Universitat Politècnica de Catalunya

Programa de Doctorat de Matemàtica Aplicada

Estructures Additives i Probabilitat en Combinatòria

Additive Structures and Randomness in Combinatorics

Doctoral thesis by

Christoph Spiegel

Thesis advisors

Juanjo Rué Perna and Oriol Serra

Department of Mathematics

Barcelona, June 2020

Christoph Spiegel

2020

Für meine Eltern

❧

Für Carlotta

Abstract

The contents of this thesis are divided into two main parts, each containing several thematically related results sitting at the intersection of Combinatorics and Number Theory. The first deals with threshold behavior in discrete random structures and positional games. In particular, we will study the distribution of solutions to arbitrary linear systems of equations in the binomial random sets as well as the thresholds for Szemerédi- and Rado-type properties. We will likewise examine Maker–Breaker games played on the hypergraph given by such solutions.

The second part of this thesis deals with the extremal behavior of additive structures, in particular with respect to their doubling and their representation function. Here we will study a generalization of Sidon sets proposed very recently by Kohayakawa, Lee, Moreira and Rödl. We will also obtain results in the same vein as the well-known Freĭman–Ruzsa Theorem for the case of particularly small doubling. Lastly, we will study Erdős–Fuchs-type results for infinite sets with near-constant representation functions.

Keywords. Additive Combinatorics, Probabilistic Combinatorics, thresholds for discrete random structures, positional games, Sidon sets, inverse results for sets of small doubling, representation functions

Contents

List of Figures and Tables

Acknowledgments

First and foremost, I want to thank both of my advisors, Juanjo Rué and Oriol Serra, for their never-ending support and wisdom during the last five years and for always guiding me towards interesting and fruitful problems. I furthermore would like to thank Tibor Szabó for making his group a second home to me and for organizing his annual workshops. I am also indebted to David Conlon and Ervin Győri for enabling me to visit their institutions.

My thanks likewise go to my co-authors, that is – if not previously already mentioned – Pablo Candela, Gonzalo Cao-Labora, Jan Corsten, David Fabian, Gregory A. Freĭman, Nina Kamčev, Christopher Kusch, Adva Mond, Alexey Pokrovskiy, Nika Salia, Casey Tompkins, Oscar Zamora and Ana Zumalacárregui, for making much of the work presented here possible and a joy to work on. I am also truly grateful to everyone else in my academic family in Barcelona, namely Pilar Cano, Gonzalo Fiz Pontiveros, Clément Requilé, Vasiliki Velona and Maximilian Wötzel, and my secret second academic family in Berlin, that is Michaelis Anastos, Anurag Bishnoi, Simona Boyadzhiyska, Shagnik Das, Codrut Grosu, Ander Lamaison, Tamás Mészáros, Patrick Morris, Alp Müyesser and Tuan Tran. May they never meet.

Lastly, I would like to thank everyone else that I have met so far during my academic career or who helped my research, be it through discussions or in other ways big and small, including Noga Alon, Simeon Ball, Małgorzata Bednarska-Bzdęga, Katy Beeler, Javier Cilleruelo, Christian Elsholtz, David Grynkiewicz, Robert Hancock, Rob Morris, Marc Noy, Arnau Padrol, Yury Person, Imre Ruzsa, Andrew Treglown, Lluis Vena, and everyone whose name I am either forgetting or are not privy to, in particular the anonymous referees of my papers that gave detailed feedback, constructive suggestions and, sometimes more often than I would have liked, a healthy dose of reality.

This thesis would not have been possible without the funding of the Ministerio de Economía y Empresa, the Berlin Mathematical School, the Fundació Ferran Sunyer i Balaguer and all other institutions that supported my work.

Preface

Arithmetic Combinatorics, Combinatorial Number Theory, Structural Additive Theory and Additive Number Theory are just some of the terms used to describe the vast field that sits at the intersection of Number Theory and Combinatorics and which will be the focus of this thesis. Its contents are divided into two main parts, each containing several thematically related results. The first deals with the question under what circumstances solutions to arbitrary linear systems of equations usually occur in combinatorial structures. For this, we will establish threshold behavior in random sets and positional games. The second deals with the extremal behavior of additive structures. Here we will study Sidon sets, sets of small doubling and sets with near-constant representation functions.

Randomness and Games Discrete random structures often posses seemingly contradictory properties that are difficult to obtain via deterministic constructions. One central point of study have been analogues of some well-established deterministic combinatorial results in sparse random structures. In particular, there is great interest in threshold behavior, that is to decide for which values of that probability a property does or does not hold asymptotically almost surely. The properties we will be interested in studying in this chapter relate to the solutions to linear systems of equations.

A first question one might ask concerns the point at which sets of a given size will typically contain a solution. In Chapter 2, we will address this question and establish a threshold. We will also study the distribution of the number of solutions at that threshold, showing that it converges to a Poisson distribution in certain cases. The parameter of that distribution depends on the volumes of certain convex polytopes arising from the linear system under study.

Van der Waerden's Theorem, stating that every finite coloring of the integers contains monochromatic arithmetic progression of arbitrary length, is by some considered to be the first result in Ramsey Theory. Rado generalized van der Waerden's result by characterizing those linear systems whose solutions satisfy a similar property and Szemerédi strengthened it to a statement concerning density rather than colorings. In Chapter 3 we will turn our attention towards versions of Rado's and Szemerédi's Theorem in random sets, extending previous work of Friedgut, Rödl, Ruciński and Schacht in the case of the former and of Conlon, Gowers and Schacht for the latter to include a larger variety of systems and solutions.

Chvátal and Erdős suggested studying games in which two opponents take turns occupying vertices of a hypergraph, with the first player winning if he occupies an edge and the second winning if he can prevent this. To address an inherent unfairness in these types of games, the second player is allowed to occupy several elements each round. These games have deep connections to the theory of random structures and in Chapter 4 we will build on work of Bednarska and Łuczak to establish the threshold for how much a large variety of games need to be biased in favor of the second player. These include games in which the first player wants to occupy a solution to some given linear system, generalizing the van der Waerden games introduced by Beck.

Additive Structures The focus of the second chapter will be on extremal questions relating to sets with interesting additive properties. In particular, we will be interested in bounds or structural descriptions for sets exhibiting some restrictions with regards to either their representation function or their sumset.

'Sidon sets' are sets of integers with pairwise unique differences. Their study also touches upon the areas of discrete geometry and theoretical computer science and they have had many practical applications. In Chapter 5, we will study a generalization of Sidon sets proposed very recently by Kohayakawa, Lee, Moreira and Rödl. In these sets the pairwise differences of its elements are not just distinct, but in fact far apart by a certain measure. We will obtain strong lower bounds for such infinite sets using an approach of Cilleruelo. As a consequence of these bounds, we will also obtain the best current lower bound for Sidon sets in randomly generated infinite sets of integers of high density.

One of the central results at the intersection of Combinatorics and Number Theory is the Freĭman–Ruzsa Theorem stating that any finite set of integers of given doubling can be efficiently covered by a generalized arithmetic progression. In the case of particularly small doubling, more precise structural descriptions exist. In Chapter 6 we will first study results going beyond Freĭman's well-known $3k-4$ Theorem in the integers. We will then see an application of these results to sets of small doubling in finite cyclic groups.

Lastly, we will turn our attention towards sets with near-constant representation functions. Interest in these functions is connected to Gauss' circle problem: Hardy and independently Landau established a lower bound on the difference between the number of integer lattice points in a circle and the area of that circle. Erdős and Fuchs obtained a slightly weaker bound for a much more general problem by showing that representation functions of arbitrary sets of integers cannot be too close to being constant. In Chapter 7 we will first extend the result of Erdős and Fuchs to ordered representation functions. We will then address a related question of Sárközy and Sós regarding weighted representation function.

Notation Most notation will be introduced at the appropriate parts of this thesis. However, we will use the usual big O and little o notation to state asymptotic results throughout. More specifically, for any two functions f and g, we will write $f = O(g)$ if there exists some constant $C > 0$ such that $f \leq Cg$, $f = \Omega(g)$ if $g = O(f)$ and $f = \Theta(g)$ if both $f = O(g)$ and $f = \Omega(g)$. We also write $f = o(g)$ if f/g tends to zero and $f = \omega(g)$ if $g = o(f)$. We will always write $\mathbb{N} = \{1, 2, 3, \ldots\}$ for the positive integers excluding zero and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ for those including zero. The set of the first n integers will be denoted by $[n] = \{1, \ldots, n\}$ and $[a, b] = \{a, a + 1, a + 2, \ldots, b\}$ will denote the discrete interval of all integers between a and b. The space of integer-valued matrices with r rows and m columns will be denoted by $\mathbb{Z}^{r \times m}$. We will use bold letter to denote row vectors and a multiplicative dot symbol to denote the matrix product, occasionally also used to denote the dilate of a set. Calligraphic letters will usually denote either a hypergraph or a subset of a finite cyclic group. $\mathbb{1}_S$ will refer to the indicator function of a given set S, that is $\mathbb{1}_S(x) = 1$ if $x \in S$ and $\mathbb{1}_S(x) = 0$ otherwise. Whenever a new definition or notion is introduced, it will be highlighted like this throughout the thesis. Proofs of major statements will end with a filled-in black square and proofs of minor statements in an outlined square.

Part I

Randomness and Games

Schur [139] proved one of the earliest results in the field that would later become known as Ramsey Theory in 1916 by establishing that for n large enough and any partition of the first n integers $[n] = \{1, \dots, n\}$ into at most c parts, at least one of the parts must contain a Schur triple, that is three elements $x, y, z \in [n]$ such that $x + y = z$. We will refer to such a partition as a c-coloring and will say that a triple that is contained in a single part is monochromatic.

Schur's Theorem. *For any $c \in \mathbb{N}$ there exists some $n_0 = n_0(c)$ such that any c-coloring of $[n_0]$ must contain a monochromatic Schur triple.*

Van der Waerden [157] proved an analogous result for k-term arithmetic progression, that is for sets of integers that can be written as $\{a, a + d, \dots, a + (k - 1)d\}$ for some $a \in \mathbb{Z}$ and $d \in \mathbb{Z} \setminus \{0\}$.

Van der Waerden's Theorem. *For any $k, c \in \mathbb{N}$ there exists some $n_0 = n_0(k, c)$ such that any c-coloring of $[n_0]$ must contain a monochromatic k-term arithmetic progression.*

Note that Schur triples can be considered as solutions to the linear homogeneous system of equations given by $A \cdot \mathbf{x}^T = \mathbf{0}^T$ where $A = (1 \quad 1 \ -1) \in \mathbb{Z}^{1\times 3}$ is an integer-valued matrix. Likewise, k-term arithmetic progression are solutions to such a system when

$$A = \begin{pmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & & \ddots & & \\ & & & 1 & -2 & 1 \end{pmatrix} \in \mathbb{Z}^{k-2\times k}.$$

This is the framework in which we will be interested throughout this chapter. Let us say that a solution $\mathbf{x} = (x_1, \dots, x_m)$ to $A \cdot \mathbf{x}^T = \mathbf{0}^T$ for some arbitrary $A \in \mathbb{Z}^{r\times m}$ is proper if its entries are pairwise distinct, that is $x_i \neq x_j$ if $i \neq j$. Using this, we state the following definition.

Definition. *An integer-valued matrix $A \in \mathbb{Z}^{r\times m}$ is* partition regular *if for any $c \in \mathbb{N}$ there exists some $n_0 = n_0(A, c)$ such that any c-coloring of $[n_0]$ must contain a monochromatic proper solution to the system $A \cdot \mathbf{x}^T = \mathbf{0}^T$.*

Rado [112] generalized both the result of Schur and that of van der Waerden by finding an exact characterization of matrices that are partition regular. In order to state it, let $\{\mathbf{c}_i \in \mathbb{Z}^r : 1 \leq i \leq m\}$ be the set of columns of A. We say that A satisfies the columns condition if there exists a partition of the column indices $1, \dots, m$ into sets $C_1, \dots, C_\ell$ such that for $\mathbf{s}_i = \sum_{j\in C_i} \mathbf{c}_j$ one has $\mathbf{s}_1 = \mathbf{0}$ and for any $2 \leq i_0 \leq \ell$ the vector $\mathbf{s}_{i_0}$ is a rational linear combination of the columns in $\{\mathbf{c}_j : j \in C_i, 1 \leq i < i_0\}$.

Rado's Theorem. *Any matrix $A \in \mathbb{Z}^{r \times m}$ is partition regular if and only if it satisfies the columns condition.*

A natural question to ask is if the property of being partition regular is merely a consequence of the fact that any c-coloring must contain a large color class of density at least $1/c$. For Schur triples this is easy to answer: clearly the set of odd integers in $[n]$ cannot contain such triples despite making up at least half of the elements of that set. For arithmetic progressions however, the answer is different. Already in 1936 Erdős and Turán conjectured [54] that any set of positive upper density in the integers must contain 3-term arithmetic progression. Roth proved this statement [117] and Szemerédi famously extended it to arithmetic progressions of arbitrary length [150].

Szemerédi's Theorem. *For any $k \in \mathbb{N}$ and $\varepsilon > 0$ there exists some $n_0 = n_0(k, \varepsilon)$ such that any subset of $[n_0]$ of size at least εn_0 must contain a k-term arithmetic progression.*

We therefore also state the following definition.

Definition. *An integer-valued matrix $A \in \mathbb{Z}^{r \times m}$ is density regular if for any $\varepsilon > 0$ there exists some $n_0 = n_0(A, \varepsilon)$ such that any subset of $[n_0]$ of size at least εn_0 must contain a proper solution to the system $A \cdot \mathbf{x}^T = \mathbf{0}^T$.*

Clearly any matrix that is density regular must also be partition regular. Frankl, Graham and Rödl [58] observed, as an easily obtained consequence of Szemerédis theorem, that a matrix $A \in \mathbb{Z}^{r \times m}$ is density regular if and only if it is invariant, that is $A \cdot \mathbf{1}^T = \mathbf{0}^T$.

Recently there has been a lot of interest in studying the 'common' behavior of subsets of combinatorial structures. In our particular case, that means we are interested in determining what combinatorial properties one can expect most subsets of the first n integers of a given size to satisfy: do they contain solutions to some given system? If so, do they satisfy a Rado- or even a Szemerédi-type property?

In order to study these types of questions, we will consider the binomial random set $[n]_p$ in which each element in $[n]$ is chosen independently with probability p. Clearly the expected size of $[n]_p$ is np and any particular set $S \subseteq [n]$ will be sampled by the binomial random set with probability $p^{|S|}(1-p)^{n-|S|}$. Here $p = p(n)$ will in fact grow with n and we will be interested in asymptotic statements as n tends to infinity. A given property will hold asymptotically almost surely if the probability of $[n]_p$ not satisfying it tends to zero as n tends to infinity.

Let us also introduce the notion of thresholds that will be central throughout this chapter. We will state it for the particular case of the binomial random set, though analogous definitions exist of course for any parameterized probability distribution over some combinatorial structure.

Definition. *Given some combinatorial property $\mathcal{P}$ of the subsets of $[n]$, we say that $p_0 = p_0(n)$ is a threshold for $\mathcal{P}$ if*

$$\lim_{n\to\infty} \mathbb{P}\left([n]_p \text{ satisfies } \mathcal{P}\right) = \begin{cases} 1 & \text{if } p = \omega(p_0), \\ 0 & \text{if } p = o(p_0). \end{cases}$$

We say that it is sharp if there in fact exist constants $c, C > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\left([n]_p \text{ satisfies } \mathcal{P}\right) = \begin{cases} 1 & \text{if } p \geq C\, p_0, \\ 0 & \text{if } p \leq c\, p_0. \end{cases}$$

Roughly speaking, when p is 'above the threshold' almost all sets satisfy the property $\mathcal{P}$ and 'below the threshold' almost none of them do. We will refer to the former as the 1-statement and the later as the 0-statement. Note that monotone properties always have thresholds in the binomial random set [14].

The binomial random set is the natural analogue to the binomial random graph $G(n,p)$ on n vertices where each possible edge gets chosen independently with probability p [47]. As is the case there, results obtained for the binomial random set can usually be transferred to the setting given by the uniform random set $[n]_M$ obtained by simply taking the uniform distribution over all subsets of $[n]$ of cardinality M. This transfer will be made explicit in some cases, but in general one may expect to get the corresponding statement in the uniform setting by simply setting M equal to the number of expected elements in $[n]_p$, that is $M = np$.

In Chapter 1 we will establish some necessary preliminaries concerning the general framework. This means that we will first introduce the notion of abundant matrices, an extension of the notions of partition and density regularity. We will then establish results concerning the enumeration of solutions to some given system in the integers. This will in fact be done not just for the homogeneous case presented in this introduction, but also for the inhomogeneous case of $A \cdot \mathbf{x}^T = \mathbf{b}^T$ for some $\mathbf{b} \neq \mathbf{0}$. We will furthermore establish a notion of symmetry between solutions as well as a notion of non-trivial solutions, that is not necessarily proper solutions that can have repeated entries. We will also introduce two notions of density for integer-valued matrices that parallel the notions of density and 2-density of graphs. Lastly, we will state some well-established results from Probability Theory and then conclude the chapter by surveying some results relating to the extremal question of solution-free sets.

In Chapter 2 we will establish the threshold for the property that the binomial random set contains a solution to some arbitrary given system. This threshold will not be sharp and we will in fact show that in particular cases the number of solutions at the thresholds follows a Poisson distribution whose parameter depends on the volumes of certain convex polytopes associated with the system as well as its symmetry. These results of this chapter will be based on the paper 'Threshold functions and Poisson convergence for systems of equations in random sets', which is joint work with Juanjo Rué and Ana Zumalacárregui [122].

In Chapter 3 we turn our attention towards thresholds for the property of satisfying a Rado- or Szemerédi-type statement. We will extend previous results of Friedgut, Rödl, Ruciński and Schacht concerning the threshold for a Rado-type property to include non-trivial solutions. The 1-statement of that proof will be based on a hypergraph container approach due to Nenadov and Steger, resulting in a much shorter proof. We will also extend results of Conlon, Gowers and Schacht concerning a Szemerédi-type property for density regular matrices to the broadest class of matrices possible, that of abundant matrices, again including non-trivial solutions. The results of this chapter will be based on the paper 'A Note on Sparse Supersaturation and Extremal Results for Linear Homogeneous Systems' [144].

Lastly, in Chapter 4 we have a change of pace and play some games: the notion of biased Maker-Breaker games was introduced by Chvátal and Erdős and has since become central to the field of positional games. As will become apparent in this chapter, they have deep connections to the theory of random structures. The main questions is to determine the smallest bias needed by Breaker to ensure that Maker ends up with an independent set in a given hypergraph. Building on a result of Bednarska and Łuczak concerning H-building games, we prove matching general winning criteria for Maker and Breaker when the game hypergraph satisfies certain 'container-type' regularity conditions. This allows one to study hypergraph generalizations of the H-building games as well as a generalization of the van der Waerden games introduced by Beck [6]. These results will be based on the paper 'On the optimality of the uniform random strategy', which is joint work with Christopher Kusch, Juanjo Rué and Tibor Szabó [101].

Chapter 1

Preliminaries for Linear Systems

For a given integer-valued matrix $A \in \mathbb{Z}^{r \times m}$ and a given integer-valued vector $\mathbf{b} \in \mathbb{Z}^r$, let us denote the set of all integer-valued solutions to the $A \cdot \mathbf{x}^T = \mathbf{b}^T$ by

$$S(A, \mathbf{b}) = \{\mathbf{x} \in \mathbb{Z}^m : A \cdot \mathbf{x}^T = \mathbf{b}^T\} \tag{1.1}$$

and the set of all proper solutions by

$$S_0(A, \mathbf{b}) = \{\mathbf{x} = (x_1, \ldots, x_m) \in S(A, \mathbf{b}) : x_i \neq x_j \text{ for } i \neq j\}. \tag{1.2}$$

Let us make a trivial observation: for any invertible matrix $P \in \mathbb{Z}^{r \times r}$, we clearly have $S(P \cdot A, P \cdot \mathbf{b}) = S(A, \mathbf{b})$ as well as $S_0(P \cdot A, P \cdot \mathbf{b}) = S_0(A, \mathbf{b})$. We may therefore freely apply elementary row operations, that is multiplying a row by a non-zero constant, adding a row to another row and switching two rows, to any given matrix without affecting the solution space. In general, none of the properties that we will introduce in this chapter will be affected by such operations. When it comes to elementary column operations, this of course no longer holds. We will however occasionally assume, without loss of generality, that the columns are ordered in an opportune way.

1.1 The notion of abundant matrices

We have already introduced the notion of partition and density regular matrices in the introduction of this chapter. Let us now give an even broader notion that, as we will later show, encompasses both of the previous ones and in some sense describes the largest class of matrices creating a solution space without 'degenerate dependencies'.

In order to state it, let A^Q denote the matrix obtained from A by keeping only the columns indexed by $Q \subseteq \{1, \ldots, m\}$. Here $A^\emptyset$ is the empty matrix of rank 0. We also denote the rank of A as $\mathrm{rk}(A)$.

Definition 1.1. *An integer-valued matrix $A \in \mathbb{Z}^{r \times m}$ is*

(i) positive if there are solutions whose entries lie in the positive integers, that is

$$S(A, \mathbf{0}) \cap \mathbb{N}^m \neq \emptyset$$

and for any two distinct column indices $i_1, i_2 \in \{1, \ldots, m\}$ there exists a solution $\mathbf{x} = (x_1, \ldots, x_m) \in S(A, \mathbf{0})$ satisfying $x_{i_1} \neq x_{i_2}$,

(ii) abundant if A has rank strictly greater than 0 and every matrix obtained from A by deleting at most two columns must be of the same rank as A, that is

$$\mathrm{rk}(A^Q) = \mathrm{rk}(A) > 0$$

for all $Q \subseteq \{1, \ldots, m\}$ satisfying $|Q| \geq m - 2$.

Note that, in order for a matrix to be abundant, we must have $m \geq 3$. In fact, the number of columns of an abundant matrix must always be at least 2 more than its rank. In order for a matrix to be non-abundant, it must have non-zero rank.

The importance of the notion of positivity is easy to justify: firstly, if there are two distinct column indices such that any solution must have the same entry at those two indices, then one should instead consider the matrix obtained by contracting the two corresponding columns. Note that this condition is sometimes stated separately from positivity in the literature and referred to as irredundancy. Here we will consider it as part of positivity to keep the list of matrix properties small.

Regarding the requirement that $S(A, \mathbf{0}) \cap \mathbb{N}^m \neq \emptyset$: if the homogeneous solution space is disjoint from the positive quadrant in which we are interested, then either it is contained in a subspace of $\mathbb{Z}^m$ or the number of solutions in $[n]$ to the inhomogeneous system will be finitely bounded for each $\mathbf{b}$ independent of n. The later case is clearly not of interest when we let n go to infinity. In the former case there must exist at least one $i \in \{1, \ldots, m\}$ such that $x_i = b_i$ for any $\mathbf{x} = (x_1, \ldots, x_m) \in S(A, \mathbf{b})$, so that the behavior of any random process or game is disproportionately determined by where b_i lands.

The second notion, that of abundancy might initially be somewhat obscure, but we will see in the following sections that it is of great importance. Non-abundant systems will have certain trivial behavior with regards to the properties studied here. The following simple lemma establishes a central aspect of this.

Lemma 1.2. *For any positive and non-abundant matrix $A \in \mathbb{Z}^{r \times m}$, there exists some invertible matrix $P \in \mathbb{Z}^{r \times r}$ and distinct column indices $i_1, i_2 \in \{1, \ldots, m\}$ such that the matrix $P \cdot A$ contains a row $\mathbf{a} = (a_1, \ldots, a_m)$ satisfying $a_i = 0$ for all $i \in \{1, \ldots, m\} \setminus \{i_1, i_2\}$ as well as $a_{i_1}, a_{i_2} \neq 0$ and $a_{i_1} + a_{i_2} \neq 0$.*

Proof. From the non-abundancy, it follows that there exist two column indices $1 \leq i_1, i_2 \leq m$ such that for $Q = \{1, \ldots, m\} \setminus \{i_1, i_2\}$ we get $\mathrm{rk}(A^Q) < \mathrm{rk}(A)$. It follows that there exist a set of basic row transformations and a row $\mathbf{a} = (a_1, \ldots, a_m)$ in the transformed matrix such that $\mathbf{a}^Q$ consists only of 0 entries and $\mathbf{a}$ can be taken as a basis vector of the space spanned by the rows of A. As $\mathrm{rk}(A^Q) < \mathrm{rk}(A)$, we cannot have $\mathbf{a} = \mathbf{0}$. If, without loss of generality, $a_{i_2} = 0$, then any solution $\mathbf{x} = (x_1, \ldots, x_m) \in S(A, \mathbf{b})$ satisfies $a_{i_1} x_{i_1} = 0$, that is $x_{i_1} = 0$ since $a_{i_1} \neq 0$, contradicting the assumption that A is positive. If $a_{i_1} + a_{i_2} = 0$ then any solution would satisfy $a_{i_1}(x_{i_1} - x_{i_2}) = 0$, that is in particular $x_{i_1} = x_{i_2}$ since $a_{i_1}, a_{i_2} \neq 0$, again contradicting the assumption that A is positive. □

Partition and density regular matrices are positive by definition. The next statement establishes that they are also abundant.

Lemma 1.3. *Any partition regular matrix $A \in \mathbb{Z}^{r \times m}$ is abundant.*

Proof. Assume that A is non-abundant. By Lemma 1.2, A can be transformed into a matrix whose last row consists only of two non-zero entries a_{i_1}, a_{i_1} satisfying $a_{i_{i_1}}, a_{i_{i_2}} \neq 0$ and $a_{i_{i_1}} + a_{i_{i_2}} \neq 0$. We will without loss of generality assume that $\{i_1, i_2\} = \{1, 2\}$. By Rado's Theorem, A must fulfill the columns condition, so let $\{\mathbf{c}_i \in \mathbb{Z}^r : 1 \leq i \leq m\}$ be the set of columns of A, $C_1, \ldots, C_\ell$ the partition of the column indices $1, \ldots, m$ such that for $\mathbf{s}_i = \sum_{j \in C_i} \mathbf{c}_j$ one has $\mathbf{s}_1 = \mathbf{0}$ and for any $2 \leq i_0 \leq \ell$ the vector $\mathbf{s}_{i_0}$ is a rational linear combination of the columns in $\{\mathbf{c}_j : j \in C_i, 1 \leq i < i_0\}$. Since $\mathbf{s}_1 = \mathbf{0}$, we may assume that the columns are arranged such that the last entry of the first column is zero. However, since we established that $a_1, a_2 \neq 0$, $a_1 + a_2 \neq 0$ and that all other entries of that row are zero, there now must exist some $2 \leq i_0 \leq \ell$ such that the last entry in $\mathbf{s}_{i_0}$ is non-zero while the last entries in $\mathbf{s}_1, \ldots, \mathbf{s}_{i_0 - 1}$ are zero, violating the requirement that $\mathbf{s}_{i_0}$ is a linear combination of the previous columns. It follows that A cannot have been partition regular. □

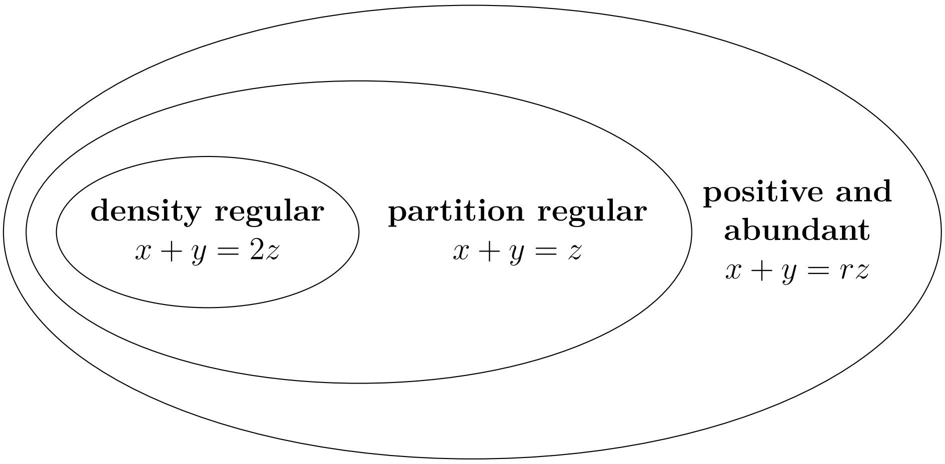


Figure 1.1: *Illustrating the relation of the notions of density regularity, partition regularity, positivity and abundancy for integer-valued matrices.*

Let us consider some examples to illustrate these three categories, see Figure 1.1: $A = (1 \quad 1 \; -2)$, that is the matrix associated with 3-term arithmetic progressions, is density regular by Roth's Theorem. It therefore is also partition regular, which was previously established by van der Waerden. $A = (1 \quad 1 \; -1)$, that is the matrix associated with Schur triples, is not density regular, but by Schur's Theorem it is still partition regular. Lastly, $A = (1 \quad 1 \; -r)$, that is the matrix associated with r-sums, is neither partition nor density regular when $r \geq 3$, but it is clearly still abundant.

1.2 Counting solutions in the inhomogeneous case

We will establish two basic bounds for the number of proper solutions. Given any matrix $A \in \mathbb{Z}^{r \times m}$ and vector $\mathbf{b} \in \mathbb{Z}^r$, we have the upper bound

$$\left| S_0(A, \mathbf{b}) \cap [n]^m \right| \leq \left| S(A, \mathbf{b}) \cap [n]^m \right| \leq n^{m - \mathrm{rk}(A)}. \tag{1.3}$$

Indeed, taking a subset $Q \subseteq \{1, \ldots, m\}$ of the column indices with $\mathrm{rk}(A) = |Q| = \mathrm{rk}(A^Q)$ and setting the $m - \mathrm{rk}(A)$ entries in $\overline{Q}$ of a solution $\mathbf{x} \in S(A, \mathbf{b})$ arbitrarily, the entries in Q are determined uniquely.

The following lemma, the proof of which is based on a construction by Janson and Ruciński [88], establishes that Equation (1.3) is tight up to a constant factor for positive matrices.

Lemma 1.4. *For every positive matrix $A \in \mathbb{Z}^{r \times m}$ and vector $\mathbf{b} \in \mathbb{Z}^r$ satisfying $S(A, \mathbf{b}) \neq \emptyset$ there exist constants $c_0 = c_0(A, \mathbf{b}) > 0$ and $n_0 = n_0(A, \mathbf{b}) \in \mathbb{N}$ such that for every $n \geq n_0$*

$$|\, S_0(A, \mathbf{b}) \cap [n]^m| \geq c_0\, n^{m - \mathrm{rk}(A)}. \tag{1.4}$$

Proof. We need to construct many proper solutions to $A \cdot \mathbf{x}^T = \mathbf{b}^T$ in $[n]$. First we prove that there exists at least one to the homogenous system. Since A is positive, we can take a solution $\mathbf{x}^* = (x_1^*, \ldots, x_m^*) \in S(A, \mathbf{0}) \cap \mathbb{N}^m$ that minimizes the number of pairs of equal entries. We claim that $\mathbf{x}^*$ is proper.

Assume to the contrary that there are two column indices, without loss of generality 1 and 2, such that the corresponding entries of $\mathbf{x}^*$ are equal, that is $x_1^* = x_2^*$. If there was a solution $\mathbf{y} = (y_1, \ldots, y_m) \in S(A, \mathbf{0}) \cap \mathbb{N}^m$ with $y_1 \neq y_2$, then consider

$$\mathbf{w} = \mathbf{x}^* + \alpha\, \mathbf{y} \in S(A, \mathbf{0}) \cap \mathbb{N}^m,$$

where $\alpha \in \mathbb{N} \setminus \{(x_r^* - x_s^*)/(y_s - y_r) : 1 \leq r, s \leq m, y_s \neq y_r\}$ is chosen arbitrarily. By the definition of α, for all r, s with $x_r^* \neq x_s^*$ we also have $w_r \neq w_s$, furthermore $w_1 \neq w_2$. Consequently $\mathbf{w}$ has less pairs of equal entries than $\mathbf{x}^*$, contradicting the choice of $\mathbf{x}^*$. It follows that if $\mathbf{x}^*$ is not proper, then for every positive solution $\mathbf{y} \in S(A, \mathbf{0}) \cap \mathbb{N}^m$ we must have $y_1 = y_2$. This contradicts the assumption that A is positive, since there must exist some vector $\mathbf{z} = (z_1, \ldots, z_m) \in S(A, \mathbf{0})$ satisfying $z_1 \neq z_2$. We may add some positive $\mathbf{y}$ a sufficient number of times to $\mathbf{z}$ to obtain a vector $\mathbf{x}$ in $\mathbb{N}^m$ satisfying $x_1 \neq x_2$, a contradiction.

It follows that there is a proper solution $\mathbf{x}^* \in S_0(A, \mathbf{0}) \cap \mathbb{N}^m$. To construct many proper positive solutions to the inhomogeneous system $A \cdot \mathbf{x}^T = \mathbf{b}^T$ in $[n]$, we choose a solution $\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_m) \in S(A, \mathbf{b})$ as well as $m - \mathrm{rk}(A)$ linearly independent solutions $\mathbf{x}_1, \ldots \mathbf{x}_{m-\mathrm{rk}(A)} \in S(A, \mathbf{0})$. Let s^* and $\hat{s}$ be the maximum absolute value of the entries of $\mathbf{x}^*$ and $\hat{\mathbf{x}}$, respectively, and s the maximum absolute value of any entry in any of the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{m-\mathrm{rk}(A)}$. Define $a(n) = \lfloor n/(s^* + 1) \rfloor$ and set

$$S_n = \left\{ \hat{\mathbf{x}} + a(n)\, \mathbf{x}^* + \sum_{i=1}^{m-\mathrm{rk}(A)} w_i\, \mathbf{x}_i : w_i \in \mathbb{Z},\ |w_i| < \frac{a(n) - 2\hat{s}}{2s(m - \mathrm{rk}(A))} \right\} \subseteq \mathbb{Z}^m.$$

Since $A \cdot \hat{\mathbf{x}}^T = \mathbf{b}^T$ and $A \cdot \mathbf{x}^{*T} = A \cdot \mathbf{x}_i^T = \mathbf{0}^T$, we have $S_n \subseteq S(A, \mathbf{b})$. Let $\mathbf{x} =$

$(x_1, \ldots, x_m) \in S_n$ and observe that for n large enough

$$x_i > \hat{x}_i + a(n)\, x_i^* - (m - \mathrm{rk}(A))\, s\, \frac{a(n) - 2\hat{s}}{2s(m - \mathrm{rk}(A))} \geq -\hat{s} + a(n) - \frac{a(n)}{2} + \hat{s} \geq 1$$

as well as

$$\begin{aligned} x_i &< \hat{x}_i + a(n)\, x_i^* + (m - \mathrm{rk}(A))\, s\, \frac{a(n) - 2\hat{s}}{2s(m - \mathrm{rk}(A))} \\ &\leq \hat{s} + n\, \frac{s^*}{s^* + 1} + n\, \frac{1}{2(s^* + 1)} - \hat{s} \leq n \end{aligned}$$

for every $i \in \{1, \ldots, m\}$. Consequently $S_n \subseteq [n]^m$ for n large enough.

Now assume without loss of generality that $x_1^* < \cdots < x_m^*$. It follows that

$$\begin{aligned} x_i &< \hat{x}_i + a(n)\, x_i^* + (m - \mathrm{rk}(A))\, s\, \frac{a(n) - 2\hat{s}}{2s(m - \mathrm{rk}(A))} \\ &= (\hat{x}_i - \hat{s}) + a(n) \left(x_i^* + \frac{1}{2} \right) \leq (\hat{x}_{i+1} + \hat{s}) + a(n) \left(x_{i+1}^* - \frac{1}{2} \right) \\ &= \hat{x}_{i+1} + a(n)\, x_{i+1}^* - (m - \mathrm{rk}(A))\, s\, \frac{a(n) - 2\hat{s}}{2s(m - \mathrm{rk}(A))} < x_{i+1} \end{aligned}$$

for every $1 \leq i \leq m - 1$, so $\mathbf{x}$ is proper. Therefore $S_n \subseteq S_0(A, \mathbf{b}) \cap [n]^m$.

Lastly observe that since $\mathbf{x}_1, \ldots \mathbf{x}_{m-\mathrm{rk}(A)}$ are linearly independent, S_n contains

$$\left(2 \left\lfloor \frac{a(n) - 2\hat{s}}{2s(m - \mathrm{rk}(A))} \right\rfloor + 1 \right)^{m-\mathrm{rk}(A)} \geq \left(\frac{1/(4s^* + 4)}{s\big(m - \mathrm{rk}(A)\big)}\, n \right)^{m-\mathrm{rk}(A)}$$

elements, where the lower bound holds for n large enough. It follows that for

$$c_0 = c_0(A, \mathbf{b}) = \left(\frac{1/(4\hat{s}_{\mathbf{0}} + 4)}{s\big(m - \mathrm{rk}(A)\big)} \right)^{m-\mathrm{rk}(A)} < 1$$

and $n_0 = \lceil 4\, \hat{s}\, (\hat{s}_{\mathbf{0}} + 1) \rceil$ we have $|\, S_0(A, \mathbf{b}) \cap [n]^m| \geq c_0\, n^{m-\mathrm{rk}(A)}$ for all $n \geq n_0$. □

1.3 Counting solutions precisely in the homogeneous case

For most of the applications in this chapter, the fact that

$$|\, S_0(A, \mathbf{b}) \cap [n]^m| \, = \Theta(n^{m-\mathrm{rk}(A)}), \tag{1.5}$$

as established by Equation (1.3) and Lemma 1.4 in the previous section, will be sufficient. In Chapter 2 however, we will be interested in determining the precise limit distribution of the number of solutions in the binomial random set for certain values of p. Here, we will restrict ourselves to the homogeneous case, that is $\mathbf{b} = \mathbf{0}$. We face two challenges: first, we need to determine the exact leading coefficient, that is we need to know the limit of $| \, S_0(A, \mathbf{b}) \cap [n]^m|/n^{m-\mathrm{rk}(A)}$ as n tends to infinity. Second, we need to establish a notion of when two solutions that are the same up to permutation of the vector entries are in fact 'the same' solution. We start with the former of the two.

In the homogeneous case, one can determine the precise leading coefficient using Ehrhart's Theory. We will introduce the necessary essentials of this theory in order to state our result, but a good basic reference for the theory of convex polytopes is [161]. For further results on lattice points in rational polytopes, see also [9, 38].

A convex polytope is the convex hull of a finite set of points in the d-dimensional space $\mathbb{R}^d$. It can likewise be described as the bounded intersection of a finite set of half-spaces. We say that a convex polytope is rational (respectively integral) if its corner points have rational (respectively integer) coordinates. Every rational polytope has a matrix representation of the form

$$\{\mathbf{x} \in \mathbb{R}^d : M \cdot \mathbf{x}^T \geq \mathbf{b}^T\}, \; M \in \mathbb{Z}^{m \times d}, \; \mathbf{b} \in \mathbb{Z}^d \tag{1.6}$$

for some $m \in \mathbb{N}$. Note that the inequalities can be easily turned into equalities through the use of slack variables. The relative dimension $\dim(\mathcal{P})$ of a polytope $\mathcal{P}$ is the dimension of the affine space

$$\operatorname{span} \mathcal{P} = \{\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) \; : \; \mathbf{x}, \mathbf{y} \in \mathcal{P}, \; \lambda \in \mathbb{R}\} \, . \tag{1.7}$$

Note that this dimension is not necessarily the same as d, but can in fact be smaller. We let $\mathrm{Vol}\,(\mathcal{P})$ be the volume of $\mathcal{P}$ as embedded in this affine space and $n \cdot \mathcal{P} = \{n\mathbf{p} : \mathbf{p} \in \mathcal{P}\}$ the n-th dilate of the polytope.

Ehrhart's Theorem [44], see also [106], gives a precise description of the number of integer points on the n-th dilate of a rational polytope: the quantity $|n \cdot \mathcal{P} \cap \mathbb{Z}^n|$ is given by a pseudo-polynomial in n of degree $\dim(\mathcal{P})$. Here a pseudo-polynomial is a function $p(n) = c_0(n) + c_1(n)\, n + \ldots c_t(n)\, n^t$ where the functions $c_0(n), c_1(n), \ldots, c_t(n)$ are periodic.

Ehrhart's Theorem. *Let $\mathcal{P}$ be a d-dimensional convex polytope. If $\mathcal{P}$ is integral, then*

$\left|n \cdot \mathcal{P} \cap \mathbb{Z}^d\right|$ *is a polynomial in* n *of degree dim*$(\mathcal{P})$*. If* $\mathcal{P}$ *is rational, then* $\left|n \cdot \mathcal{P} \cap \mathbb{Z}^d\right|$ *is a pseudo-polynomial in* n *of degree* $\dim(\mathcal{P})$*.*

One can show that the leading coefficient in both cases is equal to $\operatorname{Vol}(\mathcal{P})$. As an immediate corollary, for a rational polytope $\mathcal{P}$ of dimension $\dim(\mathcal{P})$ embedded in $\mathbb{R}^{\dim(\mathcal{P})}$, we have

$$\left|n \cdot \mathcal{P} \cap \mathbb{Z}^{\dim(\mathcal{P})}\right| = \operatorname{Vol}(\mathcal{P})\, n^{\dim(\mathcal{P})}(1 + o(1)). \tag{1.8}$$

Observe that for any positive integer-valued matrix $A \in \mathbb{Z}^{r \times m}$, the system of equations $A \cdot \mathbf{x}^T = \mathbf{0}^T$ in $\mathbb{R}^n$ together with the restrictions that $\mathbf{x} \in [0,1]^m$ defines a non-empty, convex and rational polytope of relative dimension $m - \operatorname{rk}(A)$. Note that it is just the intersection of the $m - \operatorname{rk}(A)$-dimensional solution space with the m-dimensional unit hypercube in $\mathbb{R}^n$. We will denote this polytope by $\mathcal{P}_A$.

Lemma 1.5. *For any positive matrix* $A \in \mathbb{Z}^{r \times m}$ *we have*

$$\left|S_0(A, \mathbf{0}) \cap [n]^m\right| = \operatorname{Vol}(\mathcal{P}_A)\, n^{m-\operatorname{rk}(A)}(1 + o(1)). \tag{1.9}$$

Proof. The number of lattice points in $n \cdot \mathcal{P}_A$ is precisely the number of (not necessarily proper) solutions $\mathbf{x} \in [n]^m$ to $A \cdot \mathbf{x}^T = \mathbf{0}^T$. As the intersection of the $m - \operatorname{rk}(A)$-dimensional solution space and the m-dimensional unit hypercube, the polytope $\mathcal{P}_A$ has relative dimension $m - \operatorname{rk}(A)$. By Equation (1.8) the number of lattice points in the dilate $n \cdot \mathcal{P}_A$ is simply $\operatorname{Vol}(\mathcal{P}_A)\, n^{m-\operatorname{rk}(A)}(1 + o(1))$.

Let us to therefore consider the set of solutions $\mathbf{x} \in [n]^m$ to $A \cdot \mathbf{x}^T = \mathbf{0}^T$ with some repeated coordinates and show that they have a negligible contribution to the total number of solutions. These solutions belong to the intersection of $\mathcal{P}_A$ with a subspace defined by the repetitions of coordinates. Since $\mathcal{A}$ is positive and abundant, Lemma 1.4 tells us that there exists at least one solution in the integers with no repeated coordinates and therefore also a point in $\mathcal{P}_A$ with no repeated entries; this implies that there is no subspace defined by the repetition of coordinates containing all of $\mathcal{P}_A$. Therefore, the polytope resulting from the intersection has dimension strictly smaller than $m - \operatorname{rk}(A)$. The number of solutions with those particular repeated coordinates is therefore $O(n^{m-\operatorname{rk}(A)-1}) = o(n^{m-\operatorname{rk}(A)})$. As the number of possible repeated coordinates is bounded by the number of partitions of $\{1, \ldots, m\}$, the total number of solutions with repeated coordinates is $o(n^{m-\operatorname{rk}(A)})$ and the statement follows. □

Let us now turn our attention to the fact that two proper solutions which are counted as separate by Ehrhart's Theory can be essentially the same when considering symmetry. As an easy example for this, consider that 3-term arithmetic progressions are given by $x_1 + x_3 = 2x_2$ for which $(1, 2, 3)$ and $(3, 2, 1)$ both are proper solutions that one might consider as essentially identical. However the situation is not quite as simple as grouping solutions together if they are identical up to permutation. Consider for example the system given by

$$x_1 + x_2 + x_3 = x_4 + x_5 + x_6 + x_7 \tag{1.10}$$

for which both $(2, 3, 100; 1, 4, 49, 51)$ and $(1, 4, 100; 2, 3, 49, 51)$ are again proper solutions. In this case, one should reasonably consider them to be distinct because the permutation did not just occur between coordinates with identical coefficients and hence cannot be applied to all solutions. The semicolons in the vector representations delineating the coordinates with factor -1 from those with factor 1 were added to emphasize that fact.

In order to deal with this distinction, let $\mathrm{Sym(m)}$ denote the symmetric group on m elements and $\mathbf{x}_\pi$ the vector obtained by permuting the coordinates of $\mathbf{x} \in \mathbb{Z}^m$ according to $\pi \in \mathrm{Sym(m)}$. We will use the following definition when deciding if two solutions are distinct. Note that it correctly applies to both of the previously stated examples.

Definition 1.6. *The* set of symmetries *of a matrix $A \in \mathbb{Z}^{r\times m}$ is defined as*

$$\Sigma(A) = \{\pi \in \mathrm{Sym(m)} : \mathbf{x} \in \mathrm{S(A, \mathbf{0})} \textit{ if and only if } \mathbf{x}_\pi \in \mathrm{S(A, \mathbf{0})}\} \tag{1.11}$$

and the cardinality of this set is its symmetry constant *$\sigma(A)$. We furthermore say that two solutions $\mathbf{x}, \mathbf{y} \in S(A, \mathbf{0})$ are* distinct *if and only if $\mathbf{x} \neq \mathbf{y}_\pi$ for all $\pi \in \Sigma(A)$.*

The following establishes an easy characterization of the symmetry constant relying solely on the matrix A without requiring an understand of its solution space. In order to state it, let A^π denote the matrix obtained by permuting the columns of a matrix $A \in \mathbb{Z}^{r\times m}$ according to some $\pi \in \mathrm{Sym(m)}$.

Lemma 1.7. *For any matrix $A \in \mathbb{Z}^{r\times m}$ we have*

$$\Sigma(A) = \left\{\pi \in \mathrm{Sym(m)} : \exists \textit{ invertible } \mathrm{P} \in \mathbb{Z}^{\mathrm{r\times r}} \textit{ s.t. } \mathrm{A}^\pi = \mathrm{P} \cdot \mathrm{A}\right\}. \tag{1.12}$$

Proof. The inequality $\Sigma(A) \supseteq \{\pi \in \mathrm{Sym(m)} : \mathrm{A}^\pi = \mathrm{P} \cdot \mathrm{A}\}$ trivially holds. In order to show equality, note that for every permutation $\pi \in \Sigma(A)$ we have $\ker(A) \subseteq \ker(A^{\pi^{-1}})$ and since both kernels have dimension $m - \mathrm{rk}(A)$ we have equality. The kernel of a matrix is the orthogonal of the span of its rows and therefore the rows of $A^{\pi^{-1}}$ can be obtained by linear transformation of the rows in A. □

We now have the following statement as an immediate consequence of Lemma 1.5 and the definition of the symmetry constant.

Corollary 1.8. *For any positive matrix $A \in \mathbb{Z}^{r\times m}$ there are*

$$\frac{\mathrm{Vol}\left(\mathcal{P}_A\right)}{\sigma(A)}\left(1+o(1)\right) n^{m-\mathrm{rk}(A)} \tag{1.13}$$

distinct proper solutions to $A \cdot \mathbf{x}^T = \mathbf{0}^T$ with entries in $[n]$.

1.4 Counting non-proper solutions

For singe-line equations, that is when $r = 1$, it is common to not just limit oneself to proper solutions, but to also consider solutions which may have repeated entries. Here one needs to be precise what type of solutions are allowed: in the case of arithmetic progressions for example, the only non-proper solutions are those which are constant, that is $(1 \quad 1 \; -2) \cdot (c, c, c)^T = 0$ for any $c \in \mathbb{N}$. Clearly those should not be permitted. However, in the case of the obstruction that defines Sidon sets, see Section 1.7, 3-term arithmetic progressions constitute solutions that are commonly considered to be valid, for example $(1 \quad 1 \; -1 \; -1) \cdot (1, 3, 2, 2)^T = 0$. On the other hand, $(1 \quad 1 \; -1 \; -1) \cdot (c_1, c_2, c_1, c_2)^T = 0$ clearly holds for any $c_1, c_2 \in \mathbb{N}$ and such solutions should reasonably be omitted.

Ruzsa [124] gave a formal definition for single line equations that covers these examples. in this chapter we will use a natural extension of this notion for arbitrary r. Given a vector $\mathbf{x} \in \mathbb{Z}^m$, let

$$\mathfrak{p}(\mathbf{x}) = \Big\{\{1 \le j \le m : x_i = x_j\} : 1 \le i \le m\Big\}$$

denote the set partition of the column indices $\{1, \ldots, m\}$ indicating the repeated entries in $\mathbf{x}$. Note that for $\mathbf{x} \in S_0(A, \mathbf{b})$ we have $\mathfrak{p}(\mathbf{x}) = \{\{1\}, \ldots, \{m\}\}$. Given some set partition $\mathfrak{p}$ of $\{1, \ldots, m\}$, let $A_{\mathfrak{p}}$ denote the matrix obtained by summing up the columns of A according to $\mathfrak{p}$, that is for $\mathfrak{p} = \{T_1, \ldots, T_s\}$ such that $\min(T_1) < \cdots < \min(T_s)$ for some $1 \leq s \leq m$ and $\mathbf{c}_i$ the i-th column vector of A for every $1 \leq i \leq m$, we have

$$A_{\mathfrak{p}} = \left(\sum_{i \in T_1} \mathbf{c}_i \;\middle|\; \sum_{i \in T_2} \mathbf{c}_i \;\middle|\; \cdots \;\middle|\; \sum_{i \in T_s} \mathbf{c}_i \right).$$

Note that the assumption $\min(T_1) < \cdots < \min(T_s)$ ensures that this notion is well-defined and that $A_{\mathfrak{p}} = A$ for $\mathfrak{p} = \{\{1\}, \ldots, \{m\}\}$. Using these definitions we can now define when a solution is considered to be non-trivial.

Definition 1.9. *A solution $\mathbf{x} \in S(A, \mathbf{b})$ is non-trivial if $\mathrm{rk}(A_{\mathfrak{p}(\mathbf{x})}) = \mathrm{rk}(A)$.*

We denote the set of all non-trivial solutions by

$$S_1(A, \mathbf{b}) = \{\mathbf{x} \in S(A, \mathbf{b}) : \mathrm{rk}(A_{\mathfrak{p}(\mathbf{x})}) = \mathrm{rk}(A)\}, \tag{1.14}$$

that is we have $S(A, \mathbf{b}) \supseteq S_1(A, \mathbf{b}) \supseteq S_0(A, \mathbf{b})$. Other notions of what constitutes a non-trivial solution are possible and make sense in particular contexts, but this covers the widest range of applications while still extending the notion of Ruzsa and is therefore the one we will use throughout this part of the thesis.

Now let

$$\mathfrak{P}(A) = \{\mathfrak{p} : \mathrm{rk}(A_{\mathfrak{p}}) = \mathrm{rk}(A)\} \tag{1.15}$$

denote the family of all set partitions of the column indices $\{1, \ldots, m\}$ that could stem from non-trivial solution. The following lemma gives us the necessary tool to handle non-trivial solutions with repeated entries.

Lemma 1.10. *For every $A \in \mathbb{Z}^{r \times m}$, $\mathbf{b} \in \mathbb{Z}^r$, $\mathfrak{p} \in \mathfrak{P}(A)$ and $T \subset \mathbb{N}$ we have*

$$\Big|\{\mathbf{x} \in S_1(A, \mathbf{b}) \cap T^m : \mathfrak{p}(\mathbf{x}) = \mathfrak{p}\}\Big| \leq \Big|S_0(A_{\mathfrak{p}}, \mathbf{b}) \cap T^{|\mathfrak{p}|}\Big|. \tag{1.16}$$

Proof. Write $\mathfrak{p} = \{T_1, \ldots, T_s\}$ for some $1 \leq s \leq m$ such that $\min(T_1) < \cdots < \min(T_s)$. Let $Q = \{\min(T_1), \ldots, \min(T_s)\}$. Now for every $\mathbf{x} = (x_1, \ldots, x_m) \in S_1(A, \mathbf{b}) \cap T^m$ such that $\mathfrak{p}(\mathbf{x}) = \mathfrak{p}$, we would have $\mathbf{x}^Q = (x_{\min(T_1)}, \ldots, x_{\min(T_S)}) \in T^{|\mathfrak{p}|}$ as well as $\mathbf{x}^Q \in S(A_{\mathfrak{p}}, \mathbf{b})$ as can be readily seen by the definition of $A_{\mathfrak{p}}$. Since $\mathfrak{p} = \mathfrak{p}(\mathbf{x})$, the vector $\mathbf{x}^Q$ would furthermore be proper, so that $\mathbf{x}^Q \in S_0(A_{\mathfrak{p}}, \mathbf{b}) \cap T^{|\mathfrak{p}|}$. The map taking $\mathbf{x} \in S_1(A, \mathbf{b}) \cap T^m$ with $\mathfrak{p}(\mathbf{x}) = \mathfrak{p}\}$ to $\mathbf{x}^Q \in S_0(A_{\mathfrak{p}}, \mathbf{b}) \cap T^{|\mathfrak{p}|}$ is clearly injective by design, proving the desired statement. □

In the applications presented later on, whenever we want to prove the existence of a *non-trivial* solution, we will actually always prove the existence of a *proper* solution.

Whenever we want to prove the absence of a non-trivial solution, we will simply prove the absence of proper solutions to $A_{\mathfrak{p}}$ for every partition $\mathfrak{p}$ coming from some non-trivial solution $\mathbf{x} \in S_1(A, \mathbf{b})$. Note that there is only a finite number of such partitions. Lastly, let us note that if A is positive and $\mathbf{x} \in S(A, 0) \cap \mathbb{N}^m$, then $A_{\mathfrak{p}[\mathbf{x}]}$ is clearly also positive. However, the fact that A is abundant does not imply that $A_{\mathfrak{p}(\mathbf{x})}$ needs to be abundant as well, as is illustrated by the matrix $A = (1 \quad 1 \; -3)$ and the solution $\mathbf{x} = (3, 3, 2)$.

1.5 Two notions of density

Let us introduce two notions of density for abundant matrices that parallel those of the maximum density and maximum 2-density of a graph. In order to state them, let $r_Q = r_Q(A) = \operatorname{rk}(A) - \operatorname{rk}(A^{\overline{Q}})$ for any set of column indices $Q \subseteq \{1, \ldots, m\}$ where $\overline{Q} = \{1, \ldots, m\} \setminus Q$ denotes the complement of Q in $\{1, \ldots, m\}$.

Definition 1.11. *For any positive and abundant matrix $A \in \mathbb{Z}^{r \times m}$, its maximum density is given by*

$$m(A) = \max_{\substack{Q \subseteq [m] \\ 2 \leq |Q|}} \frac{|Q|}{|Q| - r_Q} \tag{1.17}$$

and its maximum 1-density is given by

$$m_1(A) = \max_{\substack{Q \subseteq [m] \\ 2 \leq |Q|}} \frac{|Q| - 1}{|Q| - r_Q - 1}. \tag{1.18}$$

Furthermore, we say that A is strictly balanced or strictly 1-balanced if the maximum in $m(A)$ or respectively $m_1(A)$ is exclusively attained by $Q = \{1, \ldots, m\}$.

The later of these two parameters was previously introduced by Rödl and Ruciński [115] for partition regular matrices. The intuition behind why these parameters are relevant is similar to that in the graph case: the probability $p = n^{-1/m(A)}$ roughly describes the point where one would expect the binomial random set to go from not containing any solutions to $A \cdot \mathbf{x}^T = \mathbf{0}^T$ to containing few. In fact, this will precisely be the point of Chapter 2. At $p = n^{-1/m_1(A)}$, more solutions will be present, but they will still be very isolated, that is in expectation each element in the binomial random set should not be part of more than one solution. This parameter will be relevant in Chapter 3 and Chapter 4.

We have not yet shown that both parameters are indeed well-defined, that is $|Q| - r_Q - 1 > 0$ for all $Q \subseteq \{1, \ldots, m\}$ satisfying $|Q| \geq 2$ if A is abundant. We also would like to give some intuition as to why the maximization over Q is needed. In order to do both, we will first develop the notion of an induced submatrix through the following lemma. It was again originally introduced (though not explicitly referred to as such) by Rödl and Ruciński [115] for partition regular matrices. Their proofs, adapted for the full generality of abundant matrices and the inhomogeneous case, are included here for completeness. For any matrix $A \in \mathbb{Z}^{r \times m}$ and selection of row indices $R \subseteq [r]$, we let A_R denote the matrix obtained by only keeping the rows indexed by R.

Proposition 1.12. *For every integer-valued matrix $A \in \mathbb{Z}^{r \times m}$, $\mathbf{b} \in \mathbb{Z}^r$ and $Q \subseteq \{1, \ldots, m\}$ satisfying $r_Q > 0$ there exists an invertible matrix $P = P(A, Q) \in \mathbb{Z}^{r \times r}$ such that the integer-valued matrix*

$$B = B(P, A, Q) = (P \cdot A)^Q_{[r_Q]} \in \mathbb{Z}^{r_Q \times |Q|} \tag{1.19}$$

and the integer-valued vector

$$\mathbf{c} = \mathbf{c}(P, \mathbf{b}) = (P \cdot \mathbf{b})_{[r_Q]} \in \mathbb{Z}^{r_Q} \tag{1.20}$$

satisfy the following:

(i) We have $\operatorname{rk}(B) = r_Q$ *and* $\operatorname{rk}((P \cdot A)^{\overline{Q}}_{[r] \setminus [r_Q]}) = \operatorname{rk}(A) - r_Q$.

(ii) For $\mathbf{x} \in S_0(A, \mathbf{b})$ we have $\mathbf{x}^Q \in S_0(B, \mathbf{c})$.

(iii) For $\mathbf{x} \in S_1(A, \mathbf{b})$ we have $\mathbf{x}^Q \in S_1(B, \mathbf{c})$.

(iv) For $Q' \subseteq \{1, \ldots, |Q|\}$ there exists $Q'' \subseteq Q$ with $|Q''| = |Q'|$ and $r_{Q''}(A) = r_{Q'}(B)$.

(v) If A is positive or abundant, then so is B.

Note that, for any A and Q, there can of course exist multiple P satisfying these properties. We will simply fix an arbitrary such $P = P(A, Q)$ and denote $B(P, A, Q)$ by $B(A, Q)$ as well as $\mathbf{c}(P, \mathbf{b})$ by $\mathbf{c}(A, Q, \mathbf{b})$. The following block decomposition demonstrates the situation for $Q = \{1, \ldots, |Q|\}$.

$$P \cdot A = \begin{pmatrix} B & \mathbf{0} \\ X & Y \end{pmatrix} \begin{array}{l} \big] \, r_Q \\ \big] \, r - r_Q \end{array} \tag{1.21}$$

It will in general be helpful to keep Equation (1.21) in mind.

Proof of Proposition 1.12. We construct P through a standard Gaussian elimination, using elementary row operations. We denote the rows of A by $\mathbf{a}_1, \ldots, \mathbf{a}_r$. Among the rows $\mathbf{a}_1^{\overline{Q}}, \dots, \mathbf{a}_r^{\overline{Q}}$ of $A^{\overline{Q}}$ we choose $\mathrm{rk}(A^{\overline{Q}})$ linearly independent vectors and express each of the remaining $r - \mathrm{rk}(A^{\overline{Q}})$ vectors as a rational linear combination of this basis. Multiplying with the denominators we create integer linear combinations for each of these rows and then we perform the corresponding elementary row operations for each row in A. This turns each entry in the $\overline{Q}$-columns of these $r - \mathrm{rk}(A^{\overline{Q}})$ rows into a 0. Hence the dimension of these rows must be $\mathrm{rk}(A) - \mathrm{rk}(A^{\overline{Q}}) = r_Q$. We identify a set of r_Q linearly independent rows and permute them to the top of the matrix, hence obtaining the promised block decomposition in Equation (1.21).

To prove (i), write $Y = (P \cdot A)^{\overline{Q}}_{[r]\setminus[r_Q]}$ as in the block decomposition and note that the rank of B is r_Q by construction and that

$$\mathrm{rk}(Y) = \mathrm{rk}((P \cdot A)^{\overline{Q}}) = \mathrm{rk}(A^{\overline{Q}}) = \mathrm{rk}(A) - r_Q \tag{1.22}$$

by definition of r_Q.

