Delta(t^j_i)=0$, for each $j=1,\ldots,t$. Then we have a strong $\tau(m,s)$-color zero-sum generalization whenever $f_{sz}(L^t_m,\tau(m,s))=f(L^t_m,\tau(m,s))$.

When defined, the 'strong' notion of zero-sum generalization is a perhaps more pleasing notion of an r-color zero-sum generalization. On the negative side, it only gives an r-color zero-sum generalization for r equal to some $\tau(m,s)$, which increases geometrically in s; there are no current methods that seem anywhere near sufficient to handle the associated added difficulties (the determination of $\kappa(m,s)$ just for $s=2$ was considered a major triumph in itself); and it is still unclear, though no counter examples are yet known, to what extent $\tau(m,s)(m-1)+1=\kappa(m,s)$. Regardless, this thesis deals only with 'weak' r-color zero-sum generalizations, and most examples will be in the case $r=2$. Additionally, since the second part deals only with un-weighted zero-sum questions, the sequence W in Theorems 3.1 and 3.2 is always assumed to be all 1's.

Chapter 7

A Simple 5-Color Zero-Sum Generalizing Result

7.1 Discussion

The first zero-sum generalizing result (after EGZ) was the nondecreasing diameter problem introduced by Bialostocki, Erdős and Lefmann [6], corresponding to the system ND_m^2 of inequalities given by

$$x_1 < x_2 < \cdots < x_m < y_1 < y_2 < \cdots < y_m$$

$$x_m - x_1 \leq y_m - y_1.$$

They were able to show that the system ND_m^2 admitted an r-color zero-sum generalization for $r = 2$ and $r = 3$, i.e., that $f(ND_m^2, r) = f_{zs}(ND_m^2, r)$ for $r = 2, 3$, and conjectured this to be the case for all r. Later work, using the methods of this thesis, determined that the $r = 4$ case also zero-sum generalized [23].

As an attempt to capture the belief that a system zero-sum generalizes if the constraints are sufficiently weak, Bialostocki boldly conjectured that if a linear system WL_m^2 were strictly weaker than a zero-sum generalizing linear system L_m^2, meaning that any integer solution to L_m^2 was also a solution to WL_m^2, then the weaker system WL_m^2 would also zero-sum generalize.

As a test case, A. Schultz looked at the system WND_m^2, given by

$$x_1 < x_2 < \cdots < x_m < y_1 < y_2 < \cdots < y_m$$

$$2(x_m - x_1) \leq y_m - x_1,$$

which is strictly weaker than original nondecreasing diameter system ND_m^2. Together, with minimal input from myself, the methods from the first part of this thesis were easily adapted to the results obtained by Schultz to show that the system WND_m^2 admitted an r-color zero-sum generalization for $r = 2, 3, 4, 5$.

7.2 A 5-Color Zero-Sum Generalization

We begin with some helpful notation. An *m-set*, denoted $Z = z_1, \ldots, z_m$, is a sequence of m distinct positive integers such that $z_1 < \cdots < z_m$. For a pair of m-sets X and Y, we write $X \prec Y$ if $x_m < y_1$. We also adopt the following notation:

(i) $\mathrm{int}_i(Z) = z_i$ for $i \leq m$;

(ii) $\mathrm{first}_k(Z) = \{z_1, \ldots, z_{\min\{k,m\}}\}$;

(iii) $\mathrm{last}_k(Z) = \{z_{\max\{1,m-k+1\}}, \ldots, z_m\}$.

For matters of simplicity, we abbreviate $f(WND_m^2, r)$ by $g(m, r)$, and we abbreviate $f_{zs}(WND_m^2, r)$ by $g_{zs}(m, r)$. Additionally, we use string notation for describing colorings, e.g., the coloring $\Delta : [1, 10] \to \{0, 1, 2\}$ given by $\Delta([1, 3]) = 1$, $\Delta(4) = 2$, $\Delta([5, 6]) = 0$, and $\Delta([7, 10]) = 1$, is denoted $\Delta[1, 10] = 1^3 2 0^2 1^4$.

To facilitate our evaluation of $g_{zs}(m,r)$, we make the following observation.

Observation 1. *Let* $\Delta : [1,n] \to \bigsqcup_{i=1}^{k} (\mathbb{Z}/m\mathbb{Z})^{(i)}$ *(let* $\Delta : [1,n] \to \bigsqcup_{i=1}^{k} (\mathbb{Z}/m\mathbb{Z})^{(i)} \sqcup \{\infty\}$*) be a coloring, where* $k = \lfloor \frac{r}{2} \rfloor$*. If there exists a zero-sum (zero-sum or monochromatic)* m*-set* $Y \subset [r(m-1)+2, n]$ *such that* $y_m \geq 2r(m-1)+1$*, then there exists a zero-sum (zero-sum or monochromatic) solution to the system* WND_m^2*.*

Proof. From the pigeonhole principle and EGZ, it follows that there is some zero-sum or monochromatic m-set $X \subset [1, r(m-1)+1]$. If a zero-sum or monochromatic m-set $Y \subset [r(m-1)+2, n]$ exists, then $X \prec Y$. If $y_m \geq 2r(m-1)+1$, then $y_m - x_1 \geq 2r(m-1)+1-x_1 \geq 2(r(m-1)+1-x_1) \geq 2(x_m - x_1)$. $\square$

The determination of $g_{zs}(m,2)$ is a simple application of EGZ.

Theorem 7.1. *If* $m \geq 2$ *is an integer, then* $g_{zs}(m,2) = g(m,2) = 5m-4$*.*

Proof. That $g_{zs}(m,2) \geq g(m,2) \geq 5m-4$ follows from the coloring

$$21^{m-1}2^{m-1}1^{2m-3}2^{m-1}.$$

Thus it remains to show $g_{zs}(m,2) \leq 5m-4$.

By Observation 1 it is sufficient to find a zero-sum m-set $Y \subset [2m, 5m-4]$ with $y_m \geq 4m-3$. Let $P = [3m-2, 5m-4]$. Since $|P| = 2m-1$, it follows from EGZ that there exists some zero-sum m-set $Y \subset P$. Since $|P \cap [3m-2, 4m-4]| = m-1$, it follows that $y_m \geq 4m-3$. $\square$

The determination of $g_{zs}(m,3)$ will require the machinery of Chapter 3 and the following two lemmas, which will be used to handle cases that easily reduce to trichromatic colorings.

Lemma 7. *Let $m \geq 4$ be an integer, and let $\Delta : [1, 3m-4] \to \mathbb{Z}/m\mathbb{Z} \sqcup \{\infty\}$ be a coloring. If $|\Delta^{-1}(\infty)| \geq 3m - \lceil \frac{m}{2} \rceil - 2$, then there exist monochromatic m-sets $X \prec Y$ such that $y_m - x_1 \geq 2(x_m - x_1)$.*

Proof. Let $x_i = \mathrm{int}_i(\Delta^{-1}(\infty))$, let $Y = \mathrm{last}_m(\Delta^{-1}(\infty))$, and let $x_m = m + t$ (hence t is the number of integers less than x_m not colored by ∞). Since $|\Delta^{-1}(\infty)| \geq 3m - \lceil \frac{m}{2} \rceil - 2 \geq 2m$ (since $m \geq 4$), then $X \prec Y$. Also, note that $x_m - x_1 = m + t - x_1$. Hence, if the theorem is false, then $\Delta^{-1}(\infty) \cap [2(m + t - x_1) + x_1, 3m - 4] = \emptyset$, either contradicting that there can be at most $(3m - 4) - (3m - \lceil \frac{m}{2} \rceil - 2 + t) = \lceil \frac{m}{2} \rceil - 2 - t$ integers greater than x_m not colored by ∞, or contradicting that there are at most $(3m - 4) - (3m - \lceil \frac{m}{2} \rceil - 2) = \lceil \frac{m}{2} \rceil - 2$ integers less than x_m that are not colored by ∞. □

Lemma 8. *Let $m \geq 4$ be an integer. If $\Delta : [3m-1, 7m + \lfloor \frac{m}{2} \rfloor - 6] \to [1, 3]$ is a given coloring, then either:*

(i) there exists a monochromatic m-set Y such that $y_m \geq 6m - 5$,

(ii) there exist monochromatic m-sets $W \prec Y$ such that $y_m - w_1 \geq 2(w_m - w_1)$.

Proof. Let $t = |\Delta([6m - 5, 7m + \lfloor \frac{m}{2} \rfloor - 6])|$. If $t = 3$, then either (i) follows, or else there can be at most $3(m-1) < 4m - 4 + \lfloor \frac{m}{2} \rfloor$ integers colored by Δ, a contradiction. If $t = 1$, then (i) follows by letting $Y = [6m - 5, 7m - 6]$. So we may assume $t = 2$, that w.l.o.g. $\Delta([6m - 5, 7m + \lfloor \frac{m}{2} \rfloor - 6]) = \{1, 2\}$, and that

$$|\Delta^{-1}(\{1,2\}) \cap [3m-1, 6m-6]| \leq 2(m-1) - (m + \left\lfloor \frac{m}{2} \right\rfloor) = \left\lceil \frac{m}{2} \right\rceil - 2.$$

Let $W = \mathrm{first}_m(\Delta^{-1}(3))$, let $Y = \mathrm{last}_m(\Delta^{-1}(3))$, let $t_1 = |\Delta^{-1}(\{1,2\}) \cap [w_1, w_m]|$, and let $t_2 = |\Delta^{-1}(\{1,2\}) \cap [y_m, 6m-6]|$. Since $|\Delta^{-1}(3)| \geq 3m - 4 - (\lceil \frac{m}{2} \rceil - 2) \geq 2m$, it follows that $W \prec Y$. Also, $w_m - w_1 = m - 1 + t_1$ and

$$y_m - w_1 \geq (6m - 6 - t_2) - w_1 = 2(m - 1 + t_1) + 4m - 4 - (t_2 + t_1) - t_1 - w_1 \geq$$
$$2(m - 1 + t_1) + 4m - 4 - (t_2 + t_1) - t_1 - \left((3m - 1) + \left\lceil \frac{m}{2} \right\rceil - 2 - t_1 - t_2\right) =$$
$$2(m - 1 + t_1) + \left\lfloor \frac{m}{2} \right\rfloor - 1 - t_1 \geq 2(m - 1 + t_1),$$

whence (ii) follows. □

We proceed with the proof of the 3-color case.

Theorem 7.2. *If $m \geq 4$ is an integer, then $g_{zs}(m, 3) = g(m, 3) = 7m + \lfloor \frac{m}{2} \rfloor - 6$.*

Proof. That $g_{zs}(m, 3) \geq g(m, 3) \geq 7m + \lfloor \frac{m}{2} \rfloor - 6$, follows from the coloring

$$01^{m-1}2^{m-1}0^{m-1}1^{m-\lceil \frac{m}{2} \rceil - 1}2^{\lfloor \frac{m}{2} \rfloor - 1}1^{2m-2}2^{\lceil \frac{m}{2} \rceil}0^{m-1}.$$

Next we show that $g_{zs}(m, 3) \leq 7m + \lfloor \frac{m}{2} \rfloor - 6$. Let $\Delta : [1, 7m + \lfloor \frac{m}{2} \rfloor - 6] \to \mathbb{Z}/m\mathbb{Z} \sqcup \{\infty\}$ be an arbitrary coloring. By Observation 1 it is sufficient to find a zero-sum or monochromatic m-set $Y \subset [3m - 1, 7m + \lfloor \frac{m}{2} \rfloor - 6]$ with $y_m \geq 6m - 5$. For convenience let

$$P = \Delta^{-1}(\mathbb{Z}/m\mathbb{Z}) \cap [3m - 1, 7m + \left\lfloor \frac{m}{2} \right\rfloor - 6],$$

and let S be the sequence of colors from $\mathbb{Z}/m\mathbb{Z}$ associated to P. Let $k = |\Delta^{-1}(\infty) \cap [6m - 5, 7m + \lfloor \frac{m}{2} \rfloor - 6]|$. Note $k < m$ holds, else we can trivially find a monochromatic m-set $Y \subset [3m - 1, 7m + \lfloor \frac{m}{2} \rfloor - 6]$ with $y_m \geq 6m - 5$.

Suppose $k = 0$. If $|P| \geq 2m - 1$, then one may find a zero-sum m-set $Y \subset [3m - 1, 7m + \lfloor \frac{m}{2} \rfloor - 6]$ with $y_m \geq 6m - 5$ by selecting $P' \subset P$ such that $|P'| = 2m - 1$ and

$|P' \cap [6m-5, 7m+\lfloor \frac{m}{2} \rfloor - 6]| = m$ and applying EGZ. Otherwise $|P| < 2m-1$, so that $|\Delta^{-1}(\infty) \cap [3m-1, 6m-6]| \geq 3m-4-(2m-2-(m+\lfloor \frac{m}{2} \rfloor)) = 3m - \lceil \frac{m}{2} \rceil - 2$. Shifting $[3m-1, 6m-6]$ to the interval $[1, 3m-4]$ and applying Lemma 7 completes the proof. So we may assume that $k > 0$.

Since $k > 0$, it follows that $|\Delta^{-1}(\infty) \cap [3m-1, 7m+\lfloor \frac{m}{2} \rfloor - 6]| < m$, else $\text{last}_m(\Delta^{-1}(\infty))$ will satisfy (i). Hence $|P| \geq 4m + \lfloor \frac{m}{2} \rfloor - 4 - (m-1) \geq 2m$.

Suppose that $S \setminus \Delta(\max(P))$ does not have an $(|P|-m)$-set partition. Hence, since $|S \setminus \Delta(\max(P))| = |P| - 1 \geq 2m-1$, then it follows from Proposition 2.3 that there exists $\alpha \in \mathbb{Z}/m\mathbb{Z}$ such that $\Delta(x) = \alpha$ for all but at most $m-2+1 = m-1$ elements $x \in P$. For convenience, let $H = \{x \in P | \Delta(x) \neq \alpha\}$. Induce a coloring $\Delta_e : [3m-1, 7m+\lfloor \frac{m}{2} \rfloor - 6] \to [1, 3]$ defined by

$$\Delta_e(x) = \begin{cases} 1 & \text{if } x \in P \setminus H \\ 2 & \text{if } x \in H \\ 3 & \text{if } \Delta(x) = \infty. \end{cases}$$

Note, since $|\Delta_e^{-1}(2)| < m$, that any monochromatic m-set in Δ_e is either zero-sum or monochromatic in Δ. Hence the theorem follows from Lemma 8. So we may assume that $S \setminus \Delta(\max(P))$ has a $(|P|-m)$-set partition B.

Hence, since $|S| - m \geq m$, it follows that we can apply Theorem 3.2 to the subsequence $S \setminus \Delta(\max(P))$ of the sequence $S \setminus \Delta(\max(P))$ with $(|P|-m)$-set partition B. Let $A = A_1, \ldots, A_{|P|-m}$ be the resulting set partition, and note that at most $m-1$ sets A_i in A can have cardinality greater than one. Hence we can re-index so that $|A_i| = 1$ for $i \geq m$.

Suppose Theorem 3.2(ii) holds. Hence all but at most $a-2+1 = a-1$ terms of S will be from the same H_a-coset, where H_a is a proper, nontrivial subgroup of index a. Thus, as remarked in Section 3.1, it follows from EGZ that any selection of

$$m + \frac{m}{a} - 1 + (a-1) \leq \lfloor \frac{3}{2}m \rfloor$$

terms of S (or correspondingly of P) will contain an m-term zero-sum. Since $|P \cap [6m - 5, 7m + \lfloor \frac{m}{2} \rfloor - 6]| \geq \lfloor \frac{m}{2} \rfloor + 1$ we may select $P' \subset P$ with $\lfloor \frac{3}{2}m \rfloor$ elements such that $|P' \cap [6m-5, 7m + \lfloor \frac{m}{2} \rfloor - 6]| \geq \lfloor \frac{m}{2} \rfloor + 1$. As previously noted, since $|P'| \geq \lfloor \frac{3}{2}m \rfloor$, it follows that P' must contain a zero-sum m-set Y. Since $|P' \cap [3m-1, 6m-6]| < m$ it follows that $y_m \geq 6m - 5$, completing the proof. So we may instead assume Theorem 3.2(i) holds, implying (since $|A_i| = 1$ for $i \geq m$) that

$$|\sum_{i=1}^{|P|-m} A_i| = |\sum_{i=1}^{m-1} A_i| = m.$$

Hence by an appropriate selection of terms $a_i \in A_i$, with $i \leq m-1$, it follows (since $k < m$) that there exists a zero-sum m-set Y with $y_m = \max(P) \geq 6m - 5$. $\square$

Next we consider the evaluation of $g_{zs}(m, 4)$. Towards that end we need the following lemma, which will be used to handle cases that easily reduce to quadrichromatic colorings.

Lemma 9. *Let $m \geq 3$ be an integer. If $\Delta : [4m-2, 10m-9] \to [1,4]$ is a coloring, then either:*

(i) there exists a monochromatic m-set Y such that $y_m \geq 8m - 7$,

(ii) there exist monochromatic m-sets $W \prec Y$ such that $y_m - w_1 \geq 2(w_m - w_1)$.

Proof. Since $|[8m-7, 10m-9]| = 2m - 1$, it follows that $|\Delta([8m-7, 10m-9])| > 2$, else (i) follows from the pigeonhole principle. Thus, since $|\Delta^{-1}(j)| \leq m - 1$ must hold for each color j used in the interval $[8m-7, 10m-9]$, as (i) follows otherwise, and since

$|[4m-2, 10m-9]| = 6m-6 > 4m-4$, it follows that $|\Delta([8m-7, 10m-9])| = 3$. We may w.l.o.g. assume 4 is the color class with $|\Delta^{-1}(4)| \geq m$. Since 4 is not used to color any of the $2m-1$ integers in $[8m-7, 10m-9]$, and since each of the remaining three colors is used at most $m-1$ times in the interval $[4m-2, 10m-9]$, it follows that

$$|\Delta^{-1}(\{1,2,3\}) \cap [4m-2, 8m-8]| \leq (3m-3) - (2m-1) = m-2,$$

and (since $m \geq 3$) that

$$|\Delta^{-1}(4) \cap [4m-2, 8m-8]| \geq (4m-5) - (m-2) = 3m-3 \geq 2m.$$

Let $W = \text{first}_m(\Delta^{-1}(4))$ and let $Y = \text{last}_m(\Delta^{-1}(4))$. Then $w_m - w_1 \leq m-1+t$ and $y_m - w_1 \geq 4m-6-(m-2-t)$, where

$$t = |\Delta^{-1}(\{1,2,3\}) \cap [\min(Y), \max(Y)]| \leq |\Delta^{-1}(\{1,2,3\}) \cap [4m-2, 8m-8]| \leq m-2,$$

from which it follows that (ii) holds. □

Theorem 7.3. *If $m \geq 3$ is an integer, then $g_{zs}(m,4) = g(m,4) = 10m-9$.*

Proof. That $g_{zs}(m,4) \geq g(m,4) \geq 10m-9$, follows from the coloring

$$41^{m-1}2^{m-1}3^{m-1}4^{m-1}1^{m-3}2^{m-1}1^{2m-1}3^{m-1}4^{m-1}.$$

Next we show that $g_{zs}(m,4) \leq 10m-9$. Let $\Delta : [1, 10m-9] \to (\mathbb{Z}/m\mathbb{Z})^{(1)} \bigsqcup (\mathbb{Z}/m\mathbb{Z})^{(2)}$ be an arbitrary coloring, and let $\widetilde{\Delta} : [1, 10m-9] \to \{1,2\}$ be the coloring given by $\widetilde{\Delta}(x) = i$ for $\Delta(x) \in (\mathbb{Z}/m\mathbb{Z})^{(i)}$. By Observation 1 it is sufficient to find a zero-sum m-set $Y \subset$

$[4m-2, 10m-9]$ with $y_m \geq 8m-7$.

Since $|[8m-7, 10m-9]| = 2m-1$, we may w.l.o.g. assume $|\widetilde{\Delta}^{-1}(2) \cap [8m-7, 10m-9]| = m+k$ where $k \geq 0$. If $|\widetilde{\Delta}^{-1}(2) \cap [4m-2, 8m-8]| \geq m-1-k$, then by EGZ there exists a zero-sum m-set $Y \subset [4m-2, 10m-9]$ with $y_m \geq 8m-7$. So we may assume

$$|\widetilde{\Delta}^{-1}(2) \cap [4m-2, 8m-8]| \leq m-2-k. \tag{7.1}$$

Hence $k \leq m-2$. Letting $P = \widetilde{\Delta}^{-1}(1) \cap [4m-2, 10m-9]$, and letting S be the corresponding sequence of colors associated to P, it follows from (7.1) that $|P \cap [4m-2, 8m-8]| \geq 4m-5-(m-2-k) = 3m-3+k$, implying (since $0 \leq k \leq m-2$) that $|P| \geq 3m-2+k \geq 2m$.

Suppose that there does not exist a $(|P|-m)$-set partition of $S \setminus \Delta(\max(P))$. Hence, since $|S \setminus \Delta(\max(P))| = |P|-1 \geq 2m-1$, then it follows from Proposition 2.3 that there exists $\alpha \in \mathbb{Z}/m\mathbb{Z}$ such that $\Delta(x) = \alpha$ for all but at most $m-2+1 = m-1$ elements $x \in P$. For convenience, let $H = \{x \in P | \Delta(x) \neq \alpha\}$. Induce a coloring $\Delta_e : [4m-2, 10m-9] \to [1, 4]$ defined by

$$\Delta_e(x) = \begin{cases} 1 \ \text{ if } x \in P \setminus H \\ 2 \ \text{ if } x \in H \\ 3 \ \text{ if } x \in \text{first}_{m-1}(\widetilde{\Delta}^{-1}(2) \cap [4m-2, 10m-9]) \\ 4 \ \text{ if } x = \text{int}_i(\widetilde{\Delta}^{-1}(2) \cap [4m-2, 10m-9]), \ \text{with } i \geq m. \end{cases}$$

Note, since $|\Delta^{-1}(j)| \leq m-1$ for each $j \neq 1$ (since $|\widetilde{\Delta}^{-1}(2)| \leq 2m-2$), that any monochromatic m-set X in Δ_e is zero-sum in Δ. Hence the theorem follows from Lemma 9. So we may assume that $S \setminus \Delta(\max(P))$ has an $(|P|-m)$-set partition B.

Hence, since $|S|-m \geq m$, it follows that we can apply Theorem 3.2 to the subsequence

$S \setminus \Delta(\max(P))$ of the sequence $S \setminus \Delta(\max(P))$ with $(|P|-m)$-set partition B. Let $A = A_1, \ldots, A_{|P|-m}$ be the resulting set partition, and note that at most $m-1$ sets A_i in A can have cardinality greater than one. Hence we can re-index so that $|A_i| = 1$ for $i \geq m$.

Suppose Theorem 3.2(ii) holds. Hence all but at most $a-2+1 = a-1$ terms of S will be from the same H_a-coset, where H_a is a proper, nontrivial subgroup of index a. Thus, as remarked in Section 3.1, it follows from EGZ that any selection of $m + \frac{m}{a} + a - 2 \leq \lfloor \frac{3}{2}m \rfloor$ terms of S (or correspondingly of P) will contain an m-term zero-sum. However, the proof will be complete if there is an m-term zero-sum subset Y with $y_m \geq 8m-7$. Hence, since $|\widetilde{\Delta}^{-1}(1) \cap [8m-7, 10m-9]| \geq 2m-1-(m+k) = m-1-k$, we may assume that any set S of $\lfloor \frac{3}{2}m \rfloor - (m-1-k) = \lfloor \frac{m}{2} \rfloor + k + 1$ elements from $\widetilde{\Delta}^{-1}(1) \cap [4m-2, 8m-8]$ contains a zero-sum m-set, since otherwise $W \cup (\widetilde{\Delta}^{-1}(1) \cap [8m-7, 10m-9])$ must contain a zero-sum m-set with $y_m \geq 8m-7$. Hence $k \geq \lceil \frac{m}{2} \rceil - 1$, and, since $|P \cap [4m-2, 8m-8]| \geq 3m-3+k$, it follows that there exists some zero-sum m-set $W \subseteq \text{first}_{\lfloor \frac{m}{2} \rfloor + k+1}(\widetilde{\Delta}^{-1}(1) \cap [4m-2, 8m-8])$.

From (7.1) it follows that

$$t = |\widetilde{\Delta}^{-1}(2) \cap [4m-2, w_m]| \leq m-2-k, \tag{7.2}$$

so that $w_m - w_1 \leq \lfloor \frac{m}{2} \rfloor + k + t - t'$, where $w_1 = (4m-2) + t'$. Hence, if there exists a zero-sum m-set Y such that $W \prec Y$ and

$$y_m \geq 2(\left\lfloor \frac{m}{2} \right\rfloor + k + t) + 4m - 2 \geq 2(\left\lfloor \frac{m}{2} \right\rfloor + k + t - t') + 4m - 2 + t' \geq 2(w_m - w_1) + w_1, \tag{7.3}$$

then the proof will be complete. Taking $Y' = \text{last}_{\lfloor \frac{m}{2} \rfloor + k+1}(\widetilde{\Delta}^{-1}(1) \cap [4m-2, 8m-8])$, we see that there exists an m-set $Y \subset Y'$ that satisfies these requirements as follows. First, it is quickly verified from (7.1) that there are at least $2(\lfloor \frac{m}{2} \rfloor + k + 1)$ many elements from

$\widetilde{\Delta}^{-1}(1)$ in $[4m-2, 8m-8]$ for every $m \geq 3$. Hence, we have $W' \prec Y'$, from which it follows that $W \prec Y$. Second, we note that it follows in view of (7.1) and (7.2) that

$$y_m \geq 8m - 8 - |Y' \setminus Y| - (|\widetilde{\Delta}^{-1}(2) \cap [4m-2, 8m-8]| - t) \geq$$
$$8m - 8 - \Big((\Big\lfloor \frac{m}{2} \Big\rfloor + k + 1) + (m - 2 - k) - m \Big) - ((m - 2 - k) - t) =$$
$$6m + \Big\lceil \frac{m}{2} \Big\rceil - 5 + k + t \geq 2(\lfloor \frac{m}{2} \rfloor + (k+t)) + 4m - 2 + (m - 2 - (k+t)) - 1 + \Big\lceil \frac{m}{2} \Big\rceil \geq$$
$$2(\Big\lfloor \frac{m}{2} \Big\rfloor + (k+t)) + 4m - 2.$$

Hence (7.3) is satisfied. So we may instead assume that Theorem 3.2(i) holds, implying (since $|A_i| = 1$ for $i \geq m$) that

$$|\sum_{i=1}^{|P|-m} A_i| = |\sum_{i=1}^{m-1} A_i| = m.$$

Hence by an appropriate selection of terms $a_i \in A_i$, with $i \leq m-1$, it follows (since $k \leq m-2$) that there exists a zero-sum m-set Y with $y_m = \mathrm{last}(P) \geq 8m - 7$, completing the proof. □

The evaluation of $g_{zs}(m, 5)$ is actually simpler than the 3 and 4 color cases and only requires the following two lemmas, which handle cases that reduce to dichromatic or pentachromatic colorings.

Lemma 10. *Let $m \geq 2$ be an integer. If $\Delta : [5m-3, 10m-10] \to [1, 2]$ is a coloring with $|\Delta^{-1}(c)| \leq m-2$ for some $c \in [1, 2]$, then there exist monochromatic m-sets $X \prec Y$ such that $y_m - x_1 \geq 2(x_m - x_1)$.*

Proof. Assume w.l.o.g. that $c = 2$, and let $X = \mathrm{first}_m(\Delta^{-1}(1))$ and $Y = \mathrm{last}_m(\Delta^{-1}(1))$. Since $|\Delta^{-1}(2)| \leq m-2$, it follows that $|\Delta^{-1}(1)| \geq 5m - 6 - (m-2) = 4m - 4 \geq 2m$, whence

$X \prec Y$. Also, since $|\Delta^{-1}(2)| \leq m-2$, it follows that $x_m - x_1 \leq m-1+m-2 = 2m-3$ while $y_m - x_1 \geq 5m-7-(m-2) \geq 2(2m-3)$. $\square$

Lemma 11. *Let $m \geq 2$ be an integer. If $\Delta : [5m-3, 13m-12] \to [1,5]$ is a given coloring, then either:*

(i) there exists a monochromatic m-set Y such that $y_m \geq 10m-9$,

(ii) there exist monochromatic m-sets $W \prec Y$ with $y_m - w_1 \geq 2(w_m - w_1)$.

Proof. The argument is almost identical to that of Lemma 9. Since $|[10m-9, 13m-12]| = 3m-2$, it follows that $|\Delta([8m-7, 10m-9])| > 3$, else (i) follows from the pigeonhole principle. Thus, since $|\Delta^{-1}(j)| \leq m-1$ must hold for each color j used in the interval $[10m-9, 13m-12]$, as (i) follows otherwise, and since $|[5m-3, 13m-12]| = 8m-8 > 5m-5$, it follows that $|\Delta([10m-9, 13m-12])| = 4$. We may w.l.o.g. assume 5 is the color class with $|\Delta^{-1}(5)| \geq m$. Since 5 is not used to color any of the $3m-2$ integers in $[10m-9, 13m-12]$, and since each of the remaining four colors is used at most $m-1$ times in the interval $[5m-3, 13m-12]$, it follows that

$$|\Delta^{-1}(\{1,2,3,4\}) \cap [5m-3, 10m-10]| \leq (4m-4) - (3m-2) = m-2,$$

and that

$$|\Delta^{-1}(5) \cap [5m-3, 10m-10]| \geq (5m-6) - (m-2) = 4m-4 \geq 2m.$$

Let $W = \mathrm{first}_m(\Delta^{-1}(5))$ and let $Y = \mathrm{last}_m(\Delta^{-1}(5))$. Then $w_m - w_1 \leq m-1+t$ and

$y_m - w_1 \geq 5m - 7 - (m - 2 - t)$, where

$$t = |\Delta^{-1}(\{1,2,3,4\}) \cap [\min(Y), \max(Y)]| \leq |\Delta^{-1}(\{1,2,3,4\}) \cap [5m-3, 10m-10]| \leq m-2,$$

from which it follows that (ii) holds. □

We conclude the chapter with the proof of the 5-color zero-sum generalizing case.

Theorem 7.4. *If $m \geq 2$ is an integer, then $g_{zs}(m,5) = g(m,5) = 13m - 12$.*

Proof. That $g_{zs}(m,5) \geq g(m,5) \geq 13m - 12$ follows from the coloring

$$51^{m-1}2^{m-1}3^{m-1}4^{m-1}5^{m-1}1^{m-1}2^{m-1}1^{2m-3}2^{m-1}3^{m-1}4^{m-1}5^{m-1}.$$

Thus it remains to show $g_{zs}(m,5) \leq 13m - 12$.

Let $\Delta : [1, 13m-12] \to (\mathbb{Z}/m\mathbb{Z})^{(1)} \bigsqcup (\mathbb{Z}/m\mathbb{Z})^{(2)} \sqcup \{\infty\}$ be an arbitrary coloring. By Observation 1 it is sufficient to find a zero-sum or monochromatic m-set $Y \subset [5m-3, 13m-12]$ with $y_m \geq 10m-9$. If $|\Delta^{-1}(\infty) \cap [10m-9, 13m-12]| \geq m$, then this will be the case trivially, so we may assume otherwise. Hence, it follows w.l.o.g. that

$$|\widetilde{\Delta}^{-1}(2) \cap [10m-9, 13m-12]| \geq m. \tag{7.4}$$

We break the proof into two cases based on whether or not the color class ∞ is used in the interval $[10m-9, 13m-12]$.

