Sum-free Sets and Arithmetic Notions of Structure in Combinatorics

Benjamin Bedert

Magdalen College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

This thesis is dedicated to
my mum, dad, brothers and Romana,
for their unwavering support and encouragement

Acknowledgements

I would like to thank my supervisor Ben Green for his support and expert guidance, and for many helpful and inspirational discussions that have played an important role in shaping my DPhil.

I am grateful to the Engineering and Physical Sciences Research Council for providing the financial support that made this DPhil possible. I am also grateful to Magdalen College, Oxford for their additional support and for my temporary appointment as a Lecturer II.

I would further like to express my gratitude to a number of people who through collaboration, fruitful discussions or proposing problems have contributed to some of the work appearing in this thesis. This list includes Thomas Bloom, Zachary Chase, Tamás Erdélyi, Swastik Kopparty, Noah Kravitz, Vsevolod Lev, Cédric Pilatte, Julian Sahasrabudhe, Yifan Jing, and many other mathematicians that I have interacted with, particularly at the 2024 Additive Combinatorics Workshop hosted by the Erdős Center in Budapest and the 2024 Additive Combinatorics conference hosted by the ICMS in Edinburgh. Further, I was fortunate to be a part of the Analytic Number Theory group in Oxford and I am thankful to all its members for contributing to an enjoyable and stimulating academic environment.

Abstract

This thesis is concerned with four problems in additive combinatorics, each of which studies a certain notion of arithmetic structure and which properties of a general additive set, such as its density or the structure of the group in which it is contained, imply that it necessarily exhibits such structure. One of the main results of this thesis shows that any set A of N positive integers contains a subset $A' \subset A$ of size $|A'| \geqslant \frac{N}{3} + c \log\log N$ which is sum-free, meaning that there are **no** three $x, y, z \in A'$ with $x + y = z$. This answers a longstanding problem of Erdős from 1965.

Contents

Chapter 1

Introduction

This thesis is comprised of a collection of four results of additive combinatorial nature. Broadly speaking, one can view each of these results as describing various properties of certain notions of arithmetic structure, and to what extent a general additive set is guaranteed to contain this type of structure (under some natural assumptions). This is an old and central theme within the subject of additive combinatorics, going back to the landmark results of Roth [56] and Szemerédi [64] who showed that dense subsets of $\mathbf{N}$ must contain arithmetic progressions of length 3 and of length k for any k, respectively. To obtain our results, we employ a variety of techniques ranging from harmonic analysis to discrete combinatorics and graph theory. We shall also make substantial use of many of the classical tools of additive combinatorics concerning sumsets, small doubling, Freiman's theorem, additive energy, and most importantly the concepts of dissociativity and additive dimension, which we shall introduce and discuss in the next chapter. Let us now give an overview of the content of this thesis, as well as a brief outline of the proof techniques that go into each of the chapters.

In Chapter 3, we study a seminal question about sum-free sets from a 1965 paper [19] of Erdős. A set B is said to be *sum-free* if there are no $x, y, z \in B$ with $x + y = z$. Using a simple yet ingenious probabilistic argument, Erdős showed that any set of N positive integers contains a sum-free subset of size $N/3$. This lower bound turned out to be surprisingly difficult to improve upon, and the previous best bound, obtained by Bourgain [12] using an elaborate Fourier-analytic approach, was the only slightly larger quantity $(N + 2)/3$. Confirming a longstanding suspicion in additive combinatorics, we show that there exists a constant $c > 0$ such that any set A of N integers contains a sum-free subset A' of size $|A'| \geqslant N/3 + c \log\log N$. The starting point of our proof is Bourgain's Fourier-analytic approach which we develop further. One essential step consists of proving that for a set A of integers, either

its Fourier transform $\hat{1}_A(x) = \sum_{a \in A} e(ax)$ has large L^1 norm, or else it must have very small additive dimension $\dim(A)$. In the former case, Bourgain's Fourier-analytic machinery immediately shows that A has a large sum-free set. The latter case is more complicated, and we rely on several combinatorial properties of low-dimensional sets, such as that they have dense Freiman models and that their distribution modulo powers of small primes cannot be too spread out, to run a more efficient version of the Fourier-analytic argument.

Chapter 4 studies the problem of determining for a finite Abelian group G the size of its smallest subset $A \subset G$ that has no unique sum, meaning that for every two $a_1, a_2 \in A$ we can write $a_1 + a_2 = a_1' + a_2'$ for different $a_1', a_2' \in A$. Let $m(p)$ be the size of a smallest subset of $\mathbf{Z}/p\mathbf{Z}$ with no unique sum. The previous best known bounds are $\log p \ll m(p) \ll \sqrt{p}$. This chapter improves both the upper and lower bounds to $\omega(p) \log p \leqslant m(p) \ll (\log p)^2$ for some function $\omega(p)$ which tends to infinity as $p \to \infty$. We also obtain corresponding bounds on the size of the smallest subset of a general Abelian group having no unique sum. The proof of these results use several key notions in additive combinatorics such as additive dimension, compressions, rectification and the density increment method.

Chapter 5 is concerned with a problem of Erdős and Sárközy which asks how large the size of set $A \subset \{1, 2 \ldots, n\}$ can be if A satisfies what Erdős and Sárközy call property P. A set $A \subset \mathbf{N}$ is said to have *property* P if for all $x, y, z \in A$ with $x < y, z$, x does not divide $y + z$. In their original 1970 paper [29], it was stated that the authors believed that a subset $A \subset [n]$ with property P has cardinality at most $|A| \leqslant \lfloor \frac{n}{3} \rfloor + 1$. Erdős later offered \$100 for a proof or disproof of the weaker claim that $|A| \leqslant \frac{n}{3} + C$ for some absolute constant C. We resolve this problem by proving that the proposed bound is indeed true and that the stronger upper bound of $\lfloor \frac{n}{3} \rfloor + 1$ holds for all large n. The argument naturally splits into 3 cases based on the density of the set A in the crucial interval $I_n = (2n/3, n]$. The first two cases, in which A has density at least $2/3$ or between $1/2$ and $2/3$ on I_n, admit efficient and clean proofs based on (a variant of) Freiman's $3k - 4$ theorem. This theorem provides optimal structural information for sets B with very small doubling $|B+B| \leqslant 3|B|-4$, stating that B lies in a very short progression. The final case, in which A has density less than $1/2$ on I_n, is quite involved and based in part on a number of delicate ad hoc arguments. For this reason, we have decided to omit the final case from this thesis, referring the interested reader to the actual paper [5]. We do note that we manage to show in that paper that if a set A with property P falls into this final case, then

A has size $|A| \leqslant (1/3 - \varepsilon)n$ and hence cannot have size close to the optimal value of $n/3$, so there is room for some slack in this part of the argument.

The topic of study in the final chapter 6 is another well-known problem in additive combinatorics, known as Graham's rearrangement conjecture. This conjecture posits that any set $A \subset \mathbf{Z}/p\mathbf{Z}$ not containing 0 may be arranged in some order $a_1, a_2, \dots, a_n$ such that all the partial sums $a_1, a_1+a_2, \dots, a_1+a_2+\cdots+a_n$ are distinct. In this case, we say that $a_1, a_2 \dots, a_n$ is a *valid ordering* of A. After partial progress by various authors, this conjecture had only been confirmed for sets which are either very small or very large; more precisely for sets of size at most 12 or of size at least $p-3$. Recently, Kravitz [44] and Sawin independently observed that a rectification argument may be used to verify Graham's conjecture for sets of size up to $\log p/\log\log p$. Here, we manage to do significantly better and show that the conjecture is true for sets of size as large as $e^{c(\log p)^{1/4}}$, breaking the so-called rectification barrier. One important idea in the proof is to establish a structural result which decomposes any $A \subset \mathbf{Z}/p\mathbf{Z}$ into a number of large dissociated sets and one low-dimensional set. Roughly speaking, one may employ rectification techniques to find a valid ordering of the low-dimensional piece, while the lack of additive relations in dissociated sets implies that a suitably distributed random ordering of these dissociated sets is likely to be valid. Combining these two orderings into a valid ordering of A is somewhat subtle; to do this successfully, we develop an algorithm that produces, for any rectifiable set, a valid ordering whose partial sums additionally avoid most elements from prescribed sets. The content of this chapter is joint work with Noah Kravitz.

Chapter 2

Preliminaries and background

Additive combinatorics can be thought of as the area of mathematics that studies the additive properties of sets in groups, typically Abelian groups such as the integers $\mathbf{Z}$ or finite field vector spaces like $\mathbf{F}_p^n$. In this introductory chapter, we discuss some of the foundational concepts, tools and results in additive combinatorics that will be relevant in later chapters.

A central topic of investigation in additive combinatorics are the various notions of additive structure of sets and how these are related to each other. We begin by defining the most important such notions. For subsets A, B of an Abelian group, we write

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

and

$$A - B := \{a - b : a \in A, b \in B\}.$$

For a finite set A, the ratio $K = |A + A|/|A|$ is referred to as the *doubling constant* of A. A set with small doubling should be thought of as having strong additive structure. In fact, (generalised) arithmetic progressions are the canonical examples of sets with small doubling.

Definition 2.0.1. A d-dimensional *generalised arihmetic progressions* (GAP) is a set $P \subset \mathbf{Z}$ of the form

$$P := \{n_0 + \sum_{j=1}^{d} n_j d_j : 0 \leqslant n_j < L_j\},$$

for some parameters $L_j, j \in [d]$. P is said to be *proper* if the sums $\sum_{j=1}^{d} n_j d_j$ with $0 \leqslant n_j < L_j$ are pairwise distinct.

A simple check confirms that a proper d-dimensional GAP P has doubling constant at most 2^d, and hence any subset $A \subset P$ of density α has doubling constant $K \leqslant \alpha^{-1}2^d$. One of the foundational results in additive combinatorics is the following theorem of Freiman [66, Theorem 5.44], stating that every integer set with small doubling is a dense subset of some generalised arithmetic progression. From a qualitative perspective, this provides a complete description of sets with small doubling.

Theorem 2.0.2 (Freiman's Theorem)**.** *Let $K \geqslant 1$ and let $B \subset \mathbf{Z}$ be a set with doubling $|B+B| \leqslant K|B|$. Then there exists a proper generalised arithmetic progression P of dimension d such that $B \subset P$ and such that*

- *the dimension of P is bounded by $d \leqslant K^{O(1)}$,*
- *B is 'dense' in P in the sense that $|B| \geqslant e^{-K^{O(1)}}|P|$.*

If the doubling constant K is very small, then quantitatively optimal results are known such as the classical $3k-4$ theorem of Freiman [34], which shows that if $|A+A| \leqslant 3|A|-4$, then A is contained in a progression of length at most $2|A|-4$. Our arguments in Chapter 5 will make use of the following variant due to Bardaji and Grynkiewicz [4]. For a set $A \subset \mathbf{Z}$ of integers, we define $\operatorname{diam} A := \max A - \min A$ and $\gcd_*(A) := \gcd(A-A)$ to be the greatest common divisor of all differences $a-a'$ with $a, a' \in A$. Note that $\gcd_*(A)$ is the largest integer d such that A is contained in an arithmetic progression with common difference d.

Theorem 2.0.3 (Bardaji & Grynkiewicz [4], Corollary 1.2)**.** *Let A, B be non-empty subsets of $\mathbf{Z}$ with $\operatorname{diam} B \leqslant \operatorname{diam} A$ and $\gcd_*(A+B) = 1$. If*

$$|A+B| \leqslant |A| + 2|B| - 4$$

and either $\gcd_*(A) = 1$ or $|A+B| \leqslant 2|A|+|B|-3$, then $A+B$ contains an arithmetic progression with common difference 1 and length $|A|+|B|-1$.

Depending on the applications one has in mind, the doubling constant K of set may not be the best notion of additive structure of a set A. For example, K is extremely large if A contains a Sidon set, i.e. a set with no non-trivial relations of the form $x_1 + x_2 = x_3 + x_4$, of size $|A|/100$, even if most of A is highly structured. For an element x in an Abelian group G, we shall write $r_A(x)$ to denote the number of ordered pairs in A^2 whose sum equals x, so

$$r_A(x) := \left|\{(a, a') \in A^2 : a + a' = x\}\right|.$$

The additive energy of A is then defined as

$$E(A) = \sum_x r_A(x)^2,$$

and, roughly speaking, the additive energy is useful for detecting whether A contains a relatively large subset with small doubling, as in the example above. Note also that $E(A)$ simply counts the number of solutions to $a_1 + a_2 = a_3 + a_4$ with $a_j \in A$. The trivial bounds are $|A|^2 \leqslant E(A) \leqslant |A|^3$ and we think of A as having large energy if $E(A)$ is 'not much smaller' than $|A|^3$. A standard application of Cauchy-Schwarz shows that

$$E(A) \geqslant \left(\sum_x r_A(x)\right)^2 / |\operatorname{supp}(r_A)| = |A|^4/|A+A|,$$

implying that any set A with small doubling automatically has large energy. Since $E(A) \geqslant E(B)$ for all subsets $B \subset A$, more generally the additive energy of a set A is large if A contains a large subset with small doubling. The following important classical result [66, Theorem 2.27] connects high additive energy with the small doubling property in the reverse direction.

Theorem 2.0.4 (Balog-Szemerédi-Gowers Theorem)**.** *Let $K \geqslant 1$ and let $B \subset \mathbf{Z}$ have additive energy $E(B) \geqslant |B|^3/K$. Then there exists a subset $B' \subset B$ of size $|B'| \gg |B|/K^{O(1)}$ with small doubling $|B' - B'| \ll K^{O(1)}|B|$.*

We have not stated either of Freiman's or the Balog-Szemerédi-Gowers Theorems with the best currently known quantitative dependence on K. Another powerful concept in additive combinatorics is that of additive dimension.

Definition 2.0.5. Let G be an Abelian group.

- A set $D \subseteq G$ is *dissociated* if whenever

  $$\sum_{d \in D} \varepsilon_d d = 0$$

  for some $\varepsilon_d \in \{-1, 0, 1\}$, then all $\varepsilon_d = 0$. Equivalently, D is dissociated if the set of subset sums $\left\{\sum_{d \in S} d : S \subset D\right\}$ consists of $2^{|D|}$ distinct elements.

- The *additive dimension*, denoted by $\dim(A)$, is the size of the largest dissociated subset of A.

The trivial bounds are $\log_3 |A| \leqslant \dim(A) \leqslant |A|$, and again one should view the property of having low additive dimension as an indication of additive structure. The notion of additive dimension captures a different (though related) aspect of the additive structure of a set to the doubling constant or additive energy. As an example of this, it is a classical result [32] of Erdős and Turán that there are Sidon sets S of size $|S| \gg \sqrt{N}$ in the interval $\{1, 2, \dots, N\}$. Thus, essentially by definition, the doubling constant and additive energy of S are identical to those of a completely additively unstructured set such as $U := \{2^j : 1 \leqslant j \leqslant |S|\}$. However, the sumset $S+S$ is dense in the arithmetic progression $[2N]$ and hence its higher order sumsets $\underbrace{S + \dots + S}_{k}$ are highly structured, behaving very differently to $\underbrace{U + \dots + U}_{k}$. On the other hand, as $S \subset [N]$ we find that $\dim(S) \leqslant \dim([N]) \ll \log N \ll \log |S|$ is minimal up to a constant factor, and so the additive dimension strongly distinguishes S from the $|S|$-dimensional set U. In fact, our discussion below will indicate that $\dim(A)$ provides a good description of the 'scale' of the ambient interval in which an integer set A lives. We shall prove a number of powerful new results about additive dimension in this thesis. These will play a crucial role, though for rather different reasons, in our arguments in Chapters 3, 4 and 6. We defer most of the details of these results until the relevant chapters, giving only a brief overview of some properties of additive dimension here. Firstly, the use of the term 'dimension' is natural in light of the following observation.

Lemma 2.0.6. *If $A \subset G$ and $D \subset A$ is a maximal dissociated subset of size $|D| = \dim(A)$, then A is contained in the cube $\left\{\sum_{d \in D} \varepsilon_d d : \varepsilon_d \in \{-1, 0, 1\}\right\}$.*

Proof. Let $a \in A$, if $a \in D$ then a trivially lies in this additive cube. Otherwise, $D \cup \{a\}$ is a strictly larger subset of A, so not dissociated and hence we get a non-trivial relation of the form $\varepsilon_a a + \sum_{d \in D} \varepsilon_d d = 0$. As D is dissociated, $\varepsilon_a \neq 0$ and the result follows. □

One rather satisfactory result that we shall establish in Chapter 4 shows that, for any set A, the size of set $\Sigma(A) := \left\{\sum_{a \in A'} a : A' \subseteq A\right\}$ of all subset sums of A can be bounded rather precisely in terms of its additive dimension $\dim(A)$. Since all $2^{|D|}$ subset sums of a dissociated set $D \subset A$ are distinct by definition, there is a trivial lower bound $|\Sigma(A)| \geqslant 2^{\dim(A)}$. In Corollary 4.3.3, we shall prove the upper bound $|\Sigma(A)| \leqslant 2^{2\dim(A) \cdot \left(\log_2\left(\frac{|A|}{\dim(A)}\right)+2\right)}$ which matches the lower bound up to an extra factor $\log_2\left(\frac{|A|}{\dim(A)}\right)$ in the exponent (which in fact is necessary), thus showing

that A has many distinct subset sums if and only if it has large dimension. To describe the next application of additive dimension, we introduce the concept of a Freiman homomorphism.

Definition 2.0.7. Let G, G' be Abelian groups and let $A \subset G$, $A' \subset G'$. We say that a map $\phi : A \to A'$ is a *Freiman homomorphism* if whenever $a_1, a_2, a_3, a_4 \in A$ satisfy

$$a_1 + a_2 = a_3 + a_4,$$

then

$$\phi(a_1) + \phi(a_2) = \phi(a_3) + \phi(a_4).$$

We say that A and A' are *Freiman-isomorphic* if there is a bijective Freiman homomorphism $\phi : A \to A'$ so that ϕ^{-1} is also a Freiman homomorphism. In this case, we call A' a *Freiman-isomorphic copy* of A.

One can also define various closely related notions; in Chapter 3 a crucial role will be played by the stronger notion of a F_4-isomorphism which, by definition, preserves all additive relations of length at most 4, see Definition 3.2.1. It is a simple yet powerful fact that Freiman-isomorphic sets have the same additive behaviour in many ways; certainly doubling constants and additive energy are preserved under Freiman isomorphisms. In Chapters 4 and 6, the notion of Freiman isomorphism will prove to be a helpful tool through so-called rectification results. Such results show that, under certain assumptions, a subset A of an Abelian group G is Freiman-isomorphic to a set of integers. If this is the case, we say that A is *rectifiable*. The first rectification principle that we shall need states that sufficiently small subsets of $\mathbf{Z}/p\mathbf{Z}$ are always rectifiable.

Lemma 2.0.8 (Bilu-Lev-Ruzsa, [9] Theorem 3.1)**.** *Let p be prime and let $Z \subset \mathbf{Z}/p\mathbf{Z}$ have size $|Z| \leqslant \log_2 p$. Then Z is Freiman-isomorphic to a set of integers.*

The following lemma provides a short proof of such a statement based on the Dirichlet box principle, though with a slightly weaker constant. Here, we write $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ for the torus and denote the dilate of a set by $\lambda \cdot B = \{\lambda b : b \in B\}$.

Lemma 2.0.9. *If $B \subseteq \mathbf{Z}/p\mathbf{Z}$, then there is some $\lambda \in (\mathbf{Z}/p\mathbf{Z})^\times$ such that the dilate $\lambda \cdot B$ is contained in the interval $[-2p^{1-1/|B|}, 2p^{1-1/|B|}]$.*

Proof. Consider the set

$$\{(\lambda b/p)_{b\in B} : \lambda \in \mathbf{Z}/p\mathbf{Z}\} \subseteq \mathbf{T}^B$$

of p vectors in $\mathbf{T}^B$. We divide $\mathbf{T}^B$ into boxes of the form $\prod_{b\in B}[2j_b/p^{1/|B|}, 2(j_b+1)/p^{1/|B|})$ where the j_b range over the integers in $[0, p^{1/|B|}/2)$. There are $(p^{1/|B|}/2)^{|B|} < p$ boxes in total, and hence the pigeonhole principle provides distinct $\lambda_1, \lambda_2 \in \mathbf{Z}/p\mathbf{Z}$ such that the vectors $(\lambda_1 b/p)_{b\in B}$ and $(\lambda_2 b/p)_{b\in B}$ lie in the same box. This means precisely that $\|\lambda_1 b/p - \lambda_2 b/p\|_{\mathbf{T}} \leqslant 2p^{-1/|B|}$ for all $b \in B$. Take $\lambda := \lambda_1 - \lambda_2 \in (\mathbf{Z}/p\mathbf{Z})^\times$, so then $\lambda b \in [-2p^{1-1/|B|}, 2p^{1-1/|B|}]$ for all $b \in B$. □

In particular, if $B \subset \mathbf{Z}/p\mathbf{Z}$ has size $|B| < \frac{1}{10}\log p$, then we can find a dilate such that $\lambda \cdot B \subset (-p/4, p/4)$. Note that B is trivially Freiman-isomorphic to $\lambda \cdot B$ and that the sums in $\lambda \cdot B + \lambda \cdot B$ exhibit no 'wraparound'. Hence, seeing $\lambda \cdot B$ as a subset of $\mathbf{Z}$ shows that B may be rectified. We observe that for sets with small additive dimension, one can do significantly better.

Lemma 2.0.10. *If $B \subseteq \mathbf{Z}/p\mathbf{Z}$ is a subset of dimension $\dim(B) \leqslant k$, then there is some $\lambda \in (\mathbf{Z}/p\mathbf{Z})^\times$ such that $\lambda \cdot B$ is contained in the interval $[-2kp^{1-1/k}, 2kp^{1-1/k}]$.*

Proof. Let D be a maximal dissociated subset of B, so that $|D| \leqslant k$. By Lemma 2.0.9 we can find a $\lambda \neq 0$ so that $\lambda d \in [-2p^{1-1/k}, 2p^{1-1/k}]$ for all $d \in D$. Since $B \subseteq \mathrm{span}(D)$ by Lemma 2.0.6, we conclude that $\lambda \cdot B \subseteq [-2kp^{1-1/k}, 2kp^{1-1/k}]$. □

In Chapter 3, we will use this lemma to prove Proposition 3.6.2 which states that integer sets $A \subset \mathbf{Z}$ with additive dimension at most $\dim(A) \leqslant k$ have a Freiman-isomorphic copy $A' \subset [-T, T]$ where $T \ll e^{O(k\log k)}$. For reference, it is a well-known fact [43] in additive combinatorics that any set $A \subset \mathbf{Z}$ has a Freiman isomorphic copy $A' \subset [-e^{O(|A|)}, e^{O(|A|)}]$; the point being that one can obtain significantly stronger bounds for sets with small dimension. In fact, since a set $A \subset [T]$ always trivially satisfies $\dim(A) \ll \log T$ we see that the bound $T \leqslant e^{O(k\log k)}$ above is tight up to the extra factor of $\log k$ in the exponent. For this reason, it is a helpful perspective to think of $\dim(A)$, for a set of integers A, as describing the 'scale' of A as far as its additive structure is concerned. Finally, we note that Green and Ruzsa [37] proved a similar result which provides a 'dense' Freiman copy A' of A under the assumption that the set A has small doubling instead of small dimension.

2.1 Fourier analysis on the torus

This section discusses all the necessary basic results about the theory of the Fourier transform on $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ that we require in Chapter 3. Let $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ be the one-dimensional torus which we identify with the interval $[0, 1)$, and note that is a compact Abelian group with Haar probability measure given by the Lebesgue measure.

We use the notation $e(t) = e^{2\pi it}$. Throughout the first chapter, we shall write $c(x)$ for $\cos(2\pi x)$ so that $c(\cdot)$ is a well-defined (1-periodic) function on $\mathbf{T}$. For a suitably integrable function $g : \mathbf{T} \to \mathbf{C}$ we denote, for $p \in [1, \infty)$, its L^p-norm by

$$\|g\|_p := \left(\int_0^1 |g(t)|^p \, dt \right)^{1/p},$$

and its L^∞-norm $\|g\|_\infty$ is defined as the infimum over all constants M for which $|g(x)| \leqslant M$ holds almost everywhere. If $g \in L^1$, its Fourier transform is the bounded function $\hat{g} : \mathbf{Z} \to \mathbf{C}$ which is defined by $\hat{g}(n) = \int_{\mathbf{T}} g(t) e(-nt) \, dt$. Conversely, if $f : \mathbf{Z} \to \mathbf{C}$ is a function on $\mathbf{Z}$, we shall denote its Fourier transform $\hat{f} : \mathbf{T} \to \mathbf{C}$ by $\hat{f}(x) = \sum_{n \in \mathbf{Z}} f(n) e(nx)$ which is a priori simply a formal series.

For two functions $g, h \in L^2(\mathbf{T})$ we define $\langle g, h \rangle = \int_{\mathbf{T}} g(x) \overline{h(x)} \, dx$ and we shall frequently use Parseval's theorem which states that $\langle g, h \rangle = \sum_{n \in \mathbf{Z}} \hat{g}(n) \overline{\hat{h}(n)}$, in particular implying that

$$\|g\|_2 = \left(\sum_{n \in \mathbf{Z}} |\hat{g}(n)|^2 \right)^{1/2}.$$

We also define their convolution $g * h(x) = \int_{\mathbf{T}} g(x-y) h(y) \, dy$ and note that $\widehat{g * h}(n) = \hat{g}(n) \hat{h}(n)$.

In Chapter 3, we will want to study various functions $g \in L^2(\mathbf{T})$ through their Fourier expansion $\sum_{n \in \mathbf{Z}} \hat{g}(n) e(nx)$. We will only ever need that $\sum_{n \in \mathbf{Z}} \hat{g}(n) e(nx)$ converges to $g(x)$ in L^2, which is immediate from Parseval. In fact, all functions g that we consider will be step functions, and hence one could also formally justify pointwise almost everywhere convergence of all their Fourier series using the following classical result about functions of bounded total variation, see [67, Chapter II, Theorem 8.1]. A function $f : \mathbf{T} \to \mathbf{R}$ is said to be of *bounded variation* if the total variation

$$V(f) := \sup \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})|$$

is finite, where the supremum is taken over all partitions $0 = x_0 < x_1 < \cdots < x_n = 1$ of the interval $[0, 1]$.

Theorem 2.1.1 (Dirichlet–Jordan)**.** *Let $f : \mathbf{T} \to \mathbf{R}$ be a function of bounded variation. Then the Fourier series of f converges at every point $x \in \mathbf{T}$ to the average $\frac{1}{2}(f(x^+) + f(x^-))$ of the right and left limits of f at x.*

The following two standard inequalities will be used throughout. Hölder's inequality states that for $1 \leqslant p, q \leqslant \infty$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$ and functions $f \in L^p(\mathbf{T})$, $g \in L^q(\mathbf{T})$, we can bound

$$\left| \int_{\mathbf{T}} f(x) g(x)\, dx \right| \leqslant \|f\|_p \|g\|_q.$$

Young's convolution inequality states that for $1 \leqslant p, q, r \leqslant \infty$ satisfying

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$$

and functions $f \in L^p(\mathbf{T})$, $g \in L^q(\mathbf{T})$, the convolution $f * g$ lies in $L^r(\mathbf{T})$, and

$$\|f * g\|_r \leqslant \|f\|_p \|g\|_q.$$

Finally, we shall make use of some basic properties of *de la Vallée-Poussin kernels*. The degree $2T - 1$ de la Vallée-Poussin kernel is the trigonometric polynomial given by

$$V_T(x) := \sum_{n=-2T}^{-T-1} \left(1 - \frac{|T+n|}{T}\right) e(nx) + \sum_{n=-T}^{T} e(nx) + \sum_{n=T+1}^{2T} \left(1 - \frac{|T-n|}{T}\right) e(nx).$$

It is immediate from this definition that the Fourier coefficients of V_T are given by $\hat{V}_T(n) = 1$ for $|n| < T$ while $\hat{V}_T(n) = 0$ for $|n| \geqslant 2T$. The other important property that we will rely on is that the L^1-norms of the de la Vallée-Poussin kernels are uniformly bounded, see [52, Chapter VIII, (205)].

Lemma 2.1.2. *Let $T \in \mathbf{N}$, then the L^1-norm of V_T is bounded by $\|V_T\|_1 \leqslant 3$.*

The proof of this is rather elementary; one may note that $V_T(x) = 2F_{2T}(x) - F_T(x)$ can be expressed as the difference of two Fejér kernels, which are defined by

$$F_m(x) = \frac{1}{m} \left| \sum_{j=0}^{m-1} e(jx) \right|^2 = \sum_{j=-m}^{m} \left(1 - \frac{|j|}{m}\right) e(jx).$$

Since Fejér kernels are pointwise non-negative, their L^1 norms can be calculated explicitly $\|F_m\|_1 = \int_{\mathbf{T}} |F_m(x)|\, dx = \int_{\mathbf{T}} F_m(x)\, dx = \hat{F}_m(0) = 1$, implying that $\|V_T\|_1 \leqslant 3$. These properties of V_T show that, given any function g, the convolution $g * V_T(x)$ has Fourier coefficients given by

$$(g * V_T)^\wedge(n) = \hat{g}(n)\hat{V}_T(n) = \begin{cases} \hat{g}(n) & \text{if } |n| < T, \\ 0 & \text{if } |n| \geqslant 2T, \end{cases}$$

so that $g * V_T$ is a particular truncation of the Fourier series of g. Moreover, this truncation is well-behaved in the sense that, if $g \in L^p(\mathbf{T})$, then $\|g * V_T\|_p \leqslant \|g\|_p \|V_T\|_1 \leqslant 3\|g\|_p$ by Young's convolution inequality.

2.2 Notation

We use the asymptotic notation $f = O(g)$ or $f \ll g$ if there is an absolute constant C such that $|f(x)| \leqslant Cg(x)$ for all large x, and we write $f = o(g)$ if $f(x)/g(x) \to 0$ as $x \to \infty$. Further, if f is non-negative and $f = O(g)$, then we also write $g = \Omega(f)$. Finally, we write $f \asymp g$ when $f \ll g$ and $g \ll f$.

Chapter 3

Large sum-free subsets of sets of integers

3.1 Introduction

This chapter is based on the paper [7]. A set B is said to be sum-free if it contains no three elements x, y, z with $x + y = z$. The study of sum-free sets can be traced back to Schur [61], who introduced the concept in proving that Fermat's last theorem does not hold in $\mathbf{Z}/p\mathbf{Z}$. There is a substantial amount of literature on sum-free sets, most of which we will not discuss here; the interested reader may consult the survey [65] of Tao and Vu. One of the most central open questions in this area is that of determining the quantity $S(N)$ which is defined to be the largest integer S such that any set of N positive integers contains a sum-free subset of size S. Let us write $S(A)$ for the size of the largest sum-free subset of A, so that

$$S(N) := \min_{A \subset \mathbf{N} : |A| = N} S(A). \tag{3.1}$$

The following classical argument of Erdős [19] proves that $S(A) \geqslant |A|/3$ for any $A \subset \mathbf{Z} \setminus \{0\}$. Let $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ and note that the interval $(1/3, 2/3) \subset \mathbf{T}$ is sum-free. Hence, for any $x \in \mathbf{T}$ the set $A_x = \{a \in A : ax (\mathrm{mod}\, 1) \in (1/3, 2/3)\}$ is sum-free. One can conclude by observing that $\mathbb{E}_{x \in \mathbf{T}} |A_x| = \sum_{a \in A} \mathbb{P}(ax (\mathrm{mod}\, 1) \in (1/3, 2/3)) = |A|/3$ which shows that one of the sum-free sets A_x must have size at least $|A|/3$.

Despite its simplicity, only minor improvements over this lower bound have been obtained. Alon and Kleitman [2] observed that Erdős's argument may in fact be improved to give $S(N) \geqslant (N + 1)/3$. The best bound $S(N) \geqslant (N + 2)/3$ before this work was established in a celebrated paper of Bourgain [12] using an elaborate Fourier analytic approach. Recent work of Shakan [62] provides an alternative proof of Bourgain's bound using a similar method. It has been a long-standing open problem

to find a more substantial improvement over Erdős's lower bound and this question appears in the work of various authors such as [2,12,17,19,38,40,41,47,62,65,66]. The main problem is to prove the following widely believed estimate for $S(N)$ which asserts that one can improve the Erdős-Alon/Kleitman-Bourgain bounds by an arbitrarily large constant. This problem is also listed as Problem 1 on Green's list [36] of 100 open problems.

Problem 3.1.1. *Is there a function $\omega(N) \to \infty$ such that $S(N) \geqslant \frac{N}{3} + \omega(N)$?*

Our aim here is to establish the following theorem confirming this.

Theorem 3.1.2. *There exists some constant $c > 0$ such that for all finite sets $A \subset \mathbf{Z}$ we have $S(A) \geqslant \frac{|A|}{3} + c \log\log |A|$. In particular, $S(N) \geqslant \frac{N}{3} + c \log\log N$.*

We further prove a strong structural result for sets where $S(A) \leqslant N/3 + C$ which provides non-trivial information about the global structure of A, even for a much larger value of C than $\log\log N$.

Theorem 3.1.3 (99% Structure Theorem)**.** *Let $A \subset \mathbf{Z} \setminus \{0\}$ be a set of size N with $S(A) \leqslant N/3 + C$. Then A has a Freiman-isomorphic copy inside $[-N^{C^{O(1)}}, N^{C^{O(1)}}]$.*[1] *Moreover, for any parameter $K > 1$, we can find a partition*

$$A = \left(\bigcup_{j=1}^{s} A_j \right) \cup B \tag{3.2}$$

which has the following properties.

(i) For each $j \in [s]$, the set A_j has size $|A_j| \gg (KC)^{-O(1)} N$ and small doubling $|A_j - A_j| \leqslant (CK)^{O(1)} |A_j|$. In particular, A_j is contained in a generalised arithmetic progression P_j of dimension $d_j \ll (KC)^{O(1)}$ and of size $|P_j| \ll e^{(KC)^{O(1)}} |A_j|$.

(ii) The set B is small: $|B| \ll (KC)^{-10} N$.

Though not the topic of study here, we briefly discuss progress on the upper bounds for $S(N)$. This bound has been improved many times; we summarise this in the following table.

[1] By a Freiman isomorphism, we mean an F_4-isomorphism as in Definition 3.2.1.

Author	Value of c s.t. $S(N) \leqslant cN + o(N)$
Hinton [19]	7/15
Klarner [19]	3/7
Alon, Kleitman [2]	12/29
Malouf, Furedi [38, 49]	2/5
Lewko [47]	11/28
Alon [1]	$11/28 - \varepsilon$
Eberhard, Green and Manners [17]	1/3

The first five of these bounds were obtained using increasingly better explicit constructions of sets A for which $S(A) \leqslant c|A|$, where c is the corresponding constant in the bounds above. The breakthrough paper of Eberhard, Green and Manners [17] establishes an upper bound $S(N) \leqslant \frac{N}{3} + o(N)$ which matches Erdős's lower bound up to a function which is $o(N)$. Their proof employs an elegant argument based on the arithmetic regularity lemma which leads to a more-or-less ineffective bound for $o(N)$; determining a reasonable upper bound remains an interesting problem. Contrary to the arguments that came before, the sets A with $S(A) \leqslant \frac{|A|}{3} + o(|A|)$ whose existence is proved by Eberhard, Green and Manners are not explicitly constructible, but Eberhard [16] later provided explicit examples of such sets.

3.2 Setup and overview

We begin by discussing a well-known strategy for obtaining lower bounds for $S(N)$, dating back (in its simplest form) to the work of Erdős [19]. Let $A \subset \mathbf{Z}$ have size N. By simply removing 0 if it lies in A, it is sufficient to establish Theorem 3.1.2 for sets $A \subset \mathbf{Z} \setminus \{0\}$ so we assume throughout the rest of this argument that $0 \notin A$. We write $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ for the one-dimensional torus and recall that the interval $(1/3, 2/3) \subset \mathbf{T}$ is sum-free. Hence, for any $x \in \mathbf{T}$ the set $\{a \in A : ax (\mathrm{mod}\, 1) \in (1/3, 2/3)\}$ is sum-free and we deduce the following important basic estimate

$$S(A) \geqslant \max_{x \in \mathbf{T}} \sum_{a \in A} \phi(ax) = \frac{N}{3} + \max_{x} \sum_{a \in A} (\phi - 1/3)(ax),$$

where ϕ is the characteristic function of the interval $(1/3, 2/3)$. Noting that

$$\max_{x} \sum_{a \in A} (\phi - 1/3)(ax) \geqslant \int_{\mathbf{T}} \sum_{a \in A} (\phi - 1/3)(ax)\, dx = |A| \int_0^1 \phi(x)\, dx - |A|/3 = 0$$

recovers Erdős's bound and Bourgain obtained his improvement by showing that $\max_x \sum_{a \in A} (\phi - 1/3)(ax) > 1/3$ using the Fourier-theoretic properties of $\phi - 1/3$.

Before stating our main theorem, we introduce the following strong notion of Freiman isomorphism.

Definition 3.2.1. *Let G, G' be Abelian groups and let $A \subset G$, $A' \subset G'$. We say that a map $\phi : A \to A'$ is an F_4-homomorphism if whenever $a_i \in A$ and $\varepsilon_i \in \{-1, 0, 1\}$ for $i \in [4]$ satisfy*

$$\sum_{i=1}^{4} \varepsilon_i a_i = 0,$$

then

$$\sum_{i=1}^{4} \varepsilon_i \phi(a_i) = 0.$$

We say that A and A' are F_4-isomorphic if there is a bijective F_4-homomorphism $\phi : A \to A'$ so that ϕ^{-1} is also an F_4-homomorphism.

It is obvious that $S(A) = S(A')$ for F_4-isomorphic sets. Our main result is to prove the following theorem which clearly implies Theorem 3.1.2.

Theorem 3.2.2. *Let $A \subset \mathbf{Z} \setminus \{0\}$. Then there exists a set $B \subset \mathbf{Z} \setminus \{0\}$ which is F_4-isomorphic to A and which satisfies*

$$\max_x \sum_{b \in B} \left(\phi - \frac{1}{3}\right)(bx) \gg \log\log|B|.$$

Before beginning the proof, we attempt to give a broad outline of our approach and indicate where it differs from previous work due to Bourgain [12]. Bourgain proved that $S(N) \geqslant (N+2)/3$, but perhaps the most interesting part of his paper is his progress for $(3,1)$-sum-free sets, which he defines to be sets containing no x, y, z, w with $x + y + z = w$. In analogy to $S(N)$, the quantity $S_{(3,1)}(N)$ is then defined to be the largest number so that any set of N positive integers contains a $(3,1)$-sum-free subset of size $S_{(3,1)}(N)$. Noting that the interval $(1/8, 3/8) \subset \mathbf{T}$ is $(3,1)$-sum-free allows one to use Erdős's argument and obtain $S_{(3,1)}(N) \geqslant N/4$. Bourgain proved the substantially better bound $S_{(3,1)}(N) \geqslant N/4 + (\log N)^{1-o(1)}$. However, his argument rather crucially relies on the existence of another maximal $(3,1)$-sum-free subinterval of $\mathbf{T}$, namely $(5/8, 7/8) = -(1/8, 3/8)$, which allows him to combine the Fourier series of $1_{(1/8),(3/8)}$ and $1_{(5/8,7/8)}$ to get what is essentially a one-sided Fourier series (i.e. its Fourier spectrum consists of non-negative integers only). As Bourgain points out, such an approach cannot be applied to bound $S(N)$ since it is easy to see that $(1/3, 2/3)$ is the unique sum-free interval in $\mathbf{T}$ with measure $1/3$. Recently, Jing and Wu [40, 41] extended Bourgain's method and showed that $S_{(k,\ell)}(N) \geqslant N/(k+\ell) + (\log N)^{1-o(1)}$ for various other pairs (k, ℓ),[2] perhaps most interestingly $(2, 4)$ and

[2]Here, $S_{(k,l)}(N)$ is the largest number S so that any set of N positive integers contains a subset of size S with no solutions to $x_1 + \cdots + x_k = x'_1 + \cdots + x'_\ell$.

$(1,5)$, but their method still relies on the existence of asymmetric maximal (k,ℓ)-sum-free subsets of the torus. We also mention that Eberhard [16] has shown that $S_{(k,1)}(N) \leqslant N/(k+1) + o(N)$ and that Jing and Wu proved the analogous bound for all (k,ℓ).

Overview. In Section 4, we begin by recalling some of Bourgain's approach which considers the Fourier expansion

$$F_A(x) = \sum_{a\in A}(\phi - 1/3)(ax) = \sum_{a\in A}\sum_{n\geqslant 1}\frac{\chi(n)}{n}\cos 2\pi nax,$$

where χ is a character mod 3. Bourgain shows in particular that in order to establish Theorem 3.1.2, it suffices to prove that $\|F_A\|_1 \gg \log\log N$. Bourgain also observed that one can 'sift' out the contribution of $n > 1$ and that a bound $\|F_A\|_1 \gg C$ would follow if $\|\hat{1}_A\|_1 \gg C\log N$.

The first step in our argument is to establish two inverse theorems describing structural properties of sets A for which $\|\hat{1}_A\|_1 \ll C\log N$ and C is 'small'. The main result in Section 5 shows that such sets A have small additive dimension $\dim(A)$. In Section 6, we exploit such additive information about A to find a 'dense' Freiman-isomorphic copy B of A and we use this to obtain strong bounds on the Fourier coefficients and L^2 norm of (certain modifications of) F_B.

Section 7 is concerned with studying the distribution of A in residue classes modulo powers of 'small' primes $p \leqslant (\log N)^{1/2}$. First, we use a combinatorial argument to show that a large part of A must lie in a single such residue class when $\dim(A)$ is small. Secondly, we prove in Proposition 3.7.9 that either $\|F_A\|_1 \gg \log\log N$ or else that the distribution of A in these residue classes is highly structured. The final step in the proof of Theorem 3.1.2 is accomplished in Section 8 and consists of exploiting this non-Archimedean structure in A to construct an explicit test function Φ which witnesses that $\|F_A\|_1$ is large (i.e. we construct Φ s.t. $|\Phi| \leqslant 1$ and $\int_0^1 F_A(x)\Phi(x)\,dx \gg \log\log N$).

The proof of the structural result Theorem 3.1.3 is given in Section 9 and proceeds by bootstrapping our application of the inverse results from Section 5.

3.3 Prerequisites

The following seminal result is known as the 'Littlewood L^1 conjecture', which was proved by McGehee, Pigno and Smith [50] and independently by Konyagin [42].

Theorem 3.3.1 (Littlewood's L^1 conjecture)**.** *Let $a_1, a_2, \dots, a_k$ be complex numbers and $n_1 < n_2 < \cdots < n_k$ be integers. Then*

$$\left\| \sum_{j=1}^{k} a_j e(n_j t) \right\|_1 \gg \sum_{j=1}^{k} \frac{|a_j|}{j}.$$

Corollary 3.3.2. *Let $B \subset \mathbf{Z}$ be finite, then $\|\hat{1}_B\|_1 \gg \log |B|$.*

We shall not, in fact, use either of these results. Rather, we will employ various non-trivial modifications of some ideas appearing in the McGehee-Pigno-Smith proof of the Littlewood L^1 conjecture. Their proof proceeds, like much of the earlier progress on the L^1 conjecture, by constructing a function Φ such that $|\Phi| \ll 1$ and $\langle \hat{1}_B, \Phi \rangle$ is 'large'. The most basic part of their construction of such a test function is something that we will use repeatedly, so we state the following general lemma about constructing these. In fact, the construction in the following lemma already differs from that of McGehee-Pigno-Smith; the reader may notice that we do not require the one-sidedness of the Fourier spectrum (a condition that is crucial for their original construction).

Lemma 3.3.3 (M-P-S basic construction of test functions)**.** *Let $f : \mathbf{Z} \to \mathbf{C}$ be a function with finite support* $\operatorname{supp}(f)$*. Let $X_1, X_2, \dots, X_J \subset \mathbf{Z}$ be finite and define the functions*

$$g_i : \mathbf{Z} \to \mathbf{C} : g_i(n) = \begin{cases} |X_i|^{-1} \frac{f(n)}{|f(n)|} & \text{if } n \in X_i \cap \operatorname{supp} f, \\ 0 & \text{otherwise.} \end{cases}$$

We further define $Q_i(x) = e^{-|\hat{g}_i(x)|}$ and

$$\Phi_j = \hat{g}_j + \hat{g}_{j-1} Q_j + \hat{g}_{j-2} Q_{j-1} Q_j \cdots + \hat{g}_1 Q_2 \dots Q_j,$$

and note that these are functions defined on $\mathbf{T}$. Then the g_i and Φ_j satisfy the following properties

$$\begin{aligned} \|\Phi_j\|_\infty &\leqslant 10, \\ \|\hat{g}_i\|_\infty &\leqslant 1, \\ \|\hat{g}_i\|_2 &\leqslant |X_i|^{-1/2}, \\ \langle \hat{f}, \hat{g}_i \rangle &= |X_i|^{-1} \sum_{n \in X_i \cap \operatorname{supp} f} |f(n)|. \end{aligned} \tag{3.3}$$

Moreover, the functions Q_i satisfy the bounds $|Q_i| \leqslant 1$ and $|1 - Q_i| \leqslant |\hat{g}_i|$.