To prove both (ii) and (iii), note that from the block decomposition it easy to see that for any solution $\mathbf{x} \in S(A, \mathbf{b}) = S(P \cdot A, P \cdot \mathbf{b})$ we have

$$(P \cdot A)_{[r_Q]} \cdot \mathbf{x}^T = (P \cdot \mathbf{b}^T)_{[r_Q]} = \mathbf{c}^T \tag{1.23}$$

and since $(P \cdot A)^{\overline{Q}}_{[r_Q]}$ is the zero-matrix, it follows that $\mathbf{x}^Q \in S(B, \mathbf{c})$, establishing (ii). Again referring to the block decomposition, it is easy to see that if $\mathrm{rk}(B_{\mathfrak{p}(\mathbf{x}^Q)}) < \mathrm{rk}(B)$, then we must in fact also have $\mathrm{rk}(A_{\mathfrak{p}(\mathbf{x})}) < \mathrm{rk}(A)$, establishing (iii).

To prove (iv), let us assume without loss of generality that the columns are permuted such that $Q = \{1, \dots, |Q|\}$ so that we may simply choose $Q'' = Q'$. From (i) we know that we can choose a basis of the vectors space generated by the rows of A that consists of r_Q rows $\mathbf{a}_1, \dots, \mathbf{a}_{r_Q}$ from $(P \cdot A)_{[r_Q]}$ and $\mathrm{rk}(A^{\overline{Q}})$ rows $\mathbf{a}_{r_Q+1}, \dots, \mathbf{a}_{\mathrm{rk}(A)}$ from $(P \cdot A)_{[r]\setminus[r_Q]}$. By construction the vectors $\mathbf{a}^{\overline{Q}}_{r_Q+1}, \dots, \mathbf{a}^{\overline{Q}}_{\mathrm{rk}(A)}$ are linearly independent from each other, so since $Q'' \subseteq Q$ the vectors $\mathbf{a}^{\overline{Q''}}_{r_Q+1}, \dots, \mathbf{a}^{\overline{Q''}}_{\mathrm{rk}(A)}$ are as well, implying $\mathrm{rk}((P \cdot A)^{\overline{Q''}}_{[r]\setminus[r_Q]}) = \mathrm{rk}((P \cdot A)^{\overline{Q}})$. Note that again by construction any linear combination of the vectors $\mathbf{a}_1^{\overline{Q''}}$, ..., $\mathbf{a}_{r_Q}^{\overline{Q''}}$ has the last $|\overline{Q}|$ entries equal to zero and hence cannot be expressed as a linear combination of $\mathbf{a}^{\overline{Q''}}_{r_Q+1}$, ..., $\mathbf{a}^{\overline{Q''}}_{\mathrm{rk}(A)}$, as $\mathbf{a}^{\overline{Q}}_{r_Q+1}$, ..., $\mathbf{a}^{\overline{Q}}_{\mathrm{rk}(A)}$ were linearly independent. It follows that we can add $\mathrm{rk}((P \cdot A)^{\overline{Q''}}_{[r_Q]}) = \mathrm{rk}(B^{\overline{Q'}})$ linearly independent vectors from $\mathbf{a}_1^{\overline{Q''}}, \dots, \mathbf{a}_{r_Q}^{\overline{Q''}}$ to the $\mathrm{rk}((P \cdot A)^{\overline{Q}}) = \mathrm{rk}(A^{\overline{Q}})$ linearly independent vectors $\mathbf{a}^{\overline{Q''}}_{r_Q+1}, \dots, \mathbf{a}^{\overline{Q''}}_{\mathrm{rk}(A)}$ to form a basis of the row space of $\mathrm{rk}(A^{\overline{Q''}})$. This implies that $\operatorname{rk}(A^{\overline{Q''}}) = \operatorname{rk}(A^{\overline{Q}}) + \operatorname{rk}(B^{\overline{Q'}})$ from which we can conclude that

$$\begin{aligned} r_{Q''}(A) &= \operatorname{rk}(A) - \operatorname{rk}(A^{\overline{Q''}}) = r_Q + \operatorname{rk}(A^{\overline{Q}}) - \operatorname{rk}(A^{\overline{Q''}}) \\ &= \operatorname{rk}(B) - \operatorname{rk}(B^{\overline{Q'}}) = r_{Q'}(B) \end{aligned}$$

as desired.

Lastly, to prove (v) we note that the aspect of positivity immediately follows from (ii). Regarding abundancy, we again invoke the block composition to see that if there is some set of column indices $Q' \subseteq Q$, where without loss of generality we assume that $Q = \{1, \dots, |Q|\}$, satisfying $|Q'| \geq |Q| - 2$ as well as $\operatorname{rk}(B^Q) < \operatorname{rk}(B)$, then $Q'' = Q' \cup \{|Q|+1, \dots, m\}$ must satisfy both $|Q''| \geq m - 2$ and $\operatorname{rk}(A^{Q''}) < \operatorname{rk}(A)$. □

The following corollary to this lemma will now allow us to handle the case when a given matrix is, depending on the context, not strictly balanced or not strictly 1-balanced.

Corollary 1.13. *For every positive and abundant $A \in \mathbb{Z}^{r \times m}$ there exist $Q, Q_1 \subseteq \{1, \dots, m\}$ such that for $B = B(A, Q)$, $B_1 = (A, Q_1)$, $\mathbf{c} = \mathbf{c}(A, Q, \mathbf{b})$ and $\mathbf{c}_1 = \mathbf{c}(A, Q_1, \mathbf{b})$ as given by Proposition 1.12 and for any subset $T \subseteq \mathbb{N}$ we have*

(1) B and B_1 are abundant and positive,
(2) B is strictly balanced and satisfies $m(B) = m(A)$,
(3) B_1 is strictly 1-balanced and satisfies $m_1(B_1) = m_1(A)$,
(4) if $S_0(B, \mathbf{c}) \cap T^m = \emptyset$ or $S_0(B_1, \mathbf{c}_1) \cap T^m = \emptyset$ then $S_0(A, \mathbf{b}) \cap T^m = \emptyset$ and
(5) if $S_1(B, \mathbf{c}) \cap T^m = \emptyset$ or $S_1(B_1, \mathbf{c}_1) \cap T^m = \emptyset$ then also $S_1(A, \mathbf{b}) \cap T^m = \emptyset$.

Proof. We start with the case of *strictly balanced*. Choose $Q \subseteq \{1, \dots, m\}$ such that $|Q|/(|Q| - r_Q) = m(A)$ and $|Q|$ is minimal with this property. By (v) we know that B is positive and abundant. Assume that there exists $Q' \subsetneq \{1, \dots, |Q|\}$ such that $|Q'|/(|Q'| - r_{Q'}(B)) \geq |Q|/(|Q| - r_Q - 1)$. By (iv) there must exist $Q'' \subseteq \{1, \dots, m\}$ with $|Q''| = |Q'| < |Q|$ such that $r_{Q''}(A) = r_{Q'}(B)$. It follows that $|Q''|/(|Q''| - r_{Q''}(A)) \geq m_1(A)$, giving us a contradiction to our choice of Q. For the case of *strictly 1-balanced*, we simply choose $Q \subseteq \{1, \dots, m\}$ such that $(|Q| - 1)/(|Q| - r_Q - 1) = m_1(A)$ and $|Q|$ is again minimal with this property and proceed exactly the same way as before. From this (1), (2) and (3) immediately follow. Finally, (4) readily follows from (ii) and (5) from (iii). □

The following lemma now establishes some results regarding the rank of induced submatrices of abundant matrices. It also verifies that the two notions of density are indeed well-defined for abundant matrices.

Lemma 1.14. *For any abundant matrix $A \in \mathbb{Z}^{r \times m}$ and $Q \subseteq \{1, \dots, m\}$ the following holds: if $|Q| \geq 3$ then $|Q| - r_Q - 1 \geq 1$ and if $|Q| \leq 2$ then $r_Q = 0$.*

Proof. If $|Q| \leq 2$ then, since A is abundant, deleting the columns in Q does not reduce the rank of A. Hence $\text{rk}(A^{\overline{Q}}) = \text{rk}(A)$ and therefore $r_Q = 0$. Now if $|Q| \geq 3$ and $r_Q = 0$, then trivially $|Q| - r_Q - 1 \geq 2$. If $|Q| \geq 3$ and $r_Q \geq 1$, then by Proposition 1.12 $B(A, Q)$ is abundant, has rank r_Q and the number of its columns is $|Q|$. Since for any abundant matrix the number of columns must be at least two more than its rank, it follows that $r_Q \leq |Q| + 2$ and therefore $|Q| - r_Q - 1 \geq 1$. □

1.6 Tools from Probability Theory

Let us briefly survey some well established tools from Probability Theory that will be used throughout this chapter. Given a random variable X, we denote its expected value by $\mathbb{E}(X)$ and its variance by $\text{Var}(X)$.

Markov's Inequality states that for any X satisfying $\mathbb{P}(X \geq 0) = 1$ we have

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t} \tag{1.24}$$

for any $t > 0$. A straight forward consequence of Markov's Inequality is that if a sequence of random variable $(X_n)_{n \in \mathbb{N}}$ satisfies $\mathbb{P}(X_n > 0) = o(1)$, then it must also satisfy $\mathbb{E}(X_n) = o(1)$.

Chebyshev's Inequality allows one to establish concentration of a random variable around its expected value. It states that, for any random variable X with finite expected value and non-zero variance, we have

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{\text{Var}(X)}{t^2} \tag{1.25}$$

for any $t > 0$.

The **Second Moment Method** is an easy application of Chebyshev's Inequality will be used in the version given by Corollary 4.3.4 of Alon and Spencer [2]. Let

$X_n = \mathbb{1}_1 + \cdots + \mathbb{1}_n$ be the sum of n indicator random variables of some events E_i. We write $i \sim j$ if $i \neq j$ and the events E_i and E_j are not independent. Define

$$\Delta = \sum_{i \sim j} \mathbb{P}\left(E_i \wedge E_j\right). \tag{1.26}$$

If $\mathbb{E}(X_n) = \omega(1)$ and $\Delta = o\Big(\mathbb{E}(X_n)^2\Big)$, then $X_n = \mathbb{E}(X_n)\,(1 + o(1))$ asymptotically almost surely and in particular $X_n > 0$ asymptotically almost surely.

We will state **Brun's Sieve** as given by Theorem 8.3.1 of Alon and Spencer [2]. Let $X_n = \mathbb{1}_1 + \cdots + \mathbb{1}_n$ be the sum of n indicator random variables of some events E_i. Let

$$S^{(t)} = \sum_{\{i_1,\ldots,i_t\} \in \binom{[n]}{t}} \mathbb{P}\left(E_{i_1} \wedge E_{1_2} \wedge \cdots \wedge E_{i_t}\right), \tag{1.27}$$

for any $t \in \mathbb{N}$ where the sum is taken over all subsets $\{i_1, \ldots, i_t\} \subseteq [n]$ of size t. From the Inclusion-Exclusion Principle it follows that

$$\begin{aligned} \mathbb{P}\left(X_n = 0\right) &= \mathbb{P}\left(\overline{E_1} \wedge \cdots \wedge \overline{E_s}\right) \\ &= 1 - S^{(1)} + S^{(2)} - \cdots + (-1)^t S^{(t)} + \cdots. \end{aligned}$$

Suppose now there is a constant μ satisfying $\mathbb{E}(X) = S^{(1)} = \mu(1 + o(1))$ and that for every fixed $t \in \mathbb{N}$ we have $S^{(t)} \to \mu^t/t!$ as n goes to infinity. Then we have

$$\mathbb{P}\left(X_n = t\right) \to \frac{\mu^t}{t!}\, e^{-\mu}. \tag{1.28}$$

for every fixed $t \in \mathbb{N}$.

Lastly, there are many results know as **Chernoff Bounds** that give exponential bounds on the tail distribution of sums of independent random variables. We will just be interested in the case for the binomial distribution $\mathcal{B}(n,p)$ on n elements with parameter p, where we have

$$\mathbb{P}\left(\mathcal{B}(n,p) < \frac{np}{2}\right) \leq \exp\left(-\frac{n}{2p}\,(p - 1/2)^2\right). \tag{1.29}$$

1.7 A short survey of extremal results

Let us briefly survey the vast area of explicit quantitive bounds regarding the extremal question of how large a set can be without containing solutions to a given system. For some general results, see Ruzsa [124, 126] as well as Shapira [141].

In the particular case of Szemerédi's Theorem, we note that for $k = 3$ the upper bounds obtained by Roth were refined by Heath-Brown [83], Szémeredi [151], Bourgain [15, 16], Sanders [130] and Bloom [13] to its current best of $O(n \log^4 \log n / \log n)$. On the other hand, Behrend [11] gave a construction of a set free of 3-term arithmetic progressions, that was later slightly improved upon by Elkin [45], see also Green and Wolf [73] and O'Bryant [111], to give a set of size $n \log^{1/4} n / 2^{\sqrt{8 \log n}}$. Concerning the case of general k, the best current upper bounds are due to Gowers [69] and, more recently, dense constructions that lead to lower bounds for this problem were established by O'Bryant [111], building on the previous work of Behrend, Rankin [113], Elkin [45, 73] as well as Łaba and Lacey [102]. Notable progress has also been made in the study of arithmetic progressions in $\mathbb{F}_q^n$, see the recent breakthrough of Ellenberg and Gijswijt [46] and the paper of Croot, Lev and Pach [35] on which it is based.

	lower bound	upper bound
3-AP	$n \log^{1/4} n / 2^{\sqrt{8 \log n}}$	$O(n \log^4 \log n / \log n)$
k-AP	$\Omega\left(\frac{n \exp((\log \log n)/(2\lceil \log k \rceil))}{\exp(\lceil \log k \rceil 2^{(\lceil \log k \rceil - 1)/2} \log^{1/\lceil \log k \rceil} n))}\right)$	$n/(\log \log n)^{2^{-2^{k+9}}}$
Sidon	— $(1 + o(1))\, n^{1/2}$ —	
$B_h[g]$	$(1 + o(1))\, n^{1/h}$	$gh\, h!\, n^{1/h}$
3-cube	$\Omega(n^{2/3})$	$2n^{3/4}$
k-cube	$n^{1-k/(2^k-1)}$	$2n^{1-1/2^{k-1}}$
sum-free	— $\lceil n/2 \rceil$ —	
3-sum-free	— $\lceil n/2 \rceil$ —	
k-sum-free	— $\left(\frac{k(k-2)}{k^2-2} + \frac{8(k-2)}{k(k^2-2)(k^4-2k^2-4)}\right)(1 + o(1))\, n$ —	

Table 1.2: *Extremal bounds for common linear systems.*

Regarding structures other than that of arithmetic progressions, the most notable example is perhaps that of Sidon sets, also known as Golomb rulers to other parts of the Discrete Mathematics world. A set of integers $S \subset \mathbb{N}$ is called a Sidon set if all pairwise sums of its elements are distinct, that is if $x+w \neq y+z$ for any $x, y, z, w \in S$ satisfying $x < y \leq z < w$. This can be considered as sets excluding solutions associated with the matrix $A = (1 \quad 1 - 1 - 1)$. Results of Chowla [25], Erdős [48], Erdős and Turán [55] and Singer [143] established that the maximum cardinality of a Sidon set contained in $[n] = \{1, 2, \ldots, n\}$ is $(1 + o(1))\, n^{1/2}$. Sidon sets have also been generalized to $\mathsf{B_h}$ sets, that is sets where all h-fold sums of its elements are unique, and even more broadly to $\mathsf{B_h[g]}$ sets, that is sets where every integer has at most g representations as a sum of h of the elements of the set. It is known that a maximum $B_h[g]$ set in $[n]$ is of size $\Theta(n^{1/h})$ but precise constants are not know in general for the problem, see [32, 31]. Note that we will consider a question dealing with Sidon sets of infinite cardinality in Chapter 5 as well as a problem tangentially related to $B_h[g]$ sets in Chapter 7.

Another possible generalization of Sidon sets are sets free of k-dimensional Hilbert cubes, or just k-cube for short, that is configurations of the shape

$$\left\{ h_0 + \sum_{i=1}^{k} \varepsilon_i h_i \,:\, \varepsilon_i \in \{0, 1\} \right\} \tag{1.30}$$

for some positive and distinct integers $h_0, h_1, \ldots, h_k \in \mathbb{N}$. Sidon sets are sets free of 2-dimensional Hilbert cubes. Hilbert [84], in a Ramsey-type result predating even that of Schur, originally proved that any finite coloring of the positive integers contains a monochromatic k-cube. The density version of this result is known as Szemerédi's Cube Lemma and is a key ingredient in his re-proof of Roth's Theorem. Gunderson and Rödl [75] proved that any set in $[n]$ of size $2n^{1-1/2^{k-1}}$ contains a k-cube for sufficiently large n. On the other side, a probabilistic argument proves the existence of a set of size $n^{1-k/(2^k-1)}$ avoiding k-cubes. For the particular case of $k = 3$, Cilleruelo and Tesoro [33] obtained an algebraic construction of a set of size $\Omega(n^{2/3})$.

Finally, the maximum size of sets containing no solutions associated with matrices that are not density regular has also been studied. In particular, a set of integers is a k-sum-free set if it contains no solution to $x + y = kz$. The case of $k = 1$ corresponds to Schur triples and we have already seen the extremal construction: one cannot select more than $\lceil n/2 \rceil$ integers in $[n]$ without selecting all elements of a Schur triple. The case $k = 2$ corresponds to 3-term arithmetic progressions. Chung and Goldwasser [26] solved the case of $k = 3$ by getting the same estimates as for $k = 1$. The case $k \geq 4$ was also studied by Chung and Goldwasser [27] and then settled by Baltz, Hegarty, Knape, Larsson and Schoen [5]. Finally, Hancock and Treglown [79, 80] also studied the case of sets excluding solutions to more general three-variable equations $a_1 x + a_2 y = a_3 z$

where $a_1 \geq a_2 \geq a_3$ and $\gcd(a_1, a_2, a_3) = 1$.

We observe that results dealing with the extremal question often rely on smart ad-hoc arguments that strongly depend on the specific structures being considered. Answers regarding the typical behavior, as we will study in the remainder of this chapter, can often cover large groups of structures in a unified statement.

Chapter 2

The Appearance Threshold

In this chapter, we will be interested in counting the number of proper and non-trivial solutions to a linear system $A \cdot \mathbf{x}^T = \mathbf{b}^T$ in the binomial random set $[n]_p$ given by some arbitrary positive matrices $A \in \mathbb{Z}^{r\times m}$ and vectors $\mathbf{b} \in \mathbb{Z}^r$. As a first result, let us establish the threshold for the property that $[n]_p$ contains a non-trivial solution when A is abundant.

Theorem 2.1. *For any positive and abundant $A \in \mathbb{Z}^{r\times m}$ and $\mathbf{b} \in \mathbb{Z}^r$ satisfying $S(A, \mathbf{b}) \neq \emptyset$, the function $p_0(n) = n^{-1/m(A)}$ is a threshold for the property that $[n]_p$ contains a non-trivial solution to $A \cdot \mathbf{x}^T = \mathbf{b}^T$.*

In other words, almost all subsets of $[n]$ that are of size $o(n^{1-1/m(A)})$ contain no non-trivial solution and almost all that are of size $\omega(n^{1-1/m(A)})$ do. Note that $m(A)$ refers to the maximum density of A as defined in Section 1.5. The proof of Theorem 2.1 will be given in Section 2.1. When establishing the 1-statement, we will in fact prove the existence of a *proper* solution in the set, that is $p_0(n) = n^{-1/m(A)}$ is also a threshold for the property that $[n]_p$ contains a proper solution. In Section 2.1 we will also establish a statement regarding non-abundant matrices.

Note that the case of $B_h[1]$ sets was previously studied in detail by Godbole, Janson, Loncatore and Rapoport in [68] who showed that almost no set of size $\omega(n^{1/2h})$ is $B_h[1]$. Their proof is based on a tailor-made analysis on the particular shape of the equations defining $B_h[1]$ sets.

The threshold given in Theorem 2.1 is not sharp, so let us study the distribution of the number of non-trivial solutions when p asymptotically grows like the threshold. Here, we will restrict ourselves to the particular case of homogeneous systems coming from strictly balanced matrices. We will show that if $p(n) = Cn^{-1/m(A)}$ for some fixed

$C > 0$, then the number of solutions converges to a Poisson distribution. Furthermore, the parameter of that Poisson distribution only depends, besides the constant C, on the volume of the polytope $\mathcal{P}_A$ and the inherent symmetry of the system $\sigma(A)$, as defined in Section 1.3.

Theorem 2.2. *For any positive and abundant matrix $A \in \mathbb{Z}^{r\times m}$ and $p(n) = Cn^{-1/m(A)}$ for some fixed $C > 0$, the number of distinct, non-trivial solutions to $A \cdot \mathbf{x}^T = \mathbf{0}^T$ in $[n]_p$ converges in distribution to a Poisson distributed random variable with parameter*

$$\lambda = \lambda(A, C) = \operatorname{Vol}(\mathcal{P}_A)\, C^m/\sigma(A) \tag{2.1}$$

if and only if A is strictly balanced.

Note that this means that, for every non-negative integer t, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\left|[n]_p^m \cap S_1(A, \mathbf{0})\right| = t\right) = \tfrac{\lambda^t}{t!}e^{-\lambda}.$$

In particular, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\left|[n]_p^m \cap S_1(A, \mathbf{0})\right| > 0\right) = 1 - e^{-\lambda}, \tag{2.2}$$

that is the probability of $[n]_p$ containing a non-trivial solution tends to 1 with an exponential decay in C. We will prove Theorem 2.2 in Section 2.2. As previously for Theorem 2.1, the statement of Theorem 2.2 likewise holds when one restricts oneself to proper solutions.

In a direction similar to this result, Warnke [160] studied the upper tail of the number of k-term arithmetic progressions and Schur triples in random subsets, establishing exponential bounds. Let us also mention that Kohayakawa, Lee, Rödl and Samotij [96] studied the number of Sidon sets and the maximum size of Sidon set contained in a sparse random set of integers. In particular, they analyze the number of solutions to the Sidon equation when the probability lies above the threshold by means of the Kim-Vu polynomial concentration inequality [92].

The computation of the constant $\operatorname{Vol}(\mathcal{P}_A)$ for arbitrary $A \in \mathbb{Z}^{r\times m}$ is an algorithmically involved problem. One could compute this volume by means of triangulations of the polytope [39], but in dimensions greater than 3 the problem is in general NP-complete [18]. Some concrete values for specific matrices will be provided in Section 2.3. The results of Theorem 2.1, Theorem 2.2 and that section are summarized in Figure 2.1

for some of the examples previously introduced in Section 1.7. Note that all of these examples are strictly balanced.

	$r \times m$	np_0	$\mathrm{Vol}\,(\mathcal{P}_A)$	$\sigma(A)$
k-AP	$k-2 \times k$	$n^{1-2/k}$	$1/(k-1)$	2
Sidon	1×4	$n^{1/4}$	$2/3$	8
$B_h[g]$	$g \times h(g+1)$	$n^{\frac{g}{h(g+1)}}$	Section 2.3	$(g+1)!(h!)^{g+1}$
k-cube	$2^k - (k+1) \times 2^k$	$n^{1-\frac{k+1}{2^k}}$	$2^{2^k-1}/(k+1)!k!$	2^{2^k-1}
sum-free	1×3	$n^{1/3}$	$1/2$	2
k-sum-free	1×3	$n^{1/3}$	$1/k$	2

Table 2.1: *Results of this chapter for common systems.*

Lastly, let us compare the results of Theorem 2.1 to the extremal case summarized in Section 1.7: almost all sets with size $\omega(n^{1-2/k})$ contain k-term arithmetic progressions. This means that for $k = 3$ the gap between the usual and the extremal situation is very large. Most sets with size $\omega(n^{1/3})$ contain 3-term arithmetic progressions but there are examples of almost linear size avoiding this structure. Nevertheless, as k grows, the gap between the exponents tends to 0. For Sidon sets, the gap between the exponents in the extremal and usual situation is likewise very big. However, almost all sets in $[n]$ of size $o(n^{1/h-1/h(g+1)})$ are $B_h[g]$ sets, so if we fix h and let g grow to infinity, both situations once again approach each other. Lastly, the maximal size of a k-sum-free set is linear in n, but Theorem 2.1 asserts that almost all sets of size $\omega(n^{1/3})$ contain at least one solution to $x + y = kz$, for every k. In this family the parameter k does not play a role in the position of the threshold.

2.1 Proof of Theorem 2.1 – When do solutions appear?

Before proving Theorem 2.1, let us briefly establish a result concerning the case when A is non-abundant. Note that a corresponding 1-statement does not hold in general.

Proposition 2.3. *For any positive but non-abundant matrix $A \in \mathbb{Z}^{r\times m}$ and $\mathbf{b} \in \mathbb{Z}^r$ satisfying $S(A, \mathbf{b}) \neq \emptyset$, $[n]_p$ asymptotically almost surely does not contain any solution to $A \cdot \mathbf{x}^T = \mathbf{b}^T$ when $p = o(n^{-1/2})$.*

Proof. By Lemma 1.2 we may, without loss of generality, assume that the first row of A is of the shape $\mathbf{a} = (a_1, \ldots, a_m)$ where $a_i = 0$ for all $i \in \{3, \ldots, m\}$ and $a_1, a_2 \neq 0$ as well as $a_1+a_2 \neq 0$. Writing b_1 for the first entry of $\mathbf{b}$, it follows that any $\mathbf{x}(x_1,\dots,x_n) \in S(A,\mathbf{b}) \cap [n]^m$ must satisfy

$$a_1x_1 + a_2x_2 = b_1. \tag{2.3}$$

The probability of $[n]_p$ containing any solution to $A \cdot \mathbf{x}^T = \mathbf{b}^T$ is therefore bounded from above by the probability of $[n]_p$ containing two elements x_1 and x_2 satisfying Equation (2.3).

Since $a_1 + a_2 \neq 0$ either x_1 and x_2 are distinct or $x_1 = x_2 = b_1/(a_1 + a_2)$. The later case trivially occurs with probability exactly $p = o(1)$, so let us focus on the former. The number of such distinct pairs (x_1, x_2) in $[n]$ is bounded by n and the probability that both elements of a pair lie in $[n]_p$ is p^2. It follows by linearity of expectation that the probability of $[n]_p$ containing such a suitable pair is bounded by $np^2 = o(1)$. The desired statement therefore follows by Markov's Inequality. □

Let us now establish Theorem 2.1.

Proof of Theorem 2.1. We split the proof up into two separate arguments for the 0- and the 1-statement. The former will follow from Markov's Inequality and the later from the Second Moment Method, both introduced in Section 1.6.

The 0-statement. One may assume that A is strictly balanced as otherwise we can replace A and $\mathbf{b}$ with B and $\mathbf{c}$ as given by Corollary 1.13. Let us apply Markov's Inequality, that is we will show that the expected number of non-trivial solutions to $A \cdot \mathbf{x}^T = \mathbf{b}^T$ goes to zero if $p = o(n^{-1/m_1(A)})$, in order to show that the probability of $[n]_p$ containing a non-trivial solutions goes to zero as well.

We will write $S_n = S_1(A,\mathbf{b}) \cap [n]^m$ and for every $\mathbf{x} \in S_n$ we let $\mathbb{1}_\mathbf{x}$ be the indicator variable for the event that $\mathbf{x} \in [n]_p^m$. It follows that

$$X_n = \sum_{\mathbf{x} \in S_n} \mathbb{1}_\mathbf{x} \tag{2.4}$$

is the random variable counting the number of non-trivial solutions to $A \cdot \mathbf{x}^T = \mathbf{b}^T$ in $[n]_p$. Using linearity of expectation and Lemma 1.10, it follows that

$$\begin{aligned}\mathbb{E}(X_n) = \sum_{\mathbf{x} \in S_n} \mathbb{P}\left(\mathbf{x} \in [n]_p^m\right) &= \sum_{s=\mathrm{rk}(A)}^{m} \sum_{\substack{\mathfrak{p} \in \mathfrak{P}(A) \\ |\mathfrak{p}|=s}} \sum_{\substack{\mathbf{x} \in S_n \\ \mathfrak{p}(\mathbf{x})=\mathfrak{p}}} p^s \\ &\leq \sum_{s=\mathrm{rk}(A)}^{m} \sum_{\substack{\mathfrak{p} \in \mathfrak{P}(A) \\ |\mathfrak{p}|=s}} |S_0(A_\mathfrak{p},\mathbf{b}) \cap [n]^s|\, p^s.\end{aligned}$$

Applying Equation (1.3) as well as the fact that $\mathrm{rk}(A_{\mathfrak{p}}) = \mathrm{rk}(A)$ for any $\mathfrak{p} \in \mathfrak{P}(A)$, we get that

$$\mathbb{E}(X_n) \leq \sum_{s=\mathrm{rk}(A)}^{m} \sum_{\substack{\mathfrak{p}\in\mathfrak{P}(A)\\ |\mathfrak{p}|=s}} n^{s-\mathrm{rk}(A)} p^s = \sum_{s=\mathrm{rk}(A)}^{m} \sum_{\substack{\mathfrak{p}\in\mathfrak{P}(A)\\ |\mathfrak{p}|=s}} (n^{1-\mathrm{rk}(A)/s} p)^s.$$

Since we are assuming that A is strictly balanced, we have that $1-\mathrm{rk}(A)/m = 1/m(A)$. Since $|\mathfrak{p}| \leq m$, it therefore follows that

$$\mathbb{E}(X_n) \leq \sum_{s=\mathrm{rk}(A)}^{m} \sum_{\substack{\mathfrak{p}\in\mathfrak{P}(A)\\ |\mathfrak{p}|=s}} \left(\frac{p}{n^{-1/m(A)}}\right)^s = \sum_{s=\mathrm{rk}(A)}^{m} O\left(\frac{p}{n^{-1/m(A)}}\right)^s = o(1).$$

Here we have used the obvious fact that $|\mathfrak{P}(A)|$ as well as $m - \mathrm{rk}(A) + 1$ are finite, that is they do not grow with n.

The 1-statement. Let us show that $[n]_p$ not only contains a non-trivial solution asymptotically almost surely, but that it in fact contains a *proper* solution. We therefore write $S_n = S_0(A, \mathbf{b}) \cap [n]^m$ and for every $\mathbf{x} \in S_n$ we let $\mathbb{1}_{\mathbf{x}}$ again be the indicator variable for the event that $\mathbf{x} \in [n]_p^m$. It follows that $X_n = \sum_{\mathbf{x}\in S_n} \mathbb{1}_{\mathbf{x}}$ is now the random variable counting the number of proper solutions to $A \cdot \mathbf{x} = \mathbf{b}$ in $[n]_p$. We first establish that that the expected value of X_n tends to infinity. By Lemma 1.4, there exists some $c_0 = c_0(A, \mathbf{b})$ such that

$$\mathbb{E}(X_n) = \sum_{\mathbf{x}\in S_n} p^m \geq c_0 n^{m-\mathrm{rk}(A)} p^m = c_0 \left(\frac{p}{n^{-(m-\mathrm{rk}(A))/m}}\right)^m.$$

Since $p = \omega(n^{-1/m(A)})$ and by definition $1 - \mathrm{rk}(A)/m \geq 1/m(A)$, it follows that $\mathbb{E}(X_n)$ tends to infinity.

Let us now study the variance of X_n in order to apply the Second Moment Method. We know that, for $\mathbf{x} = (x_1, \ldots, x_m) \in S_n$ and $\mathbf{y} = (y_1, \ldots, y_m) \in S_n$, the two events $\mathbf{x} \in [n]_p^m$ and $\mathbf{y} \in [n]_p^m$ are dependent if and only if $\{x_1, \ldots, x_m\} \cap \{y_1, \ldots, y_m\} \neq \emptyset$. Let us write $\mathbf{x} \sim \mathbf{y}$ if this is the case. Following the notation in Equation (1.26), we need to establish that the quantity

$$\Delta_n = \sum_{\substack{\mathbf{x},\mathbf{y}\in S_n\\ \mathbf{x}\sim\mathbf{y}}} \mathbb{P}\left(\mathbf{x} \in [n]_p^m \wedge \mathbf{y} \in [n]_p^m\right) \tag{2.5}$$

is asymptotically dominated by $\mathbb{E}(X_n)^2$.

Let $\mathbf{x} \in S_n$ now be arbitrary but fixed and $M = (V, E)$ a graph that defines a non-empty, directed matching between two copies of of the column indices $\{1, \ldots, m\}$ where all edges are directed towards the same part. Note that this obviously implies that $1 \leq |E| \leq m$. Let us establish an upper bound for the number of solutions $\mathbf{y} \in S_n$ whose 'intersections' with $\mathbf{x}$ are indicated by M, that is $y_j = x_i$ if and only if $ij \in E$. Writing $E = \{i_1 j_1, \ldots, i_{|E|} j_{|E|}\}$ such that $j_1 \leq \ldots \leq j_t$ as well as $Q = Q(M) = \{j_1, \ldots, j_{|E|}\}$, it is clear that the number of such solutions is bounded from above by the number of solutions to

$$A^{\overline{Q}} \cdot \mathbf{z} = \mathbf{b}^T - A^Q \cdot (x_{i_1}, \ldots, x_{i_{|E|}})^T. \tag{2.6}$$

Since $\operatorname{rk}(A^{\overline{Q}}) = r - r_Q$, it follows by Equation (1.3) that there are at most

$$n^{|\overline{Q}| - r + r_Q} = n^{m - |Q| - r + r_Q} \tag{2.7}$$

such solutions. Writing $\mathcal{M}$ for the family of such matchings and $\mathbf{x} \sim_M \mathbf{y}$ if $\mathbf{x} \sim \mathbf{y}$ and the intersections of $\mathbf{x}$ and $\mathbf{y}$ are indicated by some $M \in \mathcal{M}$, we therefore have that

$$\begin{aligned}
\Delta_n &= \sum_{\substack{\mathbf{x},\mathbf{y} \in S_n \\ \mathbf{x} \sim \mathbf{y}}} \mathbb{P}\left(\mathbf{x} \in [n]_p^m \wedge \mathbf{y} \in [n]_p^m\right) \\
&= \sum_{\mathbf{x} \in S_n} \sum_{\emptyset \neq Q \subsetneq [m]} \sum_{\substack{M \in \mathcal{M} \\ Q(M) = Q}} \sum_{\substack{\mathbf{y} \in S_n \\ \mathbf{x} \sim_M \mathbf{y}}} p^{2m - |Q|} \\
&\leq \sum_{\mathbf{x} \in S_n} \sum_{\emptyset \neq Q \subsetneq [m]} \sum_{\substack{M \in \mathcal{M} \\ Q(M) = Q}} n^{m - |Q| - \operatorname{rk}(A) + r_Q} \, p^{2m - |Q|} \\
&= \sum_{\emptyset \neq Q \subsetneq [m]} O\left(n^{2m - 2\operatorname{rk}(A) - |Q| + r_Q} \, p^{2m - |Q|}\right) \\
&= O\left(n^{m - \operatorname{rk}(A)} p^m\right)^2 \sum_{\emptyset \neq Q \subsetneq [m]} \left(\frac{n^{-(|Q| - r_Q)/|Q|}}{p}\right)^{|Q|} = o\left(\mathbb{E}(X_n)^2\right).
\end{aligned}$$

Here we have again used that $(|Q| - r_Q)/|Q| \geq 1/m(A)$ so that $p = \omega(n^{-1/m(A)})$ also implies that $p = \omega(n^{-(|Q| - r_Q)/|Q|})$ for any non-empty $Q \subseteq \{1, \ldots, m\}$. By the Second Moment Method, it follows that $X_n > 0$ asymptotically almost surely. ■

2.2 Proof of Theorem 2.2 – Distribution at the threshold

Sufficiency. Lust us write S_n for the set of *distinct* non-trivial solutions to $A \cdot \mathbf{x}^T = \mathbf{0}^T$. Here we will simply assume that for each equivalence class of solutions in $S_1(A, \mathbf{0})$ whose elements are pairwise non-distinct, that is they are the same up to some permutation in $\Sigma(A)$, we have only one arbitrary representative in S_n. Let X_n now denote the random variable counting the number of solutions from S_n in $[n]_p$. We can easily develop the expected value of X_n using Corollary 1.8, Equation (1.3) and Lemma 1.4, that is

$$\begin{aligned}\mathbb{E}(X_n) &= \sum_{\mathbf{x} \in S_n} \mathbb{P}\left(\mathbf{x} \in [n]_p^m\right) = \sum_{s=\mathrm{rk}(A)}^{m} \sum_{\substack{\mathfrak{p} \in \mathfrak{P}(A) \\ |\mathfrak{p}|=s}} \sum_{\substack{\mathbf{x} \in S_n \\ \mathfrak{p}(\mathbf{x})=\mathfrak{p}}} p^s \\ &= \sum_{\substack{\mathbf{x} \in S_n \\ |\mathfrak{p}(\mathbf{x})|=m}} p^m + \sum_{s=\mathrm{rk}(A)}^{m-1} \sum_{\substack{\mathfrak{p} \in \mathfrak{P}(A) \\ |\mathfrak{p}|=s}} \Theta(n^{s-\mathrm{rk}(A)}) p^s \\ &= \mathrm{Vol}\left(\mathcal{P}_A\right)/\sigma(A)(1+o(1))\, n^{m-\mathrm{rk}(A)}\, p^m + \Theta\left(n^{1-\mathrm{rk}(A)/s} p\right)^s \\ &= \mu(1+o(1))\end{aligned}$$

where we have set $\mu = C^m \mathrm{Vol}\left(\mathcal{P}_A\right)/\sigma(A)$. Here we have also used the fact that A is strictly balanced, so that $1 - \mathrm{rk}(A)/s < 1/m(A)$ if $s < m$, and that $p = Cn^{-1/m(A)}$, so that $O\left(n^{1-\mathrm{rk}(A)/s} p\right) = o(1)$.

For $t \geq 2$, let $X_{t,n}$ now denote the number of ordered t-tuples of distinct non-trivial solutions in $[n]_p$. We split it up into three parts

$$X_{t,n} = X'_{t,n} + X''_{t,n} + X'''_{t,n}. \tag{2.8}$$

The first part $X'_{t,n}$ refers to t-tuples of pairwise disjoint distinct proper solutions, $X''_{t,n}$ refers to t-tuples of pairwise disjoint distinct non-trivial solutions of which at least one is not proper and $X'''_{t,n}$ refers to t-tuples of distinct non-trivial solutions in which at least two share a coordinate. We will compute the expected value of each of the parts in order to show that $\mathbb{E}(X_{t,n}) = \mu^t(1+o(1))$, so that the desired statement follows by Brun's Sieve.

In order to compute the expected number of ordered t-tuples of solutions, we need to introduce some additional notation. Consider two matrices $A \in \mathbb{Z}^{r_A \times m_A}$ and $B \in \mathbb{Z}^{r_B \times m_B}$ as well as an incomplete and directed matching $M = (V, E)$ where V consists of disjoint copies of $\{1, \ldots, m_A\}$ and $\{1, \ldots, m_B\}$ and $E = \left\{i_1 j_1, \ldots, i_{|E|} j_{|E|}\right\}$ where

$|E| < \min(m_A, m_B)$. Note the similarity to the proof of the 1-statement of Theorem 2.1. Denote the columns of A and B by $\{\mathbf{a}_i : 1 \leq i \leq m_A\}$ and $\{\mathbf{b}_j : 1 \leq j \leq m_B\}$ respectively. The compounded matrix of A, B and M is now defined as

$$A \times_M B = \begin{pmatrix} A^{[m_A]\setminus\{i_1,\dots,i_{|E|}\}} & \mathbf{a}_{i_1} & \dots & \mathbf{a}_{i_{|E|}} & 0 \\ 0 & \mathbf{b}_{j_1} & \dots & \mathbf{b}_{j_{|E|}} & B^{[m_B]\setminus\{j_1,\dots,j_{|E|}\}} \end{pmatrix}. \tag{2.9}$$

There is a clear bijection between proper solutions to the system given by $(A \times_M B) \cdot \mathbf{x}^T = \mathbf{0}^T$ and to 2-tuples of proper solutions to $A \cdot \mathbf{x}^T = \mathbf{0}^T$ and $B \cdot \mathbf{x}^T = \mathbf{0}^T$ whose coincidences are indicated by M.

For our application the actual matching and hence the concrete type of overlap will be irrelevant. What matters is whether the systems are disjoint or not, that is if the bipartite graph is empty or if there is some actual overlap. We therefore simply omit the graph in our notation and write $A \times B$ when compounding matrices and $A \dot{\times} B$ to specify when they are being compounded without overlap, that is the implied M is empty. Note that this operator is not commutative or associative (and in fact strongly depends on the size of the respective matrices) and we will write $A \times B \times C = (A \times B) \times C$.

Using this notation, we observe that finding t-tuples of pairwise disjoint distinct and proper solutions to $A \cdot \mathbf{x}^T = \mathbf{0}^T$ is equivalent to finding proper solutions to the system given by some compounded matrix $A\dot{\times}\overset{t}{\ldots}\dot{\times}A$. Note that $A\dot{\times}\overset{t}{\ldots}\dot{\times}A$ is trivially positive, abundant and of rank $t\,\mathrm{rk}(A)$. One can also easily verify that $\sigma(A\dot{\times}\overset{t}{\ldots}\dot{\times}A) = \sigma(A)^t$ and $\mathrm{Vol}\left(\mathcal{P}_{A\dot{\times}\overset{t}{\ldots}\dot{\times}A}\right) = \mathrm{Vol}\left(\mathcal{P}_A\right)^t$. By Corollary 1.8, it therefore follows that

$$\mathbb{E}\Big(X'_{t,n}\Big) = \frac{\mathrm{Vol}\left(\mathcal{P}_{A\dot{\times}\overset{t}{\ldots}\dot{\times}A}\right)}{\sigma(A\dot{\times}\overset{t}{\ldots}\dot{\times}A)} n^{tm - t\mathrm{rk}(A)} p^{tm} \left(1 + o(1)\right) = \mu^t \left(1 + o(1)\right). \tag{2.10}$$

Next, let us consider t-tuples of pairwise disjoint distinct and non-trivial solutions of which at least one is not proper, that is $X''_{t,n}$. This means we are considering all compounded matrices of the form $A_{\mathfrak{p}_1} \dot{\times} \dots \dot{\times} A_{\mathfrak{p}_t}$ where $\mathfrak{p}_i \in \mathfrak{P}(A)$ and at least one of them is not equal to $\{1, \dots, m\}$. Such a compounded matrix has $|\mathfrak{p}_1| + \dots |\mathfrak{p}_t| < tm$ columns and is of rank $t\,\mathrm{rk}(A)$. It follows by Equation (1.3) that

$$\begin{aligned} \mathbb{E}\Big(X''_{t,n}\Big) &= O\Big(\max_{\mathfrak{p}_1,\dots,\mathfrak{p}_t} n^{\sum_{i=1}^t (|\mathfrak{p}_i| - \mathrm{rk}(A))} p^{\sum_{i=1}^t |\mathfrak{p}_i|}\Big) \\ &= O\Big(\max_{\mathfrak{p}_1,\dots,\mathfrak{p}_t} n^{\mathrm{rk}(A)\left(\sum_{i=1}^t |\mathfrak{p}_i|/m - t\right)}\Big) \end{aligned}$$

$$= O(n^{-\mathrm{rk}(A)/m}) = o(1).$$

Here we have used the fact that $|\mathfrak{p}_1| + \ldots + |\mathfrak{p}_t| \leq tm - 1$ since one of the partitions does not come from a proper solution, that is there exists a partition $|\mathfrak{p}_i|$ which satisfies $|\mathfrak{p}_i| < m$.

Let us now consider t-tuples of distinct and non-trivial solution in which at least two solutions share elements. This means we are considering compounded matrices of the form $A_{\mathfrak{p}_1} \times \ldots \times A_{\mathfrak{p}_t}$ where $\mathfrak{p}_i \in \mathfrak{P}(A)$ for all $1 \leq i \leq t$ and for at least one of the compound operators the implied matching is non-empty. We consider in general terms what happens when we compound two admissible systems $A \in \mathbb{Z}^{r_A \times m_A}$ and $B \in \mathbb{Z}_{r_B \times m_B}$: as previously established, $A \dot{\times} B$ has $m_A + m_B$ columns and rank $\mathrm{rk}(A) + \mathrm{rk}(B)$. Now let us assume there is some 'overlap', that is some columns $\emptyset \neq Q \subsetneq [m_B]$ of B are matched to columns of A in the implied bipartite matching. The compounded matrix therefore has $m_A + m_B - |Q|$ columns and is of rank at least $\mathrm{rk}(A) + \mathrm{rk}(B) - r_Q(B)$ where $r_Q(B) = \mathrm{rk}(B) - \mathrm{rk}(B^{\overline{Q}})$. This follows easily since if some rows in the compounded system stemming from B are linearly dependent, then first their coordinates in $\overline{Q}$ have to be linearly dependent on other rows stemming from B. Using the above notation gives the upper bound $r_Q(B)$ for the number of rows that can become linearly dependent by compounding the two matrices.

We know that $A_{\mathfrak{p}_1} \times \ldots \times A_{\mathfrak{p}_t}$ is a matrix with $tm - \beta$ columns of rank $t\,\mathrm{rk}(A) - \alpha$ for some $\alpha, \beta \in \mathbb{N}_0$, where $\beta > 0$. We will induce over $1 \leq i \leq t$ to show that $\beta\,\mathrm{rk}(A)/m > \alpha$. Assume without loss of generality that the first two matrices $A_{\mathfrak{p}_1}$ and $A_{\mathfrak{p}_2}$ overlap in some columns $\emptyset \neq Q \subsetneq [|\mathfrak{p}_2|]$ of $A_{\mathfrak{p}_2}$. It follows that $A_{\mathfrak{p}_1} \times A_{\mathfrak{p}_2}$ has $|\mathfrak{p}_1| + |\mathfrak{p}_2| - |Q|$ columns and, by the previous observation, is of rank at least $2\mathrm{rk}(A) - r_Q(A_{\mathfrak{p}_2})$. Since A is assumed to be strictly balanced, we have by (iii) in Proposition 1.12 that

$$\frac{|Q|}{|Q| - r_Q(A_{\mathfrak{p}_2})} = \frac{|Q|}{|Q| - r_Q} < \frac{m}{m - \mathrm{rk}(A)} \tag{2.11}$$

so that $r_Q(A_{\mathfrak{p}_2}) < \mathrm{rk}(A)|Q|/m$. It follows that

$$\mathrm{rk}(A_{\mathfrak{p}_1} \times A_{\mathfrak{p}_2}) > 2\mathrm{rk}(A) - |Q|\,\mathrm{rk}(A)/m \tag{2.12}$$

and the first step of the induction is complete. Assume therefore that $A_{\mathfrak{p}_1} \times \ldots \times A_{\mathfrak{p}_k}$ has $km - \beta$ variables and is of rank $k\,\mathrm{rk}(A) - \alpha$ where $1 < k < t$ as well as $\beta\,\mathrm{rk}(A)/m > \alpha$

and $\beta > 0$. We compound $A_{\mathfrak{p}_{k+1}}$ with $A_{\mathfrak{p}_1} \times \ldots \times A_{\mathfrak{p}_k}$ where the overlap is again indicated by $Q \subsetneq [|\mathfrak{p}_{k+1}|]$ though now $Q = \emptyset$ is possible. It follows by the same arguments as before that $A_{\mathfrak{p}_1} \times \ldots \times A_{\mathfrak{p}_{k+1}}$ has $(k+1)m - (\beta + |Q|)$ columns and is of rank strictly greater than

$$k \operatorname{rk}(A) - \alpha + \operatorname{rk}(A) - r_Q(A_{\mathfrak{p}_{k+1}}) \geq (k+1) \operatorname{rk}(A) - (\alpha + \operatorname{rk}(A)|Q|/m). \tag{2.13}$$

Here we have used that $\mathfrak{p}_{k+1}$ comes from a non-trivial solution. Obviously we still have

$$\operatorname{rk}(A)(\beta + |Q|)/m > \alpha + \operatorname{rk}(A)|Q|/m \tag{2.14}$$

and therefore the induction is complete.

Now using the fact that $\beta \operatorname{rk}(A)/m > \alpha$ as well as $p = Cn^{-(m-\operatorname{rk}(A))/m}$, we can apply Equation (1.3) to obtain that

$$\mathbb{E}\Big(X'''_{t,n}\Big) = O\left(n^{(tm-\beta)-(t \operatorname{rk}(A)-\alpha)} p^{tm-\beta}\right) = O\left(n^{\alpha - \operatorname{rk}(A)\beta/m}\right) = o(1).$$

Taken together it follows that

$$\mathbb{E}(X_{t,n}) = \mathbb{E}\Big(X'_{t,n}\Big) + \mathbb{E}\Big(X''_{t,n}\Big) + \mathbb{E}\Big(X'''_{t,n}\Big) = \mu^t(1 + o(1)). \tag{2.15}$$

Since $S^{(t)} = \mathbb{E}(X_{t,n})\,/t!$, we can apply Brun's Sieve to deduce the desired statement.

Necessity. It remains to show that the requirement that A is strictly balanced is in fact necessary. If the system is balanced, but not strictly balanced, that is $m(A) = m/(m - \operatorname{rk}(A))$ but there exists some induced submatrix also attaining this value, we again split $X_{t,n}$ into three parts $X'_{t,n} + X''_{t,n} + X'''_{t,n}$ as previously. Observe that $\mathbb{E}\Big(X'_{t,n}\Big) = \mu^t(1 + o(1))$ and $\mathbb{E}\Big(X''_{t,n}\Big) = o(1)$ as before since we did not rely on the fact that the matrix is *strictly* balanced. Further continuing the notation from before, we know that the compounded matrices $A_{\mathfrak{p}_1} \times \ldots \times A_{\mathfrak{p}_t}$ considered in $\mathbb{E}\Big(X'''_{t,n}\Big)$ have $tm - \beta$ variables and are of rank $tr - \alpha$ for some $\alpha, \beta \in \mathbb{N}_0$ where $\beta > 0$. We can again show by induction that $\beta\, r/m \geq \alpha$ since the system is balanced. Note that previously we had a strict inequality. Since by assumption our system is balanced but not strictly balanced, there are compounded matrices for which $\beta\, r/m = \alpha$. Each of these contributes a term of constant order, since

$$\Theta\left(n^{(tm-\beta)-(t \operatorname{rk}(A)-\alpha)} p^{tm-\beta}\right) = \Theta\left(n^{\alpha - \operatorname{rk}(A)\,\beta/m}\right) = \Theta(1). \tag{2.16}$$

Combining these observations, it follows that

$$\mathbb{E}(X_{t,n}) = \mathbb{E}\Big(X'_{t,n}\Big) + \mathbb{E}\Big(X''_{t,n}\Big) + \mathbb{E}\Big(X'''_{t,n}\Big) = \mu^t\,(1 + o(1)) + c_t \tag{2.17}$$

for some appropriate constants $c_t > 0$. For each $n > 0$, fixed s and $0 \leq t \leq s$ the values $\mathbb{E}(X_{t,n})$ are moments of a random variable that satisfy Stieltjes condition [114]. Consequently, for each t the sequence $\mu^t\,(1 + o(1)) + c_t$ converges to the t-th moment of a certain random variable. Due to Carleman's condition, this random variable is indeed uniquely determined. Finally, the limit of the sequence X_n is determined by its moments which differ from those of a Poisson distribution. We conclude that we cannot have convergence in distribution towards a Poisson distributed random variable.

To conclude the analysis, the unbalanced case can be deduced by using a similar argument to that just developed for the balanced case by conveniently rescaling the random variable and showing that it does not converge in distribution to a Poisson random variable. The details are the same as in the proof of Theorem 5 in [119]. ■

2.3 The computation of Vol($\mathcal{P}_A$)

As mentioned in the introduction of this section, computing $\mathrm{Vol}\,(\mathcal{P}_A)$ is in general an algorithmically involved problem. Let us determine some concrete values for the examples introduced in Section 1.7.

***k*-sums.** We start with the easy example of determining the volume of the polytope associated with sum-free sets, that is $k = 1$ and we are interested in $\mathrm{Vol}\,(\mathcal{P}_{A_1})$ where $A_1 = (1 \quad 1 \; -1)$. In this case, the polytope can be described as

$$\mathcal{P}_{A_1} = \{(x_1, x_3) \in \mathbb{R}^2 : 0 \leq x_1 \leq x_3 \leq 1\}. \tag{2.18}$$

Clearly $\mathcal{P}_{A_1}$ is integral, since it is in fact the triangle given by the vertices $(0,0)$, $(0,1)$ and $(1,1)$, and trivially the volume is equal to $1/2$. Let us however obtain this value through an interpolation argument: it follows from Ehrhart's Theorem that $p(n) = |n \cdot \mathcal{P}_{A_1} \cap \mathbb{Z}^2|$ is a polynomial of degree 2 with leading coefficient $\mathrm{Vol}\,(\mathcal{P}_{A_1})$, that is $p(n) = \mathrm{Vol}\,(\mathcal{P}_{A_1})\,n^2 + bn + c$. Clearly $p(0) = |\{(0,0)\}| = 1$, so that $c = 1$ and $p(1) = p(0) + |\{(0,1),(1,1)\}| = 3$, so that $b = 2 - \mathrm{Vol}\,(\mathcal{P}_{A_1})$. It follows that

$$p(n) = \mathrm{Vol}\,(\mathcal{P}_{A_1})\,(n^2 - n) + 2n + 1. \tag{2.19}$$

Finally, since $p(2) = p(1) + |\{(0,2),(1,2),(2,2)\}| = 6$, it follows that $\mathrm{Vol}\,(\mathcal{P}_{A_1}) = 1/2$, as we wanted to show.

The case of k-sum free sets when $k > 1$ is slightly different. Here the matrix is given by $A_k = (1 \quad 1-k)$ and its associated polytope can be described as

$$\mathcal{P}_{A_k} = \{(x_1, x_3) \in \mathbb{R}^2 : 0 \leq kx_3 - x_1 \leq 1,\ 0 \leq x_1, x_3 \leq 1\}, \tag{2.20}$$

which is a parallelogram of area $1/k$. The main difference is that in this case the polytope is not integral, so by Erhart's Theorem we will obtain a pseudo-polynomial.

***k*-term arithmetic progressions.** Let us determine the volume in the case of k-term arithmetic progressions, where we can obtain a closed expression for the volume through elementary means. This family has been studied widely and the following results are also implicitly stated in [131].

Lemma 2.4. *For any integer $k \geq 3$ the number of k-term arithmetic progressions in $\{0, 1, 2, \ldots, n\}$, including trivial ones, is given by*

$$(n+1)\left(\left\lfloor \frac{n}{k-1} \right\rfloor + 1\right) - \frac{k-1}{2}\left(\left\lfloor \frac{n}{k-1} \right\rfloor^2 + \left\lfloor \frac{n}{k-1} \right\rfloor\right) = \frac{1}{2(k-1)} n^2 + O(n).$$

Proof. Observe that any k-term arithmetic progression is of the form $\{a, a+d, \ldots, a+(k-1)d\}$ where $a \in \{0, 1, 2, \ldots, n\}$ and $d \in \{0, 1, 2, \ldots, \lfloor n/(k-1) \rfloor\}$. Additionally, for a given d, we know that $\{0, d, \ldots, (k-1)d)\}$, $\{1, 1+d, \ldots, 1+(k-1)d)\}$, $\{n-(k-1)d, n-(k-2)d \ldots, n\}$ are the only k-term arithmetic progression with common difference d. The total number of k-term arithmetic progression is therefore simply given by $\sum_{d=0}^{\lfloor n/(k-1) \rfloor} n + 1 - (k-1)d$, which evaluates to the desired formula. □

As we have previously seen, the matrix associated with k-term arithmetic progressions is given by

$$A_k = \begin{pmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & & \ddots & & \\ & & & 1 & -2 & 1 \end{pmatrix} \in \mathbb{Z}^{k-2\times k}. \tag{2.21}$$

As an immediate corollary to the previous lemma, we get that $\mathrm{Vol}\,(\mathcal{P}_{A_k}) = 1/(k-1)$, where we have made use of the fact that the lemma already eliminated the symmetry constant $\sigma(A) = 2$.

Sidon and *B_h*[*g*] sets. We start by noting that the matrix associated with $B_h[g]$ sets is given by

$$A_{h,g} = \begin{pmatrix} 1 \overset{h}{\cdots} 1 \; -1 \overset{h}{\cdots} -1 & & & \\ & 1 \overset{h}{\cdots} 1 \; -1 \overset{h}{\cdots} -1 & & \\ & & \cdots & \\ & & & 1 \overset{h}{\cdots} 1 \; -1 \overset{h}{\cdots} -1 \end{pmatrix} \in \mathbb{Z}^{g \times h(g+1)}. \tag{2.22}$$

Let us prove that the polytope associated with $A_{h,g}$ is integral, so that we may apply the interpolation technique we previously demonstrated for sum-free sets.

Proposition 2.5. *The polytope $\mathcal{P}_{A_{h,g}}$ is integral.*

Proof. We recall Equation (1.6) and note that the polytope $\mathcal{P}_{A_{h,g}}$ can be written as

$$\mathcal{P}_{A_{h,g}} = \{\mathbf{x} : A_{h,g} \cdot \mathbf{x}^T \leq \mathbf{0}^T\} \cap \{\mathbf{x} : -A_{h,g} \cdot \mathbf{x}^T \leq \mathbf{0}^T\} \cap [0,1]^m \subset \mathbb{R}^m.$$

so that it is represented by $\{\mathbf{x} \in \mathbb{R}^k : P \cdot \mathbf{x} \geq \mathbf{b}\}$ where

$$P = \begin{pmatrix} A \\ -A \\ I_{h(g+1)} \\ -I_{h(g+1)} \end{pmatrix} \quad \text{and} \quad \mathbf{b} = (0, \overset{2g}{\ldots}, 0, 0, \overset{h(g+1)}{\ldots}, 0, -1, \overset{h(g+1)}{\ldots}, -1). \tag{2.23}$$

here $I_{h(g+1)}$ is the unit matrix of size $h(g+1)$. Note that a polytope associated with a unimodular matrix, that is a matrix where each quadrangular submatrix has determinant either 0 or ±1, is integral [138]. It follows that we only need to prove that P is unimodular. Observe that we can reduce our argument to minors with entries in the topmost part of the matrix, that is A. We argue by induction on the size of the minor: the result is clear for minors of size 1, as the entries of the matrix belong to $\{0, \pm 1\}$. Assume that the result is true for every minor of size at most k, and let us show that the result is also true for k. We will use the fact that every column of A has at most two elements different from 0.

Consider the first row of the minor under study. If all elements are equal to 0, the minor is equal to 0. If there exists a unique element different from 0, we apply induction by developing the determinant along the row. Finally, let us assume that there exist at least two elements different from 0 in the first row. If these two elements are equal, then the corresponding columns are linearly dependent, and the determinant is equal to 0. If these two elements are different, the column containing the 1 otherwise only contains zeros. It follows that we can develop the determinant from this point and apply induction. □

Let us determine the number of solutions in $\{0, 1, 2, \ldots, n\}$ through the inclusion-exclusion method. Given any $k \in \mathbb{N}_0$, write $k = k_1 n + k_2$ where either $0 \leq k_1 \leq h - 1$ and $1 \leq k_2 \leq n$ or $k_1 = 0$ and $k_2 = 0$. The number of solutions of the equation $x_1 + \ldots + x_h = k$ with $x_i \in \{0, \ldots, n\}$ is equal to

$$a(k_1) = \sum_{j=0}^{k_1} (-1)^j \binom{h}{j} \binom{(k_1 - j)n + k_2 - j + h - 1}{h - 1}. \tag{2.24}$$

Here we have used that the number of solutions of $x_1 + \ldots + x_h = k$ where at least j of the x_is are strictly greater than n is equal to

$$\binom{k - (n+1)j + h - 1}{h - 1}. \tag{2.25}$$

Since k is at most hn, the total number of integer points in $n \cdot \mathcal{P}_{A_{h,g}}$ is given by the polynomial

$$p_{h,g}(n) = 1 + \sum_{k_1=0}^{h-1} \sum_{k_2=1}^{n} \left(a(k_1)\right)_{g+1}, \tag{2.26}$$

where we have used the falling factorial

$$\left(a(k_1)\right)_{g+1} = a(k_1)\left(a(k_1) - 1\right) \cdots \left(a(k_1) - g\right). \tag{2.27}$$

The argument we previously used in the case of k-term arithmetic progression does not work here, as the expressions are too involved. However, we can apply an interpolation argument to obtain the dominant term of $p_{h,g}(n)$ as we did for the case of 1-sums: by Proposition 2.5 and Ehrhart's Theorem, $p_{h,g}(n)$ is a polynomial of degree $d = (h-1)(g+1) + 1$ with coefficients $a_0, a_1, \ldots, a_d$. The values $p_{h,g}(0)$, $p_{h,g}(1)$,... , $p_{h,g}(d-1)$ therefore determine $p_{h,g}(n)$ through the Vandermonde-matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^{d-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & d-1 & \cdots & (d-1)^{d-1} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix} = \begin{pmatrix} p_{h,g}(0) \\ p_{h,g}(1) \\ p_{h,g}(2) \\ \vdots \\ p_{h,g}(d-1) \end{pmatrix}. \tag{2.28}$$

By Equation (1.8) we have $\mathrm{Vol}\left(\mathcal{P}_{A_{h,g}}\right) = a_d$. These coefficients can be easily determined, using for example Cramer's rule, and some concrete values are shown in

Figure 2.2.