Case 1 ($|\Delta^{-1}(\infty) \cap [10m-9, 13m-12]| = 0$): Suppose $\widetilde{\Delta}([10m-9, 13m-12]) = \{1, 2\}$. Hence it follows that either $|\widetilde{\Delta}^{-1}(2) \cap [10m-9, 13m-12]| \geq 2m-1$ or else $|\widetilde{\Delta}^{-1}(1) \cap [10m-9, 13m-12]| \geq m$. In the former case, the proof is complete by EGZ. So we may instead assume, in view of (7.4), that $|\widetilde{\Delta}^{-1}(j) \cap [10m-9, 13m-12]| \geq m$ holds for $j = 1, 2$.

Suppose there exists a subset $S \subseteq \widetilde{\Delta}^{-1}(\{1,2\}) \cap [5m-3, 10m-10]$ with $|S| = m-1$. Thus, in view of $|\Delta^{-1}(\infty) \cap [10m-9, 13m-12]| = 0$, it follows that

$$|S \cup (\widetilde{\Delta}^{-1}(\{1,2\}) \cap [10m-9, 13m-12])| \geq m - 1 + (3m-2) = 4m-3.$$

Hence from the pigeonhole principle it follows that there is a cardinality $2m-1$ subset $S' \subset S \cup (\widetilde{\Delta}^{-1}(\{1,2\}) \cap [10m-9, 13m-12])$ either with $\widetilde{\Delta}(S') = 1$ or with $\widetilde{\Delta}(S') = 2$. Since $\text{int}_m(S) \geq 10m-9$, it follows that $\text{int}_m(S') \geq 10m-9$. Hence applying EGZ to S' yields a zero-sum m-set $Y \subseteq S'$ with $y_m \geq 10m-9$, completing the proof. So we may assume that $|\widetilde{\Delta}^{-1}(\{1,2\}) \cap [5m-3, 10m-10]| \leq m-2$.

Hence we may induce a coloring $\Delta_e : [5m-3, 10m-10] \to [1, 2]$ defined by

$$\Delta_e(x) = \begin{cases} 1 & \text{if } \Delta(x) = \infty \\ 2 & \text{if } x \in \widetilde{\Delta}^{-1}(\{1,2\} \cap [5m-3, 10m-10]. \end{cases}$$

Since any monochromatic m-set in Δ_e is also monochromatic in Δ, the result follows by Lemma 10.

Case 2 ($|\Delta^{-1}(\infty) \cap [10m-9, 13m-12]| > 0$): Suppose $|\Delta^{-1}(\infty) \cap [10m-9, 13m-12]| > 0$. Let $k = |(\Delta^{-1}(\infty) \cup \widetilde{\Delta}^{-1}(2)) \cap [10m-9, 13m-12]|$. Note that $k < 3m-2$, since otherwise the desired monochromatic or zero-sum subset Y with $y_m \geq 10m-9$ follows readily from the pigeonhole principle and EGZ.

Since $|\Delta^{-1}(\infty) \cap [10m-9, 13m-12]| > 0$, and since $|\widetilde{\Delta}^{-1}(2) \cap [10m-9, 13m-12]| \geq m$, it follows that

$$|(\Delta^{-1}(\infty) \cup \widetilde{\Delta}^{-1}(2)) \cap [5m-3, 10m-10]| < 3m-2-k, \tag{7.5}$$

since otherwise it follows, in view of EGZ and the pigeonhole principle (similar to the argument from the second paragraph of Case 1), that there exists either a zero-sum or monochromatic m-set Y with $y_m \in [10m-9, 13m-12]$. Likewise, we may assume that any set S of

$$\begin{aligned} 2m-1-|\widetilde{\Delta}^{-1}(1) \cap [10m-9, 13m-12]| &= 2m-1-(3m-2-k) \\ &= k+1-m \end{aligned}$$

elements from $\widetilde{\Delta}^{-1}(1) \cap [5m-3, 10m-10]$ contains a zero-sum m-set, since otherwise from EGZ it follows that $S \cup (\widetilde{\Delta}^{-1}(1) \cap [10m-9, 13m-12])$ will contain a zero-sum m-set with $y_m \geq 10m-9$. Consequently, $k \geq 2m-1$.

Since $k \leq 3m-3 \leq 4m-5$, then it follows from (7.5) that

$$|\widetilde{\Delta}^{-1}(1) \cap [5m-3, 10m-10]| \geq 2m-3+k \geq 2(k+1-m). \tag{7.6}$$

Thus, letting $W' = \text{first}_{k+1-m}(\widetilde{\Delta}^{-1}(1) \cap [5m-3, 10m-10])$, it follows from the conclusion of the previous paragraph that there exists a zero-sum m-set $W \subset W'$. From (7.5) we see that

$$t = |(\Delta^{-1}(\infty) \cup \widetilde{\Delta}^{-1}(2)) \cap [5m-3, w_m]| \leq 3m-3-k, \tag{7.7}$$

so that $w_m - w_1 \leq k - m + t - t'$, where $w_1 = 5m - 3 + t'$, with $t' \geq 0$. Hence, since

$$3m + 2k + 2t - 3 \geq 2(k - m + t - t') + 5m - 3 + t' \geq 2(k - m + t - t') + w_1,$$

we see that if there exists a zero-sum m-set Y such that $W \prec Y$ and such that

$$y_m \geq 3m + 2k + 2t - 3, \tag{7.8}$$

then the proof will be complete. Taking $Y' = \text{last}_{k+1-m}(\widetilde{\Delta}^{-1}(1) \cap [5m - 3, 10m - 10])$, we see that there exists an m-set $Y \subset Y'$ that satisfies these requirements as follows. First, it follows from (7.6) that $W' \prec Y'$, whence $W \prec Y$. Second, since $k + t \leq 3m - 3$ (from (7.7)), then it follows from (7.5) that

$$y_m \geq 10m - 10 - |Y' \setminus Y| - (|(\Delta^{-1}(\infty) \cup \widetilde{\Delta}^{-1}(2)) \cap [5m - 3, 10m - 10]| - t) \geq$$

$$10m - 10 - ((k + 1 - m) - m)) - ((3m - 3 - k) - t) =$$

$$9m + t - 8 \geq 9m + t - 8 - (3m - 3 - k) = 6m + k + t - 5 \geq$$

$$3m + (3m - 3) + k + t - 3 \geq 3m + (k + t) + (k + t) - 3 = 3m + 2k + 2t - 3.$$

Hence (7.8) is satisfied, completing the proof. $\square$

Chapter 8

Distinct Terms and Subsequence Sums

8.1 Discussion

We begin by introducing the notation $n \wedge S$ to denote the set of elements that can be represented as a sum of some n-term subsequence of a sequence S. Then EGZ can be rephrased to state $0 \in m \wedge S$ for a sequence S of $2m - 1$ terms from an abelian group of order m. One might wonder why zero is special, and what conditions might instead imply $g \in m \wedge S$ for a nonzero g. Informally, one explanation of zero's distinctiveness is that it is contained in every subgroup. Thus, though it is often not possible to always find $g \in m \wedge S$ no matter how large $|S|$ becomes, one might still hope for there to be some nontrivial subgroup H such that $g \in m \wedge S$ for each $g \in H$, provided $|S|$ is large enough and no term has too high a multiplicity (to avoid such degenerate sequences as those with only one distinct term).

In 1977, J. E. Olson proved a special case of Theorem 3.1; namely if S is a sequence of $2m - 1$ terms from an abelian group G of order m with every term having multiplicity at most m, then either $m \wedge S = G$, or there exists a proper, nontrivial subgroup H_a of index a such that $H_a \subseteq m \wedge S$, and all but at most $a - 2$ terms of S are from the same

H_a-coset [50]. Unfortunately, while the conclusion of Olson's Theorem was quite strong, including a structure restriction on the sequence S, it failed to cover sequences with length smaller than $2m-1$. In an effort to alleviate this restriction, Bollobás and Leader obtained a weaker version of Olson's result valid for sequences of any length; they showed that if $0 \notin m \wedge S$, then $|m \wedge S| \geq |S| - m + 1$ [8] (also an immediate consequence of Theorem 3.1). In [5], Bialostocki and Dierker proceeded to address the question of tightness in the Erdős-Ginzburg-Ziv Theorem and showed that if there were at least three distinct terms in a sequence S from the cyclic group $\mathbb{Z}/m\mathbb{Z}$, and if $|S| = 2m-2$, then $0 \in m \wedge S$. In the case of m prime, Bialostocki and Lotspeich generalized the previous result by showing that $|S| = 2m - k + 1$ guaranteed an m-term zero-sum in a sequence S with at least k distinct terms [4]. Hamidoune, Ordaz and Ortuño extended this result, in the weak Olson sense (i.e., without the structural coset condition), by showing that if $|S| = 2m - k + 1$, and if every term of S has multiplicity at most $m - k + 2$, then there exists a nontrivial subgroup H_a such that $H_a \subseteq m \wedge S$ [37]. In an attempt to further generalize the result to sequences of smaller length along lines of the Bollobás-Leader result, Hamidoune made the following conjecture [38].

Conjecture 8.1. *Let G be a cyclic group of order m, and let S be a sequence of terms from G with $|S| \geq m+1$ and at least k distinct terms. If the multiplicity of every term of S is at most $m-k+2$, then either*

(i) $|m \wedge S| \geq |S| - m + k - 1$,

(ii) there exists a nontrivial subgroup H_a such that $H_a \subseteq m \wedge S$.

Hamidoune was able to prove a weakened form of Conjecture 8.1, where the inequality in (i) was replaced by $|m \wedge S| \geq |S| - m + k - 2$, and additionally showed that result to be valid for abelian groups with cyclic or trivial 2-torsion subgroup [38].

The main result of this chapter is Theorem 8.2, which confirms Conjecture 8.1 for an arbitrary abelian group, and which gives a more complete generalization of Olson's result [50] in that it includes the corresponding structural coset condition on S. Theorem 8.2 also implies that if $|m \wedge S| < |S| - m + k - 1$, then $m \wedge S$ is periodic, a conclusion similar to the classical result of Kneser for sumsets from Chapter 1.

Theorem 8.2. *Let G be an abelian group of order m, and let S be a sequence of terms from G that has at least k distinct terms. If $|S| \geq m + 1$ and the multiplicity of each term of S is at most $m - k + 2$, then either:*

(i) $|m \wedge S| \geq \min\{m, |S| - m + k - 1\}$,

(ii) there exists a proper, nontrivial subgroup H_a of index a, such that $m \wedge S$ is H_a-periodic and $H_a \subseteq m \wedge S$, and there exists $\alpha \in G$, such that the coset $\alpha + H_a$ contains all but e terms of S, where $e \leq \min\{\lfloor \frac{|S|-m+k-2}{|H_a|} \rfloor - 1, a - 2\}$ and $|m \wedge S| \geq (e+1)|H_a|$.

8.2 Subsequence Sums

For conceptual convenience the proof of Theorem 8.2 has been divided into three sections labelled Steps 1, 2 and 3. The goal of the first is to achieve the conditions needed to apply Theorem 3.2. The goal of the second is to complete the proof minus the conclusion that $m \wedge S$ is H_a-periodic, which will then be achieved in Step 3 by an extremal argument using the results from Step 2.

Proof. Since $m \wedge S = |S| \wedge S - (|S| - m) \wedge S$ holds trivially, and since $|1 \wedge S| \geq k$, it follows that (i) holds for $|S| = m + 1$. So assume $|S| \geq m + 2$.

Step 1. Let $\epsilon = \max\{0, |S| - (2m - k + 1)\}$, let T be a subsequence of S consisting of k distinct terms including a term of S with greatest multiplicity, let $S_0 = S \setminus T$, let

$n = |S| - m$, let $n_0 = |S| - m - 1$, and let $n_1 = m - k + 1 + \epsilon$. Note that

$$\frac{|S_0| - n_1 - 1}{n_0} + 1 = \frac{|S| - m - 2 - \epsilon}{|S| - m - 1} + 1 < 2. \tag{8.1}$$

If there exists a subset $X \subseteq G$ such that $|X| = 1$ and at least $(n_1+1) = m-k+2+\epsilon$ terms of S_0 are from X, then, since the multiplicity of every term of S is at most $m - k + 2$, and since T contains a term of S with greatest multiplicity, it follows that $\epsilon = 0$ and that there are two terms of S with multiplicity $m-k+2$, whence $|S| \geq 2(m-k+2)+k-2 = 2m-k+2$, contradicting $\epsilon = 0$. So we may assume no such subset X exists. Hence, since $|S| \geq m + 2$, then in view of (8.1) and Proposition 2.3 applied to S_0, it follows that there exists an n_1-set partition $P_2, P_3, \ldots, P_{n_1+1}$ of S_0 with $|P_i| = 1$ for $i > n_0 + 1 = n$. Letting $P = P_1, \ldots, P_n$, where $P_1 = T$, and letting S' be the subsequence that P partitions, we obtain an n-set partition of the subsequence S' of S with $|S'| = |S| - (n_1 - n_0) = 2|S| - 2m + k - 2 - \epsilon$.

Apply Theorem 3.1 to the subsequence S' of S with n-set partition P, and let $A = A_1, \ldots, A_n$ be the resulting n-set partition, and H_a the corresponding subgroup. Hence, since $m \wedge S = |S| \wedge S - (|S| - m) \wedge S$, then from Theorem 3.1 it follows that we may assume,

$$((N-1)(|S| - m) + e + 1)|H_a| \leq \min\{|S| - m + k - 2,\, m - 1\}, \tag{8.2}$$

where $e = E(A, H_a)$ and $N = N(A, H_a)$, since otherwise (i) holds. Hence H_a is a proper subgroup. Observe that $|S'| - (|S| - m) + 1 \geq \min\{m,\, |S| - m + k - 1\}$. Let l be the number of distinct terms x of S such that $\phi_a(x)$ is an H_a-exception in A. Observe that $e \geq l$ and that

$$\frac{k - l}{|H_a|} \leq N, \tag{8.3}$$

hold trivially. Since $|S'| - (|S| - m) + 1 \geq \min\{m, |S| - m + k - 1\}$, then from (8.2) it follows that we may assume H_a is nontrivial and $N \geq 1$.

Let $k - |H_a| = l + \delta$, and suppose $\delta \geq 1$. Hence (8.3) implies $N|H_a| \geq |H_a| + \delta$. Thus, since $|S| \geq m + 1$, since $e \geq l$, and since $\delta \geq 1$, it follows from (8.2) that

$$k \geq (\delta - 1)(|S| - m) + |H_a|(l + 1) + 2 \geq \delta - 1 + |H_a| + l + 2 = |H_a| + l + \delta + 1,$$

contradicting the definition of δ. So we may assume

$$k - |H_a| \leq l. \tag{8.4}$$

Suppose $N > 1$. Hence (8.2), $|S| \geq m + 1$, and $e \geq l$ imply

$$(|S| - m)(|H_a| - 1) + (l + 1)|H_a| \leq k - 2,$$

which, since (8.4) implies $|H_a|(l + 1) \geq l + |H_a| \geq k$, since $|S| \geq m + 1$, and since $|H_a| \geq 2$, is impossible. So we may assume $N = 1$.

Suppose that $|S| < m + |H_a| + e$. Hence from $N = 1$ and (8.2) it follows that $e|H_a| - e \leq k - 3$. Thus, since $e \geq l$, it follows from (8.4) that $e(|H_a| - 2) \leq |H_a| - 3$, which is only possible if $e = 0$. However, if $e = 0$, then every term of S is from the same H_a-coset, say $\alpha + H_a$, and by translation we may w.l.o.g. assume $\alpha = 0$. Hence, since $\sum_{i=1}^{n} A_i$ is H_a-periodic, and since $N = 1$, it follows that $H_a \subseteq (|S| - m) \wedge S$. Since every term of S is from H_a, it follows that $|S| \wedge S \in H_a$. Thus, since $H_a \subseteq (|S| - m) \wedge S$, and since $m \wedge S = |S| \wedge S - (|S| - m) \wedge S$, it follows that $H_a \subseteq m \wedge S$. Hence, since (8.2) implies that $e \leq \min\{\lfloor \frac{|S| - m + k - 2}{|H_a|} \rfloor - 1,\, a - 2\}$, and since $e = 0$ implies $m \wedge S \subseteq H_a$, it follows that (ii) holds. So we may assume that

$$|S| \geq m + |H_a| + e. \tag{8.5}$$

Since $e \geq l$, then it follows in view of (8.5) and (8.4) that

$$|S| \geq m + k. \tag{8.6}$$

Suppose that $n < \frac{|S'|-n+1}{p} - 1$, where p is the smallest prime divisor of m. Hence, since $n = |S| - m$, and since $|S'| = 2|S| - 2m + k - 2 - \epsilon$, it follows that

$$|S| - m < \frac{|S| - m + k - 1 - \epsilon}{p} - 1. \tag{8.7}$$

Since $p \geq 2$, and since $|S| \geq m + 1$, it follows from (8.7) that $|S| - m < \frac{|S|-m+k-1-\epsilon}{2} - 1$, implying that $|S| < m + k - 3 - \epsilon$, a contradiction to (8.6). So we may assume that $n \geq \frac{|S'|-n+1}{p} - 1$.

Step 2. Since $n \geq \frac{|S'|-n+1}{p} - 1$, it follows that we can apply Theorem 3.2 to the subsequence S' of S with n-set partition A. If Theorem 3.2(i) holds, then, since $m \wedge S = |S| \wedge S - (|S| - m) \wedge S$, it follows that (i) holds. So assume that Theorem 3.2(ii) holds with proper, nontrivial subgroup H_b of index b, with coset $\beta + H_b$, with e' terms of S not from $\beta + H_b$, and with n-set partitions $A' = A'_1, \ldots, A'_n$ and $B = B_1, \ldots, B_n$, where $|\sum_{i=1}^{n} A'_i| \geq (e' + 1)|H_b|$ and $\sum_{i=1}^{n} B_i = n\beta + H_b$. Hence the inequality

$$k - |H_b| \leq l', \tag{8.8}$$

holds trivially, where l' is the number of distinct terms of S not from the coset $\beta + H_b$, and the inequality in Theorem 3.2(ii) implies

$$e' \leq \min\left\{\left\lfloor\frac{|S|-m+k-2}{|H_b|}\right\rfloor - 1,\, b-2\right\}. \tag{8.9}$$

We may w.l.o.g. assume $\beta = 0$. Hence, since $\sum\limits_{i=1}^{n} B_i = H_b$, it follows that $H_b \subseteq (|S|-m) \wedge S$. Thus, if $e' = 0$, then $|S| \wedge S \in H_b$ and $m \wedge S \subseteq H_b$, whence (ii) follows from (8.9) and $m \wedge S = |S| \wedge S - (|S|-m) \wedge S$. So $e' > 0$. Since there are at most $n + |H_b| - 1$ terms partitioned by the set partition B, it follows in view of (8.9) that there are at least

$$(e'+1)|H_b| + m - k + 2 - e' - (n + |H_b| - 1) = 2m - |S| - k + 3 + e'(|H_b| - 1) \tag{8.10}$$

terms of S from $\beta + H_b$ that are not partitioned by B.

Hence if there are at most $2m-|S|-1$ terms of S from $\beta+H_b$ that are not partitioned by B, then since $e' > 0$, and since $e' \geq l'$, it follows in view of (8.10) that $k-4 \geq e'(|H_b|-1) \geq e' + |H_b| - 2 \geq l' + |H_b| - 2$, contradicting (8.8). Consequently we may assume that there are at least $2m - |S| = m - n$ terms of S from $\beta + H_b$ that are not partitioned by B. Thus (provided $m \geq n$) we can add $m-n$ singleton sets, each containing a term of S from $\beta + H_b$ not partitioned by B, to the set partition B, to obtain an m-set partition whose sumset is H_b. Hence

$$H_b \subseteq m \wedge S, \tag{8.11}$$

if $m \geq n$, but if $n > m$, then (8.11) follows instead from Proposition 2.4. So we can assume (8.11) holds regardless.

Step 3. In view of $|\sum\limits_{i=1}^{n} A'_i| \geq (e'+1)|H_b|$, (8.8), (8.9), and (8.11), let $H_{b'}$ be a minimal cardinality nontrivial subgroup such that

$$H_{b'} \subseteq m \wedge S, \tag{8.12}$$

and there exists a coset $\gamma + H_{b'}$ satisfying

$$e'' \leq \min\left\{\left\lfloor \frac{|S| - m + k - 2}{|H_{b'}|} \right\rfloor - 1,\, b' - 2\right\}, \tag{8.13}$$

and

$$k - |H_{b'}| \leq l'', \tag{8.14}$$

and $|m \wedge S| \geq (e'' + 1)|H_{b'}|$, where b' is the index of $H_{b'}$, and e'' is the number of terms of S not from the coset $\gamma + H_{b'}$, and l'' is the number of distinct terms of S not from $\gamma + H_{b'}$.

Suppose $e'' = 0$. Hence all terms of S are from $\gamma + H_{b'}$. Thus $m \wedge S \subseteq H_{b'}$, and (ii) follows from (8.12) and (8.13). So $e'' > 0$.

Suppose $|S| < m + |H_{b'}| + e''$. Hence it follows from (8.13) that $e''|H_{b'}| - e'' \leq k - 3$. Thus, since $e'' \geq l''$, it follows from (8.14) that $e''(|H_{b'}| - 2) \leq |H_{b'}| - 3$, which is only possible if $e'' = 0$, a contradiction. So

$$|S| \geq m + |H_{b'}| + e''. \tag{8.15}$$

Let $T = (a_1, \ldots, a_m)$ be an m-term subsequence of S. To complete the proof we will show that every element from the same $H_{b'}$-coset as $\sum_{i=1}^{m} a_i$ is contained in $m \wedge S$. By reordering, we may w.l.o.g. assume $a_i \in \gamma + H_{b'}$ for $i \leq n_0$, where e_0 is the number of terms of T not from $\gamma + H_{b'}$, and $n_0 = m - e_0$. Let S_0 be the subsequence of S consisting of terms from $\gamma + H_{b'}$, and let $n_1 = |S| - e'' - |H_{b'}| + 1$. Note $e_0 \leq e''$, and hence in view of (8.13) and (8.15) it follows that both n_0 and n_1 are positive integers. Also, since $H_{b'}$ being proper and nontrivial implies $m \geq 4$, then it follows in view of (8.13) that

$$\frac{|S_0| - n_1 - 1}{n_0} + 1 = \frac{|H_{b'}| - 2}{m - e_0} + 1 < \frac{|H_{b'}|}{m - b'} + 1 \leq 2. \tag{8.16}$$

In view of (8.15) it follows that $n_1 + 1 = |S| - e'' - |H_{b'}| + 2 \geq m + 2 > m - k + 2$. Hence every term of S_0 has multiplicity at most n_1, and in view of (8.16) and Proposition 2.3, it follows that there exists an n_1-set partition $A = A_1, \ldots, A_{n_1}$ of S_0 with $|A_i| = 1$ for $i > n_0$.

Assume A is chosen such that the number of indices $i \leq n_0$ with $a_i \notin A_i$ is minimal. If there exists an index j such that $a_j \notin A_j$, then there will exist an index $j' \neq j$ with $a_j \in A_{j'}$ and, if $j' \leq n_0$, then also with $a_j \neq a_{j'}$, whence the set partition $A' = A'_1, \ldots, A'_{n_1}$ defined by letting $A'_i = A_i$ for $i \neq j$, j', and, if $|A_{j'}| = 1$, letting $A'_j = (A_j \setminus \{y\}) \cup \{a_j\}$ and $A'_{j'} = (A_{j'} \setminus \{a_j\}) \cup \{y\}$, or, if $|A_{j'}| > 1$, then letting $A'_j = A_j \cup \{a_j\}$ and $A'_{j'} = A_{j'} \setminus \{a_j\}$, where $y \in A_j$, will contradict the minimality of A. Hence we may assume $a_i \in A_i$ for all $i \leq n_0$.

Let S'_0 be the subsequence of S_0 partitioned by the n_0-set partition $A_1, \ldots, A_{n_0}$. Note $|S'_0| = |S_0| - (n_1 - n_0) = n_0 + |H_{b'}| - 1$. Hence, if $n_0 \leq \frac{|S'_0| - n_0}{p'} - 1$, where p' is the smallest prime divisor of $|H_{b'}|$, then since $e_0 \leq e''$, it follows in view of (8.13) that $m \leq |H_{b'}| + e_0 - 1 \leq \frac{m}{b'} + b' - 3 \leq \frac{m}{2} - 1$, a contradiction. So assume $n_0 \geq \frac{|S'_0| - n_0 + 1}{p'} - 1$.

We may w.l.o.g. assume $\gamma = 0$. Hence, since $n_0 \geq \frac{|S'_0| - n_0 + 1}{p'} - 1$, it follows that we can apply Theorem 3.2 to the subsequence S'_0 of S_0 with n_0-set partition $A_1, \ldots, A_{n_0}$, with group $G = H_{b'}$, and with fixed elements $a_i \in A_i$ for $i \leq n_0$. If Theorem 3.2(i) holds with corresponding set partition $A' = A'_1, \ldots, A'_{n_0}$, then since $|S'_0| = n_0 + |H_{b'}| - 1$, it follows that $\sum_{i=1}^{n_0} A'_i = H_{b'}$, whence $\left(\sum_{i=n_0+1}^{m} a_i\right) + \sum_{i=1}^{n_0} A'_i$ is $H_{b'}$-periodic, and $\sum_{i=1}^{m} a_i \in \left(\sum_{i=n_0+1}^{m} a_i\right) + \sum_{i=1}^{n_0} A'_i$.

Thus every element from the same $H_{b'}$-coset as $\sum_{i=1}^{m} a_i$ is contained in $m \wedge S$, and the proof is complete. So assume that Theorem 3.2(ii) holds and let $H_{cb'} \leq H_{b'}$ be the corresponding subgroup with $c = [H_{b'} : H_{cb'}]$, let $\gamma' + H_{cb'}$ be the corresponding coset, and let $e_0' \leq c-2$ be the number of terms of S_0 not from $\gamma + H_{cb'}$. Thus, since $|S| \geq |H_{b'}| + (m-k+2)$ follows from (8.13), then it follows from (8.13) and from $|m \wedge S| \geq (e''+1)|H_{b'}|$, as in the proof of Theorem 3.2, that there are $e''' \leq c-2+\min\{\lfloor \frac{|S|-m+k-2}{|H_{b'}|} \rfloor - 1,\, b'-2\} < \min\{\lfloor \frac{|S|-m+k-2}{|H_{cb'}|} \rfloor - 1,\, cb'-2\}$ terms of S not from the coset $\gamma' + H_{cb'}$, and that $|m \wedge S| \geq (e''' + 1)|H_{cb'}|$. Thus (8.13) holds for S with subgroup $H_{cb'}$. Furthermore, since $H_{cb'} \leq H_{b'}$, then (8.12) implies that $H_{cb'} \subseteq m \wedge S$. Finally, $k - |H_{cb'}| \leq l_0$, where l_0 is the number of distinct terms not from $\gamma + H_{cb'}$, holds trivially. Consequently, from the conclusions of the last three sentences we see that the minimality of $H_{b'}$ is contradicted by $H_{cb'}$, and the proof is complete. □

We conclude the chapter by remarking that the inequality $e \leq \min\{\lfloor \frac{|S|-m+k-2}{|H_a|} \rfloor - 1,\, a-2\}$ from Theorem 8.2(ii) implies

$$|S| \geq m - k + 2 + (e+1)|H_a| + \epsilon, \tag{8.17}$$

where e is the number of terms of S not from the coset $\alpha + H_a$, and $\epsilon = \max\{0,\, |S| - (2m - k + 1)\}$; also, as seen in the proof of Theorem 8.2, if $e > 0$, then (8.17) (which is just the inequality in (8.2) rearranged with $N = 1$) implies

$$|S| \geq m + |H_a| + e \geq m + |H_a| + l \geq m + k,$$

where l is the number of distinct terms of S not from $\alpha + H_a$.

Chapter 9

The $g(m, k)$ Function of Bialostocki and Lotspeich

9.1 Discussion

As partially remarked in Chapter 8, Bialostocki and Dierker showed that $0^{m-1}1^{m-1}$ was, up to order and affine transformation (where by *affine transformation* we mean any map of the form $x \mapsto ax + b$, with a, $b \in \mathbb{Z}/m\mathbb{Z}$ and $(a, m) = 1$), the unique sequence of $2m - 2$ terms from $\mathbb{Z}/m\mathbb{Z}$ that does not contain an m-term zero-sum. Note that m-term zero-sums are preserved under reordering and affine transformation, so when describing sequences of terms from $\mathbb{Z}/m\mathbb{Z}$ that contain no m-term zero-sum subsequence, which in this chapter we refer to as *m-zsf* sequences (m-term zero-sum subsequence free), it suffices to describe one representative from each equivalence class, up to order and affine transformation. Furthermore, since a subsequence of a m-zsf sequence is also m-zsf, it really suffices to describe for all s the set $E(m, s)$, consisting of all equivalence classes of m-zsf sequences S of length s, up to order and affine transformation, that are not a proper subsequence of another m-zsf sequence.

Keeping the above observations in mind, the essential uniqueness of the lower bound coloring for EGZ easily follows from the statement that $g(m, 3) = 2m-2$, where $g(m, k)$ denotes the least integer such that every sequence of terms from $\mathbb{Z}/m\mathbb{Z}$ with at least k distinct elements and length $g(m,k)$ must contain an m-term zero-sum subsequence. Additionally, EGZ becomes the statement that $g(m,1) = g(m,2) = 2m-1$ for $m \geq 2$.

The function $g(m,k)$ was introduced by Bialostocki and Lotspeich in [4] as a way to measure the increased ease in finding an m-term zero-sum in a sequence with more distinct terms. In the same paper, they determined that $g(m,4) = 2m-3$ for $m \geq 4$. The behavior of $g(m,k)$ for large k (relative to m) was determined by Gallardo, Grekos and Phiko [17].

Theorem 9.1. *Let m and k be integers with $m \geq k \geq 2$.*

(i) If $\frac{m}{2} + 1 < k \leq m-1$, then $g(m,k) = m+2$.

(ii) If $k = m$, then $g(m,k) = \begin{cases} m & m \text{ odd} \\ m+1 & m \text{ even.} \end{cases}$

However the behavior of $g(m,k)$ for large m (relative to k) was previously unknown, though an incorrect value had been conjectured by Bialostocki and Lotspeich [4].

One goal of this chapter is to prove a theorem establishing a correspondence between m-zsf sequences of sufficient length and certain pairs of integer partitions. This result will then be used to determine $g(m,k)$ for sufficiently large m relative to k and will also provide a way of listing all the elements of $E(m,s)$ for large values of s. Finally, by separate means involving the results of Chapter 3, the value of $g(m,5)$ for all $m \geq 5$ will be determined

9.2 Zero-Sums and Integer Partitions

Theorem 9.2 gives necessary and sufficient conditions for a sequence of sufficient length to be m-zsf in terms of several inequalities over the integers. It is important to note that the two inequalities in (9.1) are interchanged by the affine transformation that interchanges

0 and 1. Thus, more precisely, Theorem 9.2 reduces the problem of determining extremal m-zsf sequences of sufficient length to the problem of finding pairs of integer partitions of $m-u-1$ and $m-v-1$ (allowing the empty partition for 0), with all parts greater than 1, as the parameters u and v range over allowed values as given by the inequalities in (9.2).