Proof. That $\|\hat{g}_i\|_\infty \leqslant 1$ follows immediately from the fact that $|g_i(n)| = |X_i|^{-1}$ if $n \in X_i \cap \operatorname{supp} f$ and $|g_i(n)| = 0$ otherwise, and Parseval shows that $\|\hat{g}_i\|_2 \leqslant |X_i|^{-1/2}$. Since $\Phi_1 = \hat{g}_1$, the inequality $\|\Phi_1\|_\infty \leqslant 10$ holds. Assuming now that $\|\Phi_j\|_\infty \leqslant 10$, then one can observe that $\Phi_{j+1} = \hat{g}_{j+1} + \Phi_j Q_{j+1}$ so that

$$\begin{aligned} |\Phi_{j+1}(x)| &\leqslant |\hat{g}_{j+1}(x)| + 10e^{-|\hat{g}_{j+1}(x)|} \\ &\leqslant 10, \end{aligned}$$

where the final line follows from the basic fact that $y + 10e^{-y} \leqslant 10$ whenever $y \in [0, 1]$. Parseval's theorem shows that

$$\langle \hat{f}, \hat{g}_i \rangle = \sum_{n \in \mathbf{Z}} f(n)\overline{g_i}(n) = |X_i|^{-1} \sum_{n \in X_i \cap \operatorname{supp} f} |f(n)|.$$

Finally, it is trivial that $Q_i = e^{-|\hat{g}_i|}$ is 1-bounded, and the inequality $|1 - Q_i| \leqslant |\hat{g}_i|$ follows from the fact that $|e^{-x} - 1| \leqslant x$ for $x \geqslant 0$. □

We shall also make use of the following important inequality of Rudin [57]. To state Rudin's theorem, we recall Definition 2.0.5 which states that a set D is *dissociated* if whenever

$$\sum_{d \in D} \varepsilon_d d = 0$$

for some $\varepsilon_d \in \{-1, 0, 1\}$, then all $\varepsilon_d = 0$.

Theorem 3.3.4 (Rudin's inequality)**.** *Let $D \subset \mathbf{Z}$ be a dissociated set and let $f : \mathbf{Z} \to \mathbf{C}$ have $\operatorname{supp} f \subset D$. Then for any $p \in [2, \infty)$ the following bound holds:*

$$\|\hat{f}\|_p \leqslant 10\sqrt{p}\|\hat{f}\|_2.$$

3.4 The Fourier series of $\sum_{a \in A} (\phi - 1/3)(ax)$

The purpose of this section is to recall some of the Fourier-theoretic setup of Bourgain's paper [12]. We shall assume throughout that $A \subset \mathbf{Z} \setminus \{0\}$ has size N. Recall the important basic estimate

$$S(A) \geqslant \frac{N}{3} + \max_x \sum_{a \in A} \left(\phi - \frac{1}{3}\right)(ax), \tag{3.4}$$

where ϕ is the characteristic function of $(1/3, 2/3) \subset \mathbf{T}$. The function $\phi - 1/3 : \mathbf{T} \to \mathbf{R}$ has the following Fourier expansion

$$\phi(x) - 1/3 = \frac{-\sqrt{3}}{\pi} \sum_{n=1}^{\infty} \frac{\chi(n)}{n} c(nx)$$

where $c(x) = \cos(2\pi x)$ and χ is the multiplicative character given by

$$\chi(n) = \begin{cases} 0 & \text{if } n \equiv 0 (\mathrm{mod}\, 3) \\ 1 & \text{if } n \equiv 1 (\mathrm{mod}\, 3) \\ -1 & \text{if } n \equiv 2 (\mathrm{mod}\, 3). \end{cases}$$

Let μ denote the Möbius function, and recall the following fundamental property

$$\sum_{k|n} \mu(k) = 1_{n=1}.$$

For a parameter Q, we say that an integer n is Q-rough if all its prime factors are greater than Q, and we denote the set of Q-rough numbers by

$$\mathcal{R}_Q = \{n \in \mathbf{N} : (n, p) = 1 \text{ for all primes } p \leqslant Q\}.$$

The function appearing in the right hand side of (3.4) is fundamental to our approach, and it will be convenient to introduce notation for the following rescaled version

$$\begin{aligned} F(x) = F_A(x) &:= \frac{-\pi}{\sqrt{3}} \sum_{a \in A} (\phi - 1/3)(ax) \\ &= \sum_{a \in A} \sum_{n \geqslant 1} \frac{\chi(n)}{n} c(nax). \end{aligned}$$

We record some of its important basic properties. For a set $A \subset \mathbf{Z}$, we will write $c_A(x) = \sum_{a \in A} c(ax)$ for the cosine polynomial with frequencies in A.

Proposition 3.4.1 (Bourgain [12]). *Let $A \subset \mathbf{Z} \setminus \{0\}$. Then the following holds:*

(i) there exists some absolute constant $c > 0$ such that

$$S(A) \geqslant \frac{|A|}{3} + c \max_{x \in \mathbf{T}} (-F(x));$$

(ii) we have the one-sided estimate $\max_x(-F(x)) \geqslant \frac{1}{2}\|F\|_1$;

(iii) for any parameter Q we have

$$\begin{aligned} \sum_{k | \prod_{p \leqslant Q} p} \frac{\mu(k)\chi(k)}{k} F(kx) &= \sum_{a \in A} \sum_{n \in \mathcal{R}_Q} \frac{\chi(n)}{n} c(nax) \\ &= \underbrace{\sum_{a \in A} c(ax)}_{c_A(x)} + \underbrace{\sum_{\substack{1 < n \in \mathcal{R}_Q \\ a \in A}} \frac{\chi(n)}{n} c(nax)}_{R_Q(x)}, \end{aligned}$$

where the variable p runs over primes only;

(iv) for any Q we have

$$\|F\|_1 \gg \|c_A + R_Q\|_1 / \log Q.$$

Proof. Item (i) follows immediately from (3.4) with $c = \sqrt{3}/\pi$. For item (ii), we note that $\int_{\mathbf{T}} (\phi(x) - 1/3)\, dx = 0$ so that $\int F = 0$ and hence

$$\int_0^1 |F(x)|\, dx = \int (|F(x)| - F(x))\, dx = \int 2\max(-F(x), 0)\, dx \leqslant 2 \max_x(-F(x)).$$

To prove the (iii), we use the multiplicative nature of χ to calculate

$$\begin{aligned}
\sum_{k | \prod_{p \leqslant Q} p} \frac{\mu(k)\chi(k)}{k} \frac{-\pi}{\sqrt{3}} (\phi - 1/3)(kx) &= \sum_{k | \prod_{p \leqslant Q} p} \frac{\mu(k)\chi(k)}{k} \sum_{n \geqslant 1} \frac{\chi(n)}{n} c(nkx) \\
&= \sum_{m \geqslant 1} \frac{\chi(m)}{m} c(mx) \sum_{k | \gcd(m, \prod_{p \leqslant Q} p)} \mu(k) \\
&= \sum_{m \geqslant 1} \frac{\chi(m)}{m} c(mx) 1_{\{m \text{ is } Q\text{-rough}\}} \\
&= c(x) + \sum_{1 < m \text{ is } Q\text{-rough}} \frac{\chi(m)}{m} c(mx).
\end{aligned}$$

Replacing x by ax and summing over $a \in A$ yields (iii). For the final item, we use the bound

$$\begin{aligned}
\|c_A + R_Q\|_1 &= \left\| \sum_{k | \prod_{p \leqslant Q} p} \frac{\mu(k)\chi(k)}{k} F(kx) \right\|_1 \\
&\leqslant \sum_{k | \prod_{p \leqslant Q} p} \frac{1}{k} \|F\|_1 \\
&= \|F\|_1 \prod_{p \leqslant Q} (1 + 1/p) \\
&\ll \|F\|_1 (\log Q),
\end{aligned}$$

where we used the bound $1 + p^{-1} \leqslant e^{p^{-1}}$ and Mertens' estimate [51, Theorem 2.7]: $\sum_{p \leqslant Q} \frac{1}{p} \leqslant \log\log Q + O(1)$. □

Bourgain noted that an important consequence is the following bound for $S(A)$.

Proposition 3.4.2. *Let $A \subset \mathbf{Z} \setminus \{0\}$ be a set of size N. Then*

$$S(A) \geqslant \frac{N}{3} + c \frac{\|c_A\|_1}{\log N} \tag{3.5}$$

where $c > 0$ is some absolute constant. In particular, if $S(A) \leqslant N/3 + C$, then $\|\hat{1}_A + \hat{1}_{-A}\|_1 \ll C \log N$.

Proof. By combining (i),(ii) and (iv) in Proposition 3.4.1 we get $S(A) - N/3 \gg (\log Q)^{-1}\|c_A + R_Q\|_1$, so it suffices to show that $\|c_A + R_Q\|_1 \gg \|c_A\|_1$ for $Q = 100N^2$. Observe that by monotonicity of L^p norms and Parseval's identity

$$\|R_Q\|_1 \leqslant \|R_Q\|_2 \leqslant |A| \left\| \sum_{1<n\in\mathcal{R}_Q} \frac{\chi(n)}{n} c(nx) \right\|_2 \leqslant |A| \left(\sum_{n>Q} n^{-2} \right)^{1/2} \leqslant \frac{|A|}{Q^{1/2}} = 1/10.$$

Hence, $\|c_A + R_Q\|_1 \geqslant \|c_A\|_1 - 1/10$ and from the trivial lower bound $\|c_A\|_1 \geqslant 1/2$ (which can for example be proved by noting that $\int_0^1 |c_A(x)| \geqslant \int_0^1 c_A(x)e(-ax) = 1/2$ for $a \in A$) we see that $\|c_A + R_Q\|_1 \gg \|c_A\|_1$. □

Propositions 3.4.1 and 3.4.2 can be found in Bourgain's paper, but this is the limit of what his approach yields with regards to Problem 3.1.1. Now that we have introduced the function F_A and its useful relation to $S(A)$ which forms the starting point for our approach, we state the following more detailed theorem.

Theorem 3.4.3. *Let $A \subset \mathbf{Z} \setminus \{0\}$ have size N. Then there exists a set $B \subset \mathbf{Z} \setminus \{0\}$ which is F_4-isomorphic to A and satisfies $\|F_B\|_1 \gg \log\log N$ where $F_B(x) = \sum_{b\in B} \sum_{n\geqslant 1} \frac{\chi(n)}{n} \cos 2\pi nbx$.*

For clarity of exposition however, we have written our arguments in this chapter to simply give bounds for $S(B)$, where B is F_4-isomorphic to A, rather than for $\|F_B\|_1$, but one can check that it is in fact the result above that our proof gives. Note also that this theorem immediately implies Theorem 3.1.2. To see this, recall that $S(A) = S(B) \geqslant N/3 + c\|F_B\|_1$ by (i) and (ii) in Proposition 3.4.1.

3.5 The structure of sets with small L^1-norm

The goal of this section is to prove two structural theorems for sets whose Fourier transform has small L^1-norm. The first shows that if $\hat{1}_B$ has small L^1 norm for some $B \subset \mathbf{Z}$, then its additive dimension $\dim(B)$ is small. The second shows, again under the assumption that $\|\hat{1}_B\|_1$ is small, that every large subset of B has large additive energy. In fact, we shall need to prove a more general result which establishes these conclusions under a weaker condition that $\|\hat{f}\|_1$ is small for a function $f : \mathbf{Z} \to \mathbf{C}$ which satisfies $f(b) \gg 1$ for all $b \in B$. For comparison, note that we always have

$$\log|B| \ll \|\hat{1}_B\|_1 \leqslant |B|^{1/2},$$

where the lower bound follows from the Littlewood L^1 conjecture and the upper bound from the simple estimate $\|\hat{1}_B\|_1 \leqslant \|\hat{1}_B\|_2 = |B|^{1/2}$ by Parseval. Both the upper and lower bounds are tight up to a constant factor in general.[3]

Before we state the first theorem, the reader may want to recall the Definition 2.0.5 of dissociativity and additive dimension.

Theorem 3.5.1. *Let $f : \mathbf{Z} \to \mathbf{C}$ be a function with $\|\hat{f}\|_\infty = N$ and let $D \subset \mathbf{Z}$ be any dissociated set. If $\min_{n\in D} |f(n)| \geqslant 1$, then*

$$\|\hat{f}\|_1 \gg \left(\frac{|D|}{\log N}\right)^{1/2}. \tag{3.6}$$

Corollary 3.5.2. *Let $B \subset \mathbf{Z}$ be a finite set of integers. Then*

$$\dim(B) \ll \|\hat{1}_B\|_1^2 \log |B|. \tag{3.7}$$

Remark. *The bound in (3.6) is best possible up to a constant factor for general f. As an example, one can take a Fejér kernel $\hat{f} = F_N(x) = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N}\right) e(nx)$ which has $\|F_N\|_1 = 1$, while $f(n) \geqslant 1/2$ for all n in the dissociated set $\{2^j : j < \log_2 N - 1\}$.*

Proof of Theorem 3.5.1. Let $D \subset \mathbf{Z}$ be dissociated. We define the function $g : D \to \mathbf{C}$ by $g(d) = |D|^{-1} \frac{f(d)}{|f(d)|}$. We also define the corresponding 'correction' function $Q(t) = \exp(-|\hat{g}(t)|)$ and note that $|Q(t)| \leqslant 1$ and $|1 - Q(t)| \leqslant |\hat{g}(t)|$, by Lemma 3.3.3. We further know from Lemma 3.3.3 that if Φ_j is defined as follows:

$$\Phi_j(t) = \hat{g}(t)(1 + Q(t) + \cdots + Q^{j-1}(t)),$$

then $\|\Phi_j\|_\infty \leqslant 10$ for all j.

Let us fix an integer J. By a telescoping identity, we see that

$$\Phi_J(t) = J\hat{g}(t) - \sum_{j=1}^{J-1} \sum_{k=1}^{j} \hat{g}(t)(1 - Q(t))Q(t)^{k-1}.$$

Then as $\|\Phi_J\|_\infty \leqslant 10$, we have

$$\begin{aligned} \|\hat{f}\|_1 &\gg \langle \hat{f}, \Phi_J \rangle \\ &= J\langle \hat{f}, \hat{g} \rangle - E \\ &\geqslant J \min_{d\in D} |f(d)| - |E|, \end{aligned} \tag{3.8}$$

[3] In fact, it is a well-known problem to determine either of these constants.

where we used the last equation in (3.3) and we defined

$$E = \sum_{1 \leqslant k \leqslant j \leqslant J-1} \langle \hat{f}, \hat{g}(1-Q)Q^{k-1} \rangle.$$

To conclude, we bound E and optimise the choice of J. Note that by the properties of $\hat{g}$ and Q we get

$$\begin{aligned} |E| &\leqslant \sum_{1 \leqslant k \leqslant j \leqslant J-1} \langle |\hat{f}|, |\hat{g}|^2 \rangle \\ &\leqslant J^2 \|\hat{f}\|_p \|\hat{g}\|_{2q}^2, \end{aligned} \tag{3.9}$$

where we used Hölder's inequality with exponent pair (p, q) such that $\frac{1}{p} + \frac{1}{q} = 1$. To estimate the first L^p-norm in terms of the L^1-norm of $\hat{f}$, we take $p = 1 + 1/\log N$ so that $|\hat{f}|^p \ll |\hat{f}|$ because of our assumption that $\|\hat{f}\|_\infty = N$, and hence

$$\|\hat{f}\|_p \ll \|\hat{f}\|_1^{1/p} \leqslant \|\hat{f}\|_1$$

using also that $\|\hat{f}\|_1 \geqslant \langle \hat{f}, e(d\cdot) \rangle = 1$ for $d \in D$. To bound $\|\hat{g}\|_{2q}$, we use the fact that $\operatorname{supp} g \subseteq D$ is dissociated so that by Rudin's inequality in Theorem 3.3.4 we obtain

$$\|\hat{g}\|_{2q} \ll q^{1/2} \|\hat{g}\|_2 \ll q^{1/2} |D|^{-1/2},$$

where we used Parseval to evaluate $\|\hat{g}\|_2$. In total, since $q \ll \log N$ we can bound (3.9) by

$$|E| \ll J^2 \|\hat{f}\|_1 \frac{\log N}{|D|},$$

and we can substitute this back in (3.8) to deduce that

$$\|\hat{f}\|_1 + J^2 \|\hat{f}\|_1 \frac{\log N}{|D|} \gg J \min_{n \in D} |f(n)| \geqslant J.$$

Taking $J = \lfloor (|D|/\log N)^{1/2} \rfloor$ shows that $\|\hat{f}\|_1 \gg (|D|/\log N)^{1/2}$ as desired.

□

Remark. The author would like to thank Thomas Bloom for pointing out that Zygmund [67, Chapter XII, (7.6)] proves that if (n_j) is a lacunary (i.e. $n_{j+1}/n_j > c > 1$) and $\hat{f}(\log^+ |\hat{f}|)^{1/2}$ is integrable, then $\sum_j |f(n_j)|^2$ converges. Interestingly, Pichorides noted that this proof works when $\{n_j\}$ is dissociated and that a quantitative version of his argument yields the bound in Theorem 3.5.1. Pichorides [54] in fact used this to establish what was at the time the best bound towards the Littlewood L^1

conjecture, and Konyagin [42] makes use of similar ideas (but about the number of distinct 2-adic valuations of the n_j rather than dissociativity). We include the short proof above since it is different and in particular constructs an explicit test function witnessing that $\|\hat{f}\|_1$ is large.

In Proposition 3.4.2, we showed that if A is a set of N integers such that $S(A) \leqslant N/3 + C$, then $\|\hat{1}_A + \hat{1}_{-A}\| \ll C \log N$. Applying Theorem 3.5.1 with $f = 1_A + 1_{-A}$ therefore shows the following.

Corollary 3.5.3. *Let $A \subset \mathbf{Z} \setminus \{0\}$ have size N and let $S(A) \leqslant N/3 + C$. Then $\dim(A) \ll C^2(\log N)^3$.*

The fact that A has small dimension implies that A has strong additive structure as we will prove in the next section.

We first discuss the next theorem, which roughly speaking states that if $\|\hat{f}\|_1$ is small, then every large set A with $\min_{a\in A} |f(a)| \gg 1$ has large additive energy. For sets $B, B' \subset \mathbf{Z}$, we define the joint additive energy

$$E(B, B') = \#\{(b_1, b_2 \in B, b_1', b_2' \in B' : b_1 - b_2 = b_1' - b_2'\}$$

and note that the joint energy of a set with itself $E(B, B) = E(B)$ matches the notion of additive energy from Chapter 2. We also point out that this theorem is used to obtain the structure Theorem 3.1.3, but that it is not required for Theorem 3.1.2.

Theorem 3.5.4. *Let $f : \mathbf{Z} \to \mathbf{C}$ be a function with $\|\hat{f}\|_2 \ll N^{1/2}$ and let $K = 100\|\hat{f}\|_1$. Let $X_1, X_2, \dots, X_K \subset \mathbf{Z}$ be any sets such that $\min_{n\in X_i} |f(n)| \geqslant 1/2$. Then there exists distinct $j, j' \in [K]$ such that*

$$E(X_j, X_{j'}) \gg K^{-2} \frac{|X_j|^2 |X_{j'}|^2}{N}. \tag{3.10}$$

Proof. We argue by contradiction, assuming that we can find $X_1, X_2, \dots, X_K$ such that $\min_{n\in X_i} |f(n)| \geqslant 1/2$ and

$$E(X_j, X_{j'}) \leqslant cK^{-2} \frac{|X_j|^2 |X_{j'}|^2}{N}$$

for all $j < j'$, where $c > 0$ is some absolute constant to be determined later. Define the functions

$$g_i : \mathbf{Z} \to \mathbf{C} : g_i(n) = \begin{cases} |X_i|^{-1} \frac{f(n)}{|f(n)|} & \text{if } n \in X_i, \\ 0 & \text{otherwise,} \end{cases}$$

and further define $Q_i(x) = e^{-|\hat{g}_i(x)|}$ and

$$\Phi_j = \hat{g}_j + \hat{g}_{j-1}Q_j + \cdots + \hat{g}_1 Q_2 \dots Q_j.$$

We again have the basic inequalities $|Q_j(t)| \leqslant 1, |1 - Q_j(t)| \leqslant |\hat{g}_j(t)|$ and note that $\|\Phi_j\|_\infty \leqslant 10$ for all j by Lemma 3.3.3. We define

$$Z(t) = \sum_{j=1}^{K-1} \hat{g}_j(1 - Q_{j+1} \dots Q_K)$$

and hence,

$$\begin{aligned} 10\|\hat{f}\|_1 &\geqslant \langle \hat{f}, \Phi_K \rangle \\ &= \sum_{j=1}^{K} \langle \hat{f}, \hat{g}_j \rangle - \langle \hat{f}, Z \rangle \\ &\geqslant K \min_{n \in \cup_j X_j} |f(n)| - |\langle \hat{f}, Z \rangle|, \end{aligned} \tag{3.11}$$

where we used the last equation in (3.3). By a telescoping identity, we may rewrite

$$Z(t) = \sum_{j=1}^{K-1} \sum_{k=j+1}^{K} \hat{g}_j(t)(1 - Q_k(t))Q_{j+1}(t) \dots Q_{k-1}(t)$$

which yields the following upper bound

$$\begin{aligned} |\langle \hat{f}, Z \rangle| &\leqslant \sum_{1 \leqslant j < k \leqslant K} \langle |\hat{f}|, |\hat{g}_j||\hat{g}_k| \rangle \\ &\leqslant \sum_{1 \leqslant j < k \leqslant K} \|\hat{f}\|_2 \|\hat{g}_j \hat{g}_k\|_2 \end{aligned}$$

by the Cauchy-Schwarz inequality. By assumption we can bound $\|\hat{f}\|_2 \ll N^{1/2}$. As $\hat{g}_i(t) = |X_i|^{-1} \sum_{n \in X_i} \frac{f(n)}{|f(n)|} e(nt)$, the norms $\|\hat{g}_j \hat{g}_k\|_2$ can be explicitly calculated using the orthogonality of characters as follows:

$$\begin{aligned} |X_j|^2 |X_k|^2 \|\hat{g}_j \hat{g}_k\|_2^2 &= \int_0^1 \left| \sum_{n \in X_j} \frac{f(n)}{|f(n)|} e(nt) \right|^2 \left| \sum_{n \in X_k} \frac{f(n)}{|f(n)|} e(nt) \right|^2 dt \\ &= \sum_{\substack{n_1, n_2 \in X_j, n_3, n_4 \in X_k \\ n_1 - n_2 = n_3 - n_4}} \frac{f(n_1)\overline{f(n_2)f(n_3)}f(n_4)}{|f(n_1)f(n_2)f(n_3)f(n_4)|} \\ &\leqslant \#\{n_1, n_2 \in X_j, n_3, n_4 \in X_k : n_1 - n_2 = n_3 - n_4\} \\ &= E(X_j, X_k). \end{aligned}$$

As we are assuming that $E(X_j, X_k) \leqslant cK^{-2}|X_j|^2|X_k|^2/N$, we conclude that

$$|\langle \hat{f}, Z\rangle| \ll \sum_{1\leqslant j<k\leqslant K} N^{1/2}c^{1/2}K^{-1}N^{-1/2} \leqslant c^{1/2}K. \tag{3.12}$$

Recall that $K = 100\|\hat{f}\|_1$ and that $\min_{n\in\cup_j X_j}|f(n)| \geqslant 1/2$ so combining (3.11) and (3.12) produces the inequality

$$\frac{K}{10} = 10\|\hat{f}\|_1 \geqslant K \min_{n\in\cup_j X_j}|f(n)| - O(c^{1/2}K) \geqslant K/2 - O(c^{1/2}K).$$

This gives a contradiction upon choosing $c > 0$ to be sufficiently small.

□

One can apply Theorem 3.5.4 with $X_1 = \cdots = X_K = X$ to deduce the following result.

Corollary 3.5.5. *Let $f : \mathbf{Z} \to \mathbf{C}$ be a function with $\|\hat{f}\|_2 \leqslant N^{1/2}$. Let $X \subset \mathbf{Z}$ be any set such that $\min_{n\in X}|f(n)| \geqslant 1/2$. Then*

$$E(X) \gg \|\hat{f}\|_1^{-2}\frac{|X|^4}{N}. \tag{3.13}$$

3.6 Sets with small dimension have a 'dense' model

In this section we will show that sets of integers with small additive dimension are Freiman isomorphic to sets which have relatively large density on an interval, as we briefly discussed in Chapter 2. For example, we will show that a set $B \subset \mathbf{Z}$ with $\dim(B) \leqslant (\log|B|)^{O(1)}$ has a Freiman isomorphic copy B' which is contained in $[-e^{(\log|B|)^{O(1)}}, e^{(\log|B|)^{O(1)}}]$. The results in this section hold for any reasonable notion of Freiman isomorphism.

Definition 3.6.1. *Let G, G' be Abelian groups and let $A \subset G$, $A' \subset G'$. We say that a map $\phi : A \to A'$ is an F_ℓ-homomorphism if whenever $a_1, a_2, \ldots, a_\ell \in A$ satisfy*

$$\varepsilon_1 a_1 + \varepsilon_2 a_2 + \cdots + \varepsilon_\ell a_\ell = 0$$

for some $\varepsilon_j \in \{-1, 0, 1\}$, then

$$\varepsilon_1\phi(a_1) + \varepsilon_2\phi(a_2) + \cdots + \varepsilon_\ell\phi(a_\ell) = 0.$$

We say that A and A' are F_ℓ-isomorphic if there is a bijective F_ℓ-homomorphism $\phi : A \to A'$ so that ϕ^{-1} is also an F_ℓ-homomorphism.

Our first goal in this section is to prove the following theorem.

Theorem 3.6.2. *Let $A \subset \mathbf{Z}$ be a set such that any F_ℓ-isomorphic set $A' \subset \mathbf{Z}$ satisfies $\dim(A') \leqslant k$. Then A is F_ℓ-isomorphic to a subset $B \subset [-T, T]$ where $T \ll \ell^{O(k \log k)}$.*

The importance of this result in our setting is as follows.

Corollary 3.6.3 ('Dense' model I)**.** *Let $A \subset \mathbf{Z}$ be a set of size N with $S(A) \leqslant N/3 + C$. Then A is F_4-isomorphic to a set $B \subset [-T, T]$ where $T \leqslant e^{O((C \log N)^4)}$.*

Proof of Corollary 3.6.3. Let $A' \subset \mathbf{Z}$ be a set which is F_4-isomorphic to A. By Corollary 3.5.3 we have that $\dim(A') \ll C^2(\log N)^3$. The result now follows from Theorem 3.6.2 (with a somewhat stronger bound than we claimed here). □

We shall not be concerned with optimising T for now; $T = e^{O((C \log N)^4)}$ is more than sufficient for our proof of Theorem 3.1.2. In Section 3.9, we will show that if $S(A) \leqslant N/3 + C$ then A has a significantly denser F_4-model inside $[-T, T]$, where $T \leqslant N^{C^{O(1)}}$. To prove Theorem 3.6.2, we combine some combinatorial lemmas. First, we need to recall Lemma 2.0.10 from Chapter 2 which shows that if $B \subset \mathbf{Z}/p\mathbf{Z}$ has $\dim(B) \leqslant k$, then we can find a $\lambda \in (\mathbf{Z}/p\mathbf{Z})^\times$ such that $\lambda \cdot B \subset [-kp^{1-1/k}, kp^{1-1/k}]$. The next elementary lemma provides a procedure for finding F_ℓ-isomorphic copies of a set.

Lemma 3.6.4. *Let $B \subset \mathbf{Z}$ satisfy $B \subset (-p/\ell, p/\ell)$ where p is a prime. Denote by $\pi : (-p/2, p/2) \to \mathbf{Z}/p\mathbf{Z}$ the projection map. If $\lambda \in (\mathbf{Z}/p\mathbf{Z})^\times$ has that $\lambda \cdot \pi(B) \subset (-p/\ell, p/\ell) \subset \mathbf{Z}/p\mathbf{Z}$, then B is F_ℓ-isomorphic to $\pi^{-1}(\lambda \cdot \pi(B)) \subset (-p/\ell, p/\ell) \subset \mathbf{Z}$.*

Proof. B is F_ℓ-isomorphic to $\pi(B)$ precisely because the conditions $\sum_{j=1}^{\ell} \varepsilon_j x_j = 0$ and $\sum_{j=1}^{\ell} \varepsilon_j x_j \equiv 0 \pmod p$ are equivalent for integers $x_j \in (-p/\ell, p/\ell)$ and $\varepsilon_j \in \{-1, 0, 1\}$. Furthermore, it is trivial that C and any dilate $\lambda \cdot C$ are F_ℓ-isomorphic for any $C \subset \mathbf{Z}/p\mathbf{Z}$ and $\lambda \in (\mathbf{Z}/p\mathbf{Z})^\times$. □

Proof of Theorem 3.6.2. Suppose that $A \subset \mathbf{N}$ is a set such that any F_ℓ-isomorphic set $A' \subset \mathbf{Z}$ satisfies $\dim(A') \leqslant k$. We further suppose that $A' \subset \mathbf{Z}$ has that $m = m(A') := \max_{x \in A'} |x|$ is minimal over all sets A' which are F_ℓ-isomorphic to A. We may find a prime $p \in (\ell m, 2\ell m]$ and by Lemma 3.6.4, A' is F_ℓ-isomorphic to $\pi(A') \subset \mathbf{Z}/p\mathbf{Z}$. It is trivial that $\dim(\pi(A')) \leqslant \dim(A') \leqslant k$ (note that the first notion of additive dimension is taken in $\mathbf{Z}/p\mathbf{Z}$ while the second is in $\mathbf{Z}$). If it were the case

that $kp^{1-1/k} < m$, then $kp^{1-1/k} < p/\ell$ so by combining Lemmas 2.0.10 and 3.6.4 we would obtain a set A'' which is F_ℓ-isomorphic to A and moreover has

$$m(A'') \leqslant kp^{1-1/k} < m.$$

These properties contradict our choice of A' so we must have that $k(2\ell m)^{1-1/k} \geqslant kp^{1-1/k} \geqslant m$. We deduce that $m \ll (2\ell k)^k$ as desired. □

In the rest of this section, we study the structure of F_A for sets A with $S(A) \leqslant N/3 + C$ which are contained in some 'short' interval. By Corollary 3.6.3, we may indeed consider sets $A \subset [-T, T]$ where $T \leqslant e^{(\log N)^5}$ from now on, as all sets A' which do not have such a 'dense' F_4-model must satisfy $S(A') \geqslant N/3 + c(\log N)^{1/4}$. We will use the following lemma which one can think of as providing good bounds on the size of the Fourier coefficients of R_Q, which we defined in Proposition 3.4.1 as

$$R_Q(x) = \sum_{\substack{1<n\in\mathcal{R}_Q \\ a\in A}} \frac{\chi(n)}{n} c(nax).$$

We need to prove such a result in a somewhat more general setting.

Lemma 3.6.5. *Let $B_s \subset \mathbf{Z} \setminus \{0\}$, $s \in \mathcal{S}$, be a collection of disjoint sets. Let $Q > 1$ and $T \leqslant e^Q$ be parameters and let k be a non-zero integer. Define for each $s \in \mathcal{S}$ the function*

$$R_{B_s}(x) = \sum_{b\in B_s} \sum_{1<n\in\mathcal{R}_Q} \frac{\alpha(n,s)}{n} e(nkbx)$$

where the $\alpha(n,s)$ are 1-bounded complex numbers, and where $\mathcal{R}_Q$ is the set of Q-rough numbers. Then $R(x) := \sum_s R_{B_s}(x)$ satisfies the bound $|\hat{R}(m)| \ll \frac{\log T}{Q}$ for all its Fourier coefficients with frequencies $m \in [-10T, 10T]$.

Proof. This can be proved as follows:

$$\begin{aligned} |\hat{R}(m)| &\leqslant \sum_s \sum_{b\in B_s} \sum_{\substack{1<n\in\mathcal{R}_Q \\ \exists b\in B_s \text{ with } nkb=m}} \frac{1}{n} \\ &\leqslant \sum_{\substack{1<n\in\mathcal{R}_Q \\ n\mid m}} \frac{1}{n} \end{aligned}$$

which one can check by noting that each divisor $n \in \mathcal{R}_Q$ of m can contribute at most once since the sets B_s are disjoint and the relation $nkb = m$ uniquely determines b (if it exists). As $\mathcal{R}_Q$ is the set of Q-rough numbers, we can bound this by

$$|\hat{R}(m)| \ll -1 + \prod_{Q<p \text{ prime divisor of } |m|} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \dots\right)$$

$$\ll -1 + \left(1 + 2Q^{-1}\right)^{\omega(|m|)},$$

where ω counts the number of distinct prime divisors. Using the trivial bound $\omega(|m|) \ll \log T \leqslant Q$ for all $m \in [-10T, 10T]$ and the basic inequality $(1+x)^y - 1 \leqslant e^{xy} - 1 \ll xy$ which is valid for $x > 0$ and all $0 \leqslant y \ll 1/x$, we obtain the claimed bound $|\hat{R}(m)| \ll \omega(|m|)Q^{-1} \ll (\log T)Q^{-1}$. □

We show that such a result can be used to deduce strong L^2-bounds for functions of the type discussed below. For technical reasons that will become clear later, we will in practice bound the L^2-norm of a certain truncation of their Fourier series, and to obtain such truncations we recall the definition of de la Vallée-Poussin kernels.

Definition 3.6.6. We define the *de la Vallée-Poussin kernel*

$$V_T(x) = \sum_{n=-2T}^{-T-1} \left(1 - \frac{|T+n|}{T}\right) e(nx) + \sum_{n=-T}^{T} e(nx) + \sum_{n=T+1}^{2T} \left(1 - \frac{|T-n|}{T}\right) e(nx).$$

The following rather technical looking lemma provides properties that will be crucial in two later stages of the argument. We also emphasise that the variable p runs over primes only.

Lemma 3.6.7. *Let $1 \leqslant Q_1 < Q$ and $T \leqslant e^Q$ be parameters, let $(\nu_p)_{p \leqslant Q_1} \in \mathbf{N}^{\pi(Q_1)}$ and let $r \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$. Let $B_s \subset \mathbf{Z} \setminus \{0\}$, $s \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$, be a collection of sets of size at most K. Assume that k is a non-zero integer such that for each s and each $b \in B_s$ the congruence $bk \equiv s \prod_{p \leqslant Q_1} p^{\nu_p} (\mathrm{mod} \prod_{p \leqslant Q_1} p^{\nu_p + 1})$ holds. Define*

$$B_s(x) = \sum_{b \in B_s} \sum_{\substack{n \in \mathcal{R}_Q \\ n \equiv rs^{-1} (\mathrm{mod} \prod_{p \leqslant Q_1} p)}} \frac{\chi(n)}{n} e(nkbx).$$

Then

$$\sum_s B_s(x) = \sum_{b \in B_r} e(bkx) + E(x)$$

for some function $E : \mathbf{T} \to \mathbf{C}$ which satisfies $\|E * V_T\|_2 \ll \frac{\log T}{Q^{1/2}} K^{1/2}$.

Proof. Recall that $\mathcal{R}_Q$ is the set of Q-rough numbers. Trivially, the only value of $s \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ such that $n = 1$ satisfies $n \equiv rs^{-1} (\mathrm{mod} \prod_{p \leqslant Q_1} p)$ is $s = r$. Hence, if we write

$$E(x) = \sum_s \sum_{b \in B_s} \sum_{\substack{1 < n \in \mathcal{R}_Q \\ n \equiv rs^{-1} (\mathrm{mod} \prod p)}} \frac{\chi(n)}{n} e(nbkx), \tag{3.14}$$

then $\sum_s B_s(x) = \sum_{b\in B_r} e(bkx) + E(x)$. It remains to show that $\|E * V_T\|_2 \ll (\log T)Q^{-1/2}K^{1/2}$. Note that $E(x)$ is precisely of the form that we studied in the previous lemma: the congruence condition implies that the sets B_s are pairwise disjoint and we take $\alpha(n,s) = 1_{n\equiv rs^{-1} (\mathrm{mod}\ \prod p)}\chi(n)$. Hence, $\max_{|m|\leqslant 10T} |\hat{E}(m)| \ll \frac{\log T}{Q}$. Observe that $E*V_T$ is a trigonometric polynomial of degree at most $2T$ as $(E*V_T)^\wedge(n) = \hat{E}(n)\hat{V}_T(n)$, and note that $|\hat{V}_T(n)| \leqslant 1$, so by Parseval we can bound

$$\begin{aligned}\|E*V_T\|_2^2 &= \sum_{n=-2T}^{2T} |\hat{E}(n)|^2|\hat{V}_T(n)|^2 \leqslant \sum_{n=-2T}^{2T} |\hat{E}(n)|^2 \\ &\leqslant \max_{|n|\leqslant 2T} |\hat{E}(n)| \sum_{n=-2T}^{2T} |\hat{E}(n)| \ll \frac{\log T}{Q} \sum_{n=-2T}^{2T} |\hat{E}(n)|,\end{aligned}$$

so it suffices to show that $\sum_{|n|\leqslant 2T} |\hat{E}(n)| \ll K\log T$. This follows from the explicit form (3.14) of E as χ is 1-bounded:

$$\begin{aligned}\sum_{m=-2T}^{2T} |\hat{E}(m)| &\leqslant \sum_s \sum_{b\in B_s} \sum_{\substack{1<n\in\mathcal{R}_Q \\ n\equiv rs^{-1} (\mathrm{mod}\ \prod p)}} \frac{1_{|nkb|\leqslant 2T}}{n} \\ &\leqslant \sum_s |B_s| \sum_{\substack{n\in\mathcal{R}_Q \\ n\equiv rs^{-1} (\mathrm{mod}\ \prod p)}} \frac{1_{n\leqslant 2T}}{n} \\ &\leqslant K \sum_{n\leqslant 2T} \frac{1}{n} \ll K\log T,\end{aligned}$$

where we used that $\max_s |B_s| \leqslant K$. □

3.7 The distribution of A modulo small primes

We study the distribution of the set A in residue classes modulo powers of small primes under the assumption that $S(A) \leqslant N/3 + C$. Let Q_1 be a parameter. We define for $r \in \prod_{p\leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ and $(\nu_p)_{p\leqslant Q_1} \in \mathbf{N}^{\pi(Q_1)}$ the following subset of A:

$$A(r,(\nu_p)) = A \cap \left\{ n\in\mathbf{Z} : n \equiv r \prod_{p\leqslant Q_1} p^{\nu_p} (\mathrm{mod} \prod_{p\leqslant Q_1} p^{\nu_p+1}) \right\}.$$

We emphasise that the variable p always runs over primes only and that, for our purposes, $\mathbf{N}$ contains 0. We first show, using the assumption that $\dim(A)$ is small, that a large portion of A must be contained in a single such set $A(r,(\nu_p))$.

Lemma 3.7.1. *Let $A \subset \mathbf{Z}$ be a set of N integers. Then there exist $r \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ and $(\nu_p)_{p \leqslant Q_1} \in \mathbf{N}^{\pi(Q_1)}$ such that*

$$|A(r, (\nu_p))| \geqslant \frac{N}{(\dim(A))^{\pi(Q_1)} \prod_{p \leqslant Q_1} p}.$$

Proof. Let p be a prime, and for an integer n let $\nu_p(n)$ be the exact power of p dividing n. We claim that $\{\nu_p(a) : a \in A\}$ has size at most $\dim(A)$. Indeed, if $\nu_p(a_1) < \nu_p(a_2) < \cdots < \nu_p(a_k)$, then considering a possible relation $\sum_{j=1}^k \varepsilon_j a_j = 0$ with $\varepsilon_j \in \{-1, 0, 1\}$ modulo $p^{\nu_p(a_i)+1}$ for all i shows that each $\varepsilon_i = 0$, thus implying that $\{a_1, a_2, \ldots, a_k\}$ is dissociated. Hence, the image of the map

$$\rho : A \to \mathbf{N}^{\pi(Q_1)} : a \mapsto (\nu_p(a))_{p \leqslant Q_1}$$

has size at most $(\dim(A))^{\pi(Q_1)}$ and it follows that there exists a $(\nu_p)_{p \leqslant Q_1}$ such that its preimage

$$\rho^{-1}((\nu_p)) = \bigcup_{r \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times} A(r, (\nu_p))$$

has size at least $N(\dim(A))^{-\pi(Q_1)}$. Clearly, there then exists an $r \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ satisfying the conclusion of the lemma. □

This implies the following structure for sets A for which $S(A)$ is small. The exact choice of the parameter Q_1 is not important here, or in the rest of the paper (any choice $Q_1 = (\log N)^c$ for $c \in (0, 1)$ will do). For clarity, we shall from now on always take $Q_1 = (\log N)^{1/2}$.

Corollary 3.7.2. *Let $Q_1 = (\log N)^{1/2}$. Let $A \subset \mathbf{Z} \backslash \{0\}$ be a set of size N. Then either $S(A) \geqslant N/3 + c(\log N)^{1/2}$ or else the following holds. There exist $r \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ and $(\nu_p)_{p \leqslant Q_1} \in \mathbf{N}^{\pi(Q_1)}$ such that $|A(r, (\nu_p))| \geqslant_\varepsilon N^{1-\varepsilon}$.*

Proof. Corollary 3.5.3 implies that either $S(A) \geqslant N/3 + c(\log N)^{1/2}$ so that we are done, or else that $\dim(A) \ll (\log N)^4$. One can now simply use Lemma 3.7.1 and calculate that $(\dim(A))^{\pi(Q_1)} \prod_{p \leqslant Q_1} p \leqslant e^{O((\log N)^{1/2})} \ll_\varepsilon N^\varepsilon$ when $Q_1 = (\log N)^{1/2}$. □

The main result of this section is Proposition 3.7.9 which shows that the collection of sizes of the sets $A(r, (\nu_p))$ exhibits strong structure when $S(A)$ is small. The statement and proof of Proposition 3.7.9 are somewhat technical so we discuss the following model setting first. Its proof contains many of the main ideas but allows us to ignore error term contributions that show up in the general case.

Proposition 3.7.3 (Model setting). *Let $A \subset \mathbf{Z} \setminus \{0\}$ and assume that $A \subset [-T, T]$ where $T \ll e^{O((\log N)^5)}$. Let $Q_1 = (\log N)^{1/2}$ and suppose that*

$$\max_{r \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times} |A(r, (0))| \geqslant (\log N)^4,$$

where $(0) = (0)_{p \leqslant Q_1}$. Then $\|F_A\|_1 \gg \log\log N$, and in particular $S(A) \geqslant N/3 + c\log\log N$.

Proof. Suppose that $r \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ is such that

$$|A(r, (0))| = \max_{s \in \prod_p (\mathbf{Z}/p\mathbf{Z})^\times} |A(s, (0))| \geqslant (\log N)^4.$$

We recall that the function F_A is given by

$$F_A(x) = \sum_{a \in A} \sum_{n \geqslant 1} \frac{\chi(n)}{n} c(nax).$$

We define

$$\mathcal{P}_{\text{med}} = \{p \in [(\log N)^{1/2}, (\log N)^{20}] : p \text{ is prime}\}$$

and we shall think of these as 'medium'-range primes. Correspondingly, we define

$$\mathcal{R}_{\text{med}} = \{n \geqslant 1 : n \text{ is coprime to all primes in } \mathcal{P}_{\text{med}}\}.$$

We fix throughout the parameters $Q_1 = (\log N)^{1/2}$ and $Q = (\log N)^{20}$. In this proof, we will consider the function

$$F_{\text{med}}(x) = \sum_{a \in A} \sum_{n \in \mathcal{R}_{\text{med}}} \frac{\chi(n)}{n} c(nax).$$

First, we show how one may obtain F_{med} from F_A by 'sifting' out all primes in $\mathcal{P}_{\text{med}}$ with a procedure much like that in Proposition 3.4.1.

Lemma 3.7.4. *We have that $\|F_{\text{med}}\|_1 \ll \|F_A\|_1$.*

Proof of Lemma 3.7.4. We follow the proof of Proposition 3.4.1 to show that

$$\sum_{k \mid \prod_{p \in \mathcal{P}_{\text{med}}} p} \frac{\mu(k)\chi(k)}{k} F_A(kx) = F_{\text{med}}(x). \tag{3.15}$$

As χ is multiplicative, we can calculate

$$\sum_{k \mid \prod_{p \in \mathcal{P}_{\text{med}}} p} \frac{\mu(k)\chi(k)}{k} \sum_{n \geqslant 1} \frac{\chi(n)}{n} c(nkx) = \sum_{m \geqslant 1} \frac{\chi(m)}{m} c(mx) \sum_{k \mid \gcd(m, \prod_{p \in \mathcal{P}_{\text{med}}} p)} \mu(k)$$

$$= \sum_{m \in \mathcal{R}_{\mathrm{med}}} \frac{\chi(m)}{m} c(mx).$$

Replacing x by ax and summing over $a \in A$ proves (3.15). We again follow the proof of Proposition 3.4.1 to bound

$$\begin{aligned} \|F_{\mathrm{med}}\|_1 &= \left\| \sum_{k \mid \prod_{p \in \mathcal{P}_{\mathrm{med}}} p} \frac{\mu(k)\chi(k)}{k} F(kx) \right\|_1 \\ &\leqslant \|F\|_1 \prod_{p \in \mathcal{P}_{\mathrm{med}}} (1 + 1/p) \\ &\ll \|F\|_1, \end{aligned}$$

since $\sum_{p \in \mathcal{P}_{\mathrm{med}}} \frac{1}{p} = \sum_{p \in [(\log N)^{1/2}, (\log N)^{20}]} \frac{1}{p} \ll 1$ by Mertens' estimate [51, Theorem 2.7]. □

We will show that $\|F_{\mathrm{med}}\|_1 \gg \log\log N$ which, by the lemma above, then implies the desired lower bound $\|F_A\|_1 \gg \log\log N$. In order for the information about $A(r, (0))$ to be exploited, we will use the following lemma.