$h\diagdown g$	1	2	3	4	5
2	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{2}{5}$	$\frac{1}{3}$	$\frac{2}{7}$
3	$\frac{11}{20}$	$\frac{12}{35}$	$\frac{379}{1680}$	$\frac{565}{3696}$	$\frac{6759}{64064}$
4	$\frac{151}{315}$	$\frac{1979}{7560}$	$\frac{40853}{270270}$	$\frac{200267}{2223936}$	$\frac{825643615}{15084957888}$

Table 2.2: *Volumes of polytopes associated with $B_h[g]$ sets.*

Recall that a detailed study for $B_h[1]$ sets was done by Godbole et al. [68] by means of trigonometric sums and Fourier analytic methods. Their result implies that

$$\operatorname{Vol}\left(\mathcal{P}_{A_{h,1}}\right) = 2(h!)^2\kappa_h = \frac{\sum_{j=0}^{h-1}(-1)^j\binom{2h}{j}(h-j)^{2h-1}}{(2h-1)!}. \tag{2.29}$$

Closed formulas for bigger values of g seem to be much more involved.

2.4 Further remarks

The problem considered in this chapter can be rephrased in a more general setting: instead of studying vectors $\mathbf{x}$ satisfying $A \cdot \mathbf{x}^T = \mathbf{0}^T$, one could study those for which $A \cdot \mathbf{x}^T \in \mathcal{Q}^r$ for some given (possibly infinite) sequence of integers $\mathcal{Q}$. The homogeneous case is that of $\mathcal{Q} = \{0\}$. When $A = (1\ -1)$, Sárközy [132] showed that every set with positive upper density contains at least two elements whose difference is a square, see also [105]. It is conjectured that for every $\varepsilon > 0$ there exists a subset of $[n]$ of size $n^{1-\varepsilon}$ whose differences are never a square. Ruzsa [123] proved this conjecture for every $\varepsilon \geq 0.267$.

Some things can be said for the case of this particular matrix that go in the direction of the questions studied in this chapter: if $\mathcal{Q}$ is the sequence of k-th powers, that is $\mathcal{Q} = \{x^k : x \in \mathbb{N}\}$, then if we denote the set of solutions by $S_{\mathcal{Q}}(n) = \{\mathbf{x} = (x_1, x_2) \in [n]^2 : x_1 - x_2 \in \mathcal{Q}\}$, we have

$$\begin{aligned} |S_{\mathcal{Q}}(n)| &= \sum_{q \in \mathcal{Q}(n)} (n-q) = n|\mathcal{Q} \cap [n]| - \sum_{q \in \mathcal{Q} \cap [n]} q \\ &= \int_0^n x^{1/k} dx = \frac{k}{k+1} n^{1+1/k}(1+o(1)) \end{aligned}$$

by Abel's summation formula. Following the ideas of the proof of Theorem 2.1, one can easily verify that $p = n^{-(k+1)/(2k)}$ is a threshold for the property that $[n]_p$ contains some x_1, x_2 with $x_1 - x_2 \in \mathcal{Q}$.

Chapter 3

The Resilience Threshold

Given a matrix $A \in \mathbb{Z}^{r \times m}$, a set of integers $T \subset \mathbb{N}$ and some $c \in \mathbb{N}$, we write

$$T \rightarrow_c A \tag{3.1}$$

if any c-coloring of T contains a monochromatic proper solution to $A \cdot \mathbf{x}^T = \mathbf{0}^T$, that is if for every finite partition $T_1, \ldots, T_c$ of T there exists $1 \leq i \leq c$ such that $T_i^m \cap S_0(A, \mathbf{0}) \neq \emptyset$. Rödl and Ruciński [115] established the threshold for a version of Rado's Theorem in the binomial random set that was later completed by Friedgut, Rödl and Schacht [66]. Combined, their results state that the function $p_0(n) = n^{-1/m_1(A)}$ is a sharp threshold for the property that $[n]_p \rightarrow_c A$ for any partition regular matrix $A \in \mathbb{Z}^{r \times m}$. Here $m_1(A)$ is the maximum 1-density introduced in Section 1.5.

The first result of this chapter extends this result to include non-trivial solutions as described in Section 1.4. Let us extend the notation: given a set of integers T and some integer $c \in \mathbb{N}$, we write

$$T \rightarrow_c^\star A \tag{3.2}$$

if for every finite partition $T_1, \ldots, T_c$ of T there exists $1 \leq i \leq c$ such that $T_i^m \cap S_1(A, \mathbf{0}) \neq \emptyset$, that is any c-coloring of T must contain a monochromatic *non-trivial* solution to $A \cdot \mathbf{x}^T = \mathbf{0}^T$. The following result establishes that the threshold for the property that $[n]_p \rightarrow_c^\star A$ is of the same order as the threshold for the property that $[n]_p \rightarrow_c A$.

Theorem 3.1. *For every partition regular matrix $A \in \mathbb{Z}^{r \times m}$ and $c \in \mathbb{N}$ the function $p_0(n) = n^{-1/m_1(A)}$ is a sharp threshold for the property that $[n]_p \rightarrow_c^\star A$.*

Note that the 1-statement in Theorem 3.1 already follows from the previously mentioned results. However, as this proof is quite involved, we will re-prove it here using an alternative approach based on the ideas behind Nenadov and Steger's short proof of a sparse random Ramsey's Theorem [110]. This approach combines the recently developed hypergraph container framework by Balogh, Morris and Samotij [3] as well as Saxton and Thomason [135] with a supersaturation result of Frankl, Graham and Rödl [58]. The proof of Theorem 3.1 is stated in Section 3.3.

Schacht [136] as well as independently Conlon and Gowers [34] also stated a version of Szémeredi's Theorem in sparse random sets. Schacht also extended it to density regular systems as well as sum-free sets. Given a set of integers T and $\varepsilon > 0$, we write

$$T \rightarrow_{\varepsilon} A \tag{3.3}$$

if every subset S for which $|S|/|T| \geq \varepsilon$ satisfies $S^m \cap S_0(A) \neq \emptyset$, that is any subset of T of density at least ε contains a proper solution to $A \cdot \mathbf{x}^T = \mathbf{0}^T$. The second goal of this note is to extend Schacht's statement to the broadest sensible group of matrices, that is that of abundant matrices. We will again include non-trivial solutions in this consideration. Before we can state it, we will need to introduce some additional notation.

Given some matrix $A \in \mathbb{Z}^{r \times m}$, let $\operatorname{ex}(n, A)$ be the size of the largest subset of $[n]$ not containing a proper solution and define $\pi(A) = \limsup_{n \to \infty} \operatorname{ex}(n, A)/n$. Observe the clear parallels to the Turán number of a graph. Clearly density regular systems satisfy $\pi(A) = 0$ and for other systems systems we have $\pi(A) > 0$. One can easily bound this value away from 1, as we will later see in Lemma 3.3. Now, given a set of integers T and some $\varepsilon > 0$, we write

$$T \rightarrow_{\varepsilon}^{\star} A \tag{3.4}$$

if every subset S for which $|S|/|T| \geq \varepsilon$ also satisfies $S^m \cap S_1(A, \mathbf{0}) \neq \emptyset$, that is any subset of T of density at least ε contains a non-trivial solution to $A \cdot \mathbf{x}^T = \mathbf{0}^T$. These definitions and observations allow us to state the following result.

Theorem 3.2. *For every positive and abundant matrix $A \in \mathbb{Z}^{r \times m}$ and $\varepsilon > \pi(A)$ the function $p_0(n) = n^{-1/m_1(A)}$ is a sharp threshold for the property that $[n]_p \rightarrow_{\varepsilon}^{\star} A$.*

To prove the 1-statement in Theorem 3.2, we will derive a supersaturation result from a removal lemma due to Král', Serra and Vena [98] and combine it with a corollary of the hypergraph containers due to Balogh, Morris and Samotij [3]. We will in fact prove the existence of a proper solution above the threshold, so the statement of Theorem 3.2 does not change if one instead considers the property that $[n]_p \to_\varepsilon A$. The 0-statement will be established through the usual approach. The proof of Theorem 3.2 is given in Section 3.2.

Let us make a remark regarding the notion of resilience, see also [148]: a set of integers or a graph resiliently posses some property, if even after removing a certain number of elements or edges, the property still holds. Szemerédi's Theorem can therefore be interpreted as a statement about how resiliently $[n]$ has the property of containing arithmetic progressions. Theorem 3.2 can likewise be considered as a statement about how resiliently a typical subset of $[n]$ of density p has the property of containing a solution to $A \cdot \mathbf{x}^T = \mathbf{0}^T$. Hancock, Staden and Treglown [78], in simultaneous and independent work, not only also obtained Theorem 3.2, but in fact a broader statement that can be considered as a resilience version of Theorem 3.1, though in both cases their results are restricted to proper solutions. Their approach likewise consists of combining hypergraph containers with supersaturation results.

3.1 Preliminaries

Given some matrix $A \in \mathbb{Z}^{r \times m}$, we have previously defined $\mathrm{ex}(n, A)$ to be the size of the largest subset of $[n]$ not containing a proper solution and

$$\pi(A) = \limsup_{n \to \infty} \mathrm{ex}(n, A)/n.$$

Unfortunately, unlike the Erdős–Stone–Simonovits Theorem [53, 52] in the graph case, no good characterization of $\pi(A)$ is known for arbitrary matrices A, see also Section 1.7. However, the following lemma shows that one can still easily bound this value away from 1 for every positive matrix.

Lemma 3.3 (Folklore). *Every positive matrix $A \in \mathbb{Z}^{r \times m}$ satisfies $\pi(A) < 1$.*

Proof. By Lemma 1.4, there exists some $\mathbf{x} = (x_1, \dots, x_m) \in S_0(A, \mathbf{0}) \cap \mathbb{N}^m$. Clearly we also have $j \cdot \mathbf{x} = (jx_1, \dots, jx_m) \in S_0(A, \mathbf{0}) \cap \mathbb{N}^m$ for any $j \geq 1$. Now for $n \geq m \max_i(x_i)$ we observe that every $i \in [n]$ can appear in at most m of the $J = \lceil n/\max_i(x_i) \rceil$ solutions $\mathbf{x}, 2\cdot\mathbf{x}, \dots, J\cdot\mathbf{x} \in [n]^m$, so every subset of $[n]$ that avoids $S_0(A, \mathbf{0})$ is missing at least J/m elements. It follows that $\pi(A) \leq (n - J/m)/n \leq 1 - 1/(m \max_i(x_i)) < 1$. □

3.1.1 Removal Lemma and Supersaturation Results

A common ingredient to proving results in the sparse random setting are robust versions of the deterministic statement, referred to as supersaturation results. In the graph setting such a result is folklore and easy to prove. A counterpart in our setting is for example Varnavides' robust version of Szemerédi's Theorem [158], which states that a set of positive density contains not just one, but a positive proportion of all k–term arithmetic progressions. Frankl, Graham and Rödl [58] formulated such results both for partition and density regular systems.

Lemma 3.4 (Theorem 1 in [58])**.** *For a given partition regular matrix $A \in \mathbb{Z}^{r \times m}$ and $s \in \mathbb{N}$ there exists $\zeta = \zeta(A, s) > 0$ such that for n large enough and for any partition $T_1, \dots, T_s$ of $[n]$, we have*

$$|S_0(A, \mathbf{0}) \cap T_1^m| + \dots + |S_0(A, \mathbf{0}) \cap T_s^m| \geq \zeta \, |S_0(A, \mathbf{0}) \cap [n]^m|. \tag{3.5}$$

Lemma 3.5 (Theorem 2 in [58])**.** *For a given density regular matrix $A \in \mathbb{Z}^{r \times m}$ and $\delta > 0$ there exists $\zeta = \zeta(A, \delta) > 0$ such that, for n large enough, any subset $T \subseteq [n]$ satisfying $|T| \geq \delta n$ also satisfies*

$$|S_0(A, \mathbf{0}) \cap T^m| \geq \zeta \, |S_0(A, \mathbf{0}) \cap [n]^m|. \tag{3.6}$$

We will extend Lemma 3.5 to cover the scope of this note by using an Arithmetic Removal Lemma. Green [70] first formulated such a statement for linear equations in an abelian group. Later Shapira [142] as well as independently Král', Serra and Vena [98] proved a removal lemma for linear maps in finite fields. We will state it here in a simplified version.

Theorem 3.6. *Let $\mathbb{F}_q$ be the finite field of order q. Let X be a subset of $\mathbb{F}_q$ and $A \in \mathbb{F}_q^{r \times m}$ a matrix of full rank. For $\mathcal{S} = \{\mathbf{x} \in \mathbb{F}_q^m : A \cdot \mathbf{x}^T = \mathbf{0}^T\}$ and every $\varepsilon > 0$ there exists an $\eta = \eta(\varepsilon, r, m)$ such that if $|\mathcal{S} \cap X^m| < \eta \, |\mathcal{S}|$ then there exists a set $X' \subset X$ with $|X'| < \varepsilon q$ and $\mathcal{S} \cap (X \setminus X')^m = \emptyset$.*

Applying this result, we formulate the following extension of Lemma 3.5. Note that one could also obtain this result through a direct application of a removal lemma for colored hypergraphs as for example Theorem 2 in [98].

Lemma 3.7. *For any positive matrix $A \in \mathbb{Z}^{r \times m}$ and $\delta > \pi(A)$ there exists $\zeta = \zeta(\delta, A) > 0$ and $n_0 = n_0(\delta, A)$ such that, for $n \geq n_0$, any subset $T \subseteq [n]$ satisfying*

$|T| \geq \delta n$ *also satisfies*

$$|S_0(A, \mathbf{0}) \cap T^m| \geq \zeta \, |S_0(A, \mathbf{0}) \cap [n]^m|. \tag{3.7}$$

Proof. Let $p = p(A, n)$ be a prime number between $2mn \max(|A|)$ and $4mn \max(|A|)$ and $\mathbb{F}_p$ the finite field with p elements. Here $\max(|A|)$ refers to the maximal absolute entry in A. Note that such a prime number exists for example because of the Bertrand–Chebyshev Theorem. We have $\mathbb{F}_p \cong \mathbb{Z}_p$ and we can identify the integers with their corresponding residue classes in $\mathbb{F}_p$. The matrix A now defines a map from $\mathbb{F}_p^m$ to $\mathbb{F}_p^r$. A solution in $S(A, \mathbf{0})$ clearly lies in $\mathcal{S}$ and, as we have chosen p large enough, all canonical representatives from $\mathcal{S} \cap [n]^m$ also lie in $S(A, \mathbf{0}) \cap [n]^m$ for $n \geq \max |A|$.

Next, set $\delta' = (\delta + \pi(A))/2$ and let n be large enough such that any subset of density at least δ' in $[n]$ contains a proper solution. Note that $\delta > \delta' > \pi(A)$. Given a subset $T \subseteq [n]$ satisfying $|T| \geq \delta n$ consider the corresponding set X of residue classes in $\mathbb{F}_p$. One needs to remove at least $(\delta - \delta')n$ elements from T in order for T^m to avoid $S_0(A, \mathbf{0})$ in $[n]$, so one needs to remove at least an

$$\varepsilon = \frac{(\delta - \delta')n}{q} \geq \frac{(\delta - \delta')}{4m \max(|A|)} > 0$$

proportion of elements in $\mathbb{F}_p$ from X so that X^m avoids $\mathcal{S}$ in $\mathbb{F}_p$. It follows from Theorem 3.6 that $|\mathcal{S} \cap X^m| \geq \eta |\mathcal{S}|$ for some $\eta = \eta(\varepsilon, r, m)$. Since we have chosen p large enough, it follows that T contains at least an η proportion of $S(A, \mathbf{0}) \cap [n]^m$. An easy consequence of Equation (1.3) and Lemma 1.4 is that $\lim_{n \to \infty} | \, S_0(A, \mathbf{0}) \cap [n]^m| / | \, S(A, \mathbf{0}) \cap [n]^m| \geq c_0$ for $c_0 = c_0(A) > 0$ as given by Lemma 1.4. It follows that the result holds for n large enough and $\zeta = \zeta(\delta, A) = (c_0 \, \eta)/2$. □

3.1.2 Hypergraph Containers

The development of hypergraph containers by Balogh, Morris and Samotij [3] as well as independently Thomason and Saxton [135] has opened a new, easy and unified framework to proving results in the sparse random setting. Let us start by stating the Hypergraph Container Theorem as given by Balogh, Morris and Samotij.

Given a hypergraph $\mathcal{H}$ we denote its vertex set by $V(\mathcal{H})$ and its set of hyperedges by $E(\mathcal{H})$. The cardinality of these sets will be respectively denoted by $v(\mathcal{H})$ and $e(\mathcal{H})$. Given some subset of vertices $A \subseteq V(\mathcal{H})$, we denote the subgraph it induces in $\mathcal{H}$ by $\mathcal{H}[A]$ and its degree by $\deg_{\mathcal{H}}(A) = |\{e \in E(\mathcal{H}) : A \subseteq e\}|$. For $\ell \in \mathbb{N}$ we denote the maximum ℓ-degree by $\Delta_\ell(\mathcal{H}) = \max\{\deg_{\mathcal{H}}(A) : A \subseteq V(\mathcal{H}) \text{ and } |A| = \ell\}$. Let the set of independent vertex sets in $\mathcal{H}$ be denoted by $\mathcal{I}(\mathcal{H})$. Lastly, let $\mathcal{H}$ be a uniform hypergraph, $\mathcal{F}$ an increasing family of subsets of $V(\mathcal{H})$ and $\varepsilon > 0$. We say that $\mathcal{H}$ is $(\mathcal{F}, \varepsilon)$-dense if $e(\mathcal{H}[A]) \geq \varepsilon\, e(\mathcal{H})$ for every $A \in \mathcal{F}$.

Theorem 3.8 (Theorem 2.2 in [3])**.** *For every $m \in \mathbb{N}$, $c > 0$ and $\varepsilon > 0$, there exists a constant $C = C(m, c, \varepsilon) > 0$ such that the following holds. Let $\mathcal{H}$ be an m-uniform hypergraph and let $\mathcal{F} \subseteq 2^{V(\mathcal{H})}$ be an increasing family of sets such that $|A| \geq \varepsilon v(\mathcal{H})$ for all $A \in \mathcal{F}$. Suppose that $\mathcal{H}$ is $(\mathcal{F}, \varepsilon)$-dense and $p \in (0, 1)$ is such that*

$$\Delta_\ell(\mathcal{H}) \leq c\, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})}. \tag{3.8}$$

for every $\ell \in \{1, \dots, k\}$. Then there exists a family $\mathcal{T} \subseteq \binom{V(\mathcal{H})}{\leq Cp\, v(\mathcal{H})}$ and functions $f : \mathcal{T} \to \overline{\mathcal{F}}$ and $g : \mathcal{I}(\mathcal{H}) \to \mathcal{T}$ such that for every $I \in \mathcal{I}(\mathcal{H})$,

$$g(I) \subseteq I \quad \text{and} \quad I \setminus g(I) \subseteq f(g(I)). \tag{3.9}$$

The statement gives the existence of a small number of containers $\overline{\mathcal{F}}$ and some fingerprints $\mathcal{T}$ so that every independent set I in $\mathcal{H}$ is identified with a fingerprint $g(I)$ that determines a container $f(g(I))$ which contains $I \setminus g(I)$.

Next, let $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$ be a sequence of m-uniform hypergraphs and let $\alpha \in [0, 1)$. We say that $\mathcal{H}$ is α-dense if for every $\delta > 0$, there exist some $\varepsilon > 0$ such that for $U \subseteq V(\mathcal{H}_n)$ which satisfies $|U| > (\alpha + \delta)\, v(\mathcal{H}_n)$ we have $e(\mathcal{H}_n[U]) > \varepsilon\, e(\mathcal{H}_n)$ for n large enough. Balogh, Morris and Samotij proved the following consequence of their container statement.

Theorem 3.9 (Theorem 5.2 in [3])**.** *Let $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$ be a sequence of m-uniform hypergraphs, $\alpha \in [0, 1)$ and let $C > 0$. Suppose that $q = q(n)$ is a sequence of probabilities such that for all sufficiently large n and for every $\ell \in \{1, \dots, m\}$ we have*

$$\Delta_\ell(\mathcal{H}_n) \leq Cq(n)^{\ell-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}. \tag{3.10}$$

If $\mathcal{H}$ is α-dense, then for every $\delta > 0$, there exists a constant $c = c(C, \alpha, m) > 0$ such that if $p(n) > c\, q(n)$ and $p(n)v(\mathcal{H}_n) = \omega(1)$ as n tends to infinity, then asymptotically almost surely

$$\alpha\Big(\mathcal{H}_n[V(\mathcal{H}_n)_{p(n)}]\Big) \leq (\alpha + \delta)p(n)v(\mathcal{H}_n). \tag{3.11}$$

We will make use of this statement in order to obtain a proof for the 1-statement of Theorem 3.2. For a proof of the 1-statement of Theorem 3.1 such a ready-made statement does not exist and we will follow Nenadov and Steger's [110] short proof of a sparse Ramsey statement by applying Theorem 3.8.

3.2 Proof of Theorem 3.2 – Sparse density regularity

We split the proof up into two separate arguments for the 0- and the 1-statement. The former follows through an alteration argument using some elementary tools form Probability Theory established in Section 1.6 and the later follows from Theorem 3.9.

0-statement. By Corollary 1.13 we may assume that A is strictly 1-balanced, as we can otherwise replace it with B_1 as given by the corollary. Due to Lemma 1.10, we again know that

$$\mathbb{P}\left([n]_p \not\to^{\star}_{\varepsilon} A\right) \leq \mathbb{P}\left(\bigcup_{\mathfrak{p} \in \mathfrak{P}(A)} \Big([n]_p \not\to_{\varepsilon} A_{\mathfrak{p}}\Big) \right) \leq \sum_{\mathfrak{p} \in \mathfrak{P}(A)} \mathbb{P}\left([n]_p \not\to_{\varepsilon} A_{\mathfrak{p}}\right). \tag{3.12}$$

We will therefore analyze the individual probabilities $\mathbb{P}\left([n]_p \not\to_{\varepsilon} A_{\mathfrak{p}}\right)$ for each $\mathfrak{p} \in \mathfrak{P}(A)$. The constant $c = c(A, \varepsilon)$ will be define later in Equation (3.13). We start by first stating the following two propositions. The first deals with non-abundant matrices and the second restricts the statement of Theorem 3.2 to proper solutions. Before we state their proofs, we will show that the 0-statement of Theorem 3.2 follows easily from them.

Proposition 3.10. *For every positive and non-abundant matrix $A \in \mathbb{Z}^{r \times m}$ and $\varepsilon > 0$ we have $\mathbb{P}\left([n]_p \not\to^{\star}_{\varepsilon} A\right) = o(1)$ for any $p(n) = o(1)$.*

Proposition 3.11. *For every positive and abundant matrix $A \in \mathbb{Z}^{r \times m}$ and $\varepsilon > \pi(A)$ there exists $c = c(A, \varepsilon)$ such that $\mathbb{P}\left([n]_p \not\to_{\varepsilon} A\right) = o(1)$ if $p(n) \leq c\, n^{-1/m_1(A)}$.*

For $|\mathfrak{p}| < m$ we note that $A_{\mathfrak{p}}$ clearly is positive. If $A_{\mathfrak{p}}$ is non-abundant, then Proposition 3.10 states that $\mathbb{P}\left([n]_p \not\to_{\varepsilon} A_{\mathfrak{p}}\right) = o(1)$ for $p = p(n) \leq c\, n^{-1/m_1(A)} = o(1)$ independent of the constant c. If $A_{\mathfrak{p}}$ is abundant, then we can apply Proposition 3.11 to it. We have $n^{-1/m_1(A)} = o\Big(n^{-1/m_1(A_{\mathfrak{p}})}\Big)$ and therefore $\mathbb{P}\left([n]_p \not\to_{s} A_{\mathfrak{p}}\right) = o(1)$ for $p = p(n) \leq c\, n^{-1/m_1(A)}$ independent of c. Lastly, let $|\mathfrak{p}| = m$, that is $\mathfrak{p} = \{\{1\}, \dots, \{m\}\}$ and therefore $A_{\mathfrak{p}} = A$. Proposition 3.11 applies to A and therefore we obtain the desired statement with $c = c(A, \varepsilon)$ as given by Proposition 3.11.

Proof of Proposition 3.10. Since A is non-abundant but positive, Lemma 1.2 establishes that we may, without loss of generality, assume that the first row of A is of the shape $\mathbf{a} = (a_1, \ldots, a_m)$ where $a_i = 0$ for all $i \in \{3, \ldots, m\}$ and $a_1, a_2 \neq 0$ as well as $a_1 + a_2 \neq 0$. Any solution $\mathbf{x} = (x_1, \ldots, x_m) \in S(A, \mathbf{0}) \cap [n]_p^m$ therefore has to satisfy $a_1 x_1 + a_2 x_2 = 0$. It follows that we can upper bound the likelihood that $\mathbb{P}\left([n]_p \to_\varepsilon^\star A\right)$ from above by the probability that any subsets of $[n]_p$ of density at least ε contains two elements x_1 and x_2 satisfying $a_1 x_1 + a_2 x_2 = 0$. By Equation (1.3) as well as the linearity of expectation, the expected number of such tuples is at most np^2 while $\mathbb{E}(|[n]_p|) = np$. If $np = O(1)$ then the expected number of such tuples goes to zero, so that the result trivially holds by Markov's Inequality. If $np = \omega(1)$ then by Chernoff's Bound we have $|[n]_p| \geq np/2$ asymptotically almost surely. Since $p = o(1)$, the expected number of such tuples is $o(np/2)$ and therefore, by Markov's Inequality, for any given set of positive density we can remove one element per solution and still asymptotically almost surely have a solution-free set of that same density. This proves the desired result. □

Proof of Proposition 3.11. Following the alteration method as used for example by Schacht [136], we make three case distinctions. Note that we need to cover the whole range of $0 \leq p(n) \leq c\, n^{-1/m_1(A)}$ since we are not dealing with a monotone property.

Case 1. Assume that $p = o(n^{-1/m(A)})$. By Corollary 1.13, we may assume that A is strictly balanced (but not necessarily strictly 1-balanced), as we can otherwise replace it with B as given by the corollary. By Equation (1.3) we have

$$\mathbb{E}\Big(|S_0(A, \mathbf{0}) \cap [n]_p^m|\Big) \leq n^{m - \mathrm{rk}(A)}\, p^m = o(1).$$

Here we have used the assumption that A is strictly balanced. Markov's Inequality therefore implies that $\mathbb{P}\left(|S_0(A, \mathbf{0}) \cap [n]_p^m| \neq 0\right) = o(1)$, see also Chapter 2. It clearly follows that we also have $\mathbb{P}\left([n]_p \to_\varepsilon A\right) = o(1)$ for any $\varepsilon > 0$.

Case 2. Assume that $p = o(n^{-1/m_1(A)})$ but also $p = \omega(n^{-1})$. By Corollary 1.13, we may assume that A is strictly 1-balanced, as we can otherwise replace it with B_1 as given by the corollary. Since $np = \omega(1)$, we have $|[n]_p| \geq np/2$ asymptotically almost surely due to Chernoff's Bound. The expected number of solutions in $[n]_p$ now is asymptotically smaller than the number of elements, since by Equation (1.3) we have

$$\mathbb{E}\Big(|S_0(A, \mathbf{0}) \cap [n]_p^m|\Big) \leq n^{m - \mathrm{rk}(A)}\, p^m = np\left(n^{1/m_1(A)}\, p\right)^{m-1} = o(np/2).$$

Here we have used the assumption that A is strictly 1-balanced. It follows by Markov's

Inequality that for any subset of $[n]_p$ of positive density $\varepsilon > 0$ we can remove one element per solution contained in this subset, so that the resulting set is free of solutions while asymptotically almost surely still having positive density ε in $[n]_p$. It follows that $\mathbb{P}\left([n]_p \to_\varepsilon A\right) = o(1)$.

Case 3. Lastly, assume that $p \leq cn^{-1/m_1(A)}$ but also $p = \omega(n^{-1/m(A)})$, where $c = c(A, \varepsilon)$ will be given in Equation (3.13). By Corollary 1.13, we may again assume that A is strictly 1-balanced, as we can otherwise replace it with B_1 as given by the corollary. Due to Chernoff's Bound, we again have $|[n]_p| \geq np/2$ asymptotically almost surely. Let X_n denote the random variable counting the number of proper solutions in $[n]_p$, that is $X_n = |S_0(A, \mathbf{0}) \cap [n]_p^m|$ for $n \in \mathbb{N}$. For

$$c = c(A, \varepsilon) = \left(\frac{1-\varepsilon}{4}\right)^{1/(m-1)} \tag{3.13}$$

it follows by Equation (1.3) that

$$\mathbb{E}(X_n) \leq n^{m-\mathrm{rk}(A)}\, p^m \leq np \left(n^{1/m_1(A)}\, p\right)^{m-1} \leq (1-\varepsilon)\, np/4.$$

Here we have used the assumption that A is strictly 1-balanced. By Lemma 1.4, there also exists some $c_0 > 0$ such that

$$\mathbb{E}(X_n) \geq c_0\, n^{m-\mathrm{rk}(A)}\, p^m.$$

Following exactly the same ideas as in the proof of the 1-statement of Theorem 2.1, we can in fact show that X_n is concentrated around its expected value. As it was shown there, we have

$$\Delta_n = O\left(n^{m-\mathrm{rk}(A)} p^m\right)^2 \sum_{\emptyset \neq Q \subsetneq [m]} \left(\frac{n^{-(|Q|-r_Q)/|Q|}}{p}\right)^{|Q|} = o\left(\mathbb{E}(X_n)^2\right)$$

where we have used that $(|Q| - r_Q)/|Q| \geq 1/m(A)$ so that $p = \omega(n^{-1/m(A)})$ also implies that $p = \omega(n^{-(|Q|-r_Q)/|Q|})$ for any non-empty $Q \subseteq \{1, \dots, m\}$. Chebyshev's Inequality therefore gives us $\mathbb{P}\left(|X_n - \mathbb{E}(X_n)| \geq \mathbb{E}(X_n)\right) = o(1)$, so that

$$\left| S_0(A, \mathbf{0}) \cap [n]_p^m \right| \leq 2\, \mathbb{E}(X) \leq (1-\varepsilon)\, np/2$$

asymptotically almost surely. It follows that, given a set of density ε, we can remove one element from $[n]_p$ for each solution in $S_0(A,\mathbf{0}) \cap [n]_p^m$ and asymptotically almost surely still be left with a set of density ε, so that $\mathbb{P}([n]_p \to_\varepsilon A) = o(1)$. □

1-statement. Let $\mathcal{H}_n$ be the hypergraph with vertex set $V(\mathcal{H}_n) = [n]$ and edge multiset

$$E(\mathcal{H}_n) = \Big\{\Big\{\{x_1,\dots,x_m\} : (x_1,\dots,x_m) \in S_0(A,\mathbf{0}) \cap [n]^m\Big\}\Big\}. \tag{3.14}$$

Observe that H_n can be a multigraph, that is multiple edges are allowed, but the multiplicity of each edge is clearly bounded by the factorial of m. We do this to simplify counting, since this way we have $|E(\mathcal{H}_n)| = |\, S_0(A,\mathbf{0}) \cap [n]^m|$. Note that in Lemma 3.7 we are intentionally ignoring the symmetry constant $\sigma(A)$ introduced in Section 1.3 as it is not of relevance in this context. We observe that we can limit ourselves to proper solutions when proving the 1-statement.

Lemma 3.7 now states that $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$ is $\pi(A)$-dense. The statement of Theorem 3.2 follows if Theorem 3.9 can be applied. In order to apply Theorem 3.9, it remains to determine a sequence $q = q(n)$ satisfying the required condition. The following lemma gives us upper bounds for the maximum ℓ-degrees in $\mathcal{H}_n$.

Lemma 3.12. *For any $1 \leq \ell \leq m$ we have*

$$\Delta_\ell(\mathcal{H}_n) \leq \ell!\, m^\ell \max_{\substack{Q\subseteq[m]\\ |Q|=\ell}} n^{(m-\mathrm{rk}(A))-(|Q|-r_Q)}. \tag{3.15}$$

Proof. For $\mathcal{H} = (\mathcal{H}_n)$ as defined above and $\ell \in \{1,\dots,m\}$ we have

$$\begin{aligned}
\Delta_\ell(\mathcal{H}_n) &\leq \max_{x_1,\dots,x_\ell\in[n]} \left|\{\mathbf{x} \in S_0(A,\mathbf{0}) \cap [n]^m : \exists Q,\pi \text{ s.t. } \mathbf{x}^Q = (x_{\pi(1)},\dots,x_{\pi(\ell)})\}\right| \\
&\leq \ell! \binom{m}{\ell} \max_{\substack{(x_1,\dots,x_\ell)\in[n]^\ell\\ Q\subseteq[m],|Q|=\ell}} \left|\{\mathbf{x} \in [n]^{m-\ell} \mid A^{\overline{Q}} \cdot \mathbf{x}^T = -A^Q \cdot (x_1,\dots,x_\ell)^T\}\right| \\
&\leq \ell!\, m^\ell \max_{\substack{Q\subseteq[m]\\ |Q|=\ell}} \max_{\mathbf{b}\in\mathbb{Z}^r} \left| S(A^{\overline{Q}},\mathbf{b}) \cap [n]^m \right| \leq \ell!\, m^\ell \max_{\substack{Q\subseteq[m]\\ |Q|=l}} n^{|\overline{Q}|-\mathrm{rk}(A^{\overline{Q}})} \\
&= \ell!\, m^\ell \max_{\substack{Q\subseteq[m]\\ |Q|=\ell}} n^{(m-\mathrm{rk}(A))-(|Q|-r_Q)}.
\end{aligned}$$

In the first inequality we have $Q \subseteq [m]$ and $\pi \in \mathrm{Sym}(\ell)$. We have also made extensive use of the notation defined in Section 1.5 as well as as the trivial upper bound for the number of solutions stated in Equation (1.3). □

Note that $r_Q = 0$ for any $Q \subseteq \{1, \ldots, m\}$ satisfying $|Q| = 1$ due to Lemma 1.14. There also exists $c_0 = c_0(A) > 0$ such that $e(\mathcal{H}_n) \geq c_0\, n^{m-\mathrm{rk}(A)}$ due to Lemma 1.4. Using Lemma 3.12, we therefore observe that

$$\Delta_1(\mathcal{H}_n) \leq m\, n^{m-\mathrm{rk}(A)-1} \leq m/c_0\, \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.$$

For $\ell \in \{2, \ldots, m\}$ we again apply Lemma 3.12 to see that

$$\begin{aligned}
\Delta_\ell(\mathcal{H}_n) &\leq \ell!\, m^\ell \max_{Q\subseteq[m],\, |Q|=\ell} n^{(m-\mathrm{rk}(A))-(|Q|-r_Q)} \\
&= \ell!\, m^\ell \left(\max_{Q\subseteq[m],\, |Q|=\ell} n^{-\frac{|Q|-r_Q-1}{|Q|-1}} \right)^{\ell-1} n^{m-\mathrm{rk}(A)-1} \\
&\leq \ell!\, m^\ell \left(n^{-1/m_1(A)} \right)^{\ell-1} n^{m-\mathrm{rk}(A)-1} \\
&\leq (\ell!\, m^\ell)/c_0 \left(n^{-1/m_1(A)} \right)^{\ell-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.
\end{aligned}$$

Lastly we observe that $n^{-1/m_1(A)}\, v(\mathcal{H}_n) = n^{1-1/m_1(A)}$ tends to infinity since $m_1(A) > 1$. It follows that the prerequisites for Theorem 3.9 hold for $C = (m!\, m^m)/c_0$, $q = q(n) = n^{-1/m_1(A)}$ and we can choose the $c = c(A, \varepsilon)$ in Theorem 3.2 to be equal to the $c = c(C, \pi(A), m)$ as given by Theorem 3.9. ■

3.3 Proof of Theorem 3.1 – Sparse partition regularity

We split the proof up into two separate arguments for the 0- and the 1-statement. The former will follow from the result of Rödl and Ruciński and the later will be established along the ideas of Nenadov and Steger.

0-statement. By Corollary 1.13 we may assume that A is strictly 1-balanced, as we can otherwise replace it with B_1 as given by the corollary. Due to Lemma 1.10, we know that

$$\mathbb{P}\left([n]_p \to_s^\star A\right) \leq \mathbb{P}\left(\bigcup_{\mathfrak{p}\in\mathfrak{P}(A)} \left([n]_p \to_s A_\mathfrak{p} \right) \right) \leq \sum_{\mathfrak{p}\in\mathfrak{P}(A)} \mathbb{P}\left([n]_p \to_s A_\mathfrak{p}\right). \tag{3.16}$$

Let us bound the individual probabilities $\mathbb{P}\left([n]_p \to_s A_\mathfrak{p}\right)$ for each $\mathfrak{p} \in \mathfrak{P}(A)$. For $|\mathfrak{p}| = m$, that is $\mathfrak{p} = \{\{1\}, \ldots, \{m\}\}$, we know due to Rödl and Ruciński that there exists a $c = c(A, s)$ such that $\mathbb{P}\left([n]_p \to_s A\right) = o(1)$ when $p = p(n) \leq c\, n^{-1/m_1(A)}$. For $|\mathfrak{p}| < m$ we consider two separate cases: if $A_\mathfrak{p}$ is not partition regular, then $[n] \not\to_s A$

and therefore trivially $\mathbb{P}\left([n]_p \to_s A_\mathfrak{p}\right) = 0$. If $A_\mathfrak{p}$ is partition regular, then

$$m_1(A_\mathfrak{p}) \geq \frac{|\mathfrak{p}| - 1}{|\mathfrak{p}| - \mathrm{rk}(A_\mathfrak{p}) - 1} > \frac{m-1}{m - \mathrm{rk}(A) - 1} = m_1(A)$$

so that $n^{-1/m_1(A)} = o\Big(n^{-1/m_1(A_\mathfrak{p})}\Big)$ and therefore again due to Rödl and Ruciński for $p = p(n) \leq c\,n^{-1/m_1(A)}$ we have $\mathbb{P}\left([n]_p \to_s A_\mathfrak{p}\right) = o(1)$. Here we have used the assumption that A is strictly 1-balanced. The desired statement follows due to Equation (3.16).

1-statement. As stated in the introduction, this result was previously proved by Friedgut, Rödl and Schacht [66] as well as independently Conlon and Gowers [34]. The proof presented here serves as a short version that follows the short proof of a sparse Ramsey result due to Nenadov and Steger [110].

We will need two ingredients in order to prove the 1-statement of Theorem 3.1. The first will be the following easy corollary to Lemma 3.4.

Corollary 3.13. *For a given partition regular matrix $A \in \mathbb{Z}^{r \times m}$ and $s \in \mathbb{N}$ there exist $\varepsilon = \varepsilon(A, s)$ and $\delta = \delta(A, s) > 0$ such that for any $T_1, \dots, T_s \subseteq [n]$ satisfying $|S_0(A, \mathbf{0}) \cap T_i^m| \leq \varepsilon\,|S_0(A, \mathbf{0}) \cap [n]^m|$ for $1 \leq i \leq s$ we have $\Big|[n] \setminus (T_1 \cup \dots \cup T_s)\Big| \geq \delta n$ for n large enough.*

Proof. Let $\zeta = \zeta(A, s+1)$ be as in Lemma 3.4 and $\varepsilon = \varepsilon(A, s) = \zeta/2s$. Set $\tilde{T}_i = T_i \setminus \bigcup_{j=1}^{i-1} T_j$ for $1 \leq i \leq s$ and $\tilde{T}_{s+1} = [n] \setminus \bigcup_{j=1}^{r} T_j$ and consider the partition $[n] = \tilde{T}_1 \,\dot\cup\, \dots \,\dot\cup\, \tilde{T}_s \,\dot\cup\, \tilde{T}_{s+1}$. By Lemma 3.4 we have $|S_0(A, \mathbf{0}) \cap \tilde{T}_1^m| + \dots + |S_0(A, \mathbf{0}) \cap \tilde{T}_{s+1}^m| \geq \zeta\,|S_0(A, \mathbf{0}) \cap [n]^m|$ and since by assumption $|S_0(A, \mathbf{0}) \cap \tilde{T}_i^m| \leq |S_0(A, \mathbf{0}) \cap T_i^m| \leq \zeta/2s\,|S_0(A, \mathbf{0}) \cap [n]^m|$ for all $i \in \{1, \dots, s\}$, we have $|S_0(A, \mathbf{0}) \cap ([n] \setminus (T_1 \cup \dots \cup T_s))^m| \geq \zeta/2\,|S_0(A, \mathbf{0}) \cap [n]^m|$. Observe that by Lemma 3.12 every element in $[n]$ is contained in at most $m\,n^{m-\mathrm{rk}(A)-1}$ solutions and by Lemma 1.4 there exists $c_0 = c_0(A) > 0$ such that $|\,S_0(A, \mathbf{0}) \cap [n]^m| \geq c_0\,n^{m-\mathrm{rk}(A)}$ for n large enough. The result therefore follows for $\delta = \zeta c_0/2$. □

The second ingredient is stated in the following corollary that is obtained by applying the Hypergraph Container Theorem to the hyperpgraph of solutions.

Corollary 3.14. *For a given partition regular matrix $A \in \mathbb{Z}^{r \times m}$ and $\varepsilon > 0$ there exist $t = t(n)$ sets $T_1, \dots, T_t \in \binom{[n]}{\leq c_0\, n^{1-1/m_1(A)}}$ for some $c_0 > 0$ as well as sets $C_1, \dots, C_t \subseteq [n]$ such that*

$$|S_0(A, \mathbf{0}) \cap C_i^m| \leq \varepsilon\,|S_0(A, \mathbf{0}) \cap [n]^m|. \tag{3.17}$$

Furthermore, for every set $T \subseteq [n]$ *satisfying* $S_0(A, \mathbf{0}) \cap T^m = \emptyset$ *there exists* $1 \leq i \leq t$ *such that*

$$T_i \subseteq T \subseteq C_i. \tag{3.18}$$

Proof. Let $\mathcal{H}_n$ again be the hypergraph with vertex set $V(\mathcal{H}_n) = [n]$ and edge multiset

$$E(\mathcal{H}_n) = \Big\{\!\!\Big\{ \{x_1, \ldots, x_m\} : (x_1, \ldots, x_m) \in S_0(A, \mathbf{0}) \cap [n]^m \Big\}\!\!\Big\}.$$

We have previously observed that there exists a $c > 0$ such that for $1 \leq \ell \leq m$ we have $\Delta_\ell(\mathcal{H}_n) \leq c\, p(n)^{\ell-1}\, e(\mathcal{H}_n)/v(\mathcal{H}_n)$ for $p = p(n) = n^{-1/m_1(A)}$. We observe that $\mathcal{H}_n$ is trivially $(\mathcal{F}, \varepsilon)$-dense for $\mathcal{F} = \{T \subseteq [n] : |\, S_0(A, \mathbf{0}) \cap T^m| \geq \varepsilon\, |\, S_0(A, \mathbf{0}) \cap [n]^m|\}$. Applying Theorem 3.8 gives the desired statement. □

We are now ready to give a short proof of the 1-statement in Theorem 3.1. Let $\varepsilon, \delta > 0$ be as in Corollary 3.13 and let $t = t(n)$, c_0, $S_1, \ldots, S_t$ and $C_1, \ldots, C_t$ be as in Corollary 3.14. Let $C = C(A, s)$ be large enough such that

$$\left(1 + \ln\left(\frac{2^s}{sc_0}\right) + \ln(C)\right) \frac{sc_0}{C} < \frac{1}{2}.$$

Observe now that for a partition of the random set $T_1 \dot\cup \ldots \dot\cup T_s = [n]_p$ satisfying $S_0(A, \mathbf{0}) \cap T_i^m = \emptyset$ for all $i \in \{1, \ldots, s\}$ there exist $j_1, \ldots, j_s \in \{1, \ldots, t\}$ so that $S_{j_i} \subseteq T_i \subseteq C_{j_i}$ for all $i \in \{1, \ldots, s\}$. Since $T_i \subseteq [n]_p$ for $1 \leq i \leq s$ and $[n] \setminus (C_1 \cup \cdots \cup C_s) \cap [n]_p = \emptyset$ we can bound the probability of $[n]_p$ not fulfilling the partition property by

$$\mathbb{P}\left([n]_p \not\to_s A\right) \leq \sum_{j_1, \ldots, j_s \in \{1, \ldots, t\}} \mathbb{P}\left(S_{j_1}, \ldots, S_{j_s} \subseteq [n]_p \,\wedge\, [n] \setminus (C_{j_1}, \ldots, C_{j_s}) \cap [n]_p = \emptyset\right).$$

Observe that the two events $S_{j_1}, \ldots, S_{j_s} \subseteq [n]_p$ and $[n] \setminus (C_{j_1}, \ldots, C_{j_s}) \cap [n]_p = \emptyset$ are independent, so that we have

$$\mathbb{P}\left([n]_p \not\to_s A\right) \leq \sum_{j_1, \ldots, j_s \in \{1, \ldots, t\}} p^{\left|\bigcup_{j=1}^s S_j\right|} (1-p)^{\left|[n] \setminus (C_{j_1}, \ldots, C_{j_s})\right|}.$$

We bound this by choosing $k = |\bigcup_{j=1}^s S_j| \leq sc_0 n^{1-1/m_1(A)}$, then picking k elements and lastly deciding for each element in this selection in which of the S_i it is contained, so that we have

$$\begin{aligned}\mathbb{P}\left([n]_p \not\to_s A\right) &\leq (1-p)^{\delta n} \sum_{k=0}^{sc_0 n^{1-1/m_1(A)}} \binom{n}{k} \left(2^s\right)^k p^k \\ &\leq e^{-\delta np} \left(1 + \sum_{k=1}^{sc_0 C^{-1}\, np} \left(\frac{e 2^s\, np}{k}\right)^k\right).\end{aligned}$$

Lastly, we note that for $c > 0$ the function $f(x) = (c/x)^x$ is increasing for $0 < x \leq c/e$ since $d/dx\, f(x) = (c/x)^x\,(\log(c/x) - 1)$. We have chosen C large enough so that

$$\begin{aligned}\mathbb{P}\left([n]_p \not\to_s A\right) &\leq e^{-\delta np} \left(1 + (sc_0 C^{-1}\, np) \left(\frac{e 2^s\, np}{sc_0 C^{-1}\, np}\right)^{sc_0 C^{-1}\, np}\right) \\ &\leq e^{-\delta np}\, e^{\delta np/2} = o(1).\end{aligned}$$

for n large enough and therefore $[n]_p \to_s A$ asymptotically almost surely. ■

Chapter 4

The Breaker Threshold

Motivated by van der Waerden's Theorem, Beck [6] introduced k-term van der Waerden games as the positional games played on the set $[n]$ by two players, Maker and Breaker, who take turns occupying integers that have previously not been occupied by either of them. We will always assume that Maker goes first. Maker wins the game if he manages to occupy a k-term arithmetic progression. Breaker wins if he can keep Maker from achieving his goal, that is if he occupies at least one integer in every k-term arithmetic progression in $[n]$. There is no draw in this game.

If we denote the 2-color van der Waerden number by $W(k)$, that is smallest integer n such that any 2-coloring of $[n]$ must contain a k-term arithmetic progression, then Breaker cannot win the game on $[W(k)]$ without occupying a k-term arithmetic progression himself. A standard strategy stealing argument, see for example [8], implies that Breaker therefore cannot have a winning strategy on $[W(k)]$ and consequently Maker has to have one. Beck therefore defined $W^\star(k)$ to be the smallest integer n for which Maker has a winning strategy in the k-term van der Waerden game played on $[n]$. Clearly one has the trivial upper bound $W^\star(k) \leq W(k)$ for every $k \in \mathbb{N}$, but Beck actually established that the van der Waerden game number is single exponential, namely that $W^\star(k) = 2^{k(1+o(1))}$. This is in strong contrast to the enormous gap between the known upper and lower bounds for $W(k)$: there are several lower bounds of the form $2^{k(1+o(1))}$ [12, 149], while the best known upper bound is a tower function of height five [69].

One common way to even out the odds in games that disadvantage one player, is to allow the disadvantaged person to take multiple elements each round. Such biased games represent a central direction in the field of positional games, with deep connections to the theory of random structures. The notion was first suggested by Chvátal and Erdős [28] while investigating the connectivity game, hamiltonicity game and triangle-building game played on the edges of the complete graph. Given a hypergraph $\mathcal{H} = (V(\mathcal{H}), E(\mathcal{H}))$ and a positive integer bias q, we define the q-biased Maker-Breaker game $\mathbf{G}(\mathcal{H}; q)$ as follows: Maker and Breaker take turns occupying previously unoccupied vertices from $V(\mathcal{H})$, with Maker going first and occupying one vertex in each round and Breaker occupying up to q. Maker wins if his selection completely covers an edge from $\mathcal{H}$ and Breaker wins otherwise. In other words, Breaker wins if and only if the vertices of Maker form an independent set in $\mathcal{H}$. Note that, by definition, the game again cannot end in a draw. Given a hypergraph $\mathcal{H}$, one is interested in determining the threshold bias $q_0(\mathcal{H})$, defined to be the smallest integer $q \in \mathbb{N}$ for which Breaker has a winning strategy in the game $\mathbf{G}(\mathcal{H}; q)$.

The relationship of biased games to random discrete structures originates from the simple observation that if both Maker and Breaker occupy their vertices uniformly at random from the remaining free vertices, then Maker ends up occupying a uniform random set of size at least $|V(\mathcal{H})|/(q+1)$. For exactly what size a uniform random subset will likely be independent, is a central line of research in the study of random discrete structures. We previously considered it for the hypergraph of integer solutions to linear systems in Chapter 2.

Chvátal and Erdős [28] were mostly concerned about graph games where the vertex set of the game hypergraph $\mathcal{H}$ is the edge set $E(K_n)$ of the complete graph on n vertices and $\mathcal{H}$ represents a graph property. For the connectivity game, they proved the surprising phenomenon that the threshold bias for the game with 'clever' players is of the same order as the 'likely' threshold bias in the game with random players. In other words the *result* of the clever game and the random game are likely to be the same for almost all biases except an interval of length of smaller order than the value of the threshold bias. This result was later strengthened to establish the equality of the constant factors of the threshold biases and also extended to the Hamiltonicity game [67, 99].

Given some fixed graph H, the H-building game is the Maker-Breaker game where the vertex set of the game hypergraph is again the edge set $E(K_n)$ and the edges of the hypergraph correspond to all copies of H in K_n. Here it is import that, unlike in the previously mentioned graph games, H is a fixed small graph rather that does not grow with n. Chvátal and Erdős resolved the issue of the threshold bias for the triangle-building via ad-hoc strategies. They found that here the previously described phenomenon does not hold: their clever ad-hoc Breaker-strategy wins with a bias much smaller than is needed for a random Breaker playing against a random Maker. The real reason for this and the question for H-building games remained a mystery until its spectacular resolution by Bednarska and Łuczak [10]. Writing

$$m_2(H) = \max_{\substack{F \subset H \\ v(F) \geq 3}} \frac{e(H) - 1}{v(H) - 2},$$

for the 2-density of the graph H as well as $\mathcal{B}(H, n)$ for the $e(H)$-uniform hypergraph of all copies of H in K_n, their result can be stated as follows.

Theorem 4.1 (Bednarska–Łuczak)**.** *For every graph H with at least three non-isolated vertices, the threshold bias of the H-building game on K_n satisfies*

$$q_0\Big(\mathcal{B}(H, n)\Big) = \Theta(n^{1/m_2(H)}). \tag{4.1}$$

To gain an intuition for this result, it is worthwhile to investigate the general lower bound on $q_0(\mathcal{H})$, which is delivered by the uniform random strategy of Maker. Namely, if Maker occupies a free element of $V(\mathcal{H})$ uniformly at random in each round and wins with non-zero probability against a Breaker playing optimally with bias q, then clearly $q_0(\mathcal{H}) \geq q$. It is important to note that in this 'half-random' scenario Maker's random set of size $|V(\mathcal{H})|/(b+1)$ is no longer chosen uniformly at random, since it depends very much on Breaker's strategy. It turns out however that the success of the uniform random strategy of Maker can be salvaged if, for some constant $\varepsilon > 0$, a uniform random set of size $\varepsilon \, |V(\mathcal{H})|/(b+1)$ not only is expected to contain a hyperedge, but more resiliently, that *every δ-fraction of it* is expected to contain a hyperedge, for some $\delta < 1$. Note that we previously studied resilience results in Chapter 3. If Maker actually occupies at least a δ-fraction of some uniform random subset of that size, he wins. Bednarska and Łuczak [10] managed to implement this plan and couple it with an appropriate Breaker strategy.

Building on these ideas, let us extend the results of Bednarska and Łuczak to a whole range of other hypergraphs. All results will follow from two general winning criteria, one for Maker and one for Breaker, which can be applied to hypergraphs possessing certain 'container-type' regularity conditions that properly separate the maximum degree of ℓ-element vertex sets from the average degree. These hypergraphs in particular include the ones corresponding to Beck's van der Waerden games. More generally, one can also obtain tight results for a much broader class of games, which will be called Rado games and in which Maker's goal is to occupy a solution to an arbitrary given linear system of equations. One can also extend the result of Bednarska and Łuczak to hypergraph-building games for arbitrary fixed uniformity. It is worthwhile to note that the analogous extension from graphs to hypergraphs represented a significant jump in difficulty for the analogous sparse random problem [34, 136], while here one obtains it for a wider classes of hypergraphs, using a deterministic Breaker strategy. Furthermore, the container method, so effective there, only provides a Maker winning strategy with a log-factor below the optimal bias.

Let us now first give a formal statement of the two winning criteria. This will be followed by the statement of the results regarding Rado games, that is the generalization of Beck's van der Waerden games for linear systems. Lastly, we will describe the results regarding hypergraph-building games.

General Winning Criteria

In order to simplify notation, we often identify the hypergraph $\mathcal{H}$ with its edge set $E(\mathcal{H})$. We denote the number of vertices of a hypergraph $\mathcal{H}$ by $v(\mathcal{H})$, the number of edges by $e(\mathcal{H})$ and its density by $d(\mathcal{H}) = e(\mathcal{H})/v(\mathcal{H})$. Given a subset $S \subseteq V(\mathcal{H})$ of vertices, let $\deg(S) = |\{e \in \mathcal{H} : S \subset e\}|$. For any integer $\ell \in \mathbb{N}$ the maximum ℓ-degree is given by $\Delta_\ell(\mathcal{H}) = \max\{\deg(S) : S \subseteq V(\mathcal{H}), |S| = \ell\}$. Note that if $\mathcal{H}$ is simple and k-uniform, then $\Delta_k(\mathcal{H}) = 1$ and $\Delta_\ell(\mathcal{H}) = 0$ for all integers $\ell > k$. Given some sequence of hypergraphs $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$, the first statement now gives a criterion for a lower bound of the threshold biases of $\mathbf{G}(\mathcal{H}_n; q)$.

Theorem 4.2. *If $k \geq 2$ and $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$ is a sequence of k-uniform hypergraphs that satisfies*

(M1) $\Delta_1(\mathcal{H}_n) = O\Big(d(\mathcal{H}_n)\Big)$,
(M2) $\Delta_2(\mathcal{H}_n) = o\Big(d(\mathcal{H}_n)\Big)$ *and*
(M3) $d(\mathcal{H}_n) = o\Big(v(\mathcal{H}_n)^{k-1}\Big)$,

then the threshold bias of the game played on $\mathcal{H}_n$ satisfies

$$q_0(\mathcal{H}_n) = \Omega\left(\min_{2\leq\ell\leq k}\left(\frac{d(\mathcal{H})}{\Delta_\ell(\mathcal{H})}\right)^{\frac{1}{\ell-1}}\right).$$

The proof of Theorem 4.2 is based on a random strategy for Maker and will be given in Section 4.1. The second statement now gives a criterion for an upper bound on the threshold.

Theorem 4.3. *If $k \geq 2$ and $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$ is a sequence of k-uniform hypergraphs such that $v(\mathcal{H}_n)$ tends to infinity and there exists an $\varepsilon > 0$ so that*

$$\Delta_\ell(\mathcal{H}_n)^{\frac{1}{k-\ell}}\, v(\mathcal{H}_n)^{\varepsilon} = O\Big(\Delta_1(\mathcal{H}_n)^{\frac{1}{k-1}}\Big)$$

for every $2 \leq \ell \leq k-1$, then the threshold bias of the game played on $\mathcal{H}_n$ satisfies

$$q_0(\mathcal{H}_n) = O\left(\Delta_1(\mathcal{H}_n)^{\frac{1}{k-1}}\right).$$

The proof of Theorem 4.3 will be given in Section 4.2 and constructs an explicit winning strategy for Breaker through multiple applications of the biased Erdős–Selfridge Criterion of Beck together with a bias-doubling strategy that mimics a common alteration approach of the Probabilistic Method. The approach used both for the proof of Maker's criterion, as well as to devise Breaker's optimal strategy and prove its validity, follows the general lines of the proofs from [10].

Rado Games

Given some matrix $A \in \mathbb{Z}^{r\times m}$ and vector $\mathbf{b} \in \mathbb{Z}^r$, let us write

$$\mathcal{S}_0(A, \mathbf{b}, n) = \Big\{\{x_1, \ldots, x_m\} : (x_1, \ldots, x_m) \in S_0(A, \mathbf{b}) \cap [n]^m\Big\}.$$

for the m-uniform hypergraph given by the proper solutions in $[n]$ to the linear system $A \cdot \mathbf{x}^T = \mathbf{b}^T$. We will assume that $\mathcal{S}_0(A, \mathbf{b}, n)$ is simple, that is every edge occurs only once even though it may come from multiple solutions. We also denote by $\mathcal{S}_1(A, \mathbf{b}, n)$ the hypergraph containing all non-trivial solutions in $[n]$, that is

$$\mathcal{S}_1(A, \mathbf{b}, n) = \Big\{\{x_1, \ldots, x_m\} : (x_1, \ldots, x_m) \in \mathcal{S}_1(A, \mathbf{b}) \cap [n]^m\Big\},$$

and likewise assume that it is simple. Note that $\mathcal{S}_1(A, \mathbf{b}, n)$, in contrast to $\mathcal{S}_0(A, \mathbf{b}, n)$, is not necessarily uniform.

We refer to the biased Maker-Breaker game played on the hypergraph $\mathcal{S}_1(A, \mathbf{b}, n)$ as the Maker-Breaker $(A, \mathbf{b})$-game on $[n]$. This game is played on the set $[n]$ and Maker's goal is to occupy a non-trivial solution to $A \cdot \mathbf{x}^T = \mathbf{b}^T$. This notion extends van der Waerden games introduced by Beck [6] and we will therefore also call them Rado games motivated by Rado's Theorem. The main result regarding Rado games states the asymptotic behavior of the threshold bias of these games when A is abundant.

Theorem 4.4. *For every positive and abundant matrix $A \in \mathbb{Z}^{r \times m}$ and vector $\mathbf{b} \in \mathbb{Z}^r$ such that $S(A, \mathbf{b}) \neq \emptyset$, the threshold bias of the Maker-Breaker $(A, \mathbf{b})$-game on $[n]$ satisfies*

$$q_0\Big(\mathcal{S}_1(A, \mathbf{b}, n)\Big) = \Theta\left(n^{1/m_1(A)}\right). \tag{4.2}$$

We note that Maker's strategy will in fact result in occupying a *proper* instead of just a non-trivial solution. It follows that the bias threshold of the game player on $\mathcal{S}_0(A, \mathbf{b}, n)$ is of the same order as that played on $\mathcal{S}_1(A, \mathbf{b}, n)$. We will prove Theorem 4.4 in Section 4.3.

Small Hypergraph Games

As previously mentioned, Bednarska and Łuczak [10] showed that the threshold bias of the G-building game satisfies $q_0(\mathcal{B}(G, n)) = \Theta(n^{1/m_2(G)})$. We consider the following generalization: given some r-uniform hypergraph $\mathcal{G}$ on at least $r + 1$ non-isolated vertices, we define the r-density of $\mathcal{G}$ to be

$$m_r(\mathcal{G}) = \max_{\substack{\mathcal{F} \subseteq \mathcal{G} \\ v(\mathcal{F}) \geq r+1}} \frac{e(\mathcal{F}) - 1}{v(\mathcal{F}) - r}.$$

Note that this is an obvious generalization of the 2-density of a graph. Furthermore, we call $\mathcal{G}$ strictly r-balanced, if $m_r(\mathcal{G}) > (e(\mathcal{F}) - 1)/(v(\mathcal{F}) - r)$ for every subhypergraph $\mathcal{F}$ of $\mathcal{G}$ on at least $r + 1$ vertices.