Theorem 9.2. *For integers m and l, let S be a sequence of elements from $\mathbb{Z}/m\mathbb{Z}$, satisfying $|S| = 2m - l \geq 2m - \lfloor \frac{m}{4} \rfloor - 2$. The sequence S does not contain an m-term zero-sum subsequence if and only if there exists a sequence $S' = 0^u 1^v a_1 \ldots a_{w_1} b_1 \ldots b_{w_2}$, where $1 < \overline{a_i} \leq \frac{m}{2}$ and $1 \leq \overline{-b_i} < \frac{m}{2}$, that is equivalent to S up to order and affine transformation, and for which the following four inequalities are satisfied,*

$$\sum_{i=1}^{w_1} \overline{a_i} \leq m - v - 1 \quad \text{and} \quad \sum_{i=1}^{w_2} \overline{-b_i} \leq m - u - 1 - w_2, \tag{9.1}$$

$$m - 2l + 3 \leq v \leq u \leq m - 1 \quad \text{and} \quad w_1 + w_2 \leq l - 2. \tag{9.2}$$

Moreover, equality holds in both inequalities of (9.1) if and only if S belongs to an equivalence class of $E(m, 2m - l)$.

The proof of Theorem 9.2 is an easy adaption of a proof of W. Gao and Y. O. Hamidoune [21], which uses the following two results of W. Gao [20] [18] [21] [19].

Theorem 9.3. *Let l and m be positive integers satisfying $2 \leq l \leq \lfloor \frac{m}{4} \rfloor + 2$, and let S be a sequence of elements from $\mathbb{Z}/m\mathbb{Z}$ satisfying $|S| = 2m - l$. If $0 \notin m \wedge S$, then up to order and affine transformation $S = 0^u 1^v c_1 \ldots c_w$, where $m - 2l + 3 \leq v \leq u \leq m - 1$ and $w \leq l - 2$.*

Theorem 9.4. *If S is a sequence of terms from an abelian group of order m such that no term in S has greater multiplicity than 0, then $m \wedge S = \bigcup_{i=m}^{|S|} (i \wedge S)$.*

From the proof of Theorem 9.2, we will see that it is sufficient for the inequalities in

(9.1) to hold with $v \le u \le m-1$ in order for the sequence S to be m-zsf. Consequently, the inequalities in (9.1) together with $v \le u \le m-1$ imply the remaining inequalities from (9.2), a fact which can also be deduced directly by summing the inequalities from (9.1). We proceed with the proof of Theorem 9.2.

Proof. First, suppose S is a sequence of elements from $\mathbb{Z}/m\mathbb{Z}$, satisfying $|S| = 2m - l \ge 2m - \lfloor \frac{m}{4} \rfloor - 2$, and $0 \notin m \wedge S$. Hence from Theorem 9.3 it follows that S is equivalent, up to order and affine transformation, to a sequence $S' = 0^u 1^v a_1 \ldots a_{w_1} b_1 \ldots b_{w_2}$ satisfying the inequalities in (9.2), where $1 < \overline{a_i} \le \frac{m}{2}$ and $1 \le \overline{-b_i} < \frac{m}{2}$. Hence, since $l \le \lfloor \frac{m}{4} \rfloor + 2$, then $u \ge v \ge l - 2 \ge w_1 + w_2$. Thus since S is m-zsf, then it follows from Theorem 9.4 that for any given subsequence T of $a_1 \ldots a_{w_1} b_1 \ldots b_{w_2}$,

$$\text{either } \overline{\sum_{t_i \in T} t_i} \le m - v - 1 \ \text{ or } \ \overline{\sum_{t_i \in T} t_i} \ge u + 1 + |T|, \ \text{ and} \tag{9.3}$$

$$\text{either } \overline{-\sum_{t_i \in T} t_i} \le m - u - 1 - |T| \ \text{ or } \ \overline{-\sum_{t_i \in T} t_i} \ge v + 1. \tag{9.4}$$

Induction on r, in view of (9.3) and the following three inequalities, (i) $l \le \lfloor \frac{m}{4} \rfloor + 2$, (ii) $m - v - 1 \le \lfloor \frac{m}{2} \rfloor$ (follows from (9.2) and $l \le \lfloor \frac{m}{4} \rfloor + 2$), (iii) $3m - 4l + 5 \le u + 2v$ (follows from (9.2)), implies

$$\sum_{i=1}^{r} \overline{a_i} = \overline{\sum_{i=1}^{r} a_i} \le m - v - 1, \tag{9.5}$$

for every r satisfying $1 \le r \le w_1$.

Similarly, induction on r, in view of (9.4) and the inequalities (i), (ii) and (iii), and the fact that $u \ge v$, implies

$$\sum_{i=1}^{r} \overline{-b_i} = \overline{-\sum_{i=1}^{r} b_i} \le m - u - 1 - r, \tag{9.6}$$

for every r satisfying $1 \le r \le w_2$. Hence (9.5) and (9.6) imply (9.1).

Next suppose S is an arbitrary sequence of residues from $\mathbb{Z}/m\mathbb{Z}$ that satisfies (9.1) and (9.2). Actually, we will use only the fact that (9.1) is satisfied and $v \le u \le m-1$. It follows from (9.1) that any m-term zero-sum modulo m subsequence of

$$S = (m - \overline{0})^u (\overline{1})^v (\overline{a_i})_{i=1}^{w_1} (-\overline{-b_i})_{i=1}^{w_2}$$

must be zero-sum in $\mathbb{Z}$ as well. In addition, it follows from (9.1) that the longest zero-sum in $\mathbb{Z}$ subsequence of S that does not contain a zero is of length $w_2 + \sum_{i=1}^{w_2} \overline{-b_i} \le m - u - 1$. Hence any m-term zero-sum subsequence must use at least $u+1$ zeros, which exceeds the multiplicity of zero in S. Thus S is m-zsf, and as affine transformations and reordering preserve m-term zero-sum subsequences, the proof of the main part of the theorem is complete. Since the two inequalities in (9.1) are interchanged by the affine transformation that interchanges 0 and 1, then the moreover part of the theorem is easily deduced from the main part of the theorem. □

9.3 The Function $g(m, k)$ for Large m

In this section, we use Theorem 9.2 to determine the behavior of $g(m, k)$ for large m with respect to k. We begin first by giving a lower bound construction for $g(m, k)$ that uses precisely k distinct residues.

Theorem 9.5. *Let $m \ge k \ge 2$ be positive integers. If k is odd and $m \ge \frac{k^2+4k+3}{8} + 1$ or k is even and $m \ge \frac{k^2+2k}{8} + 1$, then $g(m, k) \ge 2m - \lfloor \frac{k^2-2k+5}{4} \rfloor$.*

Proof. If k is even, consider the sequence

$$S_0 = (-\frac{k-2}{2})\dots(-2)(-1)(0)^{m-\frac{k^2+2k}{8}}(1)^{m-\frac{k^2+2k}{8}}(2)(3)\dots(\frac{k}{2}),$$

and if k is odd, consider the sequence

$$S_1 = (-\frac{k-3}{2})\dots(-2)(-1)(0)^{m-\frac{k^2-1}{8}}(1)^{m-\frac{k^2+4k+3}{8}}(2)(3)\dots(\frac{k+1}{2}).$$

It follows from the hypotheses that both strings are well defined. Since both S_1 and S_2 satisfy (9.1), and since $v \le u \le m-1$, where u and v are the multiplicities of 0 and 1, respectively, it follows from the proof of the second direction of Theorem 9.2 that S_1 and S_2 are m-zsf. $\square$

Next, we use Theorem 9.2 to give a matching upper bound for slightly larger m. Again, the proof is only a minor modification of the proof used in [21].

Theorem 9.6. *Let $m \ge k \ge 2$ be positive integers. If k is even and $m \ge k^2 - 2k - 4$ or k is odd and $m \ge k^2 - 2k - 3$, then $g(m,k) = 2m - \lfloor \frac{k^2-2k+5}{4} \rfloor$.*

Proof. From Theorem 9.5, it suffices to show $g(m,k) \le 2m - \lfloor \frac{k^2-2k+5}{4} \rfloor$. Assume to the contrary that there is a sequence S of terms from $\mathbb{Z}/m\mathbb{Z}$, with $|S| = 2m - \lfloor \frac{k^2-2k+5}{4} \rfloor$, and $0 \notin m \wedge S$. From the hypotheses and the fact that $k^2 \equiv 0$ or $1 \mod (4)$, it follows that $\lfloor \frac{k^2-2k+5}{4} \rfloor \le \lfloor \frac{m}{4} \rfloor + 2$. Hence from Theorem 9.2 it follows that w.l.o.g. S satisfies (9.1) and (9.2). Let $c_1 = |\{a_1, \dots, a_{w_1}\}|$ and $c_2 = |\{b_1, \dots, b_{w_2}\}|$. It follows from the first inequality in (9.1), that $2 + 3 + \dots + (c_1 + 1) + 2(w_1 - c_1) \le m - v - 1$, implying that

$$\frac{c_1^2 - c_1}{2} + 2w_1 \le m - v - 1. \tag{9.7}$$

Likewise from the second inequality in (9.1), it follows that

$$\frac{c_2^2 - c_2}{2} + w_2 \leq m - u - 1 - w_2. \tag{9.8}$$

Inequalities (9.7) and (9.8) imply (since $c_1 \leq w_1$ and $c_2 \leq w_2$)

$$\frac{c_1^2 + c_1}{2} + \frac{c_2^2 + c_2}{2} \leq \frac{c_1^2 - c_1}{2} + \frac{c_2^2 - c_2}{2} + w_1 + w_2 \leq m - v - 1 - w_1 + m - u - 1 - w_2 = l - 2,$$

which, in turn, yields

$$l \geq \frac{(c_1 + c_2)^2}{4} + \frac{c_1 + c_2}{2} + 2 \geq \frac{(k-2)^2}{4} + \frac{k-2}{2} + 2 = \frac{k^2 - 2k + 4}{4} + 1 > l,$$

which is a contradiction, completing the proof. □

We conclude the section on the following page with a table of $E(m, s)$ for all m and s satisfying $2m - 2 \geq s \geq \max\{2m - 8,\ 2m - \lfloor \frac{m}{4} \rfloor - 2\}$. This is accomplished by fixing the length $2m - l$ of S and using the inequalities from (9.2) to determine allowed pairs of parameters u and v. Then, from Theorem 9.2, we obtain an element of $E(m, s)$, for each of these pairs u and v, and each pair of integer partitions of $m - u - 1$ and $m - v - 1$ with all parts greater than one and the total number of parts (between both partitions) equal to $2m - l - u - v$.

The values for $g(m, k)$ with $k \leq 4$ can be easily derived from the table. Since no string with four distinct residues, with $m \geq 4$, occurs with length $s \geq 2m - 3$, it follows that $g(m, 4) \leq 2m - 3$. From Theorem 9.5, it follows that $g(m, 4) \geq 2m - 3$, whence $g(m, 4) = 2m - 3$. Since no string with three distinct residues, with $m \geq 2$, occurs with length $s \geq 2m - 2$, it follows that $g(m, 3) \leq 2m - 2$. From Theorem 9.5, it follows that

$g(m,3) \geq 2m-2$ for $m \geq 4$, whence $g(m,3) = 2m-2$ for $m \geq 4$. The remaining case, $g(m,3) = 2m-3 = 3$ for $m = 3$, follows from Theorem 9.1, or else by a simple case investigation.

Table of $E(m,s)$ for Large m and s

m	s	$E(m,s)$		
$m \geq 2$	$2m-2$	$0^{m-1}1^{m-1}$		
$m \geq 4$	$2m-3$	$0^{m-1}1^{m-3}2$		
$m \geq 8$	$2m-4$	$0^{m-1}1^{m-5}2^2$	$(-1)0^{m-3}1^{m-3}2$	$0^{m-1}1^{m-4}3$
$m \geq 12$	$2m-5$	$0^{m-1}1^{m-7}2^3$	$(-1)0^{m-3}1^{m-5}2^2$	$0^{m-1}1^{m-6}23$
		$(-1)0^{m-3}1^{m-4}3$	$0^{m-1}1^{m-5}4$	
$m \geq 16$	$2m-6$	$0^{m-1}1^{m-9}2^4$	$(-1)0^{m-3}1^{m-7}2^3$	$(-1)^2 0^{m-5}1^{m-5}2^2$
		$0^{m-1}1^{m-8}2^2 3$	$(-1)0^{m-3}1^{m-6}23$	$(-2)0^{m-4}1^{m-5}2^2$
		$0^{m-1}1^{m-7}24$	$0^{m-1}1^{m-7}3^2$	$(-1)0^{m-3}1^{m-5}4$
		$(-2)0^{m-4}1^{m-4}3$	$0^{m-1}1^{m-6}5$	
$m \geq 20$	$2m-7$	$0^{m-1}1^{m-11}2^5$	$(-1)0^{m-3}1^{m-9}2^4$	$(-1)^2 0^{m-5}1^{m-7}2^3$
		$0^{m-1}1^{m-10}2^3 3$	$(-1)0^{m-3}1^{m-8}2^2 3$	$(-2)0^{m-4}1^{m-7}2^3$
		$(-1)^2 0^{m-5}1^{m-6}23$	$0^{m-1}1^{m-9}23^2$	$0^{m-1}1^{m-9}2^2 4$
		$(-1)0^{m-3}1^{m-7}24$	$(-1)0^{m-3}1^{m-7}3^2$	$(-2)0^{m-4}1^{m-6}23$
		$(-1)^2 0^{m-5}1^{m-5}4$	$0^{m-1}1^{m-8}25$	$0^{m-1}1^{m-8}34$
		$(-1)0^{m-3}1^{m-6}5$	$(-2)0^{m-4}1^{m-5}4$	$0^{m-1}1^{m-7}6$
$m \geq 24$	$2m-8$	$0^{m-1}1^{m-13}2^6$	$(-1)0^{m-3}1^{m-11}2^5$	$(-1)^2 0^{m-5}1^{m-9}2^4$
		$(-1)^3 0^{m-7}1^{m-7}2^3$	$0^{m-1}1^{m-12}2^4 3$	$(-1)0^{m-3}1^{m-10}2^3 3$
		$(-2)0^{m-4}1^{m-9}2^4$	$(-1)^2 0^{m-5}1^{m-8}2^2 3$	$(-2)(-1)0^{m-6}1^{m-7}2^3$
		$0^{m-1}1^{m-11}2^3 4$	$0^{m-1}1^{m-11}2^2 3^2$	$(-1)0^{m-3}1^{m-9}2^2 4$
		$(-1)0^{m-3}1^{m-9}23^2$	$(-2)0^{m-4}1^{m-8}2^2 3$	$(-1)^2 0^{m-5}1^{m-7}24$
		$(-1)^2 0^{m-5}1^{m-7}3^2$	$(-3)0^{m-5}1^{m-7}2^3$	$(-2)(-1)0^{m-6}1^{m-6}23$
		$0^{m-1}1^{m-10}2^2 5$	$0^{m-1}1^{m-10}234$	$0^{m-1}1^{m-10}3^3$
		$(-1)0^{m-3}1^{m-8}25$	$(-1)0^{m-3}1^{m-8}34$	$(-2)0^{m-4}1^{m-7}24$
		$(-2)0^{m-4}1^{m-7}3^2$	$(-1)^2 0^{m-5}1^{m-6}5$	$(-3)0^{m-5}1^{m-6}23$
		$0^{m-1}1^{m-9}26$	$0^{m-1}1^{m-9}35$	$0^{m-1}1^{m-9}4^2$
		$(-1)0^{m-3}1^{m-7}6$	$(-2)0^{m-4}1^{m-6}5$	$(-3)0^{m-5}16^{m-5}4$
		$0^{m-1}1^{m-8}7$		

9.4 The Erdős-Heilbronn Conjecture and $g(m,5)$

From the previous section, we know that $g(m,5) = 2m-5$ for $m \geq 12$, while from Theorem 9.1 it follows that $g(m,5) = m = 2m-5$ for $m = 5$, that $g(m,5) = m+2 = 2m-4$ for $m = 6$, and that $g(m,5) = m+2 = 2m-5$ for $m = 7$. Thus only the cases $m = 8$, 9, 10, 11 are left remaining in the function $g(m,5)$. In this section, we use the results of Chapter 3 to give an alternative derivation of $g(m,5)$ for $m \geq 7$.

Note $k = 5$ is the first value of k where the function $g(m,k)$ ceases to behave linearly. Thus the linear bound in k from Theorem 8.2 is not immediately useful. This nonlinearity in $g(m,k)$ is perhaps indicative that a result similar to Theorem 8.2 might hold with a quadratic expression in k, though this would be more difficult to show. For instance, the following theorem, of J. A. Dias da Silva and Y. O. Hamidoune [12], which gives a bound on $|n \wedge S|$ in the case when S consists of distinct terms, confirmed a long-standing conjecture of Erdős and Heilbronn.

Erdős-Heilbronn Conjecture (EHC). *Let S be a sequence of distinct elements from $\mathbb{Z}/m\mathbb{Z}$. If m is prime, then $|n \wedge S| \geq \min\{m,\, n|S| - n^2 + 1\}$.*

Unfortunately, the structure of sequences that fail to satisfy the EHC bound for composite m is still not well understood. One might hope, as in Kneser's Theorem, that if a set S failed to satisfy the EHC bound, then S would be a large subset of a periodic set (which, with a few exceptions, would imply $n \wedge S$ was periodic itself). Regardless, such a statement can be verified, with brute force, for $|S| \leq 5$, and we will need such a result for our derivation of $g(m,5)$.

Theorem 9.7. *Let S be a sequence of distinct terms from $\mathbb{Z}/m\mathbb{Z}$. Suppose $|S| = k \leq 5$. If $|2 \wedge S| < 2|S| - 3$, then there is a H-periodic set T with $S \subseteq T$ and $2(|T| - |S|) \leq |H| - 2$.*

Proof. Since $|2 \wedge S| \geq 1$, the cases $k \leq 2$ are trivial. Suppose $k = 3$. Let $S = a_1, a_2, a_3$. Then all three pairs $a_i + a_j$ must be distinct, else w.l.o.g. $a_1 + a_2 = a_2 + a_3$, implying $a_1 = a_3$, a contradiction. Hence $|2 \wedge S| \geq 3 = 2|S| - 3$. Suppose $k = 4$, and let $S = a_1, a_2, a_3, a_4$. Let $A = \{a_1 + a_2,\ a_1 + a_3,\ a_1 + a_4, , a_2 + a_3,\ a_2 + a_4,\ a_3 + a_4, \}$ be the set of all 2-sums of S. If any three of the 2-sums in A are all equal to one another, then this implies that not all the a_i are distinct, a contradiction. Hence if $|2 \wedge S| \leq 2|S| - 4 = 4$, then there must be at least 2 pairwise disjoint equalities among the 2-sums. Note that in any such equation $a_i + a_j = a_l + a_k$ we must have all indices distinct, else the distinctness of the terms in S will be contradicted. Hence w.l.o.g. by symmetry, it follows that $a_1 + a_2 = a_3 + a_4$ and $a_1 + a_3 = a_2 + a_4$. Hence $2a_2 = 2a_3$ and $2a_1 = 2a_4$ follows, implying that a_2 and a_3 are from the same $\frac{m}{2}(\mathbb{Z}/m\mathbb{Z})$-coset, and that a_1 and a_4 are also from the same $\frac{m}{2}(\mathbb{Z}/m\mathbb{Z})$-coset. Thus, since $|\frac{m}{2}(\mathbb{Z}/m\mathbb{Z})| = 2$, it follows that S is $\frac{m}{2}(\mathbb{Z}/m\mathbb{Z})$-periodic, and taking $T = S$ completes the proof. The remaining case $k = 5$ follows from the next lemma. $\square$

Lemma 12. *Let $S = a_1, a_2, a_3, a_4, a_5$ be a sequence of five distinct terms from $\mathbb{Z}/m\mathbb{Z}$. Then either $|2 \wedge S| \geq 7$ or else there exists a subgroup H of $\mathbb{Z}/m\mathbb{Z}$ of cardinality $h = 5$ or $h = 6$, and $\alpha \in G$, such that $S \subseteq \alpha + H$.*

Proof. Let A be the set

$$\{a_1 + a_2,\ a_1 + a_3,\ a_1 + a_4,\ a_1 + a_5,\ a_2 + a_3,\ a_2 + a_4,\ a_2 + a_5,\ a_3 + a_4,\ a_3 + a_5,\ a_4 + a_5\}$$

consisting of all 2-sums of S. If any three of the 2-sums in A are all equal to one another, then this implies that not all the a_i are distinct, a contradiction. Hence if $|2 \wedge S| \leq 6$, then there must be at least 4 pairwise disjoint equalities among the 2-sums. Since there are four distinct a_i's in each of the four equalities, it follows by the pigeonhole principle that one a_i

must occur in all 4 equalities, say a_1. Thus, since $a_1 + a_2 \neq a_1 + a_5$, it follows w.l.o.g. that the equalities in (9.9) and (9.10) hold. Furthermore, one of the equalities in (9.11) and one of the equalities in (9.12) hold as well.

$$a_1 + a_2 = a_3 + a_4 \tag{9.9}$$

$$a_1 + a_5 = a_2 + a_3 \tag{9.10}$$

$$a_1 + a_3 = a_2 + a_4, \qquad a_1 + a_3 = a_2 + a_5, \qquad a_1 + a_3 = a_4 + a_5 \tag{9.11}$$

$$a_1 + a_4 = a_2 + a_3, \qquad a_1 + a_4 = a_2 + a_5, \qquad a_1 + a_4 = a_3 + a_5. \tag{9.12}$$

Subsequently, we will refer to an equation in a numbered line by the number of the line followed by a letter from $a, b, c, \ldots$ in lexicographic order, e.g., (9.11)a, (9.11)b and (9.11)c correspond to the equations $a_1 + a_3 = a_2 + a_4$, $a_1 + a_3 = a_2 + a_5$ and $a_1 + a_3 = a_4 + a_5$, respectively. From (9.9) and (9.10) it follows that

$$2a_2 = a_4 + a_5. \tag{9.13}$$

Observe that (9.12)a cannot hold, since if it does, then together with (9.9) and (9.13), it will imply that $a_4 = a_5$, a contradiction. Thus either (9.12)b or (9.12)c holds. If (9.12)b holds, then together with (9.9) and (9.10), equalities (9.14)(a), (9.14)(b) and (9.14)(c) are implied; and in turn (9.14)a and (9.14)(b) imply (9.14)(d).

$$2a_5 = a_3 + a_4, \qquad 2a_1 = a_3 + a_5, \qquad 2a_2 + a_3 = 2a_1 + a_4, \qquad a_4 + 2a_1 = 3a_5. \tag{9.14}$$

If (9.12)c holds, then together with (9.9) and (9.10), equalities (9.15)(a), (9.15)(b), (9.15)(c)

and (9.15)(d) are implied; and in turn (9.15)a and (9.15)(b) imply (9.16).

$$2a_4 = a_2 + a_5, \qquad 2a_5 = a_2 + a_4, \qquad 2a_3 + a_5 = 2a_1 + a_2, \qquad 2a_1 + a_4 = 2a_3 + a_2, \quad (9.15)$$

$$3a_5 = 3a_4. \quad (9.16)$$

We proceed by considering three cases, corresponding to each of the three equalities in (9.11).

Case 1: (9.11)a holds. Then (9.11)a and (9.9) imply (9.17)(a) and (9.17)(b).

$$2a_1 = 2a_4, \qquad 2a_3 = 2a_2. \quad (9.17)$$

Suppose (9.12)c holds. Then (9.17)(a), (9.17)(b) and (9.15)d imply $3a_4 = 3a_2$, which, along with (9.16), (9.17)a and (9.17)b, implies the lemma with $h = 6$. Hence we may assume (9.12)b holds. Furthermore, (9.17)(a), (9.17)(b) and (9.14)c imply $3a_3 = 3a_4$; and (9.17)a and (9.14)d imply $3a_4 = 3a_5$. Thus from (9.17)(a), (9.17)(b) and $3a_3 = 3a_4 = 3a_5$, it follows that the lemma holds with $h = 6$.

Case 2: (9.11)b holds. Then (9.11)b and (9.10) imply $2a_3 = 2a_5$ and $2a_1 = 2a_2$, which, along with (9.14)a, implies $a_3 = a_4$, a contradiction, and which, along with (9.15)c, implies $3a_5 = 3a_2$. In the later case, we obtain the three equalities, $2a_3 = 2a_5$, $2a_1 = 2a_2$ and $3a_5 = 3a_2$, which, along with (9.16), imply the lemma with $h = 6$.

Case 3: (9.11)c holds. Then (9.11)c, (9.10) and (9.9) imply (9.18)a and (9.18)b.

$$2a_3 = a_2 + a_5, \qquad 2a_1 = a_2 + a_4. \quad (9.18)$$

Suppose that (9.12)c holds. Then (9.18)a and (9.15)a imply $2a_3 = 2a_4$. Additionally, (9.18)b and (9.15)b imply $2a_5 = 2a_1$. Furthermore, (9.15)b and (9.15)c imply $3a_5 + 2a_3 = 2a_1 + 2a_2 + a_4$, which, along with (9.18)b and $2a_3 = 2a_4$, implies $3a_2 = 3a_5$. Thus $2a_3 = 2a_4$, $2a_5 = 2a_1$, $3a_2 = 3a_5$ and (9.16) imply the lemma with $h = 6$. So we may assume (9.12)b holds. Then (9.18)b and (9.14)b imply $a_2 + a_4 = a_3 + a_5$, and we conclude in this case that there are at most 5 distinct 2-sums. Furthermore, $a_2 + a_4 = a_3 + a_5$ and (9.9) imply $2a_4 = a_5 + a_1$. Thus (9.18)a, (9.18)b, (9.14)a, (9.13) and $2a_4 = a_5 + a_1$ imply (9.19)(a), (9.19)(b) and (9.19)(c).

$$3a_4 + a_3 = 3a_5 + a_1, \qquad 3a_3 + a_4 = 3a_5 + a_2, \qquad 3a_2 = 2a_1 + a_5. \tag{9.19}$$

We proceed by combining (9.19)(a), (9.19)(b) and (9.19)(c) with (9.18), (9.13) and $2a_4 = a_5 + a_1$, yielding 20(a), 20(b) and 20(c).

$$5a_4 = 4a_5 + 2a_1 - a_3, \qquad 5a_3 = 4a_5 + 2a_2 - a_4, \qquad 5a_2 = 2a_1 + 2a_5 + a_4 \tag{9.20}$$

Next (9.20)a, (9.20)b and (9.14)c imply $5a_4 = 5a_3$. Additionally, (9.20)c and (9.14)d imply $5a_2 = 5a_5$. Furthermore, (9.20)a, (9.14)a and (9.14)d imply $5a_4 = 5a_5$. Therefore it follows that $\{a_2, a_3, a_4, a_5\}$ are four elements from a coset $\alpha + H$, where H is a subgroup of $\mathbb{Z}/m\mathbb{Z}$ of cardinality 5. Then it can be easily verified that a_1 is the fifth element of $\alpha + H$, as otherwise $|2 \wedge S| > 5$, contradicting the fact that there are at most five distinct 2-sums. Thus the lemma holds with $h = 5$, completing the proof. □

We will also need the following two results. The first is a very basic theorem, so much so that the result is sometimes referred to as the 'Caveman Theorem.'

Theorem 9.8. *Let S be a sequence of elements from a finite abelian group G. If $|S| = |G|$,* *then there exists a nonempty zero-sum subsequence consisting of consecutive terms of S.*

Proof. Let $S = s_1, \ldots, s_m$, where $|G| = m$. If the theorem is false, then by the pigeonhole principle at least two of the sums $\sum_{i=1}^{j} s_i$, for $j \in \{1, 2, \ldots, m\}$, must be equal to each other, say j_1 and j_2, with $j_1 < j_2$. However, then the sequence $s_{j_1+1}, \ldots, s_{j_2}$ satisfies the theorem. $\square$

The second is a simple result of R. Eggleton and P. Erdős [14].

Theorem 9.9. *Let S be a sequence of distinct elements from a finite abelian group. If $0 \notin \bigcup_{i=1}^{|S|}(i \wedge S)$ and $|S| \geq 4$, then $|\bigcup_{i=1}^{|S|}(i \wedge S)| \geq 2|S|$.*

We conclude the section with the derivation of $g(m, 5)$ for all m. Note that $1 \wedge S$ is just the set of distinct elements that occur as a term of S.

Theorem 9.10. *Let $m \geq 5$. Then $g(6, 5) = 8$, and if $m \neq 6$, then $g(m, 5) = 2m - 5$.*

Proof. For $m \leq 6$ the result follows from Theorem 9.1. So we may assume $m \geq 7$. The lower bound follows from Theorem 9.5. Suppose S is m-zsf and $|S| = 2m - 5$. We may w.l.o.g. assume that 0 has the greatest multiplicity in S.

Case 1: The multiplicity of 0 in S is at most $m - 2$. Applying Lemma 12 to all possible 5-sets of $1 \wedge S$ that include 0, we can either find a 5-set $A \subseteq 1 \wedge S$ such that $|2 \wedge A| = |3 \wedge A| \geq 7$ and $0 \in A$, or else there exists a subgroup H of cardinality $h = 5$ or $h = 6$ such that $1 \wedge S \subseteq H$. In the latter case, it follows from $m \geq 7$ that $m \geq 10$. Hence from $\frac{m}{h}$ applications of EGZ considering terms as elements of H, it follows that any subsequence with length $m + h - 1 \leq 2m - 5$ must contain an m-term zero-sum subsequence, a contradiction. So we may instead assume that $|3 \wedge A| \geq 7$. In view of the assumption of the case and Proposition 2.3, it follows that there exists an $(m - 3)$-set partition P of $S \setminus A$. Note that $m - 3 \geq m - 7 = |S \setminus A| - (m - 3)$. Hence we can apply Theorem 3.2

to the subsequence $S \setminus A$ of $S \setminus A$. If Theorem 3.2(i) holds, then the resulting $(m-3)$-set partition will have cardinality at least $m-6$, whence from $|3 \wedge A| \geq 7$ and from the pigeonhole principle, it follows that we can find an m-term zero-sum subsequence of S by an appropriate selection of $(m-3)$ terms from the resulting set partition and an appropriate three terms from A. Hence the proof is complete for $m \leq 8$ (since Theorem 3.2(i) trivially holds in this case, as all but one set have cardinality one).

So assume that Theorem 3.2(ii) holds with coset $\alpha + H_a$, where H_a has index a, and w.l.o.g assume $\alpha = 0$. Let P be the second $(m-3)$-set partition from Theorem 3.2(ii) whose sumset is H_a, and apply Proposition 2.4 to P to obtain an $(\frac{m}{a}-1)$-set partition P' of a subsequence Q of $S \setminus A$ of length at most $|Q| \leq 2\frac{m}{a} - 2$, whose sumset is also H_a. Then there exists a subsequence R of $S \setminus A$ of length $a-1$ whose terms are from H_a and are not used in P'. We can repeatedly apply Theorem 9.8 to a subsequence of $(S \setminus Q) \setminus R$ of length $m - \frac{m}{a} + 1$ with its terms considered as elements from $(\mathbb{Z}/m\mathbb{Z})/H_a$ to obtain a subsequence T of $(S \setminus Q) \setminus R$ whose sum is an element of H_a and of length r, where r satisfies $m - \frac{m}{a} - a + 2 \leq r \leq m - \frac{m}{a} + 1$. Since the sumset of P' is H_a, we can find $\frac{m}{a} - 1$ terms from P' that, along with T and an appropriate number of terms from R, gives an m-term zero-sum subsequence.