Lemma 3.7.5. *Let $h(x) = \sum_{n \in \mathbf{Z}} \hat{h}(n) e(nx) \in L^1(\mathbf{T})$ and let $q \in \mathbf{N}$ and $\ell \in \mathbf{Z}/q\mathbf{Z}$. Then $\| \sum_{n \equiv \ell (\mathrm{mod}\, q)} \hat{h}(n) e(nx) \|_1 \leqslant \|h\|_1$.*

Proof of Lemma 3.7.5. This follows as

$$\sum_{k \equiv \ell (\mathrm{mod}\, q)} \hat{h}(k) e(kx) = \frac{1}{q} \sum_{j=0}^{q-1} e(-\ell j/q) h(x + j/q),$$

by orthogonality of characters modulo q. Hence, by the triangle inequality and as $\|h(x + j/q)\|_1 = \|h(x)\|_1$, the left hand side has L^1-norm at most $\|h\|_1$. □

Hence, if we define

$$\mathrm{Proj}(F_{\mathrm{med}}; r, (0))(x) = \sum_{k \equiv r (\mathrm{mod}\, \prod_{p \leqslant Q_1} p)} \hat{F}_{\mathrm{med}}(k) e(kx)$$

to be the function obtained by keeping only those terms in the Fourier series of F_{med} whose frequencies are $r (\mathrm{mod} \prod_{p \leqslant Q_1} p)$, then $\|\mathrm{Proj}(F_{\mathrm{med}}; r, (0))\|_1 \leqslant \|F_{\mathrm{med}}\|_1$. We shall analyse the structure of $\mathrm{Proj}(F_{\mathrm{med}}; r, (0))$. Recall that we use the notation

$$A(s, (\mu_p)) = \{a \in A : a \equiv s \prod_{p \leqslant Q_1} p^{\mu_p} (\mathrm{mod} \prod_{p \leqslant Q_1} p^{\mu_p + 1})\}.$$

We can decompose

$$F_{\mathrm{med}}(x) = \sum_{s,(\mu_p)} \sum_{a \in A(s,(\mu_p))} \left(\sum_{n \in \mathcal{R}_{\mathrm{med}}} \frac{\chi(n)}{n} c(nax) \right) \tag{3.16}$$

and we shall consider the contribution from each $A(s,(\mu_p))$ to $\mathrm{Proj}(F_{\mathrm{med}}; r, (0))$. It is convenient to write

$$F_{\mathrm{med},s,(\mu_p)}(x) := \sum_{a \in A(s,(\mu_p))} \sum_{n \in \mathcal{R}_{\mathrm{med}}} \frac{\chi(n)}{n} c(nax)$$

so that (3.16) becomes

$$F_{\mathrm{med}} = \sum_{s,(\mu_p)} F_{\mathrm{med},s,(\mu_p)}.$$

Let $a \in A(s,(\mu_p))$ for some $(s,(\mu_p))$ and suppose that one of its terms $\frac{\chi(n)}{n} c(nax) = \frac{\chi(n)}{2n}(e(nax) + e(-nax))$ contributes to $\mathrm{Proj}(F_{\mathrm{med}}; r, (0))$. This means precisely that $n \in \mathcal{R}_{\mathrm{med}}$ satisfies

$$\pm na \equiv r (\mathrm{mod} \prod_{p \leqslant Q_1} p)$$

and in particular this implies that $\mu_p = 0$ and that $\nu_p(n) = 0$ for all $p \leqslant Q_1$ (here $\nu_p(n)$ denotes the largest integer such that p^{ν_p} divides n). We deduce that $\mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (0)\big) = 0$ unless $(\mu_p) = (0)$, and so

$$\mathrm{Proj}(F_{\mathrm{med}}; r, (0)) = \sum_{s \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times} \mathrm{Proj}\big(F_{\mathrm{med},s,(0)}; r, (0)\big). \tag{3.17}$$

We know from above that if a term $\frac{\chi(n)}{n} c(nax)$ contributes to $\mathrm{Proj}(F_{\mathrm{med}}; r, (0))$, where $n \in \mathcal{R}_{\mathrm{med}}$ and $a \in A(s,(0))$ for some $s \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$, then $(n,p) = 1$ for all $p \leqslant Q_1$. Hence, we must in fact have that $n \in \mathcal{R}_Q$, where we recall that $\mathcal{R}_Q$ is the set of Q-rough numbers. Finally, if $a \in A(s,(0))$ then $a \equiv s (\mathrm{mod} \prod p)$ by definition, so the congruence $\pm an \equiv r (\mathrm{mod} \prod p)$ is satisfied precisely when $n \equiv \pm r s^{-1} (\mathrm{mod} \prod p)$. It follows that we can precisely write down the contribution of $F_{\mathrm{med},s,(0)}$ to $\mathrm{Proj}(F_{\mathrm{med}}; r, (0))$ as[4]

$$\mathrm{Proj}\big(F_{\mathrm{med},s,(0)}; r, (0)\big) = \frac{1}{2} \sum_{a \in A(s,(0))} \sum_{\substack{n \in \mathcal{R}_Q \\ n \equiv r s^{-1} (\mathrm{mod} \prod p)}} \frac{\chi(n)}{n} e(nax) \tag{3.18}$$

[4]Intuitively, one should think of this projection onto a residue class $r (\mathrm{mod} \prod_{p \leqslant Q_1} p)$ (for which $A(r,(0))$ is large) as replacing the need to 'sift' out primes $p \leqslant Q_1$ as in Bourgain's approach in Proposition 3.4.1, in that it also restricts the sum over n to a sum over Q-rough numbers. This is ultimately the reason why we are able to gain a factor of $\log Q_1 \gg \log\log N$.

$$+\frac{1}{2}\sum_{a\in A(s,(0))}\sum_{\substack{n\in\mathcal{R}_Q\\ n\equiv -rs^{-1}(\mathrm{mod}\ \prod p)}}\frac{\chi(n)}{n}e(-nax).$$

Our objective is to obtain a strong lower bound for the L^1-norm of $\mathrm{Proj}(F_{\mathrm{med}};r,(0))$. For this purpose, we will show that the main term comes from the projection of $F_{\mathrm{med},r,(0)}$, whereas all other $F_{\mathrm{med},s,(0)}$ contribute an error term which is negligible in an L^2-sense. In practice, we will find estimates for the L^1-norm of the convolution $\mathrm{Proj}(F_{\mathrm{med}};r,(0))*V_T=\mathrm{Proj}(F_{\mathrm{med}}*V_T;r,(0))$ where $V_T(x)$ is a de la Vallée-Poussin kernel, rather than for $\mathrm{Proj}(F_{\mathrm{med}};r,(0))$. Recall that $T=e^{O((\log N)^5)}$ is such that $A\subset[-T,T]$. We also recall that by Lemma 2.1.2 and our discussion of the de la Vallée-Poussin kernels in Chapter 2, V_T has the following properties:

- $\hat{V}_T(n)=1$ for $|n|<T$ while $\hat{V}_T(n)=0$ for $|n|\geqslant 2T$.
- $\|V_T\|_1\leqslant 3$.

These imply the following useful relation between $\mathrm{Proj}(F_{\mathrm{med}};r,(0))*V_T$ and F_A.

Lemma 3.7.6. *We define* $F^*=\mathrm{Proj}(F_{\mathrm{med}};r,(0))*V_T$. *Then* $\|F^*\|_1\ll\|F_A\|_1$.

Proof of Lemma 3.7.6. Note that from (a trivial instance of) Young's convolution inequality we obtain

$$\|\mathrm{Proj}(F_{\mathrm{med}};r,(0))*V_T\|_1\leqslant\|\mathrm{Proj}(F_{\mathrm{med}};r,(0))\|_1\|V_T\|_1\ll\|F_{\mathrm{med}}\|_1$$

where we used Lemma 3.7.5 and that $\|V_T\|_1\ll 1$. The result follows upon recalling Lemma 3.7.4. □

The following lemma provides properties of the functions $\mathrm{Proj}(F_{\mathrm{med},s,(0)};r,(0))$ appearing in the expression (3.17) for $\mathrm{Proj}(F_{\mathrm{med}};r,(0))$ which are relevant for the eventual purpose of applying a McGehee-Pigno-Smith style construction to lower bound the L^1 norm.

Lemma 3.7.7. *We have that*

$$\sum_{s\in\prod_{p\leqslant Q_1}(\mathbf{Z}/p\mathbf{Z})^\times}\mathrm{Proj}\big(F_{\mathrm{med},s,(0)};r,(0)\big)(x)$$
$$=\frac{1}{2}\left(\sum_{a\in A(r,(0))}e(ax)+\sum_{a\in A(-r,(0))}e(-ax)+E_{(0)}(x)\right)$$

for some function $E_{(0)}:\mathbf{T}\to\mathbf{C}$ *satisfying* $\|E_{(0)}*V_T\|_2\ll(\log N)^{-2}|A(r,(0))|^{1/2}$.

We write $E_{(0)}$ with the subscript (0) to be consistent with the notation used in the proof of the general setting Proposition 3.7.9.

Proof of Lemma 3.7.7. This is immediate from an application of Lemma 3.6.7 with the sets $B_s = A(s,(0))$ and $k = \pm 1$, and with our choice of parameters $Q = (\log N)^{20}$ and $T = e^{O((\log N)^5)}$, by noting that the explicit Fourier expansion (3.18) of the functions $\mathrm{Proj}\big(F_{\mathrm{med},s,(0)}; r, (0)\big)$ is precisely of the type to which Lemma 3.6.7 applies. Furthermore, we note that by assumption $K = \max_s |B_s| = \max_s |A(s,(0))| = |A(r,(0))|$. □

Let us summarise what we have achieved thus far. We have shown that if $A \subset \mathbf{Z} \setminus \{0\} \subset [-T,T]$ where $T \leqslant e^{O((\log N)^5)}$, and moreover $|A(r,(0))| = \max_s |A(s,(0))|$, then by (3.17) and the lemma above, we can write the function $\mathrm{Proj}(F_{\mathrm{med}}; r, (0))$ as

$$\mathrm{Proj}\big(F_{\mathrm{med}}; r, (0)\big) = \frac{1}{2}\left(\sum_{a \in A(r,(0)) \cup -A(-r,(0))} e(ax) + E_{(0)}(x)\right),$$

and where

$$\|E_{(0)} * V_T\|_2 \ll (\log N)^{-2} |A(r,(0))|^{1/2}.$$

Let us finally consider the function $F^*(x) = \mathrm{Proj}(F_{\mathrm{med}}; r, (0)) * V_T$. By the above, we may write

$$F^*(x) = \left(\frac{1}{2}\sum_{a \in A(r,(0)) \cup -A(-r,(0))} e(ax)\right) * V_T + E^*(x)$$

where $\|E^*\|_2 \ll (\log N)^{-2} |A(r,(0))|^{1/2}$. Hence,

$$F^*(x) = \frac{1}{2}\sum_{a \in A(r,(0)) \cup -A(-r,(0))} e(ax) + E^*,$$

noting that $A \subset [-T,T]$ and that $\hat{V}_T(n) = 1$ whenever $|n| \leqslant T$ which implies that $e(\pm ax) * V_T(x) = e(\pm ax)$ for all $a \in A$.

As we are assuming that $|A(r,(0))| \geqslant (\log N)^4$, it now follows from an application of the next theorem with $B_1 = A(r,(0)), B_2 = -A(-r,(0))$, $E = E^*$ and $K = \log N$ that such a function F^* of the form above has L^1-norm at least $\|F^*\|_1 \gg \log\log N$. Note that this implies that $\|F_A\|_1 \gg \log\log N$ by Lemma 3.7.6 and hence this finishes the proof of Proposition 3.7.3.

Theorem 3.7.8. *Let $K \geqslant 1$ be a parameter. Let $B_1, B_2 \subset \mathbf{Z}$ be finite with $|B_1| \geqslant K^2$, and let $E \in L^2(\mathbf{T})$ satisfy $\|E\|_2 \leqslant |B_1|^{1/2}/K$. Then $\|\hat{1}_{B_1} + \hat{1}_{B_2} + E\|_1 \gg \log K$.*

Proof of Theorem 3.7.8. This follows from a relatively straightforward adaptation of the method of McGehee-Pigno-Smith, see Section 3.10. □

□

We now come to the main result of this section. Corollary 3.7.2 shows that either $S(A) \geqslant N/3 + c(\log N)^{1/2}$ or else one of the sets $A(r, (\nu_p))$ has size at least $N^{1/2}$ (say). The previous proposition allows us to obtain the desired bound $S(A) \geqslant N/3 + c \log\log N$ if it happens to be the case that $(\nu_p) = (0)$. We show in the following proposition that one can still deduce strong structural information about A from $A(r, (\nu_p))$ being large for a general $(\nu_p) \in \mathbf{N}^{\pi(Q_1)}$. We first need to introduce some convenient notation and we shall write

$$(\nu'_p)_{p \leqslant Q_1} \prec (\nu_p)_{p \leqslant Q_1}$$

if $\nu'_p \leqslant \nu_p$ for all $p \leqslant Q_1$ and $\nu'_p < \nu_p$ for at least one p. We shall also not repeatedly write $p \leqslant Q_1$ and it shall be clear from context what the range of p is.

Proposition 3.7.9. *Let $A \subset \mathbf{Z} \setminus \{0\}$ be a set of size N and assume that $A \subset [-T, T]$ where $T \ll e^{O((\log N)^5)}$. Then either $\|F_A\|_1 \gg \log\log N$ so that $S(A) \geqslant N/3 + c \log\log N$, or else the following holds. Let $Q_1 = (\log N)^{1/2}$. Suppose that there exist an $r \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ and $(\nu_p)_{p \leqslant Q_1} \in \mathbf{N}^{\pi(Q_1)}$ such that*

$$|A(r, (\nu_p))| = \max_{s \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times} |A(s, (\nu_p))| \geqslant (\log N)^4.$$

Then there exist an $r' \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ and $(\nu'_p)_{p \leqslant Q_1} \in \mathbf{N}^{\pi(Q_1)}$ such that

- $(\nu'_p) \prec (\nu_p)$,
- $|A(r', (\nu'_p))| \geqslant (\log N)^{-4} |A(r, (\nu_p))|$.

Proof. We argue by assuming that there exist $r, (\nu_p)$ such that $\max_s |A(s, (\nu_p))| = |A(r, (\nu_p))| \geqslant (\log N)^4$ and that $|A(r', (\nu'_p))| < (\log N)^{-4} |A(r, (\nu_p))|$ for all r' and $(\nu'_p) \prec (\nu_p)$. We shall show that under these assumptions, $\|F_A\|_1 \gg \log\log N$ and recall from Proposition 3.4.1 that such a bound implies that $S(A) \geqslant N/3 + c \log\log N$ for some absolute $c > 0$.

The proof has many similarities to the proof of the model setting above, so we shall be brief in our discussion of these parts. Let us consider, as per usual, the function

$$F_A(x) = \sum_{a \in A} \sum_{n \geqslant 1} \frac{\chi(n)}{n} c(nax).$$

We fix throughout the parameters $Q_1 = (\log N)^{1/2}$ and $Q = (\log N)^{20}$. We define $\mathcal{P}_{\text{med}} = \{p \in [Q_1, Q] : p \text{ is prime}\}$ and $\mathcal{R}_{\text{med}} = \{n \geqslant 1 : (n,p) = 1 \text{ for all } p \in \mathcal{P}_{\text{med}}\}$. We consider the function $F_{\text{med}}(x) = \sum_{a\in A}\sum_{n\in\mathcal{R}_{\text{med}}} \frac{\chi(n)}{n} c(nax)$. Exactly as in Lemma 3.7.4, we have the following relation between the L^1 norms of F_{med} and F_A.

Lemma 3.7.10. *We have that* $\|F_{\text{med}}\|_1 \ll \|F_A\|_1$.

Combining this with Lemma 3.7.5 shows the following.

Lemma 3.7.11. *Let*

$$\text{Proj}(F_{\text{med}}; r, (\nu_p))(x) = \sum_{k \equiv r \prod_{p\leqslant Q_1} p^{\nu_p} (\text{mod } \prod_{p\leqslant Q_1} p^{\nu_p+1})} \hat{F}_{\text{med}}(k) e(kx)$$

be the function obtained by keeping only those terms in the Fourier series of F_{med} *whose frequencies are* $r \prod_{p\leqslant Q_1} p^{\nu_p} (\text{mod } \prod_{p\leqslant Q_1} p^{\nu_p+1})$. *Then* $\|\text{Proj}(F_{\text{med}}; r, (\nu_p))\|_1 \ll \|F_A\|_1$.

We shall analyse the structure of $\text{Proj}(F_{\text{med}}; r, (\nu_p))$. We can decompose

$$F_{\text{med}}(x) = \sum_{s,(\mu_p)} \sum_{a\in A(s,(\mu_p))} \left(\sum_{n\in\mathcal{R}_{\text{med}}} \frac{\chi(n)}{n} c(nax) \right) \tag{3.19}$$

and we shall consider the contribution from each $A(s,(\mu_p))$ to $\text{Proj}(F_{\text{med}}; r, (\nu_p))$. It is convenient to write

$$F_{\text{med},s,(\mu_p)}(x) := \sum_{a\in A(s,(\mu_p))} \sum_{n\in\mathcal{R}_{\text{med}}} \frac{\chi(n)}{n} c(nax)$$

so that (3.19) becomes

$$F_{\text{med}} = \sum_{s,(\mu_p)} F_{\text{med},s,(\mu_p)}. \tag{3.20}$$

Let $a \in A(s,(\mu_p))$ for some $(s,(\mu_p))$ and suppose that one of its terms $\frac{\chi(n)}{n} c(nax) = \frac{\chi(n)}{2n}(e(nax) + e(-nax))$ contributes to $\text{Proj}(F_{\text{med}}; r, (\nu_p))$. This means precisely that $n \in \mathcal{R}_{\text{med}}$ satisfies

$$\pm na \equiv r \prod_{p\leqslant Q_1} p^{\nu_p} (\text{mod } \prod_{p\leqslant Q_1} p^{\nu_p+1})$$

and in particular this implies that $\nu_p \geqslant \nu_p(a) = \mu_p$ and that $\nu_p(n) = \nu_p - \mu_p$ for all $p \leqslant Q_1$ (here $\nu_p(n)$ denotes the largest integer such that p^{ν_p} divides n). We deduce that $\text{Proj}\big(F_{\text{med},s,(\mu_p)}; r, (\nu_p)\big) = 0$ unless $(\mu_p) \preceq (\nu_p)$, meaning that the only

terms in (3.20) which contribute to $\mathrm{Proj}\big(F_{\mathrm{med}}; r, (\nu_p)\big)$ come from those $(s, (\mu_p))$ with $(\mu_p) \preceq (\nu_p)$ and so

$$\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p)) = \sum_{(s,(\mu_p)):(\mu_p)\preceq(\nu_p)} \mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p)\big). \tag{3.21}$$

Let us consider $(\mu_p) \preceq (\nu_p)$. We know from our discussion above that if a term $\frac{\chi(n)}{n} c(nax)$ contributes to $\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p))$, where $n \in \mathcal{R}_{\mathrm{med}}$ and $a \in A(s, (\mu_p))$ for some $s \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$, then $\nu_p(n) = \nu_p - \mu_p$ for all $p \leqslant Q_1$. Hence, we can write $n = n' \prod_{p \leqslant Q_1} p^{\nu_p - \mu_p}$ for some $n' \in \mathcal{R}_Q$, where we recall that $\mathcal{R}_Q$ is the set of Q-rough numbers. Finally, if $a \in A(s, (\mu_p))$ then $a \equiv s \prod p^{\mu_p} (\mathrm{mod} \prod p^{\mu_p+1})$ by definition, so the congruence $\pm a n' \prod_{p \leqslant Q_1} p^{\nu_p - \mu_p} \equiv r \prod_{p \leqslant Q_1} p^{\nu_p} (\mathrm{mod} \prod_{p \leqslant Q_1} p^{\nu_p+1})$ is satisfied precisely when $n' \equiv \pm r s^{-1} (\mathrm{mod} \prod p)$. It follows that for each $(s, (\mu_p))$ with $(\mu_p) \preceq (\nu_p)$, we can precisely write down the contribution of $F_{\mathrm{med},s,(\mu_p)}$ to $\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p))$ as

$$\begin{aligned}\mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p)\big)(x) = &\sum_{a \in A(s,(\mu_p))} \sum_{\substack{n = n' \prod p^{\nu_p - \mu_p} \\ n' \in \mathcal{R}_Q \\ n' \equiv r s^{-1} (\mathrm{mod} \prod p)}} \frac{\chi(n)}{2n} e(nax) \\ &+ \sum_{a \in A(s,(\mu_p))} \sum_{\substack{n = n' \prod p^{\nu_p - \mu_p} \\ n' \in \mathcal{R}_Q \\ n' \equiv -r s^{-1} (\mathrm{mod} \prod p)}} \frac{\chi(n)}{2n} e(-nax),\end{aligned}$$

Using the multiplicative nature of χ, we can take out the factor $k(\mu_p) := \prod_{p \leqslant Q_1} p^{\nu_p - \mu_p}$ and simplify this further to

$$\begin{aligned}\mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p)\big) = &\frac{\chi(k(\mu_p))}{2k(\mu_p)} \sum_{a \in A(s,(\mu_p))} \sum_{\substack{n' \in \mathcal{R}_Q \\ n' \equiv r s^{-1} (\mathrm{mod} \prod p)}} \frac{\chi(n')}{n'} e(n' k(\mu_p) a x) \\ &+ \frac{\chi(k(\mu_p))}{2k(\mu_p)} \sum_{a \in A(s,(\mu_p))} \sum_{\substack{n' \in \mathcal{R}_Q \\ n' \equiv -r s^{-1} (\mathrm{mod} \prod p)}} \frac{\chi(n')}{n'} e(-n' k(\mu_p) a x).\end{aligned} \tag{3.22}$$

Contrary to the model case in Proposition 3.7.3, these functions $\mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p)\big)$ can contribute to $\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p))$ for all $s \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ and all $(\mu_p) \preceq (\nu_p)$, rather than only for $(\mu_p) = (\nu_p)$. For the purpose of obtaining a lower bound for the L^1-norm of $\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p))$, we will show that the main term still comes from the projection of $F_{\mathrm{med},r,(\nu_p)}$, whereas all other $F_{\mathrm{med},s,(\mu_p)}$ with $(\mu_p) \preceq (\nu_p)$ contribute an error term which is negligible in an L^2-sense due to our assumption that $\max_{(\mu_p) \prec (\nu_p)} \max_s |A(s, (\mu_p))| < (\log N)^{-4} |A(r, (\nu_p))|$. Again, we will in practice find

estimates for the L^1-norm of the convolution $\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p)) * V_T$ where $V_T(x)$ is a de la Vallée-Poussin kernel, and $T = e^{O((\log N)^5)}$ is such that $A \subset [-T, T]$. Analogously to Lemma 3.7.6, we note the following useful relation between $\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p)) * V_T$ and F_A.

Lemma 3.7.12. *We define* $F^* = \mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p)) * V_T$. *Then* $\|F^*\|_1 \ll \|F_A\|_1$.

Recall that we have an explicit expression (3.21) for $\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p))$ in terms of the contributions $\mathrm{Proj}(F_{\mathrm{med},s,(\mu_p)}; r, (\mu_p))$ which are given by (3.22), for $(\mu_p) \preceq (\nu_p)$. The following lemma provides a useful description of these $\mathrm{Proj}(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p))$.

Lemma 3.7.13. *Let* $(\mu_p) \preceq (\nu_p)$ *and define* $k(\mu_p) = \prod_{p \leqslant Q_1} p^{\nu_p - \mu_p}$. *Then*

$$\sum_s \mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p)\big)(x)$$
$$= \frac{\chi(k(\mu_p))}{2k(\mu_p)} \left(\sum_{a \in A(r,(\mu_p))} e(ak(\mu_p)x) + \sum_{a \in A(-r,(\mu_p))} e(-ak(\mu_p)x) + E_{(\mu_p)}(x) \right)$$

for some function $E_{(\mu_p)} : \mathbf{T} \to \mathbf{C}$ *satisfying* $\|E_{(\mu_p)} * V_T\|_2 \ll (\log N)^{-2} |A(r, (\nu_p))|^{1/2}$.

Proof of Lemma 3.7.13. This is immediate from an application of Lemma 3.6.7 with $B_s = A(s, (\mu_p))$ and $k = \pm k(\mu_p)$, and with our choice of parameters $Q = (\log N)^{20}$ and $T = e^{O((\log N)^5)}$, by noting that the explicit Fourier expansion (3.22) of the functions $\mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p)\big)$ is precisely of the type to which Lemma 3.6.7 applies. Furthermore, we note that by assumption $K = \max_s |B_s| = \max_s |A(s, (\mu_p))| \leqslant |A(r, (\nu_p))|$ for all $(\mu_p) \preceq (\nu_p)$. □

Let us summarise what we have achieved thus far. We have shown that if $A \subset \mathbf{Z} \setminus \{0\}$ is a set of N integers with $A \subset [-T, T]$ where $T \leqslant e^{O((\log N)^5)}$, and moreover $|A(r, (\nu_p))| \geqslant \max_s |A(s, (\mu_p))|$ for all $(\mu_p) \preceq (\nu_p)$, then by (3.21) we can write the function $\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p))$ as

$$\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p)) = \sum_{(\mu_p) : (\mu_p) \preceq (\nu_p)} \sum_s \mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p)\big), \tag{3.23}$$

and where there exists, for each $(\mu_p) \preceq (\nu_p)$, a function $E_{(\mu_p)}$ satisfying

$$\|E_{(\mu_p)} * V_T\|_2 \ll (\log N)^{-2} |A(r, (\nu_p))|^{1/2} \tag{3.24}$$

such that

$$\sum_s \mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p)\big)(x) \tag{3.25}$$

$$= \frac{\chi(k(\mu_p))}{2k(\mu_p)} \left(\sum_{a \in A(r,(\mu_p)) \cup -A(-r,(\mu_p))} e(ak(\mu_p)x) + E_{(\mu_p)}(x) \right),$$

and we recall that $k(\mu_p) = \prod_{p \leqslant Q_1} p^{\nu_p - \mu_p}$ and that s always ranges over $\prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$.

We are in addition assuming that $|A(r,(\nu_p))| = \max_s |A(s,(\nu_p))| \geqslant (\log N)^4$ while

$$\max_s |A(s,(\mu_p))| \leqslant (\log N)^{-4} |A(r,(\nu_p))| \tag{3.26}$$

for all $(\mu_p) \prec (\nu_p)$. We show that this implies that out of all terms (3.25), the contribution from $A(r,(\nu_p)) \cup -A(-r,(\nu_p))$ turns out to form, for the purpose of obtaining a lower bound for $\|\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p))\|_1$, the main term. In practice, this means that we prove that the contribution of all other terms combined is negligible in an L^2-sense. To make this precise, we define

$$E(x) := \frac{1}{2} E_{(\nu_p)}(x) + \sum_{(\mu_p) \prec (\nu_p)} \sum_s \mathrm{Proj}\big(F_{\mathrm{med},s,(\mu_p)}; r, (\nu_p)\big)(x)$$

which by (3.23) and (3.25) allows us to write

$$\mathrm{Proj}(F_{\mathrm{med}}; r, (\nu_p))(x) = \frac{1}{2} \sum_{a \in A(r,(\nu_p)) \cup -A(-r,(\nu_p))} e(ax) + E(x), \tag{3.27}$$

and we will prove the following L^2-estimate.

Lemma 3.7.14. *E satisfies the estimate $\|E * V_T\|_2 \ll (\log N)^{-1} |A(r,(\nu_p))|^{1/2}$.*

Proof of Lemma 3.7.14. By (3.24) we have the desired bound for $\|E_{(\nu_p)} * V_T\|_2$ so it suffices to consider the contribution of those terms (3.25) from all $(\mu_p) \prec (\nu_p)$ which we denote by

$$E'(x) = \sum_{(\mu_p) \prec (\nu_p)} \frac{\chi(k(\mu_p))}{2k(\mu_p)} \left(\sum_{a \in A(r,(\mu_p))} e(ak(\mu_p)x) + \sum_{a \in A(-r,(\mu_p))} e(-ak(\mu_p)x) \right)$$
$$+ \sum_{(\mu_p) \prec (\nu_p)} \frac{\chi(k(\mu_p))}{2k(\mu_p)} E_{(\mu_p)}(x).$$

Let us denote the first term on the right hand side of the equation above by E_1 and the second E_2. The term E_2 is easy to estimate upon recalling (3.24), that $k(\mu_p) = \prod_{p \leqslant Q_1} p^{\nu_p - \mu_p}$ and that χ is 1-bounded:

$$\|E_2 * V_T\|_2 \leqslant \sum_{(\mu_p) \prec (\nu_p)} \frac{1}{2 \prod_p p^{\nu_p - \mu_p}} \|E_{(\mu_p)} * V_T\|_2$$

$$\ll (\log N)^{-2}|A(r,(\nu_p))|^{1/2}\prod_{p\leqslant Q_1}(1+1/p+1/p^2+\dots)$$
$$\ll (\log N)^{-1}|A(r,(\nu_p))|^{1/2},$$

where we used that $\prod_{p\leqslant Q_1}(1-1/p)^{-1}\ll \log Q_1\ll \log\log N$ by Mertens' theorem. For the term E_1, we may note by Parseval and (3.26) that

$$\Big\|\sum_{a\in A(r,(\mu_p))}e(ak(\mu_p)x)\Big\|_2=|A(r,(\mu_p))|^{1/2}\leqslant(\log N)^{-2}|A(r,(\nu_p))|^{1/2}$$

for $(\mu_p)\prec(\nu_p)$ and that the same bound holds for the contribution from $A(-r,(\mu_p))$. By Young's inequality, this L^2 bound also holds after convolving with V_T since $\|V_T\|_1\ll 1$:

$$\Big\|\Big(\sum_{a\in A(r,(\nu_p))}e(ak(\mu_p)x)\Big)*V_T\Big\|_2\leqslant\Big\|\sum_{a\in A(r,(\mu_p))}e(ak(\mu_p)x)\Big\|_2\|V_T\|_1$$
$$\ll(\log N)^{-2}|A(r,(\nu_p))|^{1/2}.$$

Hence, summing over all $(\mu_p)\prec(\nu_p)$ gives

$$\|E_1*V_T\|_2\ll(\log N)^{-2}|A(r,(\nu_p))|^{1/2}\sum_{(\mu_p)\prec(\nu_p)}\frac{1}{\prod_p p^{\nu_p-\mu_p}}$$
$$\ll(\log N)^{-2}|A(r,(\nu_p))|^{1/2}\prod_{p\leqslant Q_1}(1+1/p+1/p^2+\dots)$$
$$\ll(\log N)^{-1}|A(r,(\nu_p))|^{1/2}.$$

□

Let us finally consider the function $F^*(x)=\mathrm{Proj}(F_{\mathrm{med}};r,(\nu_p))*V_T$. From (3.27) and Lemma 3.7.14, we may write

$$F^*(x)=\left(\frac{1}{2}\sum_{a\in A(r,(\nu_p))\cup -A(-r,(\nu_p))}e(ax)\right)*V_T+E^*(x)$$

where $\|E^*\|_2\leqslant(\log N)^{-1}|A(r,(\nu_p))|^{1/2}$. Hence,

$$F^*(x)=\frac{1}{2}\sum_{a\in A(r,(\nu_p))\cup -A(-r,(\nu_p))}e(ax)+E^*,$$

noting that $A\subset[-T,T]$ and that $\hat{V}_T(n)=1$ whenever $|n|\leqslant T$ which implies that $e(\pm ax)*V_T(x)=e(\pm ax)$ for all $a\in A$.

An application of Theorem 3.7.8 with $B_1=A(r,(\nu_p)),B_2=-A(-r,(\nu_p)),E=E^*$ and $K=\log N$ implies that that $\|F^*\|_1\gg\log\log N$ and by Lemma 3.7.12 we deduce that $\|F_A\|_1\gg\log\log N$ which finishes the proof of Proposition 3.7.9. □

3.8 Non-Archimedean test functions

In this section, we show that F_A has large L^1-norm for sets A which exhibit the strong structure provided by Proposition 3.7.9. We shall achieve this by showing that

$$c_A + R_Q = \sum_{a\in A} c(ax) + \sum_{a\in A}\sum_{1<n\in\mathcal{R}_Q} \frac{\chi(n)}{n} c(nax)$$

has large L^1-norm, where $Q = (\log N)^{20}$, and then quoting (iv) in Proposition 3.4.1. We will obtain our bound for the L^1-norm by constructing a test function in a McGehee-Pigno-Smith style fashion, but with a crucial modification: instead of constructing the test function using an increasing ordering of the Fourier spectrum as in Section 3.10, in particular relying on one-sided Fourier series, we construct the test function using an ordering of (a subset of) the spectrum based on p-adic valuations. First, we begin by showing that sets A which have the properties that Corollary 3.7.2 and Proposition 3.7.9 establish contain the following convenient structure.

Lemma 3.8.1. *Let $A \subset \mathbf{Z}$ and assume that A has the following two properties for a parameter Q_1.*

- *There exist $r^* \in \prod_{p\leqslant Q_1}(\mathbf{Z}/p\mathbf{Z})^\times$ and $(\nu_p^*)_{p\leqslant Q_1} \in \mathbf{N}^{\pi(Q_1)}$ with $|A(r^*,(\nu_p^*))| \geqslant N^{1/2}$.*
- *Whenever $r,(\nu_p)$ are such that $|A(r,(\nu_p))| = \max_{s\in\prod_{p\leqslant Q_1}(\mathbf{Z}/p\mathbf{Z})^\times}|A(s,(\nu_p))|$ and $|A(r,(\nu_p))| \geqslant (\log N)^4$, then there exist r' and $(\nu_p') \prec (\nu_p)$ satisfying $|A(r',(\nu_p'))| \geqslant (\log N)^{-4}|A(r,(\nu_p))|$.*

Then we can find an integer $J \gg (\log N)/\log\log N$, and a sequence of residues $r^{(i)} \in \prod_{p\leqslant Q_1}(\mathbf{Z}/p\mathbf{Z})^\times$ and $(\nu_p^{(i)}) \in \mathbf{N}^{\pi(Q_1)}$ such that:

(i) $|A(r^{(i)},(\nu_p^{(i)}))| = \max_s |A(s,(\nu_p^{(i)}))|$ *for all* $i \in [J]$,

(ii) $|A(r^{(i+1)},(\nu_p^{(i+1)}))| \geqslant (\log N)^4|A(r^{(i)},(\nu_p^{(i)}))|$ *for all* $i < J$,

(iii) $(\nu_p^{(i+1)}) \succ (\nu_p^{(i)})$ *for all* $i < J$.

Proof. Let r_1^* be such that $|A(r_1^*,(\nu_p^*))| = \max_s |A(s,(\nu_p^*))|$ so that $|A(r_1^*,(\nu_p^*))| \geqslant |A(r^*,(\nu_p^*))| \geqslant N^{1/2}$ by assumption. Let us take $r^{(J)} = r_1^*$ and $(\nu_p^{(J)}) = (\nu_p^*)$. This choice then clearly satisfies condition (i) as well as that $|A(r^{(J)},(\nu_p^{(J)}))| \geqslant (\log N)^{8J}$ for some $J \gg (\log N)/\log\log N$. Suppose now that we have created, for some $j \geqslant 1$, a sequence of $r^{(i)}$ and $(\nu_p^{(i)})$ for $j \leqslant i \leqslant J$ which satisfies (i), (ii) and (iii)

and has the additional property that $|A(r^{(j)}, (\nu_p^{(j)}))| \geqslant (\log N)^{8j}$. By the assumptions of the lemma, there must exist some $(\mu_p) \prec (\nu_p^{(j)})$ so that $\max_s |A(s, (\mu_p))| \geqslant (\log N)^{-4}|A(r^{(j)}, (\nu_p^{(j)}))| \geqslant (\log N)^{8j-4}$, and we may assume in addition that (μ_p) is a minimal element with respect to the partial order $\prec$ which satisfies the inequality $\max_s |A(s, (\mu_p))| \geqslant (\log N)^{8j-4}$.

We choose s_0 such that $|A(s_0, (\mu_p))| = \max_s |A(s, (\mu_p))|$. Hence, $|A(s_0, (\mu_p))| \geqslant (\log N)^{8j-4}$. Since $j \geqslant 1$, the assumptions of the lemma again imply that there exists some $(\mu_p') \prec (\mu_p)$ with

$$\max_s |A(s, (\mu_p'))| \geqslant (\log N)^{-4}|A(s_0, (\mu_p))| \geqslant (\log N)^{8(j-1)}.$$

Let s_0' be chosen such that $|A(s_0', (\mu_p'))| = \max_s |A(s, (\mu_p'))|$. We now claim that taking $r^{(j-1)} = s_0'$ and $(\nu_p^{(j-1)}) = (\mu_p')$ works. Clearly, (i) is satisfied and (iii) holds as $(\mu_p') \prec (\mu_p) \prec (\nu_p^{(j)})$. Note also that $|A(s_0', (\mu_p'))| = \max_s |A(s, (\mu_p'))| \geqslant (\log N)^{8(j-1)}$. Finally, the claimed upper bound $|A(s_0', (\mu_p'))| < (\log N)^{-4}|A(r^{(j)}, (\nu_p^{(j)}))|$ in (ii) follows by the choice of (μ_p) (being a minimal element with respect to $\prec$ which satisfies $\max_s |A(s, (\mu_p))| \geqslant (\log N)^{8j-4}$) and because $(\mu_p') \prec (\mu_p)$. □

Corollary 3.7.2 and Proposition 3.7.9 show that if $A \subset \mathbf{Z} \setminus \{0\}$ is a set of size N and $A \subset [-T, T]$ where $T \leqslant e^{O(\log N)^5}$, then either

- $\|F_A\|_1 \gg \log\log N$,
- or A satisfies the assumptions of Lemma 3.8.1 with $Q_1 = (\log N)^{1/2}$.

If the first alternative above holds, then the conclusion of Theorem 3.4.3 follows, so it only remains to prove this conclusion assuming that the following properties from the conclusion of Lemma 3.8.1 are satisfied. There exist $r^{(i)} \in \prod_{p \leqslant Q_1} (\mathbf{Z}/p\mathbf{Z})^\times$ and $(\nu_p^{(i)}) \in \mathbf{N}^{\pi(Q_1)}$ such that

(i) $|A(r^{(i)}, (\nu_p^{(i)}))| = \max_s |A(s, (\nu_p^{(i)}))|$,

(ii) $|A(r^{(i+1)}, (\nu_p^{(i+1)}))| \geqslant (\log N)^4 |A(r^{(i)}, (\nu_p^{(i)}))|$,

(iii) $(\nu_p^{(i+1)}) \succ (\nu_p^{(i)})$,

for all $i \in [J]$ and for some integer $J \gg (\log N)/\log\log N$. We show that under these assumptions, $\|F_A\|_1 \gg J/\log\log N \gg (\log N)/(\log\log N)^2$ and this finishes the proof because such a bound also confirms Theorem 3.4.3 (in fact with a significantly stronger bound).

We take $Q = (\log N)^{20}$ and recall, using (iv) in Proposition 3.4.1, that

$$\|F_A\|_1 \gg \|c_A + R_Q\|_1 / \log\log N,$$

where $c_A = \sum_{a\in A} c(ax)$ and

$$R_Q(x) = \sum_{a\in A} \sum_{1<n\in\mathcal{R}_Q} \frac{\chi(n)}{n} c(nax)$$

and where $\mathcal{R}_Q$ is the set of Q-rough numbers. Hence, to complete the proof of Theorem 3.4.3, it suffices to show that

$$\|c_A + R_Q\|_1 \gg J \tag{3.28}$$

assuming the existence of the sets $A(r^{(i)}, (\nu_p^{(i)}))$ satisfying (i), (ii) and (iii) above. To do this, we will use a test function Φ which satisfies $\langle c_A + R_Q, \Phi\rangle \gg J$ and $\|\Phi\|_\infty \ll 1$. Such a test function Φ will be constructed by 'going up' in the residue classes $r^{(i)} \prod_{p\leqslant Q_1} p^{\nu_p^{(i)}} \pmod{\prod p^{\nu_p^{(i)}+1}}$. In order for such a construction to work, we analyse again what the relevant projections onto each of these residue classes look like. Since we have obtained $c_A + R_Q$ from F_A by 'sifting' out all primes below $Q = (\log N)^{20}$, rather than only those in $((\log N)^{1/2}, (\log N)^{20}]$ as in the proof of Proposition 3.7.9, these projections are somewhat simpler. We claim that

$$\mathrm{Proj}\big(c_A + R_Q; r^{(i)}, (\nu_p^{(i)})\big)(x) = \frac{1}{2} \sum_{a\in A(r^{(i)},(\nu_p^{(i)}))\cup -A(-r^{(i)},(\nu_p^{(i)}))} e(ax) \tag{3.29}$$

$$+ \sum_s \sum_{a\in A(s,(\nu_p^{(i)}))} \left(\sum_{\substack{1<n\in\mathcal{R}_Q \\ n\equiv rs^{-1} (\mathrm{mod}\ \prod p)}} \frac{\chi(n)}{2n} e(anx) + \sum_{\substack{1<n\in\mathcal{R}_Q \\ n\equiv -rs^{-1} (\mathrm{mod}\ \prod p)}} \frac{\chi(n)}{2n} e(-anx) \right).$$

These projections of $c_A + R_Q$ are simpler to analyse than those of F_{med} in the proof of Proposition 3.7.9 in the sense that, here, only those $A(s, (\mu_p))$ with $(\mu_p) = (\nu_p^{(i)})$ contribute, rather than all $A(s, (\mu_p))$ with $(\mu_p) \preceq (\nu_p^{(i)})$. To see why (3.29) holds, note that $\mathrm{Proj}(c_A; r^{(i)}, (\nu_p^{(i)}))$ is precisely the first term on the right hand side. Further, if $a \in A$ and $1 < n \in \mathcal{R}_Q$ are such that $\frac{\chi(n)}{n} c(nax)$ contributes to $\mathrm{Proj}(R_Q; r^{(i)}, (\nu_p^{(i)}))$, then $na \equiv \pm r^{(i)} \prod_{p\leqslant Q_1} p^{\nu_p^{(i)}} \pmod{\prod_{p\leqslant Q_1} p^{\nu_p^{(i)}+1}}$ and hence, since n is Q-rough and $Q_1 = (\log N)^{1/2} < Q$, we need that $a \in \bigcup_s A(s, (\nu_p^{(i)}))$. We finally observe that if $a \in A(s, (\nu_p^{(i)}))$, then $\frac{\chi(n)}{n} c(nax)$ contributes if and only if $ns \equiv \pm r^{(i)} (\mathrm{mod} \prod_{p\leqslant Q_1} p)$.

The second term on the right hand side of (3.29) is of a form that we have studied in Lemma 3.6.7 which in this situation states exactly that we may write

$$\mathrm{Proj}(c_A + R_Q; r^{(i)}, (\nu_p^{(i)})) = \frac{1}{2} \sum_{a\in A(r^{(i)},(\nu_p^{(i)}))\cup -A(-r^{(i)},(\nu_p^{(i)}))} e(ax) + E^{(i)}(x), \tag{3.30}$$

for each $i \in [J]$ and where

$$\|E^{(i)} * V_T\|_2 \ll (\log N)^{-2} \left(\max_s |A(s, (\nu_p^{(i)}))|\right)^{1/2} = (\log N)^{-2} |A(r^{(i)}, (\nu_p^{(i)}))|^{1/2},$$

by property (i) of these sets $A(r^{(i)}, (\nu_p^{(i)}))$. Since we are also assuming that $A \subset [-T, T]$, we have that $c_A * V_T = c_A$ so that Young's convolution inequality (and that $\|V_T\|_1 \ll 1$) gives

$$\|c_A + R_Q\|_1 \gg \|(c_A + R_Q) * V_T\|_1 = \|c_A + R_Q * V_T\|_1.$$

Our final task was to prove (3.28) assuming the existence of the sets $A(r^{(i)}, (\nu_p^{(i)}))$ which satisfy (i), (ii) and (iii) above. It therefore suffices to show that $\|c_A + R_Q * V_T\|_1 \gg J$ and this we will deduce from the following theorem.

Theorem 3.8.2 (non-Archimedean variant of the McGehee-Pigno-Smith construction)**.** *Let $B \subset \mathbf{Z}$ be finite and suppose that for each $i \in [J]$ there exist $q_i \in \mathbf{N}$ and $r_i \in \mathbf{Z}/q_i\mathbf{Z}$ for which the following conditions hold.*

(i) Let $B_i = B \cap \{n \in \mathbf{Z} : n \equiv r_i (\mathrm{mod}\, q_i)\}$. Assume that $|B_{i+1}| > 10|B_i|$.

(ii) Let $q_1 | q_2 | \ldots | q_J$ and assume that the residue classes $\{n \in \mathbf{Z} : n \equiv r_i (\mathrm{mod}\, q_i)\}$ are pairwise disjoint.

If $E \in L^1(\mathbf{T})$ is a function such that $\|\mathrm{Proj}(E; r_i (\mathrm{mod}\, q_i))\|_2 \leqslant |B_i|^{1/2}/10$, then $\|\hat{1}_B + E\|_1 \gg J$.

Proof. We employ the same basic McGehee-Pigno-Smith construction as per usual and take

$$g_i : \mathbf{Z} \to \mathbf{C} : g_i(n) = |B_i|^{-1} 1_{n \in B_i}.$$

Then we define $Q_i(x) = e^{-|\hat{g}_i(x)|}$ and

$$\Phi_j(x) = \hat{g}_j + \hat{g}_{j-1} Q_j + \cdots + \hat{g}_1 Q_2 \ldots Q_j.$$

Lemma 3.3.3 states that $\|\Phi_j\|_1 \leqslant 10$ for all j. We shall need to prove some new properties of these test functions relating to the divisibility properties of its Fourier spectrum.