Let $\mathcal{B}(\mathcal{G}, n)$ denote the $|\mathcal{G}|$-uniform hypergraph of all copies of $\mathcal{G}$ in the complete r-uniform hypergraph $\mathcal{K}_n^{(r)}$. Using the general winning criteria, we generalize the result of Bednarska and Łuczak to the Maker-Breaker $\mathcal{G}$-building game on $\mathcal{K}_n^{(r)}$, that is the game in which Maker tries to occupy a copy of $\mathcal{G}$ in $\mathcal{K}_n^{(r)}$.

Theorem 4.5. *For any r-uniform hypergraph $\mathcal{G}$ on at least $r+1$ non-isolated vertices, the threshold bias of the Maker-Breaker $\mathcal{G}$-building game on $\mathcal{K}_n^{(r)}$ satisfies*

$$q_0\Big(\mathcal{B}(\mathcal{G}, n)\Big) = \Theta\left(n^{1/m_r(\mathcal{G})}\right). \tag{4.3}$$

We will prove Theorem 4.5 in Section 4.4.

4.1 Proof of Theorem 4.2 – Maker's strategy

We start by stating a strengthening of Theorem 4.2 that we will actually prove. In order to do so, we introduce the function

$$f(\mathcal{H}) = \min_{2\le\ell\le k} \left(\frac{d(\mathcal{H})}{\Delta_\ell(\mathcal{H})}\right)^{\frac{1}{\ell-1}}$$

for any given k-uniform hypergraph $\mathcal{H}$ and note that

$$1/f(\mathcal{H}) = \max_{2\le\ell\le k} \left(\Delta_\ell(\mathcal{H})/d(\mathcal{H})\right)^{1/(\ell-1)}. \tag{4.4}$$

The combinatorial winning criterion for Maker now can be stated as follows. We will see how to derive Theorem 4.2 from it immediately afterwards.

Theorem 4.6. *For every $k \ge 2$ and every positive $c_1 \ge k$ there exists $c = c(k, c_1) > 0$ and $\bar{c} = \bar{c}(k, c_1) > 0$ such that the following holds: if $\mathcal{H}$ is a k-uniform hypergraph satisfying*

$$\textit{(Mi)}\ \Delta_1(\mathcal{H}) \le c_1\, d(\mathcal{H})\,, \quad \textit{(Mii)}\ f(\mathcal{H}) > 1, \quad \textit{(Miii)}\ \frac{v(\mathcal{H})}{f(\mathcal{H})}\left(1 - \frac{1}{f(\mathcal{H})}\right) \ge \bar{c},$$

then Maker has a winning strategy in $\mathbf{G}(\mathcal{H}; q)$ provided that $q \le cf(\mathcal{H}) - 1$.

We start by showing that Theorem 4.2 is a consequence of this result.

Proof of Theorem 4.2 from Theorem 4.6. We see that (M1) implies (Mi) for n large enough. Now (M2) implies $d(\mathcal{H}_n)/\Delta_2(\mathcal{H}_n) = \omega(1)$. As $\Delta_\ell(\mathcal{H}_n) \le \Delta_2(\mathcal{H}_n)$ for $3 \le \ell \le k$ this gives us $f(\mathcal{H}_n) = \omega(1)$, implying (Mii) for n large enough. As $f(\mathcal{H}_n) = \omega(1)$ we also know that $1 - 1/f(\mathcal{H}_n) \to 1$. Now by definition of f we have $v(\mathcal{H}_n)/f(\mathcal{H}_n) \ge v(\mathcal{H}_n)/d(\mathcal{H}_n)^{1/(k-1)}$ which by (M3) goes to infinity. This gives (Miii) for n large enough and the desired result follows. □

The following notion plays a crucial role in the proof of Theorem 4.6 and is a natural generalization of the notion that a set is (δ, k)-Szeméredi as defined by Conlon and Gowers [34], see also Section 3.1.

Definition 4.7. *Let $\mathcal{F}$ be a hypergraph and $0 < \delta < 1$. We say that a subset $T \subseteq V(\mathcal{F})$ of the vertices is δ-stable if every subset of $S \subseteq T$ of size $|S| \ge \delta|T|$ contains an edge of $\mathcal{F}$.*

Equivalently, T is called δ-stable, if the subhypergraph of $\mathcal{F}$ induced by T has independence number less than $\delta\,|T|$.

Maker's strategy will now consist of *picking* (but not necessarily occupying) elements uniformly at random from among all elements he has not previously picked. With this rule it is guaranteed that the set of elements Maker picks is uniformly random. Then, if possible, Maker occupies that picked element in the game, otherwise he occupies an arbitrary free element. We will prove that with positive probability Maker wins using this strategy by showing that a δ-fraction of the elements Maker *picked* he was also able to *occupy*. If we ensure that the set of vertices occupied by Maker is δ-stable, it then follows that Maker's set of vertices contain an edge with positive probability.

Let us extend the notation previously introduced for the binomial and uniform random set of integers: given any finite set S and $0 < p < 1$, the notation S_p to will refer to the binomial random set that is obtained by picking each element of S independently with probability p. On the other hand, given $0 \leq M \leq n$, we let S_M denote the uniform random set that is obtained by sampling uniformly at random over all subsets of S of size M. The key ingredient to prove the existence of a winning strategy for Maker is the following statement saying that $V(\mathcal{H})_M$ is δ-stable for suitable M and δ if $\mathcal{H}$ satisfies the requirements of Theorem 4.6.

Theorem 4.8. *For every $k \geq 2$ and for every constant $c_1 \geq k$ there exist constants $\delta = \delta(k, c_1) < 1$ and $\tilde{c} = \tilde{c}(k, c_1) > 0$ such that the following holds: if $\mathcal{H}$ is a k-uniform hypergraph satisfying*

$$\textit{(Mi)}\ \Delta_1(\mathcal{H}) \leq c_1\, d(\mathcal{H})\,, \quad \textit{(Mii)}\ f(\mathcal{H}) > 1, \quad \textit{(Miii)}\ \frac{v(\mathcal{H})}{f(\mathcal{H})}\left(1 - \frac{1}{f(\mathcal{H})}\right) \geq \tilde{c}$$

then

$$\mathbb{P}\left(V(\mathcal{H})_M \textit{ is not } \delta\textit{-stable}\right) < 3 \exp\left(-\frac{M}{c_1\, 2^{k+2}}\right),$$

for every $M \geq 2\,\lfloor v(\mathcal{H})/f(\mathcal{H}) \rfloor$.

We start by showing how to deduce Theorem 4.6 assuming the statement of Theorem 4.8. The proof of Theorem 4.8 will be given immediately afterwards.

Proof of Theorem 4.6 from Theorem 4.8. Fix an arbitrary strategy for Breaker. We will study the following random strategy for Maker: in each round he *picks* an element uniformly at random from among all elements of $V(\mathcal{H})$ that he has not previously picked. If this element was not already occupied by either Maker or Breaker, then Maker occupies it. Otherwise he occupies an arbitrary free vertex and 'forgets' about it for the rest of the game, i.e. he doesn't consider it as picked and can potentially pick it at a later point. Note the subtle difference between picking and occupying a vertex: occupying is the act of actually choosing a vertex in the process of the game and picking can, depending on whether the vertex was already occupied by either player, be merely in the mind of Maker. We label an element picked by Maker as a *failure*, if that element was already occupied by Breaker. We will show that this random strategy succeeds with positive probability against Breaker's arbitrary strategy, implying that Breaker does not have a winning strategy and therefore Maker must have one.

Let $\delta = \delta(k, c_1) < 1$ be chosen according to Theorem 4.8 and $q \leq cf(\mathcal{H}) - 1$ where $c = (1-\delta)/4 > 0$. We will consider the first

$$M = 2\left\lfloor \frac{v(\mathcal{H})}{f(\mathcal{H})} \right\rfloor \leq \frac{1-\delta}{2}\,\frac{v(\mathcal{H})}{q+1} \tag{4.5}$$

rounds of the game. We may consider the set of elements that Maker picked in these M rounds as the uniform random set $V(\mathcal{H})_M$. Note that some of his elements may be failures. We will now upper bound the probability that Maker's i-th move, which we refer to as m_i, was a failure. Clearly this probability is upper bounded by the probability that his M-th move is a failure since in every round the number of potential failures does not decrease and the number of vertices Maker picks from strictly decreases. Note that in the first $M-1$ rounds, Maker picked exactly $M-1$ vertices. So, in round M, there are $v(\mathcal{H}) - M + 1$ available vertices to pick from. The potential failures are among the vertices occupied by Breaker and hence their number is at most $q\,(M-1)$. Using Equation (4.5) it follows that

$$\mathbb{P}\left(m_i \text{ a failure}\right) \leq \mathbb{P}\left(m_M \text{ a failure}\right) \leq \frac{q\,(M-1)}{v(\mathcal{H}) - (M-1)} \leq \frac{q\,M}{v(\mathcal{H}) - M} \leq \frac{1-\delta}{2}$$

for every $i \in \{1, \ldots, M\}$. The probability that Maker has more than $(1-\delta)\,M$ failures is now at most the probability that among M independent Bernoulli trials with failure probability $(1-\delta)/2$ there exist more than $(1-\delta)\,M$ failures, which is less than $1/2$ by Markov's Inequality. In other words, with probability at least $1/2$, at least δM elements picked by Maker are not failures.

By Theorem 4.8 the probability that $V(\mathcal{H})_M$ is not δ-stable is strictly less than $3\exp\left(-M(c_1\,2^{k+2})^{-1}\right)$. Setting $\bar{c}$ to be the maximum of $c_1 2^{k+1}(\log(3)+\log(4))+1$ and the according value of $\tilde{c}$ in Theorem 4.8, and using that $f(\mathcal{H}) \leq \bar{c}^{-1} v(\mathcal{H})$ by (Miii), one can verify that the probability that the uniform random set $V(\mathcal{H})_M$ is not δ-stable is at most 1/4. Consequently, with probability at least 1/4, the at least δM vertices occupied by Maker contain an edge of $\mathcal{H}$. □

Proof of Theorem 4.8. The heart of the proof of Theorem 4.8 is the following statement due to Janson, Łuczak and Ruciński [86], see also the Second Moment Method in Section 1.6.

Theorem 4.9 (Theorem 2.18, (ii) in [87])**.** *Let S be a finite set, $0 < p < 1$ and $\mathbb{1}_T$ the indicator random variable of the event that $T \subseteq S_p$ for a given $T \subseteq S$. If $\mathcal{S} \subset \mathcal{P}(S)$ is a family of subsets of S and $X = \sum_{T \in \mathcal{S}} \mathbb{1}_T$, then*

$$\mathbb{P}\left(X = 0\right) \leq \exp\left(- \frac{\mathbb{E}(X)^2}{\sum\limits_{\substack{(T,T') \in \mathcal{S}^2 \\ T \cap T' \neq \emptyset}} \mathbb{E}(\mathbb{1}_T \mathbb{1}_{T'})} \right).$$

We want to apply Theorem 4.9 with $S = V(\mathcal{H})$, $\mathcal{S} = E(\mathcal{H})$ and $p = 1/f(\mathcal{H})$. Note that $p < 1$ due to (Mii). For $X = \sum_{e \in \mathcal{H}} \mathbb{1}_e$ we have, by linearity of expectation, that $\mathbb{E}(X) = e(\mathcal{H})\, p^k$. Note that $\mathbb{E}(\mathbb{1}_e \mathbb{1}_{e'}) = p^{2k - |e \cap e'|}$ for $e, e' \in \mathcal{H}$ and therefore

$$\begin{aligned}
\sum_{\substack{(e,e') \in \mathcal{H}^2 \\ e \cap e' \neq \emptyset}} \mathbb{E}(\mathbb{1}_e \mathbb{1}_{e'}) &= \sum_{e \in \mathcal{H}} \sum_{\emptyset \neq T \subseteq e} \sum_{\substack{e' \in \mathcal{H} \\ e \cap e' = T}} p^{2k - |T|} \leq \sum_{e \in \mathcal{H}} \sum_{\emptyset \neq T \subseteq e} \deg(T)\, p^{2k-|T|} \\
&\leq e(\mathcal{H}) \left(2^k - 1\right) \max_{\emptyset \neq T \subseteq V(\mathcal{H})} \left(\deg(T)\, p^{2k-|T|}\right) \\
&\leq 2^k\, e(\mathcal{H}) \max_{1 \leq \ell \leq k} \left(\Delta_\ell(\mathcal{H})\, p^{2k-\ell}\right) \\
&= 2^k\, e(\mathcal{H})\, p^{2k-1} \max\left(\max_{2 \leq \ell \leq k} \left(\frac{\Delta_\ell(\mathcal{H})}{p^{\ell-1}} \right), \Delta_1(\mathcal{H}) \right) \\
&= 2^k \frac{\mathbb{E}(X)^2}{p\, v(\mathcal{H})\, d(\mathcal{H})} \max\left(\max_{2 \leq \ell \leq k} \left(\frac{\Delta_\ell(\mathcal{H})}{p^{\ell-1}} \right), \Delta_1(\mathcal{H}) \right).
\end{aligned}$$

Using (Mi) we now get

$$\begin{aligned}
\sum_{\substack{(e,e') \in \mathcal{H}^2 \\ e \cap e' \neq \emptyset}} \mathbb{E}(\mathbb{1}_e \mathbb{1}_{e'}) &\leq 2^k \frac{\mathbb{E}(X)^2}{p\, v(\mathcal{H})} \max\left(\max_{2 \leq \ell \leq k} \left(\frac{\Delta_\ell(\mathcal{H})}{d(\mathcal{H})\; p^{\ell-1}} \right), c_1 \right) \\
&\leq 2^k \frac{\mathbb{E}(X)^2}{p\, v(\mathcal{H})} \max\left(\max_{2 \leq \ell \leq k} \left(\frac{1}{\left(f(\mathcal{H})\; p\right)^{\ell-1}} \right), c_1 \right) = c_1\, 2^k \frac{\mathbb{E}(X)^2}{p\, v(\mathcal{H})}.
\end{aligned}$$

where the last inequality follows from the definition of $f(\mathcal{H})$ and the last equality follows from the fact that $f(\mathcal{H})\,p = 1$ and $c_1 \geq k$. Now, using the estimate from Theorem 4.9, we get

$$\mathbb{P}\left(V(\mathcal{H})_p \text{ contains no edge of } \mathcal{H}\right) = \mathbb{P}\left(X = 0\right) \leq \exp(-c'\, v(\mathcal{H})\, p) \tag{4.6}$$

where $c' = 1/(c_1\, 2^k)$. The following lemma will now allow us to bound the probability of a uniform random set fulfilling our desired property by the probability that a corresponding binomial random set fulfills it.

Lemma 4.10. *Let $X \sim \mathcal{B}(n,p)$ and let $\mathcal{P}$ be a monotone decreasing family of subsets of $[n]$.Then there exists a constant $C > 0$ such that if $\sqrt{np(1-p)} > C$, then $\mathbb{P}\left([n]_{\lfloor np \rfloor} \in \mathcal{P}\right) \leq 3\,\mathbb{P}\left([n]_p \in \mathcal{P}\right)$.*

Proof. Note that since $\mathcal{P}$ is monotone decreasing, we have $\mathbb{P}\left([n]_K \in \mathcal{P}\right)$ is greater than $\mathbb{P}\left([n]_L \in \mathcal{P}\right)$ whenever $K \leq L$. Thus

$$\begin{aligned}
\mathbb{P}\left([n]_p \in \mathcal{P}\right) &= \sum_{M=0}^{n} \mathbb{P}\left([n]_p \in \mathcal{P} \,|\, |[n]_p| = M\right) \mathbb{P}\left(|[n]_p| = M\right) \\
&= \sum_{M=0}^{n} \mathbb{P}\left([n]_M \in \mathcal{P}\right) \mathbb{P}\left(|[n]_p| = M\right) \\
&\geq \sum_{M=0}^{\lfloor np \rfloor} \mathbb{P}\left([n]_M \in \mathcal{P}\right) \mathbb{P}\left(|[n]_p| = M\right) \\
&\geq \mathbb{P}\left([n]_{\lfloor np \rfloor} \in \mathcal{P}\right) \sum_{M=0}^{\lfloor np \rfloor} \mathbb{P}\left(|[n]_p| = M\right).
\end{aligned}$$

Note that $\sum_{M=0}^{\lfloor np \rfloor} \mathbb{P}\left(|[n]_p| = M\right) = \mathbb{P}\left(X \leq \lfloor np \rfloor\right)$. Let $\mu_{1/2}$ be the median of X and assume first that $\lfloor np \rfloor \leq \mu_{1/2} < \lceil np \rceil$. Then $\mathbb{P}\left(X \leq \lfloor np \rfloor\right) = \mathbb{P}\left(X \leq \mu_{1/2}\right) \geq \frac{1}{2}$ and hence

$$\mathbb{P}([n]_{\lfloor np \rfloor} \in \mathcal{P}) \leq 2\,\mathbb{P}\left([n]_p \in \mathcal{P}\right).$$

It remains to be shown that the assertion follows as well if $\mu_{1/2} = \lceil np \rceil$. Note that

$$\mathbb{P}\left(X \leq \lfloor np \rfloor\right) = \mathbb{P}\left(X \leq \lceil np \rceil\right) - \mathbb{P}\left(X = \lceil np \rceil\right) \geq \frac{1}{2} - \mathbb{P}\left(X = \lceil np \rceil\right).$$

We will show that $\mathbb{P}\left(X = \lceil np \rceil\right) \leq 1/6$ which then implies that $\mathbb{P}\left([n]_{\lfloor np \rfloor} \in \mathcal{P}\right)$ is smaller than $3\,\mathbb{P}\left([n]_p \in \mathcal{P}\right)$. To do so, we will upper bound the probability that $X =$

$\lceil np \rceil$ and use the inequalities $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \leq n! \leq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e$ as follows:

$$\begin{aligned}\mathbb{P}\left(X = \lceil np \rceil\right) = \binom{n}{\lceil np \rceil} p^{\lceil np \rceil}\,(1-p)^{n-\lceil np \rceil} &= \frac{n!\,p^{\lceil np \rceil}\,(1-p)^{n-\lceil np \rceil}}{\lceil np \rceil!(n-\lceil np \rceil)!} \\ &\leq \frac{\sqrt{n}\;n^n\;e\;p^{\lceil np \rceil}\;(1-p)^{n-\lceil np \rceil}}{\sqrt{2\pi\lceil np \rceil}\;(\lceil np \rceil)^{\lceil np \rceil}\;\sqrt{n-\lceil np \rceil}\;(n-\lceil np \rceil)^{n-\lceil np \rceil}} \\ &= \frac{\sqrt{n}}{\sqrt{n-\lceil np \rceil}}\,\frac{(np)^{\lceil np \rceil}}{\lceil np \rceil^{\lceil np \rceil}}\,\frac{(n-np)^{n-\lceil np \rceil}}{(n-\lceil np \rceil)^{n-\lceil np \rceil}}\,\frac{e}{\sqrt{2\pi\lceil np \rceil}},\end{aligned}$$

Clearly we have $(np)^{\lceil np \rceil}/\lceil np \rceil^{\lceil np \rceil} \leq 1$ as well as $(n-np)^{n-\lceil np \rceil}/(n-\lceil np \rceil)^{n-\lceil np \rceil} \leq e$. Hence we get

$$\mathbb{P}\left(X = \lceil np \rceil\right) \leq \frac{e^2}{\sqrt{2\pi}}\sqrt{\frac{n}{n-np-1}\,\frac{1}{np}} \leq \frac{3}{\sqrt{(1-p)np-p}} < \frac{3}{\sqrt{C^2-1}}.$$

Choosing $C > 0$ large enough such that $\mathbb{P}\left(X = \lceil np \rceil\right) \leq 1/6$ gives the desired property. □

Since the property of 'not containing an edge of $\mathcal{H}$' is monotone decreasing, we can choose $\tilde{c}$ such that we can apply Lemma 4.10 to restate Equation (4.6) for the uniform random set model as follows:

$$\begin{aligned}&\mathbb{P}\left(V(\mathcal{H})_{\bar{M}} \text{ contains no edge of } \mathcal{H}\right) \\ \leq\,&3\,\mathbb{P}\left(V(\mathcal{H})_p \text{ contains no edge of } \mathcal{H}\right) \\ \leq\,&3\,\exp(-c'\bar{M}) \end{aligned} \tag{4.7}$$

for any $\bar{M} \geq \lfloor v(\mathcal{H})/f(\mathcal{H}) \rfloor$.

We are now ready to finish the proof. Let $M \geq 2\lfloor v(\mathcal{H})/f(\mathcal{H}) \rfloor$ and let $\delta = \delta(k, c_1) \in (1/2, 1)$ be such that $(1-\delta)(1-\ln(1-\delta)) < c'/4$. To see that this is indeed possible, note that for $x \in (0,1)$ the function $g(x) = (1-x)(1-\ln(1-x))$ satisfies $g(x) \to 0$ as $x \to 1$. Consider pairs (T, T') where $T \subset V(\mathcal{H})$ with $|T| = M$ and $T' \subseteq T$ is such that $|T'| = \delta\, M$ and T' does not contain an edge of $\mathcal{H}$. Using Equation (4.7) with $\delta\, M > \lfloor v(\mathcal{H})/f(\mathcal{H}) \rfloor$, we can estimate the number of choices for a set T' of size $\delta\, M$ that contains no edge of $\mathcal{H}$ by

$$3\,\exp(-c'\delta M)\binom{v(\mathcal{H})}{\delta\, M} \leq 3\,\exp\left(-\,c'\frac{M}{2}\right)\binom{v(\mathcal{H})}{\delta\, M}.$$

Hence, we can upper bound the number of pairs (T, T') as described above by

$$3 \exp\Big(- c' \frac{M}{2} \Big) \binom{v(\mathcal{H})}{\delta\, M} \binom{v(\mathcal{H}) - \delta M}{(1-\delta)\, M} = 3 \exp\Big(- c' \frac{M}{2} \Big) \binom{M}{(1-\delta) M} \binom{v(\mathcal{H})}{M}.$$

We can therefore upper bound the number of choices for a set T of size M containing a subset of size δM that does not contain an edge of $\mathcal{H}$ by

$$\begin{aligned} &3 \, \exp\Big(- c' \frac{M}{2} \Big) \binom{M}{(1-\delta)\, M} \binom{v(\mathcal{H})}{M} \\ \leq\, &3 \, \exp\Big(M \Big(- c'/2 + (1-\delta)(1 - \ln(1-\delta)) \Big) \Big) \binom{v(\mathcal{H})}{M}. \end{aligned}$$

Hence we get

$$\begin{aligned} &\mathbb{P}\left(V(\mathcal{H})_M \text{ is not } \delta\text{-stable } \right) \\ \leq\, &3 \, \exp\left(M(-c'/2 + (1-\delta)(1 - \ln(1-\delta))) \right) \\ \leq\, &3 \exp\left(-M \frac{c'}{4} \right), \end{aligned}$$

where the last inequality follows by choice of $\delta = \delta(k, c_1)$. ■

4.2 Proof of Theorem 4.3 – Breaker's strategy

We will derive Theorem 4.3 from the following stronger combinatorial statement.

Theorem 4.11. *For every $k \geq 2$ and $t > (2k)^k$ the following holds: if $\mathcal{H}$ is a k-uniform hypergraph, then Breaker has a winning strategy in $\mathbf{G}(\mathcal{H}; q)$ provided that*

$$q > 4 \Big((2\, v(\mathcal{H}))^{1/t} \, \Delta_1(\mathcal{H}) \, ke \Big)^{\frac{1}{k-1}}$$

as well as

$$q > 8k^2 t^3 \left(\max_{2 \leq \ell \leq k-1} \Big(\Delta_\ell(\mathcal{H}) \, \big((tk)^{tk} \, k^t \, v(\mathcal{H})^2 \big)^{\frac{k}{t^{1/k}}} \Big)^{\frac{1}{k-\ell}} + 2 \right).$$

Note that here e denotes the base of the natural logarithm and should not be confused with the number of edges. We start by giving a proof of Theorem 4.3 using Theorem 4.11. Then we define the necessary concepts for the remainder of the section. Following this, we present the two main strategies for Breaker and prove their correctness. Finally we prove Theorem 4.11 using these ingredients.

Proof of Theorem 4.3 from Theorem 4.11. Let $k \geq 2$ and $\varepsilon > 0$ be given and set $t = \log v(\mathcal{H})$. Assume that $v(\mathcal{H})$ is large enough such that $\log v(\mathcal{H}) > (2k)^k$. Using $e = v(\mathcal{H})^{1/\log v(\mathcal{H})}$, it is straightforward to check that

$$\Big((2n)^{1/t} \Delta_1(\mathcal{H})\; ke\Big)^{\frac{1}{k-1}} \leq C_1' \, \Delta_1(\mathcal{H})^{\frac{1}{k-1}}$$

for some constant $C_1' = C_1'(k) > 0$. Similarly for $v(\mathcal{H})$ sufficiently large we can upper bound the term

$$\begin{aligned} &2k^2 t^3 \left(\max_{2 \leq \ell \leq k-1} \left(2k\; \Delta_\ell(\mathcal{H}) \left(v(\mathcal{H})^2 \; (tk)^{tk} \right)^{\frac{k}{t^{1/k}}} \right)^{\frac{1}{k-\ell}} + 2 \right) \\ \leq\, & C_2' \, v(\mathcal{H})^{C_3' \frac{\log\log v(\mathcal{H})}{\log^{1/k} v(\mathcal{H})}} \max_{2 \leq \ell \leq k-1} \left(\Delta_\ell(\mathcal{H}) \right)^{\frac{1}{k-\ell}} \end{aligned}$$

for some constants $C_2' = C_2'(k) > 0$ and $C_3' = C_3'(k) > 0$. Note that

$$\log\log v(\mathcal{H}) / \log^{1/k} v(\mathcal{H}) = o(1)$$

and so for $v(\mathcal{H})$ large enough this will be at most $v(\mathcal{H})^{\varepsilon} \max_{2 \leq \ell \leq k-1} \Big(\Delta_\ell(\mathcal{H})^{\frac{1}{k-\ell}} \Big)$. Choose $C_1 = C_1(k) \geq \max(C_1', C_2', 4)$ and $v_0 = v_0(k)$ large enough, giving us the statement that Breaker has a winning strategy if

$$q \geq C_1 \max \Big(\Delta_1(\mathcal{H})^{\frac{1}{k-1}}, \max_{2 \leq \ell \leq k-1} \Big(\Delta_\ell(\mathcal{H})^{\frac{1}{k-\ell}} \Big) \, v(\mathcal{H})^{\varepsilon} \Big). \tag{4.8}$$

From this, the statement of Theorem 4.3 immediately follows. □

4.2.1 Preliminaries for the proof of Theorem 4.11

One of the most important results in the area of positional games is the Erdős–Selfridge Theorem [51], the biased version of which is due to Beck [7]. It ensures that Breaker can do at least as well as the expected outcome when both players act randomly. We will use the following consequence of it heavily in the proof of Theorem 4.11.

Biased Erdős–Selfridge Theorem [7]. *For every hypergraph $\mathcal{H}$ and integer $q \geq 1$ the following holds. If Breaker plays as the second player, he can keep Maker from* *covering more than*

$$(q+1) \sum_{H \in \mathcal{H}} \left(\frac{1}{q+1} \right)^{|H|} \tag{4.9}$$

winning sets in $\mathbf{G}(\mathcal{H}; q)$*. If he plays as the first player, then one can omit the first* $(q+1)$ *factor.*

We will also need the following simple yet powerful remark.

Remark 4.12. *If Breaker has a winning strategy for some positional game* $\mathbf{G}(\mathcal{H}; q)$ *where he is allowed to make* at most q *moves each round, then he also wins if he has to make* exactly q *moves each round. It follows that if he has a winning strategy for some game* $\mathbf{G}(\mathcal{H}_1; q_1)$ *and a winning strategy for another game* $\mathbf{G}(\mathcal{H}_2; q_2)$*, then he can combine these two strategies to define a winning strategy in* $\mathbf{G}(\mathcal{H}_1 \cup \mathcal{H}_2; q_1 + q_2)$*.*

This remark will be used extensively throughout the proof. Furthermore, we will need the following definitions, which are based on those developed in [10].

Definition 4.13 (Set-Theoretic definitions)**.** *Given some hypergraph* $\mathcal{H}$*, we define the following notions:*

1. *a* t-cluster *is a set of distinct edges* $\{H_1, \ldots, H_t\} \subset \mathcal{H}$ *satisfying* $|\bigcap_{i=1}^t H_i| \geq 2$*,*
2. *an* almost complete solution (H°, h) *is a tuple consisting of a set* $H^\circ \subseteq V(\mathcal{H})$ *as well as an element* $h \notin H^\circ$ *so that* $H = H^\circ \cup \{h\}$ *is an edge in* $\mathcal{H}$*,*
3. *a* t-fan *is a family of distinct almost complete solutions* $\{(H_1^\circ, h_1), \ldots, (H_t^\circ, h_t)\}$ *in* $\mathcal{H}$ *satisfying* $|\bigcap_{i=1}^t H_i^\circ| \geq 1$*,*
4. *a* t*-fan is is called* simple *if* $|H_i^\circ \cap H_j^\circ| = 1$ *for all* $1 \leq i < j \leq t$*,*
5. *a* t-flower *is a* t*-fan satisfying* $|\bigcap_{i=1}^t H_i^\circ| \geq 2$*.*

For each t*-fan in* $\mathcal{H}$ *we call the* h_i *the* open elements*, the* H_i° *the* major parts *and the elements of the intersection* $\bigcap_{i=1}^t H_i^\circ$ *the* common elements*.*

Definition 4.14 (Game-Theoretic Definitions)**.** *At any given point in a positional game on a given hypergraph* $\mathcal{H}$*, we call an almost complete solution* (H°, h) dangerous *if all elements of* H° *have been picked by Maker and* h *has not yet been picked by either player. A fan or flower is* dangerous *if their respective almost complete solutions are.*

Observe that for a dangerous t-fan or t-flower we must have $h_i \notin H_j^\circ$ for all $1 \leq i, j \leq t$. In the following we will always assume that Breaker plays as second player. We say that a player occupies a given t-fan or t-flower $(H_1^\circ, h_1), \dots, (H_t^\circ, h_t)$ if his selection of vertices contains $\bigcup_{i=1}^t H_i^\circ$. Similarly a player occupies a t-cluster $H_1, \dots, H_t$ if his selection of vertices contains $\bigcup_{i=1}^t H_i$.

4.2.2 Two important strategies for Breaker

The following two lemmata give us strategies that we will use to construct a larger strategy in the proof of Theorem 4.11. Note that in the statement of the lemma we do not care about which player covers the open elements of a fan.

Lemma 4.15. *For every integer $k \geq 2$ and $t \geq 1$ the following holds. If $\mathcal{H}$ is a k-uniform hypergraph, then Breaker with a bias of*

$$q > \Big((2\, v(\mathcal{H}))^{1/t}\ \Delta_1(\mathcal{H})\ ke \Big)^{1/(k-1)}$$

can prevent Maker from occupying $\binom{q}{t}/2$ simple t-fans in the game $\mathbf{G}(\mathcal{H}; q)$.

Proof. Let $\mathcal{F} = \Big\{ \bigcup_{i=1}^t H_i^\circ \mid \{(H_1^\circ, h_1), \dots, (H_t^\circ, h_t)\} \text{ simple } t\text{-fan in } \mathcal{H} \Big\}$ be the hypergraph of all simple t-fans in $\mathcal{H}$. We want to apply the Biased Erdős–Selfridge Theorem, so we estimate

$$(q+1) \sum_{F \in \mathcal{F}} \left(\frac{1}{q+1} \right)^{|F|} \leq (q+1) \left(v(\mathcal{H})\ \frac{\Delta_1(\mathcal{H})^t\ (k-1)^t}{t!} \right) \left(\frac{1}{q+1} \right)^{t(k-2)+1}.$$

This inequality holds because there are $v(\mathcal{H})$ ways to fix the common element of a simple t-fan, $\Delta_1(\mathcal{H})^t$ is an upper bound on the number of t-tuples of edges containing the fixed common element and there are $(k-1)^t$ ways of fixing the corresponding open elements. Note that an open element is never a common element by definition. Furthermore, $t!$ takes care of the symmetry and each simple t-fan is of size $t(k-2)+1$. We therefore get

$$\begin{aligned}(q+1) \sum_{F \in \mathcal{F}} \left(\frac{1}{q+1} \right)^{|F|} &\leq v(\mathcal{H}) \left(\frac{\Delta_1(\mathcal{H})\ ke}{t\ q^{k-2}} \right)^t = v(\mathcal{H}) \left(\frac{\Delta_1(\mathcal{H})\ ke}{q^{k-1}} \right)^t \left(\frac{q}{t} \right)^t \\ &< v(\mathcal{H})\ \frac{1}{2\, v(\mathcal{H})} \left(\frac{q}{t} \right)^t \leq \frac{1}{2} \binom{q}{t}.\end{aligned}$$

The claim now follows by applying the Biased Erdős–Selfridge Theorem. □

Lemma 4.16. *For every integer $k \geq 2$ and $t > (2k)^k$ the following holds: if $\mathcal{H}$ is a k-uniform hypergraph, then Breaker with a bias of*

$$q > \max_{2 \leq \ell \leq k-1} \left(\Delta_\ell(\mathcal{H}) \, ((tk)^{tk} \, k^t \, v(\mathcal{H})^2)^{\frac{k}{t^{1/k}}} \right)^{\frac{1}{k-\ell}} \tag{4.10}$$

has a strategy that prevents dangerous $t(q+1)$-flowers in $\mathbf{G}(\mathcal{H}; q)$

Proof. Let $\mathcal{F} = \left\{ \bigcup_{i=1}^t H_i \mid \{H_1, \dots, H_t\} \text{ t-cluster in } \mathcal{H} \right\}$ be the hypergraph of all t-clusters in $\mathcal{H}$. First we will show that Breaker can prevent t-clusters. Given some t-cluster $H_1, \dots, H_t$ let $\ell_i = |H_i \cap \bigcup_{j=1}^{i-1} H_j|$ for all $2 \leq i \leq t$. We call $(2, \ell_2, \dots, \ell_t)$ its intersection characteristic and observe that $2 \leq \ell_i \leq k$ for $2 \leq i \leq t$. We will set $\ell_1 = 2$ for notational convenience. For any $\boldsymbol{\ell} = (\ell_1, \dots, \ell_t) \in \{2\} \times [2,k]^{t-1}$ let $\mathcal{F}(\boldsymbol{\ell})$ denote the set of edges in $\mathcal{F}$ which come from some t-cluster with the intersection characteristic $\boldsymbol{\ell}$ and observe that it is $v(\boldsymbol{\ell})$-uniform where

$$v(\boldsymbol{\ell}) = 2 + \sum_{i=1}^t (k - \ell_i) = k + \sum_{i=2}^t (k - \ell_i). \tag{4.11}$$

This follows since given any cluster $H_1, \dots, H_t$ with intersection characteristic $\boldsymbol{\ell}$ we have $|\bigcup_{i=1}^t H_i| = v(\boldsymbol{\ell})$. There is the trivial upper bound $v(\boldsymbol{\ell}) \leq tk$ for all $\boldsymbol{\ell} \in \{2\} \times [2,k]^{t-1}$. Let $L = \{\boldsymbol{\ell} : \mathcal{F}(\boldsymbol{\ell}) \neq \emptyset\} \subseteq \{2\} \times [2,k]^{t-1}$ be the set of all intersection characteristics that actually occur in $\mathcal{H}$. Now for any $\boldsymbol{\ell} \in L$ we trivially have $t \leq \binom{v(\boldsymbol{\ell})-2}{k-2}$, which we restate as the lower bound

$$v(\boldsymbol{\ell}) \geq t^{1/k} \text{ for all } \boldsymbol{\ell} \in L. \tag{4.12}$$

Now for $\boldsymbol{\ell} = (\ell_1, \dots, \ell_t) \in L$ observe that

$$\begin{aligned} |\mathcal{F}(\boldsymbol{\ell})| &\leq \binom{v(\mathcal{H})}{2} \Delta_2(\mathcal{H}) \prod_{i=2}^t \binom{k + \sum_{j=2}^{i-1}(k-\ell_i) - 2}{\ell_i - 2} \Delta_{\ell_i}(\mathcal{H}) \\ &\leq \binom{v(\mathcal{H})}{2} \binom{v(\boldsymbol{\ell}) - 2}{k-2}^{t-1} \Delta_2(\mathcal{H}) \prod_{i=2}^t \Delta_{\ell_i}(\mathcal{H}) \leq v(\mathcal{H})^2 \, (tk)^{tk} \prod_{i=1}^t \Delta_{\ell_i}(\mathcal{H}). \end{aligned}$$

Here, the first inequality is justified by observing that there are $\binom{v(\mathcal{H})}{2}$ ways to fix two common elements and at most $\Delta_2(\mathcal{H})$ ways to choose the first edge H_1 of a t-cluster. The product counts ways to add the i-th additional edge H_i for $2 \leq i \leq t$ by first fixing the intersection with the already established parts $\bigcup_{j=1}^{i-1} H_j$ and then adding one of the at most Δ_{ℓ_i} possible ways of picking H_i. The second inequality follows since by assumption $t > (2k)^k$ so that Equation (4.12) gives us $v(\boldsymbol{\ell}) \geq 2k$ from which it follows that for all $2 \leq i \leq t$ we have

$$\binom{k + \sum_{j=2}^{i-1}(k - \ell_i) - 2}{\ell_i - 2} \leq \binom{v(\boldsymbol{\ell}) - 2}{k - 2}.$$

We now want to apply the Biased Erdős–Selfridge Theorem, so we estimate

$$\begin{aligned}(q+1) \sum_{F \in \mathcal{F}} \left(\frac{1}{q+1}\right)^{|F|} &\leq (q+1) \sum_{\ell_2 \in [2,k]} \cdots \sum_{\ell_t \in [2,k]} |\mathcal{F}(\boldsymbol{\ell})| \left(\frac{1}{q+1}\right)^{v(\boldsymbol{\ell})} \\ &\leq (tk)^{tk}\, v(\mathcal{H})^2\, (q+1) \sum_{\boldsymbol{\ell} \in L} \prod_{i=1}^{t} \Delta_{\ell_i}(\mathcal{H}) \left(\frac{1}{q+1}\right)^{v(\boldsymbol{\ell})}.\end{aligned}$$

where we have just inserted the previously stated upper bound on $|\mathcal{F}(\boldsymbol{\ell})|$. We now split up the factor $(1/(q+1))^{v(\boldsymbol{\ell})}$ using Equation (4.11) to obtain

$$(q+1) \sum_{F \in \mathcal{F}} \left(\frac{1}{q+1}\right)^{|F|} \leq (tk)^{tk}\, v(\mathcal{H})^2\, \frac{1}{q+1} \sum_{\boldsymbol{\ell} \in L} \prod_{i=1}^{t} \left(\Delta_{\ell_i}(\mathcal{H}) \left(\frac{1}{q+1}\right)^{k-\ell_i} \right).$$

Note that we have $\Delta_\ell(\mathcal{H})\,(1/(q+1))^{k-\ell} = 1$ for $\ell = k$ and $\Delta_\ell(\mathcal{H})\,(1/(q+1))^{k-\ell} < 1$ for $2 \leq \ell < k$ due to the lower bound on q. Furthermore, since $\boldsymbol{\ell} \in L$ is the intersection characteristic of a t-cluster in $\mathcal{H}$, the number of indices $1 \leq i \leq t$ for which $\ell_i < k$ must be at least $\lceil v(\boldsymbol{\ell})/k \rceil \geq \left\lceil t^{1/k}/k \right\rceil$. Now, due to Equation (4.10) it follows that

$$(q+1) \sum_{F \in \mathcal{F}} \left(\frac{1}{q+1}\right)^{|F|} \leq (tk)^{tk}\, v(\mathcal{H})^2\, k^t \left(\max_{2 \leq \ell \leq k-1} \Delta_\ell(\mathcal{H}) \left(\frac{1}{q}\right)^{k-\ell} \right)^{\frac{t^{1/k}}{k}} < 1.$$

It follows, by applying the Biased Erdős–Selfridge Theorem, that using a bias of q, Breaker has a strategy to keep Maker from fully covering any t-cluster. Following this strategy, it is easy to see that Breaker will also keep Maker from creating a dangerous $t(q+1)$-flower at any point in the game. To see this, suppose that this is not the case and that Maker succeeds in creating such a dangerous flower. By repeatedly claiming the open element of this dangerous flower which has not yet been claimed and is the open element of the most almost complete solutions in the flower, Maker would be able to cover a t-cluster, as $t(q+1)/(q+1) = t$, which is a contradiction. □

4.2.3 Proof of Theorem 4.11

In order to join the previous two strategies together, we will need the following simple auxiliary statement. We include its proof for the convenience of the reader.

Lemma 4.17. *For every $q \geq 2$ and $t \geq 2$ the following holds: if F is a graph on q vertices with $e(F) < q^2/2t^2$ then it has at least $\binom{q}{t}/2$ independent sets of size t.*

Proof. The number of subsets of $V(F)$ of size t that are not independent is upper bounded by

$$e(F)\binom{q-2}{t-2} \leq e(F)\left(\frac{t^2}{q^2}\right)\binom{q}{t} < \frac{1}{2}\binom{q}{t}$$

since $e(F) < q^2/2t^2$. □

We are now ready to prove Theorem 4.11.

Proof of Theorem 4.11. Let $k \geq 2$ and $t > (2k)^k$ be given and let

$$q > 4\left((2\,v(\mathcal{H}))^{1/t}\Delta_1(\mathcal{H})\;ke\right)^{\frac{1}{k-1}}$$

as well as

$$q > 8k^2t^3\left(\max_{2\leq\ell\leq k-1}\left(\Delta_\ell(\mathcal{H})((tk)^{tk}\;k^t\;v(\mathcal{H})^2)^{\frac{k}{t^{1/k}}}\right)^{\frac{1}{k-\ell}}+2\right).$$

Breaker will play according to the following three strategies, splitting his bias as $q = q/2+q/4+q/4$. Note that in case Breaker does not need all his moves to play according to one of the strategies, he plays them arbitrarily, which cannot hurt him.

SB1: Using $q/4$ moves, he will play according to Lemma 4.15 and thus preventing Maker from occupying $\binom{q/4}{t}/2$ simple t-fans.

SB2: Using

$$\overline{q} = \max_{2\leq\ell\leq k-1}\left(\Delta_\ell(\mathcal{H})\left((tk)^{tk}k^t\;v(\mathcal{H})^2\right)^{k/t^{1/k}}\right)^{1/(k-\ell)}+1 < q/4$$

moves, he will play according to Lemma 4.16 and hence preventing dangerous $t(\overline{q}+1)$-flowers from appearing.

SB3: Using $q/2$ moves, he will occupy all open elements of any dangerous almost complete solution.

First of all, note that Maker can play according to *SB1* and *SB2* since

$$q/4 > \left((2\,v(\mathcal{H}))^{1/t}\Delta_1(\mathcal{H})\;ke\right)^{1/(k-1)}$$

and

$$\overline{q} > \max_{2\leq \ell \leq k-1} \left(2k \; \Delta_\ell(\mathcal{H}) \; \left(v(\mathcal{H})^2 \; (tk)^{tk} \right)^{\frac{k}{t^{1/k}}} \right)^{1/(k-\ell)}.$$

We can combine these strategies due to Remark 4.12 and will now prove by induction, that after each of Breaker's moves there is no dangerous almost complete solution. Clearly this implies that Breaker's strategy is indeed a winning strategy. Initially there is obviously no dangerous almost complete solution. So suppose the result is true in round $r-1$. In round r Maker claims some element w. Then every new dangerous almost complete solution must contain w. Therefore they all belong to the same dangerous fan with common element w. In order to complete the inductive step, we have to show that the size of this dangerous fan is not more than $q/2$ as Breaker can then occupy all open elements in this dangerous fan using SB3, which completes the inductive step. Indeed, using a bias of $q/2$ Breaker has a strategy that avoids dangerous $q/2$-fans at any point in the game.

Suppose Maker succeeds in occupying a dangerous $(q/2)$-fan

$$(H_1^\circ, h_1), \dots, (H_{q/2}^\circ, h_{q/2})$$

Construct an auxiliary graph F whose vertices are the almost complete solutions of this fan and an edge between (H_i°, h_i) and (H_j°, h_j) indicates that $|H_i^\circ \cap H_j^\circ| \geq 2$ where $1 \leq i < j \leq q/2$. Recall that using $\overline{q}$ moves according to *SB2*, Breaker prevents dangerous $t(\overline{q}+1)$-flowers from appearing. Therefore the maximum degree in F is bounded by $\Delta(F) \leq (t(\overline{q}+1)-2)\binom{k-1}{2} \leq t(\overline{q}+1)k^2$ and hence

$$e(F) \leq \frac{1}{2}\frac{q}{2} t(\overline{q}+1)k^2 < \frac{1}{2}\frac{(q/2)^2}{t^2}$$

by choice of q. Therefore, by Lemma 4.17, F has at least $\binom{q/4}{t}/2$ independent sets of size t. But that means that Maker occupied $\binom{q/4}{t}/2$ simple t-fans contradicting SB1. This establishes the claim that Breaker has a strategy that avoids dangerous $q/2$-fans and completes the proof. ■

4.3 Proof of Theorem 4.4 – Rado games

The goal of this chapter is to prove the statement in Theorem 4.4, that is to show that the threshold bias of the Maker-Breaker $(A, \mathbf{b})$-game on $[n]$ satisfies $q(\mathcal{S}_1(A, \mathbf{b}, n)) = \Theta\left(n^{1/m_1(A)}\right)$ for a given positive and abundant matrix $A \in \mathbb{Z}^{r\times m}$ and vector $\mathbf{b} \in \mathbb{Z}^r$.

Before we do so, let us first prove a result regarding positive matrices not covered by the previous result. This will be much easier to prove and one can see that the simple structure of non-abundant matrices strongly favors Breaker compared to the abundant case.

Proposition 4.18. *For a positive but non-abundant matrix $A \in \mathbb{Z}^{r\times m}$ and vector $\mathbf{b} \in \mathbb{Z}^r$, the threshold bias of the Maker-Breaker $(A, \mathbf{b})$-game on $[n]$ satisfies*

$$q_0\Big(\mathcal{S}_0(A, \mathbf{b}, n)\Big) \leq 2. \tag{4.13}$$

Proof of Proposition 4.18. Let us start by noting that if $S(A, \mathbf{b}) = \emptyset$, then the game hypergraph $\mathcal{S}_0(A, \mathbf{b}, n)$ is empty and the game trivially is an immediate win for Breaker. We therefore assume thate $S(A, \mathbf{b}) \neq \emptyset$. By Lemma 1.2 we may, without loss of generality, assume that the first row of A is of the shape $\mathbf{a} = (a_1, \ldots, a_m)$ where $a_i = 0$ for all $i \in \{3, \ldots, m\}$ and $a_1, a_2 \neq 0$ as well as $a_1 + a_2 \neq 0$. Writing b_1 for the first entry of $\mathbf{b}$, it follows that whenever Maker occupies some $i \in [n]$, Breaker can simply pick $(b_1 - a_1\, i)/a_2$ and $(b_1 - a_2\, i)/a_1$, if these are indeed integer values in $[n]$, and thus block Maker's ability to cover any solution. It follows that Breaker has a winning strategy with a bias of at most 2. □

Proof of Rado Games statements

We will obtain Maker's strategy through an application of Theorem 4.2 and Breaker's strategy through an application of Theorem 4.3. In order to do so, let us first state some general observations regarding the distribution of edges in the hypergraph $\mathcal{S}_0(A, \mathbf{b}, n)$ that will be needed when applying both Maker's and Breaker's criterion. Note that we previously studied essentially the same hypergraph in Chapter 3, obtaining similar results.

Lemma 4.19. *For every positive matrix $A \in \mathbb{Z}^{r\times m}$ and vector $\mathbf{b} \in \mathbb{Z}^r$ such that $\mathcal{S}(A, \mathbf{b}) \neq \emptyset$, the average degree of $\mathcal{S}_0(A, \mathbf{b}, n)$ satisfies*

$$d(\mathcal{S}_0(A, \mathbf{b}, n)) = \Theta\Big(n^{m-\mathrm{rk}(A)-1}\Big).$$

Proof. We observe that each edge in $\mathcal{S}_0(A, \mathbf{b}, n)$ can stem from at most $m!$ solutions in $S_0(A, \mathbf{b}) \cap [n]^m$, so that we have

$$|S_0(A, \mathbf{b}) \cap [n]^m|/m! \leq e(\mathcal{S}_0(A, \mathbf{b}, n)) \leq |S_0(A, \mathbf{b}) \cap [n]^m|.$$

Using Equation (1.3) and Lemma 1.4 there therefore exists a constant $c_0 = c_0(A, \mathbf{b}) > 0$ so that

$$c_0/m!\, n^{m-\mathrm{rk}(A)-1} \leq d(\mathcal{S}_0(A, \mathbf{b}, n)) \leq n^{m-\mathrm{rk}(A)-1},$$

proving the statement. □

Lemma 4.20. *For every positive matrix $A \in \mathbb{Z}^{r\times m}$, vector $\mathbf{b} \in \mathbb{Z}^r$ and $1 \leq \ell \leq m$ the maximum ℓ-degrees in $\mathcal{S}_0(A, \mathbf{b}, n)$ satisfy*

$$\Delta_\ell(\mathcal{S}_0(A, \mathbf{b}, n)) = O\Big(\max_{Q\subseteq[m],|Q|=\ell} n^{(m-\mathrm{rk}(A))-(|Q|-r_Q)}\Big).$$

Proof. We have

$$\begin{aligned}
&\Delta_\ell(\mathcal{S}_0(A, \mathbf{b}, n))\\
\leq &\max_{(x_1,\dots,x_\ell)\in[n]^\ell} \Big|\{\mathbf{x} \in S_0(A, \mathbf{b}) \cap [n]^m : \exists\, Q \subseteq [m] \text{ s.t. } \mathbf{x}^Q = (x_1, \dots, x_\ell)\}\Big|\\
\leq &\binom{m}{\ell} \max_{\substack{(x_1,\dots,x_\ell)\in[n]^\ell\\ Q\subseteq[m],|Q|=\ell}} \Big|\{\mathbf{x} \in [n]^{m-\ell} : A^{\overline{Q}} \cdot \mathbf{x}^T = \mathbf{b} - A^Q \cdot (x_1, \dots, x_\ell)^T\}\Big|\\
\leq &m^\ell \max_{\substack{Q\subseteq[m]\\|Q|=\ell}} \max_{\mathbf{b}'\in\mathbb{Z}^r} \Big| S(A^{\overline{Q}}, \mathbf{b}') \cap [n]^{m-\ell}\Big|.
\end{aligned}$$

Using Equation (1.3) as well as the fact that $|\overline{Q}| = m - |Q|$ and $r_Q = \mathrm{rk}(A) - \mathrm{rk}(A^{\overline{Q}})$, it follows that

$$\Delta_\ell(\mathcal{S}_0(A, \mathbf{b}, n)) \leq m^\ell \max_{\substack{Q\subseteq[m]\\|Q|=\ell}} n^{|\overline{Q}|-\mathrm{rk}(A^{\overline{Q}})} = m^\ell \max_{\substack{Q\subseteq[m]\\|Q|=\ell}} n^{(m-\mathrm{rk}(A))-(|Q|-r_Q)}$$

giving us the desired statement. □

Using these results, we are now ready to provide a proof of Theorem 4.4.

Proof of Theorem 4.4. We will prove that the threshold bias satisfies

$$q_0(\mathcal{S}_1(A, \mathbf{b}, n)) = \Theta(n^{1/m_1(A)}) \tag{4.14}$$

by first showing that the criteria of Theorem 4.2 are met by $\mathcal{S}_0(A, \mathbf{b}, n)$. We will then use Remark 4.12 and show that, for B_1 and $\mathbf{c}_1$ are as given by Corollary 1.13, for any $\mathfrak{p} \in \mathfrak{P}(B_1)$ either $\mathcal{S}_0(B_{1,\mathfrak{p}}, \mathbf{c}_1, n)$ meets the criteria of Theorem 4.3 or $B_{1,\mathfrak{p}}$ is non-abundant, in which case we can apply Proposition 4.18. In fact, since we are applying

Theorem 4.2 to $\mathcal{S}_0(A, \mathbf{b}, n)$ so that Maker is guaranteed to actually win with a proper solution, this also establishes that

$$q_0(\mathcal{S}_0(A, \mathbf{b}, n)) = \Theta(n^{1/m_1(A)}). \tag{4.15}$$

Maker's Strategy. Since A is abundant, we know by Lemma 1.14 that $r_Q = 0$ for any $Q \subseteq \{1, \dots, m\}$ satisfying $|Q| \leq 2$. By Lemma 4.20, it therefore follows that $\Delta_1(\mathcal{S}_0(A, \mathbf{b}, n)) = O(n^{m-\mathrm{rk}(A)-1})$ and $\Delta_2(\mathcal{S}_0(A, \mathbf{b}, n)) = O(n^{m-\mathrm{rk}(A)-2})$. Lemma 4.19 therefore immediately implies that both (M1) and (M2) hold. As $v(\mathcal{S}_0(A, \mathbf{b}, n)) = n$ and $\mathrm{rk}(A) \geq 1$, Lemma 4.19 implies that $d(\mathcal{S}_0(A, \mathbf{b}, n)) = o(n^{m-1})$, so that (M3) holds as well. It follows that Theorem 4.2 applies and establishes the desired lower bound on the threshold bias since

$$\begin{aligned}\min_{2 \leq \ell \leq m} \left(\frac{d(\mathcal{S}_0(A, \mathbf{b}, n))}{\Delta_\ell(\mathcal{S}_0(A, \mathbf{b}, n))} \right)^{\frac{1}{\ell-1}} &= \min_{2 \leq \ell \leq m} \left(\frac{\Theta\Big(n^{m-\mathrm{rk}(A)-1}\Big)}{O\Big(\max_{\substack{Q \subseteq [m] \\ |Q| = \ell}} n^{(m-\mathrm{rk}(A))-(|Q|-r_Q)}\Big)} \right)^{\frac{1}{\ell-1}} \\ &= \min_{\substack{Q \subseteq [m] \\ |Q| \geq 2}} \Omega\left(n^{\frac{|Q|-r_Q-1}{|Q|-1}}\right) = \Omega\Big(n^{1/m_1(A)}\Big).\end{aligned}$$

Breaker's Strategy. One may assume that A is strictly 1-balanced as otherwise we can replace A and $\mathbf{b}$ with B_1 and $\mathbf{c}_1$ as given by Corollary 1.13. We will now show that for any $\mathfrak{p} \in \mathfrak{P}(A)$ either $A_\mathfrak{p}$ is non-abundant, in which case we simply apply Proposition 4.18 to establish that Breaker can take care of solutions of that type, or $A_\mathfrak{p}$ is again abundant, in which case we show that $\mathcal{S}_0(A_\mathfrak{p}, \mathbf{b}, n)$ meets the criteria of Theorem 4.3 so that by Lemma 1.10 Breaker can take care of solutions of that type. By Remark 4.12 it follows that we can combine all of these individual strategies to establish that breaker wins with a bias of size $O(n^{1/m_1(A)})$.

Let us therefore show that $\mathcal{S}_0(A_\mathfrak{p}, \mathbf{b}, n)$ meets the criteria of Theorem 4.3 for any $\mathfrak{p} \in \mathfrak{P}(A)$ assuming that $A_\mathfrak{p}$ is abundant. Note that $v(\mathcal{S}_0(A_\mathfrak{p}, \mathbf{b}, n)) = n$ so that we clearly have $v(\mathcal{S}_0(A_\mathfrak{p}, \mathbf{b}, n)) = \omega(1)$. Since $A_\mathfrak{p}$ is abundant, we know by Lemma 1.14 that $r_Q = 0$ for any $Q \subseteq \{1, \dots, m\}$ satisfying $|Q| = 1$. Lemma 4.20 combined with the fact that $\Delta_1(\mathcal{S}_0(A_\mathfrak{p}, \mathbf{b}, n)) \geq d(\mathcal{S}_0(A_\mathfrak{p}, \mathbf{b}, n))$ therefore implies that

$$\Delta_1(\mathcal{S}_0(A_\mathfrak{p}, \mathbf{b}, n))^{\frac{1}{|\mathfrak{p}|-1}} = \Theta\Big(n^{\frac{|\mathfrak{p}|-\mathrm{rk}(A_\mathfrak{p})-1}{|\mathfrak{p}|-1}}\Big). \tag{4.16}$$

We note that, as $\mathrm{rk}(A_{\mathfrak{p}}) = \mathrm{rk}(A)$ and $|\mathfrak{p}| \leq m$, we have

$$\frac{|\mathfrak{p}| - \mathrm{rk}(A_{\mathfrak{p}}) - 1}{|\mathfrak{p}| - 1} \leq \frac{m - \mathrm{rk}(A) - 1}{m - 1} = \frac{1}{m_1(A)}. \tag{4.17}$$

In the last step we have used the assumption that A is strictly 1-balanced. It follows that if we can show that if $A_{\mathfrak{p}}$ satisfies main condition of Theorem 4.3, we have established the desired bias threshold.

Since A is strictly 1-balanced, we also know that for all $Q \subseteq \{1, \ldots, m\}$ satisfying $2 \leq |Q| < m$ we have $(|Q| - 1)/(|Q| - r_Q - 1) < m_1(A)$ so that

$$\begin{aligned}\frac{(m - \mathrm{rk}(A)) - (|Q| - r_Q)}{m - |Q|} &= \frac{(m - \mathrm{rk}(A) - 1) - (|Q| - r_Q - 1)}{(m-1) - (|Q| - 1)} \\ &< \frac{(m - \mathrm{rk}(A) - 1) - (|Q| - 1)/m_1(A)}{(m-1) - (|Q| - 1)} \\ &= \frac{1}{m_1(A)} \frac{1 - (|Q| - 1)/(m-1)}{1 - (|Q| - 1)(m-1)} = \frac{1}{m_1(A)}.\end{aligned}$$

It follows that there exists some $\varepsilon = \varepsilon(A) > 0$ so that for any $2 \leq \ell \leq m$ we have by Lemma 4.20 that

$$\begin{aligned}\Delta_\ell(\mathcal{S}_0(A, \mathbf{b}, n))^{\frac{1}{m-\ell}}\, n^\varepsilon &= O\Big(\max_{Q \subseteq [m],\ |Q| = \ell} n^{(m - \mathrm{rk}(A)) - (|Q| - r_Q)} \Big)^{\frac{1}{m-\ell}}\, n^\varepsilon \\ &= O\Big(\max_{Q \subseteq [m],\ |Q| = \ell} n^{\frac{(m - \mathrm{rk}(A)) - (|Q| - r_Q)}{m - |Q|} + \varepsilon} \Big) \\ &= O\Big(n^{1/m_1(A)} \Big) = O\Big(\Delta_1(\mathcal{S}_0(A, \mathbf{b}, n))^{\frac{1}{m-1}} \Big).\end{aligned}$$

Theorem 4.3 therefore applies, concluding the proof. ■

4.4 Proof of Theorem 4.5 – Small hypergraph games

Using Theorem 4.2 and Theorem 4.3 we can easily establish Theorem 4.5.

Proof of Theorem 4.5. First, observe that if $\mathcal{G}$ is a collection of $e(\mathcal{G})$ independent edges, then Maker has a winning strategy if $q < \binom{n - r(e(\mathcal{G}) - 1)}{r} / (e(\mathcal{G}) - 1)$. We may therefore assume that this is not the case. We recall that $\mathcal{H}(\mathcal{G}, n)$ was the hypergraph of all copies of $\mathcal{G}$ in $\mathcal{K}_n^{(r)}$. We observe that $\mathcal{H}(\mathcal{G}, n)$ is $e(\mathcal{G})$-uniform and clearly satisfies

$$v(\mathcal{H}(\mathcal{G}, n)) = \binom{n}{r} = \Theta(n^r) \tag{4.18}$$

as well as $e(\mathcal{H}(\mathcal{G},n)) = \binom{n}{v(\mathcal{G})}\, v(\mathcal{G})!\,/\,\mathrm{aut}(\mathcal{G}) = \Theta(n^{v(\mathcal{G})})$. In particular, it follows that

$$d(\mathcal{H}(\mathcal{G},n)) = \Theta(n^{v(\mathcal{G})-r}). \tag{4.19}$$

Lastly observe that for $1 \leq \ell \leq e(\mathcal{G})$ we have

$$\Delta_\ell(\mathcal{H}(\mathcal{G},n)) = \Theta\left(\max_{\mathcal{F}\subseteq\mathcal{G},\, e(\mathcal{F})=\ell} n^{v(\mathcal{G})-v(\mathcal{F})}\right). \tag{4.20}$$

We will now prove that the threshold bias satisfies $q(\mathcal{H}(\mathcal{G},n)) = \Theta(n^{1/m_r(\mathcal{G})})$ by showing that the criteria of Theorem 4.2 are met by $\mathcal{H}(\mathcal{G},n)$ and that the criteria of Theorem 4.3 are met by $\mathcal{H}(\mathcal{F},n)$ where $\mathcal{F}$ will be some appropriate dense subgraph of $\mathcal{G}$. The bounds on the threshold bias obtained this way will asymptotically be the same, giving the desired statement.