Case 2: The multiplicity of 0 in S is $m-1$. Let T' be a subsequence of S that consists of 4 distinct nonzero residue classes and 3 zeros. In view of Proposition 2.3, it follows that there exists an $(m-4)$-set partition P' of $S \setminus T'$. Since $m-4 \geq m-8 = |S \setminus T'| - (m-4)$, we can apply Theorem 3.2 to P. If Theorem 3.2(i) holds, then the cardinality of the resulting $(m-4)$-set partition will be at least $m-7$. Hence applying Theorem 9.9 to $(1 \wedge T') \setminus \{0\}$, it then follows in view of the pigeonhole principle that there is an m'-term zero-sum subsequence, where $m-3 \leq m' \leq m$, consisting of an appropriate selection of $m-4$ terms from the resulting $(m-4)$-set partition and the terms from a nonempty subsequence of $(1 \wedge T') \setminus \{0\}$ (whose length must be between 1 and $4 = |(1 \wedge T') \setminus \{0\}|$). Thus adding an appropriate number of filler zeros from T' yields an m-term zero-sum. If conclusion Theorem 3.2(ii) holds instead, then since $m-4 > a-2$ implies $0 \in \alpha + H_a$, the arguments from the end of Case 1 complete the proof. $\square$

Chapter 10

A Modified Nondecreasing Diameter Problem

10.1 Discussion

As mentioned in Chapter 7, the first Ramsey-type problem considered with respect to zero-sum generalizations was the nondecreasing diameter problem introduced by Bialostocki, Erdős and Lefmann [6], defined as the system ND_m^2 given by

$$x_1 < x_2 < \cdots < x_m < y_1 < y_2 < \cdots < y_m$$

$$x_m - x_1 \leq y_m - y_1.$$

The quantity $x_m - x_1 = \max X - \min X$ is the diameter of the set X, and the solutions to ND_m^2 are pairs of disjoint m-sets X and Y, where all terms in the second set Y come after all terms in the first set X, and where the diameter of the second set Y is at least the diameter of the first set X.

One reason that r-color zero-sum generalizations with $r \leq 4$ were obtained for ND_m^2, is that, even though there are m elements in each of the two sets X and Y, the overriding property of the system ND_m^2, namely the diameter of a set, is not very restrictive. Indeed, it depends solely on the two extreme elements of the set. Given two elements w and w' with large diameter $w' - w$, any combination of $m-2$ 'filler' elements (including elements that lie outside the interval $[w, w']$ since these can only increase the diameter) yields an m-set with large diameter. The freedom with how these 'filler' elements can be chosen can then be exploited; for instance, if they have a compressed set partition (with at most $m-2$ sets having cardinality greater than one) with large cardinality sumset, then a selection of $m-2$ terms can be chosen from the compressed set partition so that their sum is the additive inverse of the sum of the colors of w and w', yielding a zero-sum set. This element of 'filler freedom' is a reoccurrence throughout Part II, since the level of development in Part I yields methods most readily suited to problems with roughly half the elements as 'filler.' Note that the overriding property of the system WND_m^2 from Chapter 7 depended, in this same sense, only on the largest element y_m, which is why the system WND_m^2 so readily fell to the methods from Part I, while the applications from Chapters 8 and 9 involved no restrictions on the zero-sum configuration, instead placing restrictions (in terms of number of distinct terms) on the considered (coloring) sequences.

In this chapter we introduce and tackle a variation on the original nondecreasing diameter problem. Namely, we consider the system obtained by replacing, in the nondecreasing diameter system, the inequality $x_1 - x_m \leq y_m - y_1$ by $x_1 - x_j \leq y_j - y_1$, for $2 \leq j \leq m$. Thus we consider a modified notion of diameter for an m-set X, given by $g_j(X) \overset{def}{=} \text{int}_j(X) - \text{int}_1(X)$. The solutions for this modified system $N_jD_m^2$ are just those pairs of m-sets X and Y with $X \prec Y$ and $g_j(X) \leq g_j(Y)$. When $j = m$, then $g_m(X)$ is just the diameter of X, and $N_mD_m^2 = ND_m^2$.

For simplicity of notation, we let $f(m, j) = f(N_jD_m^2, 2)$ and also let $f_{zs}(m, j) = f_{zs}(N_jD_m^2, 2)$. Section 10.2 will deal with giving general upper and lower bounds on $f(m, j)$

and $f_{zs}(m,j)$, while in Section 10.3 we show that a zero-sum generalization holds in the case $j = m-1$. For Section 10.3, we will consider $N_j D_m^2$ as a system whose overriding property depends only on the three elements x_1, x_{m-1} and x_m, and apply the methods of Part I.

10.2 General Upper and Lower Bounds

Before starting, we introduce the notation $S_1 \cup S_2$ to denote the concatenation of the sequences S_1 and S_2. Next, we give in Theorem 10.1 a lower bound for $f(m,j)$, and hence $f_{zs}(m,j)$ as well. Theorem 10.1 gives the bounds $f(m,m) \geq 5m-3$ and $f(m,m-1) \geq 5m-4$, both of which will be shown tight later in this chapter. However, the bound is not in general tight for sufficiently small j.

Theorem 10.1. *Let m and j be integers satisfying $2 \leq j \leq m$, and let $k = \left\lfloor \frac{-1+\sqrt{\frac{8m-9+j}{j-1}}}{2} \right\rfloor$. Then $f(m,j) \geq 4m-2+(j-1)k$.*

Proof. Consider the coloring $\Delta : [1, 4m-3+(j-1)k] \to \{0,1\}$ given by

$$0^{m-1-(j-1)\frac{k(k+1)}{2}}(1^{j-1}0^{k(j-1)})(1^{j-1}0^{(k-1)(j-1)})\cdots(1^{j-1}0^{2(j-1)})(1^{j-1}0^{j-1})1^{2m-1}0^{m-1}.$$

Using the quadratic formula, it can be easily verified that k is the greatest integer such that $\sum\limits_{i=1}^{k}(j-1)i = (j-1)\frac{k(k+1)}{2} \leq m-1$. Thus,

$$|\Delta^{-1}(0) \cap [1, m-1+(j-1)k]| = m-1,$$

and

$$|\Delta^{-1}(1) \cap [1, m-1+(j-1)k]| = (j-1)k \leq m-1.$$

Suppose there exist sets B_1 and B_2 that form a monochromatic $N_jD_m^2$ solution. Notice that $\Delta(B_1) \neq 0$, since otherwise $|[\max(B_1)+1, 4m-3+(j-1)k]| \leq m-2$. Similarly, $\Delta(B_2) \neq 0$. Thus $\Delta(B_i) = 1$ for $i = 1, 2$. Furthermore, given any m-set B with $\Delta(B) = 1$, there exists an m-set B^* with $\Delta(B^*) = 1$ satisfying $\max(B^*) \leq \max(B)$, $g_j(B^*) \leq g_j(B)$, and $(j-1)|g_j(B^*)$ (simply compress the set B inwards until the first j integers are consecutive with the exception of one gap of length $t(j-1)$ where a single block of zeroes prevents further compression). Therefore we may assume $g_j(B_1) = j-1+t(j-1)$ for some $t \in \{0, 1, \ldots, k\}$. Since $\max(B_1) < \min(B_2)$, it follows that B_2 is contained within the last $2m-1+t(j-1)-m$ integers colored by 1. Hence, since $|\Delta^{-1}(1) \cap [1, m-1+(j-1)k]| = (j-1)k \leq m-1$ forces B_2 to be contained in the block of $2m-1$ consecutive integers colored by 1, it follows that

$$g_j(B_2) \leq (j-1) + (m-1+(j-1)t) - m = (t+1)(j-1) - 1.$$

Consequently, $g_j(B_1) > g_j(B_2)$, a contradiction. $\square$

The following lemma will be used to derive the upper bound for $f(m, j)$.

Lemma 13. *Let m and j be integers satisfying $2 \leq j \leq m$. If $\Delta : [1, 3m-2] \to \{0, 1\}$ is an arbitrary coloring, then one of the following holds:*

(i) there exists a monochromatic m-set $B \subset [1, 3m-2]$ satisfying $g_j(B) \geq m+j-2$,

(ii) there exists a monochromatic $N_jD_m^2$ solution,

(iii) the coloring Δ is given (up to symmetry) by 1^r0H, with $r \in [j, m-1]$, and H a block such that there exists a monochromatic m-set $B \subseteq 0H$ for which $g_j(B) \geq m+2j-r-3$.

Proof. Assume w.l.o.g. $\Delta(1) = 1$. If $|\Delta^{-1}(1)| < m$, then $|\Delta^{-1}(0)| \geq 2m-1$, whence (i) follows. So $|\Delta^{-1}(1)| \geq m$. Let $S = [m+j-1, 3m-2]$. Since $\Delta(1) = 1$ and $|\Delta^{-1}(1)| \geq m$, it follows that if $|\Delta^{-1}(1) \cap S| \geq m-j+1$, then (i) follows. Hence $|\Delta^{-1}(1) \cap S| \leq m-j$, whence

$$|\Delta^{-1}(0) \cap S| \geq m. \tag{10.1}$$

Let $y_2 < y_3 < \cdots < y_m$ be the elements of $\text{last}_{m-1}(\Delta^{-1}(0) \cap S)$. Observe, since $|\Delta^{-1}(1) \cap S| \leq m - j$, that $y_j \geq m + 2j - 2$. Hence, if there exists $i \in [1, j]$ such that $\Delta(i) = 0$, then (i) follows. Consequently, $\Delta(i) = 1$ for $i \in [1, j]$. However, if $\Delta(i) = 1$ for $i \in [1, m]$, then (ii) follows in view of (10.1). Therefore, there exists a minimal $i \in [j + 1, m]$ such that $\Delta(i) = 0$. Define $r = i - 1$. Then the set $B = \{r + 1, y_2, \ldots, y_m\}$ satisfies $g_j(B) \geq m + 2j - 2 - (r + 1) = m + 2j - r - 3$, whence (iii) follows. □

Theorem 10.2. *Let m and j be integers satisfying $2 \leq j \leq m$. Then*

$$f(m, j) \leq f(m, m) = 5m - 3.$$

Proof. In view of the lower bound given by Theorem 10.1, it suffices to show $f(m, j) \leq 5m - 3$. Let $\Delta : [1, 5m - 3] \to \{0, 1\}$ be an arbitrary coloring. Apply Lemma 13 to the interval $[2m, 5m-3]$. If Lemma 13(ii) holds, then the proof is complete, and if Lemma 13(i) holds, then by applying the pigeonhole principle to $[1, 2m - 1]$ the proof is also complete. Thus we may assume Lemma 13(iii) holds, so that w.l.o.g.

$$\Delta[2m, 5m - 3] = 1^r 0H,$$

where r and H are as in Lemma 13(iii), and that there is a subset $B \subseteq [2m + r, 5m - 3]$ with $g_j(B) \geq m + 2j - r - 3$. Let $S = [1, 2j - 1]$.

Case 1: $|\Delta^{-1}(1) \cap S| \geq j$.

Since $r \leq m-1$, it follows that $g_j(B) \geq 2j-2$. Hence we may assume

$$|\Delta^{-1}(1) \cap [1, 2m+r-1]| \leq m-1.$$

But then since $\Delta([2m,\, 2m+r-1]) = 1$, it follows that

$$|\Delta^{-1}(1) \cap [2j, 2m-1]| \leq m-j-r-1, \tag{10.2}$$

implying, since $j \leq r$, that

$$|\Delta^{-1}(0) \cap [2j, 2m-1]| \geq m-j+r+1 \geq m.$$

Let $y_1 < y_2 < \ldots < y_m$ be the elements of $\mathrm{first}_m(\Delta^{-1}(0) \cap [2j, 2m-1])$. Then by (10.2), it follows that $B_1 = \{y_1, \ldots, y_m\}$ is a monochromatic m-set with $g_j(B_1) \leq m-r-2$, whence B_1 and B are a monochromatic $N_j D_m^2$ solution.

Case 2: $|\Delta^{-1}(0) \cap [1, 2j-1]| \geq j$.

It follows, as in Case 1, that

$$|\Delta^{-1}(0) \cap [1, 2m+r-1]| \leq m-1. \tag{10.3}$$

Let d be the positive integer such that r is contained in the interval

$$\frac{(d-1)m + dj - d + 1}{d} \leq r < \frac{dm + (d+1)j - (d+1) + 1}{d+1}; \tag{10.4}$$

note, since

$$\lim_{d \to \infty} \frac{(d-1)m + dj - d + 1}{d} = m + j - 1 > m,$$

and since in view of Lemma 13(iii) we have $j \leq r < m$, then it follows that such a d exists. Also note that if $j \geq \frac{m}{d}$, then (10.4) implies $m - 1 < r$, a contradiction. Hence we may assume $j < \frac{m}{d}$. From (10.3) and (10.4), it follows that

$$|\Delta^{-1}(1) \cap [1, 2m + r - 1]| \geq m + r \geq m + \frac{(d-1)m + dj - d + 1}{d}. \tag{10.5}$$

But, letting $b = \mathrm{int}_{m-j+1}(-(\Delta^{-1}(1) \cap [1, 2m+r-1]))$, then it follows in view of $j < \frac{m}{d}$ and (10.5) that

$$|\Delta^{-1}(1) \cap [1, b]| \geq \frac{(d-1)m + dj - d + 1}{d} + j \geq (d+1)(j-1) + 1.$$

Hence let $z_1 < z_2 < \cdots < z_{m-j}$ be the elements of $\mathrm{last}_{m-j}(\Delta^{-1}(1) \cap [1, 2m+r-1])$, and let $y_1 < y_2 < \cdots < y_{(d+1)(j-1)+1}$ be the element of $\mathrm{first}_{(d+1)(j-1)+1}(\Delta^{-1}(1) \cap [1, 2m + r - 1])$. If for some index $i \in [0, d]$

$$|\Delta^{-1}(0) \cap [y_{i(j-1)+1}, y_{(i+1)(j-1)+1}]| \leq m + j - r - 2,$$

then $B_1 = \{y_{i(j-1)+1}, y_{i(j-1)+2}, \ldots, y_{(i+1)(j-1)+1}, z_1, z_2, \ldots, z_{m-j}\}$ is a monochromatic m-set with $g_j(B_1) \leq m + 2j - r - 3 = g_j(B)$, whence B_1 and B are a monochromatic $N_jD^2_m$ solution, and the proof is complete. Therefore, we may assume that

$$|\Delta^{-1}(0) \cap [y_{i(j-1)+1}, y_{(i+1)(j-1)+1}]| \geq m + j - r - 1 \quad \text{for } i = 0, 1, \ldots, d.$$

But then the above inequalities and (10.4) imply that

$$|\Delta^{-1}(0) \cap [1, 2m - 1]| \geq (d+1)(m + j - r - 1) > m - 1,$$

contradicting (10.3), and completing the proof. □

The upper bounds for $f_{zs}(m, j)$ in Theorem 10.3 will follow almost immediately from the following two lemmas.

Lemma 14. *Let m and j be integers satisfying $2 \leq j \leq m$, and let $\Delta : [1, 4m-3] \to \mathbb{Z}/m\mathbb{Z}$ be an arbitrary coloring.*

(i) If m is prime, then there exists a zero-sum m-set $B \subset [1, 4m-3]$ with $g_j(B) \geq m + j - 2$;

(ii) If $j \geq \frac{m}{p} + p - 1$, where p is the smallest prime divisor of m, then there exists a zero-sum m-set $B \subset [1, 4m-3]$ with $g_j(B) \geq m + j - 2$.

Proof. Consider the interval $S = [m+1, 4m-3]$. If there does not exist a $(2m-2)$-set partition of the sequence ΔS with at most $m-1$ sets of cardinality greater than one, then it follows from Proposition 2.3 that there exists $a \in \mathbb{Z}/m\mathbb{Z}$ such that

$$|\Delta^{-1}(a) \cap S| \geq 2m - 1 \quad \text{and} \quad |\Delta^{-1}((\mathbb{Z}/m\mathbb{Z}) \setminus a) \cap S| \leq m - 2.$$

Let $y_1 < y_2 < \cdots < y_{2m-1}$ be elements from $\Delta^{-1}(a) \cap S$, and define

$$B = \{y_1, \ldots, y_{j-1}, y_{m+j-1}, y_{m+j}, \ldots, y_{2m-1}\}.$$

Then $g_j(B) \geq m + j - 2$, and the proof is complete. So we may assume that there exists a $(2m-2)$-set partition P of the sequence ΔS with at most $(m-1)$ sets of cardinality greater than one.

Suppose first that m is prime. Define $x_1 = 1$. Apply CDT to P, and conclude that there exist integers $x_2 < x_3 < \cdots < x_m$ from S such that $\sum_{i=2}^{m} \Delta(x_i) = -\Delta(x_1)$. Thus, $x_1, \ldots, x_m$

is zero-sum. Furthermore, by definition of the x_i's, we have $x_j \geq m+1+(j-2) = m+j-1$, so that $B = \{x_1, \ldots, x_m\}$ satisfies $g_j(B) \geq m + j - 2$, and (i) follows.

To prove (ii), suppose $j \geq \frac{m}{p} + p - 1$, where p is the smallest prime divisor of m. Note $(m-1) \geq (|S|-(m-1))-(m-1) = m-1$. Hence applying Theorem 3.2 to P, it follows that either Theorem 3.2(i) holds, and thus there exists a selection of integers $x_2, \ldots, x_m \in S$ such that $1, x_2, x_3 \ldots, x_m$ is zero-sum, whence the proof is complete as above; or else Theorem 3.2(ii) holds, and thus there exists a coset, which w.l.o.g. we may assume by translation is a proper, nontrivial subgroup, say $a(\mathbb{Z}/m\mathbb{Z}) = H_a$, such that all but at most $a-2$ terms of the sequence ΔS are elements of H_a, whence it follows from EGZ (as remarked in Section 3.1) that any subset $T \subseteq S$ satisfying $|T| \geq m + \frac{m}{a} - 1 + (a-2)$ contains a zero-sum m-tuple. Let

$$S_1 = [m+1, m + \frac{m}{p} + p - 2] \quad \text{and} \quad S_2 = [3m-1, 4m-3].$$

Since $|S_1 \cup S_2| = m + \frac{m}{p} + p - 3 \geq m + \frac{m}{a} - 1 + (a-2)$, it follows that there exist m integers $x_1 < x_2 < \cdots < x_m$ from $S_1 \cup S_2$ such that $\sum_{i=1}^{m} \Delta(x_i) = 0$. Since $|S_2| = m-1$, we must have $x_1 \in S_1$. Furthermore, since $|S_1| = \frac{m}{p} + p - 2 \leq j - 1$, we must have $x_j \in S_2$. Hence it follows that $B = \{x_1, \ldots, x_m\}$ is a zero-sum m-set satisfying $g_j(B) \geq m + j - 2$, whence (ii) is satisfied. $\square$

Lemma 15. *Let m and j be positive integers satisfying $2 \leq j \leq m$, let p be the smallest prime divisor of m, and let $\Delta : [1, 6m + \frac{m}{p} - 5] \to \mathbb{Z}/m\mathbb{Z}$ be an arbitrary coloring. Then one of the following holds:*

(i) there exists a zero-sum m-set $B \subset [1, 6m + \frac{m}{p} - 5]$ satisfying $g_j(B) \geq m + j - 2$;

(ii) there exists a zero-sum $N_j D_m^2$ solution.

Proof. Let D be the sequence $\Delta\left(m + \frac{m}{p}\right), \Delta\left(m + \frac{m}{p} + 1\right), \ldots, \Delta\left(4m + \frac{m}{p} - 4\right)$. Note that the assumption that $j \geq \frac{m}{p} + p - 1$ was used only at the very end of the proof of Lemma 14. Hence repeating the arguments from Lemma 14, applied to the interval $[m + \frac{m}{p}, 4m + \frac{m}{p} - 4]$ rather than $[m+1, 4m-3]$, we may assume that there exists a proper, nontrivial subgroup, say $H_a = a\mathbb{Z}/m\mathbb{Z}$, such that all but at most $a-2$ terms of D are all elements of H_a, and, furthermore, that there exists a $(2m-2)$-set partition P_1 of the terms of D that are elements of H_a such that the sumset of P_1 is H_a (namely the second set partition from Theorem 3.2(ii)). Finally, it follows, in view of Theorem 9.8 applied modulo H_a, that from among the sequence

$$\Delta(1), \Delta(2), \Delta(3), \cdots, \Delta(a)$$

we can find a subsequence D_1 of length $1 \leq q \leq a$ whose terms are consecutive and whose sum is an element $h \in H_a$.

Case 1: $q < j$.

Since $m - q \geq m - a \geq \frac{m}{a} - 1$, then it follows in view of Proposition 2.4(i), by selectively deleting terms from P_1, that we can find an $(m-q)$-set partition P_2 of a subsequence D_2 of D such that the sumset of P_2 is still H_a. Consequently, we can find $m - q$ terms from D_2 with sum $-h$, which, together with the terms of D_1, gives an m-element zero-sum subset B with $g_j(B) \geq m + j - 2$ (since $q < j$ assures $\Delta(\text{int}_j(B)) \notin D_1$).

Case 2: $q \geq j$.

By the arguments in Case 1, we can find an m-element zero-sum set $B_1 \subset [1, 4m + \frac{m}{p} - 4]$ that includes all $q \geq j$ consecutive elements of D_1, and hence $g_j(B_1) \leq j - 1$. From EGZ there exists an m-element zero-sum set $B_2 \subset [4m + \frac{m}{p} - 3, 6m + \frac{m}{p} - 5]$. Since B_1 and B_2 are a zero-sum $N_j D_m^2$ solution, the proof is complete. □

We can now give linear upper bounds for $f_{zs}(m, j)$.

Theorem 10.3. *Let m and j be integers satisfying $2 \le j \le m$, and let p be the smallest prime divisor of m.*

(i) If m is prime or $j \ge \frac{m}{p} + p - 1$, then $f_{zs}(m, j) \le 6m - 4$.

(ii) $f_{zs}(m, j) \le 8m + \frac{m}{p} - 6$.

Proof. Let $s \in \{6m-4, 8m + \frac{m}{p} - 6\}$, and let $\Delta : [1,\, s] \to \mathbb{Z}/m\mathbb{Z}$ be a coloring. From EGZ, it follows that there exists a zero-sum m-set $B \subset [1,\, 2m-1]$, which must satisfy $g_j(B) \le m + j - 2$. The proof of (i) is complete by letting $s = 6m - 4$ and applying Lemma 14(i) or Lemma 14(ii) to $[2m,\, s]$, respectively. To show (ii), set $s = 8m + \frac{m}{p} - 6$, and apply Lemma 15 to $[2m,\, s]$. $\square$

10.3 The Case $j = m - 1$

This section is devoted to improving the upper bound for $f_{zs}(m, m-1)$ from $6m - 4$ to $5m - 4$, which, in view of Theorem 10.1, will show that $N_{m-1}D_m^2$ zero-sum generalizes. For notational convenience, let g denote the modified diameter function g_{m-1} (since the extremal function $g(m, k)$ from Chapter 9 takes a pair of integer arguments, this should pose no great confusion). Further, we will say a $\mathbb{Z}/m\mathbb{Z}$-coloring Δ *reduces to monochromatic* if either $|\Delta(S)| \le 2$ or there exists $B \subset S$ such that $|B| \le m - 1$ and $|\Delta(S \setminus B)| = 1$. Observe that in either case there exists a natural induced coloring $\Delta^* : S \to \{0, 1\}$ such that every m-element monochromatic set under Δ^* is zero-sum under Δ. We begin with a refinement to Lemma 13 for $j = m - 1$.

Lemma 16. *Let $m \ge 3$ be an integer, and let $\Delta : [1, 3m-3] \to \{0, 1\}$ be a coloring. Then one of the following holds:*

(i) there exists a monochromatic m-set $B \subseteq [1, 3m-3]$ with $g(B) \geq 2m-4$;

(ii) there exists a monochromatic $N_{m-1}D_m^2$ solution;

(iii) the coloring Δ is given (up to symmetry) by $1^{m-1}0^{2m-3}1$ or $1^{m-1}0^{2m-4}10$.

Proof. Assume w.l.o.g. $\Delta(1) = 1$. If $|\Delta^{-1}(1)| < m$, then $|\Delta^{-1}(0)| \geq 2m-2$, whence (i) follows. So $|\Delta^{-1}(1)| \geq m$. Let $S = [2m-3, 3m-3]$. Since $\Delta(1) = 1$ and $|\Delta^{-1}(1)| \geq m$, and since $2m-3 > 1$, it follows that if $|\Delta^{-1}(1) \cap S| \geq 2$, then (i) follows. So we can assume otherwise, whence

$$|\Delta^{-1}(0) \cap S| \geq m. \tag{10.6}$$

Let $y_2 < y_3 < \cdots < y_m$ be the elements of $\mathrm{last}_{m-1}(\Delta^{-1}(0) \cap S)$.

Observe, since $|\Delta^{-1}(1) \cap S| \leq 1$, that $y_{m-1} \geq 3m-5$. Hence, if there exists $i \in [1, m-1]$ such that $\Delta(i) = 0$, then (i) follows using $i, y_2, y_3, \ldots, y_m$. Consequently, $\Delta(i) = 1$ for $i \in [1, m-1]$. However, if $\Delta(i) = 1$ for $i \in [1, m]$, then (ii) follows in view of (10.6). Thus $\Delta(m) = 0$, implying that either $\Delta(3m-3) = 1$ or $\Delta(3m-4) = 1$, since otherwise the bound on y_{m-1} will improve by one to $y_{m-1} \geq 3m-4$, whence (i) follows using $m, y_2, y_3, \ldots, y_m$.

Finally, note that if there is $j \in [m, 2m-4]$ with $\Delta(j) = 1$, then $[1, m-1] \cup \{j\}$ will be a monochromatic m-set B_1 with minimal possible modified diameter $g(B_1) = m-2$, whence (ii) follows in view of (10.6). Thus $\Delta([m, 2m-4]) = 0$, which together with $\Delta([1, m-1]) = 1$, with $|\Delta^{-1}(1) \cap S| \geq 1$, and with either $\Delta(3m-3) = 1$ or $\Delta(3m-4) = 1$ (all obtained in the previous paragraphs), implies (iii). □

The proof of the upper bound for $f_{zs}(m, m-1)$ will utilize two additional lemmas. Lemma 17 is a zero-sum version of Lemma 16, including a description of the critical cases, that employs the methods from Chapter 3 as well as ad hoc methods for the handling of the essentially monochromatic and essentially dichromatic cases. If the coloring Δ does not reduce to monochromatic, then Lemma 17 will give the hypotheses of Lemma 18, whose conclusions will rapidly give the desired upper bound. On the other hand, if the coloring Δ does reduce to monochromatic, then (in the main proof) the induced coloring Δ^* and Lemma 16 will give sufficient information about the original coloring Δ to again invoke Lemma 18 to complete the proof.

In Chapter 9, we saw that, up to order and affine transformation, the only m-zsf sequence of length $2m-3$ with at least 3 distinct residue classes is $0^{m-1}1^{m-3}2$ $(*)$. In what follows we will make repeated use of this characterization, which we will reference by $(*)$.

Lemma 17. *If $\Delta : [1, 3m-3] \to \mathbb{Z}/m\mathbb{Z}$ is a coloring with $m \geq 9$, then one of the following holds:*

(i) there exists a zero-sum m-set $B \subseteq [1, 3m-3]$ with $g(B) \geq 2m-3$,

(ii) there exists a zero-sum $N_{m-1}D_m^2$ solution,

(iii) Δ is given up to affine transformation by $1^{m-2}21^{m-2}0^m$, by $1^{m-1}21^{m-3}0^m$ or by $1^{m-3}21^{m-1}0^m$,

(iv) Δ is given up to affine transformation by $1^{m-1}0H$, where H is a block such that there exists a zero-sum m-set $B \subset 0H$ satisfying $g(B) = 2m-4$,

(v) Δ reduces to monochromatic.

Proof. Define $S_1 = \{1, 3m-4, 3m-3\}$ and observe that if there exists a zero-sum m-set that uses all the elements of S_1, then (i) follows. Let $S = [1, 3m-3] \setminus S_1$, and let D be the sequence $\Delta(2), \Delta(3), \ldots, \Delta(3m-5)$.

Case 1: $\Delta([1, 3m-3]) = \{0, 1, 2\}$ and $|\Delta^{-1}(2)| = 1$.

Note that $|\Delta^{-1}(1)| \geq m-2$, as otherwise (v) follows. Therefore there is a zero-sum m-set B satisfying $|B \cap \Delta^{-1}(2)| = 1$, $|B \cap \Delta^{-1}(1)| = m-2$, and $|B \cap \Delta^{-1}(0)| = 1$ that contains $\{1, a, b\}$ for some distinct a, $b \in [2m-2, 3m-3]$, and hence $g(B) \geq 2m-3$ yielding (i), unless every such triple $\{1, a, b\}$ has two of its elements colored by zero. However, this implies either that there exists a monochromatic m-set B with $1 \in B$ and $|B \cap [2m-2, 3m-3]| = m-1$ yielding (i) (if $\Delta(1) = 0$), or that $\Delta([2m-2, 3m-3]) = 0^m$, whence $\Delta(1) \in \{1, 2\}$. Suppose $|\Delta^{-1}(0)| = m$. Then it is easy to see that (iii) holds unless there are m consecutive 1's, in which case (ii) follows. Therefore, we may assume that $|\Delta^{-1}(0)| \geq m+1$. Then $0 \notin \Delta([1, m-1])$ as otherwise (i) follows. Thus $2 \notin \Delta([1, m])$ as otherwise (ii) follows (take for your first set $m-1$ consecutive integers from $[1, m]$ that include an integer colored by 2 along with $\mathrm{int}_1(\Delta^{-1}(0))$, and for your second set choose any other m integers colored by 0). Hence $\Delta(i) = 1$ for $i \in [1, m-1]$ and $\Delta(m) = 0$, whence (iv) follows with $B = \{m\} \cup [2m-1, 3m-3]$.

Case 2: There does not exist $Q \subseteq [1, 3m-3]$ with $|Q| = m+1$ and $|\Delta([1, 3m-3] \backslash Q)| = 1$.

Suppose there does not exist $x \in S$ such that $|\Delta(S \backslash x)| = 2$. Hence, from the assumption of the case and Proposition 2.3, it follows that there is a $(2m-5)$-set partition P' of the terms of D that has at least $(m-2)$ sets of cardinality 1, and consequently at most $m-3$ sets with cardinality greater than one. Let P be the corresponding $(m-3)$-set partition obtained by deleting $m-2$ cardinality one sets from P'. Applying Theorem 3.2 to P (with $|S| = 3m-6$, $|S'| = 2m-4$, and $n = m-3$; note $n = m-3 \geq \frac{m}{2} - 1$ so that the hypotheses of Theorem 3.2 are satisfied), we conclude that either Theorem 3.2(i) holds—whence the cardinality of the sumset of the resulting $(m-3)$-set partition will be $|\mathbb{Z}/m\mathbb{Z}|$, allowing us to choose a selection of $m-3$ terms whose sum is the additive inverse of the sum of terms from S_1, yielding (i)—or else that Theorem 3.2(ii) holds, whence all but at most $a-2+3$ of the elements of $[1, 3m-3]$ are colored by elements from the same coset $\alpha + H_a$ of $\mathbb{Z}/m\mathbb{Z}$, where H_a has index a with $1 < a < m$. Hence, as remarked in Section 3.1, it follows from EGZ that any subset of $[1, 3m-3]$ of cardinality $(m + \frac{m}{a} - 1 + a + 1)$ must contain a zero-sum

m-set. Thus there is a zero-sum m-set

$$B \subseteq [1, m-2] \cup [3m-4-a-\frac{m}{a},\ 3m-3],$$

and as $\frac{m}{a}+a+2 \leq m-1$ for $m \geq 9$, it follows that

$$g(B) \geq 3m-5-a-\frac{m}{a} \geq 2m-2,$$

whence (i) follows.