Lemma 3.8.3. *We have that $\mathrm{supp}\, (\hat{g}_i Q_{i+1} \ldots Q_j)^\wedge \subset \{n \in \mathbf{Z} : n \equiv r_i (\mathrm{mod}\, q_i)\}$ for all $j > i$.*

Proof of Lemma 3.8.3. By definition, $\mathrm{supp}(g_i) = B_i$ and hence $\mathrm{supp}(\hat{g}_i Q_{i+1} \dots Q_j)^\wedge \subset B_i + \sum_{k\in[i+1,j)} \mathrm{supp}(\hat{Q}_k)$. It therefore suffices to show that $\mathrm{supp}(\hat{Q}_k) \subset q_k \cdot \mathbf{Z}$ for all $k \in [J]$ since $q_1|q_2|\dots|q_J$. For this, we can simply observe that $\hat{g}_k(x + 1/q_k) = e(r_k/q_k)\hat{g}_k(x)$ because $\mathrm{supp}(g_k) = B_k \subset \{n : n \equiv r_k (\mathrm{mod}\, q_k)\}$ by assumption. Hence, $|\hat{g}_k|$ is a $1/q_k$-periodic function and so is $Q_k = e^{-|\hat{g}_k|}$, implying that the Fourier spectrum of Q_k consists of multiples of q_k only. □

Since $\|\Phi_J\|_1 \leqslant 10$, we can obtain a lower bound

$$\begin{aligned}\|\hat{1}_B + E\|_1 &\gg \langle \hat{1}_B + E, \Phi_J\rangle \\ &= \sum_{j=1}^{J} \langle \hat{1}_B, \hat{g}_j\rangle - \sum_{1\leqslant j<k\leqslant J} \langle \hat{1}_B, \hat{g}_j Q_{j+1}\dots Q_{k-1}(1-Q_k)\rangle + \langle E, \Phi_J\rangle.\end{aligned}$$

By Parseval, $\langle \hat{1}_B, \hat{g}_j\rangle = 1$ for all j so the first term above contributes J. By Lemma 3.8.3, we have

$$\langle \hat{1}_B, \hat{g}_j Q_{j+1}\dots Q_{k-1}(1-Q_k)\rangle = \langle \mathrm{Proj}(\hat{1}_B; r_j (\mathrm{mod}\, q_j)), \hat{g}_j Q_{j+1}\dots Q_{k-1}(1-Q_k)\rangle$$

and hence, using that by Lemma 3.3.3 we have the inequalities $|Q_{j'}| \leqslant 1$, $|1 - Q_k| \leqslant |\hat{g}_k|$, and $|\hat{g}_j| \leqslant 1$, we can use Cauchy-Schwarz to bound the second term by

$$\begin{aligned}\sum_{1\leqslant j<k\leqslant J} \|\mathrm{Proj}(\hat{1}_B; r_j (\mathrm{mod}\, q_j))\|_2 \|\hat{g}_k\|_2 &= \sum_{1\leqslant j<k\leqslant J} |B_j|^{1/2}|B_k|^{-1/2} \\ &\leqslant \sum_{1\leqslant j<k\leqslant J} 10^{-(k-j)/2} \leqslant J/(10^{1/2}-1),\end{aligned}$$

where we used the assumption that $|B_{i+1}| > 10|B_i|$. Hence, $\|\hat{1}_B + E\|_1 \gg J/2 - \langle E, \Phi_J\rangle$.

To finish the proof, we therefore only need to show that $\langle E, \Phi_J\rangle \leqslant J/10$. This will follow from the fact that $\langle E, \hat{g}_j Q_{j+1}\dots Q_J\rangle$ has size at most $1/10$ for all j. One can prove this by a single application of Cauchy-Schwarz:

$$\begin{aligned}|\langle E, \hat{g}_j Q_{j+1}\dots Q_J\rangle| &= |\langle \mathrm{Proj}(E; r_j (\mathrm{mod}\, q_j)), \hat{g}_j Q_{j+1}\dots Q_J\rangle| \\ &\leqslant \|\mathrm{Proj}(E; r_j (\mathrm{mod}\, q_j))\|_2 \|\hat{g}_j\|_2 \leqslant 1/10\end{aligned}$$

where we used that $\|\hat{g}_j\|_2 = |B_j|^{-1/2}$ and that $\|\mathrm{Proj}(E; r_j (\mathrm{mod}\, q_j))\|_2 \leqslant |B_j|^{1/2}/10$ by assumption.

□

We apply this theorem with the set $B = A \cup -A$, moduli $q_i = \prod_{p \leqslant Q_1} p^{\nu_p^{(i)}+1}$, and residues $r_i = r^{(i)} \prod_{p \leqslant Q_1} p^{\nu_p^{(i)}}$. We also take $E = R_Q * V_T$ so that by (3.30) and the inequality right after, we get that

$$B_i = A(r^{(i)}, (\nu_p^{(i)})) \cup -A(-r^{(i)}, (\nu_p^{(i)}))$$
$$\|\mathrm{Proj}(E; r_i (\mathrm{mod}\, q_i))\|_2 = \|E^{(i)} * V_T\|_2 \ll (\log N)^{-2} |B_i|^{1/2}$$

and note furthermore that $q_1 | q_2 | \dots | q_J$ by property (iii). It is also immediate from property (ii) that $|B_{i+1}| \gg (\log N)^4 |B_i|$ for all $i < J$. The theorem above now confirms the claimed bound (3.28): $\|c_A + R_Q\|_1 \gg J \gg (\log N)/\log\log N$.

3.9 The global structure of sets with $S(A) \leqslant N/3 + C$

The methods that we have introduced to prove Theorem 3.1.2 can be exploited further to provide structural information about sets of integers A with $S(A) \leqslant N/3 + C$ for values of C much larger than $\log\log N$. In this section, we shall focus on proving Theorem 3.1.3, although there also are other types of 'structure' that one may provably find in A.

Proposition 3.9.1. *Let $A \subset \mathbf{Z} \setminus \{0\}$ have size N and let $S(A) \leqslant N/3 + C$. Then there exists a set B which is F_4-isomorphic to A and which is contained in $[-T, T]$ where $T \leqslant N^{C^{O(1)}}$. Moreover, every subset $X \subset A$ satisfies the energy bound $E(X) \gg C^{-O(1)} \frac{|X|^4}{N}$.*

Proof. By Corollary 3.6.3, we may find an F_4-isomorphic copy A' of A with $A' \subset [-T_1, T_1]$ where $T_1 \leqslant e^{O((C \log N)^4)}$. By taking $Q = (C \log N)^{10}$ (say) in Proposition 3.4.1, we find that $\|c_{A'} + R_Q\|_1 \ll (\log Q)\|F_{A'}\|_1 \ll (\log Q)(S(A') - N/3) \ll C^2$ where we have used in the final inequality that $\log Q \ll C$ since $C \gg \log\log N$. In fact, it will be convenient to note that the exact same proof from Proposition 3.4.1 provides an analogous bound for every L^p-norm:

$$\|c_{A'} + R_Q\|_p \ll (\log Q)\|F_{A'}\|_p.$$

We know from Lemma 3.6.5 that

$$R_Q(x) = \sum_{a \in A'} \sum_{1 < n \in \mathcal{R}_Q} \frac{\chi(n)}{n} c(nax)$$

satisfies the bound $|\hat{R}_Q(m)| \ll (\log T) Q^{-1} \ll (C \log N)^{-5}$ for its Fourier coefficients with frequencies $m \in [-T_1, T_1]$. In particular, if we define $f(n) := 1_{A'}(n) + 1_{-A'}(n) + 2\hat{R}_Q(n)$, then we have shown that

- $\|\hat{f}\|_1 \ll C^2$,
- $f(a) \geqslant 1/2$ for all $a \in A'$.

One can also see by taking $p = \infty$ in the L^p-inequality above and by the trivial bound $|F_{A'}(x)| \leqslant N \max_x |\phi(x) - 1/3| \leqslant N$ that $\|c_{A'} + R_Q\|_\infty \ll NC$, and hence we certainly have $\|\hat{f}\|_\infty \ll N^2$. Theorem 3.5.1 states under these assumptions on f that $\dim(A') \ll C^4(\log N)$. One may now simply use Theorem 3.6.2 to find the desired 'dense' F_4-isomorphic copy $B \subset [-N^{C^{O(1)}}, N^{C^{O(1)}}]$ (again using that $\log\log N \ll C$).

Let $\psi : A \to A'$ be the F_4-isomorphism from above. For any subset $X \subset A$ we observe that $E(X) = E(\psi(X))$ precisely because ψ preserves all additive relations of length at most 4. We write $X' = \psi(X)$ and to establish the final part of this proposition, it suffices to show that $E(X') \gg C^{-O(1)}\frac{|X'|^4}{N}$. Note that the function $f = 1_{A'} + 1_{-A'} + 2\hat{R}_Q$ from above retains its crucial properties after convolving with the de la Vallée-Poussin kernel V_{T_1}:

- $\|\hat{f} * V_{T_1}\|_1 \ll C^2$,
- $(\hat{f} * V_{T_1})^\wedge(a) \geqslant 1/2$ for all $a \in A'$, since $A' \subset [-T_1, T_1]$.

Hence, $\hat{f} * V_{T_1}$ satisfies the assumptions of Corollary 3.5.5 and we deduce that

$$E(X') \gg C^{-O(1)}\frac{|X'|^4}{\|\hat{f} * V_{T_1}\|_2^2}.$$

Hence, our final task is to find a good bound for the L^2-norm of $(c_{A'} + R_Q) * V_{T_1} = c_{A'} + R_Q * V_{T_1}$. By Parseval, $\|c_{A'}\|_2 \ll N^{1/2}$. Finally, an even stronger bound $\|R_Q * V_{T_1}\|_2 \ll (\log N)^{-2}N^{1/2}$ may be deduced from Lemma 3.6.7 (with $Q_1 = 1$). □

With a result like the proposition above in hand, which shows that any large subset of A has large additive energy, one can invoke the standard tools of additive combinatorics to obtain Theorem 3.1.3; we briefly indicate how this is done. We shall make use of two central results from additive combinatorics from Chapter 2, namely the Balog-Szemerédi-Gowers Theorem 2.0.4 and Freiman's Theorem 2.0.2.

Proof of Theorem 3.1.3. Suppose that we have obtained a partial structured decomposition $A = (\cup_{i<j} A_i) \cup B$ where each A_i has size $|A_i| \gg (CK)^{-O(1)}N$ and doubling $|A_i - A_i| \ll (CK)^{O(1)}$. Freiman's Theorem implies that A_i is contained in some generalised arithmetic progression P_i of dimension at most $(CK)^{O(1)}$ and size $|P_i| \ll e^{(CK)^{O(1)}}|A_i|$. The set B either has size $|B| < (CK)^{-10}N$ in which case we have

found the desired decomposition of A, or else the energy bound from the proposition above yields

$$E(B) \gg C^{-O(1)}\frac{|B|^4}{N} \gg (CK)^{-O(1)}|B|^3.$$

The Balog-Szemerédi-Gowers theorem tells us that B contains a subset B' of size $|B'| \geqslant (CK)^{-O(1)}|B| \gg (CK)^{-O(1)}N$ with small doubling $|B'-B'| \ll (CK)^{O(1)}$. So we take $A_j := B'$ and we have obtained a new decomposition $A = (\cup_{i\leqslant j}A_i) \cup B^*$, where A_j satisfies the same properties as the A_i with $i < j$ and where $|B| - |B^*| \gg (CK)^{-O(1)}N$. Clearly this process terminates (in fact after at most $(CK)^{O(1)}$ steps), giving the desired decomposition of A. □

3.10 The McGehee-Pigno-Smith test function

The purpose of this section is to provide a proof of Theorem 3.7.8.

Proof of Theorem 3.7.8. Let us define $f(n) = 1_{B_1} + 1_{B_2}$ so that we aim to show that $\|\hat{f} + E\|_1 \gg \log K$ under the assumption that $\|E\|_2 \leqslant |B_1|^{1/2}/K$. To prove this, we use the method of McGehee-Pigno-Smith to construct a test function Φ satisfying $\|\Phi\|_\infty \ll 1$ and $\langle \Phi, \hat{f} + E\rangle \gg \log K$.

Let us define the sets $A_1, A_2, \ldots, A_J$ to be subsets of $B_1 \cup B_2$ satisfying the following:

- A_i consists of the $100^i\lfloor |B_1 \cup B_2|/K^2\rfloor$ smallest integers in $(B_1 \cup B_2) \setminus \big(\cup_{j=1}^{i-1}A_i\big)$. In particular, $|A_{i+1}| = 100|A_i|$ for all $i \in [J]$.
- $J \gg \log K$.

We define for each $i \in [J]$ the basic function $g_i : A_i \to \mathbf{C}$ by $g_i(n) = |A_i|^{-1}1_{\{n\in A_i\}}$. As usual, this has the following properties

$$\begin{aligned} \mathrm{supp}(g_i) &\subseteq A_i \\ \|\hat{g}_i\|_\infty &\leqslant 1 \\ \langle \hat{f}, \hat{g}_i\rangle &\geqslant 1. \end{aligned} \tag{3.31}$$

The function $|\hat{g}_i(x)| = \sum_{n\in\mathbf{Z}} c_i(n)e(nx)$ is even so has a Fourier series with $c_i(n) = c_i(-n)$. We then define for each $i \in [J]$ the following function

$$h_i(x) = c_i(0) + 2\sum_{n<0} c_i(n)e(nx)$$

and we also define the 'correction' function $Q_i(x) = \exp(-h_i(x))$ and we note the following properties.

Lemma 3.10.1. *The function h_i satisfies*

$$\begin{aligned} \Re\left(h_i(x)\right) &= |\hat{g}_i(x)| \\ \operatorname{supp} \hat{h}_i &\subset \mathbf{Z}_{\leqslant 0} \\ \|h_i\|_2 &\leqslant 2\|\hat{g}_i\|_2. \end{aligned}$$

The function Q_i satisfies

$$\begin{aligned} \operatorname{supp} \hat{Q}_i &\subset \mathbf{Z}_{\leqslant 0} \\ |Q_i(x)| &\leqslant 1 \\ |1 - Q_i(x)| &\leqslant |h_i(x)|. \end{aligned}$$

Proof. Consider the Fourier expansions $|\hat{g}_i(x)| = \sum_{n\in\mathbf{Z}} c_i(n)e(nx)$ and

$$h_i(x) = c_i(n) + 2\sum_{n<0} c_i(n)e(nx).$$

One can see from this and that $c_i(n) = c_i(-n)$ that $\Re\left(h_i(x)\right) = \sum_{n\in\mathbf{Z}} c_n e(nx) = |\hat{g}_i(x)|$, that $\operatorname{supp} \hat{h}_i \subset \mathbf{Z}_{\leqslant 0}$ and that $\|h_i\|_2 \leqslant 2\|\hat{g}_i\|_2$. Now let us consider $Q_i(x) = e^{-h_i(x)}$. That $|Q_i(x)| \leqslant 1$ follows from the bound $|Q_i(x)| = e^{-\Re h_i(x)} = e^{-|\hat{g}_i(x)|}$. To show that $|1 - Q_i(x)| \leqslant |h_i(x)|$ we may take $z = h_i(x)$ in the simple inequality $|1 - e^{-z}| \leqslant |z|$ which is valid for all $z \in \mathbf{C}$ with $\Re(z) \geqslant 0$.[5] Finally, $Q_i(x) = e^{-h_i(x)} = \sum_{k\geqslant 0}(-h_i(x))^k/k!$ has a Fourier expansion whose terms all have non-positive frequencies because $\operatorname{supp} \hat{h}_i \subset \mathbf{Z}_{\leqslant 0}$ and hence $\operatorname{supp}(h_i^k)^\wedge \subset \mathbf{Z}_{\leqslant 0}$. This shows that $\operatorname{supp} \hat{Q}_i \subset \mathbf{Z}_{\leqslant 0}$.

□

We now use the iterative McGehee-Pigno-Smith construction

$$\begin{aligned} \Phi_1(x) &= \hat{g}_1(x) \\ \Phi_{j+1}(x) &= \hat{g}_{j+1}(x) + Q_{j+1}\Phi_j(t), \end{aligned}$$

and the same proof as in Lemma 3.3.3 may be used to deduce that $\|\Phi_i\|_\infty \leqslant 10$ for all $j \leqslant J$. A telescoping identity shows that

$$\begin{aligned} \Phi_J(x) &= \hat{g}_J + Q_J\hat{g}_{J-1} + \cdots + Q_2 \ldots Q_J\hat{g}_1 \\ &= \sum_{j=1}^{J} \hat{g}_j(x) - \sum_{j=1}^{J-1}\sum_{k=j+1}^{J} \hat{g}_j(x)(1 - Q_k)Q_{k+1}\ldots Q_J. \end{aligned}$$

[5] One can prove this by observing that $1 - e^{-z} = \int_0^z e^{-w}\,dw$ and that $|e^{-w}| \leqslant 1$ for all w on a straight line path from 0 to z if $\Re z \geqslant 0$.

As $\|\Phi_J\|_\infty \leqslant 10$, we have

$$\begin{aligned}\|\hat{f}+E\|_1 &\gg \langle \hat{f}+E, \Phi_J\rangle \\ &= \sum_{j=1}^{J} \langle \hat{f}, \hat{g}_j\rangle - E_1 + E_2 \\ &\geqslant J - |E_1| - |E_2|, \end{aligned} \tag{3.32}$$

where we used the last equation in (3.31) to get $\langle \hat{f}, \hat{g}_j\rangle \geqslant 1$ for each j, and where we defined

$$\begin{aligned} E_1 &= \sum_{1\leqslant j<k\leqslant J} \langle \hat{f}, \hat{g}_j(1-Q_k)Q_{k+1}\dots Q_J\rangle \\ E_2 &= \sum_{j=1}^{J} \langle E, \hat{g}_j Q_{j+1}\dots Q_J\rangle. \end{aligned}$$

We proceed by bounding E_1, E_2. To bound E_2 we simply recall that $|Q_j| \leqslant 1$ so that by Cauchy-Schwarz $|\langle E, \hat{g}_j Q_{j+1}\dots Q_J\rangle| \leqslant \|E\|_2\|\hat{g}_j\|_2 \leqslant 100^{-j/2}$ as we are assuming that $\|E\|_2 \leqslant |B_1|^{1/2}/K$ while $\|\hat{g}_j\|_2 = |A_j|^{-1/2} \leqslant 100^{-j/2}|B_1\cup B_2|^{-1/2}K$ by our choice of the A_j. Hence, $|E_2| < 1$.

Bounding E_1 follows the classical McGehee-Pigno-Smith argument. By Lemma 3.10.1, the Fourier transform of $(1-Q_k)Q_{k+1}\dots Q_J$ is supported on $\mathbf{Z}_{\leqslant 0}$ and hence

$$\operatorname{supp}\left(\hat{g}_j(1-Q_k)Q_{k+1}\dots Q_J\right)^\wedge \subset \mathbf{Z}\cap(-\infty, \max A_j].$$

We therefore get

$$\begin{aligned} &\langle \hat{f}, \hat{g}_j(1-Q_k)Q_{k+1}\dots Q_J\rangle \\ &= \langle \sum_{n\in B_1 : n\leqslant \max A_j} e(nx) + \sum_{n\in B_2 : n\leqslant \max A_j} e(nx), \hat{g}_j(1-Q_k)Q_{k+1}\dots Q_J\rangle \\ &\leqslant 2|(B_1\cup B_2)\cap(-\infty, \max A_j]|^{1/2}\|h_k\|_2 \end{aligned}$$

by Cauchy-Schwarz and Lemma 3.10.1. Using that $|(B_1\cup B_2)\cap(-\infty, \max A_j]| \leqslant 101|A_j|/100$ by our choice of the A_j, and that $\|h_k\|_2 \leqslant 2|A_k|^{-1/2}$, we obtain the bound $5|A_j|^{1/2}|A_k|^{-1/2} = 5\cdot 100^{-(k-j)/2}$ for the inner product above. In total, we get

$$|E_1| \leqslant 5\sum_{1\leqslant j<k\leqslant J} 100^{-(k-j)/2} \leqslant 5J/9.$$

Finally, we may substitute this estimate in (3.32) and recall that $J \gg \log K$ to obtain

$$\|\hat{f}+E\|_1 \gg 4J/9 - 1 \gg \log K,$$

which is the required conclusion.

□

Chapter 4

On unique sums in Abelian groups

4.1 Introduction

This chapter is based on [6]. Let A be a subset of a finite Abelian group G. We say that A has a unique sum if there exist a_1, a_2 in A so that the only solutions to $x + y = a_1 + a_2$ with $x, y \in A$ are the trivial ones $(x, y) = (a_1, a_2), (a_2, a_1)$. In this case, we say that $a_1 + a_2$ is a unique sum in $A + A$. In this chapter, we will study the conditions under which a set A must contain a unique sum. In particular, given any finite Abelian group G, we want to determine the size of the smallest subset of G having no unique sum.

Definition 4.1.1. Let G be a finite Abelian group. Then we define $m(G)$ to be the size of the smallest subset of G which has no unique sum. Equivalently, $m(G)$ is the smallest integer so that any subset $B \subset G$ with size $|B| < m(G)$ has a unique sum. Of special importance is the case where $G = \mathbf{Z}/p\mathbf{Z}$ is the cyclic group of prime order p so we abbreviate the notation and write $m(p)$ for $m(\mathbf{Z}/p\mathbf{Z})$.

The question of estimating $m(p)$ was explicitly asked by S. Kopparty (open problems session, Harvard 2017) and it also appears as Problem 27 on B. Green's list of 100 open problems [36]. Questions of this type go back at least to a paper of Straus [63] in which he proved the first bounds on the size $f(p)$ of the smallest subset $A \subset \mathbf{Z}/p\mathbf{Z}$ having no unique difference. Here, we say that A contains a unique difference if there exist $a_1, a_2 \in A$ such that the only solution to $x - y = a_1 - a_2$ with $x, y \in A$ is the trivial one $(x, y) = (a_1, a_2)$. Straus proved that $f(p) \geqslant 1 + \log_4(p-1)$ and this was later improved by Browkin, Diviš and Schinzel [13] who obtained the following.

Theorem 4.1.2 (Browkin-Diviš-Schinzel, [13])**.** *Let p be prime and $A, B \subset \mathbf{Z}/p\mathbf{Z}$.*

(i) If $p > \min\left(2^{|A|+|B|-2}, |A|^{|B|-1}, |B|^{|A|-1}\right)$*, then* $A+B$ *contains a unique sum.*[1]

(ii) If $p > 2^{|A|-1}$*, then* A *has a unique difference and a unique sum.*

Their result was extended to general Abelian groups by Lev.

Theorem 4.1.3 (Lev, [46])**.** *Let* A, B *be subsets of a finite abelian group* G *and let* $p(G)$ *be the smallest prime divisor of* $|G|$*.*

(i) If $p(G) > 2^{|A|+|B|-3}$*, then* $A+B$ *contains a unique sum.*

(ii) If $p(G) > 2^{|A|-1}$*, then* A *has a unique difference and a unique sum.*

The current best bound for unique sums in $A+B$ is due to Leung and Schmidt [45], who recently proved under the same assumptions as in Theorem 4.1.3 that $A+B$ contains a unique sum if $p(G) > (\sqrt[4]{12})^{|A|+|B|-2}$. Closely related problems, such as estimating the size of the smallest $A \subset \mathbf{Z}/p\mathbf{Z}$ so that any sum in $A+A$ has at least K distinct representations, or alternatively such that $\underbrace{A+\cdots+A}_{k}$ has no unique sum, have also been studied, for a selection see [15, 39, 48, 55].

These bounds show that the size $f(G)$ of the smallest subset of G with no unique difference satisfies $f(G) \gg \log p(G)$, and examples of sets with no unique difference and size $O(\log p(G))$ do exist and already appear in Straus's original paper [63]. Hence, we have $f(G) = \Theta(\log p(G))$. For the problem of determining the size $m(G)$ of the smallest $A \subset G$ having no unique sum, the results above provide a lower bound of the shape $m(G) \geqslant C \log p(G)$ for some absolute constant $C > 0$, which is the current record lower bound. Unlike the situation for sets with no unique difference, there are no constructions known of sets with no unique sum and size $O(\log p(G))$. The following two theorems are our main results proving that such examples cannot exist and they are the first lower bounds on $m(G)$ replacing the constant C in the bound above by a function tending to infinity with p.

Theorem 4.1.4. *There is a function* $\omega(n)$ *which tends to infinity as* $n \to \infty$ *such that the following holds. Let* p *be a prime, then* $m(p) \geqslant \omega(p) \log p$*. In fact, one can take*

$$\omega(n) \gg \frac{\sqrt{\log\log\log n}}{\log\log\log\log n}.$$

In particular, if $B \subset \mathbf{Z}/p\mathbf{Z}$ *has size* $|B| < \omega(p)\log p$*, then* B *has a unique sum.*

[1] Here, we say that $A+B$ has a unique sum if there are $a \in A, b \in B$ such that the only solutions to $a+b = x+y$ with $x \in A, y \in B$ are $(x,y) = (a,b)$, and $(x,y) = (b,a)$ if $a, b \in A \cap B$.

Our goal is to obtain a lower bound with $\omega(n) \to \infty$ and we have not tried to optimise the exact shape of ω, which can certainly be improved. To analyse $m(G)$ for general Abelian groups, we begin with the simple observation that if p is a prime dividing the order of G, then G contains a cyclic group of order p as a subgroup. Thus for a general Abelian group G we have

$$m(G) \leqslant \min_{p \text{ prime, } p||G|} m(p).$$

Hence, there is no hope of proving a better lower bound on the size of a subset of G having no unique sum than those holding in cyclic groups $\mathbf{Z}/p\mathbf{Z}$ with $p||G|$. The following theorem shows that we can get a lower bound of the form of Theorem 4.1.4 in general.

Theorem 4.1.5. *Let G be a finite Abelian group and let $p(G)$ be the smallest prime factor of $|G|$. If $A \subset G$ has no unique sum, then*

$$|A| \geqslant \omega(p(G)) \log p(G),$$

where ω is the same function as in Theorem 4.1.4.

We also improve on the best known upper bound on $m(p)$ by constructing for each prime p a set A which has no unique sum and size $O((\log p)^2)$. This improves the previous best known bound $m(p) \ll \sqrt{p}$ which came from a rather easy construction of a set $A \subset \mathbf{Z}/p\mathbf{Z}$ whose sumset $A + A$ is the whole of $\mathbf{Z}/p\mathbf{Z}$.

Theorem 4.1.6. *Let p be a prime, then $m(p) \ll (\log p)^2$. That is, for every prime p there is a set A of size $O((\log p)^2)$ having no unique sum.*

It is clear that this implies the corresponding bound $m(G) \ll (\log p(G))^2$ for general Abelian groups G.

4.2 Prerequisites

In this chapter, G denotes an Abelian group and we shall always write $+$ for the group operation. We write $p(G)$ for the smallest prime factor of $|G|$. To improve readability, we omit floor and ceiling functions throughout the chapter, but it will be clear from context which quantities should be integer-valued. For an element $g \in G$, recall that $r_A(g)$ denotes the number of ordered pairs in A^2 whose sum equals g, so

$$r_A(g) := \left|\{(a, a') \in A^2 : a + a' = g\}\right|.$$

So a set $A \subset G$ has a unique sum if and only if there is some g such that $1 \leqslant r_A(g) \leqslant 2$. We continue by further recalling the rectification result 2.0.8 of Bilu, Lev and Ruzsa which allows one to rectify any subset of $\mathbf{Z}/p\mathbf{Z}$ of size at most $\log_2 p$. We require the following useful generalised rectification lemma of Lev which proves such a statement in an arbitrary Abelian group.

Lemma 4.2.1 (Lev, [46] Theorem 1)**.** *Let G be a finite Abelian group and let $p(G)$ denote the smallest prime dividing $|G|$. If $Z \subset G$ has size $|Z| \leqslant \log_2 p(G)$, then Z is Freiman-isomorphic to a set of integers.*

We show now how one can easily recover the previous best known lower bound $m(G) \gg \log p(G)$ using this lemma. Indeed, suppose that $A \subset G$ has no unique sum. Then by definition, neither does any set that is Freiman-isomorphic to A. In particular, A cannot be rectifiable since any finite set of integers A' trivially has a unique sum, namely $\max A' + \max A'$. Thus, Lemma 4.2.1 implies that $|A| > \log_2 p(G)$ as desired. Such arguments using rectification to find a unique sum go back to Straus's original paper [63], and note that one can deduce Theorem 4.1.3 from Lemma 4.2.1 in this way.

Outline of the proof. In section 3, we prove a general structural result about sets $Z \subset G$ with large additive span, by which we mean that $\Sigma(Z) = \left\{\sum_{z' \in Z'} z' : Z' \subset Z\right\}$ is large. To be precise, we show that if $|\Sigma(Z)|$ is large, then Z contains a large dissociated subset. In other words, we show in section 3 that sets with large additive span have large additive dimension. In section 4, we show that if $A \subset G$ is a set having no unique sum, then some translate $A+g$ of A has very large additive span in the sense that $\Sigma(A+g)$ contains a non-trivial subgroup of G. For the most interesting case where $G = \mathbf{Z}/p\mathbf{Z}$, this shows that $\Sigma(A+g) = \mathbf{Z}/p\mathbf{Z}$ is the whole group.

Combining the results from sections 3 and 4, we obtain that if A has no unique sum, then some translate $A+g$ has large additive dimension. Note that $A' = A+g$ also contains no unique sum. Finally, in section 5 we employ a density increment argument to prove that a set A' with no unique sum cannot contain a dense dissociated subset, i.e. we show that $\dim(A') = o_{p(G)}(1) \cdot |A'|$ as $p(G) \to \infty$. As sections 3 and 4 imply that $\dim(A')$ is large, this will yield the required lower bound $|A| = |A'| \geqslant \omega(p(G)) \log p(G)$.

4.3 Sets with small dimension have small additive span

In this short section, we will prove an inequality that holds for any subset Z of an Abelian group G. The proof of this result is self-contained and one can forget about sets having no unique sum in this whole section. We begin with four important definitions from additive combinatorics, the first two of which should be familiar from Chapter 2 .

Definition 4.3.1. *Let G be a finite Abelian group and let $S \subset G$.*

- We say that S is *dissociated* if whenever there exist $(\mu_s)_{s\in S} \in \{-1,0,1\}^S$ so that $\sum_{s\in S} \mu_s s = 0$, then $\mu_s = 0$ for all $s \in S$. Equivalently, S is dissociated if whenever $S_1, S_2 \subset S$ with $\sum_{s\in S_1} s = \sum_{s\in S_2} s$, then $S_1 = S_2$.
- We define the *additive dimension* $\dim(S)$ of S to be the size of the largest dissociated subset of S.
- We define the *additive span* of S to be the set

$$\Sigma(S) := \left\{ \sum_{s\in S} \varepsilon_s s : \varepsilon_s \in \{0,1\} \right\} = \left\{ \sum_{s\in S'} s : S' \subseteq S \right\}. \tag{4.1}$$

  This definition also makes sense when S is a finite multiset consisting of elements of G. In general, every $s \in S$ appears k times in the sum $\sum_{s\in S} \varepsilon_s s$ in (4.1) if S contains k copies of s. We say that a (multi-)set $S \subset G$ is an *additive basis* for G if $\Sigma(S) = G$. In other words, S is an additive basis if for every element $g \in G$, there is some (multi-)subset $S_g \subseteq S$ whose elements sum to g.

Our aim in this section is to find an upper bound on $|\Sigma(Z)|$. Observe that the trivial bound $|\Sigma(Z)| \leqslant 2^{|Z|}$ always holds. In general, one can of course not improve on this trivial bound as $|\Sigma(Z)| = 2^{|Z|}$ if Z is a dissociated set. Similarly, the additive span $\Sigma(Z)$ will be large if Z contains a fairly large subset which is dissociated. It is therefore natural to wonder if this is in a sense the only reason why $\Sigma(Z)$ can be large, meaning that if $\Sigma(Z)$ is large then it implies that Z contains a large dissociated subset. The following proposition shows that this result is indeed true.

Proposition 4.3.2. *Let G be an Abelian group and let Z be a finite multiset consisting of elements of G. Then*

$$|\Sigma(Z)| \leqslant \binom{|Z|}{\dim(Z)} \binom{|Z| + \dim(Z)}{\dim(Z)}. \tag{4.2}$$

Hence, if $|\Sigma(Z)|$ is large, then $\dim(Z)$ is large which means precisely that Z has a large dissociated subset. We state a bound which is more useful in practice.

Corollary 4.3.3. *Let G be an Abelian group and let Z be a finite multiset consisting of elements of G. Then*

$$|\Sigma(Z)| \leqslant 2^{2\dim(Z)\cdot\left(\log_2\left(\frac{|Z|}{\dim(Z)}\right)+2\right)} = \left(\frac{4|Z|}{\dim(Z)}\right)^{2\dim(Z)}. \tag{4.3}$$

Clearly, we always have the lower bound $|\Sigma(Z)| \geqslant 2^{\dim(Z)}$ and one may ask if the extra factor $\log_2\left(\frac{|Z|}{\dim(Z)}\right)$ in the exponent in (4.3) is necessary. The following example shows that in fact it is necessary. Pick an integer d and consider $G = \mathbf{Z}^d$ with standard generating set $\{e_1, e_2, \dots, e_d\}$. Then we can take Z to be the the multiset consisting of k copies of each e_i with $1 \leqslant i \leqslant d$. It is easy to see that $\dim(Z) = d$, but $\Sigma(Z) = \left\{\sum_{i=1}^d n_i e_i : 0 \leqslant n_i \leqslant k \text{ for each } i\right\}$ has size $(k+1)^d > \left(\frac{|Z|}{d}\right)^d$. Let us now give the proof of Proposition 4.3.2.

Proof of Proposition 4.3.2. We consider an element $y \in \Sigma(Z)$, so we may find coefficients $\varepsilon_z(y) \in \{0, 1\}$ for $z \in Z$ so that

$$y = \sum_{z\in Z} \varepsilon_z(y) z, \tag{4.4}$$

where each z in the multiset Z occurs with multiplicity in this sum. Our idea is to use a type of compression on these sums until every y can be expressed as a sum of elements in Z with small support. We will use a different type of compression in a later section, so to avoid confusion we call the type of compressions used in this section 'support-compressions'. For an expression $y = \sum_{z\in Z} n_z z$ with non-negative integers n_z which is not already maximally compressed, a 'support-compression' yields a new expression $y = \sum_z m_z z$ with smaller support, i.e. the multiset $\{z \in Z : m_z \neq 0\}$ is smaller than $\{z \in Z : n_z \neq 0\}$. Repeatedly applying this shows that every y in $\Sigma(Z)$ can be expressed as a sum of elements of Z of the form $y = \sum_{z\in Z} n_z z$ whose support is a dissociated subset of Z and with $\sum_z n_z$ not too large. A combinatorial counting argument then yields (4.2).

We make this argument precise. For each $y \in \Sigma(Z)$, define

$$a(y) := \sum_{z\in Z} \varepsilon_z(y) \in \{0, 1, \dots, |Z|\} \tag{4.5}$$

where the $\varepsilon_z(y) \in \{0,1\}$ are so that (4.4) holds (if there is more than one choice, we just pick one of these arbitrarily). We now define the following set

$$T(y) := \Big\{(n_z)_{z\in Z} \in \mathbf{N}^Z : y = \sum_{z\in Z} n_z z \text{ and } \sum_{z\in Z} n_z \leqslant a(y)\Big\}. \tag{4.6}$$

For each $|Z|-$tuple (n_z) in $T(y)$, we define its support-size as follows

$$\operatorname{supp}((n_z)_{z\in Z}) := |\{z \in Z : n_z \neq 0\}| \tag{4.7}$$

counted with multiplicity if Z is a multiset. We begin by noting that the set $T(y)$ is non-empty because $y = \sum_{z\in Z} \varepsilon_z(y) z$ by (4.4) and $\sum_z \varepsilon_z(y) = a(y)$ by (4.5) so $(\varepsilon_z(y))_{z\in Z} \in T(y)$. Hence, we can consider an element $(k_z(y))_{z\in Z} \in T(y)$ with minimal support-size $\operatorname{supp}((k_z(y))_{z\in Z})$. Let $K = K(y) = \{z \in Z : k_z(y) \neq 0\}$ be its support, so K is a multisubset of Z and we obtain the following information about K.

Lemma 4.3.4. *Let $y \in \Sigma(Z)$ and let $(k_z)_{z\in Z} \in T(y)$ be chosen to minimise $\operatorname{supp}((k_z)_{z\in Z})$ over all elements of $T(y)$. Then $K = \{z \in Z : k_z \neq 0\}$ is a dissociated subset of Z.*

Proof. Suppose for a contradiction that the multiset K is not dissociated. Hence there exist distinct multisubsets K_1, K_2 of K so that $\sum_{z\in K_1} z = \sum_{z\in K_2} z$. We may further assume that K_1 and K_2 are disjoint as removing common elements in $K_1 \cap K_2$ from both multisets does not change that both multisets have equal sum. Also assume that $|K_1| \geqslant |K_2|$ and define $k^- = \min_{z\in K_1} k_z$. Then we can write

$$\begin{aligned} y &= \sum_{z\in K} k_z z \\ &= \sum_{z\in K\setminus(K_1\cup K_2)} k_z z + \sum_{z\in K_1} (k_z - k^-) z + \sum_{z\in K_2} (k_z + k^-) z \end{aligned} \tag{4.8}$$

so we can construct a new tuple $(k'_z)_{z\in Z}$ as follows by defining:

$$k'_z = \begin{cases} k_z, & \text{if } z \in Z \setminus (K_1 \cup K_2) \\ k_z - k^-, & \text{if } z \in K_1 \\ k_z + k^-, & \text{if } z \in K_2. \end{cases} \tag{4.9}$$

We proceed by showing that $(k'_z)_{z\in Z} \in T(y)$. First, it is clear from the definition (4.9) that each k'_z is a non-negative integer as $(k_z)_{z\in Z} \in T(y)$ and $k^- \leqslant k_z$ for all $z \in K_1$. From (4.8), we see that

$$y = \sum_{z\in Z} k'_z z.$$

Finally, from (4.9) we observe that

$$\sum_{z\in Z} k'_z = \sum_{z\in Z} k_z - k^-|K_1| + k^-|K_2|$$

$$\leqslant \sum_{z\in Z} k_z$$
$$\leqslant a(y)$$

using that $|K_1| \geqslant |K_2|$. Hence, $(k'_z)_{z\in Z} \in T(y)$. Our final task to obtain the required contradiction is to show that $\mathrm{supp}((k'_z)) < \mathrm{supp}((k_z))$. This is clear from (4.9) however as we defined $k^- = \min_{z\in K_1} k_z$. We have obtained the required contradiction as (k_z) was chosen to minimise $\mathrm{supp}((k_z))$ over all sequences in $T(y)$. Hence, K must be dissociated. □

We continue with the proof of Proposition 4.3.2. By Lemma 4.3.4, every $y \in \Sigma(Z)$ can be written as

$$y = \sum_{z\in K} k_z(y) z \tag{4.10}$$

for some dissociated set $K = K(y) \subset Z$ and with $\sum_{z\in K} k_z(y) \leqslant a(y) \leqslant |Z|$. Let us write X for the set of sequences $(n_z)_{z\in Z} \in \mathbf{N}^Z$ whose support is a dissociated subset of Z and with $\sum_z n_z \leqslant |Z|$. We upper bound the size of X. Pick any sequence $(n_z)_{z\in Z} \in \mathbf{N}^Z$ in X and let N be its support. So N is dissociated and as the largest dissociated subset of Z has size $\dim(Z)$, we can fix a set N' containing N and of size exactly $\dim(Z)$. Then there are at most

$$\binom{|Z|}{\dim(Z)}$$

choices of N' over all sequences in X. Given a set N', it is a standard combinatorial fact that the number of sequences $(m_z)_{z\in N} \in \mathbf{N}^{N'}$ with $\sum_{z\in N'} m_z \leqslant |Z|$ is

$$\binom{|Z| + |N'|}{|N'|} = \binom{|Z| + \dim(Z)}{\dim(Z)},$$

and clearly the sequence $(n_z)_{z\in Z} \in \mathbf{N}^Z$ is counted here as its support N is contained in N'. Hence, in total we get that

$$|X| \leqslant \binom{|Z|}{\dim(Z)} \binom{|Z| + \dim(Z)}{\dim(Z)}. \tag{4.11}$$

On the other hand, every $y \in \Sigma(Z)$ gives rise to the sequence $(k_z(y))_{z\in Z}$ as in (4.10) whose support is dissociated and with $\sum_z k_z(y) \leqslant |Z|$. Hence, $(k_z(y))_{z\in Z}$ in X. As $\sum_z k_z(y) z = y$ holds in G, no two distinct $y, y' \in \Sigma(Z)$ can give rise to the same sequence $(k_z(y))_{z\in Z}$ so that

$$|X| \geqslant |\Sigma(Z)|\,. \tag{4.12}$$

Combining inequalities (4.11) and (4.12) yields the desired result (4.2). □

To conclude this section, we simplify the bound obtained in Proposition 4.3.2 to obtain Corollary 4.3.3.

Proof of Corollary 4.3.3. We use the following standard inequality for binomial coefficients with integers $0 \leqslant r \leqslant n$:

$$\binom{n+r}{r} \leqslant \left(\frac{e(n+r)}{r}\right)^r,$$

where e is Euler's constant. Using inequality (4.2) and the inequality above with $n = |Z|$ and $r = \dim(Z) = d$ yields the desired bound

$$\begin{aligned} |\Sigma(Z)| &\leqslant \binom{|Z|}{d}\binom{|Z|+d}{d} \\ &\leqslant \binom{|Z|}{d}\binom{2|Z|}{d} \\ &\leqslant \left(\frac{e|Z|}{d}\right)^d \left(\frac{2e|Z|}{d}\right)^d \leqslant 2^{2d\log_2\left(\frac{4|Z|}{d}\right)}. \end{aligned}$$

□

4.4 Balanced sets

Let $A \subset G$ be a subset having no unique sum. To prove Theorem 4.1.4, we want to use Proposition 4.3.2 with $Z = A$ in order to deduce a lower bound on its size $|A|$. The first step in our proof of Theorem 4.1.4 is therefore to show that the additive span $\Sigma(A)$ is large. It turns out that this step works under a weaker assumption than that A has no unique sum, and this weaker property is all that is needed for this section.

Definition 4.4.1. Let $B \subset G$ be a subset of an Abelian group G. We say that B is *balanced* if for every $b \in B$, there exist distinct $b_1, b_2 \in B$ so that $2b = b_1 + b_2$. In other words, B is balanced if every element is the midpoint of a non-trivial 3-term arithmetic progression which is contained in B. If B is balanced but does not contain two disjoint balanced subsets, then we say that B is an *irreducible balanced* set.

Note that no finite subset of the integers is balanced, since the largest element is clearly not the midpoint of a non-trivial 3-term arithmetic progression contained in the set. Lemma 2.0.8 therefore shows that if $B \subset \mathbf{Z}/p\mathbf{Z}$ is balanced, then $|B| \geqslant \log_2 p$ as B cannot be rectifiable.

Definition 4.4.2. Let G be a finite Abelian group. Then we define $b(G)$ to be the size of the smallest subset of G which is balanced. We also write $b(p)$ for $b(\mathbf{Z}/p\mathbf{Z})$.

Balanced sets have been studied in their own right in multiple papers, resulting in a very precise asymptotic for $b(p)$ which is correct up to lower order terms. In fact, $b(p) = (1 + o(1)) \log_2 p$ where the lower bound follows from the rectification argument above and the upper bound comes from a construction of Nedev [53]. It is clear that if $A \subset G$ has no unique sum, then certainly the sum $a + a = 2a$ has a different representation as a sum of two elements in A so A is also a balanced set. Since balanced sets of size $(1 + o(1)) \log_2 p$ exist, the rectification bound is in a sense the only obstruction preventing a set from being balanced. For sets having no unique sum there are further obstructions, as the proof of Theorem 4.1.5 will show.

From this section onward, all sets that we consider are proper sets (as opposed to multisets). Recall that for a proper set $S \subset G$, its additive span is simply the set of subset sums

$$\Sigma(S) := \left\{ \sum_{s \in S'} s : S' \subseteq S \right\},$$

and we say that S is an additive basis for G if $\Sigma(S) = G$. The following proposition, whose proof we postpone to the end of this section, states that balanced sets have a translate with large additive span.

Proposition 4.4.3. *Let $B \subset \mathbf{Z}/p\mathbf{Z}$ be a balanced set. Then there exists $g \in -B$ such that the translated set $B + g$ is an additive basis for $\mathbf{Z}/p\mathbf{Z}$.*

As an aside, we note that this gives a new proof of the lower bound $b(p) > \log_2 p$ which, as we mentioned before, is best possible up to lower order terms.

Corollary 4.4.4. *If $B \subset \mathbf{Z}/p\mathbf{Z}$ is balanced, then $|B| \geqslant \log_2 p + 1$.*

Proof. Let $B \subset \mathbf{Z}/p\mathbf{Z}$ be a balanced set, then there is some translate $B + g$ of B so that $\Sigma(B + g) = \mathbf{Z}/p\mathbf{Z}$ and $g \in -B$ by Proposition 4.4.3. As $0 \in B + g$, we deduce that $|\Sigma(B + g)| \leqslant 2^{|B|-1}$ and combining this with $|\Sigma(B + g)| \geqslant p$ yields the result. □

The situation in a general Abelian group G is a bit more delicate. If p is a prime dividing the order of a group G, then G contains a cyclic group of order p as a subgroup. Hence,

$$b(G) \leqslant \min_{p \text{ prime}, \, p||G|} b(p).$$

However, one might hope that if $B \subset G$ is balanced and B generates a large subgroup of G, then one can obtain an improved lower bound on $|B|$. This is not true in general as the property of being balanced is preserved under translation, so one could take a balanced subset of G of size $b(p(G))$ and then take B to be a translate of this set which generates a large subgroup of G. This is the reason for introducing the following definition. Here, for $C \subset G$ we use the standard notation $\langle C \rangle$ for the subgroup of G generated by C.