Maker's Strategy. Equation (4.20) implies that $\Delta_1(\mathcal{H}(\mathcal{G},n)) = \Theta(n^{v(G)-r})$ as well as $\Delta_2(\mathcal{H}(\mathcal{G},n)) = O(n^{v(G)-(r+1)})$ so that (M1) and (M2) follow due to Equation (4.19). Now, as we have already excluded the case that $\mathcal{G}$ is a collection of $e(\mathcal{G})$ independent edges, we have $v(\mathcal{G})/e(\mathcal{G}) < r$ so that $v(\mathcal{G}) - r < r\,(e(\mathcal{G}) - 1)$ and hence (M3) is satisfied by Equation (4.18) and Equation (4.19). It follows that Theorem 4.2 applies and establishes the desired lower bound on the threshold bias since

$$\begin{aligned}\min_{2\leq\ell\leq e(\mathcal{G})}\left(\frac{d(\mathcal{H}(\mathcal{G},n))}{\Delta_\ell(\mathcal{H}(\mathcal{G},n))}\right)^{\frac{1}{\ell-1}} &= \min_{2\leq\ell\leq e(\mathcal{G})}\left(\frac{\Theta\Big(n^{v(\mathcal{G})-r}\Big)}{\Theta\Big(\max_{\mathcal{F}\subseteq\mathcal{G},\, e(\mathcal{F})=\ell} n^{v(\mathcal{G})-v(\mathcal{F})}\Big)}\right)^{\frac{1}{\ell-1}} \\ &= \min_{\mathcal{F}\subseteq\mathcal{G},\, e(\mathcal{F})\geq 2} \Theta\left(n^{\frac{v(\mathcal{F})-r}{e(\mathcal{F})-1}}\right) = \Theta\Big(n^{1/m_r(\mathcal{G})}\Big).\end{aligned}$$

Breaker's Strategy. Note that we can restrict our attention to the case in which $\mathcal{G}$ is strictly r-balanced, as otherwise we can replace $\mathcal{G}$ with a strictly r-balanced subhypergraph $\mathcal{F} \subset \mathcal{G}$. Indeed, if Breaker can keep Maker from occupying $\mathcal{F}$, then he clearly also succeeds in keeping Maker from occupying a copy of $\mathcal{G}$. So we may assume that $m_r(\mathcal{G}) = (e(\mathcal{G}) - 1)/(v(\mathcal{G}) - r)$ and that $m_r(\mathcal{F}) = (e(\mathcal{F}) - 1)/(v(\mathcal{F}) - r) < m_r(\mathcal{G})$ for all subgraphs $\mathcal{F} \subsetneq \mathcal{G}$ on at least $r + 1$ vertices.

Clearly $v(\mathcal{H}(\mathcal{G},n)) = \omega(1)$ by Equation (4.18). We note that by Equation (4.20) as well as the assumption that $\mathcal{G}$ is strictly r-balanced we have

$$\Delta_1(\mathcal{H}(\mathcal{G},n))^{\frac{1}{e(\mathcal{G})-1}} = \Theta\Big(n^{\frac{v(\mathcal{G})-r}{e(\mathcal{G})-1}}\Big) = \Theta\Big(n^{1/m_r(\mathcal{G})}\Big). \tag{4.21}$$

Since $\mathcal{G}$ is strictly r-balanced, we also have that for every $2 \leq \ell \leq e(\mathcal{G})$ and every subhypergraph $\mathcal{F} \subset \mathcal{G}$ with $e(\mathcal{F}) = \ell$ edges

$$\frac{v(\mathcal{G}) - v(\mathcal{F})}{e(\mathcal{G}) - \ell} = \frac{(v(\mathcal{G}) - r) - (v(\mathcal{F}) - r)}{(e(\mathcal{G}) - 1) - (e(\mathcal{F} - 1))} = \frac{1}{m_r(\mathcal{F})} \frac{1 - \frac{v(\mathcal{F})-r}{v(\mathcal{G})-r}}{1 - \frac{e(\mathcal{F})-1}{e(\mathcal{G})-1}} < \frac{1}{m_r(\mathcal{G})}.$$

It follows that there exists a sufficiently small $\varepsilon = \varepsilon(r, \mathcal{G})$ such that for any $2 \leq \ell \leq e(\mathcal{G})$ we have by Equation (4.18), Equation (4.20) and Equation (4.21) that

$$\begin{aligned}\Delta_\ell(\mathcal{H}(\mathcal{G}, n))^{\frac{1}{e(\mathcal{G})-\ell}} \, v(\mathcal{H}(\mathcal{G}, n))^\varepsilon &= \Theta\left(\max_{\substack{\mathcal{F} \subseteq \mathcal{G} \\ e(\mathcal{F}) = \ell}} n^{\frac{v(\mathcal{G})-v(\mathcal{F})}{e(\mathcal{G})-\ell} + r\varepsilon} \right) \\ &= O\left(n^{\frac{1}{m_r(\mathcal{G})}}\right) = O\left(\Delta_1(\mathcal{H}(\mathcal{G}, n))^{\frac{1}{e(\mathcal{G})-1}}\right).\end{aligned}$$

We can therefore apply Theorem 4.3 and due to Equation (4.21) this establishes the desired upper bound on the threshold bias. ■

4.5 Further remarks

in this chapter, we have established general criteria for hypergraphs $\mathcal{H}$, which guarantee that the uniformly random Maker-strategy is essentially optimal in the biased Maker-Breaker game on $\mathcal{H}$, and applied the, to two natural games: Rado games as well as $\mathcal{G}$-building games. Let us state some further remarks regarding these two criteria.

4.5.1 Combining the Criteria for Maker and Breaker

We note that one can easily combine Theorem 4.2 and Theorem 4.3 to form the following statement giving the exact asymptotic behavior of the threshold bias for games with a hypergraph that is not dense, roughly regular, and has an appropriate separation of the ℓ-degrees from the degrees.

Corollary 4.21. *For every $k \geq 2$ the following holds. If $\mathcal{H} = (\mathcal{H}_n)_{n \in \mathbb{N}}$ is a sequence of k-uniform hypergraphs for which there exists an $\varepsilon > 0$ so that we have*

(I) $d(\mathcal{H}_n) = o\Big(v(\mathcal{H}_n)^{k-1}\Big)$,

(II) $\Delta_1(\mathcal{H}_n) = O\Big(d(\mathcal{H}_n)\Big)$,

(III) $\Delta_\ell(\mathcal{H}_n)^{\frac{1}{k-\ell}} \, v(\mathcal{H}_n)^\varepsilon = O\Big(\Delta_1(\mathcal{H}_n)^{\frac{1}{k-1}}\Big)$ *for every $2 \leq \ell \leq k - 1$,*

then the threshold biases of the games played on $\mathcal{H}_n$ satisfy

$$q(\mathcal{H}_n) = \Theta\left(d(\mathcal{H}_n)^{\frac{1}{k-1}}\right). \tag{4.22}$$

As the reader of the proofs of Theorem 4.4 and Theorem 4.5 will have noticed, this would only be applicable in the special case where the matrix or the hypergraph to be built is strictly 1- or r-balanced. For the proof of the full statement of these results, we needed the two separate statements as well as the argument that without loss of generality one can replace the matrix or the hypergraph with a denser substructure when determining a strategy for Breaker.

4.5.2 Obtaining constants

One might rightfully be interested in obtaining more precise statements involving the constant factors for the games studied in this chapter. For the triangle-building, game Chvátal and Erdős [28] established upper and lower bounds that are tight up to a constant factor of $\sqrt{2}$. Their upper bound was slightly improved by Balogh and Samotij [4], however the value of the right constant factor is still unknown.

Let us state some bounds for the 3-term arithmetic progression game, where we already established that the threshold bias is of the order $\sqrt{n}$.

Proposition 4.22. *For the threshold bias $q_0(n)$ of the 3-term arithmetic progression game played on $[n]$ we have*

$$\sqrt{\frac{n}{12} - \frac{1}{6}} \leq q_0(n) \leq \sqrt{3n}.$$

Proof. Let us first prove the upper bound by providing a winning strategy for Breaker if he is given a bias of $q \geq \sqrt{3n}$. The strategy will simply consist of blocking all possible 3-term arithmetic progressions containing Maker's last choice and one of its previous choices. As for each fixed pair of integers there are at most three 3-term arithmetic progressions containing them and Maker occupies at most $M = \lceil n/(q+1) \rceil$ integers during the course of the whole game, the number of 3-term arithmetic progressions to be blocked is never more than $3\,(M-1)$. Since

$$3\,(M-1) \leq q$$

for $q \geq \sqrt{3n}$, Breaker has enough moves in each round to occupy the (at most) one unoccupied element in each of the dangerous 3-term arithmetic progressions.

For the lower bound we use the generalization of a criterion that was developed by Beck for Maker's win in the unbiased van der Waerden game [6]. He later stated a biased version [8] and this is what we will apply here.

Theorem 4.23 (Biased Maker's Win Criterion [8])**.** *Let $\mathcal{H}$ be a hypergraph and $q \in \mathbb{N}$. Maker has a winning strategy as the first player in the q-biased game on $\mathcal{H}$ if*

$$\sum_{H \in \mathcal{H}} \left(\frac{1}{1+q} \right)^{|H|} > \frac{q^2}{(1+q)^3} \, \Delta_2(\mathcal{H}) \, v(\mathcal{H}). \tag{4.23}$$

For the hypergraph $\mathcal{H}_n$ of 3-term arithmetic progressions in $[n]$ we observe that $v(\mathcal{H}_n) = n$, $e(\mathcal{H}_n) \geq n^2/4 - n/2$, and $\Delta_2(\mathcal{H}_n) \leq 3$. Consequently with a bias of $q < \sqrt{n/12 - 1/6}$ Equation (4.23) holds for $\mathcal{H}_n$ and Theorem 4.23 provides the winning strategy for Maker. □

Observe that the constants $\sqrt{1/12}$ and $\sqrt{3}$ are only a factor 6 apart. It would be interesting to close this gap. It should be noted that one may also apply Theorem 4.23 to the k-term van der Waerden game and obtain a lower bound of the right order of magnitude on the the threshold bias for every $k \geq 3$. The ad-hoc argument for Breaker's win does not seem to generalize immediately. The analogous question for graph-building games has been posed by Bednarska and Łuczak [10]. For hypergraph-building games, the same question can of course also be asked.

Part II

Additive Structures

Two related notions capturing the additive structure of a set will be the central focus of this part. Given a set of non-negative integers A, its representation function is given by

$$r(A,n) = \#\Big\{(a_1,a_2) \in A^2 : a_1 + a_2 = n\Big\}, \tag{4.24}$$

that is it counts the number of ways to express some integer $n \in \mathbb{N}_0$ as a sum of two elements in $\mathcal{A}$. Note that these elements are not necessarily distinct and that they are counted as tuples, that is any two distinct elements contribute to the function twice.

Now let G be some additively written abelian group. The Minkowski sumset of finite subset $A, B \subset G$ is defined as

$$A + B = \{a + b : a \in A, b \in B\}, \tag{4.25}$$

that is it is the set containing all sums of two elements, one from A and one from B. When $B = A$ we will often denote it by $2A$, which should not be confused with the dilate $2 \cdot A = \{2a : a \in A\}$. Note that $2A$ consists exactly of those elements for which the representation function is non-zero. We will also write $A + x = A + \{x\}$ for any $x \in G$.

As an example of how both of these concepts can be used to obtain structural information about a set, consider Sidon sets: a set is commonly defined to be a Sidon set if all pairwise sums of its elements are distinct. One can likewise require that all pairwise differences of its elements are distinct. Note that we have previously already introduced Sidon sets in Section 1.7. We observe that one can also characterize them either as sets whose representation function has range $\{0, 1\}$ or as sets whose sumset is as large as possible for sets of fixed cardinality.

In Chapter 5 we will study a generalization of Sidon sets due to Kohayakawa, Lee, Moreira and Rödl, in which the pairwise sums of elements are required to not only be distinct but in fact far apart by a certain measure depending on some parameter. This notion is motivated by a strong connection to the density of the largest Sidon set contained in a random infinite subset of the integers. We will present new lower bounds that improve on previous results and which as a corollary give the best current bounds for Sidon sets in infinite random sets of certain density. The results of this chapter will be based on the paper 'On strong infinite Sidon and B_h sets and random sets of integers', which is joint work with David Fabian and Juanjo Rué [56].

In Chapter 6 we will turn our attention to sets of small doubling. A conjecture of Freĭman gives an exact formula for the largest volume of a set of integers in terms of its cardinality and the cardinality of its sumset. After a survey of some of the results working towards this conjecture, we will verify an additional case. We will then use these results to improve the bounds on another well-known conjecture regarding sets of small doubling in cyclic groups of prime order. The results of this chapter will be based on the papers 'Additive volume of sets contained in few arithmetic progressions', which is joint work with Gregory A. Freĭman and Oriol Serra [65], and 'A step beyond Freĭman's theorem for set addition modulo a prime', which is joint work with Pablo Candela and Oriol Serra [21].

Lastly, in Chapter 7 we will examine how close certain generalizations of the representation function of an infinite set of integers can come to being constant. A first result extends a result of Erdős and Fuchs to ordered representation functions, showing that for any infinite set of integers they do in fact have to be far from being constant. Then we take another step towards answering a question of Sárközy and Sós by classifying most weights for which the multivariate weighted representation function cannot become constant. The results of this chapter will be based on the papers 'An Erdős–Fuchs theorem for ordered representation functions', which is joint work with Gonzalo Cao-Labora and Juanjo Rué [22], and 'On a problem of Sárközy and Sós for multivariate linear forms', which is joint work with Juanjo Rué [121].

Chapter 5

Stronger Sidon Sets

We previously already surveyed extremal results for finite Sidon sets in Section 1.7, establishing that they are of size $(1 + o(1))\, n^{1/2}$. in this chapter however, we will be interested in studying the much less understood behavior of *infinite* Sidon sets.

Given some set $S \subset \mathbb{N}$, let us write $S(n) = \left|S \cap [n]\right|$ for its counting function. Sidon himself found an infinite Sidon set satisfying $S(n) = \Omega(n^{1/4})$ and Erdős [49] as well as previously Chowla and Mian [107] observed that the greedy approach yields a set satisfying $S(n) = \Omega(n^{1/3})$. Ajtai, Komlós and Szemerédi [1] improved that bound by a factor of $\log^{1/3}(n)$ and Ruzsa [128] finally overcame the exponent of 1/3 by proving the existence (through probabilistic arguments) of an infinite Sidon sequence with counting function $S(n) = n^{\sqrt{2}-1+o(1)}$. Cilleruelo [29] later gave an explicit construction of an infinite Sidon set with the same exponent as Ruzsa.

Regarding an upper bound, Erdős showed that any infinite Sidon set satisfies $\liminf_{n\to\infty} S(n)/\sqrt{n} = 0$, see [147]. Note that $\limsup_{n\to\infty} S(n)/\sqrt{n} \leq 1$ trivially follows from the finite case, while Erdős also proved the existence of an infinite Sidon set satisfying $\limsup_{n\to\infty} S(n)/\sqrt{n} \geq 1/2$ which was later improved to $1/\sqrt{2}$ by Krückeberg [100].

Strong Sidon sets

Kohayakawa, Lee, Moreira and Rödl introduced a generalization of infinite Sidon sets in [94] and further studied it in [95]. Given some fixed $0 \leq \alpha < 1$ and $\gamma \geq 1$, they define an (α, γ)-strong Sidon set to be an infinite set of integers $S \subset \mathbb{N}$ for which the pairwise sums of its elements are not just distinct, but in fact far apart by a certain measure depending on α and γ.

Definition 5.1. *Let $0 \leq \alpha < 1$ and $\gamma \geq 1$ be given. A set of integers $S \subset \mathbb{N}$ is called an* **(α,γ)-strong Sidon set** *if we have*

$$\Big|(x+w) - (y+z)\Big| \geq \gamma \max\{x^\alpha, y^\alpha, z^\alpha, w^\alpha\} \tag{5.1}$$

for every $x, y, z, w \in S$ satisfying $\{x, w\} \cap \{y, z\} = \emptyset$.

Note that for $\alpha = 0$ and $\gamma = 1$ one recovers the traditional notion of an infinite Sidon set. Also note that this definition is particular to *infinite* sets and that Kohayakawa et al. also proposed and studied finite (α,γ)-strong Sidon sets where Equation (5.1) is modified accordingly.

Regarding an upper bound, Kohayakawa et al. [95] used the bounds that they obtained for the finite case to show that any $(\alpha, 1)$-strong Sidon set S satisfies $S(n) \leq c\, n^{(1-\alpha)/2}$ for some constant $c = c(\alpha)$. Regarding a lower bound, they proved the existence of an $(\alpha, 1)$-strong Sidon set S that satisfies

$$S(n) \geq n^{(\sqrt{2}-1+o(1))/(1+32\sqrt{\alpha})} \tag{5.2}$$

as long as $0 \leq \alpha \leq 10^{-4}$. They furthermore noted that a simple greedy argument gives a construction satisfying

$$S(n) \geq n^{(1-\alpha)/3} \tag{5.3}$$

for any $0 \leq \alpha < 1$, so that the previous bound only constitutes an improvement when $\alpha \leq 5.75 \cdot 10^{-5}$. The following is the main result of this chapter, establishing a marked improvement over either of those two lower bounds.

Theorem 5.2. *For every $0 \leq \alpha < 1$ and $\gamma \geq 1$ there exists an (α, γ)-strong Sidon set $S \subset \mathbb{N}$ satisfying*

$$S(n) \geq n^{\sqrt{2+(\alpha/2)^2}-(1+\alpha/2)+o(1)}. \tag{5.4}$$

The approach taken here to prove this bound is different from that in [95], where the existence of infinite Sidon sets S with density $S(n) \geq n^{\sqrt{2}-1+o(1)}$, as originally proved by Ruzsa, is used as a black box. Instead, it is based on Cilleruelo's constructive proof of that same bound, making use of some particular properties of the family of sets defined by him.

Sidon sets in random sets

While strong Sidon sets are interesting in their own right, Kohayakawa et al. originally introduced them to study the maximum density of Sidon sets contained in randomly generated infinite sets of integers. For a fixed constant $0 < \delta \leq 1$, let R_δ denote the random subset of $\mathbb{N}$ obtained by picking each $m \in \mathbb{N}$ independently with probability

$$p_m = 1/m^{1-\delta}. \tag{5.5}$$

Note that $R(n) = n^{\delta+o(1)}$ with probability 1. Kohayakawa et al. were interested in finding

(a) the largest possible constant $f(\delta)$ such that, with probability 1, there is a Sidon set $S \subset R_\delta$ satisfying $S(n) \geq n^{f(\delta)+o(1)}$ and
(b) the smallest possible constant $g(\delta)$ such that, with probability 1, every Sidon set $S \subset R_\delta$ satisfies $S(n) \leq n^{g(\delta)+o(1)}$.

It is shown in [94] that the behavior of $f(\delta)$ and $g(\delta)$ markedly depends on whether δ falls into the first, second or last third of the interval $(0, 1]$. More precisely, they showed that

(i) if $0 < \delta \leq 1/3$ then $f(\delta) = g(\delta) = \delta$,
(ii) if $1/3 \leq \delta \leq 2/3$ then $f(\delta) = g(\delta) = 1/3$ and
(ii) if $2/3 \leq \delta \leq 1$ then $f(\delta) \geq \max\{1/3, \sqrt{2} - 1 - (1 - \delta)\}$ and $g(\delta) \leq \delta/2$.

It follows that there is only a gap between the current bounds for f and g in the last third where $2/3 \leq \delta \leq 1$.

In [95] Kohayakawa et al. established a connection between $(\alpha, 16)$-strong Sidon sets and Sidon sets in the infinite random set $R_{1-\alpha}$ by proving the following statement.

Theorem 5.3 (Theorem 12 in [95])**.** *Let* $1/2 \leq \delta \leq 1$. *If there exists an* $(1 - \delta, 16)$-*strong Sidon set* $S \subset \mathbb{N}$ *satisfying*

$$S(n) \geq n^{u(\delta)+o(1)} \tag{5.6}$$

then, with probability 1, *the random set* R_δ *contains a Sidon set* S^* *satisfying*

$$S^*(n) \geq n^{u(\delta)+o(1)}. \tag{5.7}$$

The following is an immediate corollary to Theorem 5.2 using this result. It establishes a strong improvement over the previously known lower bound for f when the parameter of the random set satisfies $5/6 < \delta < 1$.

Corollary 5.4. *For any $1/2 \leq \delta \leq 1$ there exists, with probability 1, a Sidon set S in the infinite random set R_δ satisfying*

$$S(n) \geq n^{\sqrt{(3-\delta)^2/4+\delta}-(3-\delta)/2+o(1)} \tag{5.8}$$

5.1 Proof of Theorem 5.2 – A lower bound

Ruzsa [128] and Cilleruelo [29] both based their approach on the observation that the set of primes $\mathcal{P}$ forms a multiplicative Sidon set, so that the set $\{\log p : p \in \mathcal{P}\}$ is a Sidon set in the reals. Both therefore considered sets of integers whose elements are indexed by the primes and which, through the removal of few elements, can be turned into a Sidon set. In Ruzsa's approach that removal happens through a probabilistic argument and in Cilleruelo's it is explicit.

Our starting point for proving Theorem 5.2 is the same family of infinite sets of integers $\{A_c : 0 < c \leq 1\}$ constructed by Cilleruelo. For completeness and clarity of the exposition, let us briefly recall its definition.

5.1.1 The construction

We start by fixing $0 < c \leq 1$, which, roughly speaking, determines both the growth and the 'Sidon-ness' of the set we are going to construct in a negatively correlated way. Next, we will fix an ordered set of non-zero integers $\bar{q} = (q_1, q_2, q_3, ...)$, which we will refer to as the generalized basis. Observe that, for any such sequence, one can uniquely express any given non-negative integer a in the form

$$a = x_1 + x_2\, q_1 + x_3\, q_1 q_2 + x_4\, q_1 q_2 q_3 + \cdots + x_k\, q_1 \cdots q_{k-1}, \tag{5.9}$$

where $0 \leq x_i < q_{i-1}$ for any $1 \leq i \leq k$, $x_k \neq 0$. We will refer to the numbers $x_i(a)$ as the digits of a in base $\bar{q}$ and we also write $\text{len}(a) = k = k(a)$ for its length. For notational convenience, we also let $x_i(a) = 0$ for any a when $i > \text{len}(a)$.

The particular basis $\bar{q}$ that will be fixed throughout the rest of this chapter for the construction of the Sidon set is any arbitrary sequence of the form

$$\bar{q} = (4q'_1,\, 4q'_2,\, 4q'_3,\, \ldots), \tag{5.10}$$

where each q_i' is a prime number satisfying the condition

$$2^{2i-1} < q_i' \leq 2^{2i+1}. \tag{5.11}$$

Observe that we can always find prime numbers satisfying this condition by Bertrand's Postulate.

Next, we will use $\mathcal{P}$ to denote the set of prime numbers. Writing $f(c,k) = ck^2/(\log k)^{1/2}$, we partition the set of primes into disjoint parts $\mathcal{P} = \bigcup_{k \geq 3} \mathcal{P}_{k,c}$, where for any $k \geq 3$

$$\mathcal{P}_{k,c} = \left\{ p \in \mathcal{P} : 2^{c(k-1)^2 - f(c,k-1)} < p \leq 2^{ck^2 - f(c,k)} \right\}. \tag{5.12}$$

The decisive property of f that will be used later is that $f(c,k) = o(k^2)$ but $f(c,k) = \omega(k)$. Note that, depending on the value of c, some of the initial parts may be empty. Finally, for the given generalized basis $\bar{q} = (4q_1', 4q_2', 4q_3', \dots)$ and for each $i \geq 1$, we also fix some primitive root

$$g_i = g_i(q_i') \in \mathbb{F}_{q_i'}^*. \tag{5.13}$$

We are now ready to define the set A_c. Its elements will be indexed by the set of primes, that is $A_c = \{a_p : p \in \mathcal{P}\}$, and each element a_p is constructed as follows: we first consider the unique subset $\mathcal{P}_{k,c}$ such that $p \in \mathcal{P}_{k,c}$. We set $\operatorname{len}(a_p) = k$ and let the digit $x_i(a_p)$ be given as the unique solution to the equation

$$g_i^{x_i(a_p)} \equiv p \mod q_i', \quad q_i' + 1 \leq x_i(a_p) \leq 2q_i' - 1 \tag{5.14}$$

for each $0 \leq i \leq k$. As previously already noted, we set $x_i(a_p) = 0$ for any $i > k$. Note that by construction the elements a_p are all distinct, that is $a_p \neq a_{p'}$ if $p \neq p'$.

5.1.2 Some auxiliary statements

Three important properties follow immediately from the definition of the sets A_c. The first one states that the length of any element a_p is determined by the part its indexing prime falls in.

Remark 5.5. *For any $a_p \in A_c$ we have $\operatorname{len}(a_p) = k$ if and only if $p \in \mathcal{P}_{k,c}$.*

The second important observation is that, due to the second condition in Equation (5.14) and the constant 4 in the construction of the generalized basis, one can sum up any two numbers in A_c without having to carry digits.

Remark 5.6. *We have*

$$x_i(a_{p_1} + a_{p_2}) = x_i(p_1) + x_i(p_2) \tag{5.15}$$

for any $p_1, p_2 \in \mathcal{P}$ and $i \geq 1$ and therefore also

$$\operatorname{len}(a_{p_1} + a_{p_2}) = \max\{\operatorname{len}(a_{p_1}), \operatorname{len}(a_{p_2})\}. \tag{5.16}$$

Lastly, we note that for any $a_{p_1}, a_{p_2} \in A_c$ and $i \geq 1$, one can distinguish the number of non-zero i-th digits of the summands in $a_{p_1} + a_{p_2}$ simply by considering the i-th digit of $a_{p_1} + a_{p_2}$.

Remark 5.7. *For any $i \geq 1$ and $a_{p_1}, a_{p_2} \in A_c$ with $\operatorname{len}(a_{p_1}) \geq \operatorname{len}(a_{p_2})$ we have*

$$x_i(a_{p_1} + a_{p_2}) \in \begin{cases} \{0\} & \textit{if } i > \operatorname{len}(a_{p_1}) \geq \operatorname{len}(a_{p_2}), \\ \{q_i' + 1, \dots, 2q_i' - 1\} & \textit{if } \operatorname{len}(a_{p_1}) \geq i > \operatorname{len}(a_{p_2}), \\ \{2q_i' + 2, \dots, 4q_i' - 2\} & \textit{if } \operatorname{len}(a_{p_1}) \geq \operatorname{len}(a_{p_2}) \geq i. \end{cases} \tag{5.17}$$

Observe that the sets in the three cases are clearly disjoint. It follows that $x_i(a_{p_1} + a_{p_2})$ determines the relation of $\operatorname{len}(a_{p_1})$ and $\operatorname{len}(a_{p_2})$ to i.

Let us show some additional auxiliary results regarding the basis $\bar{q}$.

Lemma 5.8. *For any $a \in \mathbb{N}$ with $k = \operatorname{len}(a)$ we have*

$$2^{k^2-1} < a < 2^{k^2+4k}. \tag{5.18}$$

Proof. By nature of the generalized basis, we have

$$4q_1' \cdots 4q_{k-1}' \leq a < 4q_1' \cdots 4q_k'$$

and therefore by Equation (5.11) we have

$$a < 2^{2k} \prod_{i=1}^{k} 2^{2i+1} = 2^{k^2+4k}$$

as well as

$$a > 2^{2k-2} \prod_{i=1}^{k-1} 2^{2i-1} = 2^{k^2-1},$$

proving the statement. □

Lemma 5.9. *For any* $\gamma \geq 1$, $0 \leq \alpha < 1$ *and* $a \in \mathbb{N}$ *with* $k = \operatorname{len}(a)$ *we have*

$$\operatorname{len}\Big(\lfloor \gamma a^\alpha \rfloor\Big) \leq \Big(\alpha k^2 + 4\alpha k + \log_2 \gamma\Big)^{1/2}. \tag{5.19}$$

Proof. We have

$$\gamma a^\alpha < \gamma \Big(4q_1' \cdots 4q_k'\Big)^\alpha \leq \gamma \, 2^{\alpha(k^2+4k)}$$

so that the statement follows by the lower bound in Lemma 5.8. □

We conclude this part by stating the asymptotic growth of the set A_c, which was already observed and proved in [29].

Proposition 5.10. *For any* $0 < c \leq 1$ *we have* $A_c(n) = n^{c+o(1)}$.

Proof. Let $\pi(n)$ denote the prime counting function, which by the Prime Number Theorem satisfies $\pi(n) = n/\log n \, (1 + o(1))$. Lemma 5.8 and the definition of $\mathcal{P}_{k,c}$ in Equation (5.11) therefore imply that for $k_0 = \lfloor (\log_2(n+4))^{1/2} \rfloor$ we have

$$A_c(n) \geq |\mathcal{P}_{3,c}| + \cdots + |\mathcal{P}_{k_0,c}| \geq \pi\left(2^{c(k_0-1)^2(1+o(1))}\right) = n^{c+o(1)}$$

and for $k_1 = \lceil (\log_2 n)^{1/2} \rceil - 2$ we have

$$A_c(n) \leq |\mathcal{P}_{3,c}| + \cdots + |\mathcal{P}_{k_1,c}| \leq \pi\left(2^{ck_1^2(1+o(1))}\right) = n^{c+o(1)},$$

giving the desired statement. □

5.1.3 Proof of Theorem 5.2

The following statement is central to the proof of Theorem 5.2. It is reminiscent of Proposition 3 in [29].

Proposition 5.11. *Let* $0 \leq \alpha < 1$, $\gamma \geq 1$ *and* $0 < c < 1 - \alpha$. *Assume that there are elements* $a_{p_1}, a_{p_1'}, a_{p_2}, a_{p_2'} \in A_c$ *satisfying* $a_{p_1} \geq a_{p_2}$, $a_{p_1'} \geq a_{p_2'}$, $a_{p_1} > a_{p_1'}$ *and* $a_{p_2} \neq a_{p_2'}$ *as well as*

$$\Big|(a_{p_1} + a_{p_2}) - (a_{p_1'} + a_{p_2'})\Big| < \gamma \, a_{p_1}^\alpha. \tag{5.20}$$

Writing $k_i = \operatorname{len}(a_{p_i})$ *and* $k_i' = \operatorname{len}(a_{p_i'})$ *for* $i \in \{1, 2\}$ *as well as*

$$\ell = \max\Big\{i \in \mathbb{N} : x_i(p_1) + x_i(p_2) \neq x_i(p_1') + x_i(p_2')\Big\},$$

there exists some $k_0 = k_0(c, \alpha, \gamma)$ *such that either* $k_1 < k_0$ *or* *(i)* $\ell^2 \leq \alpha k_1^2 + 9k_1 + \log_2 \gamma$,

(ii) $k_1 = k_1' \geq k_2 = k_2' \geq \ell$,

(iii) $q'_{k_2+1} \cdots q'_{k_1} \mid (p_1 - p_1)$ *and* $k_2^2 \geq (1-c)k_1^2$ *as well as*

(iv) $q'_{\ell+1} \cdots q'_{k_2} \mid (p_1 p_2 - p_1' p_2')\,(p_1 - p_1')$ *and* $\ell^2 \geq (1-c)k_2^2 - ck_1^2$.

Proof. Since $0 < c < 1 - \alpha$, we can choose $k_0 = k_0(c, \alpha, \gamma)$ large enough such that

$$\alpha k_0^2 + 9k_0 + \log_2 \gamma < (1-c)\,k_0^2. \tag{5.21}$$

Let $a_{p_1}, a_{p_1'}, a_{p_2}, a_{p_2'} \in A_c$ now be some elements satisfying the requirements of the proposition as well as $k_1 \geq k_0$. By definition of ℓ we have

$$\ell - 1 \leq \operatorname{len}\Big(|(a_{p_1} + a_{p_2}) - (a_{p_1'} + a_{p_2'})|\Big).$$

By assumption of the proposition, by the fact that $\operatorname{len}(n)$ is an increasing function in n and by Lemma 5.9, we therefore have

$$\ell \leq \operatorname{len}\Big(\lfloor \gamma a_{p_1}^{\alpha} \rfloor\Big) + 1 \leq (\alpha k_1^2 + 4\alpha k_1 + \log_2 \gamma)^{1/2} + 1,$$

which implies part (i). To see that part (ii) holds, we note that by Remark 5.7 we must have $k_i = k_i'$ if $\ell < \max\{k_i, k_i'\}$ for $i \in \{1, 2\}$. By part (i), our choice of k_0 and the assumption that $k_1 \geq k_0$, it follows that $k_1 = k_1' > \ell$. In order to prove that $\ell < \max\{k_2, k_2'\}$ to conclude part (ii), we first note that by choice of ℓ, we have

$$g_i^{x_i(p_1)} \equiv g_i^{x_i(p_1')} \mod q_i'$$

for any $\max\{\ell, k_2\} < i \leq k_1$. By the construction of the digits of the elements in our set and by the previous observation, we therefore get

$$p_1 \equiv p_1' \mod q_i'$$

for any $\max\{\ell, k_2\} < i \leq k_1$. Since the q_i' are distinct primes, it follows that

$$p_1 \equiv p_1' \mod q'_{\max\{\ell,k_2\}+1} \cdots q'_{k_1}. \tag{5.22}$$

By Equation (5.11) and Equation (5.12) it follows that

$$2^{ck_1^2} \geq |p_1 - p_1'| \geq q'_{\max\{\ell,k_2\}+1} \cdots q'_{k_1} > 2^{k_1^2 - \max\{\ell,k_2\}^2}$$

and hence

$$\max\{\ell, k_2\}^2 \geq (1-c)k_1^2. \tag{5.23}$$

By part (i), our choice of k_0 and the assumption that $k_1 \geq k_0$, it follows that $k_2 > \ell$. From this we cannot only conclude part (ii) as previously observed, but also obtain part (iii) from Equation (5.22) and Equation (5.23).

Following the same arguments as just laid out to prove part (iii), we also have

$$g_i^{x_i(p_1)+x_i(p_2)} \equiv g_i^{x_i(p_1')+x_i(p_2')} \mod q_i'$$

and therefore

$$p_1 p_2 \equiv p_1' p_2' \mod q_i'$$

for any $\ell < i \leq k_2$. It follows that

$$p_1 p_2 \equiv p_1' p_2' \mod q_{\ell+1}' \cdots q_{k_2}'. \tag{5.24}$$

Again by Equation (5.11) and Equation (5.12) it follows that

$$2^{c(k_1^2+k_2^2)} \geq |p_1 p_2 - p_1' p_2'| \geq q_{\ell+1}' \cdots q_{k_2}' > 2^{k_2^2-\ell^2}$$

and hence

$$\ell^2 \geq c k_1^2 + (1-c)k_2^2. \tag{5.25}$$

We obtain part (iv) from Equation (5.24) and Equation (5.25). □

Using this proposition, we are now ready to prove Theorem 5.2.

Proof of Theorem 5.2. Choosing

$$c = \sqrt{2 + (\alpha/2)^2} - (1 + \alpha/2), \tag{5.26}$$

the set A_c satisfies the growth stated in Theorem 5.2 by Proposition 5.10. However, it is unfortunately not guaranteed to be an (α, γ)-strong Sidon set. The plan is to therefore remove a_{p_1} for every $a_{p_1}, a_{p_1'}, a_{p_2}, a_{p_2'} \in A_c$ satisfying $\{a_{p_1} \geq a_{p_2}\} \cap \{a_{p_1'} \geq a_{p_2'}\} = \emptyset$ and $a_{p_1} > a_{p_1'}$ which violate the condition in Equation (5.1). This removal turns the initial set into an (α, γ)-strong Sidon set. Using Proposition 5.11, we will show that this alteration does not impact the asymptotic growth of the infinite set. Note that we can in fact apply Proposition 5.11, since $0 < c < 1 - \alpha$ when c is as given by

Equation (5.26).

First, let k_0 be as in Proposition 5.11 and choose k_0' large enough such that $k_0' \geq k_0$ and

$$2(1-c)(k_0'-1)^2 > (1+\alpha)k_0^2 + 9k_0 + \log_2 \gamma. \tag{5.27}$$

Note that this is possible since $2(1-c) > 1+\alpha$ when c is as given by Equation (5.26). Now for $k_1 \geq k_0'$, let $\mathcal{B}_{k_1}$ denote the set of all prime numbers $p_1 \in \mathcal{P}_{k_1,c}$ for which there exist $a_{p_1}, a_{p_1'}, a_{p_2}, a_{p_2'} \in A_c$ such that $\{a_{p_1} \geq a_{p_2}\} \cap \{a_{p_1'} \geq a_{p_2'}\} = \emptyset$, $a_{p_1} > a_{p_1'}$ and

$$\Big|(a_{p_1} + a_{p_2}) - (a_{p_1'} + a_{p_2'})\Big| < \gamma\, a_{p_1}^{\alpha}. \tag{5.28}$$

Clearly for

$$\mathcal{P}^* = \bigcup_{k_1 \geq k_0'} \Big(\mathcal{P}_{k_1,c} \setminus \mathcal{B}_{k_1}\Big)$$

the set $S = \{a_p : p \in \mathcal{P}^*\}$ is an (α, γ)-strong Sidon set. If we can show that

$$\Big|\mathcal{B}_{k_1}\Big| = o(|\mathcal{P}_{k_1,c}|) \tag{5.29}$$

as k_1 tends to infinity, then $S(n) = n^{c+o(1)}$ follows as desired, proving the statement.

Writing $Q_1 = Q_1(\ell, k_2) = q_{\ell+1}' \cdots q_{k_2}'$ and $Q_2 = Q_2(k_2, k_1) = q_{k_2+1}' \cdots q_{k_1}'$, we note that

$$p_1(p_1 - p_2') = \frac{p_1 p_2 - p_1' p_2'}{Q_1}\, Q_1 + \frac{(p_1' - p_1)p_2'}{Q_2}\, Q_2 = s_1 Q_1 + s_2 Q_2 \tag{5.30}$$

where $s_1 = s_1(p_1, p_2, p_1', p_2') = (p_1p_2 - p_1'p_2')/Q_1$ and $s_2 = s_2(p_1, p_2, p_1', p_2') = (p_1' - p_1)p_2'/Q_2$ are integers due to the divisibility statements in parts (iii) and (iv) of Proposition 5.11. They furthermore satisfy

$$1 \leq |s_i| \leq 2^{c(k_1^2+k_2^2)-f(c,k_1)}/Q_i \tag{5.31}$$

for $i \in \{1, 2\}$ by Equation (5.12). Writing

$$S_{k_1,k_2,\ell} = \Big\{s_1 Q_1 + s_2 Q_2 : 1 \leq |s_i| \leq 2^{c(k_1^2+k_2^2)-f(c,k_1)}/Q_i \text{ for } i \in \{1,2\}\Big\} \setminus \{0\} \tag{5.32}$$

as well as

$$T_{k_1} = \Big\{(k_2, \ell) \in \mathbb{N}_0^2 : (1-c)k_2^2 - ck_1^2 \leq \ell^2 \leq \alpha k_1^2 + 9k_1 + \log_2 \gamma\Big\}, \tag{5.33}$$

then by Equation (5.30) as well as parts (i), (iii) and (iv) of Proposition 5.11 we have

$$\mathcal{B}_{k_1} \subseteq \{p_1 \in \mathcal{P}_{k_1,c} : \exists (k_2, \ell) \in T_{k_1}, s \in S_{k_1,k_2,\ell} \text{ such that } p_1 \mid s\}. \tag{5.34}$$

We note that no integer $s \in S_{k_1,k_2,\ell}$ can be divided by two distinct primes $p', p'' \in \mathcal{P}_{k_1,c}$ if $(k_2, \ell) \in T_{k_1}$ as otherwise

$$2^{2c(k_1-1)^2-2f(c,k_1-1)} \leq p'p'' \leq |s| \leq 2^{c(k_1^2+k_2^2)-2f(c,k_1)+1}$$

so that

$$2c(k_1-1)^2 \leq \frac{c}{1-c}((1+\alpha)k_1^2 + 9k_1 + \log_2 \gamma)$$

contradicting Equation (5.27) since $k_1 \geq k_0'$. Using the estimate

$$Q_1 Q_2 = q'_{\ell+1} \cdots q'_{k_1} > 2^{k_1^2 - \ell^2},$$

we can therefore bound the size of $\mathcal{B}_{k_1}$ by

$$\begin{aligned} |\mathcal{B}_{k_1}| &\leq \sum_{(k_2,\ell)\in T_{k_1}} |S_{k_1,k_2,\ell}| \leq \sum_{(k_2,\ell)\in T_{k_1}} \frac{2^{2c(k_1^2+k_2^2)-2f(c,k_1)+2}}{Q_1 Q_2} \\ &\leq 2^{(2c-1)k_1^2-2f(c,k_1)+2} \sum_{(k_2,\ell)\in T_{k_1}} 2^{2ck_2^2+\ell^2} \\ &= 2^{\frac{3c-1+(1+c)\alpha}{1-c}k_1^2 + \frac{1+c}{1-c}(9k_1+\log_2\gamma)-2f(c,k_1)+O(\log k_1)+2}. \end{aligned}$$

By the Prime Number Theorem, we can estimate the size of $\mathcal{P}_{k_1,c}$ by

$$\begin{aligned} |\mathcal{P}_{k_1,c}| &= \pi\Big(2^{ck_1^2-f(c,k_1)}\Big) - \pi\Big(2^{c(k_1-1)^2-f(c,k_1-1)}\Big) \\ &= \Theta\Big(2^{ck_1^2-f(c,k_1)-\log_2(ck_1^2-f(c,k_1))}\Big). \end{aligned}$$

Since $(1-c)c = 3c-1+(1+c)\alpha$ when c satisfies Equation (5.26) and since $f(c,k_1) = \omega(k_1)$, we get that $\mathcal{B}_{k_1} = o(\mathcal{P}_{k_1,c})$, concluding the proof. ■

5.2 Further remarks

The lower bound obtained in Theorem 5.2 appears to be a natural extension of the results of Ruzsa and Cilleruelo. Consequently, any advance in improving upon the lower bound for strong infinite Sidon sets would probably have to come as a result of an improvement on the bound in the case of normal 'non-strong' Sidon sets. Unfortunately, this problem has proved surprisingly defiant despite a fair amount of attention. Regarding the upper bound, Kohayakawa et al. [95] also asked if it can be strengthened for (α, γ)-strong Sidon sets along the lines of the results of Erdős and Turán as well as Stöhr [147] mentioned in the introduction.

It should be noted that in [56] an extension of Theorem 5.2 to B_h sets is also obtained. Here one considers h-fold sums, rather than just two-fold as is the case for Sidon sets. An equivalent statement to Corollary 5.4 can be made for B_h sets as well, though in this case no other bounds are known. However, it is reasonable to believe that using results of Dellamonica et al. [40, 41] for the case of finite random sets, one can establish exact exponents whenever $0 < \delta < h/(2h-1)$. Note that Kohayakawa et al. [94] also made use of the case of finite random sets established in [96].

Chapter 6

Structure Through Small Doubling

The doubling of A is given by the quotient of the cardinalities of the sumset and the set itself, that is

$$\sigma(A) = |2A|/|A|. \tag{6.1}$$

We observe that the cardinality of the sumset of any finite set of integers $A \subset \mathbb{Z}$ satisfies

$$2|A| - 1 \leq |2A| \leq \binom{|A|}{2} + |A|, \tag{6.2}$$

so that the doubling satisfies $2 - o(1) \leq \sigma(A) \leq |A|/2 + o(1)$. Both of the bounds in Equation (6.2) are tight: it is an easy exercise to verify that sets attaining the lower bound have to be arithmetic progressions, while sets attaining the upper bound are the Sidon sets discussed in Chapter 5.

In his original 1964 monograph [60], later translated into English in 1973 [61], Freĭman proved what is now one of the central results in Additive Combinatorics. It states that any set of integers $A \subseteq \mathbb{Z}$ with doubling K can be efficiently covered by a generalized arithmetic progression, that is a set of the form

$$\{a_0 + \ell_1 a_1 + \ldots + \ell_d a_d : 0 \leq \ell_i < L_i\}. \tag{6.3}$$

for some $d \in \mathbb{N}_0$, $a_0, a_1 \ldots, a_d \in \mathbb{Z}$ and $L_1, \ldots, L_d \in \mathbb{N} \setminus \{1\}$. Here d is the dimension of the progression and $L_1 L_2 \cdots L_d$ its size. Note that the cardinality of a generalized arithmetic progression does not have to align with its size, but we say that it is proper if it does.

Freĭman's Theorem. *For any $K > 1$ there are $d = d(K)$ and $f = f(K)$ such that* *any set $A \subseteq \mathbb{Z}$ with doubling $\sigma(A) = K$ is contained in some d-dimensional arithmetic progression of size at most $f|A|$.*

Ruzsa [125] later obtained what is now considered to be the canonical proof of this result. He also obtained the first usable bounds for $d(K)$ and $f(K)$, which were then improved by Chang [24], Sanders [129] and Schoen [137], resulting in the current best bounds of $d(K) \leq K^{1+C(\log K)^{-1/2}}$ and $f(K) \leq \exp(K^{1+C(\log K)^{-1/2}})$ for some absolute constant $C > 0$. Freĭman's theorem was also generalized to arbitrary abelian groups in [72], where generalized arithmetic progressions are replaced with coset progressions, and to solvable groups of bounded derived length [153], where coset progressions are replaced with coset nilprogressions. For arbitrary non-abelian groups, the problem of establishing a complete Freĭman–type result is still very much open, see for example [17, 155, 156].

A conjecture of Freĭman concerning the additive volume

At a conference in Toronto in 2008, Freĭman proposed a precise formula for describing the largest possible volume of a set of integers $A \subset \mathbb{Z}$ in terms of a very specific parameterisation of the cardinality of its sumset [63]. In order to properly state this conjecture as well as its notion of volume and to contrast it with Freĭman's Theorem, we need to introduce some common definitions.

Given abelian groups G and G', two sets $A \subset G$ and $B \subset G'$ are F_2-isomorphic if there is a bijection $\phi : A \to B$ such that

$$x + y = z + t \quad \Leftrightarrow \quad \phi(x) + \phi(y) = \phi(z) + \phi(t) \tag{6.4}$$

for every $x, y, z, t \in A$. One can think of this as a generalization of a group isomorphism for which only operations of depth at most 2 are required to be preserved. The additive dimension $\dim(A)$ of a set of integers $A \subset \mathbb{Z}$ is defined to be the largest $d \in \mathbb{N}$ such that there exists some F_2-isomorphic $B \subset \mathbb{Z}^d$ which is not contained in a hyperplane. Consider the following examples: the set $A = \{0, 1, \ldots, k-2\} \cup \{x\} \subset \mathbb{Z}$ is F_2-isomorphic to $\{(0,0), (0,1), \ldots, (0, k-1), (1,0)\} \subset \mathbb{Z}^2$ as long as $x \geq 2k-3$. In particular, the set is 2-dimensional when $x \geq 2k-3$ and 1-dimensional when $x < 2k-3$. One can also consider subsets of groups with torsion. The set $A = \{0, 1, \ldots, k-1\} \subseteq \mathbb{Z}_n$ is F_2-isomorphic to any k-term arithmetic progression in $\mathbb{Z}$ if $n \geq 2k-1$. It is not F_2-isomorphic to any subsets of a torsion-free group if $n < 2k-1$. Note that the concept of additive dimension can also be extended to arbitrary abelian groups, though it requires the introduction of the concept of Universal Ambient Groups, see for example Section 5.5 in [154].

We note that for any d-dimensional set $A \subseteq \mathbb{Z}$ we have, by results of Freĭman [61] as well as Konyagin and Lev [97], that

$$(d+1)|A| - \binom{d+1}{2} \leq |2A| \leq \binom{|A|}{2} + d + 1. \tag{6.5}$$

We observe that the bounds in Equation (6.2) are implied by this since in general $1 \leq \dim(A) \leq |A| - 1$.

Definition 6.1. *The additive volume* $\mathrm{vol}(A)$ *of a d-dimensional set A is the minimum number of lattice points contained in the convex hull among all sets in $\mathbb{Z}^d$ that are F_2-isomorphic to A.*

Note that for any 1-dimensional set of integers $A \subset \mathbb{Z}$ satisfying $\min(A) = 0$ as well as $\gcd(A) = 1$, we have $\mathrm{vol}(A) = \max(A) + 1$. It is clear that any set of integers $A \subset \mathbb{Z}$ can be F_2-isomorphically be mapped to a set satisfying these conditions, so we may often assume them without loss of generality. In this case, we will say that the set is in normal form. This also implies that the additive volume of a 1-dimensional set is the same as its minimal covering by a (non-generalized) arithmetic progression.

We are interested in obtaining an upper bound for the additive volume of a set $A \subset \mathbb{Z}$ in terms of its cardinality $|A|$, the cardinality of its sumset $|2A|$ and its dimension $\dim(A)$. A set A is said to be extremal if its additive volume is as large as possible for a set of that cardinality, dimension and cardinality of the sumset. The following is a more general and slightly reformulated version of the previously mentioned conjecture of Freĭman. Its notable addition is that it takes the dimension of a set into consideration.

Conjecture 6.2. *Any d-dimensional set $A \subset \mathbb{Z}$ satisfies*

$$\mathrm{vol}(A) \leq 2^{c-1} \left(|A| - c + b \right) + 1, \tag{6.6}$$

where $c = c(|A|, |2A|, d)$ and $b = b(|A|, |2A|, d)$ are the unique integers satisfying

$$|2A| = (c+d)|A| - \binom{c+d+1}{2} + b + d + 1 \tag{6.7}$$

as well as $1 \leq c \leq |A| - d - 1$ and $0 \leq b \leq |A| - d - c - 1$.

We note that both this conjecture and Freĭman's Theorem are concerned with finding an efficient covering of sets with given doubling. Where they differ, besides their specificity, is firstly that the notion of covering used in the conjecture is decidedly more geometric than that of a generalized arithmetic progression. To illustrate this difference, observe that the previously mentioned 2-dimensional set $A = \{0, 1, \ldots, k - 2\} \cup \{x\} \subset \mathbb{Z}$ with $x \geq 2k - 3$ has additive volume $k = |A|$ but cannot be contained in a 2-dimensional generalized arithmetic progression of size smaller than $2|A| - 2$. Secondly, in the updated version of the conjecture presented here, the dimension of the covering structure precisely aligns with that of the set, whereas in the Freĭman's Theorem a set might be covered by an arithmetic progression of vastly bigger dimension than the set itself. For example, the set $A = \{0, 1, 2, 4, \ldots, 2^{k-2}\}$ has doubling $\sigma(A) = (k-1)/2+o(1)$, that is as large as possible for a 1-dimensional set. Both Conjecture 6.2 and Freĭman's Theorem with Schoen's bounds give a covering structure for A whose size is exponential in $k = |A|$ (as is clearly necessary) but while the generalized arithmetic progression may have dimension around $|A|/2$, the conjecture correctly states that the covering structure should be of the same dimension as the set it is covering.

There are examples showing that the bounds Conjecture 6.2 would be tight if the statement is true.

Example 6.3. *Consider the set*

$$A_{k,c,b,1} = \{0, 1, \ldots, k - c - 1\} \cup \{2^i(k - c + b) : 0 \leq i < c\} \subset \mathbb{Z}$$

of cardinality $k \geq 3$ *where* $1 \leq c \leq k - 2$ *and* $0 \leq b \leq k - c - 2$*. This example has additive dimension* 1 *and additive volume* $2^{c-1}(k - c + b) + 1$ *while the cardinality of its sumset matches Equation* (6.7)*. Regarding an examples for higher dimensions, let* $d \geq 1$ *and consider the set*

$$A_{k,c,b,d} = \{(a, 0, \ldots, 0) : a \in A_{k-d+1,c,b,1}\} \cup \{\mathbf{e}_i : 2 \leq i \leq d\} \subset \mathbb{Z}^d$$

of cardinality $k \geq d + 2$ *where where* $\mathbf{e}_i$ *is the* i*-the standard basis vector in* $\mathbb{Z}^d$ *and* $1 \leq c \leq k - d - 1$ *as well as* $0 \leq b \leq k - c - d - 1$*. The set has additive dimension* d *and additive volume* $2^{c-1}(|A| - c + b) + 1$ *while the cardinality of its sumset again matches Equation* (6.7)*.*

Furthermore, the veracity of the conjecture has been established in a few cases, most notably by Freĭman [61] for one–dimensional sets satisfying $T \leq 3k - 4$, that is either, $c = 1$ and any admissible b or $c = 2$ and $b = 0$. This is commonly referred to as Freĭman's $3k - 4$ theorem.

Theorem 6.4 (Theorem 1.9 in [61])**.** *Any $A \subseteq \mathbb{Z}$ satisfying $|2A| \leq 3|A| - 4$ is contained in an arithmetic progression $P \subseteq \mathbb{Z}$ of size at most $|2A| - |A| + 1$.*

Note that any 2-dimensional set $A \subseteq \mathbb{Z}$ satisfies $|2A| \geq 3|A| - 3$, which is why this statement did not have to take the dimension of the set into account. A more precise structural description of extremal sets satisfying the conditions of Theorem 6.4 was given in [62]. Further particular cases for which Conjecture 6.2 has been verified are the following:

1. by Freĭman [61] as well as by Hamidoune and Plagne [76] for 1-dimensional sets if $T = 3k - 3$, that is $c = 2$ and $b = 1$, with a structural description of the extremal case due to Jin [90],
2. by Freĭman [61] for 2-dimensional sets satisfying $k \geq 10$ and $T \leq 10/3\, k - 6$, that is $c = 1$ and $0 \leq b \leq k/3 - 2$, with a structural description,
3. by Jin [89] in the case of large 1-dimensional sets satisfying $T \leq (3 + \epsilon)k$, that is $c = 2$ and $0 \leq b \leq \epsilon k$, for some $\epsilon > 0$ using tools from non-standard analysis,
4. by Stanchescu [145, 146] for any d-dimensional set satisfying $c = 1$ and $b = 0$,
5. by Freĭman and Serra [64] for a class of 1-dimensional sets called chains, which can be seen as extremal sets build by a greedy algorithm.

A step beyond the $3k - 4$ theorem in the integers

As previously mentioned, Conjecture 6.2 has been proved for sets A with doubling $|2A| \leq 3|A| - 3$ and the structure of extremal sets in this range is well understood. In spite of many efforts, not much is known about the exact maximum volume of sets with doubling at least $3|A| - 2$. In particular, the following describes the next 'regime' for 1-dimensional sets in Conjecture 6.2 after that described by Theorem 6.4.

Conjecture 6.5. *For any 1-dimensional set $A \subseteq \mathbb{Z}$ satisfying $|2A| = 3|A| - 4 + b$ for some $0 \leq b \leq |A| - 4$ there exists an arithmetic progression $P \subseteq \mathbb{Z}$ of size at most $2(|A| - 2 + b) + 1$ containing A.*

In order to give further evidence towards the validity of this conjecture, let us consider sets $A \subseteq \mathbb{Z}$ that are the union of three segments, that is $A = P_1 \cup P_2 \cup P_3$ where $P_i = [a_i, b_i]$ for $i \in \{1, 2, 3\}$ for some $a_1 \leq b_1 < a_2 - 1 \leq b_2 - 1 < a_3 - 2 \leq b_3 - 2$.

We note that in this case $|2A| \leq 4|A| - 6$ where equality holds if and only if the sums of the segments pairwise distinct, in which case the set A is 3-dimensional. The case of dimension 2 is likewise easy to describe, so let us turn our attention towards verifying Conjecture 6.5 for the particular case of 1-dimensional sets consisting of three segments and satisfying $|2A| \leq 4|A| - 8$, also giving a structural description of the extremal cases.

Theorem 6.6. *Let $A \subseteq \mathbb{Z}$ consisting of three segments and satisfying $|A| \geq 8$ as well as $|2A| > 3|A| - 4$. If A is 1-dimensional, then for $b = |2A| - (3|A| - 4)$ it is isomorphic to either*

(i) $([0, |A| + b - 2] \setminus [1, b]) \cup \{2(|A| + b - 2)\}$ *or*
(ii) $([0, |A| + b + i - 3] \setminus [|A| - i - 1, |A| + b - 3]) \cup \{2(|A| + b - 2)\}$

where $1 \leq i \leq \max\{|A|/3, (|A| - 1 - b)/2\}$. If A is 2-dimensional, then for $b = |2A| - (3|A| - 3)$ it is isomorphic to

(iii) $([0, |A| + b - 2] \setminus [1, b]) \times \{0\} \cup \{(0, 1)\} \subset \mathbb{Z}^2$

and if A is 3-dimensional, then $|2A| = 4|A| - 6$ and it is isomorphic to

(iv) $([0, k_1 - 1] \times \{(0, 0)\}) \cup ([0, k_2 - 1] \times \{(0, 1)\}) \cup ([0, k_2 - 1] \times \{(1, 0)\}) \subset \mathbb{Z}^3$

where $k_1, k_2, k_3 \geq 1$ and $k_1 + k_2 + k_3 = |A|$.

Observe that the 1-dimensional examples given in Example 6.3 for $c = 2$, as well as most other constructions one would naturally come up with, are in fact F_2-isomorphic to sets consisting of exactly three segments, lending some motivation to this result. The extremal sets described in Theorem 6.6 are illustrated for $|A| = 11$ and $|2A| = 3|A| - 1$ in Figure 6.1.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
●	○	○	○	●	●	●	●	●	●	●	●	●	○	○	○	○	○	○	○	○	○	○	○	●
●	●	●	●	●	●	●	●	●	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	●
●	●	●	●	●	●	●	●	○	○	○	○	●	●	○	○	○	○	○	○	○	○	○	○	●
●	●	●	●	●	●	●	○	○	○	○	○	●	●	●	○	○	○	○	○	○	○	○	○	●

Figure 6.1: *All extremal 1-dimensional sets satisfying $|A| = 11$ and $|2A| = 3|A| - 1$.*

Unfortunately, not all extremal sets satisfying $|2A| \leq 4|A| - 8$ will be the disjoint union of three segment. As an example, consider the 1-dimensional set given by

$$A = \{0, 1, 2, 3, 5, 10, 11, 20\} \tag{6.8}$$

which fits the parameters $k = 8$, $c = 2$ and $b = 4$. Let us therefore also state the following partial step towards Conjecture 6.5 that will be derived from a statement of Deshouiller and Freĭman [42].

Proposition 6.7. *Any* 1*-dimensional set* $A \subset \mathbb{Z}$ *satisfying* $|2A| \leq 3.04|A| - 3$ *can be covered by an arithmetic progression of length at most* $10^9\,|A|$*.*

Of course, both Theorem 6.6 and Proposition 6.7 are unfortunately far removed from answering Conjecture 6.5. However, let us now turn our attention towards covering statement for sets of small doubling in groups of prime order (rather than in the integers), which will lend some additional motivation as to why one might be interested in even partial answers to Conjecture 6.5.

A step beyond the $2.4k - 8$ theorem in groups of prime order

Statements regarding the structure of subsets $\mathcal{A}$ of the cyclic group $\mathbb{Z}_p$, where p is a prime, in the spirit of Conjecture 6.2 and Theorem 6.4 have proved to be significantly harder to obtain than their corresponding integer counterparts. Even the most basic of statements in the integers, namely that $|2A| \geq 2|A| - 1$ with equality if and only if A is an arithmetic progression, requires a non-trivial amount of effort to establish in $\mathbb{Z}_p$. The lower bound was established by Davenport [37], though it was later discovered that the statement hat previously been proved in 1813 by Cauchy [23]. The structural description of sets where equality holds is due to Vosper [159].

Vosper's Theorem. *Any set* $\mathcal{A} \subseteq \mathbb{Z}_p$ *satisfies* $|2\mathcal{A}| \geq \min\{2|\mathcal{A}| - 1, p\}$ *and we have* $|2\mathcal{A}| = 2|\mathcal{A}| - 1 \leq p - 2$ *if and only if* $\mathcal{A}$ *is an arithmetic progression.*

Note that the fact that p is prime here is crucial. Comparable statements in cyclic groups of arbitrary order have for example been made by Kneser [93] and Kemperman [91].

An equivalent statement to that of Theorem 6.4 for subsets of the cyclic group $\mathcal{A} \subseteq \mathbb{Z}_p$, where p is a prime, is widely believed to hold as well, assuming certain modest restrictions regarding the cardinality of $\mathcal{A}$ with respect to p [77, 74, 19, 140]. There are examples showing that the following conjecture, if true, would be tight.

Conjecture 6.8. *Let a set $\mathcal{A} \subset \mathbb{Z}_p$ be given. If either*

(i) $0 \leq |2\mathcal{A}| - \big(2|\mathcal{A}| - 1\big) \leq \min(|\mathcal{A}| - 4, p - |2\mathcal{A}| - 2)$ *or*
(ii) $0 \leq |2\mathcal{A}| - \big(2|\mathcal{A}| - 1\big) = |\mathcal{A}| - 3 \leq p - |2\mathcal{A}| - 3$

then $\mathcal{A}$ can be covered by an arithmetic progression of length at most $|2\mathcal{A}| - |\mathcal{A}| + 1$.