So we may assume that there exists $x \in S$ such that $|\Delta(S \setminus x)| = 2$ (i.e., that S is essentially dichromatic). One of the sets $S_2 = \{2, 3m-5, 3m-6\}$, $S_3 = \{3, 3m-5, 3m-6\}$, $S_4 = \{2, 3m-7, 3m-6\}$ or $S_5 = \{2, 3m-7, 3m-5\}$, say S_3, does not contain x. Since $x \in S'$, we may apply the arguments of the preceding paragraph to $S' = [1, 3m-3] \setminus S_3$ and conclude that $[1, 3m-3] \setminus \{x\}$ must be colored by two residue classes, say α_1, α_2, since otherwise (i) or (v) follows (since, letting $\Delta(S) = \{\alpha_1, \alpha_2, 2\}$, if w.l.o.g. α_1 colors at most one term in S', then α_1 colors at most $1+|S_3|+|S_1| = 7$ integers in total, whence α_2 colors all but $8 \leq m-1$ integers, yielding (iv)). Furthermore, we conclude that $\Delta(x) = \beta \notin \{\alpha_1, \alpha_2\}$ as otherwise (v) again follows.

Let $\alpha_1 - \alpha_2 = a$. If $(a, m) \neq 1$, then EGZ (as remarked in Section 3.1) implies that any subset of $[1, 3m-3]$ of cardinality $m+\frac{m}{a}-1+1$ contains a zero-sum m-set, whence the proof is complete by the arguments at the end of the first paragraph of Case 2. So, $(a, m) = 1$, and hence by an affine transformation we may assume that $\{\alpha_1, \alpha_2\} = \{0, 1\}$. Furthermore, if $\Delta(x)$ is not equal to 2 or -1, then there will be a zero-sum m-set B satisfying $|B \cap \{x\}| = 1$, $|B \cap \Delta^{-1}(1)| = m - \overline{\Delta(x)} \geq 2$, and $|B \cap \Delta^{-1}(0)| = \overline{\Delta(x)} - 1 \geq 2$ that contains $\{1, a, b\}$ for some distinct $a, b \in [2m-1, 3m-3]$, and hence $g_j(B) \geq 2m-2$, unless every pair $\{a,\, b\}$

satisfies $\Delta(1) = \Delta(a) = \Delta(b)$, in which case $B = \{1\} \cup [2m-1, 3m-3]$ is a monochromatic m-set B with $g_j(B) \geq 2m-2$. In both cases (i) follows. Hence, by the affine transformation exchanging 0 and 1 if $\Delta(x) = -1$, this reduces to Case 1.

Case 3: There exists $Q \subseteq [1, 3m-3]$ such that $|Q| = m+1$ and $|\Delta([1, 3m-3] \setminus Q)| = 1$ (i.e, S is essentially monochromatic).

Assume w.l.o.g. $\Delta([1, 3m-3] \setminus Q) = \{0\}$. Let R denote a sequence of $m-1$ 0's. Define $C = Q \setminus \Delta^{-1}(0)$. Observe that if $|C| \leq m-1$, then (v) follows.

First assume that $|C| = m$. Let S_1 range over all possible subsequences of ΔC of length $m-2$. Hence, since $|\Delta(C)| \geq 2$ else (v) follows, then applying $(*)$ to each $S_1 \cup R$, it follows that there exists a zero-sum subset $C' \subset C$ such that $1 < |C'| \leq m-2$, unless w.l.o.g. $\Delta(C) = \{1, 2\}$ and $|\Delta^{-1}(2) \cap C| = 1$, which reduces to Case 1. So we may assume such C' exists.

Let $y_1 = \text{int}_1(\Delta^{-1}(0))$, $y_2 = \text{int}_2(-\Delta^{-1}(0))$, and $y_3 = \text{int}_1(-\Delta^{-1}(0)) = \max(\Delta^{-1}(0))$. Notice that there will be a monochromatic m-set B with $g(B) \geq 2m-3$ unless at least $m-1$ elements of C lie in $[1, y_1 - 1] \cup [y_2 + 1, 3m-3]$. Hence, since $2 \leq |C'| \leq m-2$, it follows that C' in addition to $m - |C'|$ elements colored by zero, including y_1, y_2 and y_3 (if $|C'| < m-2$) or y_1 and y_3 (if $|C'| = m-2$, $\max(C') > y_2$) or y_2 and y_3 (if $|C'| = m-2$, $\max(C') < y_2$) will form a zero-sum m-set B satisfying $g(B) \geq 2m-3$, yielding (i).

So assume that $|C| = m+1$. As above, we may assume that there exists a zero-sum subset $C' \subset C$ such that $2 \leq |C'| \leq m-2$. If $|C'| \geq 3$, then, as in the previous paragraph, it follows that C' in addition to $m - |C'|$ elements colored by zero, including y_1, y_2 and y_3 (if $|C'| < m-2$) or y_1 and y_3 (if $|C'| = m-2$, $\max(C') > y_2$) or y_2 and y_3 (if $|C'| = m-2$, $\max(C') < y_2$) will form a zero-sum m-set B satisfying $g(B) \geq 2m-3$, yielding (i). So we can assume all such zero-sum subsets C' of C have cardinality two.

Since $m-2 \geq 4$, and since all zero-sums C' have cardinality two, it follows that any two such zero-sums must intersect (else the union of disjoint ones would give a zero-sum of size $4 \leq m-2$). Suppose the intersection of all the 2-term zero-sum subsets of C is empty. Hence there must be exactly three 2-term zero-sums that pairwise intersect each other with empty three-fold intersection (there can be no more, else there are two disjoint ones, and no fewer, else we contradict the previous sentence). Since this is only possible if all three of these zero-sums are monochromatic in $\frac{m}{2}$, it follows that there are exactly three integers x_1, x_2 and x_3 colored by $\frac{m}{2}$ (there can be no more, else we have a 4-term zero-sum consisting of four elements colored by $\frac{m}{2}$). Let $Y = C \setminus \{x_1, x_2, y\}$, where $y \in C$ is such that $\Delta(y) \neq \frac{m}{2}$. Then Y is colored by at least two distinct residues, including $\frac{m}{2}$. Hence applying $(*)$ to $R \cup Y$ yields a zero-sum $C'' \subseteq Y \subseteq C$ with $2 \leq |C''| \leq |Y| = m-2$. However, since $x_1, x_2 \notin C''$, it follows that C'' must be distinct from the original three zero-sum subsets, contradicting that C contained exactly three zero-sum subsets of size at most $m-2$. So we may assume there is a term $z \in C$ such that z is contained in every zero-sum subset $C' \subseteq C$ with $2 = |C'| \leq m-2$.

Applying the arguments of the second paragraph of Case 3 to $C \setminus \{z\}$, we contract the uniqueness of $z \in C'$, or we conclude w.l.o.g. that $\Delta(C \setminus \{z\}) \subseteq \{1, 2\}$ and $|\Delta^{-1}(2) \cap (C \setminus \{z\})| \leq 1$. Since z is one element of a two element zero-sum set, it follows that we must have $\Delta(z) = -1$ or $\Delta(z) = -2$. If $\Delta(z) = -2$, then we can find C' with $\Delta C' = -21^2$, and this reduces to the case $|C'| \geq 3$. So we can assume $\Delta(z) = -1$. Furthermore, we can assume $|\Delta^{-1}(2) \cap C| = 1$, else the affine transformation exchanging 0 and 1 reduces to the hypotheses of Case 1. Thus C is colored (up to order) by the sequence $1^{m-1}(-1)2$. Let z' be the element colored by 2.

Hence the pair $\{z, c\}$ is zero-sum for every $c \in C \setminus \{z, z'\}$. Let $z_1 < z_2$ be the elements of

$\mathrm{first}_2(C)$, and let $z_3 < z_4 < z_5$ be the elements of $\mathrm{last}_3(C)$. As noted before, at least $m-1$ elements of C lie in $[1, y_1 - 1] \cup [y_2 + 1, 3m - 3]$, so that at most 2 elements of C can lie in $[y_1, y_2]$. Since $m - 1 \geq 7$, it follows that one of $[1, y_1 - 1]$ and $[y_2 + 1, 3m - 3]$ must contain at least 4 elements of C. Hence, if $[1, y_1 - 1]$ contains at least 4 elements from C, then we can choose C' so that it contains z_1 or z_2, whence C' in addition to $m - 2$ elements colored by zero, including y_1, y_2 and y_3, will form a zero-sum m-set B satisfying $g(B) \geq 2m - 3$, yielding (i). Therefore we can assume otherwise, whence $[y_2 + 1, 3m - 3]$ must contain at least 4 elements of C.

In this case, we can choose C' so that it contains one of z_5 or z_4, whence C' in addition to $m - 2$ elements colored by zero, including y_1, y_2 and y_3, will form a zero-sum m-set B satisfying $g(B) \geq 2m-3$, yielding (i), unless $|C \cap [y_1, y_2]| \leq 1$. Hence there must be at least m elements of C outside $[y_1, y_2]$, at most three less than y_1 (from the conclusion of the last paragraph), and consequently at least $m - 4 \geq 5$ elements of C greater than y_2. Thus we must have $z \leq y_2 + 2$, else we can choose C' so that it contains z and one of z_3 or z_4 or z_5 that is distinct from z, forming (as before) a zero-sum m-set B satisfying $g(B) \geq 2m - 3$, yielding (i). Hence, since there are at least five elements of C greater than y_2, it follows that at least two of y_5, y_4 and y_3 must be colored by 1, say y_{l_1} and y_{l_2}. But then the m-set consisting of y_{l_1}, y_{l_2}, $m - 4$ additional elements colored by 1, z_1, and z' (recall $\Delta(z') = 2$) forms a zero-sum subset B with $g(B) \geq 2m - 3$, completing the proof. $\square$

Lemma 18. *Let $m \geq 5$ be an integer, and let $\Delta : [1, 5m-4] \to \mathbb{Z}/m\mathbb{Z}$ be a coloring. If there is an integer $\gamma \geq 2m$ with $\Delta([\gamma, \gamma + m - 4]) = \{z\}$, and zero-sum m-sets B_i, $i = 2, 3, 4, 5$, with $B_2 \subset [\gamma, 5m - 4]$ and $g(B_2) \geq 2m - 4$, with $B_3 \subset [\gamma + 1, 5m - 4]$ and $g(B_3) \geq 2m - 5$, with $B_4 \subset [\gamma + \left\lfloor \frac{m}{2} \right\rfloor, 5m-4]$ and $g(B_4) \geq m + \left\lceil \frac{m}{2} \right\rceil - 4$, and with $B_5 \subset [r+1, 5m-4]$, where r is an integer $r \geq \gamma + m - 3$ with $\Delta(r) = z$, then there is a zero-sum $N_{m-1}D_m^2$ solution.*

Proof. We may w.l.o.g. assume $z = 0$. Let $S = [\gamma - 2m + 1, \gamma - 1]$, $S_1 = [\gamma - 2m + 2, \gamma - 1]$ and $S_2 = [\gamma - 2m + 1, \gamma - 3] \cup \{\gamma - 1\}$. Since $g(B_2) \geq 2m - 4$, we can assume that neither S_1 nor S_2 contains a zero-sum m-set (else the proof is complete), whence the result $g(m, 3) = 2m - 2$ from Chapter 9 implies that $|\Delta(S)| = 2$. Let $S_3 = [\gamma - 2m + 4, \gamma]$. Since $g(B_3) \geq 2m - 5$, we conclude that there does not exist a zero-sum m-set in S_3.

We proceed to show that w.l.o.g. $\Delta(S) = \{1, 2\}$ or $\Delta(S) = \{0, b\}$ with $b \neq 0$. If $|\Delta(S_3)| \leq 2$, then S_3 m-zsf implies that S_3 is colored by exactly two residue classes, each with multiplicity at least $m - 2 \geq 2$ (since otherwise there is a monochromatic m-term subset), one of which must be zero (since $\gamma \in S_3$ with $\Delta(\gamma) = 0$). Hence, since $|\Delta(S)| = 2$, since $|S_3 \setminus S| = 1$, and since $m - 2 \geq 2$, it follows that $\Delta(S) = \{0, b\}$ with $b \neq 0$. If $|\Delta(S_3)| \geq 3$, then from $(*)$ it follows that $|\Delta(S_3)| = 3$. If there are at least 2 integers in S_3 colored by 0, then $0 \in \Delta(S)$, whence $|\Delta(S)| = 2$ implies that $\Delta(S) = \{0, b\}$ with $b \neq 0$. Otherwise, it follows, in view of $(*)$ and an appropriate affine transformation fixing $\Delta(\gamma) = 0$, that $\Delta(S_3 \setminus \gamma) = \{1, 2\}$. Hence, since $|\Delta(S)| = 2$, and since $S_3 \setminus \gamma \subseteq S$, then it follows that $\Delta(S) = \{1, 2\}$, completing the proof of the claim.

We first handle the case when $\Delta(S) = \{1, 2\}$. Let δ be the maximal integer such that

$$s = \sum_{i=\delta}^{\gamma-1} \overline{\Delta(i)} \geq m.$$

Then $\gamma - m \leq \delta \leq \gamma - \lceil \frac{m}{2} \rceil$. Notice that $s \in \{m, m + 1\}$. Furthermore, if $s = m$, then $B_1 = \{\delta, \delta + 1, \ldots, \delta + m - 1\}$ satisfies $g(B_1) = m - 2$, whence B_1 and B_4 are a zero-sum $N_{m-1}D_m^2$ solution.

Therefore we may assume that $s = m + 1$. Suppose there exists $j \in [\delta, \gamma - 1]$ such that $\Delta(j) = 1$. If m is even, then $\delta < \gamma - \lceil \frac{m}{2} \rceil$. On the other hand, if m is odd, then since

$s = m + 1$, it follows that there are at least two integers colored by 1 in $[\delta, \gamma - 1]$, whence $\delta < \gamma - \lceil \frac{m}{2} \rceil$ as well. Thus $B_1 = \{\delta, \delta + 1, \ldots, \delta + m\} \setminus \{j\}$ is a zero-sum m-set satisfying $g(B_1) = m - 1$, which together with B_4 yields a zero-sum $N_{m-1}D_m^2$ solution

So we may assume that $\Delta(j) = 2$ for $j \in [\delta, \gamma - 1]$, whence m is odd as $s = m + 1$. We may assume that there exists a maximal integer $\gamma - m \leq \beta \leq \gamma - 1$ such that $\Delta(\beta) = 1$, since otherwise $B_1 = \{\gamma - m, \gamma - m + 1, \ldots, \gamma - 1\}$ is a zero-sum m-set satisfying $g(B_1) \leq m - 2$, and the proof is complete as in the preceding paragraph. If $\beta \geq \gamma - m + 1$, then the set $B = \{\beta\} \cup [\gamma - \frac{m-1}{2}, \gamma - 1 + \frac{m-1}{2}]$ is a zero-sum m-set $B \subset [\beta, \gamma - 1 + \frac{m-1}{2}]$ satisfying $g(B) \leq \frac{3m-7}{2}$. But then B and B_4 are a zero-sum $N_{m-1}D_m^2$ solution. Therefore, we may assume that $\beta = \gamma - m$, whence $\Delta[\gamma - m + 1, \gamma - 1] = 2^{m-1}$. Hence, since $B_2 \subset [\gamma, 5m - 4]$ is such that $g(B_2) \geq 2m - 4$, it follows that $\Delta(j) = 1$ for $j \in [\gamma - 2m + 3, \gamma - m]$. But then $B_1 = [\gamma - 2m + 3, \gamma - m + 1] \cup \{\gamma\}$ satisfies $g(B_1) = m - 2$, whence B_1 and B_4 form a zero-sum $N_{m-1}D_m^2$ solution, completing the case when $\Delta(S) = \{1, 2\}$.

Next we handle the remaining case when $\Delta(S) = \{0, b\}$. From the pigeonhole principle it follows that there exists a monochromatic m-set $B \subset [\gamma - 2m + 1, \gamma - 1]$. Since $g(B_2) \geq 2m - 4$, we may assume that $g(B) = 2m - 3$, whence $\Delta(\gamma - 2m + 1) = \Delta(\gamma - 2) = \Delta(\gamma - 1)$ and $|\Delta^{-1}(\Delta(\gamma - 1))| = m$. If $\Delta(\gamma - 1) = 0$, then $B_1 = \{\gamma - 2, \gamma - 1, \ldots, \gamma + m - 4, r\}$ and B_5 are a zero-sum $N_{m-1}D_m^2$ solution. So we may assume that $\Delta(\gamma - 1) = b$. Let $y_1 < y_2 < \cdots < y_{m-1}$ be elements from $\Delta^{-1}(0) \cap [\gamma - 2m + 1, \gamma - 2]$. Then $B_1 = \{y_1, y_2, \ldots, y_{m-1}, \gamma\}$ and B_3 are a zero-sum $N_{m-1}D_m^2$ solution. □

We conclude the chapter with the main result of this section.

Theorem 10.4. *If $m \geq 9$ is an integer, then $f_{zs}(m, m - 1) = f(m, m - 1) = 5m - 4$.*

Proof. From Theorem 10.1 it follows that $5m - 4 \leq f(m, m - 1) \leq f_{zs}(m, m - 1)$. It remains to show that $f_{zs}(m,m-1) \leq 5m-4$. Let $\Delta : [1,5m-4] \to \mathbb{Z}/m\mathbb{Z}$ be an arbitrary coloring. From EGZ, it follows that there exists a zero-sum m-set $B \subset [1,2m-1]$ with $g(B) \leq 2m-3$. Therefore, applying Lemma 17 to $S = [2m, 5m-4]$, we may assume that neither (i) nor (ii) hold. If (iii) holds, then the proof is complete by Lemma 18 with $\gamma = 2m$. If (iv) holds, then the proof is again complete by Lemma 18 with $\gamma = 2m$ (let $B = B_i$ for all $i \in [2,5]$). Thus, we may assume that conclusion (v) of Lemma 17 holds when applied to S. Let $\Delta^* : S \to \{0,1\}$ be the natural induced coloring whose monochromatic m-sets are all zero-sum under Δ.

Then we may apply Lemma 16 to S and Δ^* and assume that conclusion (ii) does not hold. Suppose first that conclusion (iii) of Lemma 16 holds. Then

$$\Delta^*(S) = 0^{m-1}1^{2m-4}01 \qquad \text{or} \qquad \Delta^*(S) = 0^{m-1}1^{2m-3}0,$$

implying w.l.o.g., since each color class is used at least m times, that

$$\Delta(S) = 0^{m-1}a^{2m-4}0a \qquad \text{or} \qquad \Delta(S) = 0^{m-1}a^{2m-3}0, \tag{10.7}$$

where $a \in \mathbb{Z}/m\mathbb{Z}$ is nonzero. From (10.7) it follows that there is a monochromatic in a subset D_1 of $[3m-1, 5m-4]$ with $g(D_1) \geq 2m-5$. Hence applying $(*)$ to $[m+2, 3m-2]$, it follows that the proof is complete unless either $\Delta([m+2, 2m-1]) = b$, where $b \neq 0$, or else w.l.o.g. $\Delta([m+2, 2m-1] \setminus \{x\}) = 1$, and $\Delta(x) = 2$, for some $x \in [m+2, 2m-1]$. In the latter case, it can be checked that there is an m-element zero-sum subset $B' \subset [m+2, 3m-1]$ with $3m-1 \in B'$, and $g(B') \leq m-1$ (using, up to order, the zero-sum sequence $1^{m-\overline{a}}a0^{\overline{a}-1}$, if $\overline{a} \geq 3$, or $1^{m-4}20^2a$, if $a = 2$, or $1^{m-3}20a$, if $a = 1$). Likewise, in the former case if $b \neq a$, then it can be checked that there is an m-element zero-sum subset $B' \subset [m+2, 3m-1]$

with $g(B') \leq m-2$ (since $\Delta[m+2, 3m-1] = b^{m-2}0^{m-1}a$, with $a \neq b$, then applying the result $g(m,3) = 2m-2$ from Chapter 9 yields an m-term zero-sum, which can be chosen so that all elements colored by b or 0 are consecutive). Hence, since in view of (10.7) there is a monochromatic (in a) m-set D_2 with $g(D_2) \geq 2m-6 \geq m-1$ and $\min D_1 > 3m-1$, it follows that the proof is complete. So we may assume $\Delta([m+2, 2m-1]) = a$.

If $[5, m+1] \cap \Delta^{-1}(a) \neq \emptyset$, then there will be an m-element monochromatic in a subset $B' \subset [5, 3m-1]$, with $g(B') \leq 2m-6$, which together with D_2 completes the proof. Hence from $(*)$ applied to $[5, 2m+1]$, it follows that either $\Delta([5, m+1]) = 0$, or else there exists an m-element zero-sum subset $B' \subseteq [5, 2m+1]$ with $g(B') \leq 2m-5$. In the latter case B' and D_1 complete the proof. Therefore we may assume that $\Delta([5, m+1]) = 0$. Likewise, if $\Delta([1,4]) \nsubseteq \{0, a\}$, then the proof will be complete by applying $(*)$ to both $[1, 2m-4] \cup \{2m\}$ and $[1, 2m-3]$. So we can conclude $\Delta([1, 2m-1]) \subseteq \{0, a\}$.

If there exist integers $j_1 < j_2$ from $[1,4]$ such that $\Delta(j_i) = 0$ for $i = 1$ and $i = 2$, then $B_1 = \{j_1, j_2, 5, 6, 7, \ldots, m+1, 2m\}$ is a monochromatic m-set with $g(B_1) \leq m$, which along with D_1 once more completes the proof. Therefore, we can assume that there exist integers $j_1 < j_2 < j_3$ from $[1,4]$ such that $\Delta(j_i) = a$ for $i = 1, 2, 3$, whence

$$B_1 = \{j_1, j_2, j_3, m+2, m+3, \ldots, 2m-2\}$$

is a monochromatic m-set with $g(B_1) \leq 2m-4$. However, since $\Delta(2m-1) = a$, it follows from (10.7) that there exists a monochromatic m-set $B_2 \subset \{2m-1\} \cup [4m-3, 5m-4]$ such that $g(B_2) \geq 2m-4$, and the proof is complete.

So we may assume that conclusion (i) of Lemma 16 holds. We consider two cases.

Case 1: There exists $c \in \{0, 1\}$ such that $|\Delta^{*-1}(c)| \leq m-1$.

Without loss of generality $c = 1$. It follows that $|\Delta^{*-1}(0)| \geq 2m - 2$. Furthermore, we may assume that the first $2m-3$ of the integers colored by 0 are consecutive, since otherwise under Δ we obtain a zero-sum m-set B_2 satisfying $g(B_2) \geq 2m - 3$, which together with B completes the proof. Applying Lemma 18 with $\gamma = \min\{\Delta^{-1}(b) \cap S\}$, where b is the color such that $\Delta^{-1}(b) \geq 2m - 2$, completes Case 1.

Case 2: There does not exist $c \in \{0, 1\}$ such that $|\Delta^{*-1}(c)| \leq m - 1$.

In this case $|\Delta(S)| \leq 2$ and w.l.o.g. we may assume $\Delta(S) = \{0, a\}$ and that there exist two integers $i_1, i_2 \in [5m - 6, 5m - 4]$ such that $\Delta(i_1) = \Delta(i_2) = a$. Hence $x = \min\{\Delta^{-1}(a) \cap S\}$ satisfies $x \geq 3m-2$, as otherwise there will be an m-set B_2 monochromatic in a satisfying $g(B_2) \geq 2m - 3$, which along with B completes the proof. Notice that $x \leq 3m - 1$ as otherwise $[2m, 3m - 1]$ is a monochromatic m-set that, along with any m elements colored by a, forms a zero-sum $N_{m-1}D_m^2$ solution. But then since conclusion (i) of Lemma 16 holds for $[2m, 5m - 4]$, and since $(5m - 6) - (3m - 1) = 2m - 5 \geq m + \lceil \frac{m}{2} \rceil - 4$, it follows, in view of Lemma 18 with $\gamma = 2m$, that the proof is complete. $\square$

Chapter 11

The Erdős-Ginzburg-Ziv Theorem and Hypergraphs

11.1 Discussion

The original nondecreasing diameter problem of Bialostocki, Erdős and Lefmann, as well as the zero-sum problems from Chapters 7 and 10, all dealt with finding not just one zero-sum subset, but a pair of such sets, each individually zero-sum. However, in all cases the paired zero-sum subsets were disjoint. One might also wonder about zero-sum generalizations for multiple m-sets with a prescribed intersection structure.

If we think of the sequence S of length n (in which we are trying to find the collection of zero-sum subsequences) as being a $\mathbb{Z}/m\mathbb{Z}$-coloring of the vertices of the complete m-uniform hypergraph $\mathcal{K}_n^m$, then the edges of $\mathcal{K}_n^m$ correspond to the m-term subsequences of S. A collection of m-term subsequences with a prescribed intersection structure is then just some m-uniform hypergraph $\mathcal{H}$, whose vertex set we denote by $V(\mathcal{H})$, and whose edge set we denote by $E(\mathcal{H})$. If every $e \in E(\mathcal{H})$ satisfies $\sum\limits_{v \in e} \Delta(v) = 0$, then we say the m-uniform hypergraph $\mathcal{H}$ is zero-sum.

Armed with this notation, we can define what it would mean for a given m-uniform hypergraph $\mathcal{H}$ to zero-sum generalize. Let $f(\mathcal{H})$ (let $f_{zs}(\mathcal{H})$) be the least integer n such that for every 2-coloring (coloring with the elements of $\mathbb{Z}/m\mathbb{Z}$) of the vertices of $\mathcal{K}_n^m$, there exists a subhypergraph $\mathcal{K}$ isomorphic to $\mathcal{H}$ such that every edge e in $\mathcal{K}$ is monochromatic (such that $\mathcal{K}$ is zero-sum). From the pigeonhole principle it is clear that $f(\mathcal{H}) \leq 2|V(\mathcal{H})|-1$, with equality holding if H is connected. Then the m-uniform hypergraph $\mathcal{H}$ zero-sum generalizes if $f_{zs}(\mathcal{H}) = f(\mathcal{H})$, which in the connected case simply means $f_{zs}(\mathcal{H}) = 2|V(\mathcal{H})| - 1$. The Erdős-Ginzburg-Ziv Theorem is then the statement that there is a zero-sum generalization for the m-uniform hypergraph consisting of a single edge.

Not every hypergraph zero-sum generalizes. For instance, a complete m-uniform hypergraph on $k > m$ vertices is easily seen to require $m(k-1)+1$ vertices to guarantee a zero-sum copy of itself (which will necessarily be monochromatic). Note that $m(k-1)+1 > 2k-1$, for $m > 2$, and so no zero-sum generalization is present. However, the goal of this chapter is to show that a zero-sum generalization does occur provided the hypergraph has very little intersection structure. More concretely, we will be able to show a zero-sum generalization for any m-uniform hypergraph on two edges, and any hypergraph with 'many' monovalent vertices (vertices contained in precisely one edge). The proofs are simple applications of the combined machinery of Chapters 3, 4 and 6, and were the original motivation for developing the results from Chapters 4, 5 and 6.

11.2 EGZ in Hypergraphs

Theorem 11.1 below can be used to show a zero-sum generalization for an m-uniform hypergraph that can be iteratively constructed by first starting with a zero-sum generalizing hypergraph (like a single edge or pair of edges), and then adding edges, one by one, so that each added edge has—at the time of its addition—at least half its vertices monovalent.

Theorem 11.1. *Let $\mathcal{H}$ be a finite m-uniform hypergraph, let $e \in E(\mathcal{H})$, and let $\mathcal{H}'$ be the subhypergraph obtained by removing the edge e and all monovalent vertices contained in e. If $f_{zs}(\mathcal{H}') \leq 2|V(\mathcal{H}')| - 1$ and e has at least $\lceil \frac{m}{2} \rceil$ monovalent vertices, then $f_{zs}(\mathcal{H}) \leq 2|V(\mathcal{H})| - 1$.*

Proof. Let S denote the sequence given by a coloring $\Delta : V \to \mathbb{Z}/m\mathbb{Z}$, where $n = |V(\mathcal{H})|$ and $V = V(K^m_{2n-1})$. Let s be the number of non-monovalent vertices in e. Note that by assumption $s \leq \lfloor \frac{m}{2} \rfloor$. We may assume that the multiplicity of each term in S is at most $n-1$, else there will be a zero-sum copy of $\mathcal{H}$ with all edges monochromatic. Hence, if there exists a subset $X \subseteq V$ such that $|X| \leq s - 2 \leq \lfloor \frac{m}{2} \rfloor - 2$ and $|\Delta(V \setminus X)| \leq 2$, then setting aside $n - m$ terms colored by a_i for each of the two $a_i \in \Delta(V \setminus X)$ and applying Theorem 4.1 to the remaining $2m - 1$ terms, it follows that there exists an edge-wise zero-sum copy of $\mathcal{H}$ with the vertices of e colored by the zero-sum sequence given by Theorem 4.1 and all other edges monochromatic. Otherwise, since $s \leq \lfloor \frac{m}{2} \rfloor$, then it follows from Proposition 2.3 that there exists an $(2n - m)$-set partition P' of S with at least $2n - 2m + s$ cardinality one sets. Let P be the $(m - s)$-set partition obtained from P' by removing $2n - 2m + s$ cardinality one sets. Since $s \leq \lfloor \frac{m}{2} \rfloor$, it follows that $m - s \geq \frac{m}{2} - 1$, whence we can apply Theorem 3.2 to P, yielding two cases.

If Theorem 3.2(i) holds, then let A be the corresponding $(m-2)$-set partition given by (i). Applying Theorem 6.1 to the set partition A yields an $(m - s)$-set partition A' that contains at most $2(m - s)$ terms of S, and whose sumset is $\mathbb{Z}/m\mathbb{Z}$. This leaves at least $2n - 1 - 2(m - s) = 2(n - m + s) - 1 \geq 2|V(\mathcal{H}')| - 1$ vertices not contained in any term of A'. Thus, since $f_{zs}(\mathcal{H}') \leq 2|V(\mathcal{H}')| - 1$, it follows that there exists an edge-wise zero-sum copy of $\mathcal{H}'$ not containing any vertices contained in A'. Hence, since the sumset of terms in the $(m - s)$-set partition A' is $\mathbb{Z}/m\mathbb{Z}$, it follows that we can find $m - s$ vertices from A'

which together with the vertices of $\mathcal{H}'$ form an edge-wise zero-sum copy of $\mathcal{H}$.