Definition 4.4.5. Let $C \subset G$, then we define $\mathrm{minspan}(C) := \min_{g \in G} |\langle C + g \rangle|$.

One may note that $\mathrm{minspan}(C) = \langle C - C \rangle$. Further, one can see that the union of any two balanced sets is balanced. This again could lead to small balanced sets in G for which any translate generates a large subgroup of G. To avoid this, we work with irreducible balanced sets and in doing so, we obtain the following generalisation of Proposition 4.4.3, and it is the only result from this section that we need for the proof of Theorem 4.1.5.

Proposition 4.4.6. *Let G be a finite Abelian group and let $B \subset G$ be an irreducible balanced set. Then there exists $g \in -B$ such that $\Sigma(B+g) = \langle B+g \rangle$, i.e. the translated set $B+g$ is an additive basis for $\langle B+g \rangle$.*

Remark. *There do in fact exist non-irreducible balanced sets $B \subset G$ for which no translate $B+g$ is an additive basis for $\langle B+g \rangle$. As an example one can consider*

$$B = \mathbf{Z}/3\mathbf{Z} \times \{0,1\} \subset \mathbf{Z}/3\mathbf{Z} \times \mathbf{Z}/p\mathbf{Z}. \tag{4.13}$$

Then B is balanced, but any translate contains an element of order $3p$, so $\langle B+g \rangle$ has size at least $3p$ and cannot have an additive basis of size $|B| = 6$ for p large. Note that in this example, B is not irreducible as it is the disjoint union of two balanced sets of size 3.

We deduce the following corollary, giving an improved lower bound for sets with large minspan.

Corollary 4.4.7. *Let G be an Abelian group. If $B \subset G$ is an irreducible balanced set, then $|B| \geqslant \log_2(\mathrm{minspan}(B)) + 1$.*

Proof. By Proposition 4.4.6, there is some translate $B+g$ of B which contains 0 and is an additive basis of $\langle B+g \rangle$. As $0 \in B+g$, we deduce that $|\Sigma(B+g)| \leqslant 2^{|B+g|-1} = 2^{|B|-1}$. As $B+g$ is an additive basis of $\langle B+g \rangle$, we also get that $|\Sigma(B+g)| \geqslant |\langle B+g \rangle| \geqslant \mathrm{minspan}(B)$. Combining these two inequalities gives the result. □

If B is a balanced set which is not irreducible, then this result breaks down. In fact, one cannot obtain any lower bound growing with $\mathrm{minspan}(B)$ without the assumption that B is irreducible as the example defined in (4.13) shows. Hence, the following result is best possible for a general balanced set (up to lower order terms).

Corollary 4.4.8. *If $B \subset G$ is balanced, then $|B| \geqslant \log_2 p(G) + 1$.*

Proof. Let $B \subset G$ be a balanced set, then it contains an irreducible balanced subset $B' \subset B$. Any balanced set clearly has at least two distinct elements, so any translate $B' + g$ contains a non-zero element of G. Hence, we see that $\langle B' + g \rangle$ is a non-zero subgroup of G so has size at least $p(G)$ by Lagrange's theorem. So $\mathrm{minspan}(B') \geqslant p(G)$ and using the result from Corollary 4.4.7 gives the required lower bound $|B| \geqslant |B'| \geqslant \log_2 p(G) + 1$. □

Remark. *The main purpose of this section is to prove the auxiliary result Proposition 4.4.6 for the proof of Theorem 4.1.5, but we stated some of its corollaries which are interesting in their own right as they yield a new strongest lower bound on the size of a balanced set in a general Abelian group:*

$$|B| \geqslant \min\left(\log_2 \mathrm{minspan}(B),\, 2\log_2 p(G) + 1\right) + 1.$$

This follows by applying Corollary 4.4.7 if B is irreducible, and applying Corollary 4.4.8 to the two disjoint balanced sets contained in B that exist if B is not irreducible.

Let us now give the proof of Propositions 4.4.3 and 4.4.6. First we give the simple deduction of Proposition 4.4.3 from Proposition 4.4.6.

Proof of Proposition 4.4.3 assuming Proposition 4.4.6. Let p be prime and $B \subset \mathbf{Z}/p\mathbf{Z}$ be a balanced set. Note that B contains an irreducible balanced subset $B' \subset B$. By Proposition 4.4.6, as B' is an irreducible balanced set, there exists $g \in -B' \subset -B$ so that $B' + g$ is an additive basis for $\langle B' + g \rangle$. Any balanced set clearly has at least two distinct elements, so $B' + g$ contains a non-zero element of $\mathbf{Z}/p\mathbf{Z}$ so that $\langle B' + g \rangle = \mathbf{Z}/p\mathbf{Z}$. Hence,

$$\Sigma(B + g) \supset \Sigma(B' + g) \supset \mathbf{Z}/p\mathbf{Z}.$$

□

Proof of Proposition 4.4.6. Let G be a finite Abelian group and let $B \subset G$ be an irreducible balanced subset. We will show that there exists an element $g \in -B$ for which the translated set $B + g$ is an additive basis for the subgroup $\langle B + g \rangle \leqslant G$. We

will pick such a $g \in -B$ later and then consider the translated set $B+g$. Clearly, $B+g$ is still an irreducible balanced set since this property is preserved under translation. Now let us pick any element $y \in \langle B+g \rangle$ and as the ambient group G is finite, we can find non-negative integers $n_b(y)$ for $b \in B$ so that $y = \sum_{b \in B} n_b(y)(b+g)$. If it is the case that each such $n_b(y)$ is either 0 or 1, then we immediately deduce the desired conclusion that y lies in the additive span $\Sigma(B+g)$ of $B+g$. If not, there is some $b \in B$ with $n_b(y) \geqslant 2$ and we will then use the relation $2b = b_1 + b_2$ to decrease $n_b(y)$. By applying such 'compressions' in a certain order, we obtain a way to write y as a sum of the form $\sum_{b \in B} m_b(b+g)$ with $m_b \in \{0,1\}$ so that $y \in \Sigma(B+g)$ as desired. For this, we will use ´weight-compressions' which, for an expression $y = \sum_{b \in B} n_b(b+g)$ with non-negative integers n_b that is not already maximally compressed, yield a new expression $y = \sum_b m_b(b+g)$ with larger weight. To define a weight function on the finite set B, it is convenient to use the language of graph theory.

Note that B contains a balanced subset B' which is minimal in the sense that no proper subset of B' is balanced.[2] Let H be a directed graph with vertex set B and for each vertex $b \in B$ we have two outgoing edges $b \to b_1$ and $b \to b_2$ where $b_1, b_2 \in B$ are distinct with $2b = b_1 + b_2$. As B is balanced, we can always find such b_1, b_2. If there is more than one choice of b_1, b_2 then we just pick one of them arbitrarily, except when $b \in B'$ in which case we always choose $b_1, b_2 \in B'$. So every vertex in H has outdegree exactly 2 and every vertex in the induced subgraph $H[B']$ also has outdegree 2. We need the following lemma.

Lemma 4.4.9. *Let H be the directed graph defined above. Then for any vertex $g' \in B'$ and any vertex $b \in V(H) = B$, there is a directed path from b to g' in H.*

Proof. Define for a vertex $h \in V(H) = B$ the set $R(h)$ to be the set of vertices h_1 in H for which there exists a directed path (possibly consisting of a single vertex) from h to h_1 in H. We show that $R(h) \subset B$ is itself a balanced set for all $h \in B$. Consider an element $x \in R(h)$ so there exists a directed path P in H going from h to x. As B is balanced, we can find distinct x_1, x_2 in B with $2x = x_1 + x_2$ and so that $x \to x_1$ and $x \to x_2$ are edges of H. But then concatenating P with each of these edges gives directed paths from h to x_1 and from h to x_2. Hence, $x_1, x_2 \in R(h)$ and we conclude that $R(h)$ is indeed a balanced set itself. This shows that for each $b' \in B'$, the set $R(b') \subset B'$ is a balanced subset of B' so that as B' was assumed to be a minimal balanced set, we get that $R(b') = B'$. Thus, we have shown that for any two vertices

[2] Recall that B is irreducible, meaning that B does not contain two disjoint balanced subsets. In general, an irreducible balanced set B can still contain a proper subset B' which is also balanced.

$b'_1, b'_2 \in B'$, there is a directed path from b'_1 to b'_2 in H. Finally, for any $b \in B$, the two sets B' and $R(b)$ are balanced subsets of B. As B is irreducible, this means that B' and $R(b)$ intersect so there is a directed path from b to a vertex in B'. We have shown that any two vertices in B' are connected by a directed path so that $B' \subset R(b)$ as desired. □

We continue with the proof of Proposition 4.4.6. Pick any $g \in -B'$ and we will show that the translated set $B + g$ has the desired properties, meaning that $\Sigma(B + g) = \langle B + g \rangle$. Let $g' = -g$ so $g' \in B'$ and we are ready to define our weight function. For each vertex $b \in B$, let the number $s(b)$ denote the length of the shortest directed path from b to g' in H, so $s(g') = 0$ for example. By Lemma 4.4.9, $s(b)$ is finite for every b. We now define a weight function $w : B \to \mathbf{R}$ on B by $w(b) := 2^{-s(b)}$ for each $b \in B$. Let $y \in \langle B + g \rangle$ so we can write $y = \sum_{b \in B} n_b(y)(b + g)$ for some non-negative integers $n_b(y)$. Let $N_y = \sum_{b \in B} n_b(y) \in \mathbf{N}$ and consider the following set of $|B|$-tuples of non-negative integers:

$$S_{B,g}(y) := \left\{ (m_b)_{b \in B} \in \mathbf{N}^B : \sum_{b \in B} m_b(b + g) = y \text{ and } \sum_{b \in B} m_b = N_y \right\}.$$

For each $|B|$-tuple $(m_b)_{b \in B}$ in $S = S_{B,g}(y)$, we define its weight

$$w\left((m_b)_{b \in B}\right) := \sum_{b \in B} m_b w(b) \in [0, \infty).$$

Now S contains the tuple $(n_b(y))_{b \in B}$ so it is certainly non-empty. Further, the number of $|B|-$tuples $(m_b)_{b \in B} \in \mathbf{N}^B$ with $\sum_b m_b = N_y$ is finite, so S is a finite set. The idea is now to consider the tuple $(k_b)_{b \in B}$ in S with maximal weight and we show that this forces each k_b with $b \neq g'$ to be either 0 or 1. So let $(k_b)_{b \in B}$ be a tuple in S with maximal weight and suppose for a contradiction that there is some $b \in B \setminus \{g'\}$ with $k_b \geqslant 2$. Then let $P = b, b_1, \ldots, g'$ be a shortest path from b to g' in H, of length $s(b) \geqslant 1$. By definition of the edges in H this means that there exists $b_2 \in V(H) = B$ so that $2b = b_1 + b_2$. Now let $(k'_c)_{c \in B}$ be a new tuple of non-negative integers defined by $k'_c = k_c$ for all $c \in B \setminus \{b, b_1, b_2\}$, $k'_b = k_b - 2 \geqslant 0$, and $k'_{b_i} = k_{b_i} + 1$ for $i = 1, 2$. We show that $(k'_c)_{c \in B} \in S$. First note that $\sum_c k'_c = \sum_c k_c = N_y$. As $2b = b_1 + b_2$, we also have that

$$\begin{aligned} y &= y + (b_1 + g) + (b_2 + g) - 2(b + g) \\ &= \Big(\sum_c k_c(c + g) \Big) + (b_1 + g) + (b_2 + g) - 2(b + g) \end{aligned}$$

$$= \sum_c k'_c(c+g).$$

So $(k'_c) \in S$, but we show that its weight is strictly larger than the weight of (k_c) giving the required contradiction:

$$\begin{aligned} w\left((k'_c)\right) &= \sum_c k'_c w(c) \\ &= (k_b - 2)w(b) + (k_{b_1}+1)w(b_1) + (k_{b_2}+1)w(b_2) + \sum_{c \in B\setminus\{c,b_1,b_2\}} k_c w(c) \\ &= w\left((k_c)\right) - 2w(b) + w(b_1) + w(b_2) \\ &= w\left((k_c)\right) - 2\cdot 2^{-s(b)} + 2^{-s(b_1)} + 2^{-s(b_2)} \\ &\geqslant w\left((k_c)\right) - 2^{-s(b)+1} + 2^{-s(b)+1} + 2^{-s(b_2)} \\ &> w\left((k_c)\right), \end{aligned}$$

where we used that $s(b_1) \leqslant s(b) - 1$ as b_1 comes after b in the shortest path P from b to g', and that $w(b_2) = 2^{-s(b_2)} > 0$.

We conclude that if $(k_b)_{b\in B}$ is an element of S with largest weight, then $k_b \in \{0,1\}$ for all $b \in B \setminus \{g'\}$. Hence, the following equality shows that $y \in \Sigma(B+g)$ as desired:

$$\begin{aligned} y &= \sum_{b\in B} k_b(b+g) \\ &= \sum_{b\in B\setminus\{g'\}} k_b(b+g) + k_{g'}(g'+g) \\ &= \sum_{b\in B\setminus\{g'\}} k_b(b+g) \in \Sigma(B+g), \end{aligned}$$

as we chose $g = -g'$. Since $y \in \langle B+g\rangle$ was arbitrary, we have shown that $\langle B+g\rangle = \Sigma(B+g)$. □

Remark. *Note that in the proof, the compressions can be used repeatedly on the original sum $y = \sum_b n_b(b+g)$ until we obtain a sum $y = \sum_b k_b(b+g)$ for which almost all the weight is placed at the vertex $g' \in V(H) = B$. Since this element contributes $k_{g'}(g'+g) = 0 \in G$ to the sum, it does not matter that $k_{g'}$ is generally not in $\{0,1\}$. This shows the advantage of working with a translate $B+g$ instead of B, and in fact this is necessary as B does not need to be an additive basis for $\langle B\rangle$.*

Remark. *Consider an arithmetic progression $Q = \{a, 2a, \ldots ka\}$ of length k in $G = \mathbf{Z}/p\mathbf{Z}$. Then every element of Q except a and ka is the midpoint of a non-trivial 3-term arithmetic progression in Q. Taking $k = 200$ for example, one can get that*

that Q is 'almost' balanced, in the sense that 99% of the elements $b \in Q$ have that $2b = b_1 + b_2$ for distinct $b_1, b_2 \in Q$. However, Q is very far from being balanced in that one needs to add at least $\log_2 p - 200$ more elements to make it balanced. This observation shows that any proof of a logarithmic lower bound must have an algebraic flavour in the sense that one must crucially use that every single b satisfies a balanced relation, as opposed to almost every b.

We finish this section on balanced sets by using them to construct small sets in $\mathbf{Z}/p\mathbf{Z}$ having no unique sum, thus proving our new upper bound on $m(G)$ from Theorem 4.1.6. In order to construct a set $A \subset G$ having no unique sum, it is natural to try using sets with a gridlike structure. Indeed, let C, D be any subsets of the finite Abelian groups G and G' respectively. Then consider the Cartesian product $C \times D \subset G \times G'$ and let $(c_1, d_1) + (c_2, d_2)$ be any sum in its sumset $C \times D + C \times D$. Then we have that

$$(c_1, d_1) + (c_2, d_2) = (c_1, d_2) + (c_2, d_1). \tag{4.14}$$

This trivial observation shows that such a sum can only be unique if $c_1 = c_2$ or if $d_1 = d_2$. To fix the fact that sums as in (4.14) where $c_1 = c_2$ or $d_1 = d_2$ can be unique in general, we consider the set $A = B \times B \subset (\mathbf{Z}/p\mathbf{Z})^2$ for a balanced set $B \subset \mathbf{Z}/p\mathbf{Z}$. It is easy to check that A has no unique sum. Let $b, b', c, c' \in B$, so the sum $(b, b') + (c, c')$ is not unique if $b \neq c$ and $b' \neq c'$ by (4.14). On the other hand, if $b = c$ then we can write $b + c = 2b = b_1 + b_2$ for distinct $b_1, b_2 \in B$ as B is balanced. Then the equality $(b, b') + (c, c') = (b_1, b') + (b_2, c')$ shows that the sum is not unique. Finally, the same argument shows that $(b, b') + (c, c')$ is not unique if $c = c'$. Using this observation and [53], Theorem 4.1.6 easily follows.

Proof of Theorem 4.1.6. By Theorem 1 in [53], one can find a balanced set $B \subset \mathbf{Z}/p\mathbf{Z}$ of size

$$|B| \leqslant (1 + o(1)) \log_2 p.$$

From the paragraph above, $A = B \times B \subset (\mathbf{Z}/p\mathbf{Z})^2$ is a set having no unique sum of size $|A| = |B|^2$. It is clear that if T and T' are Freiman-isomorphic sets, then T has no unique sum if and only if T' has no unique sum. Hence, all that remains is to find a Freiman isomorphism from $A \subset (\mathbf{Z}/p\mathbf{Z})^2$ into $\mathbf{Z}/p\mathbf{Z}$. Let $r \in \mathbf{Z}/p\mathbf{Z}$ and define $\phi_r : A \to \mathbf{Z}/p\mathbf{Z} : (b, b') \mapsto b + rb'$. Then ϕ_r is a Freiman homomorphism for all values of r, and can only fail to be a Freiman isomorphism if $r \in (2B - 2B)/(2B - 2B)$. As this set contains at most $|B|^8 \leqslant (1 + o(1))(\log_2 p)^8 < p$ elements for p large, the map ϕ_r is a Freiman isomorphism for some r, thus giving a subset of $\mathbf{Z}/p\mathbf{Z}$ of size $|B|^2 = O((\log p)^2)$ with no unique sum. □

Remark. *One can improve the implied constant by a factor of* 2 *by noting that if* B *is a balanced set in* $\mathbf{Z}/p\mathbf{Z}$*, then the set* $A = B+B \subset \mathbf{Z}/p\mathbf{Z}$ *has no unique sum and size at most* $\binom{|B|+1}{2}$*. We opted to give the proof based on a Freiman isomorphism from* $B \times B$ *into* $\mathbf{Z}/p\mathbf{Z}$ *as it makes clear that we are using a certain two-dimensional structure in order to get non-unique sums. In correspondence with Kopparty, the author found out that essentially the same construction for this upper bound also appears in the thesis of Scheinerman [59].*

4.5 Sets with no unique sum have small dimension

In this final section, we combine all our results to prove a lower bound on the size of a set $A \subset G$ having no unique sum. Our argument begins by applying the inequality in Proposition 4.3.2 to $A + g$ and plugging in the lower bound on $|\Sigma(A + g)|$ from Proposition 4.4.6. We introduce the following convenient notation.

Definition 4.5.1. *Let* $Z \subset G$*, then define the number*

$$K(Z) := \frac{|Z|}{\dim(Z)}. \tag{4.15}$$

Then simply rewriting the inequality in Corollary 4.3.3 using that $\dim(Z) = \frac{|Z|}{K(Z)}$ gives the following.

Proposition 4.5.2. *Let* G *be an Abelian group and let* Z *be a finite multiset consisting of elements of* G*. Then*

$$|Z| \geqslant \frac{K(Z)}{2\,(2 + \log_2 K(Z))} \cdot \log_2 |\Sigma(Z)|\,. \tag{4.16}$$

This is the form of the inequality that will be useful for our purpose. In the previous section, we have proven Proposition 4.4.6 which showed that if $B \subset G$ is a balanced set, then it contains an irreducible balanced subset B' which has a translate $B' + g$ that forms an additive basis for $\langle B' + g\rangle$. Since $B' + g$ contains a non-zero element in G, $\langle B' + g\rangle$ is a non-trivial subgroup of G so that by Lagrange's Theorem we get

$$|\Sigma(B + g)| \geqslant |\Sigma(B' + g)| \geqslant |\langle B' + g\rangle| \geqslant p(G), \tag{4.17}$$

where $p(G)$ denotes the smallest prime divisor of G. Using this in inequality (4.16) shows that for any balanced set B in an Abelian group G, we have the lower bound

$$|B| = |B + g| \geqslant \frac{K(B + g)}{2\,(2 + \log_2 K(B + g))} \cdot \log_2 p(G). \tag{4.18}$$

At this point we make an interesting observation. As we noted in the previous section, Nedev [53] showed that G contains a balanced subset of size at most $(1 + o(1)) \log_2 p(G)$. Now looking at (4.18), if B is a balanced set of size $C \log_2 p(G)$ then we must have that $K(B + g) = O_C(1)$ is bounded by a constant depending only on C. This means precisely that any such balanced set B in G has a translate with large additive dimension $\dim(B + g) = \frac{|B|}{K(B+g)} \gg_C |B|$, i.e. it contains a dense dissociated subset. The goal in this section is to show that the situation is different for sets having no unique sum. We show that if $A \subset G$ has no unique sum, then A does not contain a dense dissociated subset and in fact we have $\dim(A) = o_{p(G)}(1) \cdot |A|$. Equivalently, we prove that $K(A) \to \infty$ as $p(G) \to \infty$, and plugging this into (4.18) will then give the desired lower bound

$$|A| \geqslant \omega(p(G)) \log_2 p(G).$$

Remark. *Note that it is not true that* $\dim(A) = o(|A|)$ *as* $|A| \to \infty$ *for sets* A *containing no unique sum. What we do show is that* $\dim(A) = o_{p(G)}(1) \cdot |A|$. *For an example, we can consider* $A_1 = B \times \mathbf{Z}/3\mathbf{Z} \subset \mathbf{Z}/p\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$ *where* $B \subset \mathbf{Z}/p\mathbf{Z}$ *is a balanced set of size* $C \log_2 p$ *with* $\dim(B) \geqslant \frac{|B|}{C'}$ *(such* B *exist by the discussion above). Then* A_1 *is a product of two balanced sets and hence has no unique sum. However,* A_1 *contains* $B \times \{0\}$ *so* $\dim(A_1) \geqslant \dim(B) \geqslant \frac{|B|}{C'} \geqslant \frac{|A_1|}{3C'}$. *By taking* p *large, it follows that there are arbitrarily large sets having no unique sum but which do contain a dense dissociated subset. The issue in this example, of course, is that the smallest prime factor of* $|\mathbf{Z}/p\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}|$ *is bounded.*

The following proposition gives a precise statement of the fact that sets with no unique sum have small additive dimension.

Proposition 4.5.3. *Let* G *be a finite Abelian group with* $p(G)$ *being the smallest prime dividing* $|G|$. *Let* $A \subset G$ *have no unique sum. Then*

$$K(A) \geqslant \omega_1(p(G)), \tag{4.19}$$

for some function $\omega_1 : \mathbf{N} \to (0, \infty)$ *with* $\omega_1(n)$ *tending to infinity as* $n \to \infty$. *Moreover, one can take*

$$\omega_1(n) \gg \sqrt{\log \log \log n}, \tag{4.20}$$

where the implied constant is absolute.

Assuming this proposition for the moment, we can put everything together and complete the proof of Theorem 4.1.5.

Proof of Theorem 4.1.5 assuming Proposition 4.5.3. Let $A \subset G$ have no unique sum. In particular, this means that A is balanced so (4.18) gives that

$$|A| \geqslant \frac{K(A+g)}{2\,(2+\log_2 K(A+g))} \cdot \log_2 p(G) \tag{4.21}$$

for some element $g \in G$. As $A \subset G$ does not have a unique sum, neither does the translated set $A+g$. By Proposition 4.5.3, we then get that

$$K(A+g) \geqslant \omega_1(p(G)).$$

Plugging this in (4.21) gives

$$|A| \geqslant \omega(p(G)) \log_2 p(G)$$

if we define

$$\omega(n) := \frac{\omega_1(n)}{2(2+\log_2 \omega_1(n))}.$$

Note that by Proposition 4.5.3, we have that $\omega_1(n) \gg \sqrt{\log\log\log n}$ so

$$\omega(n) \gg \frac{\sqrt{\log\log\log n}}{\log\log\log\log n}.$$

This concludes the proof of Theorem 4.1.5. □

Remark. *Theorem 4.1.5 is rather delicate in the sense that the result is false if one only assumes that A is balanced and that most sums in A are not unique. To see this, consider $A_1 = B \times Q \subset (\mathbf{Z}/p\mathbf{Z})^2 = G$ with B a minimal balanced set and Q an arithmetic progression of size 200. Then A_1 is balanced as B is, and A_1 has that 99% of its sums are non-unique. However, $|A_1| \leqslant 200(1+o(1))\log_2 p(G)$.*

Our final task, then, is to prove Proposition 4.5.3. The main tool in the proof is the following proposition.

Proposition 4.5.4. *Let G be a finite Abelian group and let $A \subset G$ have no unique sum. There exists an absolute constant C such that the following holds. Suppose that D is a dissociated subset of A with $|D| \geqslant 10$ (say) and that $S \subset G$ contains 0. If*

$$|S| \leqslant \min\left(\log_2 p(G), \left(\frac{|D|^6}{C|A|^5}\right)^{\frac{1}{4}}\right), \tag{4.22}$$

then there exists a set $S' \subset G$ containing 0 and of size

$$|S'| \leqslant \max(2|S|, |S|^3) \tag{4.23}$$

so that

$$|(D+S') \cap A| \geqslant |(D+S) \cap A| + \frac{|D|^2}{36|A|}. \tag{4.24}$$

Roughly speaking, this proposition states that if a set A having no unique sum contains a dissociated subset D so that few translates of D (namely the set $D+S$) contains a certain fraction of all elements of A, then there exists a slightly larger set of translates S' so that $D+S'$ contains a significantly bigger fraction of A. Assuming this proposition, we show how to deduce Proposition 4.5.3 using a density increment argument.

Proof of Proposition 4.5.3 assuming Proposition 4.5.4. Let $A \subset G$ have no unique sum, and let $D \subset A$ be a dissociated subset of A of largest possible size. So $|D| = \dim(A)$ and $|D| = \frac{|A|}{K(A)}$. Let $p = p(G)$. Our goal is to prove that $K(A) \gg \sqrt{\log\log\log p}$. If $|D| < 10$, then $K(A) \geqslant \frac{|A|}{10} \gg \log p$ by Corollary 4.4.8 as A is balanced, so we are done. Hence, we may assume that $|D| \geqslant 10$ so that D satisfies the assumption of Proposition 4.5.4. We apply Proposition 4.5.4 with $D \subset A$ and $S = S_0 := \{0\}$. Then either $|S_0| > \min\left(\log_2 p, \left(\frac{|D|^6}{C|A|^5}\right)^{\frac{1}{4}}\right)$ or else the assumption (4.22) is satisfied and we deduce that there exists a set $S_1 \subset G$ containing 0 of size at most $\max(2|S_0|, |S_0|^3) = 2$ so that $|(D+S_1)\cap A| \geqslant |(D+S_0)\cap A| + \frac{|A|}{36K(A)^2}$. Suppose that after i steps we have a set $S_i \subset G$ containing 0 and of size at most

$$|S_i| \leqslant 2^{3^i}$$

so that $D+S_i$ contains at least $\frac{i|A|}{36K(A)^2}$ of the elements in A. Then either we have that $|S_i| > \min\left(\log_2 p, \left(\frac{|D|^6}{C|A|^5}\right)^{\frac{1}{4}}\right)$ or else (4.22) is satisfied so we can again apply Proposition 4.5.4 to find a set $S_{i+1} \subset G$ also containing 0 and of size

$$|S_{i+1}| \leqslant \max(2|S_i|, |S_i|^3) \leqslant 2^{3^{i+1}}$$

so that $D+S_{i+1}$ contains a fraction of at least $\frac{i+1}{36K(A)^2}$ of the elements of A. Now it is clear that $|(D+S_i)\cap A| \leqslant |A|$ for all i, and hence the iterative procedure described above must fail for some $j \leqslant 36K(A)^2$. By Proposition 4.5.4, the only way that this can happen is if $|S_j| > \min\left(\log_2 p, \left(\frac{|D|^6}{C|A|^5}\right)^{\frac{1}{4}}\right)$. First, if $|S_j| > \log_2 p$, then as $|S_j| \leqslant 2^{3^j}$ we deduce that $36K(A)^2 \geqslant j \gg \log\log|S_j| \gg \log\log\log p$ and we have proven (4.19).

Finally, if $|S_j| > \left(\frac{|D|^6}{C|A|^5}\right)^{\frac{1}{4}}$, then recalling that $|S_j| \leqslant 2^{3^j}$ and that we defined $|D| = \frac{|A|}{K(A)}$ gives

$$36K(A)^2 \geqslant j \gg \log\log|S_j| \geqslant \log\left(\frac{1}{4}\log\left(\frac{|A|}{CK(A)^6}\right)\right)$$

$$= \log\left(\frac{1}{4}\log_2(|A|) - \frac{1}{4}\log_2(CK(A)^6)\right)$$
$$\geqslant \log\left(\frac{1}{4}\log_2\log_2 p - \frac{1}{4}\log_2(CK(A)^6)\right),$$

where here we used the weak bound $|A| \geqslant \log_2 p(G)$ which by Corollary 4.4.8 holds even for balanced sets. We deduce that $K(A) \gg \sqrt{\log\log\log p}$ as desired. □

To complete the proof, it now only remains to prove Proposition 4.5.4. So far, we have not yet used that A has no unique sum but only the much weaker assumption that A is balanced. As we saw, the conclusion of Proposition 4.5.3 fails completely for a general balanced set, even if most sums are not unique. So our proof of Proposition 4.5.4 here must inevitably use that A has no unique sum. This will make the proof somewhat technical, but this seems unavoidable at this stage. We begin with a useful lemma.

Lemma 4.5.5. *Let G be a finite Abelian group and let $S \subset G$ have size at most $\log_2 p(G)$. Then we can assign an element $s_X \in X$ to each non-empty subset $X \subset S$ such that the following holds. Let $X, Y \subset S$ be non-empty subsets, then the only solution to $x + y = s_X + s_Y$ with $x \in X$ and $y \in Y$ is the trivial solution $(x, y) = (s_X, s_Y)$.*

Proof. S is rectifiable by Lemma 4.2.1 and let $\phi : S \to S' \subset \mathbf{N}$ be a Freiman isomorphism to a subset of $\mathbf{N}$. Let $X \subset S$ be non-empty. We claim that the choice $s_X = \phi^{-1}(\max \phi(X))$ works. Let $X, Y \subset S$ be non-empty. As ϕ is a Freiman isomorphism, it is enough to show that the only solution $(x, y) \in \phi(X) \times \phi(Y)$ to $x + y = \max \phi(X) + \max \phi(Y)$ is the trivial one. This is clear since $x + y \leqslant \max \phi(X) + \max \phi(Y)$ is an inequality in the integers, where equality only holds if $(x, y) = (\max \phi(X), \max \phi(Y))$. □

We are now ready to begin the proof of Proposition 4.5.4.

Proof of Proposition 4.5.4. Let G be a finite Abelian group and let $A \subset G$ have no unique sum. Suppose that D is a dissociated subset of A. Let $S \subset G$ contain 0 and have size $|S| \leqslant \min\left(\log_2 p(G), \left(\frac{|D|^6}{C|A|^5}\right)^{\frac{1}{4}}\right)$. Finally, suppose that the set $D + S$ contains a fraction α of all elements of A, meaning that

$$|(D + S) \cap A| = \alpha|A|. \tag{4.25}$$

The idea is that since D is dissociated and the set of shifts S is small, the set $(D+S)\cap A$ still has many unique sums. Since A has no unique sum, we will show that this forces A to contain a large part of a larger set of translates $D + S'$.

We introduce some notation. For each $d \in D$, we define

$$S_d = \{s \in S : d + s \in A\} \tag{4.26}$$

and note that $0 \in S_d$ as D is a subset of A. Hence, each S_d is a non-empty subset of S and by applying Lemma 4.5.5, we can find an element $s_d \in S_d$ for each $d \in D$ so that whenever $d, d' \in D$, then the only solution $(x, y) \in S_d \times S_{d'}$ to

$$x + y = s_d + s_{d'} \tag{4.27}$$

is the trivial solution $(x, y) = (s_d, s_{d'})$. Also, let us define the following set of elements of D:

$$B^{(1)}(D) := \{d \in D : \text{there is a } v \in (2S - 2S) \setminus \{0\} \text{ so that } d + v \in D\}. \tag{4.28}$$

We think of $B^{(1)}(D)$ as the set of 'bad' elements in D as they will not be useful for our later argument. We shall prove later that this set $B^{(1)}(D)$ of bad elements of D is rather small and hence we can simply remove it from D, so we will work with the set $G^{(1)}(D) := D \setminus B^{(1)}(D)$ of 'good' numbers. It turns out that $G^{(1)}(D)$ can still contain some bad pairs which leads to the following definition. Let $B^{(2)}(D)$ be the set of unordered pairs $\{d, d'\} \in \binom{G^{(1)}(D)}{2}$ for which there exist $e, e' \in D$ and $s, s' \in S$ so that

$$d + s_d + d' + s_{d'} = e + s + e' + s' \tag{4.29}$$

and $\{e, e'\} \neq \{d, d'\}$.[3] We think of $B^{(2)}(D)$ as the set of 'bad' pairs in $\binom{G^{(1)}(D)}{2}$ because one can see how if (4.29) holds, then the sum $(d + s_d) + (d' + s_{d'}) \in (D + S) + (D + S)$ does not yield unique sum in $D + S$. We will see later that $B^{(2)}(D)$ is also not too large so we can also remove such pairs. Hence, we define the set of 'good' pairs in $\binom{D}{2}$ as follows:

$$G^{(2)} = G^{(2)}(D) := \binom{G^{(1)}(D)}{2} \setminus B^{(2)}(D). \tag{4.30}$$

The following lemma shows what we mean by a pair $\{d, d'\} \in G^{(2)}$ being 'good', namely that $(d + s_d) + (d' + s_{d'})$ is a unique sum in $(D + S) \cap A$.

Lemma 4.5.6. *Let $\{d, d'\} \in G^{(2)}(D)$. Then the only solutions $(x, y) \in ((D + S) \cap A)^2$ to*

$$x + y = (d + s_d) + (d' + s_{d'}) \tag{4.31}$$

are the trivial ones $(x, y) = (d + s_d, d' + s_{d'}), (d' + s_{d'}, d + s_d)$.

[3] Here, we use the notation $\binom{E}{2}$ to denote $\{\{x, y\} : x, y \in E, x \neq y\}$, the set of unordered pairs of elements of E.

Proof. Let $\{d, d'\} \in G^{(2)}(D)$ and suppose that $(x, y) \in ((D+S) \cap A)^2$ is a solution to $x + y = (d + s_d) + (d' + s_{d'})$. We show that (x, y) is a trivial solution. As $x, y \in (D+S) \cap A$, there exist $e, e' \in D$ and $s, s' \in S$ so that $x = e + s$ and $y = e' + s'$. Further note that as $x, y \in A$, this means that $s \in S_e$ and $s' \in S_{e'}$ by the definition (4.26) of the sets $S_e, S_{e'}$. We have that

$$d + s_d + d' + s_{d'} = x + y = e + s + e' + s' \tag{4.32}$$

so we have an equation of the form (4.29) and recalling that $\{d, d'\} \in G^{(2)}(D)$ is not in $B^{(2)}(D)$, we must then have that $\{d, d'\} = \{e, e'\}$. By reordering x and y (which does not affect whether or not (x, y) is a trivial solution) we may therefore assume that $d = e$ and that $d' = e'$. Plugging this into (4.32) gives that $s_d + s_{d'} = s + s'$. But as $s \in S_d$ and $s' \in S_{d'}$ and as $s_d, s_{d'}$ were chosen to satisfy the conclusion of Lemma 4.5.5, this is only possible if $s = s_d$ and $s' = s_{d'}$. Hence, $x = e + s = d + s_d$ and $y = e' + s' = d' + s_{d'}$ is a trivial solution to (4.31), as desired. □

We have shown the desirable property that pairs in $G^{(2)}(D)$ yield unique sums, and we now show that we have not removed too many elements in constructing $G^{(2)}(D)$ from $\binom{D}{2}$, i.e. there are few 'bad' pairs. Here, and several more times in the proof it will be convenient to have the following lemma at our disposal. It is an easy consequence of the skew version of Bollobás's Two Families Theorem (see for example [33]), although as all sets involved have size at most 2 one could also give an elementary direct proof.

Lemma 4.5.7 (Two Families Theorem)**.** *There exists an absolute constant C_1 such that the following holds. Let $P_1, P_2, \dots, P_k$ and $Q_1, Q_2, \dots, Q_k$ be sets of size $n \leqslant 2$ so that*

- *$P_i \cap Q_i = \emptyset$ for all $1 \leqslant i \leqslant k$.*
- *For every $i, j \in \{1, \dots, k\}$ with $i \neq j$ we have that $P_i \cap Q_j \neq \emptyset$ or $P_j \cap Q_i \neq \emptyset$.*

Then $k \leqslant C_1$.

The first important point is that $B^{(1)}(D)$ is rather small since D is dissociated.

Lemma 4.5.8. *We have that*

$$\left|B^{(1)}(D)\right| \leqslant C_1 |S|^4. \tag{4.33}$$

Proof. Let us prove this claim by assuming for a contradiction that $\left|B^{(1)}(D)\right| > C_1|S|^4 > C_1\left|(2S-2S)\setminus\{0\}\right|$. From (4.28), we deduce that there must be some non-zero $v \in (2S-2S)$ so that for at least C_1+1 distinct elements $d_i \in D$, we have that $d_i + v \in D$ for $1 \leqslant i \leqslant C_1+1$. Hence we can find $e_i \in D$ so that $d_i + v = e_i$ and note that

$$d_i \neq e_i \tag{4.34}$$

as $v \neq 0$. Hence, we can apply Lemma 4.5.7 with $P_i = \{d_i\}$ and $Q_i = \{e_i\}$ to find distinct i, j with $d_i \neq e_j$ and $d_j \neq e_i$. Without loss of generality, let $i = 1, j = 2$ so we deduce that

$$d_1 + e_2 = d_1 + v + d_2 = e_1 + d_2$$

and we have shown that $\{d_1, e_2\}, \{e_1, d_2\}$ are two subsets of D with equal sum.[4] As D is dissociated, these sets are equal so by (4.34) we must have that $d_1 = d_2$. This is the required contradiction as d_1, d_2 were distinct. Hence, $\left|B^{(1)}(D)\right| \leqslant C_1|S|^4$ as desired. □

We now show that $B^{(2)}(D)$ is also not too large.

Lemma 4.5.9. *We have that*

$$\left|B^{(2)}(D)\right| \leqslant C_1|S|^4 + |D||S|^4 \tag{4.35}$$

Proof. Assume for a contradiction that $\left|B^{(2)}(D)\right| > C_1|S|^4 + |D||S|^4$. For every pair $\{d, d'\} \in B^{(2)}(D)$ we can find $e, e' \in D$ with $\{e, e'\} \neq \{d, d'\}$ and $s, s' \in S$ such that (4.29) holds. So to each pair $\{d, d'\} \in B^{(2)}(D)$, we can associate a $4-$tuple $(s_d, s_{d'}, s, s') \in S^4$ (there may be more than one choice of such a tuple, but in this case we pick one arbitrarily). As we are assuming for a contradiction that $\left|B^{(2)}(D)\right| > C_1|S|^4 + |D||S|^4$, there must be some such $4-$tuple in S^4 that is associated to $C_1 + |D| + 1$ distinct pairs $\{d_i, d_i'\} \in B^{(2)}(D)$. Let $e_i, e_i' \in D$ and $s_i, s_i' \in S$ be so that

$$d_i + s_{d_i} + d_i' + s_{d_i'} = e_i + s_i + e_i' + s_i' \tag{4.36}$$

with

$$\{e_i, e_i'\} \neq \{d_i, d_i'\} \tag{4.37}$$

[4] To show that $\{d_1, e_2\}, \{d_2, e_1\}$ are subsets of D as opposed to a multisubsets, we used the skew Two Families Theorem. For this application, a very simple direct argument would have sufficed, but we will make similar applications of the Two Families Theorem later and hence try to do this in a consistent manner.

for $i = 1, \dots, C_1 + |D| + 1$, where e_i, e_i', s_i, s_i' exist by definition of $B^{(2)}(D)$. Then the assumption that all $\{d_i, d_i'\}$ are associated to the same $4-$tuple in S^4 means precisely that there exist fixed $s_d, s_{d'}, s, s' \in S$ so that

$$s_{d_i} = s_d,\ s_{d_i'} = s_{d'},\ s_i = s,\ s_i' = s' \tag{4.38}$$

for all i. Clearly there are at most $|D|$ of these indices i for which $e_i = e_i'$, so suppose without loss of generality that $e_i \neq e_i'$ for $i = 1, \dots, C_1 + 1$. We will apply Lemma 4.5.7 with $P_i = \{d_i, d_i'\}$ and $Q_i = \{e_i, e_i'\}$ for $i = 1, \dots, C_1 + 1$ and we first show that the condition $P_i \cap Q_i = \emptyset$ is satisfied. Indeed, if $d_i = e_i$ for a contradiction (the other cases when $P_i \cap Q_i \neq \emptyset$ can be handled similarly), then (4.36) would imply that $d_i' + (s_{d_1} + s_{d_1'} - s_1 - s_1') = e_i'$. In other words, $d_1' + v = e_1'$ for some $v \in 2S - 2S$ and note that $v \neq 0$ as else $d_1' = e_1'$ but, as $d_1 = e_1$, this would contradict (4.37). However, the existence of a non-zero $v \in 2S - 2S$ so that $d_1' + v \in D$ means precisely that $d_1' \in B^{(1)}(D)$ and this is impossible because we removed $B^{(1)}(D)$ from D to obtain $G^{(1)}(D)$. Lemma 4.5.7 therefore gives distinct i, j such that $P_i \cap Q_j = \emptyset = P_j \cap Q_i$ and without loss of generality, let $i = 1, j = 2$. Plugging (4.38) in (4.29) then gives

$$d_1 + d_1' - e_1 - e_1' = s_1 + s_1' - s_{d_1} - s_{d_1'} = s_2 + s_2' - s_{d_2} - s_{d_2'} = d_2 + d_2' - e_2 - e_2'.$$

So we obtain two subsets $Q = \{d_1, d_1', e_2, e_2'\}$ and $R = \{e_1, e_1', d_2, d_2'\}$ of the dissociated set D having equal sum, where we noted that these are not multisets as $P_1 \cap Q_2 = \emptyset = P_2 \cap Q_1$ by our application of Lemma 4.5.7 and that $e_i \neq e_i'$ for $i = 1, \dots, C_1 + 1$. We conclude that $Q = R$. As $\{d_1, d_1'\}, \{d_2, d_2'\}$ were distinct pairs, we may (after potentially relabeling these elements) assume that $d_1 \notin \{d_2, d_2'\}$. Then $d_1 \in Q \setminus \{d_2, d_2'\} = R \setminus \{d_2, d_2'\} = \{e_1, e_1'\}$ which gives the desired contradiction as we showed above that $P_i \cap Q_i = \emptyset$ for all i. □

From (4.33) and (4.35) we deduce that the set $G^{(2)}$ of good pairs is still large:

$$\begin{aligned}
\left|G^{(2)}\right| &\geqslant \binom{\left|G^{(1)}(D)\right|}{2} - \left|B^{(2)}(D)\right| \\
&= \binom{|D| - \left|B^{(1)}(D)\right|}{2} - \left|B^{(2)}(D)\right| \\
&\geqslant \binom{|D| - C_1|S|^4}{2} - C_1\,|S|^4 - |D||S|^4 \\
&\geqslant \frac{|D|^2}{3},
\end{aligned} \tag{4.39}$$

where in the final line we used that $|D| \geqslant 10$ and the assumption (4.22) so that $|S|^4 \leqslant \frac{|D|^6}{C|A|^5} \leqslant \frac{|D|}{C}$ as $D \subset A$ so $|D| \leqslant |A|$, and we can take C to be a sufficiently large

constant. This result that there are many good pairs in $(D+S)\cap A$, i.e. many pairs giving a unique sum in $(D+S)\cap A$, is the only result out of all the work we did in this proof so far that will be needed for the rest of the argument.

For every pair $\{d,d'\}\in G^{(2)}$, we have that $(d+s_d)+(d'+s_{d'})$ is a sum in $A+A$ and therefore it must allow for a non-trivially different representation as a sum of two elements $x,y\in A$. By Lemma 4.5.6, it cannot be the case that both of x,y lie in $(D+S)\cap A$, so we can define $x(d,d'),y(d,d')\in A$ so that $(d+s_d)+(d'+s_{d'})=x(d,d')+y(d,d')$ and $x(d,d')\in A\setminus(D+S)$.[5] We introduce some further notation. Define for each $a\in A$ the set

$$N(a):=\left\{\{d,d'\}\in G^{(2)}:x(d,d')=a\right\}. \tag{4.40}$$

so from (4.39) we obtain the inequality

$$\sum_{a\in A\setminus(D+S)}|N(a)|=\left|G^{(2)}\right|\geqslant\frac{|D|^2}{3} \tag{4.41}$$

as each pair $\{d,d'\}\in G^{(2)}$ appears in exactly one $N(a)$ with $a\notin D+S$. We now pick out those $a\in A\setminus(D+S)$ for which $N(a)$ is large. We define

$$\mathcal{N}:=\left\{a\in A\setminus(D+S):|N(a)|\geqslant\frac{|D|^2}{6|A|}\right\}. \tag{4.42}$$

We show that by a simple averaging argument, $\mathcal{N}$ is fairly large.

Lemma 4.5.10. *We have that*

$$|\mathcal{N}|\geqslant\frac{|D|^2}{6|A|}.$$

Proof. First we prove that for every $a\in A$, we have the upper bound $|N(a)|\leqslant|A|$. In fact, we show that if $\{d_1,d_1'\},\{d_2,d_2'\}\in N(a)$ are distinct, then $y(d_1,d_1')\neq y(d_2,d_2')$. As $y(d,d')\in A$ always holds, there can then be at most $|A|$ pairs in $N(a)$. Now let $\{d_1,d_1'\},\{d_2,d_2'\}\in N(a)$ with $y(d_1,d_1')=y(d_2,d_2')$, then as $x(d_1,d_1')=x(d_2,d_2')=a$ we get that

$$(d_1+s_{d_1})+(d_1'+s_{d_1'})=a+y(d_1,d_1')=a+y(d_2,d_2')=(d_2+s_{d_2})+(d_2'+s_{d_2'})$$

so that $\{d_1,d_1'\}=\{d_2,d_2'\}$ are not distinct by Lemma 4.5.6.