However, such a statement has turned out to be significantly more difficult to prove. It was Freĭman himself who first made progress towards this conjecture, by showing that the covering property holds for any set $\mathcal{A} \subset \mathbb{Z}_p$ satisfying $|2\mathcal{A}| \leq 2.4|\mathcal{A}| - 3$ and $|\mathcal{A}| < p/35$, see [59]. Rødseth [116] later showed that the density requirement can be weakened to $p/10.7$.

A more general result of Green and Ruzsa [71] immediately gives the same conclusion for all sets satisfying $|2\mathcal{A}| \leq 3|\mathcal{A}| - 4$ as long as they also satisfy the rather strong density requirement $|\mathcal{A}| < p/10^{215}$. Serra and Zémor [140] obtained a result with the same covering conclusion and no restrictions regarding the size of $|\mathcal{A}|$ itself, but assuming that $|2\mathcal{A}| \leq (2+\varepsilon)|\mathcal{A}|$ with $\varepsilon < 10^{-4}$. The latter bound was relaxed to $\varepsilon < 0.1368$, under the mild additional assumption $|2\mathcal{A}| \leq 3p/4$ by Candela, González-Sanchez and Grynkiewicz [20].

Here we will present a result in the spirit of those of Freĭman, Rødseth as well as Lev and Shkredov: similar to all of these statements, it will make uses a Fourier-analytic rectification argument that allows one to transplant a significant part of the set into the integers, where Theorem 6.4 is applied. The resulting structure of that significant part of $\mathcal{A}$ in fact allows one to argue that the whole set has to behave as if it were in the integers. Where this result differs however, is that we will allow the cardinality of the sumset of that significant part to go past the $3|A| - 4$ barrier in the integers, where we make us of Proposition 6.7. This also implies that, unlike in the original approach, we have to take the dimension of our sets into consideration.

Theorem 6.9. *If a set $\mathcal{A} \subseteq \mathbb{Z}_p$ satisfies $|2\mathcal{A}| \leq 2.48|\mathcal{A}| - 7$ and $|\mathcal{A}| < p/10^{10}$, then there exists an arithmetic progression $\mathcal{P} \subseteq \mathbb{Z}_p$ of size at most $|2\mathcal{A}| - |\mathcal{A}| + 1$ containing $\mathcal{A}$.*

Very recently, Lev and Shkredov improved the requirements to $|2\mathcal{A}| < 2.59|\mathcal{A}| - 3$ and $|\mathcal{A}| < 0.0045p$ using the properties of higher energies. However, it should be noted that their approach can very be combined with the one we will take to establish Theorem 6.9 to get a statement with the requirements $|2\mathcal{A}| < 2.62|\mathcal{A}| - 3$ and $|\mathcal{A}| < p/10^{10}$. Furthermore, Theorem 6.9 is very much conditional on progress towards Conjecture 6.5.

Any improvement on Proposition 6.7 that gets closer towards Conjecture 6.5, either in the size of the covering or the doubling, would yield a proportional improvement towards Conjecture 6.8 using the proof methodology behind Theorem 6.9.

6.1 Proof of Theorem 6.6 – Sets on few intervals

Konyagin and Lev [97] established a formula for the dimension of a given set $A \subset \mathbb{Z}^m$ of cardinality k. For $1 \leq i \leq k$, let $\mathbf{e}_i$ denote the vector in $\mathbb{R}^k$ that has a one at coordinate i and zero everywhere else. We denote by M_A the integer–valued matrix with k columns obtained by listing as its rows all vectors $\mathbf{e}_{i_1} + \mathbf{e}_{i_2} - \mathbf{e}_{i_3} - \mathbf{e}_{i_4}$ for which $a_{i_1} + a_{i_2} = a_{i_3} + a_{i_4}$ holds and for which we do not have $i_1 = i_2 = i_3 = i_4$. The result of Konyagin and Lev that derives the dimension of A from the rank of M_A can be stated as follows.

Theorem 6.10 (Theorem 4 in [97]). *For any set $A \subseteq \mathbb{Z}^m$ we have*

$$\dim(A) = |A| - 1 - \mathrm{rk}(M_A). \tag{6.9}$$

Now let $A \subset \mathbb{Z}$ be a set which is the union of s disjoint segments $P_1, \dots, P_s$. Given such a set A, we denote by S_A the integer–valued matrix with s columns obtained by listing in its rows all vectors $\mathbf{e}_{j_1} + \mathbf{e}_{j_2} - \mathbf{e}_{j_3} - \mathbf{e}_{j_4}$ for which $(P_{j_1} + P_{j_2}) \cap (P_{j_3} + P_{j_4}) \neq \emptyset$ and for which we do not have $j_1 = j_2 = j_3 = j_4$. We derive the following result from Theorem 6.10.

Proposition 6.11. *Any $A \subset \mathbb{Z}$ that is the union of $1 \leq s \leq |A| - 1$ disjoint segments satisfies*

$$\dim(A) = s - \mathrm{rk}(S_A). \tag{6.10}$$

Proof. Every row in M_A is associated with up to four (not all equal) elements a_{i_1}, a_{i_2}, a_{i_3} and a_{i_4} such that $a_{i_1} + a_{i_2} = a_{i_3} + a_{i_4}$. Let $a_{i_1} \in P_{j_1}$, $a_{i_2} \in P_{j_2}$, $a_{i_3} \in P_{j_3}$ and $a_{i_4} \in P_{j_4}$ where we may assume $j_1 \leq j_2$ and $j_3 \leq j_4$ as well as $\min\{j_1, j_2\} \leq \min\{j_3, j_4\}$. Furthermore, let $0 \leq y = \#\{j : |P_j| = 1\} < s$ denote the number of segments that are singletons. We distinguish the following cases.

Case 1. $\#\{j_1, j_2, j_3, j_4\} = 1$, that is, $j_1 = j_2 = j_3 = j_4 = j$ for some $1 \leq j \leq s$. Since P_j is one dimensional, by Theorem 6.10 there are a total of $\max\{|P_j| - 2, 0\}$ linearly independent equations of this type for each P_j. As these equations only involve elements in P_j and the segments are disjoint, it is clear that each equation is linearly independent from those of other segments, so we get a total of $|A| - 2s + y$ linearly independent equations of this type in M_A. On the other hand this case does not contribute to the rank of S_A

Case 2. $\#\{j_1, j_2, j_3, j_4\} = 2$. We distinguish two further cases.

Case 2.1. $j_1 = j_3 < j_4 = j_2$. Segments of length one can only give trivial equations of this type not contributing to the rank of M_A. If $P_{i_0} < P_{i_1} < \cdots < P_{i_{s-y}}$ are the $s - y$ segments which are not singletons, then equations of this type give us a total of $s - y - 1$ new linear independent ones on top of the ones given by Case 1, one for each pair P_{i_0}, P_{i_j} and $j = 2, \dots s - y$, the remaining ones being linearly dependent with these. Moreover, this case does not contribute to the rank of S_A.

Case 2.2. $j_1 < j_2 = j_3 = j_4$ or $j_1 = j_2 = j_3 < j_4$. Each pair $P_j, P_{j'}$ with $j \neq j'$ for which an equation of this type exists implies that $P_j \cap P_{j'}$ intersects either $2P_j$ or $2P_{j'}$, contributes one additional linear independent equation in M_A on top of the above ones and contributes to one additional linear equation in S_A as well.

Case 3. $\#\{j_1, j_2, j_3, j_4\} \geq 3$, that is, $j_1 < j_3 \leq j_4 < j_2$. This implies that $P_{j_1} + P_{j_2}$ intersects $P_{j_3} + P_{j_4}$. Each such intersection contributes with one additional linear independent equation in M_A and also on S_A on top of the above ones.

Taken together, it follows that $\mathrm{rk}(M_A) = (|A| - 2s + y) + (s - y - 1) + \mathrm{rk}(S_A)$ and therefore, by Theorem 6.10, $\dim(A) = s - \mathrm{rk}(S_A)$. □

We note that for $s \geq 6$ there are sets for which $s = |A|$, that is all segments consist of a single element, which are not covered by Proposition 6.11. We will also need a generalization of Theorem 6.4 due to Lev and Smeliansky [104], that handles the sumset of two distinct sets, as well as another result by Freĭman [61].

Theorem 6.12 (Theorem 2 in [104]). *Any two finite sets $A, B \subset \mathbb{Z}$ in normal form for which* $\max(A) > \max(B)$ *satisfy*

$$|A + B| \geq \min\{|A| + 2|B| - 2, \max(A) + |B|\}. \tag{6.11}$$

Theorem 6.13 (Lemma 1.15 in [61]). *Let A be a 2-dimensional set of cardinality $|A| > 6$ with $|2A| = 3|A| - 3 + b$. If A can not be covered by a set consisting of two lines with volume at most $|A| + b$ then $b \geq |A|/3 - 2$.*

We will also use the following Lemma which handles the case of two segments.

Lemma 6.14. *Let A be an extremal set that is the union of two segments. If A is 1-dimensional and $|2A| = 2|A| - 1 + b$ for some $1 \leq b \leq |A| - 3$, then* $\mathrm{vol}(A) = |A| + b$ *and A is in fact F_2-isomorphic to*

$$[0, |A| + b - 1] \setminus [1, b]. \tag{6.12}$$

If A is 2-dimensional and $|2A| = 3|A| - 3$, then $\mathrm{vol}(A) = k$ *and it is in fact F_2-isomorphic to some*

$$\big([0, k_1 - 1] \times \{0\}\big) \cup \big([0, k_2 - 1] \times \{1\}\big) \tag{6.13}$$

where $k_1 + k_2 = |A|$ and $k_1, k_2 \geq 1$.

Proof. If $\dim(A) = 2$, then there must be no relation in the matrix S_A in Proposition 6.11 and we are done. Suppose therefore that $\dim(A) = 1$ and $A = P_1 \cup P_2$, say $P_1 = [0, k_1 - 1]$ and $P_2 = [k_1 + \ell_1, k_1 + \ell_1 + k_2 - 1]$ for some $k_1, k_2 \geq 1$, where $k = k_1 + k_2$ and $\ell_1 \geq 1$. Since $\dim(A) = 1$, we may also assume that $(P_1 + P_2) \cap 2P_2 \neq \emptyset$, so that $2A$ consists of the interval $[0, 2(k + \ell_1 - 1)]$ with a hole of some length $h \geq 0$. Since A is extremal and has volume $vol(A) = k + \ell_1$ we have $|2A| = 2k - 1 + \ell_1$, so that $h = \ell_1$. Therefore

$$\ell_1 = \min(P_1 + P_2) - \max(2P_1) - 1 = \ell_1 + 1 - k_1,$$

which implies $k_1 = 1$. □

We are now ready to prove Theorem 6.6.

Proof of Theorem 6.6. Let A consist of three segments P_1, P_2, and P_3 where $P_i = [0, k_1 - 1]$, $P_2 = [k_1 + \ell_1, k_1 + k_2 + \ell_1 - 1]$ and $P_3 = [k_1 + k_2 + \ell_1 + \ell_2, k_1 + k_2 + k_3 + \ell_1 + \ell_2 - 1]$, so that $k_i = |P_i|$ for $i \in \{1, 2, 3\}$ and the intervals are separated by intervals of holes $L_1 = [k_1, k_1 + \ell_1 - 1]$ and $L_2 = [k_1 + k_2 + \ell_1, k_1 + k_2 + \ell_1 + \ell_2 - 1]$ of size $\ell_i = |L_i|$ for $i \in \{1, 2\}$. Lastly, we also write $a = \max(A) = k + \ell - 1$, where $k = k_1 + k_2 + k_3 = |A|$ and $\ell = \ell_1 + \ell_2$. We now consider three cases according to the dimension of A.

Case 1. $\dim(A) = 3$. By Proposition 6.11 we must have $\mathrm{rank}(S_A) = 0$, that is, all $P_i + P_j$ are disjoint for $1 \leq i, j \leq 3$, and we are led to (iv).

Case 2. $\dim(A) = 2$. It follows from Proposition 6.11 that the matrix S_A has $\mathrm{rank}(S_A) = 1$. Up to isomorphisms, we have two possibilities for the only independent relation in S_A.

Case 2.1: If $(P_1 + P_3) \cap 2P_2 \neq \emptyset$, then we may assume that

$$\begin{aligned} P_1 &= \{(0,0), \ldots, (0, k_1 - 1)\}, \\ P_2 &= \{(1,0), \ldots, (1, k_2 - 1)\} \text{ and} \\ P_3 &= \{(2,\ell), \ldots, (2, \ell + k_3 - 1)\} \end{aligned}$$

for some $\ell \in \mathbb{N}_0$. We know that $|2A| = 4k - 6 - |(P_1 + P_3) \cap 2P_2|$ so that $b = k - 3 - |(P_1 + P_3) \cap 2P_2|$. We also have $\mathrm{vol}(A) = k + \max(\lfloor (k_1 + \ell + k_3)/2 \rfloor - k_2, 0)$. Since $\dim(A) = 2$ we must have $|(P_1 + P_3) \cap 2P_2| > 0$ and therefore $\ell \leq 2k_2 - 2$. Now if $0 \leq \ell \leq 2k_2 - k_1 - k_3$ then $b = k - 3 - (k_1 + k_3 - 1) = k_2 - 2$ and $\mathrm{vol}(A) = k$. If A is extremal, we therefore have $k_2 = 2$ so that $k_1 + k_3 \leq 4$ and hence $k \leq 6$. If $\max(2k_2 - k_1 - k_3, 0) < \ell \leq 2k_2 - 2$ then $b = k - 3 - (2k_2 - 2 - \ell + 1) = k - 2 + 2k_2 - \ell > \max(\lfloor (k_1 + \ell + k_3)/2 \rfloor - k_2, 0)$ so the set cannot be extremal.

Case 2.2: If $2P_1 \cap (P_1 + P_2) \neq \emptyset$, or likewise $(P_1 + P_2) \cap 2P_2 \neq \emptyset$, we may assume that

$$\begin{aligned} P_1 &= \{(0,0), (0,1), \cdots, (0, k_1 - 1)\}, \\ P_2 &= \{(0, k_1 + \ell_1), (0, k_1 + \ell_1 + 1), \ldots, (0, k_1 + \ell_1 + k_2 - 1)\} \text{ and} \\ P_3 &= \{(1,0), \ldots, (1, k_3 - 1)\} \end{aligned}$$

where $k_1 \geq \ell_1 + 2$. Let $A_0 = P_0 \cup P_1$ and $A_1 = P_2$, so that

$$2A = 2A_0 \cup (A_0 + A_1) \cup 2A_1,$$

the union being disjoint. We have $|2A_1| = 2k_3 - 1$ and, by Theorem 6.4, we also have $|2A_0| \geq 2(k_1 + k_2) - 1 + \ell_1$. Moreover, it can be readily checked that

$$|A_0 + A_1| = \begin{cases} k_1 + \ell_1 + k_2 + (k_3 - 1) = k + \ell_1 - 1, & \text{if } k_3 > \ell_1 + 1, \\ (k_1 + k_3 - 1) + (k_2 + k_3 - 1) = k + k_3 - 2, & \text{otherwise.} \end{cases}$$

It follows that

$$|2A| \geq 3k - 3 + \ell_1 + \min\{k_3 - 1, \ell_1\}. \tag{6.14}$$

As $\mathrm{vol}(A) = k + \ell_1$, the set can only be extremal if $\min(k_3 - 1, \ell_1) = 0$, which implies $k_3 = 1$ (as $\ell_1 \geq 1$) and there is equality in Equation (6.14), namely, if $|2A_0| = 2(k_1 + k_2) - 1 + \ell_1$. Applying Lemma 6.14 to A_0 leads to (iii).

Case 3. $\dim(A) = 1$. The following are the six segments in $2A$:

$$\begin{aligned}
2P_1 &= [0, 2k_1 - 2],\\
P_1 + P_2 &= (k_1 + \ell_1) + [0, k_1 + k_2 - 2],\\
2P_2 &= 2(k_1 + \ell_1) + [0, 2k_2 - 2],\\
P_1 + P_3 &= k_1 + \ell_1 + k_2 + \ell_2 + [0, k_1 + k_3 - 2],\\
P_2 + P_3 &= 2(k_1 + \ell_1) + (k_2 + \ell_2) + [0, k_2 + k_3 - 2],\\
2P_3 &= 2(k_1 + \ell_1 + k_2 + \ell_2) + [0, 2k_3 - 2].
\end{aligned}$$

Since A is extremal, we have

$$a \geq 2(k + b - 2), \tag{6.15}$$

so that

$$\ell \geq k + 2b - 3. \tag{6.16}$$

We will use the following facts.

Claim 6.15. *If* $\max(k_1, k_2) < \ell_1 + 2$ *then* $2P_1$, $P_1 + P_2$ *and* $2P_2$ *are pairwise disjoint. If* $\max(k_1, k_2) \geq \ell_1 + 2$ *then* $P_1 \cup P_2$ *is* 1*-dimensional and* $2(P_1 \cup P_2)$ *is a segment with a hole of length*

$$h = \max\Big\{\ell_1 - \min(k_1, k_2) + 1, 0\Big\}.$$

Proof. We note that $2P_1$ does not intersect $P_1 + P_2$ if and only if $\max(2P_1) < \min(P_1 + P_2)$, which is equivalent to $k_1 < \ell_1 + 2$. Likewise, $P_1 + P_2$ does not intersect $2P_2$ if and only if $k_2 < \ell_1 + 2$, establishing the first part of the claim.

Assume without loss of generality that $k_1 \leq k_2$ and $k_2 \geq \ell_1 + 2$. Then $(P_1 + P_2) \cup 2P_2$ is a segment since the two parts intersect. In particular, $P_1 \cup P_2$ is 1-dimensional. Moreover, either $2P_1 \cup (P_1 + P_2) \cup 2P_2$ is a segment or a segment with a hole of length $h = \min(P_1 + P_2) - \max(2P_1) - 1 = \ell_1 - k_1 + 1$, establishing the second part of the claim. □

By the above Claim, if both $\max(k_1, k_2) < \ell_1 + 2$ and $\max(k_2, k_3) < \ell_2 + 2$, then the five segments in

$$2P_1 \cup (P_1 + P_2) \cup 2P_2 \cup (P_2 + P_3) \cup 2P_3$$

are pairwise disjoint. Using Proposition 6.11 it follows that $\text{rank}(S_A) \leq 1$ and hence $\dim(A) \geq 2$, contradicting the assumption of this case.

We will therefore without loss of generality assume that $\max(k_1, k_2) \geq \ell_1 + 2$. In this case we have the following.

Claim 6.16. $\max\{k_2, k_3\} < \ell_2 + 2$.

Proof. Suppose on the contrary that $\max\{k_2, k_3\} \geq \ell_2 + 2$. Then, using Equation (6.16),

$$k + 2b - 3 \leq \ell \leq \max\{k_1, k_2\} + \max\{k_2, k_3\} - 4$$

which implies $\max\{k_1, k_2\} = \max\{k_2, k_3\} = k_2$. It follows that $\max(2P_2) = 2(k_1 + l_1 + k_2 - 1) \geq 2(k_1 + \ell_1) + k_2 + \ell_2 = \min(P_2 + P_3)$. Hence, the sets $2(P_1 \cup P_2)$ and $2(P_2 \cup P_3)$ overlap and, by Claim 6.15, $2A$ consists of the interval $[0, 2a]$ with two holes of total length at most

$$\max\{\ell_1 - k_1 + 1, 0\} + \max\{\ell_2 - k_3 + 1, 0\} \leq \ell.$$

Therefore, by using $a = k + \ell - 1$ and Equation (6.16), we obtain $|2A| \geq 2a - \ell + 1 \geq 3k + 2b - 4$ and therefore A is not extremal. □

It follows from Claim 6.15 and Claim 6.16 that the three segments $2P_2, P_2 + P_3, 2P_3$ are pairwise disjoint. Since A is one–dimensional, $2(P_1 \cup P_2)$ must intersect $P_1 + P_3$. In particular, $\max(2P_2) \geq \min(P_1 + P_3)$ which yields

$$k_1 + \ell_1 + k_2 \geq \ell_2 + 2. \tag{6.17}$$

Claim 6.17. $k_3 = 1$.

Proof. Suppose on the contrary that $k_3 > 1$. We then have $\ell_1 > 1$, since otherwise Equation (6.17) and Equation (6.16) give $k_1 + k_2 \geq \ell_2 + 1 \geq k + 2b - 3$ and we get $k_3 \leq 1$.

Let $B = 2(P_1 \cup P_2) \cup (P_1 + P_3) \cup (P_2 + P_3)$. We can write $2A$ as the disjoint union

$$2A = B \cup 2P_3.$$

Consider now the set A' obtained from A by replacing $\min(P_3)$ with $\max(P_1) + 1$ if $k_1 \geq k_2$ and with $\min(P_2) - 1$ otherwise. The resulting set is still composed of three disjoint segments, $A' = P_1' \cup P_2' \cup P_3'$ with $\ell_1' = \ell_1 - 1$, $\ell_2' = \ell_2 + 1$ and $\min\{k_1', k_2'\} = \min\{k_1, k_2\}$. We can write $2A'$ as the disjoint union

$$2A' = B' \cup 2P_3',$$

where $B' = 2(P_1' \cup P_2') \cup (P_1' + P_3') \cup (P_2' + P_3')$. We have $|2P_3'| = |2P_3| - 2$. Let us show that $|B'| \leq |B| + 1$.

By Claim 6.15, $|2(P_1' \cup P_2')| \leq |2(P_1 \cup P_2)| + 1$. If $k_1 \geq k_2$ then $P_2' = P_2$ and $|P_2' + P_3'| = |P_2 + P_3| - 1$, while $P_1' + P_3' = (P_1 + P_3) + 1$. If $P_1 + P_3$ and $P_2 + P_3$ are disjoint, then the two last modifications compensate each other, while if they intersect then there is no change in the cardinality of their union. Similarly, if $k_1 < k_2$ then $P_1' = P_1$ and we loose one unit in $P_1' + P_3'$ while $P_2' + P_3'$ is translated one unit to the right from $P_2 + P_3$, and again there is no change in the cardinality of the union of these two segments.

In either case, we get $|2A'| < |2A|$ so that, if A' is one–dimensional it would have the same volume as A, contradicting that A is extremal. It follows that A' must be 2-dimensional. This implies $\max(2P_2') < \min(P_1' + P_3')$. Since $\max(2P_2) \geq \min(P_1 + P_3)$, we have equality in the last inequality. Therefore,

$$|2A| = |2(P_1 \cup P_2)| + |P_3 + A| - 1. \tag{6.18}$$

By Theorem 6.12 we have

$$|P_3 + A| \geq |A| + 2|P_3| - 2 = k + 2k_3 - 2 \tag{6.19}$$

and therefore

$$\begin{aligned}|2A| &= |2(P_1 \cup P_2)| + |P_3 + A| \\ &\geq (\max(2P_2) + 1 - \ell_1) + (k + 2k_3 - 2) \\ &= 2(k - k_3 + \ell_1 - 1) - \ell_1 + k + 2k_3 - 2 \\ &= 3k + \ell_1 - 4,\end{aligned}$$

so that $\ell_1 \leq b$. But then, by Equation (6.17), we have

$$\ell = \ell_1 + \ell_2 \leq b + (k - k_3 + b - 2) = k - k_3 + 2b - 2, \tag{6.20}$$

contradicting Equation (6.16). It follows that A could not have been extremal. □

We can therefore assume $P_3 = \{a\}$. It follows that

$$2A = 2(P_1 \cup P_2) \cup (a + A).$$

Moreover,

$$\min(P_2 + P_3) - \max(P_1 + P_3) = \ell_1 - k_3 + 2 = \ell_1 + 1 > 1.$$

We next consider two cases.

Case 3.1: If $k_1 \leq k_2$, then the sumset $2A$ can be written as the disjoint union

$$2A = B \cup (P_2 + P_3) \cup 2P_3,$$

where $B = 2(P_1 \cup P_2) \cup (P_1 + P_3)$ is an interval with a hole of length $h = \max\{\ell_1 - k_1 + 1, 0\}$. Such a 1-dimensional set with $k_1 > 1$ cannot be extremal since, by exchanging $\max(P_1)$ by $\min(P_2) - 1$ we get a 1-dimensional set with the same volume and smaller doubling. It follows that $k_1 = 1$. By using Equation (6.17), we get $\max(2P_2) - \max(P_1 + P_3) = \ell_1 + k_2 - \ell_2 - 1 \geq 0$. In this case $2(k + \ell_1 - 2) = \max(2P_2) \geq a = \max(P_1 + P_3)$ and, again by extremality, equality holds. We thus have $|2A| = (a - \ell_1 + 1) + (k - 2) + 1 = 3k + \ell_1 - 4$, leading to (i).

Case 3.2: If $k_1 > k_2$, then from

$$\begin{aligned} 3k - 4 + b = |2A| &= |2(P_1 \cup P_2)| + |a + A| - |2(P_1 \cup P_2) \cap (a + A)| \\ &= 2(k - 1 + \ell_1) - 1 - \max\{\ell_1 - k_2 + 1, 0\} + k \\ &\quad - |2(P_1 \cup P_2) \cap (a + A)|, \end{aligned}$$

we obtain

$$|2(P_1 \cup P_2) \cap (a + A)| = 2\ell_1 + 1 - \max\{\ell_1 - k_2 + 1, 0\} - b. \tag{6.21}$$

For this equality to hold, a necessary condition is

$$\max(2P_2) - a + 1 \geq 2\ell_1 + 1 - \max\{\ell_1 - k_2 + 1, 0\} - b. \tag{6.22}$$

By using $\max(2P_2) = 2(k - 1) + 2\ell_1 - 2$ and $a = k + \ell - 1$ in Equation (6.22), we obtain

$$\ell \leq k + b + \max\{\ell_1 - k_2 + 1, 0\} - 3. \tag{6.23}$$

By Equation (6.16) we have $\ell_1 \geq k_2 + b - 1$. On the other hand, since $|2(P_1 \cup P_2) \cap (a + A)| \leq |2P_2| = 2k_2 - 1$, it follows from Equation (6.21) that $\ell_1 \leq b + k_2 - 1$. Hence

$\ell_1 = b + k_2 - 1$, there is equality in Equation (6.23) and $2P_2$ must be included in $P_1 + P_3$, so that $2k_2 - 1 = |2P_2| \leq |P_1 + P_3| = k_1$. This gives (ii). ■

6.2 Proof of Proposition 6.7 – A partial $4k - 8$ Theorem

In what follows we will usually use the notation $\mathcal{A}$ to refer to sets in some cyclic group $\mathbb{Z}_m$ and the notation A to refer to sets in the integers. Often $\mathcal{A}$ will refer to the canonical projection from $\mathbb{Z}$ to some $\mathbb{Z}_m$ of some $A \subset \mathbb{Z}$. We will also say that a subset A of an arbitrary abelian group is said to be rectifiable if it is F_2-isomorphic to a set of integers. As an example, we previously saw in the introduction of this section, that $A = \{0, 1, \ldots, k-1\} \subseteq \mathbb{Z}_n$ is rectifiable if and only if $n \geq 2k - 1$.

Deshouillers and Freĭman stated the following result regarding covering properties of subsets of $\mathbb{Z}_m$ with very small sumset. Note that – unlike Conjecture 6.8 – this result concerns arbitrary $\mathbb{Z}_m$, that is the integer m does not have to be prime. This explains the weaker bounds and more complex statement.

Theorem 6.18 (Theorem 1 in [42]). *For any set $\mathcal{A} \subset \mathbb{Z}_m$ satisfying $|2\mathcal{A}| \leq 2.04|\mathcal{A}|$ and $|\mathcal{A}| \leq n/10^9$ there exists a proper subgroup $H < \mathbb{Z}_m$ such that the following holds:*

1. *If $\mathcal{A}$ is included in one coset of H then $|\mathcal{A}| > |H|/10^9$.*
2. *If $\mathcal{A}$ meets exactly 2 or at least 4 cosets of H then it is included in an ℓ-term arithmetic progression of cosets of H where*
$$(\ell - 1)|H| \leq |2\mathcal{A}| - |\mathcal{A}|. \tag{6.24}$$
3. *If $\mathcal{A}$ meets exactly three cosets of H then it is included in an ℓ-term arithmetic progression of cosets of H where*
$$(\min(\ell, 4) - 1)|H| \leq |2\mathcal{A}| - |\mathcal{A}|. \tag{6.25}$$

Furthermore, if $\ell \geq 2$ then there exists a coset of H containing at least $2/3\,|H|$ elements from $\mathcal{A}$.

Note that if m is prime, then the subgroup H in the statement has to be the trivial group $\{0\}$ and $\mathcal{A}$ clearly meets exactly $|\mathcal{A}|$ cosets of it, so we are in case 2 of the statement as long as $|\mathcal{A}| \geq 4$. The conclusion in this case is the same as that of Theorem 6.9. We will also need the following straightforward observation in order to distinguish between integer sets of different additive dimension.

Lemma 6.19. *Let $A \subset \mathbb{Z}$ be given in normal form with $|A| \geq 3$ and $m > 1$ such that $m \mid \max(A)$. If the canonical projection of A into $\mathbb{Z}_m$ is rectifiable, then* $\dim(A) \geq 2$.

Proof. Let $\varphi : \mathbb{Z} \to \mathbb{Z}_m$ denote the canonical projection. Note that

$$\{(a, \varphi(a)) : a \in A\} \subset \mathbb{Z} \times \mathbb{Z}_m$$

is F_2-isomorphic to A, since for any $a_1, a_2, a_3, a_4 \in A$ we have $a_1 + a_2 = a_3 + a_4$ if and only if $(a_1, \varphi(a_1)) + (a_2, \varphi(a_2)) = (a_3, \varphi(a_3)) + (a_4, \varphi(a_4))$. As $\mathcal{A} = \varphi(A)$ is rectifiable, there exists some F_2-isomorphism f mapping $\mathcal{A}$ into the integers. By the same argument as before, it follows that $\{(a, \varphi(a)) : a \in A\}$ and hence also A is F_2-isomorphic to $\{(a, f(\varphi(a))) : a \in A\} \subset \mathbb{Z}^2$. We may without loss of generality assume that $f(0) = 0$ and note that since A is in normal form and $|A| \geq 3$, there must exist some $a' \in A$ such that $\varphi(a') \neq 0$ and hence also $f(\varphi(a')) \neq 0$. Using the requirement that $m \mid \max(A)$, we observe that the three points $(0, f(\varphi(0))) = (0, 0)$, $(\max(A), f(\varphi(\max(A))) = (\max(A), 0)$ and $(a', f(\varphi(a'))) \neq (a', 0)$ do not lie in a hyperplane of $\mathbb{Z}^2$ and therefore $\dim(A) \geq 2$ as desired. □

We can now state and prove Proposition 6.7. It should be noted that the proof has some slight similarities with the proof of Freĭman's $3|A| - 4$ Theorem in the integers by modular reduction, see [104]. However, there is a new component in the argument here, consisting of taking into account the Freĭman dimension of the set.

Proof of Proposition 6.7. Let $A \subset \mathbb{Z}$ satisfy $|2A| \leq 3.04|A| - 3$ as well as $\max(A) \geq 10^9|A|$, and assume without loss of generality that A is in normal form. We will show that we must have $\dim(A) \geq 2$, which contradicts the assumption that A is 1-dimensional. Let $\varphi : \mathbb{Z} \to \mathbb{Z}_{\max(A)}$ denote the canonical projection and observe that $\mathcal{A} = \varphi(A)$ satisfies $|\mathcal{A}| = |A| - 1 < \max(A)/10^9$. Let B denote the set of elements $x \in 2A$ such that $x + \max(A)$ is also in $2A$. Since 0 and $\max(A)$ are both in A we have $B \supset A$, whence $|2A| = |2\mathcal{A}| + |B| \geq |2\mathcal{A}| + |A|$, and so

$$|2\mathcal{A}| \leq |2A| - |A| \leq 2.04|A| - 3 \leq 2.04|\mathcal{A}|.$$

We can therefore apply Theorem 6.18, obtaining that $\mathcal{A}$ is covered by some small arithmetic progression of cosets of some proper subgroup $H < \mathbb{Z}_{\max(A)}$. Let us go through the cases given by this theorem. In the following $\mathbb{1}_C$ will denote the indicator function of some given set C.

1. As A is in normal form, $\mathcal{A}$ cannot be contained in a single coset of H.
2. If $\mathcal{A}$ meets exactly 2 or at least 4 cosets of H then it is included in an ℓ-term arithmetic progression of cosets of H, where by Equation (6.24) we have $\ell \leq 1.04|\mathcal{A}|/|H| + 1 \leq (1.04 + 3/2)|\mathcal{A}|/|H|$, the last equality following from the last sentence in Theorem 6.18. Using that $|\mathcal{A}| < 10^{-9}\max(A)$ we deduce that
$$\ell \leq 3|\mathcal{A}|/|H| < \tfrac{1}{2}\left|\mathbb{Z}_{\max(A)/|H|}\right|. \tag{6.26}$$
Letting $m = \max(A)/|H|$ we now observe that, since A is in normal form, its canonical projection into $\mathbb{Z}_m$ cannot be contained in a proper subgroup of $\mathbb{Z}_m$. It follows that the common difference of the ℓ-term arithmetic progression covering this projection of A does not divide m, whence we can dilate by the inverse mod m of this common difference, and it follows that the projection of A is F_2-isomorphic to some subset of an interval of size $m/2$ in $\mathbb{Z}_m$. This projection is therefore rectifiable, so by Lemma 6.19 we have $\dim(A) \geq 2$.
3. If $\mathcal{A}$ meets exactly 3 cosets of H, then we argue in a way similar to case 2, considering the projection of A to $\mathbb{Z}_m$ where $m = \max(A)/|H|$. Here, however, we distinguish two cases, according to whether the 3 cosets are in arithmetic progression or not.
Assume that these cosets are in arithmetic progression with difference d. If we can rectify the 3-term progression formed by the cosets' representatives, then we can rectify the projection of A into $\mathbb{Z}_m$. By applying Lemma 6.19 as in case 2, we again obtain the contradiction $\dim(A) \geq 2$. If we cannot rectify the 3-term progression, then we must have either $m < 6$, $d = m/3$ or $d = m/4$. We certainly have $m \geq 6$ since by Equation (6.25) we have $|H| \leq (|2\mathcal{A}| - |\mathcal{A}|)/2 \leq 0.52|\mathcal{A}|$, and as noted above we also have $\max(A) \geq 10^9|\mathcal{A}|$, so $m \geq 10^9$. Furthermore, if $d = m/3$ or $d = m/4$ then m is multiple of d and clearly A cannot have been in normal form.
If the cosets do not form an arithmetic progression, then we have $\mathcal{A} \subseteq H \cup (H + c_1) \cup (H + c_2)$ for some $c_1, c_2 \in \mathbb{Z}_{\max(A)}$ satisfying $c_2 \not\equiv 2c_1$, $c_1 \not\equiv 2c_2$ and $c_1 + c_2 \not\equiv 0$ in $\mathbb{Z}_{\max(A)}/H$. Moreover, we may assume that either $2c_1 \not\equiv 0$ or $2c_2 \not\equiv 0$ in $\mathbb{Z}_{\max(A)}/H$ as otherwise $\mathcal{A}$ would only meet 2 cosets of H. We therefore assume without loss of generality that $2c_2 \not\equiv 0$ in $\mathbb{Z}_{\max(A)}/H$. If furthermore $2c_2 \not\equiv 2c_1$ in $\mathbb{Z}_{\max(A)}/H$, then $\{\mathbb{1}_{H+c_2}(\varphi(a)) : a \in A\}$ is F_2-homomorphic to A and therefore

$\dim(A) \geq 2$ as A is F_2-isomorphic to

$$\Big\{(\mathbb{1}_{H+c_2}(\varphi(a)), a) : a \in A\Big\} \subset \mathbb{Z}^2 \tag{6.27}$$

which is not contained in some hyperplane of $\mathbb{Z}^2$ as $\varphi(0) = \varphi(\max(A)) \in H$ but $\varphi(a') \in H + c_2$ for some $a' \in A$. If however $2c_1 \equiv 2c_2 \not\equiv 0$ in $\mathbb{Z}_{\max(A)}/H$, then likewise we can argue that $\dim(A) \geq 2$ as A now is F_2-isomorphic to

$$\Big\{(\mathbb{1}_{H}(\varphi(a)), a) : a \in A\Big\} \subset \mathbb{Z}^2 \tag{6.28}$$

which for the same reason is also not contained in any hyperplane of $\mathbb{Z}^2$.

It follows that $\dim(A) \geq 2$, contradicting the assumption that $\dim(A) = 1$. ■

6.3 Proof of Theorem 6.9 – The cyclic setting

Let us state the following result of Freĭman [61] regarding 2-dimensional sets of very small doubling.

Theorem 6.20 (Theorem 1.17 in [61])**.** *Let $A \subset \mathbb{Z}^2$ be a 2-dimensional set that cannot be embedded in any straight line and that satisfies $|2A| < {}^{10}\!/_3\,|A| - 5$ and $|A| \geq 11$. Then A is contained in a set which is isomorphic to*

$$\{(0,0),(0,1),(0,2),\ldots,(0,k_1-1),(1,0),(2,0),\ldots,(1,k_2-1)\} \tag{6.29}$$

where $k_1, k_2 \geq 1$ and $k_1 + k_2 \leq |2A| - 2|A| + 3$.

We shall use the following consequence.

Corollary 6.21. *Any 2-dimensional set $A \subset \mathbb{Z}$ satisfying $|2A| \leq {}^{10}\!/_3\,|A| - 5$ is contained in the union of two arithmetic progressions P_1 and P_2 with the same common difference such that $|P_1 \cup P_2| \leq |2A| - 2|A| + 3$. Furthermore, the sumsets $2P_1$, $P_1 + P_2$ and $2P_2$ are disjoint.*

A proof of this can be immediately derived from the following statement.

Lemma 6.22. *Given a finite d-dimensional set $A \subset \mathbb{Z}^d$ not contained in a hyperplane, we can extend any Freĭman-isomorphism φ mapping A to some $A' \subset \mathbb{Z}$ to an affine linear map.*

Proof. Assume to the contrary that φ is not affine linear. As $\dim(A) = d$, there exist d elements $a_1, \ldots, a_d \in A$ spanning $\mathbb{Z}^d$. Let φ_e denote the affine linear map $\mathbb{Z}^d \to \mathbb{Z}$ determined by $a_1, \ldots, a_d \in \mathbb{Z}^d$ as well as 0, that is $\varphi_e(a_i) = \varphi(a_i)$ for $i = 1, \ldots, d$ and $\varphi_e(0) = \varphi(0)$. As φ is not affine linear, we must have $\varphi_e(x) \neq \varphi(x)$ for some $x \in A \setminus \{a_1, \ldots, a_d, 0\}$. It follows that $A'' = \{(a, \varphi_e(a) - \varphi(a)) : a \in A\} \subset \mathbb{Z}^{d+1}$ cannot be contained in a hyperplane, that is $\dim(A'') \geq d + 1$. However, one can easily verify that A'' is Freĭman-isomorphic to A', giving us a contradiction. □

Proof of Corollary 6.21. Let $A' \subset \mathbb{Z}^2$ denote a set that is F_2-isomorphic to A and not contained in a line. By Theorem 6.20 we can assume that A' is contained in two lines of combined size less than $|2A| - 2|A| + 3$. By Lemma 6.22 the F_2-isomorphism φ mapping A' to A can be extended to an affine linear map, implying the desired statement. □

The Fourier-analytic Rectification

It is obvious that at least half of any set $\mathcal{A} \subset \mathbb{Z}_p$ can be rectified. It is reasonable to expect that if $\mathcal{A}$ is 'concentrated' in some sense, then one should be able to rectify significantly more than just half of the set. Freĭman stated such a result using the language of large Fourier coefficients. In the following $\widehat{\mathbb{1}}_{\mathcal{A}}(x) = \sum_{a \in \mathcal{A}} e^{2\pi i a x/p}$ will denote the Fourier transform of the indicator function of some set $\mathcal{A} \subset \mathbb{Z}_p$.

Theorem 6.23 (Section 2.1 in [61])**.** *For any $\mathcal{A} \subset \mathbb{Z}_p$ and $d \in \mathbb{Z}_p^\star$ there exists $u \in \mathbb{Z}_p$ such that*

$$\Big|[u, u + p/2) \cap d \cdot \mathcal{A}\Big| \geq \frac{|\mathcal{A}| + |\widehat{\mathbb{1}}_{\mathcal{A}}(d)|}{2}. \tag{6.30}$$

It should be noted that an improved version of this result can be obtained using a result of Lev [103]. However, we will stick to using Theorem 6.23 when proving our main statement, as the improvement that would follow from using Lev's result is negligible in our case. Lastly, it was also Freĭman who noted that a small sumset implies the existence of a large Fourier coefficient and hence a certain 'concentration' of the set. We state this observation in the following form. The proof follows by a standard application of the Cauchy-Schwarz inequality.

Lemma 6.24 (Section 2.1 in [61])**.** *For any $\mathcal{A} \subset \mathbb{Z}_p$ there exists $d \in \mathbb{Z}_p^\star$ such that*

$$\Big|\widehat{\mathbb{1}}_{\mathcal{A}}(d)\Big| \geq \left(\frac{p/|2\mathcal{A}| - 1}{p/|\mathcal{A}| - 1}\right)^{1/2} |\mathcal{A}|. \tag{6.31}$$

Proof. We start by observing that

$$\sum_{a=0}^{p-1} \widehat{\mathbb{1}}_{\mathcal{A}}(a)^2 \, \overline{\widehat{\mathbb{1}}_{2\mathcal{A}}(a)} = \sum_{a=0}^{p-1} \sum_{x_1,x_2\in\mathcal{A}} \sum_{x_3\in 2\mathcal{A}} e^{2\pi i a(x_1+x_2-x_3)/p} = |\mathcal{A}|^2 p.$$

Now if $|\widehat{\mathbb{1}}_{\mathcal{A}}(a)| \leq \theta \, |\mathcal{A}|$ for all $a \neq 0 \mod p$ and

$$\theta < \left(\frac{p/|2\mathcal{A}| - 1}{p/|\mathcal{A}| - 1} \right)^{1/2}$$

then using Cauchy-Schwarz one would get the contradiction

$$\begin{aligned} \sum_{a=0}^{p-1} \widehat{\mathbb{1}}_{\mathcal{A}}(a)^2 \, \overline{\widehat{\mathbb{1}}_{2\mathcal{A}}(a)} &= |\mathcal{A}|^2 \, |2\mathcal{A}| + \sum_{a=1}^{p-1} \widehat{\mathbb{1}}_{\mathcal{A}}(a)^2 \, \overline{\widehat{\mathbb{1}}_{2\mathcal{A}}(a)} \\ &\leq |\mathcal{A}|^2 \, |2\mathcal{A}| + \theta|\mathcal{A}| \left(\sum_{a=1}^{p-1} |\widehat{\mathbb{1}}_{\mathcal{A}}(a)|^2 \right)^{1/2} \left(\sum_{a=1}^{p-1} |\widehat{\mathbb{1}}_{2\mathcal{A}}(a)|^2 \right)^{1/2} \\ &= |\mathcal{A}|^2 \, |2\mathcal{A}| + \theta|\mathcal{A}| \left(|\mathcal{A}|p - |\mathcal{A}|^2 \right)^{1/2} \left(|2\mathcal{A}|p - |2\mathcal{A}|^2 \right)^{1/2} < |\mathcal{A}|^2 p. \end{aligned}$$

The desired statement follows. □

Proof of Theorem 6.9. Note that throughout the proof we will simplify notation by just writing $p/2$ and $p/3$ rather than the correct rounded version. In all cases there will be an appropriate amount of slack that justifies this simplification.

Let $d \in \mathbb{Z}_p^\star$ and $u \in \mathbb{Z}_p$ be such that $\mathcal{A}_1 = [u, u + p/2) \cap d \cdot \mathcal{A}$ satisfies

$$|\mathcal{A}_1| = \max_{u',d'} |[u', u' + p/2) \cap d' \cdot \mathcal{A}|. \tag{6.32}$$

We assume without loss of generality that $d = 1$ and $u = 0$. By Theorem 6.23 and Lemma 6.24 we have that

$$|\mathcal{A}_1| \geq \left(1 + \left(\frac{p/|2\mathcal{A}| - 1}{p/|\mathcal{A}| - 1} \right)^{1/2} \right) \frac{|\mathcal{A}|}{2} > 0.8175 \, |\mathcal{A}|. \tag{6.33}$$

We note that $\mathcal{A}_1$ satisfies $|2\mathcal{A}_1| \leq 3.04|\mathcal{A}_1| - 7$ as otherwise we would get the contradiction

$$2.48 \, |\mathcal{A}| - 7 \geq |2\mathcal{A}| \geq |2\mathcal{A}_1| > 3.04|\mathcal{A}_1| - 7 > 2.484 \, |\mathcal{A}| - 7. \tag{6.34}$$

As $\mathcal{A}_1$ is contained in an interval of size less than $p/2$, it is rectifiable and hence there exists some F_2-isomorphic set $A_1 \subset \mathbb{Z}$. We note that due to Equation (6.5) we have $\dim(A_1) \in \{1, 2\}$. Let us distinguish between these two cases.

Case 1. If $\dim(A_1) = 1$, then by Proposition 6.7 it is contained in an arithmetic progression of size less than $10^9|\mathcal{A}_1|$. If the common difference r of this progression is not 1, we may dilate by $r^{-1} \mod p$ and translate once more, so that we may assume that $\mathcal{A}_1 \subset [0, 10^9|\mathcal{A}_1|]$. Since $|\mathcal{A}_1|$ is by assumption the most elements any $p/2$-segment can contain of any dilate of $\mathcal{A}$, it follows that $\mathcal{A} \subset [0, 10^9|\mathcal{A}_1|] \cup [p/2, p/2 + 10^9|\mathcal{A}_1|]$. Therefore $\left(2 \cdot \mathcal{A}\right) \subset [0, 2 \cdot 10^9\,|\mathcal{A}_1|] \subset [0, p/2)$. Hence all of $\mathcal{A}$ can be rectified, so the $3|A| - 4$ statement in the integers gives the desired covering.

Case 2. If $\dim(A_1) = 2$ then we apply Corollary 6.21, obtaining progressions P_1, P_2 with union covering A_1, with same common difference r. We claim that we can assume without loss of generality that $\mathcal{A}_1 \subset [0, 3|\mathcal{A}|) \cup [c, c + 3|\mathcal{A}|) \subset [0, p/2)$ with $0, c + 3|\mathcal{A}| - 1 \in \mathcal{A}$ for some $c \in \mathbb{Z}_p$ and $|\mathcal{A}_1 \cap [0, 3|\mathcal{A}|)| \geq |\mathcal{A}_1|/2$. Indeed, if $r \neq 1$, then we can dilate by $r^{-1} \mod p$ so as to ensure that $r^{-1} \cdot \mathcal{A}_1 \subseteq [0, 3|\mathcal{A}|) \cup [c, c + 3|\mathcal{A}|)$ with $0, c + 3|\mathcal{A}| - 1 \in r^{-1} \cdot \mathcal{A}$. If $c + 3|\mathcal{A}| < p/2$, then the first two requirements are met and we can ensure that $|r^{-1} \cdot \mathcal{A}_1 \cap [0, 3|\mathcal{A}|)| \geq |\mathcal{A}_1|/2$ by multiplying the set with -1 and translating if necessary. If $p/2 - 3|\mathcal{A}| \leq c \leq p/2 + 3|\mathcal{A}|$, then $2 \cdot r^{-1} \cdot \mathcal{A}_1$ must lie in $[-6|\mathcal{A}|, 6|\mathcal{A}|]$ and arguing as in case 1 we conclude that $2 \cdot r^{-1} \cdot \mathcal{A} \subset \{0, p/2\} + [-6|\mathcal{A}|, 6|\mathcal{A}|]$, so $4 \cdot r^{-1} \cdot \mathcal{A} \subset [-12|\mathcal{A}|, 12|\mathcal{A}|]$ and again we can rectify all of $\mathcal{A}$ and complete the argument this way. Lastly, if $p/2 + 3|\mathcal{A}| < c$ then we simply translate the set by $-c$ to meet the first two requirements and again multiply by -1 if necessary. This proves our claim.

Now, let $\mathcal{S}' = [0, 3|\mathcal{A}|)$, $\mathcal{S}'' = [c, c + 3|\mathcal{A}|)$, $\mathcal{A}_1' = \mathcal{A}_1 \cap \mathcal{S}'$ and $\mathcal{A}_1'' = \mathcal{A}_1 \cap \mathcal{S}''$. By the claim above we have $|\mathcal{A}_1'| \geq |\mathcal{A}_1|/2$ and $\mathcal{S}' \cup \mathcal{S}'' \subset [0, p/2)$. We now show that

$$\mathcal{R} = \mathcal{A} \setminus \mathcal{A}_1 = \mathcal{A} \setminus (\mathcal{S}' \cup \mathcal{S}'') \subset [2c - 3|\mathcal{A}|, 2c + 6|\mathcal{A}|) = 2\mathcal{S}'' + [-3|\mathcal{A}|, 0]. \qquad (6.35)$$

We start by observing that by assumption $\mathcal{A}_1 = \mathcal{A} \cap (\mathcal{S}' \cup \mathcal{S}'')$ was the most of $\mathcal{A}$ we could rectify. It follows that $[0, p/2) \setminus (\mathcal{S}' \cup \mathcal{S}'')$ does not contain any elements of $\mathcal{A}$. Next, let us assume that there exists $a \in \mathcal{A}$ satisfying $a \in [-3|\mathcal{A}|, 0)$. Since $\mathcal{A}_1 \cup \{a\}$ cannot be rectified, we must have $c + 3|\mathcal{A}| > p/2 - 3|\mathcal{A}|$. This implies that $[c, c + 3|\mathcal{A}|) \subset [p/2 - 6|\mathcal{A}|, p/2)$, whence

$$2 \cdot (\mathcal{A}_1 \cup \{a\}) \subset [-12|\mathcal{A}|, 6|\mathcal{A}|) \subset [0, p/2) - 12|\mathcal{A}|,$$

which contradicts our maximality assumption about $\mathcal{A}_1$. It follows that we must have $\mathcal{A} \cap [-3|\mathcal{A}|, 0) = \emptyset$. Arguing similarly, we see that $\mathcal{A} \cap [c+3|\mathcal{A}|, c+6|\mathcal{A}|) = \emptyset$: certainly $\mathcal{A} \cap [c+3|\mathcal{A}|, p/2) = \emptyset$, and if there is $a \in \mathcal{A} \cap [p/2, c+6|\mathcal{A}|) \subset [p/2, p/2+3|\mathcal{A}|)$ then $[c, c+3|\mathcal{A}|) \subset [p/2-6|\mathcal{A}|, p/2)$, and so $2 \cdot (\mathcal{A}_1 \cup \{a\}) \subset [-12|\mathcal{A}|, 12|\mathcal{A}|)$, again contradicting our maximality assumption.

Next, we note that

$$2\mathcal{A}_1 \subset [0, 6|\mathcal{A}|) \cup [c, c+6|\mathcal{A}|) \cup [2c, 2c+6|\mathcal{A}|).$$

It follows that if there exists $a \in \mathcal{A}$ satisfying

$$a \in [p/2, 0) \setminus \big([-3|\mathcal{A}|, 6|\mathcal{A}|) \cup [c-3|\mathcal{A}|, c+6|\mathcal{A}|) \cup [2c-3|\mathcal{A}|, 2c+6|\mathcal{A}|)\big),$$

then $a + \mathcal{A}_1'$ does not intersect $2\mathcal{A}_1$ and we get the contradiction

$$\begin{aligned} |2\mathcal{A}| &\geq |2\mathcal{A}_1| + |a + \mathcal{A}_1'| \\ &\geq (2|\mathcal{A}_1'| - 1) + (2|\mathcal{A}_1''| - 1) + (|\mathcal{A}_1'| + |\mathcal{A}_1''| - 1) + |\mathcal{A}_1'| \\ &\geq 3.5|\mathcal{A}_1| - 3 > 2.48|\mathcal{A}| - 7. \end{aligned}$$

Note that we have used that $2\mathcal{A}_1' \cap (A_1' + A_1'') = \emptyset$ as well as $2\mathcal{A}_1'' \cap (A_1' + A_1'') = \emptyset$ as given by Corollary 6.21. Using the previous observations, it follows that Equation (6.35) is established and we have $\mathcal{A} = \mathcal{A}_1' \cup \mathcal{A}_1'' \cup \mathcal{R}$ where $\mathcal{R} = \mathcal{A} \cap [2c-3|\mathcal{A}|, 2c+6|\mathcal{A}|)$. Note that we may assume that $|\mathcal{R}| \geq 0.17|\mathcal{A}|$ as otherwise $|\mathcal{A}_1| \geq 0.83|\mathcal{A}|$ and in crefeq:A1sumset we would in fact get $|2\mathcal{A}_1| \leq 3|\mathcal{A}| - 4$, which due to Equation (6.5) would contradict our assumption that $\mathcal{A}_1$ is 2-dimensional.

We note that $2\mathcal{A} \supseteq 2\mathcal{A}_1 \cup (\mathcal{A}_1'' + \mathcal{R})$ and that trivially $|\mathcal{A}_1'' + \mathcal{R}| \geq |\mathcal{R}|$. It follows that $\mathcal{A}_1'' + \mathcal{R}$ must intersect $2\mathcal{A}_1$ since otherwise we would get the contradiction

$$|2\mathcal{A}| \geq |2\mathcal{A}_1| + |\mathcal{R}| \geq 3.17|\mathcal{A}_1| - 2 > 2.48|\mathcal{A}| - 7.$$

It follows that one of the following must hold:

(i) If $(\mathcal{A}_1'' + \mathcal{R}) \cap 2\mathcal{A}_1'' \neq \emptyset$, then we must have

$$3c + 9|\mathcal{A}| - p \geq 2c \quad \text{and} \quad 3c - 3|\mathcal{A}| - p \leq 2c + 6|\mathcal{A}|,$$

and therefore $c \in [p-9|\mathcal{A}|, p+9|\mathcal{A}|]$. However, we know that $c \leq p/2$ and that the cardinality of $\mathcal{A}$ is sufficiently small with respect to p, so we get a contradiction.

(ii) If $(\mathcal{A}_1'' + \mathcal{R}) \cap (\mathcal{A}_1' + \mathcal{A}_1'') \neq \emptyset$, then we must have

$$3c + 9|\mathcal{A}| - p \geq c \quad \text{and} \quad 3c - 3|\mathcal{A}| - p \leq c + 6|\mathcal{A}|,$$

and therefore $c \in [p/2 - {}^{9}\!/\!{}_{2}\,|\mathcal{A}|, p/2 + {}^{9}\!/\!{}_{2}\,|\mathcal{A}|]$. Consequently, in this case $\mathcal{A}_1'$ and $\mathcal{R}$ are focused around 0 and $\mathcal{A}_1''$ is focused around $p/2$. It follows that a dilation by a factor of 2 focuses all parts of $\mathcal{A}$ around 0, that is

$$2 \cdot \mathcal{A} \subset [-24|\mathcal{A}|, 30|\mathcal{A}|) \subset -24|\mathcal{A}| + [0, p/2).$$

This means that all of $\mathcal{A}$ can be rectified and we can just apply the $3|\mathcal{A}| - 4$ statement in the integers to get the desired covering property.

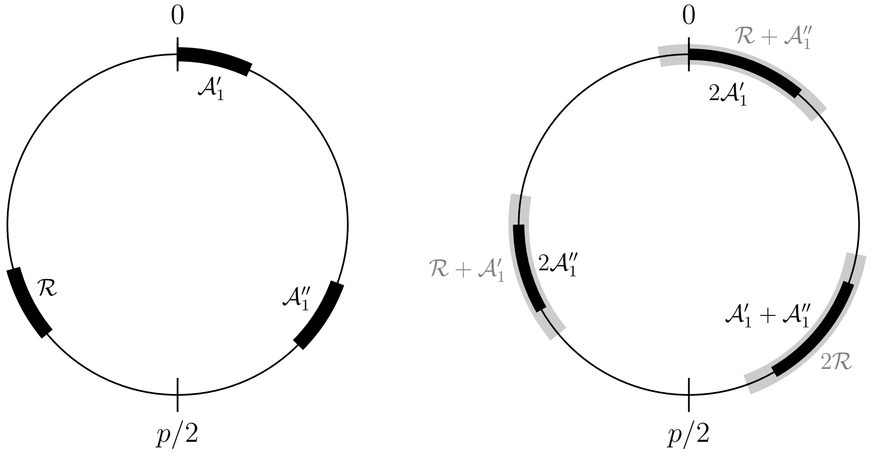


Figure 6.2: *Distribution of $\mathcal{A}$ and $2\mathcal{A}$ in $\mathbb{Z}_p$ in case (iii).*

(iii) If $(\mathcal{A}_1'' + \mathcal{R}) \cap 2\mathcal{A}_1' \neq \emptyset$, then we must have

$$3c + 9|\mathcal{A}| - p \geq 0 \quad \text{and} \quad 3c - 3|\mathcal{A}| - p \leq 6|\mathcal{A}|,$$

and therefore $c \in [p/3 - 3|\mathcal{A}|, p/3 + 3|\mathcal{A}|]$. Consequently, in this case $\mathcal{A}_1', \mathcal{A}_1''$ and $\mathcal{R}$ (or rather the intervals containing them) are roughly 'equally distributed' in $\mathbb{Z}_p$, that is they are respectively focused around $0, p/3$ and $2p/3$ as illustrated in Figure 6.2. It follows that a dilation by a factor of 3 focuses all parts of $\mathcal{A}$ around

0, that is

$$3 \cdot \mathcal{A} \subset [-27|\mathcal{A}|, 54|\mathcal{A}|) \subset -27|\mathcal{A}| + [0, p/2).$$

This again means that all of $\mathcal{A}$ can be rectified and we can just apply the $3|\mathcal{A}|-4$ statement in the integers to get the desired covering property.

It follows that we have proved the statement of Theorem 6.9. ■

6.4 Further remarks

It is probably unreasonable to expect that the rectification methodology used to prove Theorem 6.9 will lead to a proof of Conjecture 6.8. Even if all other ingredients existed in their ideal form, the rectification argument through a large Fourier coefficient appears to imply an inherent loss in the density. This is a problem concerning not only Freĭman's original approach and the result presented in this chapter, but also the broader result of Green and Ruzsa.

In fact, the more natural direction seems to be to apply covering results in the cyclic group in order to prove covering statements in the integers. Both Lev and Smeliansky's proof of Freĭman's $3|A|-4$ statement in the integers as well as the proof of Proposition 6.7 fall into that category. To further strengthen this argument, let us show that Conjecture 6.8 would also imply Conjecture 6.5 for the, admittedly rather odd, assumption that the maximum element of the normalization of the set in the integers is prime.

Corollary 6.25. *Assume that Conjecture 6.8 holds and let $A \subset \mathbb{Z}$ be a 1-dimensional set in normal form for which $\max(A)$ is prime. If $|2A| = 3|A| - 4 + b \leq 4|A| - 8$, then A can be covered by an arithmetic progression of length at most $2(|A| + b - 2) + 1$. If $|2A| = 4|A| - 7$, then A can be covered by an arithmetic progression of length at most $4|A| - 8$.*

Proof. Let $\mathcal{A}$ denote the canonical embedding of A into $\mathbb{Z}_{\max(A)}$. We start with the case $|2A| \leq 4|A| - 8$ and note as in the proof of Proposition 6.7 that

$$|2\mathcal{A}| \leq |2A| - |A| \leq 2|\mathcal{A}| - 1 + (b-1) \leq 3|\mathcal{A}| - 5.$$

Setting $x = |2\mathcal{A}| - (2|\mathcal{A}| - 1) \leq b - 1 \leq |\mathcal{A}| - 4$ it would follow from case (i) of Conjecture 6.8 that either $|2\mathcal{A}| > \max(A) - (x + 2)$ and therefore we get the desired covering property for A, or that $\mathcal{A}$ can be contained in an arithmetic progression of length at most $|\mathcal{A}| + x \leq (\max(A) + 1)/2$, implying that $\mathcal{A}$ is rectifiable and therefore by Lemma 6.19 contradicting the requirement that A is 1-dimensional.

Now if $|2A| = 4|A| - 7$ then $x = |2\mathcal{A}| - (2|\mathcal{A}| - 1) \leq |\mathcal{A}| - 3$ and we again either get the desired covering property from Conjecture 6.8, or we get a contradiction to the requirement that A is 1-dimensional. □

Note that this proof is essentially the same as that of Proposition 6.7. To prove such a statement without the primality requirement, one would need an analogue of Conjecture 6.8 in general $\mathbb{Z}_m$, that is a strengthening of the results of Kemperman [91] or Deshouillers and Freĭman [42]. Such a conjecture has not been explicitly formulated and might in fact be very intricate to state.

Chapter 7

Near-Constant Representation Functions

One of the most prominent conjectures in the theory of representation functions is that of Erdős and Turán for additive bases [55].