If Theorem 3.2(ii) holds, then there exists a proper nontrivial subgroup H_a of index a such that all but at most $a-2$ terms of S are from the coset $\alpha + H_a$, and w.l.o.g. by translation we may assume $\alpha = 0$; furthermore, there exists a subsequence S' of S of length at most $m-s+\frac{m}{a}-1$ with an $(m-s)$-set partition $P' = P'_1, \ldots, P'_{m-s}$ satisfying $\sum_{i=1}^{m-s} P'_i = H_a$. Hence it follows that there are at least $2n-1-(m-s+\frac{m}{a}-1)-(a-2) \geq 2n-1-2(m-s)$ terms of S that are not used in the set partition P', and which are from H_a, whence the proof is complete as it was in the previous paragraph. □

A simple corollary of Theorem 11.1 is the following result.

Theorem 11.2. *Let $\mathcal{H}$ be a connected, finite m-uniform hypergraph. If every subhypergraph $\mathcal{H}'$ of $\mathcal{H}$ contains an edge with at least half of its vertices monovalent in $\mathcal{H}'$, then $\mathcal{H}$ zero-sum generalizes.*

Proof. If $\mathcal{H}$ has one edge, this is precisely a restatement of the Erdős-Ginzburg-Ziv Theorem. Hence the upper bound for Theorem 11.2 follows from Theorem 11.1 and induction on the number of edges (relaxing the connectedness condition), while the lower bound for connected $\mathcal{H}$ is trivial. □

The final zero-sum generalizing result of this section will require the following simple proposition, easily proved by induction on s.

Proposition 11.3. *Let m and s be positive integers, and let S be a sequence of elements from an abelian group of order m. If $|S| \geq m+2s-1$, then there exist two disjoint s-term subsequences of S whose sums are equal.*

Theorem 11.4. *If $\mathcal{H}$ is a hypergraph that consists of two intersecting m-sets, then $\mathcal{H}$ zero-sum generalizes.*

Proof. Let S denote the sequence given by a coloring $\Delta : V \to \mathbb{Z}_m$, where $n = |V(\mathcal{H})|$ and $V = V(K_{2n-1}^m)$. Let the two edges of $\mathcal{H}$ be A and B. If $|A \cap B| < \lceil \frac{m}{2} \rceil$, then the proof is complete by Theorem 11.2. So we may assume $|A \cap B| \geq \lceil \frac{m}{2} \rceil$. Let $s = m - |A \cap B|$. Note $n = m + s$, $|S| = 2m + 2s - 1$, and $s \leq \lfloor \frac{m}{2} \rfloor$.

We may also assume that the multiplicity of each term in S is at most $n - 1$, else there will be a zero-sum copy of $\mathcal{H}$ with all edges monochromatic. Hence, if there exists a subset $X \subseteq V$ such that $|X| \leq \lceil \frac{m}{2} \rceil - 2$ and $|\Delta(V \setminus X)| \leq 2$, then setting aside s terms colored by a_i for each of the two $a_i \in \Delta(V \setminus X)$, and applying Theorem 4.1 to the remaining $2m-1$ terms, it follows that there exists an edge-wise zero-sum copy of $\mathcal{H}$ with the vertices of A colored by the zero-sum sequence given by Theorem 4.1, and with $V(\mathcal{H}) \setminus (A \cap B)$ monochromatic. Otherwise, it follows from Proposition 2.3 that there exists an $(m + 2s)$-set partition P' of S with at least $\lceil \frac{m}{2} \rceil + 2s$ cardinality one sets. Let P be the $\lfloor \frac{m}{2} \rfloor$-set partition obtained from P' by removing $\lceil \frac{m}{2} \rceil + 2s$ cardinality one sets. Applying Theorem 3.2 to P yields two cases.

If Theorem 3.2(i) holds, then let A' be the set partition given by (i). Applying Theorem 6.1 to the set partition A' yields an $\lfloor \frac{m}{2} \rfloor$-set partition A'' that contains at most m terms of S, and whose sumset is $\mathbb{Z}/m\mathbb{Z}$. This leaves at least $m + 2s - 1$ vertices not contained in any term of A''. Hence from Proposition 11.3, it follows that there are two disjoint s-term subsequences S_1 and S_2, none of whose terms are contained in a term of A'', and whose sums are equal to (say) t. Since $s \leq \lfloor \frac{m}{2} \rfloor$, then let T be a subsequence of length $m - s - \lfloor \frac{m}{2} \rfloor$ whose terms are not contained in S_1, S_2, nor any term of A''. Let t' be the sum of the terms in T if T is nonempty, and otherwise let $t' = 0$. Since $s \leq \lfloor \frac{m}{2} \rfloor$, and since the sumset of A'' is $\mathbb{Z}/m\mathbb{Z}$, it follows that we may choose $\lfloor \frac{m}{2} \rfloor$ terms of S from A'' whose sum is $-(t + t')$, which along with S_1, S_2 and T yields a zero-sum copy of $\mathcal{H}$ with the terms from A'' and T contained in $A \cap B$.

If Theorem 3.2(ii) holds, then there exists a proper nontrivial subgroup H_a of index a such that all but at most $a-2$ terms of S are from the coset $\alpha + H_a$, and w.l.o.g. by translation we may assume $\alpha = 0$; furthermore, there exists a subsequence S' of S of length at most $\lfloor \frac{m}{2} \rfloor + \frac{m}{a} - 1$ with an $\lfloor \frac{m}{2} \rfloor$-set partition $P' = P'_1, \ldots, P'_{\lfloor \frac{m}{2} \rfloor}$ satisfying $\sum_{i=1}^{\lfloor \frac{m}{2} \rfloor} P'_i = H_a$. Hence, since $\lceil \frac{m}{2} \rceil \leq m - s$, then by appending on $m - s - \lfloor \frac{m}{2} \rfloor$ singleton sets to P', each with their element from H_a, it follows that there exists a subsequence S'' of S', satisfying $|S''| \leq m - s + \frac{m}{a} - 1$, and which has an $(m-s)$-set partition P'' the sumset of whose terms is H_a (that there are enough terms from H_a to accomplish this follows from the calculation of the next sentence). Hence it follows that there are at least $2m + 2s - 1 - (a-2) - (m - s + \frac{m}{a} - 1) = m + 3s - \frac{m}{a} - a + 2 \geq \frac{m}{a} + 2s - 1 > 0$ terms of S that are not used in the set partition P'' and which are from H_a, whence the proof is complete as it was in the previous paragraph. $\square$

We remark that the arguments used in this section to obtain upper bounds for colorings with $\mathbb{Z}/m\mathbb{Z}$ work equally well for colorings with any abelian group G of order m, although in the noncyclic case the matching lower bound constructions do not hold.

We conclude by giving an example of a fairly simple hypergraph on $(\lfloor \frac{m}{2} \rfloor + 3)(\lceil \frac{m}{2} \rceil - 1)$ vertices with every edge having at least $\lceil \frac{m}{2} \rceil - 2$ monovalent vertices, but which does not zero-sum generalize, showing that the $\lceil \frac{m}{2} \rceil$ bound given in Theorems 11.1 and 11.2 can be improved at best to $\lceil \frac{m}{2} \rceil - 1$. Let X be a set of $\lfloor \frac{m}{2} \rfloor + 3$ vertices, and for each $\lfloor \frac{m}{2} \rfloor + 2$ subset X' of X, define an edge of the hypergraph $\mathcal{H}$ to be X' along with $\lceil \frac{m}{2} \rceil - 2$ monovalent vertices disjoint from X. For the coloring of the complete graph, let Δ consist entirely of an equal number of vertices colored by 0 and 1, and one vertex colored by $\lceil \frac{m}{2} \rceil$. Hence, since the only non-monochromatic m-term zero-sum sequence is

$$(\underbrace{0,\ldots,0}_{\lceil\frac{m}{2}\rceil-1},\underbrace{1,\ldots,1}_{\lfloor\frac{m}{2}\rfloor},\left\lceil\frac{m}{2}\right\rceil), \tag{11.1}$$

it follows that any zero-sum copy $\mathcal{H}'$ of $\mathcal{H}$ must have one of its edges, say e, use the coloring given by (11.1). Since $|e\cap X| = \lfloor\frac{m}{2}\rfloor+2$, then it follows from the pigeonhole principle that $e\cap X$ must contain an element x colored by 1 as well as an element y colored by 0. However, from the definition of $\mathcal{H}$ and Δ we can then find an edge of $\mathcal{H}'$ that contains both x and y but not the single element colored by $\lceil\frac{m}{2}\rceil$, which, since there can be no non-monochromatic zero-sum edge using only the colors 0 and 1, cannot be zero-sum, contradicting that $\mathcal{H}'$ is zero-sum.

Chapter 12

The Multiplicity of Zero-Sums in Sequences of Small Length

12.1 Discussion

We know from EGZ that any sequence of $2m-1$ terms from an abelian group of order m must contain an m-term zero-sum subsequence. For sequences whose length is greater than $2m-1$, a natural question to ask is *how many* m-term zero-sum subsequences can one expect. If the sequence S has length n and consists of at most two distinct terms, then there will be at least $\binom{\lceil \frac{n}{2} \rceil}{m} + \binom{\lfloor \frac{n}{2} \rfloor}{m}$ m-term monochromatic subsequences. Thus if the best way to avoid m-term zero-sum subsequences were still to use only two distinct residues from $\mathbb{Z}/m\mathbb{Z}$, then one would expect there to always be at least $\binom{\lceil \frac{n}{2} \rceil}{m} + \binom{\lfloor \frac{n}{2} \rfloor}{m}$ m-term zero-sum subsequences. This was conjectured by Bialostocki in 1989 [3] and later appeared in [4].

Conjecture 12.1. *If S is a sequence of n terms from $\mathbb{Z}/m\mathbb{Z}$, then S has at least $\binom{\lceil \frac{n}{2} \rceil}{m}+\binom{\lfloor \frac{n}{2} \rfloor}{m}$ m-term zero-sum subsequences.*

A few years after the conjecture was made, Kisin verified Conjecture 12.1 in the case $m = p^{\alpha}$ and $m = p^{\alpha}q$, where p and q are primes and $\alpha \geq 1$, and expressed reasons why the conjecture might fail for m not of this form [42]. At the same time, Füredi and Kleitman showed that Conjecture 12.1 held for sufficiently large n (of order m^{6m}), as well as for m of the form $m = pq$, where p and q are distinct primes, and showed that $2\binom{\lfloor \frac{n}{2} \rfloor}{m} - m^2\binom{\lfloor \frac{n}{2} \rfloor - 1}{m-1}$ was a general lower bound on the number of m-term zero-sum subsequences [16]. Their results, contrary to those of Kisin, led them to strongly believe the conjecture of Bialostocki to be true for $n > 4m$. Unfortunately, the lower bound shown by Füredi and Kleitman, while being very nice asymptotically for large n and fixed m, tells us very little for small n, particularly if m is also large.

The aim of this Chapter is to give a proof, using the machinery of Chapters 3 and 6, of the following general bound on the number of m-term zero-sum subsequences.

Theorem 12.2. *If S is a sequence of n terms from an abelian group G of order $m \geq 30$, then S contains at least* $\min\left\{\binom{\lceil \frac{n}{2} \rceil}{m} + \binom{\lfloor \frac{n}{2} \rfloor}{m}, \binom{n-m}{\lceil \frac{2m-1}{3} \rceil}\right\}$ *m-term zero-sum subsequences.*

Unlike the general bound of Füredi and Kleitman, the bound given by Theorem 12.2 is much more accurate for sequences of small length, and, as will be shown in section 12.2, verifies Conjecture 12.1 for $n \leq 6\frac{1}{3}m$. Ironically, this confirms the conjecture of Bialostocki for those cases least thought to be true. Theorem 12.2 also gives a bound for more general abelian groups in addition to cyclic groups.

12.2 The Multiplicity of Zero-Sums

In view of the results of Kisin [42] mentioned in the discussion, it follows that Conjecture 12.1 is known for $m < 30$, as well as for $m = 2^5 = 32$, $m = 5 \cdot 7 = 35$ and $m = 2 \cdot 19 = 38$. We begin by proving several lemmas relating the sizes of two different binomial coefficients. In view of the first sentence of this section, note that Lemma 20 and Theorem 12.2 together imply Conjecture 12.1 for $n \leq 6\frac{1}{3}m$. Both Lemmas 19 and 20 are straightforward computations, best done with machine assistance, but for the benefit of the reader we include many of the details. Their proofs will make use of the following well-known and basic proposition bounding the real roots of a polynomial with real coefficients.

Proposition 12.3. *Let $P(x)$ be a polynomial with real coefficients and positive leading coefficient, and let a be a real number. If $a > 0$, and all nonzero terms of $P(x)/(x-a)$, including remainder (computed by polynomial division), are positive, then a is an upper bound for all real roots of $P(x)$.*

Proof. Let $P(x) = Q(x)(x-a) + r$, with $r \in \mathbb{R}$. Since all nonzero terms of $P(x)/(x-a)$, including remainder (computed by polynomial division), are positive, it follows that $r \geq 0$ and $Q(x) > 0$ for all real $x > 0$. Thus, since for $x > a > 0$ we have $x - a > 0$, it follows that $P(x) = Q(x)(x-a) + r > 0$ for $x > a$. $\square$

Lemma 19. *If $m \geq 30$ and n are integers with $2m-1 \leq n \leq 3m + \lceil \frac{2m-1}{3} \rceil - 2$, then $\binom{n-m}{\lceil \frac{m}{2} \rceil} > 2\binom{\lceil \frac{n}{2} \rceil}{m}$.*

Proof. Let $R(n,m) = \binom{n-m}{\lceil \frac{m}{2} \rceil}/2\binom{\frac{n+1}{2}}{m} = \frac{(n-m)\ldots(n-m-\lceil \frac{m}{2} \rceil+1)(m)\ldots(\lceil \frac{m}{2} \rceil+1)}{2(\frac{n+1}{2})\ldots(\frac{n+1}{2}-m+1)}$. Since $\binom{\frac{n+1}{2}}{m} \geq \binom{\lceil \frac{n}{2} \rceil}{m}$, then it suffices to show $R(n,m) > 1$. We begin by showing that $R(n,m) \geq R(n+2,m)$.

Let $Q(n,m) = \frac{(n-m-\frac{m+1}{2}+2)(n-m-\frac{m+1}{2}+1)(\frac{n+1}{2}+1)}{(n-m+2)(n-m+1)(\frac{n+1}{2}-m+1)} \leq R(n,m)/R(n+2,m)$. To show $R(n,m) \geq R(n+2,m)$, we will show that $Q(n,m) \geq 1$, i.e., (by multiplying out the denominator, and expanding and collecting terms) that

$$4(m-1)n^2 - (11m^2 - 12m + 17)n + (8m^3 - 9m^2 + 16m - 15) \geq 0.$$

This will occur if both roots of the above polynomial are imaginary, which by the quadratic formula occurs when

$$m^4 - \frac{8}{7}m^3 - \frac{118}{7}m^2 - \frac{88}{7}m - 7 > 0. \tag{12.1}$$

However, in view of Proposition 12.3 it follows that the roots of the polynomial $m^4 - \frac{8}{7}m^3 - \frac{118}{7}m^2 - \frac{88}{7}m - 7$ are bounded from above by 6. Consequently (12.1) holds for $m \geq 7$, and we can assume $R(n, m) \geq R(n + 2, m)$.

Since $R(n, m) \geq R(n + 2, m)$, it suffices to show $R(3\frac{2}{3}m + b, m) > 1$ for $b = -2 + (\lceil \frac{2m-1}{3} \rceil - \frac{2}{3}m)$ and $b = -3 + (\lceil \frac{2m-1}{3} \rceil - \frac{2}{3}m)$. Note $b \in \{-\frac{5}{3}, -\frac{6}{3}, -\frac{7}{3}, -\frac{8}{3}, -\frac{9}{3}, -\frac{10}{3}\}$. Let $S(m) = R(3\frac{2}{3}m + b, m)$. Next we show that $S(m + 6) \geq S(m)$. Note that computing $S(m)$ for each $m \in \{30, \dots, 35\}$ and both possible values for b shows that $S(m) > 1$ for $m \in \{30, \dots, 35\}$. Hence the proof will be complete once we have shown that $S(m + 6) \geq S(m)$.

Let $P(m) = \frac{(\frac{8}{3}m+b+16)...(\frac{8}{3}m+b+1)(m+6)...(m+1)(\frac{5}{6}m+\frac{b+1}{2}+5)...(\frac{5}{6}m+\frac{b+1}{2}+1)}{(\frac{11}{6}m+\frac{b+1}{2}+11)...(\frac{11}{6}m+\frac{b+1}{2}+1)(\frac{m+1}{2}+3)...(\frac{m+1}{2}+1)(\frac{13}{6}m+b+13)...(\frac{13}{6}m+b+1)} \leq S(m + 6)/S(m)$. To see that $S(m + 6) \geq S(m)$, we will show that $P(m) \geq 1$. By multiplying out denominators, bringing all terms to the left hand side, expanding and collecting terms, and rounding coefficients down, it follows that it suffices to show $-3 \cdot 10^{17} - 4 \cdot 10^{18} m - 3 \cdot 10^{19} m^2 - 2 \cdot 10^{20} m^3 - 4 \cdot 10^{20} m^4 - 7 \cdot 10^{20} m^5 - 2 \cdot 10^{21} m^6 - 2 \cdot 10^{21} m^7 - 2 \cdot 10^{21} m^8 - 2 \cdot 10^{21} m^9 - 8 \cdot 10^{20} m^{10} - 5 \cdot 10^{20} m^{11} - 2 \cdot 10^{20} m^{12} - 8 \cdot 10^{19} m^{13} - 3 \cdot 10^{19} m^{14} - 7 \cdot 10^{18} m^{15} - 2 \cdot 10^{18} m^{16} - 4 \cdot 10^{17} m^{17} - 6 \cdot 10^{16} m^{18} - 7 \cdot 10^{15} m^{19} - 7 \cdot 10^{14} m^{20} - 5 \cdot 10^{13} m^{21} - 2 \cdot 10^{12} m^{22} - 4 \cdot 10^{10} m^{23} + 7 \cdot 10^8 m^{24} + 2 \cdot 10^8 m^{25} + 10^7 m^{26} + 3 \cdot 10^5 m^{27} > 0$, in order to show $P(m) \geq 1$ (the rounded polynomial just given is strictly less, for positive m, than the corresponding polynomial for each value of $b \in \{-\frac{5}{3}, -\frac{6}{3}, -\frac{7}{3}, -\frac{8}{3}, -\frac{9}{3}, -\frac{10}{3}\}$ obtained by algebraic manipulation). However, in view of Proposition 12.3, it follows that the roots of the polynomial from the previous sentence are all bounded from above by 23, implying that the inequality from the last sentence holds for $m \geq 24$, which completes the proof. $\square$

Lemma 20. *If $m \geq 30$ and n are integers either with $2m-1 \leq n \leq 6\frac{1}{3}m$, $m \neq 32$, $m \neq 35$, and $m \neq 38$, or else with $2m-1 \leq n \leq 6\frac{1}{3}m - 6$, then $\binom{n-m}{\lceil \frac{2m-1}{3} \rceil} > 2\binom{\lceil \frac{n}{2} \rceil}{m}$.*

Proof. Let $R(n,m) = \binom{n-m}{\lceil \frac{2m-1}{3} \rceil}/2\binom{\frac{n+1}{2}}{m} = \frac{(n-m)...(n-m-\lceil \frac{2m-1}{3} \rceil+1)(m)...(\lceil \frac{2m-1}{3} \rceil+1)}{2(\frac{n+1}{2})...(\frac{n+1}{2}-m+1)}$. Since we have $\binom{\frac{n+1}{2}}{m} \geq \binom{\lceil \frac{n}{2} \rceil}{m}$, then it suffices to show $R(n,m) > 1$. We begin by showing that $R(n,m) \geq R(n+2,m)$.

Let $Q(n,m) = \frac{(n-m-\frac{2m+1}{3}+2)(n-m-\frac{2m+1}{3}+1)(\frac{n+1}{2}+1)}{(n-m+2)(n-m+1)(\frac{n+1}{2}-m+1)} \leq R(n,m)/R(n+2,m)$. To show $R(n,m) \geq R(n+2,m)$, we will show that $Q(n,m) \geq 1$, i.e., (by multiplying out the denominator, and expanding and collecting terms) that

$$3(m-1)n^2 - (10m^2 - 5m + 13)n + (9m^3 - 3m^2 + 6m - 12) \geq 0.$$

This will occur if both roots of the above polynomial are imaginary, which by the quadratic formula occurs when

$$m^4 - \frac{11}{2}m^3 - \frac{177}{8}m^2 - \frac{43}{4}m - \frac{25}{8} > 0. \tag{12.2}$$

However, in view of Proposition 12.3 it follows that the roots of the polynomial $m^4 - \frac{11}{2}m^3 - \frac{177}{8}m^2 - \frac{43}{4}m - \frac{25}{8}$ are bounded from above by 9. Consequently (12.2) holds for $m \geq 10$, and we can assume $R(n,m) \geq R(n+2,m)$.

First assume that $n \leq 6\frac{1}{3}m$ with $m \neq 32$, $m \neq 35$, and $m \neq 38$. Since $R(n,m) \geq R(n+2,m)$, it suffices to show $R(6\frac{1}{3}m + b, m) > 1$ for $b = (\lfloor 6\frac{1}{3}m \rfloor - 6\frac{1}{3}m)$ and $b = -1 + (\lfloor 6\frac{1}{3}m \rfloor - 6\frac{1}{3}m)$. Note $b \in \{0, -\frac{1}{3}, -\frac{2}{3}, -\frac{3}{3}, -\frac{4}{3}, -\frac{5}{3}\}$. Let $S(m) = R(6\frac{1}{3}m + b, m)$. Next we show that $S(m+6) \geq S(m)$ for $m \geq 43$. Note that computing $S(m)$ for each $m \leq 48$, $m \neq 32$, $m \neq 35$, $m \neq 38$, and both possible values for b shows that $S(m) > 1$ for $m \leq 48$, $m \neq 32$, $m \neq 35$, $m \neq 38$. Hence the first part of the lemma will be complete once we have shown that $S(m+6) \geq S(m)$ for $m \geq 43$.

Define $P(m) = \frac{(\frac{16}{3}m+b+32)\ldots(\frac{16}{3}m+b+1)(m+6)\ldots(m+1)(\frac{13}{6}m+\frac{b+1}{2}+13)\ldots(\frac{13}{6}m+\frac{b+1}{2}+1)}{(\frac{19}{6}m+\frac{b+1}{2}+19)\ldots(\frac{19}{6}m+\frac{b+1}{2}+1)(\frac{2m+1}{3}+4)\ldots(\frac{2m+1}{3}+1)(\frac{14}{3}m+b+\frac{1}{3}+28)\ldots(\frac{14}{3}m+b+\frac{1}{3}+1)}$, and note that $P(m) \leq S(m+6)/S(m)$ for $m \geq 43$. To see that $S(m+6) \geq S(m)$, it suffices to show $P(m) \geq 1$. The proof now proceeds as in the previous lemma. The case with $n \leq 6\frac{1}{3}m - 6$ can be handled similarly. $\square$

Lemma 21. *Let n, m and x be positive integers. If $n \geq \frac{3}{2}m - 1$, then $3^x\binom{n}{m} \geq \binom{n+x}{m}$.*

Proof. Observe that the following binomial identity holds:

$$\binom{n}{m} = \frac{n-m+1}{m}\binom{n}{m-1}. \tag{12.3}$$

Since $n \geq \frac{3}{2}m - 1$, then (12.3) implies that $2\binom{n+x'}{m} \geq \binom{n+x'}{m-1}$, for $x' \geq 0$. Hence from the Pascal Identity, it follows that

$$3\binom{n+x'}{m} \geq \binom{n+x'}{m} + \binom{n+x'}{m-1} = \binom{n+x'+1}{m},$$

for $x' \geq 0$. Iterating the above inequality for $x' = 0, \ldots, x-1$ yields $3^x\binom{n}{m} \geq \binom{n+x}{m}$. $\square$

We now proceed with the proof of Theorem 12.2.

Proof. The proof will be divided into several steps. For our main method to work, we will need the existence of a sufficiently compressed $\lceil \frac{n}{2} \rceil$-set partition. Thus we will first handle several special and highly restrictive sequences S that do not admit such a compressed set partition.

Let $\mathcal{Z}_m(S)$ denote the number of m-term zero-sum subsequences of S. Note that from the Erdős-Ginzburg-Ziv Theorem it follows trivially that $\mathcal{Z}_m(S) \geq n - 2m + 2$. Thus $\mathcal{Z}_m(S) \geq \binom{\lceil \frac{n}{2} \rceil}{m} + \binom{\lfloor \frac{n}{2} \rfloor}{m}$ holds for $n \leq 2m$. Consequently, inductively assume $\mathcal{Z}_m(S') \geq$

$\min\{\binom{\lceil \frac{n}{2}\rceil}{m} + \binom{\lfloor \frac{n}{2}\rfloor}{m}, \binom{n-m}{\lceil \frac{2m-1}{3}\rceil}\}$ holds for any sequence S' of n' terms from an abelian group of order m provided $n' < n$, and also assume that $n \geq 2m+1$. In view of the mentioned results of Kisin [42], we may assume that m is composite.

Step 1 (S essentially monochromatic): Suppose that there is a term x of S with multiplicity at least $\lceil \frac{n}{2}\rceil$. Then there will be at least $\binom{\lceil \frac{n}{2}\rceil-1}{m-1}$ m-term monochromatic (and hence also zero-sum) subsequences of S that include the term x. By induction hypothesis there are at least $\min\{\binom{\lceil \frac{n-1}{2}\rceil}{m}+\binom{\lfloor \frac{n-1}{2}\rfloor}{m}, \binom{n-m-1}{\lceil \frac{2m-1}{3}\rceil}\}$ m-term zero-sum subsequences that do not include the term x. Hence there are in total at least $\min\{\binom{\lceil \frac{n}{2}\rceil-1}{m-1} + \binom{\lceil \frac{n-1}{2}\rceil}{m} + \binom{\lfloor \frac{n-1}{2}\rfloor}{m}, \binom{\lceil \frac{n}{2}\rceil-1}{m-1} + \binom{n-m-1}{\lceil \frac{2m-1}{3}\rceil}\}$ m-term zero-sum subsequences. By the Pascal Identity for binomial coefficients, it follows that

$$\binom{\lceil \frac{n}{2}\rceil - 1}{m-1} + \binom{\lceil \frac{n-1}{2}\rceil}{m} + \binom{\lfloor \frac{n-1}{2}\rfloor}{m} = \binom{\lceil \frac{n}{2}\rceil - 1}{m-1} + \binom{\lceil \frac{n}{2}\rceil - 1}{m} + \binom{\lfloor \frac{n}{2}\rfloor}{m} = \binom{\lceil \frac{n}{2}\rceil}{m} + \binom{\lfloor \frac{n}{2}\rfloor}{m}.$$

Thus the proof is complete unless

$$\binom{n-m-1}{\lceil \frac{2m-1}{3}\rceil} < \binom{\lceil \frac{n-1}{2}\rceil}{m} + \binom{\lfloor \frac{n-1}{2}\rfloor}{m}, \tag{12.4}$$

and

$$\binom{\lceil \frac{n}{2}\rceil - 1}{m-1} + \binom{n-m-1}{\lceil \frac{2m-1}{3}\rceil} < \binom{n-m}{\lceil \frac{2m-1}{3}\rceil}.$$

From the above inequality and the Pascal Identity, it follows that

$$\binom{\lceil \frac{n}{2}\rceil - 1}{m-1} < \binom{n-m-1}{\lceil \frac{2m-1}{3}\rceil - 1}. \tag{12.5}$$

From (12.4) and Lemma 20, it follows that $n-1 > 6\frac{1}{3}m-6$. Applying the binomial identity given in (12.3) to (12.5), it follows that

$$\binom{\lceil \frac{n}{2} \rceil - 1}{m-1} < \frac{\lceil \frac{2m-1}{3} \rceil}{(n-m-\lceil \frac{2m-1}{3} \rceil)} \binom{n-m-1}{\lceil \frac{2m-1}{3} \rceil}, \tag{12.6}$$

and that

$$\binom{\lceil \frac{n}{2} \rceil - 1}{m} < \frac{(\lceil \frac{n}{2} \rceil - m)}{m} \cdot \frac{\lceil \frac{2m-1}{3} \rceil}{(n-m-\lceil \frac{2m-1}{3} \rceil)} \binom{n-m-1}{\lceil \frac{2m-1}{3} \rceil}. \tag{12.7}$$

If n is odd, then (12.4) implies $\binom{n-m-1}{\lceil \frac{2m-1}{3} \rceil} < \binom{\lceil \frac{n}{2} \rceil - 1}{m} + \binom{\lceil \frac{n}{2} \rceil - 1}{m}$, and if n is even, then (12.4) and the Pascal Identity imply $\binom{n-m-1}{\lceil \frac{2m-1}{3} \rceil} < \binom{\lceil \frac{n}{2} \rceil}{m} + \binom{\lceil \frac{n}{2} \rceil - 1}{m} = \binom{\lceil \frac{n}{2} \rceil - 1}{m} + \binom{\lceil \frac{n}{2} \rceil - 1}{m-1} + \binom{\lceil \frac{n}{2} \rceil - 1}{m}$. Hence from (12.6) and (12.7), it follows that

$$\binom{n-m-1}{\lceil \frac{2m-1}{3} \rceil} < \left(2 \cdot \frac{(\lceil \frac{n}{2} \rceil - m)}{m} \cdot \frac{\lceil \frac{2m-1}{3} \rceil}{(n-m-\lceil \frac{2m-1}{3} \rceil)} + \frac{\lceil \frac{2m-1}{3} \rceil}{(n-m-\lceil \frac{2m-1}{3} \rceil)}\right) \binom{n-m-1}{\lceil \frac{2m-1}{3} \rceil},$$

which in turn implies that

$$1 < \frac{2(\frac{n+1}{2} - m) \cdot \frac{2m+1}{3} + m \cdot \frac{2m+1}{3}}{m \cdot (n - m - \frac{2m+1}{3})}.$$

From the above inequality, it follows that $(m-1)n < 3m^2 + 2m + 1$, implying $n < 3m + 5 + \frac{6}{m-1}$, which contradicts that $n - 1 > 6\frac{1}{3}m - 6$ and $m \geq 30$. So we may assume that the multiplicity of every term x of S is at most $\lceil \frac{n}{2} \rceil - 1$.

Step 2 (S essentially dichromatic): Suppose that every term of S, with at most $\max\{m - \frac{m}{p}, \lfloor \frac{2m-4}{3} \rfloor\}$ exceptions if $n \geq 3m + \lceil \frac{2m-1}{3} \rceil - 1$, and with at most $m - \frac{m}{p}$ exceptions if $n \leq 3m + \lceil \frac{2m-1}{3} \rceil - 2$, is equal to one of two elements $x, y \in G$, where p is the smallest prime divisor of m. Let n_x and n_y denote the respective multiplicities of x and y in S. Rearrange the terms of S so that all the terms equal to x precede all the terms equal to y, which in turn precede all terms equal to neither x nor y, and let $x_1, \ldots, x_n$ be the resulting sequence. For $i \in \{1, \ldots, \lfloor \frac{n}{2} \rfloor\}$, let $A_i = \{x_i, x_{i+\lceil \frac{n}{2} \rceil}\}$, and if n is odd, then let $A_{\lceil \frac{n}{2} \rceil} = \{x_{\lceil \frac{n}{2} \rceil}\}$. Then in view of Step 1, it follows that $A = A_1, \ldots, A_{\lceil \frac{n}{2} \rceil}$ is an $\lceil \frac{n}{2} \rceil$-set partition of S such that either $x \in A_i$ or $y \in A_i$ holds for every set A_i.