[5] Technically, one would have to write something like $x(\{d,d'\})$ as $\{d,d'\}$ is an unordered pair, but this is too cumbersome.

Using that $|N(a)| \leqslant |A|$ for $a \in \mathcal{N}$ and that $|N(a)| \leqslant \frac{|D|^2}{6|A|}$ for all other a, we get from (4.41) that

$$\begin{aligned}\frac{|D|^2}{3} &\leqslant \sum_{a \in A \setminus (D+S)} |N(a)| \\ &\leqslant |A|\,|\mathcal{N}| + \frac{|D|^2}{6|A|}\,|A \setminus \mathcal{N}| \\ &\leqslant |A|\,|\mathcal{N}| + \frac{|D|^2}{6}\end{aligned}$$

so that $|\mathcal{N}| \geqslant \frac{|D|^2}{6|A|}$ as desired. □

Next, we show that for at least half the elements $a \in \mathcal{N}$, there are many unordered pairs in $N(a)$ that intersect in a common element. Let us define

$$\mathcal{N}(1/3) := \left\{ a \in \mathcal{N} : \exists d(a) \in D \text{ so that } d(a) \in P \text{ for at least } \frac{|N(a)|}{3} \text{ many } P \in N(a) \right\}.$$

Lemma 4.5.11. *We have that* $|\mathcal{N}(1/3)| \geqslant \frac{|\mathcal{N}|}{2}$.

Proof. We again argue by contradiction, so assume that $|\mathcal{N}(1/3)| < \frac{|\mathcal{N}|}{2}$. Then by definition, for every $a \in \mathcal{N} \setminus \mathcal{N}(1/3)$ and every $d \in D$, d lies in less than $\frac{|N(a)|}{3}$ of all pairs in $N(a)$. Then pick any $a \in \mathcal{N} \setminus \mathcal{N}(1/3)$ and any pair $P = \{d, d'\} \in N(a)$. The number of pairs $Q \in N(a)$ which intersect P is at most $\frac{2|N(a)|}{3}$ as there are fewer than $\frac{|N(a)|}{3}$ pairs containing d, and similarly for d'. For this lemma only, we define T to be the set

$$T := \{(a, P, Q) : a \in \mathcal{N} \setminus \mathcal{N}(1/3) \text{ and } P, Q \in N(a) \text{ are disjoint}\}.$$

We have just shown that if $a \in \mathcal{N} \setminus \mathcal{N}(1/3)$, then for any $P \in N(a)$ there are at least $\frac{|N(a)|}{3}$ distinct $Q \in N(a)$ which are disjoint from P, so we get that

$$\begin{aligned}|T| &\geqslant |\mathcal{N} \setminus \mathcal{N}(1/3)| \left(\min_{a \in \mathcal{N}} |N(a)| \right) \left(\min_{a \in \mathcal{N}} \frac{|N(a)|}{3} \right) \\ &\geqslant \frac{|D|^2}{12|A|} \left(\frac{|D|^2}{6|A|} \right) \left(\frac{|D|^2}{18|A|} \right) \\ &> \frac{|D|^6}{2^{11}|A|^3}\end{aligned}$$

using Lemma 4.5.10 to get $|\mathcal{N} \setminus \mathcal{N}(1/3)| \geqslant \frac{|\mathcal{N}|}{2} \geqslant \frac{|D|^2}{12|A|}$, and that by definition of $\mathcal{N}$, $|N(a)| \geqslant \frac{|D|^2}{6|A|}$ for $a \in \mathcal{N}$. Now we can assign a sum to each element of T as follows. For each element $(a, P, Q) \in T$, define $\sigma(a, P, Q) := \sum_{x \in P} x - \sum_{x \in Q} x$. The point is

that $\sigma : T \to G$ takes each value at most C_1 times, where C_1 is the absolute constant in Lemma 4.5.7. Indeed, let us assume for a contradiction that $(a_i, P_i, Q_i) \in T$ for $i = 1, 2, \dots C_1 + 1$ all have the same image under σ. Then as $P_i \cap Q_i = \emptyset$ for all i by definition of T, Lemma 4.5.7 gives two distinct i, j so that $P_i \cap Q_j = \emptyset = P_j \cap Q_i$ and we also have $\sum_{x \in P_i} x - \sum_{x \in Q_i} x = \sigma(a_i, P_i, Q_i) = \sigma(a_j, P_j, Q_j) = \sum_{x \in P_j} x - \sum_{x \in Q_j} x$. This rearranges to $\sum_{x \in P_i \cup Q_j} x = \sum_{x \in P_j \cup Q_i} x$. But D is a dissociated set and as $P_i \cap Q_j = \emptyset = P_j \cap Q_i$, the sets $P_i \cup Q_j$ and $P_j \cup Q_i$ are subsets (and not multisubsets) of D with equal sum so we conclude that $P_i \cup Q_j = P_j \cup Q_i$. By definition of T, P_i and Q_i are disjoint and so are P_j and Q_j so we must have that $P_i = P_j$ and $Q_i = Q_j$. Finally, this implies that $a_i = a_j$ (as $P_i \in N(a_i)$ and $P_i = P_j \in N(a_j)$ but by definition (4.40) each pair P lies in exactly one $N(a)$). So we have a contradiction as we assumed that $(a_i, P_i, Q_i), (a_j, P_j, Q_j)$ were distinct. Hence, we conclude that σ takes each value at most C_1 times.

On the other hand, if $(a, P, Q) \in T$ then $P, Q \in N(a)$ which means precisely that after writing $P = \{d_1, d_1'\}$ and $Q = \{d_2, d_2'\}$, we have $x(d_i, d_i') = a$ so that

$$a + y(d_i, d_i') = (d_i + s_{d_i}) + (d_i' + s_{d_i'}),$$

for $i = 1, 2$. Subtracting this equation with $i = 2$ from that with $i = 1$ shows that

$$\begin{aligned} \sigma(a, P, Q) &= d_1 + d_1' - d_2 - d_2' \\ &= y(d_1, d_1') - y(d_2, d_2') - s_{d_1} - s_{d_1'} + s_{d_2} + s_{d_2'} \\ &\in A - A + 2S - 2S. \end{aligned}$$

Hence, $\sigma : T \to A - A + 2S - 2S$ is a map from a set of size $|T| > \frac{|D|^6}{2^{11}|A|^3}$ to a set of size $|A - A + 2S - 2S| \leqslant |A|^2|S|^4$ which takes each value at most C_1 times. We deduce that $\frac{|D|^6}{2^{11}|A|^3} < C_1|A|^2|S|^4$ and rearranging gives that

$$|S| > \left(\frac{|D|^6}{2^{11}C_1|A|^5}\right)^{\frac{1}{4}},$$

which is the required contradiction as we assumed (4.22) and we can take $C > 2^{11}C_1$. □

Let us see now what it means that for many $a \in \mathcal{N}$, namely for all $a \in \mathcal{N}(1/3)$, lots of unordered pairs in $N(a)$ contain a common element. So pick $a \in \mathcal{N}(1/3)$, then we can find an element $d(a) \in D$ and distinct pairs $P_1, P_2, \dots, P_m \in N(a)$ with $m \geqslant \frac{|N(a)|}{3} \geqslant \frac{|D|^2}{18|A|}$ (recall that $|N(a)| \geqslant \frac{|D|^2}{6|A|}$ by definition (4.42) of $\mathcal{N}$) so that each

P_i contains $d(a)$. Hence, we can write $P_i = \{d(a), d_i(a)\}$. By definition of $N(a)$, we therefore get the following list of equations

$$\begin{aligned} a + y(d(a), d_1(a)) &= (d(a) + s_{d(a)}) + (d_1(a) + s_{d_1(a)}) \\ a + y(d(a), d_2(a)) &= (d(a) + s_{d(a)}) + (d_2(a) + s_{d_2(a)}) \\ &\vdots \\ a + y(d(a), d_m(a)) &= (d(a) + s_{d(a)}) + (d_m(a) + s_{d_m(a)}), \end{aligned} \tag{4.43}$$

with $m \geqslant \frac{|D|^2}{18|A|}$. So for any $a \in \mathcal{N}(1/3)$, we get many equations like this which have a common term on the left hand side, and a common term on the right hand side. We are now almost ready to find a larger set of translates S' so that $|(D + S') \cap A| \geqslant |(D + S) \cap A| + \frac{|D|^2}{36|A|}$ and hence finish the proof. There are two cases that we need to consider based on whether many of the elements $y(d(a), d_i(a))$ are in $A \setminus (D + S)$ or in $D + S$.

The first case is straightforward now that we have (4.43). Indeed, suppose that for a single $a \in \mathcal{N}(1/3)$, at least half of the elements $y(d(a), d_i(a))$ appearing in (4.43) lie in $A \setminus (D + S)$. Without loss of generality, we may assume that $y(d(a), d_i(a)) \in A \setminus (D + S)$ for $i = 1, 2, \ldots, \frac{m}{2}$ with $m \geqslant \frac{|D|^2}{18|A|}$. Then if we set $t = d(a) + s_{d(a)} - a$, the equations (4.43) give that

$$\begin{aligned} y(d(a), d_i(a)) &= (d(a) + s_{d(a)} - a) + (d_i(a) + s_{d_i(a)}) \\ &= t + (d_i(a) + s_{d_i(a)}) \in D + (S + t) \end{aligned}$$

for $i = 1, 2, \ldots, \frac{m}{2}$. Then we can take $S' = S \cup (S + t)$ so that $|S'| \leqslant 2|S|$ and

$$|(D + S') \cap A| \geqslant |(D + S) \cap A| + \frac{m}{2} \geqslant |(D + S) \cap A| + \frac{|D|^2}{36|A|}$$

since $y(d(a), d_i(a)) \in (A \cap (D + S')) \setminus (D + S)$ for $i = 1, 2, \ldots, \frac{m}{2}$. This is the desired conclusion.

In the final case, we may assume that for every $a \in \mathcal{N}(1/3)$, at least half of the elements $y(d(a), d_i(a))$ appearing in (4.43) lie in $D + S$. Without loss of generality, we may assume that $y(d(a), d_i(a)) \in (D + S)$ for $i = 1, 2, \ldots, \frac{m}{2}$ with $m \geqslant \frac{|D|^2}{18|A|}$. Hence, we can find, for each $a \in \mathcal{N}(1/3)$ and each $1 \leqslant i \leqslant \frac{m}{2}$, the elements $e_i(a) \in D$ and $s_i(a) \in S$ so that

$$y(d(a), d_i(a)) = e_i(a) + s_i(a). \tag{4.44}$$

Recall also equation (4.43) which says that, for each such $a \in \mathcal{N}(1/3)$ and each $i = 1, 2, \ldots, \frac{m}{2}$, we have

$$a + y(d(a), d_i(a)) = (d(a) + s_{d(a)}) + (d_i(a) + s_{d_i(a)}). \tag{4.45}$$

We need one more lemma showing that, under the assumptions of this final case, for every $a \in \mathcal{N}(1/3)$, the element $e_i(a)$ must coincide with $d_i(a)$ for some i.

Lemma 4.5.12. *Assume that for every $a \in \mathcal{N}(1/3)$, we have that $y(d(a), d_i(a)) \in (D + S)$ for $i = 1, 2, \ldots, \frac{m}{2}$ and that $m \geqslant \frac{|D|^2}{18|A|}$. Then for every $a \in \mathcal{N}(1/3)$, there exists an $i \in \{1, 2, \ldots, \frac{m}{2}\}$ so that $e_i(a) = d_i(a)$.*

Assuming this lemma for the moment, we can finish the proof. Rewriting $y(d(a), d_i(a))$ using (4.44) in the equation (4.45) gives

$$a + e_i(a) + s_i(a) = (d(a) + s_{d(a)}) + (d_i(a) + s_{d_i(a)}) \tag{4.46}$$

for all $a \in \mathcal{N}(1/3)$ and $i = 1, 2, \ldots, \frac{m}{2}$. By Lemma 4.5.12, for each $a \in \mathcal{N}(1/3)$, we can find some $i \leqslant \frac{m}{2}$ so that $e_i(a) = d_i(a)$. Plugging this into (4.46) and cancelling $d_i(a) = e_i(a)$ on both sides gives

$$a + s_i(a) = d(a) + s_{d(a)} + s_{d_i(a)}$$

so that $a = d(a) + s_{d(a)} + s_{d_i(a)} - s_i(a) \in D + 2S - S$ for all $a \in \mathcal{N}(1/3)$. Hence, taking our new set of translates to be $S' = (2S - S) \cup S = 2S - S$, we get that

$$|(D + S') \cap A| \geqslant |(D + S) \cap A| + |\mathcal{N}(1/3)| \geqslant |(D + S) \cap A| + \frac{|D|^2}{12|A|}$$

as $|\mathcal{N}(1/3)| \geqslant \frac{|\mathcal{N}|}{2} \geqslant \frac{|D|^2}{12|A|}$ by Lemmas 4.5.10 and 4.5.11 and as $\mathcal{N} \subset A$ is disjoint from $D + S$ by definition (4.42). This is the desired conclusion. So we only need to prove Lemma 4.5.12.

Proof of Lemma 4.5.12. Suppose for a contradiction that the lemma is false. Then there exists some $a \in \mathcal{N}(1/3)$ so that

$$e_i(a) \neq d_i(a) \tag{4.47}$$

for all $i = 1, 2, \ldots, \frac{m}{2}$. Rewriting $y(d(a), d_i(a))$ using (4.44) in the equation (4.45) gives

$$a + e_i(a) + s_i(a) = (d(a) + s_{d(a)}) + (d_i(a) + s_{d_i(a)}) \tag{4.48}$$

for this supposed counterexample $a \in \mathcal{N}(1/3)$, and every $i = 1, 2, \ldots, \frac{m}{2}$. Hence, if we write $t' = a - (d(a) + s_{d(a)})$, then for each such i we have that

$$\begin{aligned} d_i(a) - e_i(a) &= a - (d(a) + s_{d(a)}) + s_i(a) - s_{d_i(a)} \\ &= t' + s_i(a) - s_{d_i(a)} \in t' + S - S. \end{aligned}$$

However, the set $t' + S - S$ has at most $|S|^2 \leqslant |S|^4 \leqslant \frac{|D|^6}{C|A|^5} < \frac{|D|^2}{36C_1|A|} \leqslant \frac{m}{2C_1}$ many elements by assumption (4.22), as $|D| \leqslant |A|$ since $D \subset A$, and by choosing C sufficiently large in terms of the absolute constant C_1. By the pigeonhole principle, out of all $\frac{m}{2}$ possible indices i there exist $C_1 + 1$ distinct such indices, say $i = 1, \ldots, C_1 + 1$, so that

$$d_i(a) - e_i(a) = d_j(a) - e_j(a) \tag{4.49}$$

for all $1 \leqslant i, j \leqslant C_1 + 1$. Since $e_i(a) \neq d_i(a)$ by (4.47), the sets $P_i = \{e_i(a)\}, Q_i = \{d_i(a)\}$ satisfy $P_i \cap Q_i = \emptyset$ so we can apply Lemma 4.5.7 to deduce that, without loss of generality, $P_1 \cap Q_2 = \emptyset = P_2 \cap Q_1$. Rearranging (4.49) and the fact that D is dissociated then yield $\{d_i(a), e_j(a)\} = \{d_j(a), e_i(a)\}$ so by (4.47) we conclude that $d_i(a) = d_j(a)$. This is the required contradiction (as for a fixed a, the elements $d_1(a), \ldots, d_m(a)$ that we defined for the equations (4.43) are all distinct). This finishes the proof of the lemma. □

This concludes the proof of Proposition 4.5.4. □

Chapter 5

On a problem of Erdős and Sárközy about sequences with no term dividing the sum of two larger terms

5.1 Introduction

This chapter is based on [5]. Let us begin with the following definition from a 1970 paper [29] of Erdős and Sárközy.

Definition 5.1.1. Let $A \subset \mathbf{N}$. We say that A has *property* P if there are no three numbers $x, y, z \in A$ with $z < x, y$ and $z|x + y$.

So a sequence has property P precisely if it contains no term which divides the sum of two larger terms. The main result of [29] states that an infinite sequence A of positive integers having property P must have density 0. The density of infinite sequences with property P has been studied in greater detail by various authors, see [3],[18],[60]. In the original paper [29], Erdős and Sárközy mention the following finite version of the problem.

Problem 5.1.2 (Erdős - Sárközy [29])**.** *Let $n \geqslant 1$ be an integer and $A \subset \{1, 2, \ldots, n\}$ have property P. Must it be the case that*

$$|A| \leqslant \left\lfloor \frac{n}{3} \right\rfloor + 1?$$

The paper mentions that Szemerédi proved that if a set $A \subset [n]$ has size $|A| > \left\lfloor \frac{n}{3} \right\rfloor + 1$, then it contains distinct $x, y, z \in A$ with $z|x+y$ and $\frac{x+y}{z} \neq 2$. This conclusion

is significantly weaker than what would be needed to contradict A having property P however, and the author is not aware of any results in the literature improving on this partial result. In fact, Erdős mentioned this particular problem in a large number of his open problem papers [20–28, 30] written between 1970 and 1996. In the 1973 paper [20] and several later papers, Erdős asks for a proof of the slightly weaker claim that $|A| \leqslant \frac{n}{3} + O(1)$ if $A \subset [n]$ is a set with property P. In his final open problems paper [27], Erdős offers \$100 for a resolution of this problem.

Problem 5.1.3 (Erdős)**.** *Is there is an absolute constant C such that if $n \geqslant 1$ and $A \subset \{1, 2, \ldots, n\}$ has property P, then $|A| \leqslant \frac{n}{3} + C$?*

In some of these papers, Erdős states that the bound $|A| \leqslant \left\lfloor \frac{n}{3} \right\rfloor + 1$, if true, is optimal in light of the example $A = \left\{ \left\lceil \frac{2n}{3} \right\rceil, \ldots, n \right\}$. This is a typo however, and the exact bound should be $|A| \leqslant \left\lceil \frac{n}{3} \right\rceil$ with the corresponding tight example being $A = \left\{ \left\lfloor \frac{2n}{3} \right\rfloor + 1, \ldots, n \right\}$ which is easily seen to have property P. The previous best known bound existing in the literature seems to be Erdős's observation that $|A| \leqslant \left\lceil \frac{n}{2} \right\rceil$ which is more or less trivial.[1]

We will establish the following resolution of these problems.

Theorem 5.1.4. *There is an absolute constant C such that for all $n \in \mathbf{N}$, if $A \subset \{1, 2, \ldots, n\}$ has property P, then $|A| \leqslant \frac{n}{3} + C$.*

Theorem 5.1.5. *For all sufficiently large $n \in \mathbf{N}$, if $A \subset \{1, 2, \ldots, n\}$ has property P, then $|A| \leqslant \left\lceil \frac{n}{3} \right\rceil$. Moreover, this bound is tight for all such n since $\left\{ \left\lfloor \frac{2n}{3} \right\rfloor + 1, \ldots, n \right\}$ is a subset of $[n]$ with property P and size $\left\lceil \frac{n}{3} \right\rceil$.*

Theorem 5.1.5 in fact implies Theorem 5.1.4 by choosing C sufficiently large. In order to give a streamlined and relatively easy to read version of the argument, we make no effort to optimise the value of C.

5.2 Notation

For a set $X \subset \mathbf{N}$ we define the dilated set $q \cdot X = \{qx : x \in X\}$. For real numbers $\alpha \leqslant \beta$, we define

$$(\alpha, \beta] := \{m \in \mathbf{N} : \alpha < m \leqslant \beta\}$$

[1] A set with property P certainly cannot contain two numbers with one dividing the other, and it is well-known and not hard to prove that a subset of $[n]$ with this property has size at most $\left\lceil \frac{n}{2} \right\rceil$.

and similarly for other types of intervals. We also write $[\beta]$ for $[1, \beta]$. Throughout the paper, whenever we write the word 'interval', we mean a set of consecutive integers. For a set $X \subset \mathbf{Z}$, we define $\operatorname{diam} X := \max X - \min X$ and $\gcd_*(X) := \gcd(X - X)$ is the greatest common divisor of all differences $x - x'$ with $x, x' \in X$. Equivalently, $\gcd_*(X)$ is the largest integer d such that X is contained in an arithmetic progression with common difference d.

In our proofs of Theorems 5.1.4 and 5.1.5, we will work with a set $A \subset [n]$ having property P. We will need to consider parts of A lying in various subintervals of $[n]$ and hence we will use the following notation for real numbers $0 \leqslant \alpha < \beta \leqslant 1$:

$$A_{(\alpha,\beta]} := A \cap (\alpha n, \beta n],$$

and similarly for other types of intervals. The value of n is hidden in this notation, but it will be clear from context. Finally, for a positive integer q and a residue a mod q, we write

$$A^{a(q)}_{(\alpha,\beta]} := A_{(\alpha,\beta]} \cap (a + q \cdot \mathbf{N}) = A \cap (\alpha n, \beta n] \cap (a + q \cdot \mathbf{N}).$$

5.3 Preliminaries

We begin with some easy but useful observations.

Lemma 5.3.1. *If A has property P, then*

(1) A is disjoint from $k \cdot A$ for any integer $k \geqslant 2$.

(2) In particular, for any integer $k \geqslant 2$, the sets $A, k \cdot A, k^2 \cdot A, k^3 \cdot A, \ldots$ are pairwise disjoint.

(3) $2 \cdot A$ is disjoint from $k \cdot A$ for any integer $k \geqslant 3$.

Proof. First, if $A \cap (k \cdot A)$ is non-empty for some $k \geqslant 2$, then there exist $a, a' \in A$ with $a = ka'$ so that $a > a'$ and $a' | 2ka' = a + a$ contradicting that A has property P. This proves (1) from which (2) follows immediately. For (3), we need to show that $2 \cdot A$ and $k \cdot A$ are disjoint for $k \geqslant 3$. If $(2 \cdot A) \cap (k \cdot A) \neq \emptyset$, then there exist $a, a' \in A$ with $2a = ka'$ so $a > a'$ and $a' \mid ka' = a + a$ giving a contradiction. □

Lemma 5.3.2. *Let $A \subset [n]$ have property P, let k, a, q be positive integers and let $0 \leqslant \alpha \leqslant 1$. Suppose that B is a set such that $k \cdot B$ consists exclusively of integer multiples of numbers in $A_{[\alpha]} = A \cap [\alpha n]$, that every number in $k \cdot B$ is congruent to a mod q and that I is an interval so that $k \cdot B \subset I$. Then*

$$|B| + \left|\left(A_{(\alpha,1]} + A_{(\alpha,1]}\right) \cap I \cap (a + q \cdot \mathbf{N})\right| < \frac{|I|}{q} + 1. \tag{5.1}$$

Proof. We show that $k \cdot B$ and $\left(A_{(\alpha,1]} + A_{(\alpha,1]}\right) \cap I \cap (a + q \cdot \mathbf{N})$ are disjoint sets. If not, there exists some $b \in B$ so that $kb \in A_{(\alpha,1]} + A_{(\alpha,1]}$ contradicting that A has property P as kb is a multiple of some number in $A_{[\alpha]}$ by assumption. Hence, $k \cdot B$ and $\left(A_{(\alpha,1]} + A_{(\alpha,1]}\right) \cap I \cap (a + q \cdot \mathbf{N})$ are disjoint sets contained in $I \cap (a + q \cdot \mathbf{N})$. As I is an interval, we have the bound $|I \cap (a + q \cdot \mathbf{N})| < \frac{|I|}{q} + 1$ so (5.1) follows. □

We will crucially make use of the Theorem 2.0.3 of Bardaji and Grynkiewicz from Chapter 2. It is a lesser-known version of Freiman's $3k - 4$ theorem and concludes from a set S having doubling $|S+S| \leqslant 3|S|-4$ that $S+S$ contain a long progression, rather than that S is itself contained in a short progression as in the standard $3k-4$ theorem. We also note that Freiman already proved that if S is a set of integers with doubling $|S + S| \leqslant 3|S| - 4$, then $S + S$ contains an arithmetic progression of length $2|S| - 1$, so that Theorem 2.0.3 can be viewed as an extension for sumsets with distinct summands. For convenience, we state the following corollary which is enough for all applications in this chapter.

Theorem 5.3.3. *Let S, T be non-empty subsets of $\mathbf{Z}$. Then one of the following conclusions holds:*

(1) $|S + T| \geqslant |S| + |T| + \min(|S|, |T|) - 3$.

(2) *$S + T$ contains an arithmetic progression with common difference* $\gcd_*(S + T)$ *and length* $|S| + |T| - 1$.

Proof. Let S, T be non-empty subsets of $\mathbf{Z}$ and, after translating, we may assume that $\min S = \min T = 0$. Suppose that (1) does not hold so that $|S+T| \leqslant |S|+|T|+\min(|S|, |T|) - 4$. Let $d = \gcd_*(S + T)$ and note that d divides $\gcd_*(S)$ and $\gcd_*(T)$ so that each of the three sets S, T and $S + T$ lies in an arithmetic progression with common difference d. As $\min S = \min T = 0$, we see that S and T consist of multiples of d only. Define $S' = \frac{1}{d} \cdot S$ and $T' = \frac{1}{d} \cdot T$ so S' and T' are non-empty subsets of

$\mathbf{Z}$ with $\gcd_*(S'+T')=1$ and $|S'+T'|=|S+T|\leqslant |S'|+|T'|+\min(|S'|,|T'|)-4$. Theorem 2.0.3 then implies that $S'+T'$ contains an arithmetic progression with common difference 1 and length $|S'|+|T'|-1$. As $S+T=d\cdot(S'+T')$, item (2) holds as desired. □

5.4 The proof

We are now ready to begin the proof of Theorems 5.1.4 and 5.1.5. We will use induction on n to prove the following theorem which simultaneously implies both Theorem 5.1.4 and Theorem 5.1.5.

Theorem 5.4.1. *There exist absolute constants $\delta>0$ and C such that the following holds. Let $n\in\mathbf{N}$ and let $A\subset[n]$ be a set with property P. Then*

$$|A|\leqslant \max\left(\left\lceil\frac{n}{3}\right\rceil,\left(\frac{1}{3}-\delta\right)n+C\right). \tag{5.2}$$

Throughout the paper, we assume that C is a sufficiently large constant. First note that when $n\leqslant C$, the bound (5.2) holds trivially. So from now on we may assume that $n>C$ is sufficiently large. Our induction hypothesis is that for all $m<n$, the upper bound $\max\left(\left\lceil\frac{m}{3}\right\rceil,\left(\frac{1}{3}-\delta\right)m+C\right)$ holds for subsets of $[m]$ having property P. Assume henceforth that $A\subset[n]$ has property P. To prove (5.2), we split the argument into three cases based on the size of the set

$$A_{(\frac{2}{3},1]}:=A\cap\left(\frac{2n}{3},n\right].$$

The first two cases, in which A has density at least $2/3$ or between $1/2$ and $2/3$ on I_n, admit efficient and clean proofs based on the above variant of Freiman's $3k-4$ theorem. The final case, in which A has density less than $1/2$ on I_n, is quite involved and based in part on a number of delicate ad hoc arguments. This is due to the fact that if A has density less than $1/2$, then most of $A\cap I_n$ may now be contained in a proper subprogression of I_n. This potentially makes the crucial sumset $A\cap I_n+A\cap I_n$ automatically avoid multiples of most numbers in $[1,2n/3)$, making it generally impossible to obtain good bounds on $|A|$ based solely on the fact that this sumset does not contain any multiples of numbers in $A\cap[1,2n/3]$ when A has property P. Hence, we are forced to consider various sumsets $A\cap J_n+A\cap J_n$ for other progression J_n than I_n. For this reason, we have decided to omit the final case from this thesis,

referring the interested reader to the actual paper [5]. We do note that we manage to show in that paper that if a set A with property P falls into this final case, then A has size $|A| \leqslant (1/3 - \varepsilon)n$ and hence cannot have size close to the optimal value of $n/3$, so there is room for some slack.

The three cases of our proof can be handled independently; we shall give the proof of the super dense case, which we shall refer to as Case 1, in which $A_{(\frac{2}{3},1]} := A \cap \left(\frac{2n}{3}, n\right]$ has size at least $\frac{2n}{9} + \frac{4}{3}$ in section 5, and that of the moderately dense case, which we call Case 2, in which $\frac{n}{6} + 24 \leqslant |A_{(\frac{2}{3},1]}| < \frac{2n}{9} + \frac{4}{3}$ in section 6. It is interesting to note that for large enough n, the only examples of sets with property P and size very close to $\left\lceil \frac{n}{3} \right\rceil$ seem to be sets containing almost all of $\left(\frac{2n}{3}, n\right]$.[2] One might therefore expect that the 'super dense' case where $\left|A_{(\frac{2}{3},1]}\right|$ is large would cause the most trouble in the proof. This does not seem to be true however, and the first case that we consider, where $A_{(\frac{2}{3},1]}$ has density at least $\frac{2}{3}$ on $\left(\frac{2n}{3}, n\right]$, has a far simpler proof than the remaining cases.

5.5 The super dense case

In this case, we assume that $\left|A_{(\frac{2}{3},1]}\right| \geqslant \frac{2n}{9} + \frac{4}{3}$. Note that $A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]} \subset \left(\frac{4n}{3}, 2n\right]$ so we get the trivial estimate

$$\left|A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}\right| \leqslant \left\lceil \frac{2n}{3} \right\rceil.$$

As $3\left|A_{(\frac{2}{3},1]}\right| \geqslant \frac{2n}{3} + 4$ by assumption, we get $3\left|A_{(\frac{2}{3},1]}\right| - 4 \geqslant \left\lceil \frac{2n}{3} \right\rceil \geqslant \left|A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}\right|$ and hence $A_{(\frac{2}{3},1]}$ has small enough doubling so that item (1) in Theorem 5.3.3 does not hold when $S = T = A_{(\frac{2}{3},1]}$. Hence, Theorem 5.3.3 implies that conclusion (2) from the same theorem holds and $A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$ must contain a long progression. We show that $\gcd_*\left(A_{(\frac{2}{3},1]}\right) = 1$. Indeed, $A_{(\frac{2}{3},1]}$ is contained in a progression with common difference $\gcd_*\left(A_{(\frac{2}{3},1]}\right) = d$ so if $d \geqslant 2$, then $\left|A_{(\frac{2}{3},1]}\right| < \frac{n}{6} + 1$ since $A_{(\frac{2}{3},1]} \subset \left(\frac{2n}{3}, n\right]$ which would contradict the assumption of Case 1. Hence, conclusion (2) in Theorem 5.3.3 implies that $A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$ contains an interval Q of length at least $2\left|A_{(\frac{2}{3},1]}\right| - 1 > \frac{4n}{9} + 1$. Now note that any integer $x \in \left[\frac{4n}{9} + 1\right]$ has an integer multiple in every interval of length x and hence also in $Q \subset A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$, so we deduce that $x \notin A$ as A has property P. Consider $s = \min A$, and we observe that $s > \frac{4n}{9} + 1$ by what we just showed. Without loss of generality, we may assume that $s \leqslant \left\lfloor \frac{2n}{3} \right\rfloor$ as $A \subset [s, n]$ so the desired

[2] In fact, this is true and it can be proved with similar arguments to those used in the proof of Theorem 5.4.1.

bound (5.2) holds trivially otherwise. Next, we find that $A - \{s\}$ cannot contain two numbers summing to 0 mod s as else there would be two numbers in A larger than $s = \min A$ with s dividing their sum, violating property P. Hence $\left|A \cap (s, 2s]\right| \leqslant \left\lfloor \frac{s-1}{2} \right\rfloor$. If $\frac{n}{2} \leqslant s \leqslant \left\lfloor \frac{2n}{3} \right\rfloor$, this gives in total

$$|A| \leqslant 1 + \left\lfloor \frac{s-1}{2} \right\rfloor \leqslant 1 + \left\lfloor \frac{\left\lfloor \frac{2n}{3} \right\rfloor - 1}{2} \right\rfloor \leqslant \left\lceil \frac{n}{3} \right\rceil.$$

In the remaining case where $\frac{4n}{9} + 1 < s < \frac{n}{2}$, we trivially bound the number of elements of A in $(2s, n]$ by $n - 2s$. In total we get

$$|A| \leqslant 1 + \frac{s-1}{2} + n - 2s = n + \frac{1}{2} - \frac{3s}{2} \leqslant \frac{n}{3} - 1,$$

as $s > \frac{4n}{9} + 1$. Thus, we have proved the desired bound (5.2) in Case 1.

5.6 The moderately dense case

In Case 2, we assume that $\frac{n}{6} + 24 \leqslant \left|A_{(\frac{2}{3},1]}\right| < \frac{2n}{9} + \frac{4}{3}$. So $A_{(\frac{2}{3},1]}$ has density roughly between $\frac{1}{2}$ and $\frac{2}{3}$ on the interval $\left(\frac{2n}{3}, n\right]$. We begin with a useful lemma about sets which have density greater than half on an interval.

Lemma 5.6.1. *Let $U \subset [k+1, k+m]$ be a set of integers, let q be a positive integer and a be a residue modulo q. If $|U| \geqslant \frac{m}{2} + \frac{q}{2}$, then the number of integers in the sumset $U + U$ that are $a \pmod q$ is at least $\frac{2}{q}|U| - 1$.*

Proof. Let $U_i = U \cap (i + q\cdot\mathbf{N})$ and pair the sets U_i, U_{a-i} (some of the U_i may be paired with themselves). Now look at the pair for which $|U_i| + |U_{a-i}|$ is maximal, say it is U_j, U_{a-j}. Then certainly $|U_j| + |U_{a-j}| \geqslant \frac{2}{q}|U| \geqslant \frac{m}{q} + 1$ and in particular both U_j, U_{a-j} are non-empty as we can trivially bound $|U_i| < \frac{m}{q} + 1$ for all i. Using the well-known trivial lower bound $|X + Y| \geqslant |X| + |Y| - 1$ for the sumset of two non-empty sets of integers X, Y, we get that $|U_j + U_{a-j}| \geqslant |U_j| + |U_{a-j}| - 1 \geqslant \frac{2}{q}|U| - 1$ as desired. □

Note also that the same result holds true when we consider subsets U of an arithmetic progression with common difference $d > 1$ as long as there is no obvious modular reason preventing it.

Lemma 5.6.2. *Let $U \subset \{k+d, k+2d, \ldots, k+md\}$, let q be a positive integer coprime to d and a be a residue modulo q. If $|U| \geqslant \frac{m}{2} + \frac{q}{2}$, then the number of integers in the sumset $U + U$ that are $a \pmod q$ is at least $\frac{2}{q}|U| - 1$.*

Proof. The proof is the same as that of Lemma 5.6.1, except that we have to use the assumption that d and q are coprime to deduce the upper bound $|U_i| < \frac{m}{q} + 1$ for all i, where $U_i = U \cap (i + q \cdot \mathbf{N})$. □

We return to the main analysis of Case 2. First we will construct from A an auxiliary set B_1 as follows. For every number $a \in A$ with $a \leqslant \frac{2n}{3}$ there is a unique power of 2, say 2^{j_a}, so that $2^{j_a} a \in \left(\frac{n}{3}, \frac{2n}{3}\right]$. Call B_1 the set of all numbers obtained in this way, so

$$B_1 = \left\{ 2^{j_a} a : a \in A \cap \left[\frac{2n}{3}\right] \right\} \tag{5.3}$$

and note that B_1 is a subset of $\left(\frac{n}{3}, \frac{2n}{3}\right]$. Observe that

$$|A| = |B_1| + \left|A_{(\frac{2}{3},1]}\right|, \tag{5.4}$$

because $|B_1| = \left|A \cap \left[\frac{2n}{3}\right]\right|$ since coincidences of the form $2^{j_a} a = 2^{j_b} b$ with $a \neq b$ are impossible by conclusion (2) in Lemma 5.3.1. In other words, the map $a \mapsto 2^{j_a} a$ is an injection. Also observe that any number in B_1 is a multiple of a number in $A \cap \left[\frac{2n}{3}\right]$ so as A has property P, we retain the property that $A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$ contains no multiples of any element in B_1. Our basic proof strategy in Case 2 is to show that many numbers in $\left(\frac{n}{3}, \frac{2n}{3}\right]$ do have a multiple in the sumset $A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$, so as these numbers cannot lie in B_1, we get an upper bound on $|B_1|$ which hopefully is strong enough to let us conclude the desired bound on $|A|$ using (5.4).

We cover $B_1 \subset \left(\frac{n}{3}, \frac{2n}{3}\right]$ with the following sets. On the left half of the interval $\left(\frac{n}{3}, \frac{2n}{3}\right]$ we split up B_1 into residue classes modulo 3, so let

$$B_1^{\mathrm{L},i(3)} = B_1 \cap \left(\frac{n}{3}, \frac{n}{2}\right] \cap (i + 3 \cdot \mathbf{N})$$

for $i = 0, 1, 2$. On the right half of the interval $\left(\frac{n}{3}, \frac{2n}{3}\right]$ we do the same but with residue classes modulo 4, so let

$$B_1^{\mathrm{R},i(4)} = B_1 \cap \left(\frac{n}{2}, \frac{2n}{3}\right] \cap (i + 4 \cdot \mathbf{N})$$

for $i = 0, 1, 2, 3$. As we are in Case 2, we have that $\left|A_{(\frac{2}{3},1]}\right| \geqslant \frac{n}{6} + 24$ so that $A_{(\frac{2}{3},1]} \subset \left(\frac{2n}{3}, n\right]$ satisfies the assumption of Lemma 5.6.1 with modulus $q = 12$. Applying Lemma 5.6.1 to the set $A_{(\frac{2}{3},1]}$, we conclude that for each $0 \leqslant j < 12$:

$$\left|\left(A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}\right) \cap (j + 12 \cdot \mathbf{N})\right| \geqslant \frac{1}{6}\left|A_{(\frac{2}{3},1]}\right| - 1, \tag{5.5}$$

and note that $A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]} \subset \left(\frac{4n}{3}, 2n\right]$. Further, for each i observe that $4 \cdot B_1^{\mathrm{L},i(3)} \subset \left(\frac{4n}{3}, 2n\right] \cap (4i + 12 \cdot \mathbf{N})$ and that $3 \cdot B_1^{\mathrm{R},i(4)} \subset \left(\frac{4n}{3}, 2n\right] \cap (3i + 12 \cdot \mathbf{N})$. By definition (5.3) of B_1, the sets $3 \cdot B_1$ and $4 \cdot B_1$ consist of multiples of numbers in $A \cap \left[\frac{2n}{3}\right]$ so we can apply Lemma 5.3.2 with $\alpha = \frac{2}{3}$, $q = 12$, $I = \left(\frac{4n}{3}, 2n\right]$ and $B = B_1^{\mathrm{L},i(3)}, B_1^{\mathrm{R},i(4)}$ for each i. Plugging in the lower bound (5.5) in the inequality (5.1) from Lemma 5.3.2 gives

$$\frac{1}{6}\left|A_{(\frac{2}{3},1]}\right| - 1 + \left|B_1^{\mathrm{L},i(3)}\right| \leqslant \frac{n}{18} + 1,$$
$$\frac{1}{6}\left|A_{(\frac{2}{3},1]}\right| - 1 + \left|B_1^{\mathrm{R},i(4)}\right| \leqslant \frac{n}{18} + 1.$$

Hence we conclude

$$\left|A_{(\frac{2}{3},1]}\right| + 6\left|B_1^{\mathrm{L},i(3)}\right| \leqslant \frac{n}{3} + 12, \tag{5.6}$$
$$\left|A_{(\frac{2}{3},1]}\right| + 6\left|B_1^{\mathrm{R},i(4)}\right| \leqslant \frac{n}{3} + 12. \tag{5.7}$$

Having obtained the two inequalities above, we may assume for the remainder of the argument in Case 2 that $\left|B_1^{\mathrm{L},i(3)}\right| < \frac{|B_1|}{6} + 2$ and $\left|B_1^{\mathrm{R},i(4)}\right| < \frac{|B_1|}{6} + 2$ for all i as otherwise we could plug in (5.6) or (5.7) in (5.4) to conclude that $|A| = \left|A_{(\frac{2}{3},1]}\right| + |B_1| \leqslant \frac{n}{3}$. We will use these extra assumptions in the final part of the argument in Case 2.

The main idea behind our proof in Case 2 is to apply Theorem 5.3.3 in a suitable way to $A_{(\frac{2}{3},1]}$. It is tempting to try applying Theorem 5.3.3 to the set $A_{(\frac{2}{3},1]}$ directly. This does not seem to be enough however, and we first split $A_{(\frac{2}{3},1]}$ into the sets E and O consisting of the even/odd numbers in $A_{(\frac{2}{3},1]}$. The idea is then to apply Theorem 5.3.3 to whichever of the two sets E or O contains most of $A_{(\frac{2}{3},1]}$. We shall continue under the assumption that $|O| \geqslant \frac{\left|A_{(\frac{2}{3},1]}\right|}{2}$, but the same proof works when $|E| \geqslant \frac{\left|A_{(\frac{2}{3},1]}\right|}{2}$ (after interchanging the roles of E and O in what follows). Note that $\gcd_*(O)$ is even, and if it is at least 4 then we would get $|O| < \frac{n}{12} + 1$ since $O \subset \left(\frac{2n}{3}, n\right]$. Because we are assuming in Case 2 that $\left|A_{(\frac{2}{3},1]}\right| \geqslant \frac{n}{6} + 24$, we must have that $|O| \geqslant \frac{n}{12} + 12$ so that $\gcd_*(O) = 2$. We apply Theorem 5.3.3 to the sumset $O + O$ to deduce that either this sumset has size at least $3|O| - 3$, or else that it contains a long arithmetic progression, and we prove the desired bound on $|A|$ in both cases.