Erdős–Turán Conjecture on Additive Bases. *If for a given set $\mathcal{A} \subseteq \mathbb{N}_0$ the representation function satisfies $r(\mathcal{A}, n) > 0$ for n large enough, then*

$$\limsup_{n\to\infty} r(\mathcal{A}, n) = \infty. \tag{7.1}$$

This conjecture remains unsolved, though some related statements are known about the behavior of $r(\mathcal{A}, n)$: a classical result of Erdős and Fuchs [50] establishes that $r(\mathcal{A}, n)$ cannot be 'focused' around some constant $c > 0$.

Erdős–Fuchs Theorem. *For any $c > 0$ there does not exist any set $\mathcal{A} \subseteq \mathbb{N}_0$ satisfying*

$$\sum_{j=0}^{n} (r(\mathcal{A}, j) - c) = o\left(n^{1/4} \log^{-1/2} n\right). \tag{7.2}$$

Jurkat (seemingly unpublished) and later Montgomery and Vaughan [108] improved upon the result of Erdős and Fuchs by replacing the right-hand term with $o(n^{1/4})$. On the other hand, Ruzsa [127] showed that a sequence satisfying

$$\sum_{j=0}^{n} \Big(r(\mathcal{A}, j) - c\Big) = O\Big(n^{1/4} \log n\Big) \tag{7.3}$$

does in fact exist through a probabilistic argument.

It should be noted that these statements are connected to Gauss' well-known circle problem, in which one is interested in determining the asymptotic behavior of the difference between the number of integer lattice points in a circle of radius r and the area of that circle. Hardy [81] and independently Landau proved that it cannot be of the order $o(r^{1/2}\log^{1/4} r)$. By choosing the set $\mathcal{A}$ to be all squares of integers, the result of Erdős and Fuchs and the subsequent improvement due to Jurkat as well as Montgomery and Vaughan result in an only slightly weaker bound than that of Hardy and Landau while dealing with a much more general problem.

There have been several generalizations of the result of Erdős and Fuchs, for example to the case of more than two summands, see [82, 118, 152], and also to the case where the summands come from different sequences, see [36, 85, 133]. In the remainder of this section, we will be interested in studying two specific variants of representation functions: the first are *ordered representation functions* with multiple summands. For these we will establish an Erdős–Fuchs-type result. The second are the much more involved *weighted representation functions*, again with multiple summands. The problem here, as stated by Sárközy and Sós [134], is to characterize the weights for which the representation function cannot become constant. However, before stating the results regarding these two variants, let us first introduce the language in which one usually approaches these types of problems, which goes back to Dirac [43].

Rephrasing the problems through generating functions

The generating function of a set $\mathcal{A} \subseteq \mathbb{N}_0$ is the formal power series

$$f_{\mathcal{A}}(z) = \sum_{a\in\mathcal{A}} z^a. \tag{7.4}$$

Writing $[z^k]f(z)$ for the coefficient c_k of a given formal power series $f(z) = \sum_{k\geq 0} c_k z^k$, we note that $[z^k]f_{\mathcal{A}}(z)$ is the indicator function $\mathbb{1}_{\mathcal{A}}(k)$ of $\mathcal{A}$. We also observe that $f_{\mathcal{A}}(z)$ is analytic in the open disc $\mathbb{D}$ and that it is a strictly increasing in the real interval $[0, 1)$.

Let us make some light use of the Symbolic Method, a common technique for counting combinatorial objects. A more detailed introduction to this can be found in the book of Flajolet and Sedgewick [57]. For the particular case of the representation function $r(\mathcal{A}, n)$, one would first define the combinatorial class $(\mathcal{A}, |\cdot|)$ where the size of an element $a \in \mathcal{A}$ is simply $|a| = a$. The ordinary generating function of this class is just the previously introduced generating function $f_{\mathcal{A}}(z)$ of $\mathcal{A}$. The combinatorial class we are then interested in enumerating is that of $(\mathcal{A}^2, |\cdot|)$ where the size of an element in $\mathcal{A}^2$ is simply $|(a_1, a_2)| = |a_1| + |a_2| = a_1 + a_2$. The ordinary generating function of this class is therefore $\sum_{n=0}^{\infty} r(\mathcal{A}, n) z^n$. Since $(\mathcal{A}^2, |\cdot|)$ is the two-fold cartesian product of $(\mathcal{A}, |\cdot|)$, it follows that

$$\sum_{n=0}^{\infty} r(\mathcal{A}, n)\, z^n = f_{\mathcal{A}}(z)^2. \tag{7.5}$$

This observation has been the basis of many proofs regarding the representation function $r(\mathcal{A}, n)$ and its generalizations and variants. Of course, in this particular case one can easily convince oneself of the veracity of Equation (7.5) without using the Symbolic Method. However, especially when dealing with the ordered representation function, this will be an important framework to keep in mind. Let us now introduce the main results of this section.

Ordered representation functions

For a fixed $k \geq 2$, we define the ordered representation functions of a set $\mathcal{A} \subseteq \mathbb{N}_0$ as

$$r_k^{\leq}(\mathcal{A}, n) = \#\Big\{(a_1, \ldots, a_k) \in \mathcal{A}^k : a_1 \leq \cdots \leq a_k,\ a_1 + \cdots + a_k = n\Big\} \tag{7.6}$$

as well as

$$r_k^{<}(\mathcal{A}, n) = \#\Big\{(a_1, \ldots, a_k) \in \mathcal{A}^k : a_1 < \cdots < a_k,\ a_1 + \cdots + a_k = n\Big\}. \tag{7.7}$$

They both count the number of ways to express some integer $n \in \mathbb{N}_0$ as a sum of k elements in $\mathcal{A}$ sorted in increasing order, which is equivalent to counting sets of elements rather than tuples. The former of the two functions allows for repetition whereas the later requires the elements to be distinct.

In order to place these functions in some context to the representation function defined at the beginning of this section, we note that clearly

$$\Big(r(\mathcal{A}, n) - 1\Big)/2 \leq r_2^{<}(\mathcal{A}, n) \leq r_2^{\leq}(\mathcal{A}, n) \leq r(\mathcal{A}, n) \tag{7.8}$$

holds for any $n \in \mathbb{N}_0$. The following result now establishes that the ordered representation functions also satisfy an Erdős–Fuchs-type result.

Theorem 7.1. *Let $k \geq 2$, $c > 0$ and $\star \in \{\leq, <\}$. There does not exist any set $\mathcal{A} \subseteq \mathbb{N}_0$*

satisfying

$$\sum_{j=0}^{n} (r_k^\star(\mathcal{A}, j) - c) = o\left(n^{1/4} \log^{-1/2} n\right).$$

The proof of this result is distinct in two ways: first, as already mentioned, one needs to use some additional tools from the Symbolic Method in order to obtain an expression of $\sum_{n=0}^{\infty} r_k^\star(\mathcal{A}, n)\, z^n$ in terms of the generating function of $\mathcal{A}$ along the lines of Equation (7.5). Second, while the proof of Erdős and Fuchs is based on integrating Equation (7.5) along a small arc, the proof of Theorem 7.1 uses an integration with a smoothing function around the whole circle, which is reminiscent of the techniques used in [133].

The tools developed to prove Theorem 7.1 also allow one to prove an equivalent statement to another result of Erdős and Fuchs, namely Theorem 2 in [50]. For any $c \geq 0$ and $\mathcal{A} \subseteq \mathbb{N}_0$, let

$$E_{k,c}^\star(\mathcal{A}, n) = \frac{1}{n} \sum_{j=0}^{n} (r_k^\star(\mathcal{A}, j) - c)^2$$

denote the mean squared error. The following statement is in to spirit of that lesser known result of Erdős and Fuchs.

Theorem 7.2. *Let $k \geq 2$, $c \geq 0$, $\star \in \{\leq, <\}$ and $\mathcal{A} = \{a_1 < a_2 < a_3 < \ldots\} \subseteq \mathbb{N}_0$. If either $c > 0$ or the set $\{a_s/s^k\}_{s\in\mathbb{N}}$ is bounded, then $\limsup_{n\to\infty} E_{k,c}^\star(\mathcal{A}, n) > 0$.*

Weighted representation functions

For some fixed integer $d \geq 2$ and a vector of weights $\mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{N}^d$, we now define the weighted representation function of a set $\mathcal{A} \subseteq \mathbb{N}_0$ as

$$r_{\mathbf{k}}(\mathcal{A}, n) = r(\mathcal{A}, n; k_1, \ldots, k_d) = \#\Big\{(a_1, \ldots, a_d) \in \mathcal{A}^d : k_1 a_1 + \cdots + k_d a_d = n\Big\}.$$

Note that clearly $r(\mathcal{A}, n; 1, 1) = r(\mathcal{A}, n)$. Studying these functions is much more involved than those previously introduced and they display a much larger range of behavior depending on the weights. Consequently, Sárközy and Sós asked a much simpler question regarding the weighted representation functions [134, Problem 7.1.]: for which values of $k_1, \ldots, k_d$ can one find an infinite set $\mathcal{A}$ such that the function $r(\mathcal{A}, n; k_1, \ldots, k_d)$ becomes constant for n large enough?

The result of Erdős and Fuchs establishes that $r(\mathcal{A}, n)$ not only cannot become constant but is in fact very far from it. However, there is a much easier way to answer the particular question we are interested in for weighted representation functions in the case of $r(\mathcal{A}, n)$: the function is odd whenever $n = 2a$ for some $a \in \mathcal{A}$ and even otherwise and therefore cannot become constant. For $k \geq 2$, Moser [109] constructed a set $\mathcal{A}$ such that $r(\mathcal{A}, n; 1, k) = 1$ for all $n \in \mathbb{N}_0$. The study of bivariate linear forms, that is the case of $d = 2$, was finally completely settled by Cilleruelo and Rué [30] by showing that the only cases in which $r(\mathcal{A}, n; k_1, k_2)$ may become constant are those considered by Moser.

The multivariate case where $d > 2$ is less well studied. If $\gcd(k_1, \ldots, k_d) > 1$, then one trivially observes that $r(n; k_1, \ldots, k_d)$ cannot become constant. The only non-trivial case so far was studied by the Rué [120], showing that if in the d-tuple of coefficients $(k_1, \ldots, k_d)$ each element is repeated exactly m times, then there cannot exist an infinite set $\mathcal{A}$ such that $r(\mathcal{A}, n; k_1, \ldots, k_d)$ becomes constant for n large enough. This for example covers the case $(k_1, k_2, k_3, k_4, k_5, k_6) = (2, 4, 6, 2, 4, 6)$. Observe that each coefficient in this example is repeated twice, that is $m = 2$.

Here we will go a step beyond this result and show that whenever the set of coefficients is pairwise co-prime, then there does not exist any infinite set $\mathcal{A}$ for which $r(n; k_1, \ldots, k_d)$ is constant for n large enough. In fact, the following statement covers an even wider range of weights than those that are pairwise co-prime.

Theorem 7.3. *Let $q_1, \ldots, q_m \geq 2$ be pairwise co-prime integers and $b(i, j) \in \{0, 1\}$, so that for each $1 \leq i \leq d$ there exists some $1 \leq j \leq m$ such that $b(i, j) = 1$. If $k_i = q_1^{b(i,1)} \cdots q_m^{b(i,m)}$ for $1 \leq i \leq d$, then for every infinite set $\mathcal{A} \subseteq \mathbb{N}_0$ the function $r(\mathcal{A}, n; k_1, \ldots, k_d)$ cannot become constant.*

In particular, if $m = d$ and $b(i, j) = 1$ if and only if $i = j$, then this represents the case where $k_1, \ldots, k_d \geq 2$ are pairwise co-prime numbers. Other new cases covered by this result are for instance $(k_1, k_2, k_3) = (2, 3, 2 \cdot 3)$ as well as $(k_1, k_2, k_3, k_4) = (2^2 \cdot 3, 2^2 \cdot 5, 3 \cdot 5, 2^2 \cdot 3 \cdot 5)$. The proof of this result starts with some ideas introduced in [30] dealing with generating functions and cyclotomic polynomials. The main new idea is to use an inductive argument in order to be able to show that a certain multivariate recurrence relation is not possible to be satisfied unless some initial condition is trivial.

7.1 Preliminaries for Theorem 7.1 and Theorem 7.2

In this part we will establish some common preliminaries that will be required for the proofs of Theorem 7.1 and Theorem 7.2. The first is the previously already mentioned issue of finding a way to encode the problem using the generating functions of the sets. Once we have established an equivalent to Equation (7.5) for ordered representation functions, we will turn our attention to obtaining a bound when integrating that encoding.

7.1.1 Encoding the ordered representation functions

Expressing $\sum_{n=0}^{\infty} r_k^{\star}(\mathcal{A}, n)\, z^n$ in terms of $f_{\mathcal{A}}(z)$ for $\star \in \{\leq, <\}$ is slightly more involved than it is for $\sum_{n=0}^{\infty} r(\mathcal{A}, n)\, z^n$, since we can no longer use the Cartesian product construction. For $r_k^{\leq}(\mathcal{A}, n)$ we are instead interested in enumerating the combinatorial class given by the multiset construction of $(\mathcal{A}, |\cdot|)$. Here we will need another variable u to keep track of the size of each set. Adapting Theorem I.3. in [57] to our setting therefore gives the expressions

$$\sum_{n=0}^{\infty} r_k^{\leq}(\mathcal{A}, n)\, z^n = [u^k] \exp\left(\sum_{i=1}^{\infty} \frac{1}{i} u^i f_{\mathcal{A}}(z^i)\right). \tag{7.9}$$

For $r_k^{<}(\mathcal{A}, n)$ on the other hand we need the powerset construction of $(\mathcal{A}, |\cdot|)$, where we also use an additional variable u to keep track of the size of the sequences. Again adapting Theorem I.3. in [57] to our setting gives us the expressions

$$\sum_{n=0}^{\infty} r_k^{<}(\mathcal{A}, n)\, z^n = [u^k] \exp\left(\sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} u^i f_{\mathcal{A}}(z^i)\right). \tag{7.10}$$

Writing

$$S(k) = \Big\{ \mathbf{i} = (i_1, \ldots, i_m) \in \mathbb{N}^m : 1 \leq m \leq k \text{ and } i_1 + \cdots + i_m = k \Big\}$$

and expanding the Taylor series of expressions Equation (7.9) and Equation (7.10), we get that

$$\sum_{n=0}^{\infty} r_k^{\leq}(\mathcal{A}, n)\, z^n = \sum_{\mathbf{i} \in S(k)} \frac{f_{\mathcal{A}}(z^{i_1}) \cdots f_{\mathcal{A}}(z^{i_m})}{i_1 \cdots i_m \cdot m!} \tag{7.11}$$

as well as

$$\sum_{n=0}^{\infty} r_k^{<}(\mathcal{A}, n)\, z^n = \sum_{\mathbf{i} \in S(k)} (-1)^{m+k} \frac{f_{\mathcal{A}}(z^{i_1}) \cdots f_{\mathcal{A}}(z^{i_m})}{i_1 \cdots i_m \cdot m!}. \tag{7.12}$$

Let us generalize our notation and write $\varepsilon_{\leq}(\mathbf{i}) = 1/(i_1 \cdots i_m \cdot m!)$ as well as $\varepsilon_{<}(\mathbf{i}) = (-1)^{m+k}/(i_1 \cdots i_m \cdot m!)$ for any $\mathbf{i} = (i_1, \ldots, i_m) \in S(k)$, so that Equation (7.11) and

Equation (7.12) both become

$$\sum_{n=0}^{\infty} r_k^{\star}(\mathcal{A},n)\, z^n = \sum_{\mathbf{i}\in S(k)} \varepsilon_{\star}(\mathbf{i})\, f_{\mathcal{A}}(z^{i_1})\cdots f_{\mathcal{A}}(z^{i_m}) \tag{7.13}$$

for $\star \in \{\leq, <\}$. We have therefore found our equivalent to Equation (7.5) for ordered representation functions. Finally, in both of the proofs of Theorem 7.1 and Theorem 7.2 we will argue that the term coming from the k-tuple $\mathbf{1} = (1,1,\ldots,1) \in S(k)$ will asymptotically be dominant and we therefore let

$$S_0(k) = \Big\{\mathbf{i} = (i_1,\ldots,i_m) \in \mathbb{N}^m : 1 \leq m < k \text{ and } i_1 + \cdots + i_m = k\Big\}$$

denote the set of all remaining terms. Also note that $\varepsilon_{\leq}(\mathbf{1}) = \varepsilon_{<}(\mathbf{1}) = 1/k!$.

7.1.2 The dominant term under integration

We start by noting that the variable with respect to which any asymptotic statements are made will either be $n \in \mathbb{N}$ tending to infinity or $r \in (1/2, 1)$ tending to 1. It will always specified with respect to which of the two any asymptotic statements are made through an indexed n or r. We will also assume that $r > 1/2$ simply to avoid any unimportant behavior that occurs when r is close to 0. The variable z will always lie in

$$\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}, \tag{7.14}$$

the open disk of radius 1. Integrals along

$$\mathbb{S}_r = \{z \in \mathbb{C} : |z| = r\}, \tag{7.15}$$

that is the circle of radius r, will be taken with respect to the measure

$$d\mu = |dz|/(2\pi r) = d\theta/(2\pi)$$

where $z = re^{i\theta}$. In particular, this implies that $\int_{\mathbb{S}_r} d\mu = 1$. We will also use the smoothing function

$$h_M(z) = 1 + z + \ldots + z^{M-1} = \frac{1 - z^M}{1 - z} \tag{7.16}$$

where $M \in \mathbb{N}$. We will furthermore write $h_0(z) = 1$ in order to simplify notation.

The goal of this part is to formalise the fact that, when integrating the right-hand side of Equation (7.13) over $\mathbb{S}_r$, the dominant term as r tends to 1 will come from the k-tuple $\mathbf{1} \in S(k)$ whereas all terms coming from $S_0(k)$ will be negligible. We first prove the following lemma which establishes an application of Parseval's identity as well as Hölder's inequality that will be used throughout the rest of this section. It is this technique that allows us to extend previously established Erdős–Fuchs-type results to the case of $k > 2$.

Lemma 7.4. *If $g(z) = \sum_{n=0}^{\infty} b_n z_n$ has non-negative integer coefficients and is analytic in $\mathbb{D}$, then*

$$\int_{\mathbb{S}_r} |g(z)|^2 d\mu = \sum_{n=0}^{\infty} |b_n|^2 r^{2n} \tag{7.17}$$

and if $k \geq 2$, then

$$\int_{\mathbb{S}_r} |g(z)|^k d\mu \geq g(r^2)^{k/2}. \tag{7.18}$$

Proof. For $k = 2$, Parseval's identity gives us

$$\begin{aligned} \int_{\mathbb{S}_r} |g(z)|^2 d\mu = \int_{\mathbb{S}_r} g(z)\overline{g(z)} d\mu &= \sum_{n,m\geq 0} b_n \overline{b_m} \int_{\mathbb{S}_r} z^n \overline{z^m} d\mu \\ &= \sum_{n=0}^{\infty} |b_n|^2 \, r^{2n} \geq \sum_{n=0}^{\infty} b_n \left(r^2\right)^n = g\left(r^2\right). \end{aligned}$$

When $k > 2$, then Hölder's inequality and the observation for $k = 2$ establish that

$$\left(\int_{\mathbb{S}_r} |g(z)|^k d\mu \right)^{2/k} \left(\int_{\mathbb{S}_r} d\mu \right)^{(k-2)/k} \geq \int_{\mathbb{S}_r} |g(z)|^2 d\mu \geq g\left(r^2\right).$$

Noting that $\int_{\mathbb{S}_r} d\mu = 1$ and raising the previous inequality to the $k/2$-th power gives us the result. □

We are now ready to prove the main auxiliary statement of this part in three steps. We start with the following lemma.

Lemma 7.5. *For any $M, k, m \in \mathbb{N}$ satisfying $m \leq k$ and any $\mathcal{A} \subseteq \mathbb{N}_0$ we have*

$$\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k \, h_M(z)^2| d\mu \geq \left(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^m \, h_M(z)^2| d\mu \right) \left(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^2| d\mu \right)^{(k-m)/2}.$$

Proof. We start by considering three different cases concerning k and m.

Case 1. If $k = m + 2$ and k is even (and therefore m is even), then let

$$a_n^{M,m} = [z^n] f_{\mathcal{A}}(z)^{m/2}\, h_M(z),$$

that is $f_{\mathcal{A}}(z)^{m/2}\, h_M(z) = \sum_{n=0}^{\infty} a_n^{M,m} z^n$. Note that

$$f_{\mathcal{A}}(z)^{k/2}\, h_M(z) = \Big(f_{\mathcal{A}}(z)^{m/2}\, h_M(z)\Big)\, f_{\mathcal{A}}(z)$$

and therefore

$$\begin{aligned} a_n^{M,k} &= [z^n]\Big(f_{\mathcal{A}}(z)^{m/2}\, h_M(z)\Big)\, f_{\mathcal{A}}(z) \\ &= \sum_{i+j=n} [z^i] f_{\mathcal{A}}(z)^{m/2}\, h_M(z)\, [z^j] f_{\mathcal{A}}(z) = \sum_{i+j=n} a_i^{M,m} \mathbb{1}_{\mathcal{A}}(j) \end{aligned}$$

for any $n \geq 0$. It follows that

$$\left|a_n^{M,k}\right|^2 \geq \sum_{i+j=n} \left|a_i^{M,m}\right|^2 |\mathbb{1}_{\mathcal{A}}(j)|^2.$$

Using this inequality as well as Equation (7.17) in Lemma 7.4, we conclude that

$$\begin{aligned} \int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k\, h_M(z)^2| d\mu &= \sum_{n=0}^{\infty} |a_n^{M,k}|^2\, r^{2n} \geq \sum_{n=0}^{\infty} \sum_{i+j=n} |a_i^{M,m}|^2\, |\mathbb{1}_{\mathcal{A}}(j)|^2\, r^{2i+2j} \\ &= \left(\sum_{i=0}^{\infty} |a_i^{M,m}|^2\, r^{2i}\right) \left(\sum_{j=0}^{\infty} |\mathbb{1}_{\mathcal{A}}(j)|^2\, r^{2j}\right) \\ &= \int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^m\, h_M(z)^2| d\mu \int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^2| d\mu. \end{aligned}$$

Case 2. If $k = m + 1$ and k is even, then applying Cauchy–Schwarz and Case 1, we get that

$$\begin{aligned} &\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^m\, h_M(z)^2| d\mu \\ &\leq \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{m+1} h_M(z)^2| d\mu\Big)^{1/2} \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{m-1} h_M(z)^2| d\mu\Big)^{1/2} \\ &\leq \int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k\, h_M(z)^2| d\mu \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^2| d\mu\Big)^{-1/2}. \end{aligned}$$

Passing the last integral to the left-hand side establishes the statement in this case.

Case 3. If $k = m + 1$ and k is odd, then applying Cauchy–Schwarz and Case 2, we get that

$$\begin{aligned}
&\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^m\, h_M(z)^2| d\mu \\
&\leq \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{m+1} h_M(z)^2| d\mu \Big)^{1/2} \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{m-1} h_M(z)^2| d\mu \Big)^{1/2} \\
&\leq \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{k} h_M(z)^2| d\mu \Big)^{1/2} \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{m}\, h_M(z)^2| d\mu \Big)^{1/2} \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)|^2 d\mu \Big)^{-1/4}.
\end{aligned}$$

Passing the last two integrals to the left-hand side and squaring establishes the statement in this case.

Having established these three cases, the statement of the lemma now follows through an induction on k. Clearly the statement holds for $m = k = 1$. Assume now that it holds for $k-1$ and let us show that it then must also hold for k. The statement trivially holds for $m = k$ and Cases 2 and 3 establish that it also holds for $m = k-1$. For any $1 \leq m < k-1$ we can simply use the inductive assumption, since

$$\begin{aligned}
&\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k\, h_M(z)^2| d\mu \\
&\geq \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{k-1}\, h_M(z)^2| d\mu \Big) \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^2| d\mu \Big)^{1/2} \\
&\geq \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{m}\, h_M(z)^2| d\mu \Big) \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^2| d\mu \Big)^{(k-1-m)/2} \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^2| d\mu \Big)^{1/2} \\
&= \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{m}\, h_M(z)^2| d\mu \Big) \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^2| d\mu \Big)^{(k-m)/2}.
\end{aligned}$$

This proves the desired result. □

The following is a slight generalization of the previous lemma, allowing us to consider exponents in the arguments.

Lemma 7.6. *For any $M, k, m, i \in \mathbb{N}$ satisfying $i \mid M$ and $m \leq k$, and for any $\mathcal{A} \subseteq \mathbb{N}_0$, we have*

$$\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z^i)^m\, h_M(z)^2| d\mu \leq \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k\, h_M(z)^2| d\mu \Big)\, i^2 \Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)|^2 d\mu \Big)^{-(k-m)/2}.$$

Proof. Let

$$a_n^{M,m} = [z^n] f_{\mathcal{A}}(z)^{m/2} h_M(z),$$

that is $f_{\mathcal{A}}(z)^{m/2} h_M(z) = \sum_{n=0}^{\infty} a_n^{M,m} z^n$. Since M is a multiple of i, we can set $N = M/i$

and note that

$$h_M(z) = h_N(z)\, h_i(z^N) = h_N(z)(1 + z^N + \ldots + z^{(i-1)N}),$$

so that $f_{\mathcal{A}}(z)^{m/2}\, h_M(z) = f_{\mathcal{A}}(z)^{m/2} h_N(z)\,(1 + z^N + \ldots + z^{(i-1)N})$. In terms of coefficients, this implies that

$$a_n^{M,m} = a_n^{N,m} + a_{n-N}^{N,m} + a_{n-2N}^{N,m} + \ldots + a_{n-(i-1)N}^{N,m} \geq a_n^{N,m},$$

where we let $a_n^{N,m} = 0$ for $n < 0$ and use the fact that all the coefficients are non-negative as both the coefficients of f and h_N are non-negative.

Using that $h_M(z) = h_N(z)\, h_i(z^N)$, $|h_i(z)| \leq i$, Equation (7.17) in Lemma 7.4 and Section 7.1.2, we get

$$\begin{aligned} \int_{\mathbb{S}_r} |f_{\mathcal{A}}(z^i)^m\, h_M(z)^2| d\mu = \int_{\mathbb{S}_r} |f_{\mathcal{A}}(z^i)^m h_N(z^i)^2|\, |h_i(z)^2| d\mu \leq i^2 \sum_{n=0}^{\infty} |a_n^{N,m}|^2\, r^{2in} \\ \leq i^2 \sum_{n=0}^{\infty} |a_n^{M,m}|^2\, r^{2n} = i^2 \int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^m\, h_M(z)^2| d\mu. \end{aligned}$$

We conclude the statement by applying Lemma 7.5 to this inequality. □

We can now establish the main statement of this part, which is written in ready-to-use form for the proof of both Theorem 7.1 and Theorem 7.2.

Proposition 7.7. *For any $M \in \mathbb{N}_0$, $k, m, i \in \mathbb{N}$ satisfying $i \mid M$ and $m < k$, and for any infinite set $\mathcal{A} \subseteq \mathbb{N}_0$, we have*

$$\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z^i)^m\, h_M(z)^2| d\mu = o_r\Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k\, h_M(z)^2| d\mu \Big).$$

Proof. We will distinguish based on whether $M \geq 1$ or $M = 0$. The former of these is a consequence of Lemma 7.6 whereas the latter is established using only Lemma 7.4.

Case 1. If $M \geq 1$, then by Equation (7.17) in Lemma 7.4 we have $\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)|^2 d\mu = \sum_{a \in \mathcal{A}} r^{2a}$ which tends to infinity as r tends to 1 since $\mathcal{A}$ is infinite. Therefore, the result directly follows from Lemma 7.6.

Case 2. If $M = 0$, then let us first assume that m is even. Let $f_{\mathcal{A}}(z)^{m/2} = \sum_{n=0}^{\infty} b_n z^n$ and note that $b_n \geq 0$ for all $n \in \mathbb{N}_0$. Using Equation (7.17) in Lemma 7.4 and Hölder's inequality, we obtain

$$\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z^i)^m| d\mu = \sum_{n=0}^{\infty} b_n^2 r^{2ni} \leq \sum_{n=0}^{\infty} b_n^2 r^{2n}$$
$$= \int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^m| d\mu \leq \left(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k| d\mu \right)^{m/k}.$$

Again by Hölder's inequality we know that $\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)|^k d\mu \geq \left(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)|^2 d\mu \right)^{k/2} = \omega_r(1)$, so that the statement follows for even m since $k < m$.

Now assume that m is odd. Let $f_{\mathcal{A}}(z)^{(m+1)/2}(z) = \sum_{n=0}^{\infty} b'_n z^n$ and again note that $b'_n \geq 0$ for all $\mathbb{N}_0$. Applying Equation (7.17) in Lemma 7.4, we get

$$\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z^i)^{m+1}| d\mu = \sum_{n=0}^{\infty} b'^2_n r^{2ni} \leq \sum_{n=0}^{\infty} b'^2_n r^{2n}$$
$$= \int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^{m+1}| d\mu = O_r\Big(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k| d\mu \Big). \tag{7.19}$$

The last equality is trivial if $m+1 = k$ and follows from the even case if $m+1 < k$. Using Cauchy-Schwarz, the even case for $m-1$ as well as Equation (7.19), we obtain

$$\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z^i)^m| d\mu \leq \left(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z^i)^{m-1}| d\mu \right)^{1/2} \left(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z^i)^{m+1}| d\mu \right)^{1/2}$$
$$= o_r \left(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k| d\mu \right)^{1/2} O_r \left(\int_{\mathbb{S}_r} |f_{\mathcal{A}}(z)^k| d\mu \right)^{1/2},$$

proving the statement for m odd. □

7.2 Proof of Theorem 7.1 – An Erdős–Fuchs-type result

Let us establish three general asymptotic bounds that will be needed in the proof of Theorem 7.1.

Lemma 7.8. *Let $i \in \mathbb{N}$ and suppose that $g(z) = \sum_{n=0}^{\infty} b_n z^n$ has non-negative coefficients and is analytic in $\mathbb{D}$. If $\sum_{n=0}^{\infty} b_n = \infty$, then*

$$g(r^i) = O_r\Big(g(r)^i\Big).$$

Proof. Since $g(z)$ has non-negative coefficients, $g(r)$ is monotone increasing in r and tends to infinity as r tends to 1. Let $(r_n)_{n \geq 1}$ be an arbitrary increasing sequence tending to 1 in $(1/2, 1)$. Since $r_n^i < r_n$, we have $g(r_n^i) < g(r_n) = o\Big(g(r_n)^i\Big)$ for all n due to the monotonicity of g. □

Lemma 7.9. *We have*

$$\int_{\mathbb{S}_r} \frac{d\mu}{|1-z|} = O_r\Big(-\log(1-r)\Big).$$

Proof. Let $z = re^{i\theta}$. Using the fact that $|\sin\theta| \geq |\theta|/2$ for any $\theta \in [-\pi/2, \pi/2]$, we can bound the integral as

$$\begin{aligned}
\int_{\mathbb{S}_r} \frac{d\mu}{|1-z|} &= \frac{1}{2\pi r}\int_{-\pi/2}^{3\pi/2} \frac{d\theta}{\sqrt{(1-r\cos\theta)^2 + (r\sin\theta)^2}} \\
&\leq \frac{1}{2\pi r}\int_{-\pi/2}^{\pi/2} \frac{d\theta}{\sqrt{(1-r)^2 + (r\sin\theta)^2}} + \frac{1}{2\pi r}\int_{\pi/2}^{3\pi/2} \frac{d\theta}{\sqrt{1+r^2}} \\
&\leq \frac{1}{2\pi r}\int_{-\pi/2}^{\pi/2} \frac{d\theta}{\sqrt{(1-r)^2 + r^2\theta^2/4}} + 1 \leq \frac{\sqrt{2}}{\pi r}\int_{0}^{\pi/2} \frac{d\theta}{\big(1-r+r\theta/2\big)} + 1 \\
&\leq \frac{2\sqrt{2}}{\pi r^2}\log\left(\frac{\pi/4}{1-r}+1\right) + 1 = O_r\Big(-\log(1-r)\Big),
\end{aligned}$$

where we have used the inequality of the root-mean square and the arithmetic mean. □

Lemma 7.10. *For any sequence of real numbers e_n satisfying*

$$e_n = o_n\Big(n^{1/4}\log^{-1/2} n\Big),$$

we have

$$\sum_{n=0}^{\infty} e_n^2\, r^{2n} = o_r\left(\frac{-1}{(1-r)^{3/2}\log(1-r)}\right).$$

Proof. We start by showing that that there exists some $C_0 > 0$ such that

$$\sum_{n=2}^{\infty}\Big(r^n \log^{-1/2}(n)\, n^{1/4}\Big)^2 \leq C_0\,\frac{-1}{(1-r)^{3/2}\log(1-r)}. \tag{7.20}$$

Let

$$N_0 = N_0(r) = \Big\lfloor (1-r)^{-1/2}\Big\rfloor = \omega_r(1).$$

Since $r < 1$ and $\log^{-1/2}(n)$ is decreasing, we can see that

$$\sum_{n=2}^{N_0}\Big(r^n \log^{-1/2}(n)\, n^{1/4}\Big)^2 \leq \log^{-1}(2)\sum_{n=2}^{N_0} n^{1/2} \leq 2\,{N_0}^{3/2} \leq 2\,(1-r)^{-3/4}$$

$$= o_r\left(\frac{-1}{(1-r)^{3/2}\log(1-r)}\right). \tag{7.21}$$

In order to deal with the remainder of the sum, we start by noting that for any $\delta > 1$ we have

$$\begin{aligned}(1-x)^{-\delta} = \sum_{n=0}^{\infty} c_n x^n &= \sum_{n=0}^{\infty} (-1)^n \frac{-\delta\,(-\delta-1)\cdots(-\delta-n+1)}{n!}\,x^n \\ &= \sum_{n=0}^{\infty} \frac{\delta\,(\delta+1)\cdots(\delta+n-1)}{n!}\,x^n.\end{aligned}$$

The asymptotic behavior of the logarithm of the coefficients of the Taylor expansion of $(1-x)^{-\delta}$ is therefore given by

$$\log c_n = \sum_{i=1}^{n} \log\left(\frac{\delta+i-1}{i}\right) \geq \int_1^{n+1} \log\left(\frac{\delta+x-1}{x}\right) dx$$

where we have used that that $1 + (\delta-1)/i$ is decreasing in i since $\delta > 1$. Using that

$$\int \log\left(\frac{\delta-1}{x}+1\right) dx = (\delta-1)\log(\delta-1+x) + x\log\left(\frac{\delta-1}{x}+1\right)$$

and $x\log\Big(1+(\delta-1)/x\Big) > 0$ for any $x > 0$, it follows that

$$\log c_n \geq (\delta-1)\,\log n - (\delta-1)\log\delta - \log\delta.$$

We therefore have for some appropriate constant $C_1 = C_1(\delta)$ such that

$$n^{\delta-1} \leq C_1 c_n$$

for any $n \geq 1$. Using this estimate with $\delta = 3/2$, we get

$$\begin{aligned}\sum_{n=N_0+1}^{\infty} \Big(r^n\,\log^{-1/2}(n)\,n^{1/4}\Big)^2 &\leq \log^{-1}(N_0)\,C_1 \sum_{n=1}^{\infty} c_n r^{2n} \\ &= C_1\,\frac{1}{(1-r^2)^{3/2}\log(N_0)} \\ &= O_r\left(\frac{-1}{(1-r)^{3/2}\log(1-r)}\right).\end{aligned} \tag{7.22}$$

Collecting the estimate for $2 \leq n \leq N_0$ in Equation (7.21) and the estimate for

$n \geq N_0 + 1$ in Equation (7.22) establishes Equation (7.20). Now let $\varepsilon > 0$ and choose $N_1 \geq 2$ large enough such that

$$e_n \leq \frac{\varepsilon^{1/2}}{C_0^{1/2}} \, n^{1/4} \log^{-1/2}(n)$$

for all $n \geq N_1$. From Equation (7.20) it follows that

$$\sum_{n=N}^{\infty} e_n^2 \, r^{2n} \leq \frac{-\varepsilon}{(1-r)^{3/2} \log(1-r)}. \tag{7.23}$$

On the other hand, there exists $r_0 = r_0(N_1)$ close enough to 1 so that for any $r \geq r_0$

$$\sum_{n=0}^{N-1} e_n^2 r^{2n} \leq \sum_{n=0}^{N-1} e_n^2 \leq \frac{-\varepsilon}{(1-r)^{3/2} \log(1-r)} \tag{7.24}$$

since the right-hand side tends to ∞ as r tends to 1 while the left-hand side remains constant. Combining Equation (7.23) and Equation (7.24) and letting ε tend to 0 as r tends to 1, the statement of the lemma follows. □

We are now ready to prove Theorem 7.1.

Proof of Theorem 7.1. Throughout this proof $k \geq 2$, $c > 0$, $\star \in \{\leq, <\}$ and $\mathcal{A} \subseteq \mathbb{N}_0$ are fixed. Furthermore, z will always lie in $\mathbb{D}$ and r in the open interval $(1/2, 1)$. We write

$$e_n = \sum_{j=0}^{n} \Big(r_k^\star(\mathcal{A}, j) - c \Big) \tag{7.25}$$

for $n \geq 0$ and assume that, counter to the statement of Theorem 7.1, we have $e_n = O_n\Big(n^{1/4} \log^{-1/2} n\Big)$. Multiplying Equation (7.25) by z^n and summing over $n \geq 0$ gives us

$$\sum_{n=0}^{\infty} e_n z^n + \frac{c}{(1-z)^2} = \sum_{n=0}^{\infty} \sum_{j=0}^{n} r_k^\star(\mathcal{A}, j) \, z^n = \frac{1}{1-z} \sum_{n=0}^{\infty} r_k^\star(\mathcal{A}, n) \, z^n$$

where we have used the fact that $\sum_{n=0}^{\infty} (n+1) \, z^n = 1/(1-z)^2$. Applying Equation (7.13), it follows that

$$(1-z) \sum_{n=0}^{\infty} e_n \, z^n + \frac{c}{1-z} = \sum_{\mathbf{i} \in S(k)} \varepsilon_\star(\mathbf{i}) \, f_{\mathcal{A}}(z^{i_1}) \cdots f_{\mathcal{A}}(z^{i_m}). \tag{7.26}$$

Multiplying by the smoothing function $h_M(z)^2$, for some $M = M(r)$ to be determined later, taking absolute values and integrating along $\mathbb{S}_r$, we obtain

$$\begin{aligned}&\int_{\mathbb{S}_r} \Big|(1-z)h_M(z)^2 \sum_{n=0}^{\infty} e_n z^n\Big| d\mu + \int_{\mathbb{S}_r} \Big|\frac{c\, h_M(z)^2}{1-z}\Big| d\mu \\ \geq &\int_{\mathbb{S}_r} \Big| h_M(z)^2 \sum_{S(k)} \varepsilon_\star(\mathbf{i})\, f_{\mathcal{A}}(z^{i_1}) \cdots f_{\mathcal{A}}(z^{i_k})\Big| d\mu. \end{aligned} \tag{7.27}$$

Bounding the left-hand side. We note that $(1-z)\, h_M(z)^2 = (1-z^M)\, h_M(z)$ as well as $|h_M(z)| \leq M$ and $|1-z^M| \leq 2$. Applying Cauchy–Schwarz, we therefore get that the left-hand side of Equation (7.27) is at most

$$2 \Big(\int_{\mathbb{S}_r} |h_M(z)|^2 d\mu\Big)^{1/2} \Big(\int_{\mathbb{S}_r} \Big|\sum_{n=0}^{\infty} e_n\, z^n\Big|^2 d\mu\Big)^{1/2} + \int_{\mathbb{S}_r} \frac{cM^2}{|1-z|} d\mu.$$

Applying Equation (7.17) in Lemma 7.4 and then Lemma 7.10 to the first term as well as Lemma 7.9 to the second term, we can further bound the left-hand side of Equation (7.27) by

$$\begin{aligned}&2 \Big(\sum_{n=0}^{M-1} r^{2n}\Big)^{1/2} \Big(\sum_{n=0}^{\infty} |e_n|^2\, r^{2n}\Big)^{1/2} + M^2\, O_r\Big(-\log(1-r)\Big) \\ \leq\; & M^{1/2}\, o_r\bigg(\frac{-1}{(1-r)^{3/4} \log^{1/2}(1-r)}\bigg) + M^2\, O_r\Big(-\log(1-r)\Big). \end{aligned} \tag{7.28}$$

Bounding the right-hand side. Let us now observe that the right-hand side of Equation (7.27) is at least

$$\frac{1}{k!} \int_{\mathbb{S}_r} \Big| h_M(z)^2 f_{\mathcal{A}}^k(z)\Big| d\mu - \sum_{\mathbf{i} \in S_0(k)} \int_{\mathbb{S}_r} \Big| h_M(z)^2 \varepsilon_\star(\mathbf{i})\, f_{\mathcal{A}}(z^{i_1}) \cdots f_{\mathcal{A}}(z^{i_m})\Big| d\mu.$$

If M is a multiple of $\operatorname{lcm}\{1, 2, \ldots, k\}$, then we can use Cauchy–Schwarz and Proposition 7.7 to upper bound the individual summands of the second term (where $m < k$) by

$$\begin{aligned}&|\varepsilon_\star(\mathbf{i})| \Big(\int_{\mathbb{S}_r} h_M(z)^2 f_{\mathcal{A}}(z^{i_1})^m d\mu\Big)^{1/m} \cdots \Big(\int_{\mathbb{S}_r} h_M(z)^2 f_{\mathcal{A}}(z^{i_m})^m d\mu\Big)^{1/m} \\ =\, & o_r\Big(\int_{\mathbb{S}_r} \Big| h_M(z)^2 f_{\mathcal{A}}(z)^k \Big| d\mu\Big)\end{aligned}$$

as r tends to 1. It follows that the term with $m = k$ is the dominant one and the terms coming from $S_0(k)$ are negligible. Using Equation (7.18) in Lemma 7.4, we therefore know that the right-hand side of Equation (7.27) is at least

$$\left(\frac{1}{k!} + o_r(1)\right) \int_{\mathbb{S}_r} \left|h_M(z)^2 f_{\mathcal{A}}(z)^k\right| d\mu \geq \left(\frac{1}{k!} + o_r(1)\right) f_{\mathcal{A}}\Big(r^2\Big)^{k/2} h_M\Big(r^2\Big).$$

In order to estimate $f_{\mathcal{A}}(r^2)^{k/2}$, we first note that by Lemma 7.8 we have

$$\begin{aligned}\sum_{\mathbf{i} \in S(k)} \varepsilon_\star(\mathbf{i}) f_{\mathcal{A}}(r^{2i_1}) \cdots f_{\mathcal{A}}(r^{2i_m}) &= f_{\mathcal{A}}(r^2)^k + \sum_{\mathbf{i} \in S_0(k)} O_r\Big(f_{\mathcal{A}}(r^2)^m\Big) \\ &= (1 + o(1))\, f_{\mathcal{A}}(r^2)^k.\end{aligned}$$

We secondly note, using Lemma 7.10, that

$$\sum_{n=0}^{\infty} e_n r^{2n} \leq \sum_{n=0}^{\infty} (1 + e_n^2) r^{2n} \leq \frac{1}{1 - r^2} + o_r\left(\frac{-1}{(1-r)^{3/2} \log(1-r)}\right).$$

Substituting $z = r^2$ in Equation (7.26), noting that $(1 - r^2) = (2 + o_r(1))\,(1 - r)$ and using the two previous equations, it follows that

$$\begin{aligned}f_{\mathcal{A}}(r^2)^k &= (k! + o_r(1)) \left(\frac{c}{2(1-r)} + o_r\left(\frac{-1}{(1-r)^{1/2} \log(1-r)}\right)\right) \\ &= (k! + o_r(1))\, \frac{c}{2(1-r)}.\end{aligned}$$

Let us now choose

$$M = k! \left\lceil \varepsilon \, \frac{-\log^{-1}(1-r)}{(1-r)^{1/2}} \right\rceil \tag{7.29}$$

for some fixed $\varepsilon > 0$. This choice satisfies both $M(r) = \omega_r(1)$ and $r^{M(r)} = \Omega_r(1)$. Note also that $\operatorname{lcm}\{1, 2, \ldots k\}$ divides M as previously required. It follows that $h_M(r^2) \geq M/C_2$ for some $C_2 > 1$ and therefore the right-hand side of Equation (7.27) is at least

$$\left(\frac{M\sqrt{c}}{\sqrt{2k!}C_2} + o_r(1)\right) (1-r)^{-1/2} = M\, \Omega_r\Big((1-r)^{-1/2}\Big). \tag{7.30}$$

Obtaining the contradiction. Combining our bounds for the left- and right-hand sides of Equation (7.27), that is Equation (7.28) and Equation (7.30), we obtain

$$M\, \Omega_r\Big((1-r)^{-1/2}\Big) \leq M^{1/2}\, o_r\left(\frac{-\log^{-1/2}(1-r)}{(1-r)^{3/4}}\right) + M^2\, O_r\Big(-\log(1-r)\Big).$$

Inserting Equation (7.29), we therefore have

$$\varepsilon\,\Omega_r\left(\frac{-\log^{-1}(1-r)}{1-r}\right) \leq \varepsilon^{1/2} o_r\left(\frac{-\log^{-1}(1-r)}{1-r}\right) + \varepsilon^2 O_r\left(\frac{-\log^{-1}(1-r)}{1-r}\right),$$

with the constants in Ω_r, o_r and O_r independent of $\varepsilon > 0$. Therefore, we have that $C_3\varepsilon \leq o_r(1) + C_4\varepsilon^2$, for some $C_3, C_4 > 0$. This leads a contradiction taking any $\varepsilon < C_3/C_4$, so the assumption $e_n = o_n\Big(n^{1/4}\log^{-1/2}(n)\Big)$ was not possible. ■

7.3 Proof of Theorem 7.2 – The mean squared error

We will again need some general asymptotic bounds in order to prove Theorem 7.2. The first is Lemma 7.9, which we repeat here for the readers convenience.

Lemma 7.9. *We have*

$$\int_{\mathbb{S}_r} \frac{d\mu}{|1-z|} = O_r\Big(-\log(1-r)\Big).$$

Lemma 7.11. *For any $D > 0$ and $k \geq 1$, we have*

$$\sum_{s=1}^{\infty} r^{Ds^k} = \Omega_r\left((1-r)^{-1/k}\right).$$

Proof. Since $0 < r < 1$, we note that the function $f(x) = r^{Dx^k}$ is strictly decreasing in $[0,\infty)$. It follows, that we may lower bound the series $\sum_{s=1}^{\infty} r^{Ds^k}$ by its corresponding integral, that is

$$\sum_{s=1}^{\infty} r^{Ds^k} \geq \int_1^{\infty} r^{Dx^k} dx = \int_1^{\infty} e^{-D|\log r| x^k} dx = \frac{\int_1^{\infty} e^{-y^k} dy}{D^{1/k}|\log r|^{1/k}} = \Omega_r\left((1-r)^{-1/k}\right),$$

where we have used that $\int_1^{\infty} e^{-y^k} dy < \infty$. □

We are now ready to prove Theorem 7.2.

Proof of Theorem 7.2. Throughout this proof $k \geq 2$, $c \geq 0$, $\star \in \{\leq, <\}$ and $\mathcal{A} \subseteq \mathbb{N}_0$ are fixed. Furthermore, z will always lie in $\mathbb{D}$ and r in the open interval $(1/2, 1)$. We start by noting that if c is not an integer, then the statement immediately follows since

$$\Big(r_k^\star(\mathcal{A}, n) - c\Big)^2 \geq \max\{\Big(c - \lfloor c \rfloor\Big)^2, \Big(\lceil c \rceil - c\Big)^2\} > 0.$$

We can therefore assume that $c \in \mathbb{N}_0$. We now prove that we can assume that there exist $D > 0$ such that $a_s < Ds^k$. If $c = 0$ this is given by the statement of the theorem, so let us consider $c \geq 1$. Using the fact that c and $r_k^\star(\mathcal{A}, n)$ are integers, we have

$$\begin{aligned} nE_{k,c}^\star(\mathcal{A}, n) &= \sum_{j=0}^{n} \Big(r_k^\star(\mathcal{A}, j) - c \Big)^2 \\ &\geq \sum_{j=0}^{n} |r_k^\star(\mathcal{A}, j) - c| \geq \Big| c(n+1) - \sum_{j=0}^{n} r_k^\star(\mathcal{A}, j) \Big|. \end{aligned} \tag{7.31}$$

Since $a_{i_1} + a_{i_2} + \ldots + a_{i_k} \leq a_s$ trivially implies that every i_j is at most s for any $s, i_1, \ldots, i_k \in \mathbb{N}$, it follows that $\sum_{j=0}^{a_s} r_k^\star(\mathcal{A}, j) \leq s^k$. Taking $n = a_s$ in Equation (7.31), we therefore obtain

$$E_{k,c}^\star(\mathcal{A}, a_s) \geq \frac{1}{a_s}\Big(ca_s + c - s^k \Big) = c + \frac{c}{a_s} - \frac{s^k}{a_s}.$$

Either the desired statement holds, or $\limsup_{s\to\infty} E_{k,c}^\star(\mathcal{A}, a_s) = 0$ implying that

$$\limsup_{s\to\infty} \left(c - \frac{s^k}{a_s} \right) \leq 0.$$

It follows that we can assume $a_s \leq Ds^k$ for some appropriate $D > 0$. By Equation (7.17) in Lemma 7.4 as well as Equation (7.13), we now have

$$\begin{aligned} \left(\sum_{n=0}^{\infty} \Big(r_k^\star(\mathcal{A}, n) - c \Big)^2 r^{2n} \right)^{1/2} &= \left(\int_{\mathbb{S}_r} \Big| \sum_{n=0}^{\infty} r_k^\star(\mathcal{A}, n)\, z^n - \frac{c}{1-z} \Big|^2 d\mu \right)^{1/2} \\ &\geq \int_{\mathbb{S}_r} \Big| \sum_{\mathbf{i} \in S(k)} \varepsilon_*(\mathbf{i}) f_{\mathcal{A}}(z^{i_1}) \cdots f_{\mathcal{A}}(z^{i_m}) - \frac{c}{1-z} \Big| d\mu. \end{aligned}$$

Note that the terms with $\mathbf{i} \in S_0(k)$ are negligible. Using Cauchy-Schwarz and Proposition 7.7, we have that

$$\begin{aligned} &\int_{\mathbb{S}_r} \Big| \sum_{\mathbf{i} \in S_0(k)} \varepsilon_*(\mathbf{i}) f_{\mathcal{A}}(z^{i_1}) \cdots f_{\mathcal{A}}(z^{i_m}) \Big| d\mu \\ &\leq \sum_{\mathbf{i} \in S_0(k)} \left(\int_{\mathbb{S}_r} \Big| f_{\mathcal{A}}(z^{i_1}) \Big|^m \right)^{1/m} \cdots \left(\int_{\mathbb{S}_r} \Big| f_{\mathcal{A}}(z^{i_m}) \Big|^m \right)^{1/m} \\ &= o_r \left(\int_{\mathbb{S}_r} \Big| f_{\mathcal{A}}(z)^k \Big| d\mu \right). \end{aligned}$$

Now, Equation (7.18) in Lemma 7.4, gives us

$$\begin{aligned}\int_{\mathbb{S}_r} \Big| \sum_{\mathbf{i} \in S(k)} \varepsilon_*(\mathbf{i}) f_{\mathcal{A}}(z^{i_1}) \cdots f_{\mathcal{A}}(z^{i_m}) \Big| d\mu &\geq \Big(\frac{1}{k!} + o_r(1) \Big) \int_{\mathbb{S}_r} \Big| f_{\mathcal{A}}(z)^k \Big| d\mu \\ &\geq \Big(\frac{1}{k!} + o_r(1) \Big) f_{\mathcal{A}}(r^2)^{k/2},\end{aligned}$$

so that by Lemma 7.9

$$\begin{aligned}&\int_{\mathbb{S}_r} \Big| \sum_{\mathbf{i} \in S(k)} \varepsilon_*(\mathbf{i}) f_{\mathcal{A}}(z^{i_1}) \cdots f_{\mathcal{A}}(z^{i_m}) - \frac{c}{1-z} \Big| d\mu \\ &\geq \Big(\frac{1}{k!} + o_r(1) \Big) \, f_{\mathcal{A}}\Big(r^2\Big)^{k/2} - O_r\Big(- \log(1-r) \Big).\end{aligned}$$

Now, taking into account that $a_s \leq D s^k$ and using Lemma 7.11, we have

$$f_{\mathcal{A}}(r^2)^{k/2} = \left(\sum_{s=1}^{\infty} r^{2a_s} \right)^{k/2} \geq \left(\sum_{s=1}^{\infty} r^{2Ds^k} \right)^{k/2} = \Omega_r\Big((1-r)^{-1/2} \Big) .$$

Collecting all the bounds, it follows that

$$\sum_{n=0}^{\infty} \Big(r_k^{\star}(\mathcal{A}, n) - c \Big)^2 r^{2n} = \Omega_r\Big((1-r)^{-1/2} \Big) .$$

Therefore,

$$\begin{aligned}\sum_{n=0}^{\infty} n E_{k,c}^{\star}(\mathcal{A}, n) r^{2n} &= \frac{1}{1-r^2} \sum_{n=0}^{\infty} \Big(r_k^{\star}(\mathcal{A}, n) - c \Big)^2 r^{2n} \\ &= \frac{1}{1-r^2} \, \Omega_r \left(\frac{1}{1-r^2} \right) \geq C \sum_{n=0}^{\infty} n r^{2n}\end{aligned}$$

for some appropriate constant $C > 0$. It follows that infinitely many of the coefficients $n E_{k,c}^{\star}(\mathcal{A}, n)$ must be greater than $Cn/2$ and hence $\limsup_{n \to \infty} E_{k,c}^{\star}(\mathcal{A}, n) \geq C/2 > 0$ as desired. ■

7.4 Preliminaries for Theorem 7.3

In this part we will establish some preliminaries that will be required for the proof of Theorem 7.3. The first will again be the issue of encoding the problem using the generating functions of the sets. Once we have established an equivalent to Equation (7.5) for weighted representation functions, we will then introduce cyclotomic polynomials, which will play an essential role in the proof of Theorem 7.3.

7.4.1 Encoding the weighted representation functions

For each $1 \leq i \leq d$, we define the combinatorial class $(\mathcal{A}, |\cdot|_i)$ where the size is simply $|a|_i = k_i a$. Note that ordinary generating function of this class is $f_{\mathcal{A}}(z^{k_i})$. The combinatorial class we are then interested in enumerating is that of $(\mathcal{A}^d, |\cdot|)$ where the size of an element in $\mathcal{A}^d$ is $|(a_1, \ldots, a_d)| = |a_1|_1 + \ldots + |a_d|_d = k_1 a_1 + \ldots + k_d a_d$. The ordinary generating function of this class is therefore $\sum_{n=0}^{\infty} r(\mathcal{A}, n; k_1, \ldots, k_d) z^n$. Since $(\mathcal{A}^d, |\cdot|)$ is the cartesian product of $(\mathcal{A}, |\cdot|_1)$, ..., $(\mathcal{A}, |\cdot|_d)$, it follows that

$$\sum_{n=0}^{\infty} r(\mathcal{A}, n; k_1, \ldots, k_d)\, z^n = f_{\mathcal{A}}(z^{k_1}) \cdots f_{\mathcal{A}}(z^{k_d}). \tag{7.32}$$

This is the equivalent to Equation (7.5) for weighted representation functions. If $r(\mathcal{A}, n; k_1, \ldots, k_d)$ becomes constant for n large enough, that is $r(\mathcal{A}, n; k_1, \ldots, k_d) = c$ for any $n \geq n_0$ for some $n_0 \in \mathbb{N}_0$ and $c > 0$, then

$$\sum_{n=0}^{\infty} r(\mathcal{A}, n; k_1, \ldots, k_d)\, z^n = \sum_{n=0}^{n_0 - 1} r(\mathcal{A}, n; k_1, \ldots, k_d)\, z^n + \frac{c z^{n_0}}{1 - z}. \tag{7.33}$$

Writing $P(z) = (1 - z) \sum_{n=0}^{n_0 - 1} r(\mathcal{A}, n; k_1, \ldots, k_d)\, z^n + c z^{n_0}$, this would imply that

$$f_{\mathcal{A}}(z^{k_1}) \cdots f_{\mathcal{A}}(z^{k_d}) = \frac{P(z)}{1 - z}. \tag{7.34}$$

To simplify notation, we will generally consider the d-th power of this equation, that is for $F_{\mathcal{A}}(z) = f_{\mathcal{A}}^d(z)$ we have

$$F_{\mathcal{A}}(z^{k_1}) \cdots F_{\mathcal{A}}(z^{k_d}) = \frac{P^d(z)}{(1 - z)^d}. \tag{7.35}$$

This is the starting point of the proof of Theorem 7.3. We note that, since $f_{\mathcal{A}}(z)$ is a formal power series with coefficients in $\{0, 1\}$ that is analytic in the open complex disc

$$\mathcal{D} = \{z \in \mathbb{C} : |z| < 1\}, \tag{7.36}$$

$F_{\mathcal{A}}(z)$ is likewise a formal power series with positive coefficients that is analytic in $\mathcal{D}$. We also note that P is a polynomial with integer coefficients satisfying $P(1) = c \neq 0$.

7.4.2 Cyclotomic polynomials

The proof of Theorem 7.3 relies on studying the behavior of either side of Equation (7.35) when factoring out certain functions, called cyclotomic polynomials. The cyclotomic polynomial of order n is defined as

$$\Phi_n(z) = \prod_{\xi \in \phi_n} (z - \xi) \in \mathbb{Z}[z] \tag{7.37}$$

where

$$\begin{aligned} \phi_n &= \left\{ e^{\frac{2\pi i \ell}{n}} : 0 \leq \ell < n \text{ satisfying } (\ell, n) = 1 \right\} \\ &= \left\{ \xi \in \mathbb{C} : \xi^k = 1 \text{ iff } k \equiv 0 \mod n \right\} \end{aligned} \tag{7.38}$$

denotes the set of primitive roots of unity of order $n \in \mathbb{N}$. It is well known that $\Phi_n(z) \in \mathbb{Z}[z]$, that is it has integer coefficients. Cyclotomic polynomials have the property of being irreducible over $\mathbb{Z}[z]$ and therefore it follows that for any polynomial $P(z) \in \mathbb{Z}[z]$ and $n \in \mathbb{N}$ there exists an integer $s_n \in \mathbb{N}_0$ such that

$$P_n(z) = P(z)\, \Phi_n^{-s_n}(z) \tag{7.39}$$

is a polynomial in $\mathbb{Z}[z]$ satisfying $P_n(\xi) \neq 0$ for all $\xi \in \phi_n$. We will say that we have factored $\Phi_n(z)$ out of $P(z)$ with multiplicity s_n. Note that the multiplicity is trivially unique. The following lemma illustrates this concept and will be needed in the next section.

Lemma 7.12. *Given $k, n \in \mathbb{N}$ such that $k \mid n$, we have $\phi_{n/k} = \{\xi^k : \xi \in \phi_n\}$. Furthermore, we can factor $\Phi_n(z)$ out of $\Phi_{n/k}(z^k)$ with multiplicity* 1.

Proof. To see equality between the two sets, observe that

$$\begin{aligned} \left\{ \xi^k : \xi \in \phi_n \right\} &= \left\{ \xi^k : \xi^\ell = 1 \text{ iff } \ell \equiv 0 \mod n \right\} \\ &= \left\{ \xi^k : (\xi^k)^{\ell/k} = 1 \text{ iff } \ell \equiv 0 \mod n \right\} \\ &= \left\{ \xi^k : (\xi^k)^{\ell} = 1 \text{ iff } \ell \equiv 0 \mod n/k \right\} = \phi_{n/k}. \end{aligned}$$

As $\Phi_{n/k}(z^k)$ is a polynomial in $\mathbb{Z}[z]$ and $\Phi_{n/k}(\xi^k) = 0$ for any $\xi \in \phi_n$ via the previous observation, it follows that we can factor out $\Phi_n(z)$. The multiplicity is equal to 1 since all roots of $\Phi_{n/k}(z^k)$ are simple. □

Unfortunately, we are not guaranteed to be able to factor cyclotomic polynomials out of arbitrary non-polynomial functions. In particular, our function $F_{\mathcal{A}}(z)$ is not even analytic at roots of unity and it can also be shown that even the radial limit of $F_{\mathcal{A}}(z)$, where z approaches some root of unit ξ radially from within $\mathcal{D}$, may not exist in general. However, we can extend this notion in a natural way that will be applicable to our function $F_{\mathcal{A}}(z)$.

Definition 7.13. *We say that a function $g : [0,1) \to [0,\infty)$ is* restricted *if*

$$\limsup_{z \to 1^-} g(z) \neq \infty \quad \text{and} \quad \liminf_{z \to 1^-} g(z) \neq 0. \tag{7.40}$$

Given $n \in \mathbb{N}$ and some function $F(z)$ analytic in $\mathcal{D}$, we say that we can factor *$\Phi_n(z)$* out of *$F(z)$* with multiplicity *r_n if $|F(z\,\xi)\,\Phi_n^{-r_n}(z\,\xi)|$ is restricted for any $\xi \in \phi_n$.*

Here $\limsup_{z\to 1^-}$ and $\liminf_{z\to 1^-}$ refer to the left-hand limits. In fact, throughout the rest of this chapter we will always assume that $z \in [0,1)$. Note that, by continuity, this notion is a true extension of the previous one for polynomials. It is also again easy to verify that the multiplicity, if it exists, is uniquely determined.