There are $\binom{\lfloor \frac{n}{2} \rfloor}{m}$ ways to choose m sets A_i from A all with $|A_i| = 2$, and (in case n odd) there are $\binom{\lceil \frac{n}{2} \rceil - 1}{m-1}$ ways to choose m sets A_i from A that include the set $A_{\lceil \frac{n}{2} \rceil}$ of cardinality one. Consequently, if we can show that any such selection $A_{i_1}, \ldots, A_{i_m}$ has a set A_{i_k} such that $0 \in z + \sum\limits_{\substack{j=1 \\ j \neq k}}^{m} A_{i_j}$ for every $z \in A_{i_k}$ (in which case we will say that the selection $A_{i_1}, \ldots, A_{i_m}$ is *good*), then there will be (in case n even) at least $2\binom{\lfloor \frac{n}{2} \rfloor}{m} = \binom{\lfloor \frac{n}{2} \rfloor}{m} + \binom{\lceil \frac{n}{2} \rceil}{m}$ m-term zero-sum subsequences, and (in case n odd), in view of the Pascal Identity, at least $2\binom{\lfloor \frac{n}{2} \rfloor}{m} + \binom{\lceil \frac{n}{2} \rceil - 1}{m-1} = \binom{\lfloor \frac{n}{2} \rfloor}{m} + \binom{\lceil \frac{n}{2} \rceil - 1}{m} + \binom{\lceil \frac{n}{2} \rceil - 1}{m-1} = \binom{\lfloor \frac{n}{2} \rfloor}{m} + \binom{\lceil \frac{n}{2} \rceil}{m}$ m-term zero-sum subsequences, whence the proof is complete. We proceed to show this is the case, except for a highly restrictive sequence that we handle separately afterwards.

If the selection $A_{i_1}, \ldots, A_{i_m}$ contains the set $A_{\lceil \frac{n}{2} \rceil}$ and n is odd, then let $A_{i_k} = A_{\lceil \frac{n}{2} \rceil}$, and otherwise let A_{i_k} be a set $A_{i_j} = \{x, y\}$ (such a set exists, since at most $\max\{m - \frac{m}{p}, \lfloor \frac{2m-4}{3} \rfloor\} < m$ terms of S are equal to neither x nor y). If $|\sum\limits_{\substack{j=1 \\ j \neq k}}^{m} A_{i_j}| \geq \sum\limits_{\substack{j=1 \\ j \neq k}}^{m} |A_j| - (m-1) + 1 = m$, then for each $z \in A_{i_k}$ we can select a term from each of the A_{i_j}, $j \neq k$, so that the sum of the $m-1$ selected terms from the A_{i_j}, $j \neq k$, is the additive inverse of z, whence we see that the selection $A_{i_1}, \ldots, A_{i_m}$ is good. Otherwise, from Kneser's Theorem it follows that $\sum\limits_{\substack{j=1 \\ j \neq k}}^{m} A_{i_j}$ is maximally H_a-periodic, with H_a of index a and $1 < a < m$.

Suppose that $\phi_a(x) = \phi_a(y)$, i.e., that x and y are from the same H_a-coset. Hence, since every set A_{i_j} contains either x or y, it follows that every set A_{i_j} contains a representative from the coset $x + H_a$. Hence, since $\sum\limits_{\substack{j=1 \\ j \neq k}}^{m} A_{i_j}$ is H_a-periodic, it follows that $0 \in H_a =$

$mx+H_a\subset z+\sum_{\substack{j=1\\j\neq k}}^{m}A_{i_j}$ for $z\in A_{i_k}\subseteq\{x,y\}$, and the proof is again complete. So we may assume that $\phi_a(x)\neq\phi_a(y)$.

If there are at most $m-\frac{m}{p}$ terms of S equal to neither x nor y, then there must be at least $a-1$ sets A_{i_j}, $j\neq k$, with $A_{i_j}=\{x,y\}$, and hence, since $\phi_a(x)\neq\phi_a(y)$, at least $a-1$ sets A_{i_j} with $|\phi_a(A_{i_j})|=2$. On the other hand, if there are at most $\lfloor\frac{2m-4}{3}\rfloor$ terms of S equal to neither x nor y, then either there likewise must be at least $a-1$ sets A_{i_j} with $|\phi_a(A_{i_j})|=2$, or else $|H_a|=2$, and there are at least $\frac{m}{2}+2$ sets A_{i_j} with $A_{i_j}\neq\{x,y\}$ and A_{i_j} contained in an H_a-coset. If the former holds, then from Kneser's Theorem it follows that $|\sum_{\substack{j=1\\j\neq k}}^{m}A_j|\geq|H_a|(\sum_{\substack{j=1\\j\neq k}}^{m}|\phi_a(A_j)|-(m-1)+1)\geq m$, and the proof is again complete. Therefore we may instead assume the latter. Consequently we can assume that $n\geq 3m+\lceil\frac{2m-1}{3}\rceil-1$, that m is even, and that there are at least $m-\frac{m}{p}+1=\frac{m}{2}+1$ terms t of S with $t\notin\{x,y\}$.

Suppose that $x-y$ generates a proper subgroup H_b of index b (this is the one case we do not handle by showing an arbitrary selection $A_{i_1},\ldots,A_{i_m}$ is good). Since there are at most $\lfloor\frac{2m-4}{3}\rfloor$ terms of S equal to neither x nor y, and since there are at least $\frac{m}{2}+1$ sets A_i with $A_i\neq\{x,y\}$ and A_i an H_a-coset, then we can re-index the sets A_i so that $A_i=\{x,y\}$ for $i\leq\lfloor\frac{n}{2}\rfloor-\lfloor\frac{2m-4}{3}\rfloor$, and so that A_i is an H_a-coset for $\lfloor\frac{n}{2}\rfloor-\lfloor\frac{2m-4}{3}\rfloor+1\leq i\leq\lfloor\frac{n}{2}\rfloor-\lfloor\frac{2m-4}{3}\rfloor+\frac{m}{2}+1$. Let $A_{i'_1},\ldots,A_{i'_m}$ be a selection of m sets A_i all with $i\leq\lfloor\frac{n}{2}\rfloor-\lfloor\frac{2m-4}{3}\rfloor+\frac{m}{2}+1=\lfloor\frac{n}{2}\rfloor-\lfloor\frac{m-8}{6}\rfloor+1$.

If $A_{i'_j}=\{x,y\}$ for all j, then $\sum_{j=1}^{\frac{m}{b}-1}A_{i'_j}$ is an H_b-coset, whence there will be at least $2^{m-\frac{m}{b}+1}\geq 2^{\frac{m}{2}}$ ways to select a term from each set $A_{i'_j}$ and get an m-term zero-sum subsequence. Next suppose that at least one of the $A_{i'_j}$, say w.l.o.g. $A_{i'_1}$, is an H_a-coset. Since at most $\frac{m}{2}+1$ of the sets $A_{i'_j}$ can be H_a-cosets, it follows that there are at least $\frac{m}{2}-1$ indices j with $A_{i'_j}=\{x,y\}$. Re-index so that $A_{i'_j}=\{x,y\}$ for $2\leq i\leq\frac{m}{2}$. Hence $\sum_{j=1}^{\frac{m}{2}}A_{i'_j}$ is an (H_a+H_b)-coset. Thus, since every set $A_{i'_j}$ contains either x or y, then it follows that every set $A_{i'_j}$ is contained in the same (H_a+H_b)-coset $x+H_a+H_b$, whence it follows that there will also be at least $2^{\frac{m}{2}}$ ways to a select a term from each set $A_{i'_j}$ and get an m-term zero-sum subsequence. Thus we conclude that there are at least

$$2^{\frac{m}{2}}\binom{\lfloor\frac{n}{2}\rfloor-\lfloor\frac{m-8}{6}\rfloor+1}{m} \tag{12.8}$$

m-term zero-sum subsequences. Since $n\geq 3m+\lceil\frac{2m-1}{3}\rceil-1$, then it follows in view of Lemma 21 that $3^x\binom{\lfloor\frac{n}{2}\rfloor-\lfloor\frac{m-8}{6}\rfloor+1}{m}\geq\binom{\lfloor\frac{n}{2}\rfloor-\lfloor\frac{m-8}{6}\rfloor+1+x}{m}$. Hence from (12.8) it follows that there are at least

$$2^{\frac{m}{2}}\binom{\lfloor\frac{n}{2}\rfloor-\lfloor\frac{m-8}{6}\rfloor+1}{m}\geq 2\cdot 4^{\lfloor\frac{m-2}{4}\rfloor}\binom{\lfloor\frac{n}{2}\rfloor-\lfloor\frac{m-8}{6}\rfloor+1}{m}\geq$$
$$2\cdot 3^{\lfloor\frac{m-2}{4}\rfloor}\binom{\lfloor\frac{n}{2}\rfloor-\lfloor\frac{m-8}{6}\rfloor+1}{m}\geq 2\binom{\lfloor\frac{n}{2}\rfloor-\lfloor\frac{m-8}{6}\rfloor+\lfloor\frac{m-2}{4}\rfloor+1}{m}\geq$$
$$2\binom{\lceil\frac{n}{2}\rceil}{m}\geq\binom{\lceil\frac{n}{2}\rceil}{m}+\binom{\lfloor\frac{n}{2}\rfloor}{m}$$

m-term zero-sum subsequences, whence the proof is complete. So we may assume that $x-y$ generates G, implying G is cyclic of order m.

Suppose $n_x\leq\lfloor\frac{n}{2}\rfloor-\frac{m}{2}$. Re-index the terms x_i in the sequence $x_1,\ldots,x_n$ with $x_i\notin\{x,y\}$ (leaving unchanged the terms $x_i\in\{x,y\}$) so that all terms x_i with $x_i\notin\{x,y,y+\frac{m}{2}\}$ occur in a consecutive block at the very end of the sequence. Then, since in a cyclic group there is a unique subgroup of order two, it follows that either every set A_i will contain a representative from the common H_a-coset $y+H_a$, or else every set A_i contained in an $H_{a'}$-coset with $|H_{a'}|=2$ and $i\leq\lfloor\frac{n}{2}\rfloor$ must contain x. In the latter case, since $n_x\leq\lfloor\frac{n}{2}\rfloor-\frac{m}{2}$, it follows that there are at most $\lfloor\frac{2m-4}{3}\rfloor-\frac{m}{2}+1<\frac{m}{2}+2$ sets A_i contained in an $H_{a'}$-coset with $|H_{a'}| = 2$, which shows the selection is good by a case handled in the fifth paragraph of Step 2. Therefore we may assume the former case holds. From previous work, we know that any selection $A_{i_1}, \ldots, A_{i_m}$ is good unless $\sum_{\substack{j=1\\j\neq k}}^{m} A_{i_j}$ is maximally $H_{a'}$-periodic with $|H_{a'}| = 2$ and $|\sum_{\substack{j=1\\j\neq k}}^{m} A_{i_j}| < \sum_{\substack{j=1\\j\neq k}}^{m} |A_{i_j}| - (m-1) + 1$. However, since there is a unique subgroup H_a of order two, it follows that $H_{a'} = H_a$. Hence, since every set A_i contains a representative from the common H_a-coset $y + H_a$, and since $\sum_{j=1}^{m} A_{i_j}$ is H_a-periodic, it follows that $0 \in \sum_{j=1}^{m} A_{i_j}$. Since $|\sum_{\substack{j=1\\j\neq k}}^{m} A_{i_j}| < \sum_{\substack{j=1\\j\neq k}}^{m} |A_{i_j}| - (m-1) + 1$, and since $|A_{i_j}| = 2$ for $j \neq i_k$, it follows from Proposition 2.4 that there exists A_{i_l} with $l \neq k$ such that $|\sum_{\substack{j=1\\j\neq l}}^{m} A_{i_j}| = |\sum_{j=1}^{m} A_{i_j}|$, whence it follows that every $z \in \sum_{j=1}^{m} A_{i_j}$ can be represented in at least two different ways, including $0 \in \sum_{j=1}^{m} A_{i_j}$. Thus every selection $A_{i_1}, \ldots, A_{i_m}$ is good, completing the proof. So we may assume that $n_x \geq \lfloor \frac{n}{2} \rfloor - \frac{m}{2} + 1$.

Re-index the terms x_i in the sequence $x_1, \ldots, x_n$ with $x_i \notin \{x, y\}$ (leaving unchanged the terms $x_i \in \{x, y\}$) so that all terms x_i with $x_i = x + \frac{m}{2}$ occur in a consecutive block at the very end of the sequence. Since $n_x \geq \lfloor \frac{n}{2} \rfloor - \frac{m}{2} + 1$, and since there are at least $m - \frac{m}{p} + 1 = \frac{m}{2} + 1$ terms t with $t \notin \{x, y\}$, it follows that $A_{n_x} = \{x, t\}$ with $t \notin \{x, y\}$. If n is odd, then modify the definition of the set partition $A_1, \ldots, A_{\lceil \frac{n}{2} \rceil}$ by swapping the term equal to x in A_{n_x} with the term equal to y in $A_{\lceil \frac{n}{2} \rceil}$. The proof now proceeds as in the above paragraph with the roles of x and y interchanged, completing Step 2. So we may assume that given any two elements x, $y \in G$, there are at least $m - \frac{m}{p} + 1$ terms of S equal to neither x nor y, and, if $n \geq 3m + \lceil \frac{2m-1}{3} \rceil - 1$, then there are at least $\lfloor \frac{2m-1}{3} \rfloor$ terms of S equal to neither x nor y.

Step 3 ($|S| \leq 3m + \lceil \frac{2m-1}{3} \rceil - 2$): Suppose that $n \leq 3m + \lceil \frac{2m-1}{3} \rceil - 2$. In view of

Steps 1 and 2 and Proposition 2.3 applied with $n' = n - m + 1$ and $n = \lfloor \frac{m}{2} \rfloor$, it follows that there exists an $(n-m+1)$-set partition $P = P_1, \ldots, P_{n-m+1}$ of S with $|P_i| = 1$ for $i > \lfloor \frac{m}{2} \rfloor$. Let $P' = P_1, \ldots, P_{\lfloor \frac{m}{2} \rfloor}$, and let S' be the subsequence partitioned by the $\lfloor \frac{m}{2} \rfloor$-set partition P'. Apply Theorem 3.1 to the subsequence S' of S with $\lfloor \frac{m}{2} \rfloor$-set partition P', and let $A = A_1, \ldots, A_{\lfloor \frac{m}{2} \rfloor}$ be the resulting set partition and H_a the corresponding subgroup of index a.

Suppose that $|\sum_{i=1}^{\lfloor \frac{m}{2} \rfloor} A_i| \geq m = \sum_{i=1}^{\lfloor \frac{m}{2} \rfloor} |A_i| - \lfloor \frac{m}{2} \rfloor + 1$. Then applying Theorem 6.1 to A and S' yields a subsequence S'' of S' of length m with an $\lfloor \frac{m}{2} \rfloor$-set partition $A' = A'_1, \ldots, A'_{\lfloor \frac{m}{2} \rfloor}$ satisfying $|\sum_{i=1}^{\lfloor \frac{m}{2} \rfloor} A'_i| \geq m$. Then given any $\lceil \frac{m}{2} \rceil$-term subsequence T of $S \setminus S''$, we can find a selection of $\lfloor \frac{m}{2} \rfloor$ terms from the $A'_1, \ldots, A'_{\lfloor \frac{m}{2} \rfloor}$ that sum to the additive inverse of the sum of the terms from T. Consequently, there will be at least $\binom{n-m}{\lceil \frac{m}{2} \rceil}$ m-term zero-sum subsequences. Thus, since $n \leq 3m + \lceil \frac{2m-1}{3} \rceil - 2$, it follows in view of Lemma 19 that the proof is complete. So we may assume that $|\sum_{i=1}^{\lfloor \frac{m}{2} \rfloor} A_i| < m = \sum_{i=1}^{\lfloor \frac{m}{2} \rfloor} |A_i| - \lfloor \frac{m}{2} \rfloor + 1$.

Thus from Theorem 3.1 it follows that $N(A', H_a) = 1$ and $E(A', H_a) \leq a - 2$, with H_a a nontrivial, proper subgroup. Hence all but at most $a-2$ terms of S are from the same H_a-coset, say $\alpha + H_a$. Hence, let H_b be a minimal cardinality nontrivial, proper subgroup of index b such that all but at most $b-2$ terms of S are all from the same H_b-coset, say $\beta + H_b$, and such that there exists an $(n-m+1)$-set partition $B = B_1, \ldots, B_{n-m+1}$ of the terms of S from $\beta + H_a$ with $|B_i| = 1$ for $i > \lfloor \frac{m}{2} \rfloor$ (in view of the previous two sentences, and taking $B_i = A'_i \cap (\alpha + H_a)$ for $i \leq \lfloor \frac{m}{2} \rfloor$, and appending on an additional $n - m + 1 - \lfloor \frac{m}{2} \rfloor$ singleton sets using the terms from $S \setminus S''$, it follows that such a subgroup exists). We may w.l.o.g. by translation assume $\beta = 0$. Let S_b be the subsequence of S consisting of terms from H_b, and let S'_b be the subsequence of S_b partitioned by the set partition $B' = B_1, \ldots, B_{\lfloor \frac{m}{2} \rfloor}$. Apply Theorem 3.1 to the subsequence S'_b of S_b with $\lfloor \frac{m}{2} \rfloor$-set partition

B' and with $G = H_b$, and let $B'' = B'_1, \ldots, B'_{\lfloor \frac{m}{2} \rfloor}$ be the resulting set partition and H_{kb} the corresponding subgroup with $[H_b : H_{kb}] = k$. If $N(B'', H_{kb}) = 1$ and $E(B'', H_{kb}) \leq k-2$, with H_{kb} a nontrivial, proper subgroup, then all but at most $k - 2 + b - 2 \leq kb - 2$ terms of S will all be from the same H_{kb}-coset, contradicting the minimality of H_b (the needed $(n-m+1)$-set partition can be induced from the set partition B'' as it was done for showing the existence of B). Therefore we may assume otherwise, whence from Theorem 3.1 it follows that $|\sum_{i=1}^{\lfloor \frac{m}{2} \rfloor} B'_i| \geq \min\{\frac{m}{b}, |S'_b| - \lfloor \frac{m}{2} \rfloor + 1\} = \frac{m}{b}$. Thus applying Proposition 2.4 to B', it follows that there exists a $\lfloor \frac{m}{2} \rfloor$-set partition $B''_1, \ldots, B''_{\lfloor \frac{m}{2} \rfloor}$ of a subsequence S''_b of S'_b with $|S''_b| \leq \lfloor \frac{m}{2} \rfloor + \frac{m}{b} - 1$, such that $|\sum_{i=1}^{\lfloor \frac{m}{2} \rfloor} B''_i| = \frac{m}{b}$. Consequently, as in the previous paragraph, it follows that there are at least $\binom{n-(\lfloor \frac{m}{2} \rfloor + \frac{m}{b} - 1) - (b-2)}{\lceil \frac{m}{2} \rceil} \geq \binom{n-m}{\lceil \frac{m}{2} \rceil}$ m-term zero-sum subsequences. Thus, since $n \leq 3m + \lceil \frac{2m-1}{3} \rceil - 2$, it follows in view of Lemma 19 that the proof is complete. So we may assume that $n \geq 3m + \lceil \frac{2m-1}{3} \rceil - 1$.

Step 4 (S essentially trichromatic): Suppose that every term of S, with at most $\lfloor \frac{m-4}{3} \rfloor$ exceptions, is equal to one of three elements x, y, $z \in G$. Let n_x, n_y, n_z be the respective multiplicities of x, y and z in S, and w.l.o.g. assume $n_x \geq n_y \geq n_z$. Let $l \leq \lfloor \frac{m-4}{3} \rfloor$ be the number of terms t of S with $t \notin \{x, y, z\}$. In view of steps 2 and 3, it follows for $w \in \{x, y, z\}$ that there are at least $\lfloor \frac{2m-1}{3} \rfloor - \lfloor \frac{m-4}{3} \rfloor \geq \lfloor \frac{m-4}{3} \rfloor + 2 \geq l + 2$ terms of S equal to w.

Claim 1. We proceed to show that if $n_x \leq \lfloor \frac{n}{2} \rfloor - l$, then for each $w \in \{x, y, z\}$ there exists an $\lceil \frac{n}{2} \rceil$-set partition $A^{(w)} = A_1, \ldots, A_{\lceil \frac{n}{2} \rceil}$ of S into cardinality at most two sets, such that if either $t \in A_j$ with $t \notin \{x, y, z\}$, or if $|A_j| = 1$, then $w \in A_j$. Since $n_w \geq l + 2$, then for i with $\lfloor \frac{n}{2} \rfloor - l + 1 \leq i \leq \lfloor \frac{n}{2} \rfloor$, let $A_i = \{w, t_i\}$, where the t_i are the terms with $t_i \notin \{x, y, z\}$, and if n is odd, then let $A_{\lceil \frac{n}{2} \rceil} = \{w\}$. Let S' be the subsequence of S obtained by deleting all terms contained in the A_i with $i \geq \lfloor \frac{n}{2} \rfloor - l + 1$. To show the claim it suffices to show S' has an $(\lfloor \frac{n}{2} \rfloor - l)$-set partition with all sets of cardinality at most two. However, from the construction in Proposition 2.3, this will be the case provided no term of S' has multiplicity at least $\lceil \frac{n}{2} \rceil - l + 1$, which we have by assumption of Claim 1. Thus the claim is established.

Claim 2. Next, we proceed to show that if $n_x \geq \lfloor \frac{n}{2} \rfloor - l + 1$, then for each $w \in \{y, z\}$ there exists an $\lceil \frac{n}{2} \rceil$-set partition $A^{(w)} = A_1, \ldots, A_n$ of S into cardinality at most two sets, such that either $x \in A_j$ or $w \in A_j$ for all j, such that if $|A_j| = 1$, then $A_j = \{w\}$, and such that for all j, $A_j \neq \{y, z\}$. Let w' be the remaining element in $\{y, z\} \setminus \{w\}$. Rearrange the sequence S so that all the terms equal to x precede all the terms equal to w, which proceed all the terms equal to w', which precede all the terms t with $t \notin \{x, y, z\}$, and let $x_1, \ldots, x_n$ be the resulting sequence. Let $A_i = \{x_i, x_{i+\lceil \frac{n}{2} \rceil}\}$ for $i \leq \frac{n}{2}$, and if n is odd, then let $A_{\lceil \frac{n}{2} \rceil} = \{x_{\lceil \frac{n}{2} \rceil}\}$. In view of Step 1 it follows that $n_x \leq \lceil \frac{n}{2} \rceil - 1$. Hence, since $n_w \geq \lfloor \frac{m-4}{3} \rfloor + 2 \geq l + 2$, and since $n_x \geq \lfloor \frac{n}{2} \rfloor - l + 1$, then it follows that the set partition $A^{(w)} = A_1, \ldots, A_{\lceil \frac{n}{2} \rceil}$ satisfies the claim.

Let $A^{(w)} = A_1, \ldots, A_{\lceil \frac{n}{2} \rceil}$ be the respective $\lceil \frac{n}{2} \rceil$-set partition constructed using w from Claim 1 (if $n_x \leq \lfloor \frac{n}{2} \rfloor - l$) or from Claim 2 (if $n_x \geq \lfloor \frac{n}{2} \rfloor - l + 1$), and w.l.o.g. re-index $A^{(w)}$ such that if n is odd, then $|A_{\lceil \frac{n}{2} \rceil}| = 1$, and such that $A_j \subsetneq \{x, y, z\}$ holds precisely for j satisfying $\lfloor \frac{n}{2} \rfloor - l + 1 \leq j \leq \lfloor \frac{n}{2} \rfloor$.

If $n_x \leq \lfloor \frac{n}{2} \rfloor - l$, then suppose for some $w \in \{x, y, z\}$ that difference of elements in $\{x, y, z\} \setminus \{w\}$ generates a subgroup H_b of index $b \leq 2$, and if $n_x \geq \lfloor \frac{n}{2} \rfloor - l + 1$, then suppose that for some $w \in \{y, z\}$ that difference of elements in $\{x, y, z\} \setminus \{w\}$ generates a subgroup H_b of index $b \leq 2$. Let $A_{i_1}, \ldots, A_{i_m}$ be a selection of m sets A_i from $A^{(w)}$.

First suppose that $b = 1$. As seen in Step 2, it is sufficient to show that any such selection $A_{i_1}, \ldots, A_{i_m}$ is good. We proceed to show this claim. If $|\sum_{j=1}^{m} A_{i_j}| \geq m$, then the selection is good in view of Proposition 2.4. Therefore we may assume that $|\sum_{j=1}^{m} A_{i_j}| < m$, whence from Kneser's theorem, it follows that $\sum_{j=1}^{m} A_{i_j}$ is maximally H_a-periodic for some proper, nontrivial subgroup H_a of index a, and that $|A_{i_j}| > |\phi_a(A_{i_j})|$ holds for at least $m-1-(a-2)$ sets A_{i_j}. Hence, since there are at most $\lfloor \frac{m-1}{3} \rfloor < m-a+1$ sets A_i with either $|A_i| = 1$ or $A_i \subsetneq \{x, y, z\}$, it follows that $|A_{i_{j'}}| < |\phi_a(A_{i_{j'}})|$ holds for some $A_{i_{j'}}$ with $A_{i_{j'}} \subseteq \{x, y, z\}$ and $|A_{i_{j'}}| = 2$. Hence, since the difference of the pair from $\{x, y, z\}$ not containing w generates G, it follows that $w \in A_{i_{j'}}$. Thus it follows from the pigeonhole principle and the definition of $A^{(w)}$ that every set A_{i_j} will contain a representative from the common H_a-coset $w + H_a$ (the representative being either w or the other element from $A_{i_{j'}}$, which under the case of Claim 2 will be x). If n is odd, then let $A_{i_k} = A_{\lceil \frac{n}{2} \rceil}$. Otherwise, since there are at least $m - a + 1 \geq \frac{m}{a}$ sets A_{i_j} with $|A_{i_j}| > |\phi_a(A_{i_j})| = 1$, then it follows, in view of Proposition 2.4 applied to these $\frac{m}{a}$ sets, that there is a set A_{i_k} with $|A_{i_k}| > |\phi_a(A_{i_k})|$ such that $\sum_{\substack{j=1 \\ j \neq k}}^{m} A_{i_j} = \sum_{j=1}^{m} A_{i_j}$. Thus since A_{i_k} is a subset of the H_a-coset $w + H_a$, since every set A_{i_j} contains a representative from the common H_a-coset $w + H_a$, and since $\sum_{j=1}^{m} A_{i_j}$ is H_a-periodic, it follows that $0 \in H_a = mw + H_a \subseteq t + \sum_{\substack{j=1 \\ j \neq k}}^{m} A_{i_j}$ for every $t \in A_{i_k}$, whence the selection is good. So we may assume that $b = 2$ and, consequently from the definition of $A^{(w)}$, that the difference of elements from every set A_i with $A_i \subseteq \{x, y, z\}$ generates a proper subgroup.

If $|\sum_{j=1}^{m} A_{i_j}| \geq m$, then as seen in the previous paragraph, it follows that the selection $A_{i_1}, \ldots, A_{i_m}$ is good. If this is not the case, then in view of Kneser's Theorem it follows that $\sum_{j=1}^{m} A_{i_j}$ is maximally H_a-periodic with H_a a nontrivial, proper subgroup of index a. Also, if there is a set $A_{i_j} \subseteq \{x, y, z\}$ with $w \in A_{i_j}$ and $|A_{i_j}| > |\phi_a(A_{i_j})|$, then, as in the previous paragraph, it follows that every set A_{i_j} will contain a representative from the common H_a-coset $w + H_a$ implying that the selection $A_{i_1}, \ldots, A_{i_m}$ is again good. Hence if a selection is not good, then all sets A_{i_j} with $|\phi_a(A_{i_j})| = 1$ must satisfy one of the following conditions: (a) $|A_{i_j}| = 1$, or (b) $A_{i_j} \subsetneq \{x,y,z\}$, or (c) $A_{i_j} = \{x,y,z\} \setminus \{w\}$. Since $|\sum_{j=1}^{m} A_{i_j}| < m$, then from Kneser's Theorem it follows that there can be at most $a-2$ sets A_{i_j} with $|\phi_a(A_{i_j})| = 2$ and, consequently, in view of the previous sentence, at most $a-2$ sets A_{i_j} with $A_{i_j} \subseteq \{x,y,z\}$, $|A_{i_j}| = 2$, and $w \in A_{i_j}$.

Since there are at most $\lfloor \frac{m-1}{3} \rfloor < m-a+2$ sets A_i satisfying (a) or (b), and since there are at least $m-a+2$ sets A_{i_j} with $|\phi_a(A_{i_j})| = 1$, it follows that there must be at least one set A_{i_j} that is contained in an H_a-coset and that satisfies (c). Hence $|\phi_a(\{x,y,z\} \setminus \{w\})| = 1$, implying that subgroup H_b generated by the difference of elements in $\{x,y,z\} \setminus \{w\}$ is a subgroup of H_a. Hence, since H_a is a proper subgroup, and since H_b has index $b = 2$, it follows that $H_b = H_a$. Consequently, as noted in the previous paragraph, it follows that there can be at most $a-2 = b-2 = 0$ sets A_{i_j} with $A_{i_j} \subseteq \{x,y,z\}$, $|A_{i_j}| = 2$, and $w \in A_{i_j}$.