Assume first that $|O+O| \leqslant 3|O|-4$, then conclusion (2) in Theorem 5.3.3 holds so that $O+O$ contains an arithmetic progression Q with common difference $\gcd_*(O) = 2$ and size $2|O| - 1 \geqslant \left|A_{(\frac{2}{3},1]}\right| - 1$. Also note that $O + O \subset \left(\frac{4n}{3}, 2n\right]$ is fully contained within the even integers. Hence we can find an even integer t so that

$$Q = \left\{t + 4, t + 6, \ldots, t + 2\left|A_{(\frac{2}{3},1]}\right|\right\} \subset O + O \subset \left(\frac{4n}{3}, 2n\right]. \tag{5.8}$$

So Q contains at least $\frac{|Q|-1}{2}$ multiples of 4, and at least $\frac{|Q|-2}{3}$ multiples of 6. Now, we divide all the multiples of 4 in Q by 4 and note that the set of the resulting quotients is contained in $\frac{1}{4}\cdot Q \subset \left(\frac{n}{3}, \frac{n}{2}\right]$ by (5.8). Similarly we divide all the multiples of 3 in Q by 3 and in this case the resulting quotients lie in $\frac{1}{3}\cdot Q \subset \left(\frac{4n}{9}, \frac{2n}{3}\right]$. Let A' be the set of all the quotients obtained in this way from Q, so

$$A' := \left(\frac{1}{3}\cdot Q \cup \frac{1}{4}\cdot Q\right) \cap \mathbf{N}. \tag{5.9}$$

Note that A' is a subset of $\left(\frac{n}{3}, \frac{2n}{3}\right]$ and that each element of A' has an integer multiple in $Q \subset A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$. We now show that $\frac{1}{3}\cdot Q$ and $\frac{1}{4}\cdot Q$ are disjoint. Suppose for a contradiction that $\frac{1}{3}\cdot Q$ and $\frac{1}{4}\cdot Q$ intersect, then it would have to be the case that $\max \frac{1}{4}\cdot Q \geqslant \min \frac{1}{3}\cdot Q$ so that plugging in the values of $\max Q$ and $\min Q$ from (5.8) would give

$$\frac{t}{4} + \frac{\left|A_{(\frac{2}{3},1]}\right|}{2} \geqslant \max \frac{1}{4}\cdot Q \geqslant \min \frac{1}{3}\cdot Q \geqslant \frac{t}{3} + \frac{4}{3}$$

whence $\left|A_{(\frac{2}{3},1]}\right| \geqslant \frac{t}{6} + \frac{8}{3} > \frac{2n}{9} + 2$ since $t > \frac{4n}{3} - 4$ by (5.8). This however contradicts our Case 2 assumption $\left|A_{(\frac{2}{3},1]}\right| \leqslant \frac{2n}{9} + \frac{4}{3}$. Hence we deduce

$$|A'| = \left|\left(\frac{1}{4}\cdot Q\right) \cap \mathbf{N}\right| + \left|\left(\frac{1}{3}\cdot Q\right) \cap \mathbf{N}\right| \geqslant \frac{|Q|-1}{2} + \frac{|Q|-2}{3} \geqslant \frac{5\left|A_{(\frac{2}{3},1]}\right|}{6} - 2, \tag{5.10}$$

because $|Q| \geqslant \left|A_{(\frac{2}{3},1]}\right| - 1$ by the definition (5.8) of Q. Next, as $\left|A_{(\frac{2}{3},1]}\right| \geqslant \frac{n}{6} + 24$ by the assumptions of Case 2, Lemma 5.6.1 gives that

$$\left|\left(A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}\right) \cap (3 + 6\cdot\mathbf{N})\right| \geqslant \frac{\left|A_{(\frac{2}{3},1]}\right|}{3} - 1. \tag{5.11}$$

So we can find many numbers in $A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$ that are 3 mod 6 and consider the set $\frac{1}{3}\cdot\left(\left(A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}\right) \cap (3 + 6\cdot\mathbf{N})\right)$ of quotients obtained by dividing these numbers by 3. Clearly, this set is contained in $\frac{1}{3}\cdot\left(A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}\right) \subset \frac{1}{3}\cdot\left(\frac{4n}{3}, 2n\right] = \left(\frac{4n}{9}, \frac{2n}{3}\right]$. Moreover, this set of quotients consists of odd numbers only so it is disjoint from $\left(\frac{1}{3}\cdot Q\right) \cap \mathbf{N}$ because Q is a subset of $O + O$ and therefore contains only even numbers. Since this set of quotients $\frac{1}{3}\cdot\left(\left(A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}\right) \cap (3 + 6\cdot\mathbf{N})\right)$ is contained in $\left(\frac{4n}{9}, \frac{2n}{3}\right]$, its intersection with $\left(\frac{1}{4}\cdot Q\right) \cap \mathbf{N}$ trivially has size at most $\left|\left(\frac{4n}{9}, \frac{n}{2}\right] \cap (1 + 2\cdot\mathbf{N})\right| < \frac{n}{36} + 1$ because $\frac{1}{4}\cdot Q \subset \left(\frac{n}{3}, \frac{n}{2}\right]$ by (5.8). We now add this set of quotients to A' to obtain a larger set A'' defined by

$$A'' := A' \cup \frac{1}{3}\cdot\left(\left(A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}\right) \cap (3 + 6\cdot\mathbf{N})\right)$$

$$= \left(\left(\frac{1}{3} \cdot Q \cup \frac{1}{4} \cdot Q \right) \cap \mathbf{N} \right) \cup \frac{1}{3} \cdot \left(\left(A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]} \right) \cap (3 + 6 \cdot \mathbf{N}) \right).$$

By (5.11), we see that we have added at least

$$|A'' \setminus A'| > \left| \left(A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]} \right) \cap (3 + 6 \cdot \mathbf{N}) \right| - \frac{n}{36} - 1 \geqslant \frac{\left| A_{(\frac{2}{3},1]} \right|}{3} - \frac{n}{36} - 2$$

new elements to A' to obtain A''. Combining this with the lower bound (5.10) gives

$$|A''| > \frac{5\left| A_{(\frac{2}{3},1]} \right|}{6} - 2 + \frac{\left| A_{(\frac{2}{3},1]} \right|}{3} - \frac{n}{36} - 2 = \frac{7\left| A_{(\frac{2}{3},1]} \right|}{6} - \frac{n}{36} - 4. \tag{5.12}$$

As we noted right after the definition (5.9) of A', A' is a subset of $\left(\frac{n}{3}, \frac{2n}{3}\right]$ and each of its members has a multiple in $A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$. The same is true for A'' as this set is obtained from A' by adding the set $\frac{1}{3} \cdot \left(\left(A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]} \right) \cap (3 + 6 \cdot \mathbf{N}) \right) \subset \left(\frac{4n}{9}, \frac{2n}{3}\right]$ and all of its elements also have a multiple in $A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$. As A has property P and B_1 consists of multiples of numbers in $A \cap \left[\frac{2n}{3}\right]$ by definition (5.3), A'' must be disjoint from B_1 and since both B_1 and A'' are subsets of $\left(\frac{n}{3}, \frac{2n}{3}\right]$ we deduce that $\left\lceil \frac{n}{3} \right\rceil \geqslant |A''| + |B_1|$. Using the lower bound (5.12) for $|A''|$ in this inequality then yields the desired bound for $|A|$ as follows

$$\begin{aligned} \left\lceil \frac{n}{3} \right\rceil &\geqslant |A''| + |B_1| \\ &> \frac{7\left| A_{(\frac{2}{3},1]} \right|}{6} - \frac{n}{36} - 4 + |B_1| \\ &\geqslant \left| A_{(\frac{2}{3},1]} \right| + |B_1| \\ &= |A|, \end{aligned}$$

where for the third inequality we used that $\frac{\left| A_{(\frac{2}{3},1]} \right|}{6} \geqslant \frac{n}{36} + 4$ because we assume that $\left| A_{(\frac{2}{3},1]} \right| \geqslant \frac{n}{6} + 24$ in Case 2, and for the final equality we used (5.4).

This leaves us with alternative (1) in Theorem 5.3.3, and hence we can now assume that

$$|O + O| \geqslant 3|O| - 3 \geqslant \frac{3\left| A_{(\frac{2}{3},1]} \right|}{2} - 3. \tag{5.13}$$

Note that the sumset $O + O$ is fully contained in the set of even numbers in $\left(\frac{4n}{3}, 2n\right]$. Recall that by the discussion following inequalities (5.6) and (5.7) we may assume that $\left| B_1^{\mathrm{R},i(4)} \right| < \frac{|B_1|}{6} + 2$ for all i. Hence we get that

$$\left| B_1^{\mathrm{L},0(3)} \right| + \left| B_1^{\mathrm{L},1(3)} \right| + \left| B_1^{\mathrm{L},2(3)} \right| + \left| B_1^{\mathrm{R},2(4)} \right| = |B_1| - \left| B_1^{\mathrm{R},0(4)} \right| - \left| B_1^{\mathrm{R},1(4)} \right| - \left| B_1^{\mathrm{R},3(4)} \right|$$

$$> \frac{|B_1|}{2} - 6. \tag{5.14}$$

Furthermore, the dilated sets $4\cdot B_1^{\mathrm{L},0(3)}, 4\cdot B_1^{\mathrm{L},1(3)}, 4\cdot B_1^{\mathrm{L},2(3)}$ and $3\cdot B_1^{\mathrm{R},2(4)}$ are all sets of even numbers contained in $\left(\frac{4n}{3}, 2n\right]$ and they are pairwise disjoint since they all lie in distinct residue classes modulo 12 (to be precise they lie in the classes $0, 4, 8$ and $6 \mod 12$ respectively). These dilated sets also consist only of multiples of numbers in B_1 so they are all disjoint from $O + O \subset A_{(\frac{2}{3},1]} + A_{(\frac{2}{3},1]}$ as A has property P. We conclude that the sets $4\cdot B_1^{\mathrm{L},0(3)}, 4\cdot B_1^{\mathrm{L},1(3)}, 4\cdot B_1^{\mathrm{L},2(3)}, 3\cdot B_1^{\mathrm{R},2(4)}$ and $O + O$ are pairwise disjoint sets of even numbers in $\left(\frac{4n}{3}, 2n\right]$. As there are less than $\frac{n}{3} + 1$ even numbers in $\left(\frac{4n}{3}, 2n\right]$, we get

$$\begin{aligned}\frac{n}{3} + 1 &> |O + O| + \left|4\cdot B_1^{\mathrm{L},0(3)}\right| + \left|4\cdot B_1^{\mathrm{L},1(3)}\right| + \left|4\cdot B_1^{\mathrm{L},2(3)}\right| + \left|3\cdot B_1^{\mathrm{R},2(4)}\right| \\ &> \frac{3\left|A_{(\frac{2}{3},1]}\right|}{2} - 3 + \frac{|B_1|}{2} - 6,\end{aligned}$$

using the lower bounds (5.13) and (5.14). From rearranging this inequality we obtain the bound $|B_1| < \frac{2n}{3} - 3\left|A_{(\frac{2}{3},1]}\right| + 20$, so after using that $\left|A_{(\frac{2}{3},1]}\right| + |B_1| = |A|$ by (5.4), we get in total

$$|A| = \left|A_{(\frac{2}{3},1]}\right| + |B_1| < \frac{2n}{3} - 2\left|A_{(\frac{2}{3},1]}\right| + 20 < \frac{n}{3},$$

where in the final inequality we used the assumption that $\left|A_{(\frac{2}{3},1]}\right| \geqslant \frac{n}{6} + 24$ in Case 2. This finishes the proof of Case 2.

5.7 The low density case

The proof in the final case of the argument where $\left|A_{(\frac{2}{3},1]}\right| < \frac{n}{6} + 24$ is given in [5].

Chapter 6

Graham's rearrangement conjecture beyond the rectification barrier

6.1 Introduction

Main result

This chapter is based on [8]. Let A be a finite subset of an abelian group. We say that an ordering $a_1, \dots, a_{|A|}$ of A is *valid* if the partial sums $a_1, a_1+a_2, \dots, a_1+a_2+\cdots+a_{|A|}$ are all distinct. In 1971, Graham conjectured that every set of non-zero elements of $\mathbb{F}_p$ has a valid ordering.

Conjecture 6.1.1 ([35]). *Let p be a prime. Then every subset $A \subseteq \mathbb{F}_p \setminus \{0\}$ has a valid ordering.*

This conjecture also appeared in a 1980 book of Erdős and Graham [31], and a very similar conjecture for finite cyclic groups is due to Alspach (see [10]).

The main avenue of attack on Graham's conjecture has been to show that its conclusion holds when A is small. Until recently, the published world record had established Graham's conjecture for sets A of size at most 12 (see, e.g., the discussion in [14, 44]). Earlier this year, Noah Kravitz [44] used a simple rectification argument to show that Graham's conjecture holds for all sets A of size $|A| \leqslant \log p/\log\log p$; Will Sawin [58] had independently proven a comparable bound, using roughly similar ideas, in a 2015 MathOverflow post. The purpose of the present chapter, which is joint work with N. Kravitz, is to prove Graham's conjecture for sets A of up to quasi-polynomial size.

Theorem 6.1.2. *The following holds for every constant $c > 0$. Let p be a large prime. Then every subset $A \subseteq \mathbb{F}_p \setminus \{0\}$ of size*

$$|A| \leqslant e^{c(\log p)^{1/4}}$$

has a valid ordering.

We have not made a serious effort to optimize the exponent 1/4, but the quasi-polynomial shape of this bound does appear as a natural barrier in several parts of our argument. We also mention that the conclusion of Theorem 6.1.2 still holds, with a nearly identical proof, if $\mathbb{F}_p$ is replaced by any abelian group with no non-zero elements of order strictly smaller than p.

The proof strategy for Theorem 6.1.2 is motivated by the argument in [44]. (The argument in [58] seems less well-suited to generalization.) Two new ingredients are the theory of dissociated sets (from additive combinatorics) and probabilistic tools. One of our intermediate results (see Theorem 6.3.3 below) is a structure theorem involving dissociated sets, which may be of independent interest.

Proof sketch and organization

We say that an ordering of A is *two-sided valid* if no proper nonempty subinterval sums to zero; this condition is slightly stronger than A being valid. As in [44], we will prove Theorem 6.1.2 with two-sided valid orderings.

Let us briefly recall the main ideas of [44]. Let $A \subseteq \mathbb{F}_p \setminus \{0\}$ be a subset of size $|A| \leqslant \log p / 2 \log\log p$. Using the pigeonhole principle, one can find some $\lambda \in \mathbb{F}_p^\times$ such that the dilate $\lambda \cdot A$ is contained in the interval $(-p/|A|, p/|A|)$. Since sums of elements of $\lambda \cdot A$ have no "wrap-around", we can interpret $\lambda \cdot A$ as a subset of $\mathbb{Z} \setminus \{0\}$; this process is known as "rectification". Finally, in the integer setting, one can use induction on $|A|$ to find a two-sided valid ordering in which all of the positive elements appear before all of the negative elements.

Our proof of Theorem 6.1.2 proceeds in four main steps. The first step is showing that every subset of $\mathbb{F}_p$ can be decomposed into a union of large dissociated sets and a rectifiable residual set. (A dissociated set is a set all of whose subset sums are distinct; see below.) The residual set can be broken into "positive" and "negative" sets. We will aim to find a two-sided valid ordering consisting of the positive elements, then the elements of the dissociated sets, then the negative elements.

The second step is ordering the positive and negative elements. Following [44], we inductively construct these orderings in order to avoid zero-sum intervals that

begin in the positive region and end in the negative region. We take advantage of some flexibility in the argument from [44] in order to prepare for "potential" zero-sum intervals with one endpoint in the positive region or negative region and the other endpoint very close to one of the edges of the dissociated region.

The third and fourth steps concern ordering the elements of the dissociated sets. The main idea is that in a uniformly random ordering of a dissociated set of size R, the sum of the first k elements is uniformly distributed on $\binom{R}{k}$ different values. Since the probability of assuming any particular value is very small, the probability of this initial segment forming the end of a zero-sum interval is also very small. This naïve random strategy essentially works for handling sets A of size up to $(\log p)^{3/2}$ (which breaks the "rectification barrier" of [44]), but we must employ a more elaborate random procedure in order to reach the threshold in Theorem 6.1.2. In particular, it becomes important to distinguish between the "borders" and "interiors" of the orderings of the dissociated sets. The third step of the proof is randomly splitting and then reordering the dissociated sets, and the fourth step is choosing a (suitably) random ordering for the elements within each dissociated set.

We carry out these four steps in Sections 6.3, 6.4, 6.5, and 6.6, respectively, and then we make some concluding remarks and pose several open problems in Section 6.7.

6.2 Notation and parameters

Before jumping into the proofs, we set a few pieces of notation.

- We denote the restricted sumset by $B\hat{+}B := \{b + b' : b, b' \in B \text{ and } b \neq b'\}$.
- Let $\sum_{=M}(S) := \{\sum_{s\in S'} s : S' \subseteq S, |S'| = M\}$ denote the set of all sums of exactly M elements of S. Likewise, let $\sum_{\leqslant M}(S) := \{\sum_{s\in S'} s : S' \subseteq S, |S'| \leqslant M\}$ denote the set of all sums of at most M elements of S, and let $\sum_{\geqslant M}(S) := \{\sum_{s\in S'} s : S' \subseteq S, |S'| \geqslant M\}$ denote the set of all sums of at least M elements of S.
- For a sequence $\mathbf{b} = b_1, \dots, b_r$, let $\mathrm{IS}(\mathbf{b}) := \{b_1 + \cdots + b_j : 0 \leqslant j \leqslant r\}$ denote the set of initial segment sums of $\mathbf{b}$, and let $\overline{\mathbf{b}} := b_r, \dots, b_1$ denote the reverse of $\mathbf{b}$.

When there is no risk of confusion, we sometimes omit floor functions in calculations for typographical clarity.

Let us also record a few parameters that we will carry through our proofs.

- Our set A will have size $|A| \leqslant e^{c(\log p)^{1/4}}$ for some absolute constant $c > 0$.

- We define the *rectification threshold* for a subset $A \subseteq \mathbb{F}_p$ to be

$$R = R(A) := c_1 \max\left((\log p)^{1/2}, \frac{\log p}{\log |A|}\right),$$

  where $c_1 > 0$ is a sufficiently small absolute constant.

- The *border width* is $K := c_2 R^{1/3}$, for yet another absolute constant $c_2 > 0$.

- We will use s (and later u) to denote the number of dissociated sets in our decomposition of the set A. The precise values of s, u, which are of no importance (besides the trivial bound $s, u \leqslant |A|$), will vary over the course of the proofs.

- We will always use δ_j to denote the sum of the elements of the set D_j, and we will always use τ_j to denote the sum of the elements of the set T_j; when applicable, we will also write $\delta := \sum_j \delta_j$.

Finally, we reiterate that a sequence $b_1, \dots, b_t$ is *two-sided valid* if

$$b_i + \cdots + b_j \neq 0 \quad \text{for all } 1 \leqslant i < j \leqslant t \text{ with } (i,j) \neq (1,t).$$

6.3 Structure theorem

The reader may wish to recall the Definition 2.0.5 of dissociativity and additive dimension. The following lemma says that sets of sufficiently small dimension can always be "rectified", where rectification in this context means that for a given $B \subset \mathbb{F}_p$ we hope to find a set $B' \subset \mathbb{Z}$ which is $F_{|B|}$-isomorphic to B as in Definition 3.6.1 (i.e. sums of length up to $|B|$ behave the same in B as in B'). To make this precise, we define for each (nonempty) subset $A \subseteq \mathbb{F}_p$ the parameter

$$R = R(A) := c_1 \max\left((\log p)^{1/2}, \frac{\log p}{\log |A|}\right), \tag{6.1}$$

where c_1 is a sufficiently small absolute constant. The following lemma is a slight variation of Lemma 2.0.9.

Lemma 6.3.1. *If $B \subseteq \mathbb{F}_p$ is a nonempty subset of dimension $\dim(B) < R = R(B)$, then there is some $\lambda \in \mathbb{F}_p^\times$ such that the dilate $\lambda \cdot B$ is contained in the interval $(-\frac{p}{100|B|}, \frac{p}{100|B|})$.*

Proof. Let D be a maximal dissociated subset of B, so that $|D| = \dim(B)$. Lemma 2.0.6 tells us that $B \subseteq \mathrm{span}(D)$. Consider the set

$$\{(\lambda d/p)_{d\in D} : \lambda \in \mathbb{F}_p\} \subseteq (\mathbb{R}/\mathbb{Z})^{\dim(B)}.$$

The pigeonhole principle provides some distinct $\lambda_1, \lambda_2 \in \mathbb{F}_p$ such that $\|\lambda_1 d/p - \lambda_2 d/p\|_{\mathbb{R}/\mathbb{Z}} \leqslant p^{-1/\dim(B)}$ for all $d \in D$. Set $\lambda := \lambda_1 - \lambda_2 \in \mathbb{F}_p^\times$, so that $\lambda d \in [-p^{1-1/\dim(B)}, p^{1-1/\dim(B)}]$ for all $d \in D$. Since $B \subseteq \mathrm{span}(D)$, we have

$$\lambda \cdot B \subseteq [-\dim(B)p^{1-1/\dim(B)}, \dim(B)p^{1-1/\dim(B)}].$$

It remains only to show that $\dim(B)p^{1-1/\dim(B)} < p/(100|B|)$, i.e., that $100|B|\dim(B) < p^{1/\dim(B)}$, as long as c_1 is chosen to be sufficiently small. When $\log|B| < (\log p)^{1/2}$, this inequality follows from $\dim(B) \leqslant R = c_1 \log p/\log|B|$. When $\log|B| \geqslant (\log p)^{1/2}$, the desired inequality follows from $\dim(B) \leqslant R = c_1(\log p)^{1/2}$ and $|B| \leqslant 3^{\dim(B)}$. □

Remark 6.3.2. For applications in this chapter, we will always work with sets of size at most $e^{c(\log p)^{1/4}}$, in which case the previous lemma says that every set B of size at most $e^{c(\log p)^{1/4}}$ with $\dim(B) < c_1(\log p)^{3/4}$ is rectifiable. We opted to prove Lemma 6.3.1 for arbitrary sets $B \subseteq \mathbb{F}_p$, however, so that we could state the structural results in the rest of this section in full generality. These results are nontrivial for sets B of all sizes since the rectification threshold always satisfies $R(B) \gg (\log p)^{1/2}$.

We can combine these two lemmas to obtain a decomposition of *any* subset of $\mathbb{F}_p$ into large dissociated sets and a residual set that (after suitable dilation) is contained in a small interval around 0. We shall from now on simply write R for $R(A)$. The following theorem bears many similarities to an argument of Bourgain [11] from a different context.

Theorem 6.3.3. *Every subset $A \subseteq \mathbb{F}_p$ can be partitioned as*

$$A = D_1 \cup \cdots \cup D_s \cup E, \tag{6.2}$$

where the following holds:

(i) each D_j is a dissociated set of size $|D_j| \asymp R$;

(ii) $|E| \geqslant R/2$ if $s > 0$;

(iii) there is some $\lambda \in \mathbb{F}_p^\times$ such $\lambda \cdot (E \cup \{\delta\}) \subseteq (-\frac{p}{90(|E|+1)}, \frac{p}{90(|E|+1)})$, where $\delta := \sum_{j=1}^s \sum_{d\in D_j} d$ is the sum of all of the elements in the dissociated sets.

Proof. Start with $E = A$. As long as $\dim(E) \geqslant R$, iteratively remove a dissociated subset of size $R/2$, so that at each step the set E of remaining elements has size $|E| \geqslant R/2$. Once we reach a residual set E of dimension smaller than R, Lemma 6.3.1 (applied to $E \cup \{\delta\}$, where δ is the sum of all of the dissociated elements removed) provides the desired $\lambda \in \mathbb{F}_p^\times$. □

We will, of course, apply this theorem to the set A for which we are trying to find a two-sided valid ordering. If the number s of dissociated sets happens to be 0, then the entire set A is rectifiable and therefore has a two-sided valid ordering by [44] (see the discussion in the proof sketch). Thus, we will restrict our attention to the case where $s \geqslant 1$ (so in particular $|E| \gg R$ from (ii)). The presence of a large dissociated set allows us to obtain a more detailed structural result.

Proposition 6.3.4. *For every nonempty subset $A \subseteq \mathbb{F}_p \setminus \{0\}$, there is some $\lambda \in \mathbb{F}_p^\times$ such that $\lambda \cdot A$ can be partitioned as*

$$\lambda \cdot A = P \cup N \cup (\cup_{j=1}^{s} D_j),$$

where

(i) the "positive" set P is contained in $(0, \frac{p}{4|P \cup N|})$, the "negative set" N is contained in $(-\frac{p}{4|P \cup N|}, 0)$, and the element $\delta := \sum_{j=1}^{s} \sum_{d \in D_j} d$ is contained in $(-\frac{p}{4}, \frac{p}{4})$;

and the following also holds if $s > 0$:

(ii) $P \cup N$ is nonempty, and each D_j is a dissociated set of size $|D_j| \asymp R$, where the implied constants are absolute;

(iii) $\delta \notin \{0\} \cup -P \cup -N$, and moreover $\delta \neq -\sum_{p \in P} p$ if N is nonempty and $\delta \neq -\sum_{n \in N} n$ if P is nonempty;

(iv) $D_1 \cup D_s \cup \{\delta\}$ is a dissociated set;

(v) $|D_1| = |D_s|$.

Before proving this proposition, we make a simple but powerful observation about absorbing elements into dissociated sets.

Lemma 6.3.5. *Let G be an abelian group, and let $D_1 \cup D_2$ be a partition of a dissociated subset of G. For every element $x \in G \setminus \{0\}$, either $D_1 \cup \{x\}$ or $D_2 \cup \{x\}$ is dissociated.*

Proof. Assume for the sake of contradiction that neither $D_1 \cup \{x\}$ nor $D_2 \cup \{x\}$ is dissociated. Since D_1 is dissociated, the failure of $D_1 \cup \{x\}$ to be dissociated implies that $x \in \mathrm{span}(D_1)$; similarly, $x \in \mathrm{span}(D_2)$. So $\mathrm{span}(D_1) \cap \mathrm{span}(D_2)$ contains a non-zero element, contradicting the assumption that $D_1 \cup D_2$ is dissociated. □

Iterating this observation, we find that if B is a set of size t and $D_1 \cup \cdots \cup D_{t+1}$ is a partition of a dissociated set, then it is always possible to add the elements of B to the dissociated sets D_j in such a way that the sets remain dissociated.

We will also make use of the trivial lower bound for the size of a restricted sumset in $\mathbb{Z}$: If $B \subseteq \mathbb{Z}$ is a finite set, then $|B\hat{+}B| \geqslant 2|B| - 3$. We are now ready to prove Proposition 6.3.4. The choice of numerical constants appearing in the proof is not important.

Proof of Proposition 6.3.4. To start, Theorem 6.3.3 provides some $\lambda \in \mathbb{F}_p^\times$ and a decomposition

$$\lambda \cdot A = D_1 \cup \cdots \cup D_s \cup E,$$

where each D_j is a dissociated set of size $\asymp R$ and we have $E \cup \{\delta\} \subseteq (-\frac{p}{90(|E|+1)}, \frac{p}{90(|E|+1)})$, for $\delta := \sum_{j=1}^{s} \sum_{d \in D_j} d$. Set $P := E \cap (0, p/4|E|)$ and $N := E \cap (-p/4|E|, 0)$. If $s = 0$, then we have already obtained the desired decomposition of $\lambda \cdot A$, so for the remainder of the proof we assume that $s \geqslant 1$. By replacing λ with $-\lambda$ if necessary, we may assume that $|P| \geqslant |N|$. In particular, since $|E| \gg R$, this implies that $|P| \gg R$.

We remark that once we have a decomposition satisfying conditions (i)–(iii), we can modify the decomposition to satisfy (iv) and (v) as follows. Split D_1 into 2 parts $D_1^{(1)}, D_1^{(2)}$ each of size $\asymp R$. Lemma 6.3.5 ensures that either $D_1^{(1)} \cup \{\delta\}$ or $D_1^{(2)} \cup \{\delta\}$ is dissociated; without loss of generality, assume that $D_1^{(1)} \cup \{\delta\}$ is dissociated. Then further split $D_1^{(1)}$ into 2 parts $D_1^{(3)}, D_1^{(4)}$ each of size $\lfloor |D_1^{(1)}|/2 \rfloor \asymp R$, add the leftover element of $D_1^{(1)}$ to $D_1^{(2)}$ if $|D_1^{(1)}|$ was odd, and replace the sequence of sets $D_1, \ldots, D_s$ by the sequence $D_1^{(3)}, D_1^{(2)}, D_2, D_3, \ldots, D_s, D_1^{(4)}$. This new sequence satisfies (iv) and (v). The remainder of the proof is devoted to finding a decomposition satisfying conditions (i)–(iii).

We will later apply sumset inequalities involving P, N, and we will need P, N to be not-too-small so that we have "room" for sumsets to expand. In anticipation of this, we begin by reducing to the case where N is either empty or of size at least 10. Suppose that $0 < |N| < 10$. Note that $\sum_{n \in N} n \in (-p/90, p/90)$. Split D_1 into $|N| + 1 \leqslant 10$ sets each of size $\asymp R$; the remark before the proof ensures that we can absorb all of the elements of N into these dissociated sets, and this procedure changes

the value of δ by at most $p/90$. Notice that each newly formed D_j still has size $\asymp R$, and that we still have $P \subseteq (0, \frac{p}{40|P \cup N|})$, $N \subseteq (-\frac{p}{40|P \cup N|}, 0)$, and $\delta \in (-p/40, p/40)$.

We now consider two cases depending on the size of N. First, suppose that $N = \emptyset$, and recall that $|P| \gg R$. Since P is rectifiable (i.e., Freiman-isomorphic to a subset of $\mathbb{Z}$), the trivial lower bound for restricted sumsets in integers gives

$$|P \hat{+} P| \geqslant 2|P| - 3 > |P| + 2,$$

and hence we can find distinct $p_1, p_2 \in P$ such that $p_1 + p_2 + \delta \notin \{0, -\sum_{n \in N} n\} \cup -P$. Splitting D_1 and absorbing p_1, p_2 with the help of Lemma 6.3.5 as above yields a new decomposition of A where the sum of all of the elements in the dissociated sets is $p_1 + p_2 + \delta$ and hence conditions (ii) and (iii) are satisfied. This procedure also changes the value of δ by at most $p/90$ (say), so we obtain the desired decomposition of $\lambda \cdot A$.

Finally, suppose that $|N| \geqslant 10$, and recall that we also have $|P| \geqslant |N| \geqslant 10$. By absorbing 5 arbitrary elements of N into D_1, we may assume that $|P| \geqslant |N| - 5$. Since we still have $|N| > 1$, there is some $n_1 \in N$ such that $\delta + n_1 \neq -\sum_{p \in P} p$; as above, we absorb n_1 into D_1, so that the final part of condition (iii) is satisfied. Now we have

$$|P \hat{+} P| \geqslant 2|P| - 3 > |P| + |N| + 2,$$

so there are distinct $p_1, p_2 \in P$ such that

$$\delta + p_1 + p_2 \notin \{0, -\sum_{n \in N} n\} \cup -P \cup -N.$$

Absorbing p_1, p_2 into D_1 gives the desired decomposition of $\lambda \cdot A$. (For the final part of condition (iii), note that this last step preserves the property $\delta + \sum_{p \in P} p \neq 0$.) $\square$

6.4 Ordering P and N

With Proposition 6.3.4 in hand, we can say a bit more about the remainder of the proof of Theorem 6.1.2. We will aim to find orderings $\mathbf{p}$ of P, $\mathbf{n}$ of N, and $\mathbf{d}$ of $\cup_j D_j$ such that $\overline{\mathbf{p}}, \mathbf{d}, \mathbf{n}$ is a two-sided valid ordering of A. Of course, we will need each of the three orderings to be two-sided valid on its own, and we will need to avoid creating zero-sum intervals when we concatenate them.

Condition (i) from Proposition 6.3.4 states that $P \cup N \subset (-\frac{p}{4|P \cup N|}, \frac{p}{4|P \cup N|})$ and that $\delta \in (-p/4, p/4)$, which means that the problem of constructing $\mathbf{p}$ and $\mathbf{n}$ naturally lives in the integers rather than in $\mathbb{F}_p$, as follows. Identify δ and the elements of $P \cup N$

with elements of $(-p/4, p/4) \subseteq \mathbb{Z}$ in the natural way, and note that sums of these elements can be computed equivalently in $\mathbb{F}_p$ and in $(-p/2, p/2) \subseteq \mathbb{Z}$ because the sums in $\mathbb{F}_p$ do not exhibit any wrap-around. Likewise, for any ordering $\mathbf{d}$ of $\cup_j D_j$, we can identify $\mathrm{IS}(\mathbf{d})$ and $\mathrm{IS}(\overline{\mathbf{d}})$ with subsets of $(-p/2, p/2) \subseteq \mathbb{Z}$. Now we observe that the ordering $\overline{\mathbf{p}}, \mathbf{d}, \mathbf{n}$ is two-sided valid if and only if the ordering $\overline{\mathbf{p}}, \delta, \mathbf{n}$ is two-sided valid in the integers, the ordering $\mathbf{d}$ is two-sided valid in $\mathbb{F}_p$, and $\mathrm{IS}(\mathbf{p}) \cap -\mathrm{IS}(\mathbf{d}) = \mathrm{IS}(\mathbf{n}) \cap -\mathrm{IS}(\overline{\mathbf{d}}) = \emptyset$; the key point is that the first condition lives entirely in the integers, the second condition does not concern $\mathbf{p}$ and $\mathbf{n}$, and the third condition lives in the integers for each fixed choice of $\mathbf{d}$.

We will later choose $\mathbf{d}$ randomly, but it turns out that we can model $-\mathrm{IS}(\mathbf{d}), -\mathrm{IS}(\overline{\mathbf{d}})$ by somewhat larger deterministic sets that encode all of the "potentially important" intersections with $\mathrm{IS}(\mathbf{p}), \mathrm{IS}(\mathbf{n})$ (respectively); it will suffice to ensure that $\mathrm{IS}(\mathbf{p}), \mathrm{IS}(\mathbf{n})$ have fairly small intersections with these deterministic sets. With this in mind, the main result of this section is as follows.

Proposition 6.4.1. *Let $P \subseteq (0, \infty)$ and $N \subseteq (-\infty, 0)$ be finite sets of integers, and let $\delta > 0$ be a positive integer not contained in $-N$; moreover, assume that $\delta \neq -\sum_{n \in N} n$ if $P \neq \emptyset$. Let $Y_1^+, \dots, Y_m^+, Y_1^-, \dots, Y_m^- \subseteq \mathbb{Z}$ be finite sets. Then there are orderings $\mathbf{p}$ of P and $\mathbf{n}$ of N such that $\overline{\mathbf{p}}, \delta, \mathbf{n}$ is two-sided valid and we have*

$$|\mathrm{IS}(\mathbf{p}) \cap Y_j^+| \leqslant \inf_{L \in \mathbb{N}} \left(\frac{|Y_j^+|}{L} + L + 4 + 4 \sum_{i=1}^{j-1} |Y_i^+| \right) \tag{6.3}$$

and

$$|\mathrm{IS}(\mathbf{n}) \cap Y_j^-| \leqslant \inf_{L \in \mathbb{N}} \left(\frac{|Y_j^-|}{L} + L + 4 + 4 \sum_{i=1}^{j-1} |Y_i^-| \right) \tag{6.4}$$

for all $1 \leqslant j \leqslant m$.

In [44], the first author presented a simple algorithm for inductively constructing a two-sided valid ordering of any finite subset of $\mathbb{Z} \setminus \{0\}$. Let us quickly review this algorithm since it forms the basis for the proof of Proposition 6.4.1. Suppose $P \subseteq (0, \infty)$ and $N \subseteq (-\infty, 0)$ are finite sets of integers; we want to produce orderings $\mathbf{p}$ of P and $\mathbf{n}$ of N such that $\overline{\mathbf{p}}, \mathbf{n}$ is two-sided valid. Note that we may assume that P, N are non-empty, as else this is trivial. We will construct the sequences $\mathbf{p} = p_1, \dots, p_{|P|}$ and $\mathbf{n} = n_1, \dots, n_{|N|}$ from the larger indices to the smaller indices. For the first step, consider the sign of $\sum_{n \in N} n + \sum_{p \in P} p$. Suppose that this sum is non-negative; we will choose the value of $p_{|P|}$ as follows. There is some $p^* \in P$

such that $\sum_{n\in N} n + \sum_{p\in P\setminus\{p^*\}} p \neq 0$, and we choose this p^* to be our $p_{|P|}$. This choice ensures that if $\overline{\mathbf{p}'}, \mathbf{n}$ is a two-sided valid ordering of the remaining elements $P \setminus \{p_{|P|}\}, N$, then $p_{|P|}, \overline{\mathbf{p}'}, \mathbf{n}$ is the desired two-sided valid ordering of P, N: Any interval containing $p_{|P|}$ and not containing any of $\mathbf{n}$ clearly has strictly positive sum; any proper interval containing both $p_{|P|}$ and some elements of $\mathbf{n}$ must contain all of P and hence has strictly positive sum by our assumption that $\sum_{n\in N} n + \sum_{p\in P} p \geqslant 0$; the intervals strictly contained in $\overline{\mathbf{p}'}, \mathbf{n}$ are all non-zero-sum by assumption; and the interval consisting of all of $\overline{\mathbf{p}'}, \mathbf{n}$ has non-zero sum by our choice of p^*. If instead $\sum_{n\in N} n + \sum_{p\in P} p < 0$, then we choose $n_{|N|}$ analogously. With this first step complete, we throw out the already-chosen element $p_{|P|}$ or $n_{|N|}$ and repeat this process with the remaining elements. This procedure produces the desired orderings $\mathbf{p}, \mathbf{n}$.

The above algorithm has a lot of slack, in the sense that at each step there are many possible choices. To exploit this slack and prove Proposition 6.4.1, we will employ a modified algorithm that greedily avoids partial sums of $\mathbf{p}$ lying in the Y_j^+'s and partial sums of $\mathbf{n}$ lying in the Y_j^-'s.

Proof of Proposition 6.4.1. We will construct the sequences $\mathbf{p} = p_1, \ldots, p_{|P|}$ and $\mathbf{n} = n_1, \ldots, n_{|N|}$ from the larger indices to the smaller indices. Suppose that we have already chosen the values of $p_{|P|}, p_{|P|-1}, \ldots, p_{k+1}$ and $n_{|N|}, n_{|N|-1}, \ldots, n_{\ell+1}$. At the next step, we will choose the value of either p_k or n_ℓ depending on the sign of the sum of all of the remaining elements. Let

$$P_k := P \setminus \{p_{|P|}, \ldots, p_{k+1}\} \quad \text{and} \quad N_\ell := N \setminus \{n_{|N|}, \ldots, n_{\ell+1}\}$$

be the sets of remaining elements of P and N, and define the quantities

$$\pi_k := \sum_{p\in P_k} p \quad \text{and} \quad \nu_\ell := \sum_{n\in N_\ell} n.$$

As in the algorithm from [44], we will succeed in constructing a two-sided valid ordering as long as p_k, n_ℓ avoid a few particular potential values (for more details, see Claim 6.4.3 below). If we are choosing p_k, then we want p_k not to be equal to $\pi_k + \delta + \nu_\ell$, since this choice of p_k would lead to a zero-sum interval $p_{k-1} + \cdots + p_1 + \delta + n_1 + \cdots + n_\ell = 0$. Likewise, if we are choosing n_ℓ, then we want n_ℓ not to be equal to either $\delta + \nu_\ell$ or $\pi_k + \delta + \nu_\ell$, since these choices of n_ℓ would lead to zero-sum intervals $\delta + n_1 + \cdots + n_{\ell-1} = 0$ and $p_k + \cdots + p_1 + \delta + n_1 + \cdots + n_{\ell-1} = 0$. With this in mind, we define the sets

$$P_k' := P_k \setminus \{\pi_k + \delta + \nu_\ell\} \quad \text{and} \quad N_\ell' := N_\ell \setminus \{\delta + \nu_\ell, \pi_k + \delta + \nu_\ell\}$$

of "allowable" choices for p_k and n_ℓ. Due to the assumptions in Proposition 6.4.1, it will always transpire that the set P'_k or N'_ℓ under consideration is nonempty.

Again as in the algorithm from [44], the sign of the quantity $\pi_k + \delta + \nu_\ell$ will determine whether we choose the value of p_k or the value of n_ℓ next:[1]

1. Suppose that $\pi_k + \delta + \nu_\ell \geqslant 0$ and $k > 0$. Then we will choose $p_k \in P'_k$ as follows. If there is some $p^* \in P'_k$ such that $\pi_k - p^* \notin \cup_j Y_j^+$, then choose p_k to be this p^* and say that the current step is a *skip-step for* P.

   Now, consider the case where $\pi_k - P'_k \subseteq \cup_j Y_j^+$. Let i be minimal such that $\pi_k - P'_k$ intersects Y_i^+, and say that the current step is an *i-step for* P. If $\pi_k - P'_k \subseteq Y_i^+$, then let p_k be the largest element of P'_k. If $\pi_k - P'_k \not\subseteq Y_i^+$, then let p_k be the largest $p^* \in P'_k$ such that $\pi_k - p^* \notin Y_i^+$.

2. Suppose that $\pi_k + \delta + \nu_\ell \geqslant 0$, $k = 0$ (i.e., we have already chosen all of $\mathbf{p}$), and $\ell > 0$, or that $\pi_k + \delta + \nu_\ell < 0$ and $\ell > 0$. Then we will choose $n_\ell \in N'_\ell$ as follows. If there is some $n^* \in N'_\ell$ such that $\nu_\ell - n^* \notin \cup_j Y_j^-$, then choose n_ℓ to be this n^* and say that the current step is a *skip-step for* N.

   Now, consider the case where $\nu_\ell - N'_\ell \subseteq \cup_j Y_j^-$. Let i be minimal such that $\nu_\ell - N'_\ell$ intersects Y_i^-, and say that the current step is an *i-step for* N. If $\nu_\ell - N'_\ell \subseteq Y_i^-$, then let n_ℓ be the smallest (i.e., most negative) element of N'_ℓ. If $\nu_\ell - N'_\ell \not\subseteq Y_i^-$, then let n_ℓ be the smallest (i.e., most negative) $n^* \in N'_\ell$ such that $\nu_\ell - n^* \notin Y_i^-$.

We begin with $(k, \ell) = (|P|, |N|)$ and run the above procedure until we reach $(k, \ell) = (0, 0)$. To establish Proposition 6.4.1, we must show three things: that the algorithm actually runs and produces orderings $\mathbf{p}$ of P and $\mathbf{n}$ of N; that the resulting ordering $\mathbf{p}, \delta, \overline{\mathbf{n}}$ is two-sided valid; and that IS($\mathbf{p}$) and IS($\mathbf{n}$) have small intersections with the Y_j^+'s and Y_j^-'s (respectively).

Claim 6.4.2. *The above algorithm runs all the way to $(k, \ell) = (0, 0)$ and produces orderings $\mathbf{p}$ of P and $\mathbf{n}$ of N.*

Proof. Note that as long as $(k, \ell) \neq (0, 0)$, we fall into one of the two cases. Indeed, when $\pi_k + \delta + \nu_\ell \geqslant 0$, we fall into case (1) or case (2) according to whether $k > 1$ or $k = 0$. When $\pi_k + \delta + \nu_\ell < 0$, we must have $\ell > 0$ since $\delta > 0$ and $\pi_k \geqslant 0$, so we fall into case (2).

[1] The apparent asymmetry in the cases arises from the assumption in Proposition 6.4.1 that $\delta > 0$.

It remains to show that the sets P'_k, N'_ℓ are always nonempty when needed. First, consider case (1). It is clear that $P'_k \neq \emptyset$ as long as $k > 1$. When $k = 1$, the set P_1 consists of a single element p_1, and we have $\pi_1 = p_1$. We must show that $\pi_1 + \delta + \nu_\ell \neq p_1$, i.e., that $\delta \neq -\nu_\ell$. When $\ell = |N|$, this is precisely the assumption in Proposition 6.4.1 that $\delta \neq -\sum_{n \in N} n$. For $\ell < |N|$, recall that $n_{\ell+1}$ was chosen to be an element of $N'_{\ell+1}$, which by construction does not contain $\delta + \nu_{\ell+1}$. It follows that $\nu_\ell = \nu_{\ell+1} - n_{\ell+1} \neq \nu_{\ell+1} - (\delta + \nu_{\ell+1}) = -\delta$, as desired.

Now, consider case (2). We begin with the subcase where $\pi_k + \delta + \nu_\ell \geqslant 0$ and $k = 0$. Note that $\pi_0 = 0$ and hence $\pi_0 + \delta + \nu_\ell = \delta + \nu_\ell$. It is clear that $N'_\ell \neq \emptyset$ as long as $\ell > 1$. When $\ell = 1$, the set N_1 consists of a single element n_1, and we have $\nu_1 = n_1$. The assumption $\delta > 0$ ensures that $\delta + \nu_1 = \delta + n_1 \neq n_1$, so $N'_1 \neq \emptyset$.

Finally, we treat the subcase where $\pi_k + \delta + \nu_\ell < 0$. It is clear that $N'_\ell \neq \emptyset$ as long as $\ell > 2$. When $\ell = 2$, the set N_2 consists of two elements n_1, n_2, and we have $\nu_2 = n_1 + n_2$. Then $\delta + \nu_2 = \delta + n_1 + n_2 \notin N_2$ by the assumption in Proposition 6.4.1 that $\delta \notin -N$ so neither of n_1, n_2 is equal to $-\delta$, and it follows that $N'_2 \neq \emptyset$. When $\ell = 1$, the set N_1 consists of a single element n_1, and we have $\nu_1 = n_1$. Then $\delta + \nu_1 = \delta + n_1 \neq n_1$ since $\delta > 0$, and $\pi_k + \delta + \nu_1 = \pi_k + \delta + n_1 \neq n_1$ since $\pi_k + \delta > 0$. Thus $N'_1 \neq \emptyset$, and this concludes the proof. □

Claim 6.4.3. *The ordering $\overline{\mathbf{p}}, \delta, \mathbf{n}$ is two-sided valid.*

Proof. Since any zero-sum interval must contain both positive and negative numbers, we can restrict our attention to intervals of the form $p_k + \cdots + p_1 + \delta + n_1 + \cdots + n_\ell$ (with sum $\pi_k + \delta + \nu_\ell$) and $\delta + n_1 + \cdots + n_\ell$ (with sum $\delta + \nu_\ell$).

Let us first consider the sums $\pi_k + \delta + \nu_\ell$. Note that we do not need to worry about $(k, \ell) = (|P|, |N|)$, since the corresponding interval is the entire sequence $\overline{\mathbf{p}}, \delta, \mathbf{n}$, so we may assume that either $k < |P|$ or $\ell < |N|$. Let (k^*, ℓ^*) be the earliest step in the algorithm where $k^* \leqslant k$ and $\ell^* \leqslant \ell$. Then the previous step in the algorithm was either $(k^* + 1, \ell^*)$ or $(k^*, \ell^* + 1)$; without loss of generality assume that it was the former, since the argument for the latter is very similar. Then $k^* = k$ and $\ell^* \leqslant \ell$. If $\ell^* = \ell$, then

$$\pi_k + \delta + \nu_\ell = (\pi_{k+1} - p_{k+1}) + \delta + \nu_\ell$$

is nonzero because we chose $p_{k+1} \in P'_{k+1}$ and the set P'_{k+1} does not contain $\pi_{k+1} + \delta + \nu_\ell$. If instead $\ell^* < \ell$, then there is some $k' > k$ such that n_ℓ was chosen at step (k', ℓ). It follows that

$$\pi_k + \delta + \nu_\ell < \pi_{k'} + \delta + \nu_\ell < 0,$$

so $\pi_k + \delta + \nu_\ell$ is nonzero, as desired.

Let us now consider the sums $\delta + \nu_\ell$. We can again quickly dispose of the case $\ell = |N|$. Indeed, if $P = \emptyset$, then the corresponding interval is the entire sequence $\overline{\mathbf{p}}, \delta, \mathbf{n}$, and if $P \neq \emptyset$, then $\delta \neq -\sum_{n \in N} n = -\nu_{|N|}$ by assumption. So we may assume that $\ell < |N|$, and we conclude by noting that $\delta + \nu_\ell = \delta + (\nu_{\ell+1} - n_{\ell+1}) \neq 0$ since $N'_{\ell+1}$ does not contain $\delta + \nu_{\ell+1}$. □

Claim 6.4.4. *The orderings* $\mathbf{p}$ *and* $\mathbf{n}$ *satisfy*

$$|\operatorname{IS}(\mathbf{p}) \cap Y_j^+| \leqslant \inf_{L \in \mathbb{N}} \left(\frac{|Y_j^+|}{L} + L + 2 + 4 \sum_{i=1}^{j-1} |Y_i^+| \right)$$

and

$$|\operatorname{IS}(\mathbf{n}) \cap Y_j^-| \leqslant \inf_{L \in \mathbb{N}} \left(\frac{|Y_j^-|}{L} + L + 2 + 4 \sum_{i=1}^{j-1} |Y_i^-| \right)$$

for all $1 \leqslant j \leqslant m$.