Since $n_w \geq l+2$, it follows that there exists a subset $A_k \subseteq \{x,y,z\}$ with $w \in A_k$ and $|A_k| = 2$. In view of the previous paragraph, any selection $A_{i_1}, \ldots, A_{i_m}$ that includes the set A_k will be a good selection. Thus there are at least, in case n even, $2\binom{\lfloor \frac{n}{2} \rfloor - 1}{m-1} = \binom{\lfloor \frac{n}{2} \rfloor - 1}{m-1} + \binom{\lceil \frac{n}{2} \rceil - 1}{m-1}$, and in case n odd, $2\binom{\lfloor \frac{n}{2} \rfloor - 1}{m-1} + \binom{\lfloor \frac{n}{2} \rfloor - 1}{m-2} = \binom{\lfloor \frac{n}{2} \rfloor - 1}{m-1} + \binom{\lceil \frac{n}{2} \rceil - 1}{m-1}$, m-term zero-sum subsequences that use one of the two terms contained in A_k. Hence by induction hypothesis it follows that there are at least

$$\binom{\lfloor \frac{n}{2} \rfloor - 1}{m-1} + \binom{\lceil \frac{n}{2} \rceil - 1}{m-1} + \min\left\{ \binom{\lfloor \frac{n}{2} \rfloor - 1}{m} + \binom{\lceil \frac{n}{2} \rceil - 1}{m}, \binom{n-m-2}{\lceil \frac{2m-1}{3} \rceil} \right\} \tag{12.9}$$

m-term zero-sum subsequences. In view of the Pascal Identity, it follows that $\binom{\lfloor \frac{n}{2} \rfloor - 1}{m-1} + \binom{\lceil \frac{n}{2} \rceil - 1}{m-1} + \binom{\lfloor \frac{n}{2} \rfloor - 1}{m} + \binom{\lceil \frac{n}{2} \rceil - 1}{m} = \binom{\lfloor \frac{n}{2} \rfloor}{m} + \binom{\lceil \frac{n}{2} \rceil}{m}$. Hence in view of (12.9), it follows that the proof will be complete unless

$$\binom{n-m-2}{\lceil\frac{2m-1}{3}\rceil} < \binom{\lceil\frac{n}{2}\rceil-1}{m} + \binom{\lfloor\frac{n}{2}\rfloor-1}{m}, \tag{12.10}$$

and

$$\binom{\lceil\frac{n}{2}\rceil-1}{m-1} + \binom{\lfloor\frac{n}{2}\rfloor-1}{m-1} + \binom{n-m-2}{\lceil\frac{2m-1}{3}\rceil} < \binom{n-m}{\lceil\frac{2m-1}{3}\rceil}.$$

From the above inequality and the Pascal Identity, it follows that

$$\binom{\lceil\frac{n}{2}\rceil-1}{m-1} + \binom{\lfloor\frac{n}{2}\rfloor-1}{m-1} < \binom{n-m-2}{\lceil\frac{2m-1}{3}\rceil-1} + \binom{n-m-1}{\lceil\frac{2m-1}{3}\rceil-1}. \tag{12.11}$$

From (12.10) it follows that $n \geq 2m+2$. Hence applying to (12.11) the binomial identity given in (12.3), as well as the binomial identity $\binom{n}{m} = \frac{n}{n-m}\binom{n-1}{m}$, it follows that

$$\binom{\lceil\frac{n}{2}\rceil-1}{m-1} + \binom{\lfloor\frac{n}{2}\rfloor-1}{m-1} < \frac{\lceil\frac{2m-1}{3}\rceil}{(n-m-\lceil\frac{2m-1}{3}\rceil-1)}(1+\frac{n-m-1}{n-m-\lceil\frac{2m-1}{3}\rceil})\binom{n-m-2}{\lceil\frac{2m-1}{3}\rceil}.$$

Applying (12.3) to the above inequality yields

$$\binom{\lceil\frac{n}{2}\rceil-1}{m} + \binom{\lfloor\frac{n}{2}\rfloor-1}{m} <$$
$$\frac{\lceil\frac{n}{2}\rceil-m}{m} \cdot \frac{\lceil\frac{2m-1}{3}\rceil}{(n-m-\lceil\frac{2m-1}{3}\rceil-1)}(1+\frac{n-m-1}{n-m-\lceil\frac{2m-1}{3}\rceil})\binom{n-m-2}{\lceil\frac{2m-1}{3}\rceil}.$$

Hence from (12.10) it follows that

$$1 < \frac{\frac{n+1}{2}-m}{m} \cdot \frac{\frac{2m+1}{3}}{(n-m-\frac{2m+1}{3}-1)}(1+\frac{n-m-1}{n-m-\frac{2m+1}{3}}),$$

implying that $3(m-1)n^2 - (10m^2+7m+1)n + 9m^3 + 17m^2 + 8m + 2 < 0$. Hence from the quadratic formula, it follows that $8m^4 - 44m^3 - 177m^2 - 86m - 25 \leq 0$, else the square root of the discriminant will be imaginary. However, from Proposition 12.3 it follows that the roots of $8m^4 - 44m^3 - 177m^2 - 86m - 25$ are bounded from above by 9, whence $8m^4 - 44m^3 - 177m^2 - 86m - 25 > 0$ holds for $m > 10$, a contradiction. So we may assume that if $n_x \leq \lfloor \frac{n}{2} \rfloor - l$, then none of $x - z$, $x - y$, and $y - z$ generates a subgroup H_b of index $b \leq 2$, and if $n_x \geq \lfloor \frac{n}{2} \rfloor - l + 1$, then none of $x - y$ and $x - z$ generates a subgroup H_b of index $b \leq 2$ in G.

For $t \in \{x, y, z\}$ if $n_x \leq \lfloor \frac{n}{2} \rfloor - l$, and for $t \in \{y, z\}$ if $n_x \geq \lfloor \frac{n}{2} \rfloor - l + 1$, let H_{b_t} be the subgroup of index b_t generated by the difference of the elements in $\{x, y, z\} \setminus \{t\}$. From the conclusion of the previous paragraph, it follows that $b_t > 2$ for each t. Thus given any selection $A_{i_1}, \ldots, A_{i_m}$ with all A_{i_j} satisfying $A_{i_j} \subseteq \{x, y, z\}$ and $|A_{i_j}| = 2$, it follows from the pigeonhole principle that there are at least $\frac{m}{b_t} - 1$ sets A_{i_j} equal to $\{x, y, z\} \setminus \{t\}$ for some t. Note that $\sum_{i=1}^{\frac{m}{b_t}-1} \{x, y, z\} \setminus \{t\}$ is an H_{b_t}-coset, implying that $\sum_{j=1}^{m} A_{i_j}$ is maximally H_a-periodic with $H_{b_t} \leq H_a$. Thus in view of Proposition 2.4 applied with elements considered modulo H_{b_t}, it follows that there exists a re-indexing such that

$$|\sum_{j=1}^{\frac{m}{b_t}-1+b_t-1} A_{i_j}| = |\sum_{j=1}^{m} A_{i_j}|; \tag{12.12}$$

furthermore, from Kneser's Theorem it follows that $|\phi_a(A_{i_j})| = 1$ for $i > \frac{m}{b_t} + b_t - 2$, since otherwise $\left|\phi_a\left(\sum_{j=1}^{\frac{m}{b_t}+b_t-2} A_{i_j}\right)\right| < \left|\phi_a\left(\sum_{j=1}^{\frac{m}{b_t}+b_t-2} A_{i_j}\right) + \phi_a\left(\sum_{j=\frac{m}{b_t}+b_t-1}^{m} A_{i_j}\right)\right|$, implying

$$\left|\sum_{j=1}^{\frac{m}{b_t}+b_t-2} A_{i_j}\right| < \left|\left(\sum_{j=1}^{\frac{m}{b_t}+b_t-2} A_{i_j}\right) + \sum_{j=\frac{m}{b_t}+b_t-1}^{m} A_{i_j}\right| = |\sum_{j=1}^{m} A_{i_j}|,$$

which contradicts (12.12). Since $A_{i_j} \subseteq \{x, y, z\}$ with $|A_{i_j}| = 2$ holds for all j, and since

$|\phi_{b_t}(\{x,y,z\}\setminus\{t\})|=1$ implies $|\phi_a(\{x,y,z\}\setminus\{t\})|=1$ (since $H_{b_t}\leq H_a$), it follows in view of the pigeonhole principle that every set A_{i_j} contains a representative from the H_a-coset $\{x,y,z\}\setminus\{t\}+H_a$, whence from (12.12) and the previous sentence it follows that there are at least $2^{m-\frac{m}{b_t}-b_t+2}>0$ ways to select a term from each A_{i_j} and have the resulting m-term sequence be zero-sum. Thus we conclude that there are at least $2^{m-\frac{m}{b_t}-b_t+2}\binom{\lfloor\frac{n}{2}\rfloor-\lfloor\frac{m-4}{3}\rfloor}{m}$ m-term zero-sum subsequences. If $b_t\neq\frac{m}{2}$ for every such selection $A_{i_1},\ldots,A_{i_m}$, then in view of $b_t>2$, it follows for $m\geq 30$ that $2^{m-\frac{m}{b_t}-b_t+2}\geq 2^{\frac{2}{3}m-1}=2\cdot 4^{\frac{m}{3}-1}\geq 2\cdot 3^{\lfloor\frac{m-1}{3}\rfloor}$, whence the proof is complete in view of Lemma 21 and Step 3. So we may assume $b_t=\frac{m}{2}$ for some such selection $A_{i_1},\ldots,A_{i_m}$, and it suffices to further show that each selection $A_{i_1},\ldots,A_{i_m}$, with all A_{i_j} satisfying $A_{i_j}\subseteq\{x,y,z\}$ and $|A_{i_j}|=2$, and with $b_t=\frac{m}{2}$, also has at least $2\cdot 3^{\lfloor\frac{m-1}{3}\rfloor}$ ways to select an m-term zero-sum subsequence. We proceed to show this, which will complete the proof of Step 4.

Since $b_t=\frac{m}{2}$, we may w.l.o.g. by translation assume $\{x,y,z\}\setminus\{t\}=\{0,s\}$, where s has order 2. Since $t-0=t$ does not generate a subgroup with index $b\leq 2$, implying the order of t is strictly less than $\frac{m}{2}$, and since $|G/H_{b_t}|=\frac{m}{2}$, it follows that $\phi_{b_t}(t)$ generates a proper subgroup $H_{b'}$ of G/H_{b_t} with index $b'\geq 2$ in G/H_{b_t}.

Suppose that there are at least $2\lfloor\frac{m-1}{3}\rfloor+2$ sets A_{i_j} with $|\phi_{b_t}(A_{i_j})|=1$. Then, since $|H_{b_t}|=2$ implies that $|A_{i_{j_1}}+A_{i_{j_2}}|=|A_{i_{j_1}}|$ when $|\phi_{b_t}(A_{i_{j_1}})|=|\phi_{b_t}(A_{i_{j_2}})|=1$, it follows that we can re-index such that $|\sum\limits_{j=1}^{m-(2\lfloor\frac{m-1}{3}\rfloor+1)}A_{i_j}|=|\sum\limits_{j=1}^{m}A_{i_j}|$, with $|\phi_{b_t}(A_{i_j})|=1$ for $j>m-(2\lfloor\frac{m-1}{3}\rfloor+1)$. Since there are at least $2^{m-\frac{m}{b_t}-b_t+2}>0$ ways to select an m-term zero-sum from the selection $A_{i_1},\ldots,A_{i_m}$, it follows that $0\in\sum\limits_{j=1}^{m}A_{i_j}$. Thus, since $|\sum\limits_{j=1}^{m-(2\lfloor\frac{m-1}{3}\rfloor+1)}A_{i_j}|=|\sum\limits_{j=1}^{m}A_{i_j}|$, it follows that there will be at least $2^{2\lfloor\frac{m-1}{3}\rfloor+1}\geq 2\cdot 3^{\lfloor\frac{m-1}{3}\rfloor}$ ways to select an m-term zero-sum subsequence from the selection $A_{i_1},\ldots,A_{i_m}$, completing the proof as noted earlier. So, we may assume there are at least $m-(2\lfloor\frac{m-1}{3}\rfloor+1)\geq\lceil\frac{m-1}{3}\rceil\geq\frac{m}{2b'}-1=\frac{b_t}{b'}-1$

sets A_{i_j} with $|\phi_{b_t}(A_{i_j})| = 2$.

Hence, since $|\phi_{b_t}(\{0,s\})| = 1$, and since $\phi_{b_t}(\{0,t\}) = \phi_{b_t}(\{s,t\})$ (since $|\phi_{b_t}(\{0,s\})| = 1$ implies $\phi_{b_t}(0) = \phi_{b_t}(s)$), it follows that there are at least $\frac{b_t}{b'} - 1$ sets A_{i_j} that modulo H_{b_t} have the difference of their elements generating the subgroup $H_{b'} = \langle \phi_{b_t}(t) \rangle$. Note that $\sum_{i=1}^{\frac{b_t}{b'}-1} \phi_{b_t}(\{0,t\}) = H_{b'}$. Hence, since there are at least $\frac{b_t}{b'} - 1$ sets A_{i_j} with $|\phi_{b_t}(A_{i_j})| = 2$, and since there are at least $\frac{m}{b_t} - 1$ sets A_{i_j} equal to $\{x,y,z\} \setminus \{t\} = \{0,s\}$, it follows in view of Proposition 2.4 applied with elements considered in $(G/H_{b_t})/H_{b'}$, that there exists a re-indexing such that $|\sum_{j=1}^{\frac{m}{b_t}-1+\frac{b_t}{b'}-1+b'-1} A_{i_j}| = |\sum_{j=1}^{m} A_{i_j}|$; furthermore, from Kneser's Theorem it follows that $|\phi_a(A_{i_j})| = 1$ for $i > \frac{m}{b_t} - 1 + \frac{b_t}{b'} - 1 + b' - 1 = \frac{b_t}{b'} + b' - 1$, since otherwise

$$|\sum_{j=1}^{\frac{b_t}{b'}+b'-1} A_{i_j}| < \left| \left(\sum_{j=1}^{\frac{b_t}{b'}+b'-1} A_{i_j} \right) + \sum_{j=\frac{b_t}{b'}+b'}^{m} A_{i_j} \right|$$

will hold, a contradiction. Thus, since $\frac{b_t}{b'} + b' - 1 \leq \frac{m}{4} + 1$, and since $0 \in \sum_{j=1}^{m} A_{i_j}$, it follows that there will be at least $2^{\frac{3}{4}m-1} \geq 2 \cdot 3^{\lfloor \frac{m-1}{3} \rfloor}$ ways to select an m-term zero-sum subsequence from the selection $A_{i_1}, \ldots, A_{i_m}$, completing the proof of Step 4 as noted earlier. So we may assume that given any $x, y, z \in G$, there are at least $\lfloor \frac{m-1}{3} \rfloor$ terms of S not equal to x or y or z.

Step 5 (The general case): In view of Steps 1, 2, 3 and 4, and Proposition 2.3, it follows that there exists an $(n-m+1)$-set partition $P = P_1, \ldots, P_{n-m+1}$ of S with $|P_i| = 1$ for $i > \lceil \frac{m-1}{3} \rceil$. Let $P' = P_1, \ldots, P_{\lceil \frac{m-1}{3} \rceil}$, and let S' be the corresponding subsequence partitioned by the set partition P'. Apply Theorem 3.1 to the subsequence S' of S with $\lceil \frac{m-1}{3} \rceil$-set partition P', and let S'' be the resulting subsequence, H_a the resulting subgroup of index a, and $A = A_1, \ldots, A_{\lceil \frac{m-1}{3} \rceil}$ the resulting set partition of S''.

Suppose that $|\sum_{i=1}^{\lceil\frac{m-1}{3}\rceil} A_i| \geq m = \sum_{i=1}^{\lceil\frac{m-1}{3}\rceil}|A_i| - \lceil\frac{m-1}{3}\rceil + 1$. Then applying Theorem 6.1 to A and S'', it follows that there exists a subsequence T of S'' of length at most m with a set partition $B = B_1, \ldots, B_{\lceil\frac{m-1}{3}\rceil}$ such that $|\sum_{i=1}^{\lceil\frac{m-1}{3}\rceil} B_i| \geq m$. Then given any subsequence T' of $S \setminus T$ of length $m - \lceil\frac{m-1}{3}\rceil = \lceil\frac{2m-1}{3}\rceil$, we can find a selection of $\lceil\frac{m-1}{3}\rceil$ terms from T, one from each of the $B_1, \ldots, B_{\lceil\frac{m-1}{3}\rceil}$, that sum to the additive inverse of the sum of the terms from the $\lceil\frac{2m-1}{3}\rceil$-term subsequence T'. Consequently, there will be at least $\binom{n-m}{\lceil\frac{2m-1}{3}\rceil}$ m-term zero-sum subsequences, completing the proof. So we can assume that

$$|\sum_{i=1}^{\lceil\frac{m-1}{3}\rceil} A_i| < m = \sum_{i=1}^{\lceil\frac{m-1}{3}\rceil}|A_i| - \lceil\frac{m-1}{3}\rceil + 1. \tag{12.13}$$

Hence from Theorem 3.1 it follows that H_a is a proper, nontrivial subgroup, and that either $N(A, H_a) = 1$ and $E(A, H_a) \leq a - 2$, or else $N(A, H_a) = 2$, $|H_a| = 2$, and $E(A, H_a) \leq \frac{m}{2} - \lceil\frac{m-1}{3}\rceil - 2 \leq \lfloor\frac{m-10}{6}\rfloor$. The case $N(A, H_a) = 1$ and $E(A, H_a) \leq a - 2$ can be handled by a minor modification of the arguments from the third paragraph of Step 3 (simply replace $\lfloor\frac{m}{2}\rfloor$ by $\lceil\frac{m-1}{3}\rceil$ where appropriate). Therefore we may assume the latter case holds.

Since $N(A, H_a) = 2$, choose $x, y \in G$ so that $\phi_a(x), \phi_a(y) \in G/H_a$ are the two elements from $\phi_a\left(\bigcap_{i=1}^{\lceil\frac{m-1}{3}\rceil}(A_i + H_a)\right)$. Suppose first that $\phi_a(x - y)$ generates a proper subgroup $H_{a'}/H_a$ of G/H_a. If there does not exist a set $A_{j'}$ such that $(\{x, y\} + H_a) \subseteq A_{j'}$, then there will be at least $\lceil\frac{m-1}{3}\rceil = \lceil\frac{m-1}{3}\rceil(|H_a| - 1)$ holes contained among the sets A_{i_j}, which, in view of the comments before Section 3.2, implies that (12.13) cannot hold, a contradiction. Therefore we may assume that there exists a set $A_{j'}$ with $(\{x, y\} + H_a) \subseteq A_{j'}$.

For $i = j'$, let $B_{j'} = (\{x, y\} + H_a) \cap A_{j'} = \{x, y\} + H_a$, and for $i \neq j'$, let B_i be a cardinality two subset of $A_i \cap (\{x, y\} + H_a)$ with $|\phi_a(B_i)| = 2$. Then, since $\phi_a(x - y)$ generates a proper subgroup $H_{a'}/H_a$, and since $\lceil\frac{m-1}{3}\rceil \geq \frac{m}{4} \geq |G/H_{a'}|$, it follows that

$\sum_{i=1}^{\lceil\frac{m-1}{3}\rceil} B_i$ is an $H_{a'}$-coset. Observe that all but at most $E(A, H_a) \leq \frac{m}{2} - \lceil\frac{m-1}{3}\rceil - 2$ terms of S are from the same $H_{a'}$-coset $x + H_{a'}$. Let T be the subsequence of S partitioned by the set partition $B = B_1, \ldots, B_{\lceil\frac{m-1}{3}\rceil}$. Hence, since $B_i \subseteq x + H_{a'}$ for all i, and since $\sum_{i=1}^{\lceil\frac{m-1}{3}\rceil} B_i$ is an $H_{a'}$-coset, it follows that given any $\lceil\frac{2m-1}{3}\rceil$-term subsequence T' of $S \setminus T$ with all terms from the coset $x + H_{a'}$, then we can find a selection of $\lceil\frac{m-1}{3}\rceil$ terms from T, one from each $B_1, \ldots, B_{\lceil\frac{m-1}{3}\rceil}$, that sums to the additive inverse of the sum of terms from T'. Hence, since there are at least $n-(2\lceil\frac{m-1}{3}\rceil+2+E(A', H_a)) \geq n-(2\lceil\frac{m-1}{3}\rceil+2+\frac{m}{2}-\lceil\frac{m-1}{3}\rceil-2) \geq n-m$ terms of $S \setminus T$ from the coset $x + H_{a'}$, it follows that there are at least $\binom{n-m}{\lceil\frac{2m-1}{3}\rceil}$ m-term zero-sum subsequences, completing the proof. So we may assume that $\phi_a(x-y)$ generates G/H_a.

Let x' be the other element from the coset $x + H_a$, and let y' be the other element from the coset $y + H_a$. Let n_x, $n_{x'}$, n_y and $n_{y'}$ be the respective multiplicities of x, x', y and y' in S. Since, as noted previously, there is a set $A_{j'}$ such that $(\{x, y\} + H_a) \subseteq A_{j'}$, it follows that $n_x \geq 1$, $n_{x'} \geq 1$, $n_y \geq 1$ and $n_{y'} \geq 1$. We may w.l.o.g. assume that $n_x + n_{x'} \geq n_y + n_{y'}$, that $n_x \geq n_{x'}$ and that $n_y \geq n_{y'}$. Remove two terms from S, one equal to x and one equal to x', and let the resulting sequence be T. Let B_0 be the set consisting of the two removed terms. Rearrange the terms of T so that all terms equal to x precede all terms equal to x', which precede all terms equal to y, which precede all terms equal to y', which precede all terms t with $t \notin \{x, x', y, y'\}$, and let $x_1, \ldots, x_{n-2}$ be the resulting sequence. Let $B_i = \{x_i, x_{i+\lceil\frac{n}{2}\rceil-1}\}$ for $i = 1, \ldots, \lfloor\frac{n}{2}\rfloor - 1$, and, in case n odd, let $B_{\lceil\frac{n}{2}\rceil-1} = \{x_{\lceil\frac{n}{2}\rceil-1}\}$. In view of Step 1, it follows that $B = B_1, \ldots, B_{\lceil\frac{n}{2}\rceil-1}$ is an $(\lceil\frac{n}{2}\rceil-1)$-set partition of T. As seen in the ninth paragraph of Step 4, it suffices by induction hypothesis to show that any selection $B_0, B_{i_1}, \ldots, B_{i_{m-1}}$ containing B_0 is good. We proceed to show this.

If $|\sum_{j=1}^{m-1}\phi_a(B_{i_j})| \geq \frac{m}{2}$, then, since B_0 is an H_a-coset, it follows that $|B_0 + \sum_{j=1}^{m-1} B_{i_j}| \geq m$, whence from Proposition 2.4 it follows that the selection $B_0, B_{i_1}, \ldots, B_{i_{m-1}}$ is good. Hence we may assume that

$$|\sum_{j=1}^{m-1}\phi_a(B_{i_j})| < \frac{m}{2}. \tag{12.14}$$

Suppose that $n_x + n_{x'} > \lceil \frac{n}{2} \rceil$. Then every set B_{i_j} will contain a representative from the common H_a-coset $x + H_a$. Hence, since B_0 is H_a-periodic, it follows that $0 \in B_0 + \sum_{j=1}^{m-1} B_{i_j}$.

Suppose further that $|\phi_a(B_{i_k})| = 1$ holds for some B_{i_k} with $i_k \geq 1$. Then $B_0 + \sum_{\substack{j=1\\ j\neq k}}^{m-1} B_{i_j} = B_0 + \sum_{j=1}^{m-1} B_{i_j}$, and it follows that either $|B_{i_k}| = 1$, or else there will be at least two ways to represent every $x \in B_0 + \sum_{j=1}^{m-1} B_j$. Hence, since $0 \in B_0 + \sum_{j=1}^{m-1} B_j$, it follows that the selection $B_0, B_{i_1}, \ldots, B_{i_{m-1}}$ is good, completing the proof as noted earlier. So we may assume $|\phi_a(B_{i_k})| = 2$ for all $i_k \geq 1$.

Since $|\phi_a(B_{i_k})| = 2$ for all $i_k \geq 1$, and since $|\phi_a(\{x, x'\})| = 1$, it follows that there does not exist a set B_{i_j} with $i_j \geq 1$ and $B_{i_j} = \{x, x'\}$. Hence, since there are at most $E(A, H_a) \leq \frac{m-10}{6}$ terms t with $t \notin \{x, x', y, y'\}$, and since every set B_{i_j} contains either x or x', it follows that there are at least $m - 2 - \frac{m-10}{6} \geq \frac{m}{2}$ sets B_{i_j} with the difference of terms in B_{i_j} equal modulo H_a to $\phi_a(x-y)$. Thus, since $\phi_a(x-y)$ generates G/H_a, it follows that (12.14) cannot hold, a contradiction. So we may assume that $n_x + n_{x'} \leq \lceil \frac{n}{2} \rceil$.

Since $n_x + n_{x'} \leq \lceil \frac{n}{2} \rceil$, since $n_x + n_{x'} \geq n_y + n_{y'}$, and since all but at most $E(A', H_a) \leq \frac{m}{2} - \lceil \frac{m-1}{3} \rceil - 2 \leq \lfloor \frac{m-10}{6} \rfloor$ terms of S are equal to one of x, x', y or y', it follows that at least $(m-3) - \frac{m-10}{6} \geq \frac{m}{2}$ sets B_{i_j} have $\phi_a(B_{i_j}) = \{\phi_a(x), \phi_a(y)\}$. Hence, since $\phi_a(x) - \phi_a(y)$ generates G/H_a, it follows that $|\sum_{j=1}^{m-1}\phi_a(B_{i_j})| \geq \frac{m}{2}$, contradicting (12.14) again, and completing the proof. $\square$

Bibliography

[1] N. Alon, A. Bialostocki and Y. Caro, The extremal cases in the Erdős-Ginzburg-Ziv Theorem, unpublished.

[2] N. Alon, M. B. Nathanson and I. Ruzsa, The polynomial method and restricted sums of congruence classes, *J. Number Theory*, 56 (1996), no. 2, 404–417.

[3] A. Bialostocki, Some combinatorial number theory aspects of Ramsey Theory, research proposal, 1989.

[4] A. Bialostocki and M. Lotspeich, Developments of the Erdős-Ginzburg-Ziv Theorem I, *Sets, graphs and numbers*, Budapest (1991), 97–117.

[5] A. Bialostocki and P. Dierker, On the Erdős-Ginzburg-Ziv Theorem and the Ramsey numbers for stars and matchings, *Discrete Math.*, 110 (1992), no. 1–3, 1–8.

[6] A. Bialostocki, P. Erdős and H. Lefmann, Monochromatic and zero-sum sets of nondecreasing diameter, *Discrete Math.*, 137 (1995), 19–34.

[7] A. Bialostocki, P. Dierker, D. Grynkiewicz and M. Lotspeich, On some developments of the Erdős-Ginzburg-Ziv Theorem II, *Acta. Arith.*, 110 (2003), no. 2, 173–184.

[8] B. Bollobás and I. Leader, The number of k-sums modulo k, *J. Number Theory*, 78 (1999), 27–35.

[9] Y. Caro, Zero-sum problems—a survey, *Discrete Math.*, 152 (1996), no. 1–3, 93–113.

[10] A. L. Cauchy, Recherches sur les nombres, *J. École polytech.*, 9 (1813), 99–116.

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

[12] J. A. Dias da Silva and Y. O. Hamidoune, A note on the minimal polynomial of the Kronecker sum of two linear operators, *Linear Algebra Appl.*, 141 (1990), 283–287.

[13] P. Dubreil and C. Piscot, *Algébre et théorie des nombres. Séminaire 1955–56*, Paris, 1956.

[14] R. B. Eggleton and P. Erdős, Two combinatorial problems in group theory, *Acta Arithmetica*, 21 (1972), 111–116.

[15] P. Erdős, A. Ginzburg and A. Ziv, Theorem in additive number theory, *Bull. Res. Council Israel*, 10F (1961), 41–43.

[16] Z. Füredi and D. J. Kleitman, The minimal number of zero sums; in: *Combinatorics, Paul Erdős is eighty*, Vol. 1, Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 1993, 159–172.

[17] L. Gallardo, G. Grekos and J. Pihko, On a variant of the Erdős-Ginzburg-Ziv problem, *Acta Arith.*, 89 (1999), no. 4, 331–336.

[18] W. D. Gao, Addition theorems for finite abelian groups, *J. Number Theory*, 53 (1995), no. 2, 241–246.

[19] W. D. Gao, A combinatorial problem on finite abelian groups, *J. Number Theory*, 58 (1996), no. 1, 100–103.

[20] W. D. Gao, An addition theorem for finite cyclic groups, *Discrete Math.*, 163 (1997), no. 1–3, 257–265.

[21] W. D. Gao and Y. O. Hamidoune, Zero sums in abelian groups, *Combin. Prob. Comput.*, 7 (1998), no. 3, 261–263.

[22] W. Gao and X. Jin, Weighted sums in finite cyclic groups, *Discrete Math.*, 283 (2004), no. 1–3, 243–247.

[23] D. Grynkiewicz, On four colored sets with nondecreasing diameter and the Erdős-Ginzburg-Ziv Theorem, *J. Combin. Theory Ser. A*, 100 (2002), 44–60.

[24] D. Grynkiewicz, On a partition analog of the Cauchy-Davenport Theorem, *Acta Math. Hungar.*, 107 (2005), no. 1–2, 161–174.

[25] D. Grynkiewicz, Quasi-Periodic decompostions and the Kemperman Structure Theorem, *European J. Combin.*, 26 (2005), no. 5, 543–781.

[26] D. Grynkiewicz, On a conjecture of Hamidoune for subsequence sums, to appear in *Integers.*

[27] D. Grynkiewicz, An extension of the Erdős-Ginzburg-Ziv Theorem to Hypergraphs, to appear in *European J. Combin.*

[28] D. Grynkiewicz, A weighted Erdős-Ginzburg-Ziv Theorem, to appear in *Combinatorica.*

[29] D. Grynkiewicz and A. Schultz, A 5-color zero-sum generalization, submitted (2002).

[30] D. Grynkiewicz and R. Sabar, Monochromatic and zero-sum sets of nondecreasing modified-diameter, submitted (2002).

[31] D. Grynkiewicz, On the number of m-term zero-sum subsequences, submitted (2005).

[32] H. Halberstam and K. F. Roth, *Sequences*, Springer-Verlag, New York, 1983.

[33] Y. O. Hamidoune, On weighted sequence sums, *Comb. Prob. Comput.*, 4 (1995), 363–367.

[34] Y. O. Hamidoune, On weighted sums in abelian groups, *Discrete Math.*, 162 (1996), 127–132.

[35] Y. O. Hamidoune, Subsets with small sums in abelian groups I: The Vosper property, *European J. Combin.*, 18 (1997), no. 5, 541–556.

[36] Y. O. Hamidoune, Some results in additive number theory I: The critical pair theory, *Acta Arith.*, 96 (2000), no. 2, 97–119.

[37] Y. O. Hamidoune, Subsets with a small sum II: The critical pair problem, *European J. Combin.*, 21 (2000), no. 2, 231–239.

[38] Y. O. Hamidoune, Subsequence sums, *Combin. Prob. Comput.*, 12 (2003), 413–425.

[39] X. Hou, K. Leung and Q. Xiang, A generalization of an addition theorem of Kneser, *J. Number Theory*, 97 (2002), 1–9.

[40] J. H. B. Kemperman, On complexes in a semi-group, *Indag. Math.*, 18 (1956), 247–254.

[41] J. H. B. Kemperman, On small sumsets in an abelian group. *Acta Math.*, 103 (1960), 63–88.

[42] M. Kisin, The number of zero sums modulo m in a sequence of length n, *Mathematika*, 41 (1994), no. 1, 149–163.

[43] M. Kneser, Abschätzung der asymptotischen dichte von summenmengen, *Math. Z.*, 58 (1953), 459–484.

[44] M. Kneser, Ein satz über abelsche gruppen mit anwendungen auf die geometrie der zahlen, *Math. Z.*, 64 (1955), 429–434.

[45] V. Lev, On small sumsets in abelian groups, *Astérisque*, no. 258 (1999), xv, 317-321.

[46] H. B. Mann, An addition theorem for sets of elements of an abelian group, *Proc. Am. Math. Soc.*, 4 (1953), 423.

[47] H. B. Mann, *Addition theorems: the addition theorems of group theory and number theory*, Interscience, New York, 1965.

[48] H. B. Mann, Two addition theorems, *J. Combinatorial Theory*, 3 (1967), 233–235.

[49] M. B. Nathanson, *Additive number theory: inverse problems and the geometry of sumsets*, Graduate Texts in Mathematics, 165, Springer-Verlag, New York, 1996.

[50] J. E. Olson, An addition theorem for finite abelian groups, *J. Number Theory*, 9 (1977), 63–70.

[51] C. Reiher, On Kemnitz's conjecture concerning lattice points in the plane, preprint.

[52] P. Scherk, Distinct elements in a set of sums, *Amer. Math. Monthly*, 62 (1955), 46–47.

[53] T. Tao, An uncertainty principle for cyclic groups of prime order, preprint.