Proof. We will prove the statement only for $|\operatorname{IS}(\mathbf{n}) \cap Y_j^-|$ since the argument for $|\operatorname{IS}(\mathbf{p}) \cap Y_j^+|$ is essentially identical. Recall that $\operatorname{IS}(\mathbf{n}) = \{\nu_\ell = \sum_{i=1}^{\ell} n_i : 0 \leqslant \ell \leqslant |N|\}$. For $0 \leqslant \ell < |N|$, write $\nu_\ell = \nu_{\ell+1} - n_{\ell+1}$. This quantity can lie in Y_j^- only when the choice of $n_{\ell+1}$ is a j-step or an i-step for some $i < j$. We will bound these two contributions separately. Note that skip-steps and i-steps for $i > j$ never contribute.

We first consider the contribution of j-steps. Notice that the partial sums ν_ℓ are strictly increasing (becoming less negative) as ℓ decreases. Suppose that the choice of $n_{\ell+1}$ is a j-step and $\nu_\ell = \nu_{\ell+1} - n_{\ell+1} \in Y_j^-$. Then we must have $\nu_{\ell+1} - N'_{\ell+1} \subseteq Y_j^-$. Since $n_{\ell+1}$ is the smallest (most negative) element of $N'_{\ell+1}$, the other $|N'_{\ell+1} \setminus \{n_{\ell+1}\}| \geqslant \ell + 1 - 3 = \ell - 2$ elements of $\nu_{\ell+1} - N'_{\ell+1} \subseteq Y_j^-$ lie in the interval $(\nu_{\ell+1}, \nu_\ell)$; it follows that these elements are "skipped" and can never appear in $\operatorname{IS}(\mathbf{n})$. In particular, from such j-steps with $\ell \geqslant L + 1$ we obtain at most $|Y_j^-|/L$ elements of $\operatorname{IS}(\mathbf{n}) \cap Y_j^-$. From j-steps with $\ell \leqslant L$ we trivially obtain at most $L + 1$ elements of $\operatorname{IS}(\mathbf{n}) \cap Y_j^-$.

We now consider the contribution of i-steps with $i < j$. We will trivially bound this contribution by the total number of i-steps with $i < j$. We claim that the number of i-steps is at most $4|Y_i^-|$ for each i. For each i-step ℓ, let $y(\ell)$ denote the largest (least negative) element of $(\nu_\ell - N'_\ell) \cap Y_i^-$. It suffices to show that each $y \in Y_i^-$ appears as $y(\ell)$ for at most 4 different i-steps ℓ. If $y(\ell)$ is not the largest element of $\nu_\ell - N'_\ell$, then it is distinct from $y(\ell')$ for all $\ell' < \ell$ since

$$\nu_{\ell'} \geqslant \nu_{\ell-1} = \nu_\ell - n_\ell > y(\ell)$$

by the definition of an i-step. If $y(\ell)$ is the largest element of $\nu_\ell - N'_\ell$, then it is one of the three largest elements of $\nu_\ell - N_\ell$. Notice that the largest element of $\nu_\ell - N_\ell$ is strictly increasing as ℓ decreases, the second-largest element of $\nu_\ell - N_\ell$ is strictly increasing as ℓ decreases, and the third-largest element of $\nu_\ell - N_\ell$ is strictly increasing as ℓ decreases; it follows that each y can appear at most three times as one of the three largest elements of $\nu_\ell - N_\ell$. Thus we have shown that each i-step ℓ in the algorithm is associated with some number $y(\ell) \in Y_i^-$ and moreover that any given $y \in Y_i^-$ appears as $y(\ell)$ for at most 4 different i-steps ℓ, so we conclude that the total number of i-steps is at most $4|Y_i^-|$. This establishes the claim.

Combining these contributions (and adding 1 for $\nu_{|N|}$) gives the desired upper bound.[2] □

These three claims together imply Proposition 6.4.1. □

6.5 Splitting the dissociated sets

In this section we manipulate the dissociated sets D_j in order to make their sums suitably generic; this will avoid "bad" scenarios in the random orderings of the D_j's that we will consider in the next section. Recall that if D is a dissociated set, then all of the subset sums of D are distinct. In particular, if we choose a uniformly random partition of D into parts $D^{(1)}, D^{(2)}, D^{(3)}, D^{(4)}$ of equal size (up to rounding), then (omitting floor functions) for each $1 \leqslant i \leqslant 4$ the $\binom{|D|}{|D|/4}$ possible values of $\sum_{d \in D^{(i)}} d$ are all achieved with equal probability; likewise, each of the quantities $\sum_{d \in D^{(1)} \cup D^{(2)}} d, \sum_{d \in D^{(2)} \cup D^{(3)}} d, \sum_{d \in D^{(3)} \cup D^{(4)}}$ is uniformly distributed on $\binom{|D|}{|D|/2}$ possible values, and each of the quantities $\sum_{d \in D^{(1)} \cup D^{(2)} \cup D^{(3)}} d, \sum_{d \in D^{(2)} \cup D^{(3)} \cup D^{(4)}} d$ is uniformly distributed on $\binom{|D|}{3|D|/4}$ possible values. Since $\binom{|D|}{|D|/4}, \binom{|D|}{|D|/2}, \binom{|D|}{3|D|/4}$ are all $e^{\Omega(|D|)}$, we obtain very strong anti-concentration for the sums under consideration. We record this simple but important fact in the following lemma.

Lemma 6.5.1. *Let $D \subset G$ be a dissociated set, and let $D = D^{(1)} \cup D^{(2)} \cup D^{(3)} \cup D^{(4)}$ be a uniformly random partition of D into four sets of equal size (up to rounding). Then for every nonempty proper interval $I \subseteq [4]$ and every $x \in G$, we have*

$$\mathbb{P}\left(\sum_{i \in I} \sum_{d \in D^{(i)}} d = x\right) \leqslant e^{-\Omega(|D|)}.$$

[2] To bound the intersection between $\mathrm{IS}(\mathbf{p})$ and Y_j^+, one simply interchanges "smaller" and "larger" throughout the proof. Since we have $|P_k \setminus P'_k| \leq 1$ instead of $|N_\ell \setminus N'_\ell| \leq 2$, we could replace $\ell - 2$ with $k - 1$ in the second paragraph and replace $4|Y_j^-|$ with $3|Y_j^+|$ in the third paragraph to obtain even a slightly tighter bound.

Let $D_1, \dots, D_s$ be the dissociated sets appearing in the structural decomposition of A from Proposition 6.3.4. We will split and reorder these dissociated sets as follows. For each $j \in [1, s]$, we partition $D_j = \cup_{i=1}^4 D_j^{(i)}$ into four sets of equal size uniformly at random as in Lemma 6.5.1, and we require that $|D_1^{(1)}| = |D_s^{(4)}|$. We do all of these splittings independently. Next, we place these newly formed dissociated sets in the order

$$D_1^{(1)}, D_1^{(2)}, D_2^{(1)}, D_2^{(2)}, \dots, D_s^{(1)}, D_s^{(2)}, D_1^{(3)}, D_1^{(4)}, D_2^{(3)}, D_2^{(4)}, \dots, D_s^{(3)}, D_s^{(4)} \tag{6.5}$$

and note that of course the decomposition

$$\lambda \cdot A = P \cup N \cup (\cup_{i=1}^4 \cup_{j=1}^s D_j^{(i)})$$

still holds (with the same value of δ). For notational convenience, write $T_1, T_2, \dots, T_u$ (with $u = 4s$) for the new sequence of dissociated sets in (6.5), and let $\tau_j := \sum_{t \in T_j} t$.

Let us pause at this point and describe the remainder of the strategy for proving Theorem 6.1.2. We will eventually construct a two-sided valid ordering of A of the form

$$\overline{\mathbf{p}}, \mathbf{t}_1, \dots, \mathbf{t}_u, \mathbf{n},$$

where each $\mathbf{t}_i$ is an ordering of T_i chosen randomly according to a certain distribution. Our task will be to show that such an ordering $a_1, \dots, a_{|A|}$ is likely to avoid zero-sum subintervals, namely, proper nonempty intervals $I \subset [|A|]$ with $\sum_{i \in I} a_i = 0$. For the remainder of the chapter, we will refer to proper nonempty intervals as simply "intervals". We divide such intervals I into two "types", which we will treat using different arguments. Recall that $K = c_2 R^{1/3}$.

Definition 6.5.2. *Let $I \subset [|A|]$ be a proper nonempty interval. We say that I is* Type II *if it contains between K and $|T_j| - K$ elements of some T_j, and otherwise we say that it is* Type I.

We will refer to the first K elements in an ordering $\mathbf{t} = t_1, \dots, t_m$ as its left border and to the final K elements as its right border. The remaining elements $t_{K+1}, \dots, t_{m-K}$ make up the interior region of $\mathbf{t}$. In this language (and ignoring intervals contained in a single T_j, which can never be zero-sum), a Type II interval is an interval with at least one endpoint in the interior region of one of the orderings $\mathbf{t}_j$, and a Type I interval is an interval with each endpoint in $\overline{\mathbf{p}}$, $\mathbf{n}$, or a border region of some $\mathbf{t}_j$. One should think of Type II intervals as generic and of Type I intervals as exceptional. (Obviously the identification of intervals $I \subset [|A|]$ as Type I and Type II does not depend on the random choices of the $\mathbf{t}_j$'s).

The main benefit of the above splitting-and-rearranging procedure is that it lets us dispose of nearly all Type I intervals even before we choose the random orderings $\mathbf{t}_j$. The following lemma makes this precise. We say that an event holds *with high probability* if it holds with probability tending to 1 as p tends to infinity.

Lemma 6.5.3. *Let $c > 0$ be any constant. Let $1 \leqslant s \leqslant e^{c(\log p)^{1/4}}$, and let $D_1, \dots, D_s \subseteq \mathbb{F}_p$ be dissociated sets each of size $\asymp R$, with the property that $D_1 \cup D_s \cup \{\delta\}$ is dissociated. Let $\mathbf{p}$ and $\mathbf{n}$ be sequences over $\mathbb{F}_p$ each of length at most $e^{c(\log p)^{1/4}}$, and assume that $\overline{\mathbf{p}}, \delta, \mathbf{n}$ is a two-sided valid ordering. If the sequence $T_1, \dots, T_u$ of dissociated sets is chosen randomly as described above, then each $|T_j| \asymp R$ and each $T_{2j-1} \cup T_{2j}$ is dissociated, and the following holds with high probability:*

(i) for each proper nonempty interval $I = [i, j] \subseteq [u]$, we have that

$$0 \notin \left(\sum_{\leqslant K} (T_{i-1}) \cup - \sum_{\leqslant K} (T_i) \right) + \tau_i + \cdots + \tau_j + \left(- \sum_{\leqslant K} (T_j) \cup \sum_{\leqslant K} (T_{j+1}) \right)$$

(with the convention that $T_0 = T_{u+1} = \emptyset$);

(ii) for each $1 \leqslant j \leqslant u - 1$, we have that

$$0 \notin \mathrm{IS}(\mathbf{p}) + \tau_1 + \cdots + \tau_j + \left(- \sum_{\leqslant K} (T_j) \cup \sum_{\leqslant K} (T_{j+1}) \right);$$

and for each $2 \leqslant j \leqslant u$, we have that

$$0 \notin \mathrm{IS}(\mathbf{n}) + \tau_u + \cdots + \tau_j + \left(- \sum_{\leqslant K} (T_j) \cup \sum_{\leqslant K} (T_{j-1}) \right);$$

(iii) the ordering $\overline{\mathbf{p}}, \tau_1 \dots, \tau_u, \mathbf{n}$ is two-sided valid.

Three remarks are in order before we proceed to the proof.

1. To see how this lemma pertains to Type I intervals containing nearly all (i.e., at least $|T_j| - K$ elements) of some T_j, simply note the identity $\sum_{\geqslant |T_j| - K} (T_j) = \tau_j - \sum_{\leqslant K} (T_j)$.

2. Items (i)–(iii) handle all Type I intervals except for the following:

   - intervals fully contained in a single T_j;
   - intervals starting in the left border of $\mathbf{t}_1$ and ending in the right border of $\mathbf{t}_u$;

- intervals beginning in the right border of $\mathbf{t}_j$ and ending in the left border of $\mathbf{t}_{j+1}$ for some j;
- intervals with one endpoint in $\overline{\mathbf{p}}$ or $\mathbf{n}$ and the other endpoint in the left border of $\mathbf{t}_1$ or the right border of $\mathbf{t}_u$.

Moreover, the first case cannot lead to zero-sum intervals because each T_j is dissociated; likewise, there cannot be zero-sum intervals in the second case because of the assumption that $D_1 \cup D_s \cup \{\delta\}$ (and a fortiori $T_1 \cup T_u \cup \{\delta\}$) is dissociated. In the third case, we never have to worry about zero-sum intervals with j odd since each $T_{2k-1} \cup T_{2k}$ is dissociated.

3. The lemma would continue to hold with K as large as a small constant times R, but we will not have occasion to make use of this fact.

Proof of Lemma 6.5.3. We begin with the crucial observation that if $I \subset [u]$ is any proper nonempty subinterval and $x \in \mathbb{F}_p$ is any element, then we have the anti-concentration inequality

$$\mathbb{P}\left(\sum_{i \in I} \tau_i = x\right) = e^{-\Omega(R)}.$$

Indeed, there is some $j \in [s]$ such that $\{T_i : i \in I\}$ contains at least one but not all of $D_j^{(1)}, \dots, D_j^{(4)}$. Suppose that it contains $D_j^{(1)}$ but none of $D_j^{(2)}, D_j^{(3)}, D_j^{(4)}$ (the remaining cases are analogous). Since the splitting of D_j is independent of the splittings of the other D_k's, Lemma 6.5.1 gives

$$\begin{aligned}
\mathbb{P}\left(\sum_{i \in I} \tau_i = x\right) &= \sum_{z \in \mathbb{F}_p} \mathbb{P}\left(\sum_{i \in I \setminus \{2j-1\}} \tau_i = z \quad \text{and} \quad \tau_{2j-1} = x - z\right) \\
&= \sum_{z \in \mathbb{F}_p} \mathbb{P}\left(\sum_{i \in I \setminus \{2j-1\}} \tau_i = z\right) \mathbb{P}\left(\sum_{d \in D_j^{(1)}} d = x - z\right) \\
&\leqslant \sum_{z \in \mathbb{F}_p} \mathbb{P}\left(\sum_{i \in I \setminus \{2j-1\}} \tau_i = z\right) e^{-\Omega(|D_j|)} = e^{-\Omega(R)}.
\end{aligned}$$

With this observation in hand, we proceed to the main body of the proof. Note that (iii) holds whenever (i) and (ii) hold since $0 \in \sum_{\leqslant K}(T_j)$ and as we assumed that $\overline{\mathbf{p}}, \delta, \mathbf{n}$ is two-sided valid. So, by the union bound, it suffices to show that each of (i) and (ii) holds with high probability.

We begin with (i). Fix some proper nonempty interval $I = [i, j] \subseteq [u]$. The assertion of (i) for this I is that

$$\tau_i + \cdots + \tau_j \notin \left(-\sum_{\leqslant K} (T_{i-1}) \cup \sum_{\leqslant K} (T_i) \right) + \left(\sum_{\leqslant K} (T_j) \cup -\sum_{\leqslant K} (T_{j+1}) \right).$$

The set on the right-hand side has size at most

$$\left(\left| \sum_{\leqslant K} (T_{i-1}) \right| + \left| \sum_{\leqslant K} (T_i) \right| \right) \left(\left| \sum_{\leqslant K} (T_j) \right| + \left| \sum_{\leqslant K} (T_{j+1}) \right| \right) \leqslant e^{O\left(R \times H\left(O\left(\frac{K}{R} \right) \right) \right)},$$

where $H(x) := -x \log_2(x) - (1-x) \log_2(1-x)$ is the binary entropy function. The definition of K ensures that $H(O(K/R)) = o(1)$ (with a lot of room to spare), and then the observation from the beginning of the proof tells us that (i) fails for I with probability at most $e^{o(R) - \Omega(R)} = e^{-\Omega(R)}$. A union bound over the (at most u^2) choices of I shows that (i) fails with probability at most

$$u^2 e^{-\Omega(R)} \leqslant e^{2c(\log p)^{1/4} - \Omega(c_1 (\log p)^{3/4})} = o(1),$$

again with plenty of room to spare.

The proof of (ii) is nearly identical and we omit it; we remark that the bounds $|\operatorname{IS}(\mathbf{p})|, |\operatorname{IS}(\mathbf{n})| \leqslant |A| \leqslant e^{c(\log p)^{1/4}}$ hold because we fixed $\mathbf{p}$ and $\mathbf{n}$ in advance.

□

As noted in remark (2) following Lemma 6.5.3, there remain two sorts of Type I intervals to address. The first is Type I intervals contained in $T_{2k} \cup T_{2k+1}$ for some k. We can avoid zero-sums here by picking the orderings $\mathbf{t}_{2k}, \mathbf{t}_{2k+1}$ according to a suitable joint distribution which we will describe in section 6.6. The second is Type I intervals with one endpoint in $\overline{\mathbf{p}}$ or $\mathbf{n}$ and the other endpoint in the left border of $\mathbf{t}_1$ or the right border of $\mathbf{t}_u$. The crucial ingredient for dealing with these will turn out to be the last part of Proposition 6.4.1. Since Proposition 6.4.1 must be applied prior to the random splitting procedure described in this section, it is a bit of a nuisance that the input sets Y_j^+, Y_j^- must be described in terms of the sets D_j rather than the sets T_j. The following lemma will let us remedy this issue.

Lemma 6.5.4. *Let $T_1 = D_1^{(1)}$ and $T_u = D_s^{(4)}$ be the random sets from (6.5). Then with probability at least $1/2$, we have for all $1 \leqslant j \leqslant K$ that*

$$\left| \sum_{=j} (T_1) \cap (-\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \right| \leqslant 4K \frac{\binom{|T_1|}{j}}{\binom{|D_1|}{j}} \left| \sum_{=j} (D_1) \cap (-\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \right|,$$

$$\left|\sum_{=j}(T_u)\cap(-\operatorname{IS}(\mathbf{p})\cup(\delta+\operatorname{IS}(\mathbf{n})))\right| \leqslant 4K\frac{\binom{|T_u|}{j}}{\binom{|D_s|}{j}}\left|\sum_{=j}(D_s)\cap(-\operatorname{IS}(\mathbf{p})\cup(\delta+\operatorname{IS}(\mathbf{n})))\right|.$$

Proof. Since D_1 is dissociated, the quantity $\left|\sum_{=j}(D_1)\cap(-\operatorname{IS}(\mathbf{p})\cup(\delta+\operatorname{IS}(\mathbf{n})))\right|$ simply counts the subsets $S\subseteq D_1$ of size $|S|=j$ with $\sum_{d\in S}d\in-\operatorname{IS}(\mathbf{p})\cup(\delta+\operatorname{IS}(\mathbf{n}))$, and likewise for T_1. As $T_1=D_1^{(1)}$ is chosen uniformly from all subsets of D_1 of size $|D_1|/4$, we have

$$\mathbb{E}\left(\left|\sum_{=j}(T_1)\cap(-\operatorname{IS}(\mathbf{p})\cup(\delta+\operatorname{IS}(\mathbf{n})))\right|\right)=\frac{\binom{|T_1|}{j}}{\binom{|D_1|}{j}}\left|\sum_{=j}(D_1)\cap(-\operatorname{IS}(\mathbf{p})\cup(\delta+\operatorname{IS}(\mathbf{n})))\right|,$$

and Markov's Inequality implies that the first bound in the conclusion of the lemma fails for each j with probability at most $1/4K$. The same argument applies with D_s, T_u in place of D_1, T_1, and the conclusion of the lemma follows from a union bound over $1\leqslant j\leqslant K$. □

6.6 Randomizing the dissociated sets

We are finally ready to describe how we will construct a two-sided valid ordering of A. Suppose that $A\subseteq\mathbb{F}_p\setminus\{0\}$ has size $|A|\leqslant e^{c(\log p)^{1/4}}$. After we replace A by a suitable dilate (which is harmless with regard to finding two-sided valid orderings), Proposition 6.3.4 provides a decomposition

$$A=P\cup N\cup(\cup_{j=1}^{s}D_j)$$

satisfying conditions (i)-(v) of that proposition and $\delta=\sum_j\sum_{d\in D_j}d>0$. Now, with $K=c_2R^{1/3}$ for a suitably small constant $c_2>0$, set

$$Y_j^+:=-\sum_{=j}(D_1)\cup\left(-\delta+\sum_{=j}(D_s)\right)\quad\text{and}\quad Y_j^-:=-\sum_{=j}(D_s)\cup\left(-\delta+\sum_{=j}(D_1)\right) \tag{6.6}$$

for each $1\leqslant j\leqslant K$, and apply Proposition 6.4.1. This provides orderings $\mathbf{p}$ of P and $\mathbf{n}$ of N such that the sequence $\overline{\mathbf{p}},\delta,\mathbf{n}$ is two-sided valid and such that (6.3),(6.4) hold. Finally, we can use Lemmas 6.5.3 and 6.5.4 to obtain dissociated sets $T_1\ldots,T_u$ from $D_1,\ldots,D_s$ such that $A=P\cup N\cup(\cup_iT_i)$ and the conclusions of these two lemmas are simultaneously satisfied; fix such a choice of $T_1,\ldots,T_u$. Recall that $\tau_i=\sum_{t\in T_i}t$, and write $m_i:=|T_i|$. The two-sided valid ordering of A that we will construct will be of the form

$$\overline{\mathbf{p}},\mathbf{t}_1,\ldots,\mathbf{t}_u,\mathbf{n},$$

where the $\mathbf{t}_i$'s are orderings of the T_i's chosen randomly according to certain distributions, whose description and analysis occupies the remainder of this section.

Recall that if T is a dissociated set, then all of the subset sums of T are distinct. In particular, in a uniformly random ordering of the elements of T, the sum of the first k elements is uniformly distributed on $\binom{|T|}{k}$ values. As long as k is not too close to 0 or $|T|$, this sum is very anti-concentrated, and so with very high probability it will avoid any particular small set of values. It follows that uniformly random orderings $\mathbf{t}_i$ would with high probability avoid zero-sum Type II intervals. We can ignore most Type I intervals due to Lemma 6.5.3, but the remaining Type I intervals, as described in remark (2) following that lemma, still cause issues. We will show that each of these potential zero-sum Type I intervals can be avoided "locally" by introducing some non-uniformity into the distributions determining the orderings $\mathbf{t}_i$.

We begin with the orderings $\mathbf{t}_1, \mathbf{t}_u$. Say that an ordering $t_1, \dots, t_{m_1}$ of T_1 is *acceptable* if

$$t_1 + \cdots + t_k \notin -\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n})) \quad \text{for all } 1 \leqslant k \leqslant K,$$

and say that an ordering $t_1, \dots, t_{m_u}$ of T_u is *acceptable* if

$$t_1 + \cdots + t_k \notin -\operatorname{IS}(\mathbf{n}) \cup (\delta + \operatorname{IS}(\mathbf{p})) \quad \text{for all } 1 \leqslant k \leqslant K.$$

Using Proposition 6.4.1 and the fact that T_1, T_u satisfy the conclusion of Lemma 6.5.4, we can show that uniformly random orderings of T_1, T_u are acceptable with large probability.

Lemma 6.6.1. *A uniformly random ordering of T_1 is acceptable with probability at least 0.98, and a uniformly random ordering of T_u is acceptable with probability at least 0.98.*

Proof. We prove only the statement for T_1 since the argument for T_u is identical. Let $t_1, \dots, t_{|T_1|}$ be our uniformly random ordering of T_1. By the union bound, it suffices to show that $\mathbb{P}(t_1 \in -\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \leqslant 0.01$ and that

$$\mathbb{P}(t_1 + \cdots + t_k \in -\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \leqslant 0.01K^{-1}$$

for each $2 \leqslant k \leqslant K$. Fix some $1 \leqslant k \leqslant K$. Then the quantity $t_1 + \cdots + t_k$ is uniformly distributed on the set $\sum_{=k}(T_1)$, which has size $\binom{|T_1|}{k}$. Recall that we applied Proposition 6.4.1 with the sets Y_j^+, Y_j^- as in (6.6). Since $|D_1| = |D_s|$ and

$D_1 \cup D_s \cup \{\delta\}$ is dissociated by Proposition 6.3.4, we have $|Y_j^+| = |Y_j^-| = 2\binom{|D_1|}{j}$ for all j. Then the conclusion of Proposition 6.4.1, with $L := \lfloor |Y_k^+|^{1/2} \rfloor$, gives

$$\left| \sum_{=k} (D_1) \cap (-\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \right| \ll |Y_k^+|^{1/2} + 1 + \sum_{j<k} |Y_j^+|.$$

For $k = 1$, this gives (recall that $|D_1| \asymp R$)

$$\left| \sum_{=1} (D_1) \cap (-\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \right| \ll |Y_1^+|^{1/2} \ll \binom{|D_1|}{1} \cdot R^{-1/2},$$

and for $2 \leqslant k \leqslant K$ (recall that $K = c_2 R^{1/3}$) it gives

$$\left| \sum_{=k} (D_1) \cap (-\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \right| \ll \binom{|D_1|}{k} \cdot \frac{K}{|D_1|} \ll \binom{|D_1|}{k} \cdot c_2^3 K^{-2}.$$

Since the conclusion of Lemma 6.5.4 also holds, we can "transfer" this bound from D_1 to T_1. In particular, we obtain that

$$\left| \sum_{=1} (T_1) \cap (-\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \right| \ll 4K \binom{|T_1|}{1} \cdot R^{-1/2}$$

and that

$$\left| \sum_{=k} (T_1) \cap (-\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \right| \ll 4K \binom{|T_1|}{k} \cdot \frac{K}{|D_1|} \ll 4K \binom{|T_1|}{k} \cdot c_2^3 K^{-2}$$

for $2 \leqslant k \leqslant K$. It follows that

$$\mathbb{P}(t_1 \in -\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \ll R^{-1/6}$$

is certainly at most 0.01, and for $2 \leqslant k \leqslant K$ we see that

$$\mathbb{P}(t_1 + \cdots + t_k \in -\operatorname{IS}(\mathbf{p}) \cup (\delta + \operatorname{IS}(\mathbf{n}))) \ll c_2^3 K^{-1}$$

is at most $0.01K^{-1}$ as long as c_2 is sufficiently small. This completes the proof.

□

We now choose $\mathbf{t}_1, \mathbf{t}_u$ independently such that $\mathbf{t}_1, \overline{\mathbf{t}_u}$ are uniformly random acceptable orderings of T_1, T_u, respectively. We deduce from Lemma 6.6.1 that the random variables $\mathbf{t}_1, \mathbf{t}_u$ are highly anti-concentrated in the sense that the probability of $\mathbf{t}_1$ assuming any particular ordering is at most $\ll 1/m_1!$, and likewise for $\mathbf{t}_u$. Notice that the constraint that $\mathbf{t}_1, \overline{\mathbf{t}_u}$ are acceptable precisely guarantees the absence of zero-sum

Type I intervals with one endpoint in $\overline{\mathbf{p}}$ or $\mathbf{n}$ and the other endpoint in the left border of $\mathbf{t}_1$ or the right border of $\mathbf{t}_u$.

For each $1 \leqslant j \leqslant u/2 - 1$, we choose the pair of orderings $\mathbf{t}_{2j}, \mathbf{t}_{2j+1}$ as follows. Recall that $|T_{2j}| = m_{2j}$ and $|T_{2j+1}| = m_{2j+1}$ both have size $\asymp R$. Say that a pair of partial orderings $t_1, \dots, t_k$ of T_{2j} and $t'_1, \dots, t'_\ell$ of T_{2j+1} is *permissible* if

$$t_1 + \cdots + t_i + t'_1 + \cdots + t'_j \neq 0 \quad \text{for all } (i, j).$$

Let $N(k, \ell)$ denote the number of permissible pairs with lengths (k, ℓ). Note that each permissible pair with lengths (k, ℓ) can be extended to at least $m_{2j} - k - \ell$ permissible pairs of lengths $(k+1, \ell)$ and to at least $m_{2j+1} - k - \ell$ permissible pairs of lengths $(k, \ell+1)$. It follows that

$$N(k,k) \geqslant (m_{2j})(m_{2j+1}-1)(m_{2j}-2)(m_{2j+1}-3)\cdots(m_{2j}-2k+2)(m_{2j+1}-2k+1).$$

The choice of K ensures that $N(K, K) \geqslant m_{2j}^K m_{2j+1}^K/2$ (say), which means that the permissible pairs comprise at least a constant fraction of the total pairs. (In fact, this bound would continue to hold with K as large as a small constant times $R^{1/2}$.)

We now choose $\mathbf{t}_{2j}, \mathbf{t}_{2j+1}$ to be a uniformly random pair of orderings of T_{2j}, T_{2j+1} conditional on the length-K prefixes of $\overline{\mathbf{t}_{2j}}, \mathbf{t}_{2j+1}$ forming a permissible pair of length (K, K). Equivalently, we let $t_1, \dots, t_K$ and $t'_1, \dots, t'_K$ be a uniformly random permissible pair of orderings of T_{2j}, T_{2j+1}, and then we let $\mathbf{t}_{2j}$ be a uniformly random ordering of T_{2j} conditional on $\overline{\mathbf{t}_{2j}}$ beginning with $t_1, \dots, t_K$, and we let $\mathbf{t}_{2j+1}$ be a uniformly random ordering of T_{2j+1} conditional on $\mathbf{t}_{2j+1}$ beginning with $t'_1, \dots, t'_K$. We make these choices independently for different values of j, and independently of the choices of $\mathbf{t}_1, \mathbf{t}_u$. The following lemma shows that even though the random variables $\mathbf{t}_{2j}, \mathbf{t}_{2j+1}$ are dependent, they are "conditionally anti-concentrated" in the sense that if we condition on $\mathbf{t}_{2j}$ being any particular ordering, then the the probability of $\mathbf{t}_{2j+1}$ assuming any particular ordering is still very small, and vice versa.

Lemma 6.6.2. *Choose $\mathbf{t}_{2j}, \mathbf{t}_{2j+1}$ according to the distribution described above. Then, conditional on $\mathbf{t}_{2j}$ assuming any particular ordering, the probability of $\mathbf{t}_{2j+1}$ assuming any particular ordering is $\ll 1/m_{2j+1}!$; likewise, conditional on $\mathbf{t}_{2j+1}$ assuming any particular ordering, the probability of $\mathbf{t}_{2j}$ assuming any particular ordering is $\ll 1/m_{2j}!$.*

Proof. We prove only the first statement. Let $\mathbf{u}_{2j}$ be any fixed ordering of T_{2j}. Let $\mathbf{u}_{2j}^{(K)}$ denote the ordering consisting of the first K elements of $\mathbf{u}_{2j}$. The number of permissible pairs $\mathbf{u}_{2j}^{(K)}\mathbf{u}_{2j+1}^{(K)}$ with $\mathbf{u}_{2j+1}^{(K)}$ of length K is at least

$$(m_{2j+1} - K)^K \gg m_{2j+1}^K,$$

by our choices of R, K (see the above discussion of $N(k,\ell)$). Thus, the number of orderings $\mathbf{u}_{2j+1}$ of T_{2j+1} such that $\mathbf{u}_{2j}^{(K)}, \mathbf{u}_{2j+1}^{(K)}$ is a permissible pair is

$$\gg m_{2j+1}^K \cdot (m_{2j+1} - K)! \gg m_{2j+1}!.$$

The lemma now follows since each of these $\gg m_{2j+1}!$ orderings of T_{2j+1} is equally likely to occur as $\mathbf{t}_{2j+1}$, after we condition on $\mathbf{t}_{2j} = \mathbf{u}_{2j}$. □

Notice that the constraints on the pairs $\mathbf{t}_{2j}\mathbf{t}_{2j+1}$ guarantee the absence of zero-sum Type I intervals beginning in the right border of $\mathbf{t}_{2j}$ and ending in the left border of $\mathbf{t}_{2j+1}$.

We will show that if the orderings $\mathbf{t}_1, \dots, \mathbf{t}_u$ of $T_1, \dots, T_u$ are chosen randomly as above, then with high probability the ordering

$$a_1, \dots, a_{|A|} := \overline{\mathbf{p}}, \mathbf{t}_1, \dots, \mathbf{t}_u, \mathbf{n} \tag{6.7}$$

of A is two-sided valid, i.e., we have $\sum_{i\in I} a_i \neq 0$ for every nonempty proper interval $I \subseteq [|A|]$. The output of Lemma 6.5.3 and the constraints on the random orderings $\mathbf{t}_j$ together guarantee that there are no zero-sum Type I intervals in the ordering (6.7); the reader can refer to remark (2) following Lemma 6.5.3 to see how we have covered all possible cases. It remains to verify that with high probability there are no zero-sum Type II intervals. The key point is that the sum $\sum_{i\in I} a_i$ for each Type II interval I is highly anti-concentrated because there is still enough randomness in the orderings $\mathbf{t}_j$; the following lemma makes this observation precise.

Lemma 6.6.3. *Let $I \subset [1, |A|]$ be a Type II interval, and let $a_1, \dots, a_{|A|} = \overline{\mathbf{p}}, \mathbf{t}_1, \dots, \mathbf{t}_u, \mathbf{n}$ be the random ordering (6.7) of A. Then for every $x \in \mathbb{F}_p$ we have*

$$\mathbb{P}\left(\sum_{i\in I} a_i = x\right) \leqslant e^{-\Omega(K \log R)}.$$

Proof. By definition, there exists some j such that I contains exactly k elements of T_j, where $K \leqslant k \leqslant |T_j| - K$. As in the first step of the proof of Lemma 6.5.3, we break the sum over I into the sum over the part intersecting T_j and the part not intersecting T_j and condition on $\mathbf{t}_i$ for all $i \neq j$. Lemma 6.6.2 ensures that even after this conditioning, the probability of $\mathbf{t}_j$ assuming any particular ordering is $\ll 1/m_j!$. Since the k-element subsets of T_j all have distinct sums, we see that the sum over the part of I intersecting T_j assumes each particular value with probability

$$\ll \binom{m_j}{k}^{-1} \leqslant \binom{m_j}{K}^{-1} \leqslant e^{-\Omega(K \log R)},$$

and the lemma follows.

□

Recall that $K = c_2 R^{1/3}$ and that $R = R(A) \gg c_1 (\log p)^{3/4}$ holds when $|A| \leqslant e^{c(\log p)^{1/4}}$ (see (6.1)). Since the number of Type II intervals is trivially at most $|A|^2 \leqslant e^{2c(\log p)^{1/4}}$, Lemma 6.6.3 and the union bound imply that the probability of (6.7) containing a zero-sum Type II interval is at most

$$e^{2c(\log p)^{1/4} - \Omega(K \log R)} = o(1),$$

again with plenty of room to spare. From this and the above observations about the absence of zero-sum Type I intervals, we conclude that (6.7) is two-sided valid with high probability; in particular, for p sufficiently large (in terms of c), there is at least one two-sided valid ordering of A. This proves Theorem 6.1.2.

One can in fact take c to grow as, e.g., $\asymp \log \log p$, but we are not concerned with such lower-order terms since we have not even seriously optimized the exponent $1/4$ in Theorem 6.1.2.

6.7 Remarks

We make a couple of remarks about our proof of Theorem 6.1.2.

- The union bound in Lemma 6.5.4 is one of the main bottlenecks for the value of the exponent $1/4$ in Theorem 6.1.2. Improving the argument around this lemma would likely let one take K to be a larger power of R, which in turn would let one increase $1/4$ (perhaps to $1/3$) in Theorem 6.1.2.

- In proposition 6.3.4, we can also obtain the extra property that each of P, N is either empty or of size at least $100s$ (say), by splitting each dissociated set into 201 parts and then absorbing up to 100 elements of each of P, N if P, N are small. This property was useful in an earlier version of our proof and may be of interest in the future.

Bibliography

[1] N. Alon. Paul Erdős and probabilistic reasoning. In *Erdős centennial*, volume 25 of *Bólyai Soc. Math. Stud.*, pages 11–33. János Bolyai Math. Soc., Budapest, 2013.

[2] N. Alon and D. J. Kleitman. Sum-free subsets. In *A tribute to Paul Erdős*, pages 13–26. Cambridge Univ. Press, Cambridge, 1990.

[3] S. Baier. A note on P-sets. *Integers*, 4:A13, 6, 2004.

[4] I. Bardaji and D. J. Grynkiewicz. Long arithmetic progressions in small sumsets. *Integers*, 10:A28, 335–350, 2010.

[5] B. Bedert. On a problem of Erdős and Sárközy about sequences with no term dividing the sum of two larger terms. *Online preprint available at https://arxiv.org/abs/2301.07065*, 2022.

[6] B. Bedert. On unique sums in abelian groups. *Combinatorica*, 44(2):269–298, 2024.

[7] B. Bedert. Large sum-free subsets of sets of integers via l^1-estimates for trigonometric series. *Online preprint available at https://arxiv.org/abs/2502.08624*, 2025.

[8] B. Bedert and N. Kravitz. Graham's rearrangement conjecture beyond the rectification barrier. *Israel Journal of Mathematics (forthcoming), online preprint available at https://arxiv.org/abs/2409.07403*, 2025.

[9] Y. F. Bilu, V. F. Lev, and I. Z. Ruzsa. Rectification principles in additive number theory. *Discrete Comput. Geom.*, 19(3, Special Issue):343–353, 1998.

[10] J.-P. Bode and H. Harborth. Directed paths of diagonals within polygons. *Discrete Mathematics*, 299:3–10, 2005.

[11] J. Bourgain. On arithmetic progressions in sums of sets of integers. In *A Tribute to Paul Erdős*, pages 105–109. Cambridge University Press, 1990.

[12] J. Bourgain. Estimates related to sumfree subsets of sets of integers. *Israel J. Math.*, 97:71–92, 1997.

[13] J. Browkin, B. Diviš, and A. Schinzel. Addition of sequences in general fields. *Monatsh. Math.*, 82(4):261–268, 1976.

[14] S. Costa and M. A. Pellegrini. Some new results about a conjecture by brian alspach. *Archiv der Mathematik (Basel)*, 115:479–488, 2020.

[15] T. Do Duc and B. Schmidt. Unique differences in symmetric subsets of $\mathbb{F}_p$. *Combinatorica*, 37(2):167–182, 2017.

[16] S. Eberhard. Følner sequences and sum-free sets. *Bull. Lond. Math. Soc.*, 47(1):21–28, 2015.

[17] S. Eberhard, B. Green, and F. Manners. Sets of integers with no large sum-free subset. *Ann. of Math. (2)*, 180(2):621–652, 2014.

[18] C. Elsholtz and S. Planitzer. On Erdős and Sárközy's sequences with Property P. *Monatsh. Math.*, 182(3):565–575, 2017.

[19] P. Erdős. Extremal problems in number theory. *Proc. Sympos. Pure Math.*, VIII AMS:181–189, 1965.

[20] P. Erdős. Problems and results on combinatorial number theory. In *A survey of combinatorial theory (Proc. Internat. Sympos., Colorado State Univ., Fort Collins, Colo., 1971)*, pages 117–138, 1973.

[21] P. Erdős. Problems and results in combinatorial number theory. In *Journees Arithmétiques de Bordeaux (Conf., Univ. Bordeaux, Bordeaux, 1974)*, pages 295–310. Astérisque, Nos. 24–25. 1975.

[22] P. Erdős. Problems in number theory and combinatorics. In *Proceedings of the Sixth Manitoba Conference on Numerical Mathematics (Univ. Manitoba, Winnipeg, Man., 1976)*, Congress. Numer., XVIII, pages 35–58. Utilitas Math., Winnipeg, Man., 1977.

[23] P. Erdős. A survey of problems in combinatorial number theory. *Ann. Discrete Math.*, 6:89–115, 1980.

[24] P. Erdős. Problems and results on extremal problems in number theory, geometry, and combinatorics. In *Proceedings of the 7th Fischland Colloquium, I (Wustrow, 1988)*, number 38, pages 6–14, 1989.

[25] P. Erdős. Some old and new problems on additive and combinatorial number theory. In *Combinatorial Mathematics: Proceedings of the Third International Conference (New York, 1985)*, volume 555 of *Ann. New York Acad. Sci.*, pages 181–186. New York Acad. Sci., New York, 1989.

[26] P. Erdős. Some problems and results on combinatorial number theory. In *Graph theory and its applications: East and West (Jinan, 1986)*, volume 576 of *Ann. New York Acad. Sci.*, pages 132–145. New York Acad. Sci., New York, 1989.

[27] P. Erdős. Problems in number theory. *New Zealand J. Math.*, 26(2):155–160, 1997.

[28] P. Erdős. Some old and new problems in various branches of combinatorics. volume 165/166, pages 227–231. 1997. Graphs and combinatorics (Marseille, 1995).

[29] P. Erdős and A. Sárközy. On the divisibility properties of sequences of integers. *Proc. London Math. Soc. (3)*, 21:97–101, 1970.

[30] P. Erdős. Some of my forgotten problems in number theory. *Hardy-Ramanujan J.*, 15:34–50 (1993), 1992.

[31] P. Erdős and R. L. Graham. *Old and new problems and results in combinatorial number theory*. L'Enseignement mathématique, Université de Genève, 1980.

[32] P. Erdős and P. Turán. On a problem of Sidon in additive number theory, and on some related problems. *J. London Math. Soc.*, 16:212–215, 1941.

[33] P. Frankl. An extremal problem for two families of sets. *European J. Combin.*, 3(2):125–127, 1982.

[34] G. A. Freĭman. The addition of finite sets. I. *Izv. Vysš. Učebn. Zaved. Matematika*, 1959(6(13)):202–213, 1959.

[35] R. Graham. On sums of integers taken from a fixed sequence. In *Proceedings, Washington State University Conference on Number Theory*, pages 22–40, 1971.

[36] B. Green. 100 open problems. *Manuscript, available online at https://people.maths.ox.ac.uk/greenbj/papers/open-problems.pdf*.

[37] B. Green and I. Z. Ruzsa. Sets with small sumset and rectification. *Bull. London Math. Soc.*, 38(1):43–52, 2006.

[38] R. K. Guy. Unsolved problems in number theory. *Problem Books in Mathematics*, Springer-Verlag, New York, 2004.

[39] M. Jańczak. A note on a problem of Hilliker and Straus. *Electron. J. Combin.*, 14(1):N, 23, 8, 2007.

[40] Y. Jing and S. Wu. The largest (k, ℓ)-sum-free subsets. *Trans. Amer. Math. Soc.*, 374(7):5163–5189, 2021.

[41] Y. Jing and S. Wu. A note on the largest sum-free sets of integers. *J. Lond. Math. Soc. (2)*, 109(1):Paper No. e12819, 19, 2024.

[42] S. V. Konyagin. On the Littlewood problem. *Izv. Akad. Nauk SSSR Ser. Mat.*, 45(2):243–265, 463, 1981.

[43] S. V. Konyagin and V. F. Lev. Combinatorics and linear algebra of Freiman's isomorphism. *Mathematika*, 47(1-2):39–51, 2000.

[44] N. Kravitz. Rearranging small sets for distinct partial sums, 2024. Preprint arXiv:2407.01835v2.

[45] K. H. Leung and B. Schmidt. Unique sums and differences in finite abelian groups. *J. Number Theory*, 233:370–388, 2022.

[46] V. F. Lev. The rectifiability threshold in abelian groups. *Combinatorica*, 28(4):491–497, 2008.

[47] M. Lewko. An improved upper bound for the sum-free subset constant. *J. Integer Seq.*, 13(8):Article 10.8.3, 15, 2010.

[48] T. Luczak and T. Schoen. On a problem of Konyagin. *Acta Arith.*, 134(2):101–109, 2008.

[49] J. L. Malouf. Combinatorial approaches to integer sequences. *ProQuest LLC, Ann Arbor, MI*, 1994. Thesis (Ph.D.)–University of Illinois at Urbana-Champaign.

[50] O. C. McGehee, L. Pigno, and B. Smith. Hardy's inequality and the L^1 norm of exponential sums. *Ann. of Math. (2)*, 113(3):613–618, 1981.

[51] H. L. Montgomery and R. C. Vaughan. Multiplicative number theory. I. Classical theory. *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge, 2007.

[52] I. P. Natanson. Constructive function theory. Vol. I. Uniform approximation. *Frederick Ungar Publishing Co.*, New York, 1964. Translated from the Russian by Alexis N. Obolensky.

[53] Z. Nedev. An algorithm for finding a nearly minimal balanced set in $\mathbb{F}_p$. *Math. Comp.*, 78(268):2259–2267, 2009.

[54] S. K. Pichorides. On the L^1 norm of exponential sums. *Ann. Inst. Fourier (Grenoble)*, 30(2):v, 79–89, 1980.

[55] M. Radziejewska. A note on the minimal number of representations in $A + A$. *Integers*, 12(4):705–708, 2012.

[56] K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.

[57] W. Rudin. Trigonometric series with gaps. *J. Math. Mech.*, 9:203–227, 1960.

[58] W. Sawin. Comment on the post "Ordering subsets of the cyclic group to give distinct partial sums". https://mathoverflow.net/q/202857, 2015. MathOverflow.

[59] D. Scheinermann. *Several problems in linear algebraic and additive combinatorics*. PhD thesis, Rutgers University, 2019.

[60] T. Schoen. On a problem of Erdős and Sárközy. *J. Combin. Theory Ser. A*, 94(1):191–195, 2001.

[61] I. Schur. Über die kongruenz $x^m+y^m = z^m \pmod p$. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 25:114–116, 1917.

[62] G. Shakan. On the largest sum-free subset problem in the integers. *preprint, arXiv:2207.14210*, 2022.

[63] E. G. Straus. Differences of residues (mod p). *J. Number Theory*, 8(1):40–42, 1976.

[64] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.

[65] T. Tao and V. Vu. Sum-free sets in groups: a survey. *J. Comb.*, 8(3):541–552, 2017.

[66] T. Tao and V. Vu. Additive combinatorics. *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge, 2010.

[67] A. Zygmund. Trigonometric series. Vol. I, II. *Cambridge Mathematical Library*, Cambridge University Press, Cambridge, 2002.