Lemma 7.14. *If we can factor $\Phi_n(z)$ out of $F(z)$, then the multiplicity is uniquely determined.*

Proof. Assume that we can factor $\Phi_n(z)$ out of $F(z)$ with multiplicity r_n. For $z \in [0,1)$, $\xi \in \phi_n$ and $\alpha \neq 0$ we have

$$|F(z\,\xi)\,\Phi_n^{-r_n+\alpha}(z\,\xi)| = |F(z\,\xi)\,\Phi_n^{-r_n}(z\,\xi)|\,|\Phi_n^{\alpha}(z\,\xi)|. \tag{7.41}$$

As $\Phi_n(\xi) = 0$, $\limsup_{z\to 1^-} |F(z\,\xi)\,\Phi_n^{-r_n}(z\,\xi)| \neq \infty$ and $\liminf_{z\to 1^-} |F(z\,\xi)\,\Phi_n^{-r_n}(z\,\xi)| \neq 0$, the lim sup of the right-hand side of Equation (7.41) must tend to ∞ if $\alpha < 0$ and the lim inf to 0 if $\alpha > 0$. It follows we cannot also factor $\Phi_n(z)$ out of $F(z)$ with multiplicity $r_n + \alpha$ if $\alpha \neq 0$, so the multiplicity is uniquely determined. □

Note that, when dealing with lim sup or lim inf, we are not guaranteed to have a product rule. However, the following lemma will establish a version of such a rule for the specific case and particular type of function that will be needed for the proof of Theorem 7.3. Here it is crucial that the functions we are studying are formal power series with positive coefficients.

Lemma 7.15. *Let $k_1, \ldots, k_d$ be positive integers and F a formal power series with positive coefficients that is analytic in $\mathcal{D}$. Writing $F_{\mathbf{0}}(z) = F(z)(1-z)$, if $|F_{\mathbf{0}}(z)|$ is not restricted, then neither is $|F_{\mathbf{0}}(z^{k_1})| \cdots |F_{\mathbf{0}}(z^{k_d})|$.*

Proof. We note that for $\alpha > 1$ and $z \in [0,1)$ we have

$$\begin{aligned}\Big|F_{\mathbf{0}}(z) - F_{\mathbf{0}}(z^\alpha)\Big| &= \Big|\sum_n a_n\, z^n\,(1-z) \;-\; \sum_n a_n\, z^n\,(1-z^\alpha)\Big| \\ &\leq \sum_n \Big|a_n\, z^n\,(1-z)\Big|\, \Big|1 - z^{(\alpha-1)n}\, \frac{1-z^\alpha}{1-z}\Big| \leq \alpha\, |F_{\mathbf{0}}(z)| \qquad (7.42)\end{aligned}$$

where we have used that $0 \leq z^{(\alpha-1)n} < 1$ and $0 \leq (1-z^\alpha)/(1-z) \leq \alpha$ since $0 \leq z < 1$ and $\alpha > 1$.

If there exists a sequence $(z_k)_{k\in\mathbb{N}}$ tending to 1 such that $\lim_{k\to\infty} |F_{\mathbf{0}}(z_k)| = 0$, then by Equation (7.42) it would follow that $\lim_{k\to\infty} |F_{\mathbf{0}}(z_k^{k_i})| = 0$ for any $1 \leq i \leq d$ and hence $|F_{\mathbf{0}}(z^{k_1})| \cdots |F_{\mathbf{0}}(z^{k_d})|$ would not be restricted. Assume therefore that there exists some sequence $(z_k)_{k\in\mathbb{N}}$ tending to 1 such that $\lim_{k\to\infty} |F_{\mathbf{0}}(z_k)| = \infty$, but that $\liminf_{z\to 1^-} |F_{\mathbf{0}}(z)| \neq 0$ and $|F_{\mathbf{0}}(z^{k_1})| \cdots |F_{\mathbf{0}}(z^{k_d})|$ is still restricted. Considering the sequence given by $y_k = z_k^{1/k_1}$ for $k \in \mathbb{N}$, it follows that $\lim_{k\to\infty} |F_{\mathbf{0}}(y_k^{k_1})| = \lim_{k\to\infty} |F_{\mathbf{0}}(z_k)| = \infty$. Since we are assuming that $|F_{\mathbf{0}}(z^{k_1})| \cdots |F_{\mathbf{0}}(z^{k_d})|$ is restricted, it follows that we must have that $\lim_{k\to\infty} |F_{\mathbf{0}}(y_k^{k_i})| = 0$ for some $2 \leq i \leq d$, contradicting the fact that $\liminf_{z\to 1^-} |F_{\mathbf{0}}(z)| \neq 0$. □

7.5 Proof of Theorem 7.3 – A question of Sárközy and Sós

Let us introduce some short-hand notation that will be needed in this part. If $q_1, \ldots, q_m$ are fixed co-prime integers as given by Theorem 7.3 and $\mathbf{j} = (j_1, \ldots, j_m) \in \mathbb{N}_0^m$, then we write

$$\Phi_{\mathbf{j}}(z) = \Phi_{q_1^{j_1} \cdots q_m^{j_m}}(z), \quad \phi_{\mathbf{j}} = \phi_{q_1^{j_1} \cdots q_m^{j_m}}, \quad s_{\mathbf{j}} = s_{q_1^{j_1} \cdots q_m^{j_m}} \quad \text{and} \quad r_{\mathbf{j}} = r_{q_1^{j_1} \cdots q_m^{j_m}}.$$

The main strategy of the proof will be to show that for any $\mathbf{j} \in \mathbb{N}_0^m$ we can factor $\Phi_{\mathbf{j}}(z)$ out of our hypothetical function $F_{\mathcal{A}}(z) = f_{\mathcal{A}}^d(z)$ satisfying Equation (7.35) and that the multiplicities $r_{\mathbf{j}}$ have to fulfill certain relations between themselves. The goal will be to find a contradiction in these relations, negating the possibility of such a function and therefore such a set $\mathcal{A}$ existing in the first place.

7.5.1 The recurrence relations

We can now give the statement and proof establishing that we can factor any $\Phi_{\mathbf{j}}(z)$ out of $F_{\mathcal{A}}(z)$ and that the multiplicities satisfy certain relations. We will in fact state this for any $k_1, \ldots, k_d \geq 2$ and later derive a contradiction from these relations in the specific case stated in Theorem 7.3.

For any $a, b \in \mathbb{N}_0$, $\mathbf{j} = (j_1, \ldots, j_m) \in \mathbb{N}_0^m$ and $\mathbf{b} = (b_1, \ldots, b_m) \in \mathbb{N}_0^m$, we will use the notation

$$a \ominus b = \max\{a - b, 0\} \quad \text{and} \quad \mathbf{j} \ominus \mathbf{b} = (j_1 \ominus b_1, \ldots, j_m \ominus b_m).$$

Let us state the central proposition of this section.

Proposition 7.16. *Let $q_1, \ldots, q_m \geq 2$ be pairwise co-prime integers and*

$$k_i = q_1^{b(i,1)} \cdots q_m^{b(i,m)}$$

for $1 \leq i \leq d$ where $b(i,j) \in \mathbb{N}_0$. Furthermore, let $P(z) \in \mathbb{Z}[z]$ be a polynomial satisfying $P(1) \neq 0$ and $F(z)$ a formal power series with positive coefficients that is analytic in $\mathcal{D}$ such that

$$F(z^{k_1}) \cdots F(z^{k_d}) = \frac{P^d(z)}{(1-z)^d}. \tag{7.43}$$

Then for all $\mathbf{j} \in \mathbb{N}_0^m$ there exist $r_{\mathbf{j}} \in \mathbb{N}_0$ so that we can factor $\Phi_{\mathbf{j}}$ out of F with multiplicity $r_{\mathbf{j}}$. Moreover, writing $\mathbf{b}_i = (b(i,1), \ldots, b(i,m))$ for $1 \leq i \leq m$ as well as $s_{\mathbf{j}} \in \mathbb{N}_0$ for the integer satisfying $P(\xi)\,\Phi_{\mathbf{j}}^{-s_{\mathbf{j}}}(\xi) \neq 0$ for any $\xi \in \phi_{\mathbf{j}}$, these multiplicities satisfy the relations

$$r_{\mathbf{0}} = -1 \quad \textit{and} \quad r_{\mathbf{j} \ominus \mathbf{b}_1} + \cdots + r_{\mathbf{j} \ominus \mathbf{b}_d} = d s_{\mathbf{j}} \quad \text{for all } \mathbf{j} \in \mathbb{N}_0^m \setminus \{\mathbf{0}\} \tag{7.44}$$

and we have $r_{\mathbf{i}} \equiv -1 \mod d$ for all $\mathbf{i} \in \mathbb{N}_0^m$.

Proof. We start by assuming that the set of multiplicities $\{r_{\mathbf{j}} : \mathbf{j} \in \mathbb{N}_0^m\}$ exists and show that the relations given by Equation (7.44) must be satisfied. After this, we will show that there is a way to recursively determine the values $\{r_{\mathbf{j}} : \mathbf{j} \in \mathbb{N}_0^m\}$, proving their existence.

Let us start with $r_{\mathbf{0}} = -1$. For $F_{\mathbf{0}}(z) = F(z)(1-z)$ we wish to show that $|F_{\mathbf{0}}(z)|$, as a function with domain $[0,1)$, is restricted. Inserting the equality $F(z) = (1-z)^{-1} F_{\mathbf{0}}(z)$

into Equation (7.43), we get that $F_{\mathbf{0}}(z)$ satisfies

$$\prod_{\ell=1}^{d} F_{\mathbf{0}}(z^{k_\ell})\,(1-z^{k_\ell})^{-1} = \frac{P^d(z)}{(1-z)^d}.$$

Taking absolut values and observing that $(1-z^{k_\ell})/(1-z) = (1+z+\cdots+z^{k_\ell-1})$, it follows that

$$|F_{\mathbf{0}}(z^{k_1})| \cdot |F_{\mathbf{0}}(z^{k_2})| \cdots |F_{\mathbf{0}}(z^{k_d})| = |P^d(z)| \cdot \prod_{\ell=1}^{d} \left|(1+z+\cdots+z^{k_\ell-1})\right|.$$

Since $P^d(1) \neq 0$ as well as $(1+1+\cdots+1^{k_\ell-1}) = k_\ell \neq 0$ for $1 \leq \ell \leq d$, it follows that

$$\limsup_{z\to 1^-} |F_{\mathbf{0}}(z^{k_1})| \cdots |F_{\mathbf{0}}(z^{k_d})| \neq \infty$$

and $\liminf_{z\to 1^-} |F_{\mathbf{0}}(z^{k_1})| \cdots |F_{\mathbf{0}}(z^{k_d})| \neq 0$ and hence by Lemma 7.15 it follows that $|F_{\mathbf{0}}(z)|$ must be restricted.

Next, let us show that if for a given $\mathbf{j} \in \mathbb{N}_0^m \setminus \{\mathbf{0}\}$ the values $r_{\mathbf{j}\ominus\mathbf{b}_1}, \ldots, r_{\mathbf{j}\ominus\mathbf{b}_d}$ exist, then they must satisfy the relation given by Equation (7.44). For $1 \leq i \leq d$ let

$$F_{\mathbf{j}\ominus\mathbf{b}_i} = F(z)\,\Phi_{\mathbf{j}\ominus\mathbf{b}_i}^{-r_{\mathbf{j}\ominus\mathbf{b}_i}}$$

and rewrite Equation (7.43) as

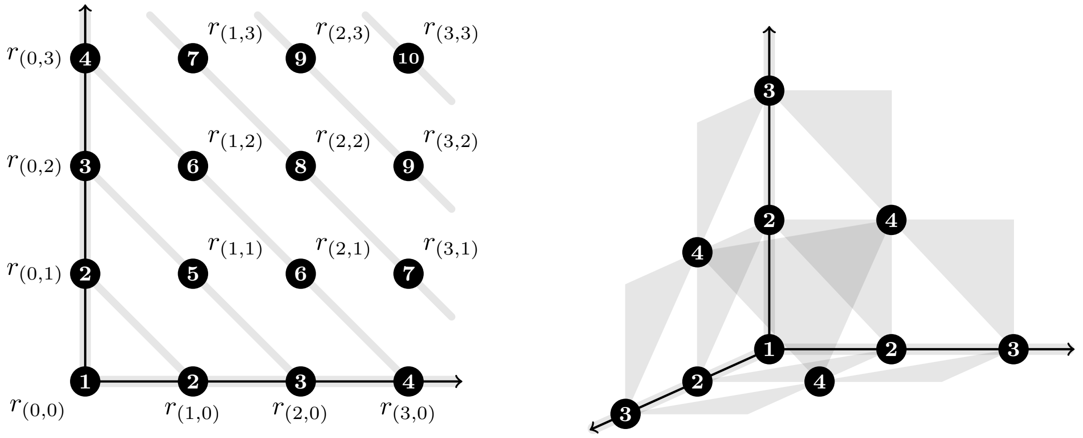
$$\Phi_{\mathbf{j}\ominus\mathbf{b}_1}^{r_{\mathbf{j}\ominus\mathbf{b}_1}}(z^{k_1})\,F_{\mathbf{j}\ominus\mathbf{b}_1}(z^{k_1}) \;\cdots\; \Phi_{\mathbf{j}\ominus\mathbf{b}_d}^{r_{\mathbf{j}\ominus\mathbf{b}_d}}(z^{k_d})\,F_{\mathbf{j}\ominus\mathbf{b}_d}(z^{k_d}) = \frac{\Phi_{\mathbf{j}}^{ds_{\mathbf{j}}}(z)\,P_{\mathbf{j}}^d(z)}{(1-z)^d}. \tag{7.45}$$

Writing $R_{\mathbf{j},i}(z) = \Phi_{\mathbf{j}\ominus\mathbf{b}_i}(z^{k_i})\,\Phi_{\mathbf{j}}^{-1}(z)$ and taking absolute values, it follows that

$$\begin{aligned} &\left|\Phi_{\mathbf{j}}^{r_{\mathbf{j}\ominus\mathbf{b}_1}+\cdots+r_{\mathbf{j}\ominus\mathbf{b}_d}-ds_{\mathbf{j}}}(z)\right| \\ &= \frac{|P_{\mathbf{j}}^d(z)|}{|(1-z)^d|}\left(|R_{\mathbf{j},1}^{r_{\mathbf{j}\ominus\mathbf{b}_1}}(z)|\,|F_{\mathbf{j}\ominus\mathbf{b}_1}(z^{k_1})| \;\cdots\; |R_{\mathbf{j},d}^{r_{\mathbf{j}\ominus\mathbf{b}_d}}(z)|\,|F_{\mathbf{j}\ominus\mathbf{b}_d}(z^{k_d})|\right). \end{aligned} \tag{7.46}$$

We observe that, by assumption as well as Lemma 7.12, if we substitute $z\,\xi$ into Equation (7.46) where $\xi \in \phi_{\mathbf{j}}$ and $z \in [0,1)$, then the involved factors on the right-hand side are restricted by assumption. As $\Phi_{\mathbf{j}}(\xi) = 0$, it follows that the exponent on the left-hand side must be 0, that is the desired relation must hold.

It remains to be shown that the values $r_{\mathbf{j}}$ actually exist for any $\mathbf{j} \in \mathbb{N}_0$. We will do



(a) *$b_1 = (1, 0)$ and $b_2 = (0, 1)$.*

(b) *$b_1 = (1, 0, 0)$, $b_2 = (0, 1, 0)$ and $b_3 = (0, 0, 1)$.*

Figure 7.1: *Illustrating how the values $r_{\mathbf{j}}$ recursively define each other through Equation (7.47). The numbers indicate the order in which the illustrated points are established.*

so recursively with the base case of $r_{\mathbf{0}} = -1$ already having been established. From now on, let us – for simplicities sake – redefine the value $s_{\mathbf{0}}$ (which previously was 0 as $P(0) \neq 0$) to be $s_{\mathbf{0}} = -1$, so that the initial relation $r_{\mathbf{0}} = -1$ is now included in the general relation for the case $\mathbf{j} = \mathbf{0}$. Assume that for some $1 \leq \ell \leq d$ all values $r_{\mathbf{j} \ominus \mathbf{b}_1}, \ldots, r_{\mathbf{j} \ominus \mathbf{b}_d}$ except for $r_{\mathbf{j} \ominus \mathbf{b}_\ell}$ have already been shown to exist. Then, through the already established Equation (7.46), it is clear that setting

$$r_{\mathbf{j} \ominus \mathbf{b}_\ell} = d s_{\mathbf{j}} - \sum_{i \neq \ell} r_{\mathbf{j} \ominus \mathbf{b}_i} \tag{7.47}$$

would give the desired property, that is $|F(z\,\xi)\,\Phi_{\mathbf{j} \ominus \mathbf{b}_\ell}^{-r_{\mathbf{j} \ominus \mathbf{b}_\ell}}(z\,\xi)|$ is restricted for any $\xi \in \phi_{\mathbf{j} \ominus \mathbf{b}_\ell}$. We therefore wish to show inductively that for all $\mathbf{i} \in \mathbb{N}_0^m$ there exists $\mathbf{j} \in \mathbb{N}_0^m$ and $1 \leq \ell \leq d$ such that $\mathbf{i} = \mathbf{j} \ominus \mathbf{b}_\ell$ and all other involved values $\mathbf{j} \ominus \mathbf{b}_1, \ldots, \mathbf{j} \ominus \mathbf{b}_{\ell-1}, \mathbf{j} \ominus \mathbf{b}_{\ell+1}, \ldots, \mathbf{j} \ominus \mathbf{b}_d$ have already been determined by the inductive hypothesis.

For this we will give the indices $\mathbf{j} \in \mathbb{N}_0^m$ inducing these relations an appropriate ordering. More precisely, for each $\mathbf{j} = (j_1, \ldots, j_m) \in \mathbb{N}_0^m$ let $\mathbf{j}^{\leq} = (j_1^{\leq}, \ldots, j_m^{\leq})$ denote the ordered version, that is $j_1^{\leq} \leq j_2^{\leq} \leq \cdots \leq j_m^{\leq}$ and there exists some permutation σ on m letters such that $\mathbf{j} = \left(j_{\sigma(1)}^{\leq}, \ldots, j_{\sigma(m)}^{\leq}\right)$. Consider the ordering on $\mathbb{N}_0^m$ given by $\mathbf{j} \prec \mathbf{j}'$ if $\mathbf{j}^{\leq}$ lexicographically comes before $\mathbf{j}'^{\leq}$. In this situation, ties are broken arbitrarily. We want to show that going through the indices $\mathbf{j}$ in that order and considering the relation $r_{\mathbf{j} \ominus \mathbf{b}_1} + \cdots + r_{\mathbf{j} \ominus \mathbf{b}_d} = d s_{\mathbf{j}}$, at most one of the $r_{\mathbf{j} \ominus \mathbf{b}_\ell}$ will not have occurred in any of the previous relations given by some $\mathbf{j}' \prec \mathbf{j}$. This is illustrated for some initial values of two different structures in Figure 7.1.

Assume to the contrary that there exist $\mathbf{i} \neq \mathbf{i}' \in \mathbb{N}_0^m$ such that, for both of them, $\mathbf{j} \in \mathbb{N}_0^m$ is the first index for which there exist $1 \leq \ell, \ell' \leq d$ satisfying $\mathbf{i} = \mathbf{j} \ominus \mathbf{b}_\ell$ and $\mathbf{i}' = \mathbf{j} \ominus \mathbf{b}_{\ell'}$. Note that $\mathbf{b}_\ell \neq \mathbf{b}_{\ell'}$ and therefore at least one of the two statements $\mathbf{j} \ominus (\mathbf{b}_\ell - \mathbf{b}_{\ell'}) \prec \mathbf{j}$ and $\mathbf{j} \ominus (\mathbf{b}_{\ell'} - \mathbf{b}_\ell) \prec \mathbf{j}$ must hold. To see this, assume without loss of generality that $\mathbf{j} = (j_1, \ldots, j_m)$ is already in ordered form. Note that $\mathbf{b}_\ell - \mathbf{b}_{\ell'} \neq \mathbf{0}$ as $\mathbf{i} \neq \mathbf{i}'$. Writing $\mathbf{b}_\ell = (b_1, \ldots, b_m)$ and $\mathbf{b}_{\ell'} = (b'_1, \ldots, b'_m)$, and letting $1 \leq i \leq m$ be the first index such that $b_i \neq b'_i$ and $j_i > 0$, then we clearly have that either

$$j_i \ominus (b_i - b_{i'}) = \max\{j_i - (b_i - b_{i'}), 0\} < j_i$$

or

$$j_i \ominus (b_{i'} - b_i) = \max\{j_i + (b_i - b_{i'}), 0\} < j_i,$$

meaning that at least one of the two values $\mathbf{j} \ominus (\mathbf{b}_\ell - \mathbf{b}_{\ell'})$ and $\mathbf{j} \ominus (\mathbf{b}_{\ell'} - \mathbf{b}_\ell)$ must lexicographically come before $\mathbf{j}$. Note that such index i must exist since if $j_i = 0$ whenever $b_i - b'_i \neq 0$ then we would have had $\mathbf{i} = \mathbf{j} \ominus \mathbf{b}_\ell = \mathbf{b}_{\ell'} = \mathbf{i}'$ in contradiction to our assumption that $\mathbf{i} \neq \mathbf{i}'$.

Assume now without loss of generality that $\mathbf{j} \ominus (\mathbf{b}_\ell - \mathbf{b}_{\ell'}) \prec \mathbf{j}$. Since for $a, b, c \geq 0$ we trivially have that $\max\{\max\{a - b + c, 0\} - c, 0\} = \max\{\max\{a - b, -c\}, 0\} = \max\{a - b, 0\}$, it follows that

$$\Big(\mathbf{j} \ominus (\mathbf{b}_\ell - \mathbf{b}_{\ell'})\Big) \ominus \mathbf{b}_{\ell'} = \mathbf{j} \ominus \mathbf{b}_\ell = \mathbf{i}.$$

This is however in contradiction to the requirement that $\mathbf{j}$ was the smallest index with respect to the ordering $\prec$ for which the relation given by Equation (7.44) involves $r_{\mathbf{i}}$, giving us the desired result.

Finally, note that from the previous argument it also inductively follows that $r_{\mathbf{i}} \equiv -1 \mod d$ for all $\mathbf{i} \in \mathbb{N}_0^m$ as in the base case we have that $r_{\mathbf{0}} = -1$. □

7.5.2 The contradiction

We will now use the proposition established in the previous section to prove Theorem 7.3 by contradiction. We start by introducing some necessary notation and definitions. For $1 \leq i \leq d$ we write $\mathbf{c}_i = (c(i,1), \ldots, c(i,m)) \in \mathbb{N}_0^m$ and for any $1 \leq \ell \leq m$

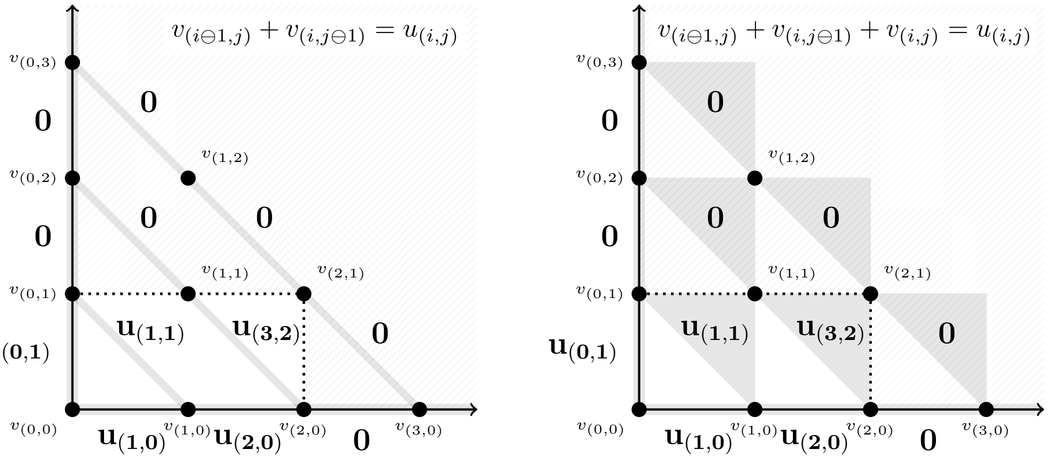
we use the notation

$$S_\ell = \{1 \le i \le d : c(i,\ell) = 0\} \quad \text{and} \quad S'_\ell = \{1,\dots,d\} \setminus S_\ell.$$

We will also use the following notation: for any $\mathbf{i} = (i_1,\dots,i_{m-1}) \in \mathbb{N}_0^{m-1}$ and $1 \le \ell \le m$ let

$$\Delta_{\mathbf{i},\ell} = v_{(i_1,\dots,i_{\ell-1},1,i_\ell,\dots,i_{m-1})} - v_{(i_1,\dots,i_{\ell-1},0,i_\ell,\dots,i_{m-1})}.$$

Finally, for $1 \le \ell \le m$, we write $\mathbb{1}_\ell \in \mathbb{N}_0^m$ for the vector whose entries are all equal to 0 except for the ℓ-th entry, which is equal to 1. using this notation we can define the notion of an m-structure, which will be illustrated in Figure 7.2.



(a) $\mathbf{c}_1 = (1,0)$ *and* $\mathbf{c}_2 = (0,1)$.

(b) $\mathbf{c}_1 = (1,0)$, $\mathbf{c}_2 = (0,1)$ *and* $\mathbf{c}_3 = (0,0)$.

Figure 7.2: *Illustrating the relations of type Equation* (7.48) *in two different* 2*-structures, both homogenous outside* $\mathbf{t} = (2,1)$ *but only the one on the left is regular.*

Definition 7.17. *For $m \ge 1$, we define an m-*structure *to be any set of values $\{v_\mathbf{j} \in \mathbb{Q} : \mathbf{j} \in \mathbb{N}_0^m\}$ for which there exist $\mathbf{c}_1,\dots,\mathbf{c}_d \in \mathbb{N}_0^m$ and $\{u_\mathbf{j} \in \mathbb{Z} : \mathbf{j} \in \mathbb{N}_0^m \setminus \{\mathbf{0}\}\}$ so that the values satisfy the relation*

$$v_{\mathbf{j}\ominus\mathbf{c}_1} + \dots + v_{\mathbf{j}\ominus\mathbf{c}_d} = u_\mathbf{j} \quad \textit{for all } \mathbf{j} \in \mathbb{N}_0^m \setminus \{\mathbf{0}\}. \tag{7.48}$$

Additionally, we define the following:

1. *An m-structure is* regular *if $\mathbf{c}_1,\dots,\mathbf{c}_d \in \{0,1\}^m \setminus \{\mathbf{0}\}$.*

2. *An m-structure is* **homogeneous outside** $\mathbf{t} = (t_1, \dots, t_m) \in \mathbb{N}_0^m$ *if* $u_{\mathbf{j}} = 0$ *for all* $\mathbf{j} \in \mathbb{N}_0^m \setminus [0, t_1] \times \cdots \times [0, t_m]$.

Note that by Equation (7.44), the values $r_{\mathbf{j}}$ together with $ds_{\mathbf{j}}$ and the vectors $\mathbf{b}_i$ form an m-structure. Furthermore, the restrictions stated in Theorem 7.3 imply that this structure is regular. Lastly, since P is a polynomial, only a finite number of the values $s_{\mathbf{j}}$ are non-zero, so the structure is homogeneous outside some appropriate $\mathbf{t}$.

Before we can apply these observations to prove Theorem 7.3, let us establish two more crucial lemmas regarding m-structures. The first enables us to follow an inductive approach by giving us the tool needed reduce the value of m. The second applies this tool and establishes that in any regular and homogeneous m-structure, the values in the homogeneous part must be zero, that is $v_{\mathbf{i}} = 0$ for $\mathbf{i} \in \mathbb{N}_0^m \setminus [0, t_1 - 1] \times \cdots \times [0, t_m - 1]$.

Lemma 7.18. *For any regular m-structure $\{v_{\mathbf{j}} \in \mathbb{Q} : \mathbf{j} \in \mathbb{N}_0^m\}$ with*

$$\{\mathbf{c}_i = (c(i,1), \dots, c(i,m)) : 1 \le i \le d\} \tag{7.49}$$

that is homogeneous outside $\mathbf{t} = (t_1, \dots, t_m) \in \mathbb{N}_0^m$ and any $1 \le \ell \le d$, the values

$$\left\{\Delta_{\mathbf{i},\ell} : \mathbf{i} \in \mathbb{N}_0^{m-1}\right\}$$

define a regular $(m-1)$-structure with

$$\{\mathbf{c}'_i = (c(i,1), \dots, c(i,\ell-1), c(i,\ell+1), \dots, c(i,m)) : i \in S_\ell\} \tag{7.50}$$

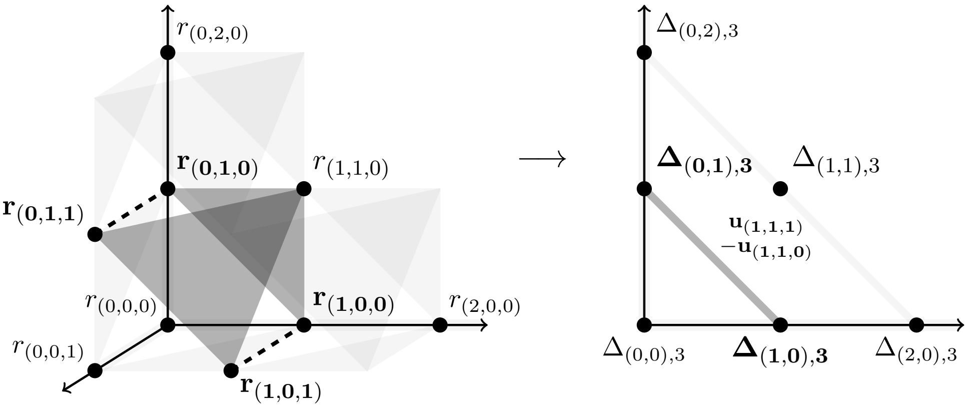
that is homogeneous outside $\mathbf{t}_\ell = (t_1, \dots, t_{\ell-1}, t_{\ell+1}, \dots, t_m)$.

Proof. For $\mathbf{j}' = (j_1, \dots, j_{m-1}) \in \mathbb{N}_0^{m-1}$ let $\mathbf{j} = (j_1, \dots, j_{\ell-1}, 0, j_\ell, \dots, j_{m-1})$. Using this notation, we set

$$\{u_{\mathbf{j}'} = u_{\mathbf{j}+\mathbb{1}_\ell} - u_{\mathbf{j}} : \mathbf{j}' \in \mathbb{N}_0^{m-1}\}$$

where $\{u_{\mathbf{j}} \in \mathbb{Z} : \mathbf{j} \in \mathbb{N}_0^m \setminus \{\mathbf{0}\}\}$ refers to the values of the m-structure. It follows that

$$\begin{aligned}\sum_{i \in S_\ell} \Delta_{\mathbf{j}' \ominus \mathbf{c}'_i, \ell} &= \sum_{i \in S_\ell} v_{(\mathbf{j}+\mathbb{1}_\ell) \ominus \mathbf{c}_i} - \sum_{i \in S_\ell} v_{\mathbf{j} \ominus \mathbf{c}_i} \\ &= \Big(u_{\mathbf{j}+\mathbb{1}_\ell} - \sum_{i \in S'_\ell} v_{(\mathbf{j}+\mathbb{1}_\ell)\ominus\mathbf{c}_i}\Big) - \Big(u_{\mathbf{j}} - \sum_{i \in S'_\ell} v_{\mathbf{j}\ominus\mathbf{c}_i}\Big) = u_{\mathbf{j}+\mathbb{1}_\ell} - u_{\mathbf{j}} = u_{\mathbf{j}'}.\end{aligned}$$


$r_{(0,2,0)}$
$\mathbf{r}_{(\mathbf{0},\mathbf{1},\mathbf{0})}$
$r_{(1,1,0)}$
$\mathbf{r}_{(\mathbf{0},\mathbf{1},\mathbf{1})}$
$r_{(0,0,0)}$
$\mathbf{r}_{(\mathbf{1},\mathbf{0},\mathbf{0})}$
$r_{(2,0,0)}$
$r_{(0,0,1)}$
$\mathbf{r}_{(\mathbf{1},\mathbf{0},\mathbf{1})}$
$\longrightarrow$
$\Delta_{(0,2),3}$
$\boldsymbol{\Delta}_{(\mathbf{0},\mathbf{1}),\mathbf{3}}$
$\Delta_{(1,1),3}$
$\mathbf{u}_{(\mathbf{1},\mathbf{1},\mathbf{1})}$
$-\mathbf{u}_{(\mathbf{1},\mathbf{1},\mathbf{0})}$
$\Delta_{(0,0),3}$
$\boldsymbol{\Delta}_{(\mathbf{1},\mathbf{0}),\mathbf{3}}$
$\Delta_{(2,0),3}$

(a) $c_1 = (1,0,0)$, $c_2 = (0,1,0)$ *and* $c_3 = (0,0,1)$.

(b) $c_1 = (1,0)$ *and* $c_2 = (0,1)$.

Figure 7.3: *Illustrating the reduction from a* 3*-structure to a* 2*-structure, both of them regular.*

Here we have used the fact that $\Delta_{\mathbf{j}'\ominus\mathbf{c}'_i,\ell} = v_{(\mathbf{j}+\mathbb{1}_\ell)\ominus\mathbf{c}_i} - v_{\mathbf{j}\ominus\mathbf{c}_i}$ since $c(i,\ell) = 0$ for $i \in S_\ell$ and that $(\mathbf{j} + \mathbb{1}_\ell) \ominus \mathbf{c}_i = \mathbf{j} \ominus \mathbf{c}_i$ since $c(i,\ell) \neq 0$ for $i \in S'_\ell$.

It follows that the values $\left\{\Delta_{\mathbf{i},\ell} : \mathbf{i} \in \mathbb{N}_0^{m-1}\right\}$ form an $(m-1)$-structure with $\{\mathbf{c}'_i : i \in S_\ell\}$ and $\{u_{\mathbf{j}'} : \mathbf{j}' \in \mathbb{N}_0^{m-1} \setminus \{\mathbf{0}\}\}$. As $u_{\mathbf{j}'} = u_{\mathbf{j}+\mathbb{1}_\ell} - u_{\mathbf{j}} = 0$ for $\mathbf{j}' \in \mathbb{N}_0^{m-1} \setminus [0,t_1] \times \cdots \times [0,t_{m-1}]$, it follows that the $(m-1)$-structure is homogeneous outside $\mathbf{t}_\ell$. Lastly, note that since the m-structure was regular, we have $\mathbf{c}'_i \in \{0,1\}^{m-1}$ as well as $\mathbf{c}'_i \neq \mathbf{0}$ since $\mathbf{c}_i \neq \mathbf{0}$ and $c(i,\ell) = 0$ for $i \in S_\ell$. It follows that the $(m-1)$-structure is regular as well. □

Lemma 7.19. *A regular m-structure $\{v_\mathbf{j} \in \mathbb{Q} : \mathbf{j} \in \mathbb{N}_0^m\}$ that is homogeneous outside $\mathbf{t} = (t_1, \ldots, t_m) \in \mathbb{N}_0^m$ and for which $S_\ell \neq \emptyset$ for any $1 \leq \ell \leq m$ satisfies $v_\mathbf{i} = 0$ for all $\mathbf{i} \in \mathbb{N}_0^m \setminus [0, t_1 - 1] \times \cdots \times [0, t_m - 1]$.*

Proof. We will prove the statement by induction on m. Let us start by showing the statement for $m = 1$. In this case, $\mathbf{c}_1, \ldots, \mathbf{c}_d$ are non-zero, positive integers satisfying $\mathbf{c}_1 = \cdots = \mathbf{c}_d = 1$ as the structure is regular. Note that $d \geq 1$ as $S_1 \neq \emptyset$. It follows that the relations defining the structure are of the type $d\, v_{\mathbf{j}\ominus 1} = u_\mathbf{j}$ for all $\mathbf{j} \in \mathbb{N}$. Since $u_\mathbf{j} = 0$ for $\mathbf{j} > \mathbf{t} = t_1$, we have $v_\mathbf{i} = 0$ for all $\mathbf{i} \in \mathbb{N}_0 \setminus [0, t_1 - 1]$ as desired.

Now assume that the statement is true for all regular $(m-1)$-structures and let us show that then it must also hold for any regular m-structure. Lemma 7.18 shows that $\{\Delta_{\mathbf{i},\ell} : \mathbf{i} \in \mathbb{N}_0^{m-1}\}$ is an $(m-1)$-structure that is homogeneous outside $\mathbf{t}_\ell$ for any

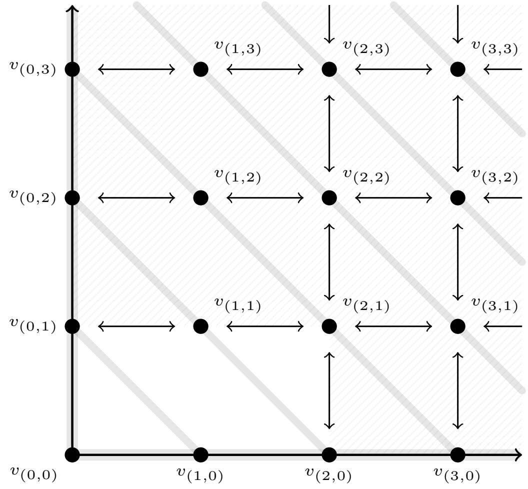


Figure 7.4: *Illustrating the equalities established by Equation* (7.53) *in Lemma 7.19 in a regular* 2*-structure given by* $\mathbf{c}_1 = (1,0)$ *and* $\mathbf{c}_2 = (0,1)$ *and homogenous outside* $(2,1)$.

$1 \leq \ell \leq m$. By the inductive assumption it follows that

$$\Delta_{\mathbf{i},\ell} = 0 \tag{7.51}$$

for all $\mathbf{i} \in \mathbb{N}_0^{m-1} \setminus [0, t_1 - 1] \times \cdots \times [0, t_\ell - 1] \times [0, t_\ell - 1] \times \cdots \times [0, t_m - 1]$. Furthermore, $\{v'_\mathbf{i} = v_{\mathbf{i}+\mathbb{1}_\ell} : \mathbf{i} \in \mathbb{N}_0^m\}$ forms an m-structure where the corresponding $\{u'_\mathbf{j} : \mathbf{j} \in \mathbb{N}_0^m\}$ satisfy

$$u'_\mathbf{j} = \begin{cases} u_{\mathbf{j}+\mathbb{1}_\ell} & \text{for } \mathbf{j} = (j_1, \ldots, j_m) \text{ s.t. } j_\ell \neq 0, \\ u_{\mathbf{j}+\mathbb{1}_\ell} + \sum_{i \in S'_\ell} \Delta_{\mathbf{j} \ominus \mathbf{c}_i, \ell} & \text{for } \mathbf{j} = (j_1, \ldots, j_m) \text{ s.t. } j_\ell = 0. \end{cases} \tag{7.52}$$

Note that this structure is again homogeneous outside $\mathbf{t}$, or in fact even homogenous outside $(t_1, \ldots, t_m - 1)$ if $t_m > 0$. Repeated application of this principle on the resulting m-structure gives us that that for any $1 \leq \ell \leq m$ we have

$$v_\mathbf{i} = v_{\mathbf{i}+\mathbb{1}_\ell} \text{ for all } \mathbf{i} = (i_1, \ldots, i_m) \in \mathbb{N}_0^m \text{ satisfying } i_j \geq t_j \text{ for some j } \neq \ell. \tag{7.53}$$

These relations are illustrated in Figure 7.4.

Now let $c = \max\{c(i,j) : 1 \leq i \leq d, 1 \leq j \leq m\}$ and $\mathbf{j} = (t_1 + c, \ldots, t_m + c)$. By Equation (7.53) we have $v_\mathbf{j} = v_{\mathbf{j} \ominus \mathbf{c}_i}$ for any $1 \leq i \leq d$ and $\mathbf{j} \in \mathbb{N}_0$. Considering the

Equation (7.48) given by $\mathbf{j}$, we therefore have

$$d\, v_{\mathbf{j}} = v_{\mathbf{j} \ominus \mathbf{c}_1} + \cdots + v_{\mathbf{j} \ominus \mathbf{c}_d} = u_{\mathbf{j}} = 0.$$

It follows that $v_{\mathbf{j}} = 0$ and hence, again by Equation (7.53), it follows that $v_{\mathbf{i}} = 0$ for all $\mathbf{i} \in \mathbb{N}_0^m \setminus [0, t_1 - 1] \times \cdots \times [0, t_m - 1]$ as desired. □

We are now ready to prove Theorem 7.3.

Proof of Theorem 7.3. Recall that $F_{\mathcal{A}}(z) = f_{\mathcal{A}}(z)^d$ and that the existence of a set $\mathcal{A}$ for which $r(\mathcal{A}, n; k_1, \ldots, k_d)$ is a constant function for n large enough would imply the existence of some polynomial $P(z) \in \mathbb{Z}[z]$ satisfying $P(1) \neq 0$ such that

$$F_{\mathcal{A}}(z^{k_1}) \cdots F_{\mathcal{A}}(z^{k_d}) = \frac{P^d(z)}{(1-z)^d}.$$

Using Proposition 7.16 we see that if a such a function $F_{\mathcal{A}}(z)$ were to exist, then the values $\{r_{\mathbf{i}} : \mathbf{i} \in \mathbb{N}_0^m\}$ together with $\mathbf{b}_1, \ldots, \mathbf{b}_m$ and $\{s_{\mathbf{j}} : \mathbf{j} \in \mathbb{N}_0^m \setminus \{\mathbf{0}\}\}$ would define an m-structure. By the requirements of the theorem we have $\mathbf{b}_i \in \{0,1\}^m$ and since $k_1, \ldots, k_d \geq 2$ we have $\mathbf{b}_i \neq \mathbf{0}$. We may also assume that $S_\ell \neq \emptyset$ for all $1 \leq \ell \leq d$ as otherwise there exists some ℓ' such that $q_{\ell'} \mid k_i$ for all $1 \leq i \leq d$, in which case the representation function clearly cannot become constant, so that this m-structure would be regular. It would also be homogeneous outside some appropriate $\mathbf{t} \in \mathbb{N}_0^m$ as $P(z)$ is a polynomial and hence $s_{\mathbf{j}} \neq 0$ only for finitely many $\mathbf{j} \in \mathbb{N}_0^m$. Finally, since $r_{\mathbf{i}} \equiv -1 \mod d$ for all $\mathbf{i} \in \mathbb{N}_0^m$, this would contradict the statement of Lemma 7.19, proving Theorem 7.3. ■

7.6 Further remarks

With Theorem 7.1 we have established an Erdős–Fuchs-type result for ordered representation functions, showing that an error term of the form $o(n^{1/4} \log^{-1/2} n)$ is not possible. It would be of interest to adapt the techniques in [108], see also [82], in order to rule out an error term of the form $o(n^{1/4})$. However, the fact that one has to introduce several extra terms in the encoding of $r_k^\star(\mathcal{A}, n)$ complicates this approach.

In a different direction, writing

$$r_k(\mathcal{A}, n) = \#\Big\{(a_1, \ldots, a_d) \in \mathcal{A}^d : a_1 + \ldots + a_d = n\Big\},$$

the work of [118], building on a previous construction [127, 36], shows that for any $k \geq 2$ there exists a set $\mathcal{A}_k$ and a constant $c > 0$ such that $\sum_{j=0}^{n}(r_k(\mathcal{A}_k, n) - c) = O(n^{1-3/(2k)})$, leaving a gap for $k \geq 3$. It would be very interesting to try to improve this bound for $k \geq 3$, both in the ordered and in the unordered case.

In Theorem 7.3 we have shown that under very general conditions for the coefficients $k_1, \ldots, k_d$ the weighted representation function cannot become constant. However, there are cases that are not covered by this result, including those where at least one of the k_i is equal to 1.

On the other side, let us point out that Moser's construction [109] can be trivially generalized to the case where $k_i = k^{i-1}$ for some integer value $k \geq 2$. In view of this construction and the results of this section, the following conjecture seems reasonable.

Conjecture 7.20. *There exists some infinite set of positive integers $\mathcal{A}$ such that $r_{\mathcal{A}}(n; k_1, \ldots, k_d)$ is constant for n large enough if and only if, up to permutation of the indices, $(k_1, \ldots, k_d) = (1, k, k^2, \ldots, k^{d-1})$, for some $k \geq 2$.*

The most likely candidates for a possible counterexample to this conjecture might be those where (k_1, k_2, k_3) is either $(1, 2, 6)$ or $(1, 2, 8)$. One could possibly try to generalize Moser's approach to these scenarios, e.g. by using generalized bases. Understanding these cases would most likely indicate a path towards completely settling the question of Sárközy and Sós.

Lastly, it would be of great interest to obtain an Erdős–Fuchs-type result for weighted representation functions in at least some of the cases covered by Theorem 7.3. See also [120], where some Erdős–Fuchs-type results were obtained in this setting for certain types of weights.

Bibliography

[1] Miklós Ajtai, János Komlós, and Endre Szemerédi. A dense infinite Sidon sequence. *European Journal of Combinatorics*, 2(1):1–11, 1981.

[2] Noga Alon and Joel H. Spencer. *The Probabilistic Method.* Wiley-Interscience series in Discrete Mathematics and Optimization. John Wiley & Sons, 3rd edition, 2008.

[3] József Balogh, Robert Morris, and Wojciech Samotij. Independent sets in hypergraphs. *Journal of the American Mathematical Society*, 28(3):669–709, 2015.

[4] József Balogh and Wojciech Samotij. On the Chvátal-Erdős triangle game. *The Electronic Journal of Combinatorics*, 18(P72), 2011.

[5] Andreas Baltz, Peter Hegarty, Jonas Knape, Urban Larsson, and Tomasz Schoen. The structure of maximum subsets of $\{1, \ldots, n\}$ with no solutions to $a + b = kc$. *Electronic Journal of Combinatorics*, 12:R19, 2005.

[6] József Beck. Van der Waerden and Ramsey type games. *Combinatorica*, 1(2):103–116, 1981.

[7] József Beck. Random graphs and positional games on the complete graph. In *Random Graphs '83*, volume 29 of *Annals of Discrete Mathematics*, pages 7–13. Elsevier, 1985.

[8] József Beck. *Combinatorial Games: Tic-Tac-Toe Theory*, volume 114 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, 2008.

[9] Matthias Beck and Sinai Robins. *Computing the Continuous Discretely.* Undergraduate Texts in Mathematics. Springer Science+Business Media, 2nd edition, 2007.

[10] Małgorzata Bednarska and Tomasz Łuczak. Biased positional games for which random strategies are nearly optimal. *Combinatorica*, 20(4):477–488, 2000.

[11] Felix A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proceedings of the National Academy of Sciences of the United States of America*, 32(12):331–332, 1946.

[12] Elwyn R. Berlekamp. A construction for partitions which avoid long arithmetic progressions. *Canadian Mathematical Bulletin*, 11(3):409–414, 1968.

[13] Thomas F. Bloom. A quantitative improvement for Roth's theorem on arithmetic progressions. *Journal of the London Mathematical Society*, 93(3):643–663, 2016.

[14] Béla Bollobás and Andrew G. Thomason. Threshold functions. *Combinatorica*, 7(1):35–38, 1987.

[15] Jean Bourgain. On triples in arithmetic progression. *Geom. Funct. Anal.*, 9(5):968–984, 1999.

[16] Jean Bourgain. Roth's theorem on progressions revisited. *Journal d'Analyse Mathématique*, 104(1):155–192, 2008.

[17] Emmanuel Breuillard, Ben Green, and Terence Tao. The structure of approximate groups. *Publications mathématiques de l'IHÉS*, 116(1):115–221, 2012.

[18] Benno Büeler, Andreas Enge, and Komei Fukuda. Exact volume computation for polytopes: a practical study. In *Polytopes – Combinatorics and computation*, volume 29 of *DMV Seminar*, pages 131–154. Springer Basel AG, 2000.

[19] Pablo Candela and Anne De Roton. On sets with small sumset in the circle. *The Quarterly Journal of Mathematics*, 70(1):49–69, 2019.

[20] Pablo Candela, Diego González-Sánchez, and David J. Grynkiewicz. On sets with small sumset and m-sum-free sets in $\mathbb{Z}/p\mathbb{Z}$. 2019. arXiv preprint arXiv:1909.07967.

[21] Pablo Candela, Oriol Serra, and Christoph Spiegel. A step beyond Freĭman's theorem for set addition modulo a prime. *Journal de Théorie des Nombres de Bordeaux*, to appear.

[22] Gonzalo Cao-Labora, Juanjo Rué, and Christoph Spiegel. An Erdős–Fuchs theorem for ordered representation functions. 2019. arXiv preprint arXiv:1911.12313.

[23] Augustin L. B. Cauchy. Recherches sur les nombres. *Journal de l'École Polytechnique*, 9:99–116, 1812.

[24] Mei-Chu Chang. A polynomial bound in Freiman's theorem. *Duke Mathematical Journal*, 113(3):399–419, 2002.

[25] Sarvadaman Chowla. Solution of a problem of Erdős and Turán in additive-number theory. *Proceedings of the National Academy of Sciences, India Section A*, 14(1-2):5–4, 1944.

[26] Fan R. K. Chung and John L. Goldwasser. Integer sets containing no solutions to $x + y = 3k$. In *The Mathematics of Paul Erdős*, volume 13 of *Algorithms and Combinatorics*, pages 267–277. Springer–Verlag, 1996.

[27] Fan R. K. Chung and John L. Goldwasser. Maximum subsets of $(0, 1]$ with no solutions to $x + y = kz$. *Electronic journal of Combinatorics*, 3(1):R1, 1996.

[28] Vašek Chvátal and Paul Erdős. Biased positional games. In *Algorithmic Aspects of Combinatorics*, volume 2 of *Annals of Discrete Mathematics*, pages 221–229. Elsevier, 1978.

[29] Javier Cilleruelo. Infinite Sidon sequences. *Advances in Mathematics*, 255:474–486, 2014.

[30] Javier Cilleruelo and Juanjo Rué. On a question of Sárközy and Sós for bilinear forms. *Bulletin of the London Mathematical Society*, 41(2):274–280, 2009.

[31] Javier Cilleruelo, Imre Ruzsa, and Carlos Vinuesa. Generalized Sidon sets. *Advances in Mathematics*, 225(5):2786–2807, 2010.

[32] Javier Cilleruelo, Imre Z. Ruzsa, and Carlos Trujillo. Upper and lower bounds for finite $B_h[g]$ sequences. *Journal of Number Theory*, 97(1):26–34, 2002.

[33] Javier Cilleruelo and Rafael Tesoro. On sets free of sumsets with summands of prescribed size. *Combinatorica*, 38(3):511–546, 2019.

[34] David Conlon and William T. Gowers. Combinatorial theorems in sparse random sets. *Annals of Mathematics*, 184(2):367–454, 2016.

[35] Ernie Croot, Vsevolod F. Lev, and Péter P. Pach. Progression-free sets in $\mathbb{Z}\mathbb{Z}_4^n$ are exponentially small. *Annals of Mathematics*, 185(1):331–337, 2017.

[36] Li-Xia Dai and Hao Pan. Inverse Erdős–Fuchs theorem for k-fold sumsets. *Journal of Number Theory*, 140:1–12, 2014.

[37] Harold Davenport. On the addition of residue classes. *Journal of the London Mathematical Society*, s1-10(1):30–32, 1935.

[38] Jesús A. De Loera. The many aspects of counting lattice points in polytopes. *Mathematische Semesterberichte*, 52(2):175–195, 2005.

[39] Jesús A. De Loera, Jörg Rambau, and Francisco Santos. *Triangulations*, volume 25 of *Algorithms and Computation in Mathematics.* Springer Science+Business Media, 2010.

[40] Domingos Dellamonica, Yoshiharu Kohayakawa, Sang J. Lee, Vojtěch Rödl, and Wojciech Samotij. On the number of B_h-sets. *Combinatorics, Probability and Computing*, 25(1):108–129, 2016.

[41] Domingos Dellamonica, Yoshiharu Kohayakawa, Sang J. Lee, Vojtěch Rödl, and Wojciech Samotij. The number of B_h-sets of a given cardinality. *Proceedings of the London Mathematical Society*, 116(3):629–669, 2018.

[42] Jean-Marc Deshouillers and Gregory A. Freiman. A step beyond Kneser's theorem for abelian finite groups. *Proceedings of the London Mathematical Society*, 86(1):1–28, 2003.

[43] Gabriel A. Dirac. Note on a problem in additive number theory. *Journal of the London Mathematical Society*, s1-26(4):312–313, 1951.

[44] Eugène Ehrhart. Sur les polyèdres homothétiques bordés à n dimensions. *Comptes rendus de l'Académie des Sciences*, 254:988–990, 1962.

[45] Michael Elkin. An improved construction of progression-free sets. *Israel Journal of Mathematics*, 184(1):93–128, 2011.

[46] Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Annals of Mathematics*, 185(1):339–343, 2017.

[47] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.

[48] Pál Erdős. On a problem of Sidon in additive number theory and on some related problems (addendum). *Journal of the London Mathematical Society*, 19(76 Part 4):208, 1944.

[49] Pál Erdős. Solved and unsolved problems in combinatorics and combinatorial number theory. *Congressus Numerantium*, 32:49–62, 1981.

[50] Pál Erdős and Wolfgang H. J. Fuchs. On a problem of additive number theory. *Journal of the London Mathematical Society*, s1-31(1):67–73, 1956.

[51] Pál Erdős and John L. Selfridge. On a combinatorial game. *Journal of Combinatorial Theory, Series A*, 14(3):298–301, 1973.

[52] Pál Erdős and Miklós Simonovits. A limit theorem in graph theory. *Studia Scientiarum Mathematicarum Hungarica*, 1, 1966.

[53] Pál Erdős and Arthur H. Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52(12):1087–1091, 1946.

[54] Pál Erdős and Pál Turán. On some sequences of integers. *Journal of the London Mathematical Society*, s1-11(4):261–264, 1936.

[55] Pál Erdős and Pál Turán. On a problem of Sidon in additive number theory, and on some related problems. *Journal of the London Mathematical Society*, s1-16(4):212–215, 1941.

[56] David Fabian, Juanjo Rué, and Christoph Spiegel. On strong infinite Sidon and B_h sets and random sets of integers. 2019. arXiv preprint arXiv:1911.13275.

[57] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge: Cambridge University Press, 2009.

[58] Peter Frankl, Ronald Lewis Graham, and Vojtěch Rödl. Quantitative theorems for regular systems of equations. *Journal of Combinatorial Theory, Series A*, 47(2):246–261, 1988.

[59] Gregory A. Freiman. Inverse problems in additive number theory. Addition of sets of residues modulo a prime. *Doklady Akademii Nauk SSSR*, 141(3):571–573, 1961.

[60] Gregory A. Freĭman. On the addition of finite sets. *Doklady Akademii Nauk SSSR*, 158(5):1038–1041, 1964.

[61] Gregory A. Freiman. Foundations of a structural theory of set addition. *Translation of Mathematical Monographs*, 37, 1973.

[62] Gregory A. Freiman. Inverse additive number theory. XI. Long arithmetic progressions in sets with small sumsets. *Acta Arithmetica*, 137(4):325–331, 2009.

[63] Gregory A. Freiman. On the additive volume of sets of integers. 2014. arXiv preprint arXiv:1412.5082.

[64] Gregory A. Freiman and Oriol Serra. On doubling and volume: chains. *Acta Arithmetica*, 186:37–59, 2018.

[65] Gregory A. Freĭman, Oriol Serra, and Christoph Spiegel. Additive volume of sets contained in few arithmetic progressions. *INTEGERS: Electronic Journal of Combinatorial Number Theory*, 19:A34, 2019.

[66] Ehud Friedgut, Vojtěch Rödl, and Mathias Schacht. Ramsey properties of random discrete structures. *Random Structures & Algorithms*, 37(4):407–436, 2010.

[67] Heidi Gebauer and Tibor Szabó. Asymptotic random graph intuition for the biased connectivity game. *Random Structures & Algorithms*, 35(4):431–443, 2009.

[68] Anant P. Godbole, Svante Janson, Nicholas W. Locantore Jr., and Rebecca Rapoport. Random Sidon sequences. *Journal of Number Theory*, 75(1):7–22, 1999.

[69] William T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.

[70] Ben Green. A Szemerédi-type regularity lemma in abelian groups, with applications. *Geometric & Functional Analysis*, 15(2):340–376, 2005.

[71] Ben Green and Imre Z. Ruzsa. Sets with small sumset and rectification. *Bulletin of the London Mathematical Society*, 38(1):43–52, 2006.

[72] Ben Green and Imre Z. Ruzsa. Freiman's theorem in an arbitrary abelian group. *Journal of the London Mathematical Society*, 75(1):163–175, 2007.

[73] Ben Green and Julia Wolf. A note on Elkin's improvement of Behrend's construction. In *Additive number theory*, pages 141–144. Springer, 2010.

[74] David J. Grynkiewicz. *Structural additive theory*, volume 30. Springer Science+Business Media, 2013.

[75] David S. Gunderson and Vojtěch Rödl. Extremal problems for affine cubes of integers. *Combinatorics, Probability and Computing*, 7(1):65–79, 1998.

[76] Yahya O. Hamidoune and Alain Plagne. A generalization of Freiman's $3k-3$ theorem. *Acta Arithmetica*, 103(2):147–156, 2002.

[77] Yahya O. Hamidoune, Oriol Serra, and Gilles Zémor. On the critical pair theory in $\mathbb{Z}/p\mathbb{Z}$. *Acta Arithmetica*, 121(2):99 – 115, 2006.

[78] Robert Hancock, Katherine Staden, and Andrew Treglown. Independent sets in hypergraphs and Ramsey properties of graphs and the integers. *SIAM Journal on Discrete Mathematics*, 33(1):153–188, 2019.

[79] Robert Hancock and Andrew Treglown. On solution-free sets of integers. *European Journal of Combinatorics*, 66:110–128, 2017.

[80] Robert Hancock and Andrew Treglown. On solution-free sets of integers (II). *Acta Arithmetica*, 180:15–33, 2017.

[81] Godfrey H. Hardy. On the expression of a number as the sum of two squares. *Quarterly Journal of Mathematics*, 46:263–283, 1915.

[82] Elmer K. Hayashi. Omega theorems for the iterated additive convolution of a nonnegative arithmetic function. *Journal of Number Theory*, 13(2):176–191, 1981.

[83] David R. Heath-Brown. Integer sets containing no arithmetic progressions. *Journal of the London Mathematical Society*, s2-35(3):385–394, 1987.

[84] David Hilbert. Ueber die Irreducibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten. *Journal für die reine und angewandte Mathematik*, 110:104–129, 1892.

[85] G. Horváth. An improvement of an extension of a theorem of Erdős and Fuchs. *Acta Mathematica Hungarica*, 104(1-2):27–37, 2004.

[86] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In *Random Graphs '87*, pages 73–87. John Wiley & Sons, 1990.

[87] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random Graphs*. Wiley-Interscience series in Discrete Mathematics and Optimization. John Wiley & Sons, 2000.

[88] Svante Janson and Andrzej Ruciński. Upper tails for counting objects in randomly induced subhypergraphs and rooted random graphs. *Arkiv för Matematik*, 49(1):79–96, 2011.

[89] Renling Jin. Freiman's inverse problem with small doubling property. *Advances in Mathematics*, 216(2):711–752, 2007.

[90] Renling Jin. Detailed structure for Freiman's 3k-3 theorem. *INTEGERS: Electronic Journal of Combinatorial Number Theory*, 15A:A9, 2015.

[91] Johannes H. B. Kemperman. On small sumsets in an abelian group. *Acta Mathematica*, 103(1-2):63–88, 1960.

[92] Jeong H. Kim and Van H. Vu. Concentration of multivariate polynomials and its applications. *Combinatorica*, 20(3):417–434, 2000.

[93] Martin Kneser. Abschätzung der asymptotischen Dichte von Summenmengen. *Mathematische Zeitschrift*, 58(1):459–484, 1953.

[94] Yoshiharu Kohayakawa, Sang J. Lee, Carlos G. Moreira, and Vojtěch Rödl. Infinite Sidon sets contained in sparse random sets of integers. *SIAM Journal on Discrete Mathematics*, 32(1):410–449, 2018.

[95] Yoshiharu Kohayakawa, Sang J. Lee, Carlos G. Moreira, and Vojtěch Rödl. On strong Sidon sets of integers. 2019. Available online at `ime.usp.br/~yoshi/`.

[96] Yoshiharu Kohayakawa, Sang J. Lee, Vojtěch Rödl, and Wojciech Samotij. The number of Sidon sets and the maximum size of Sidon sets contained in a sparse random set of integers. *Random Structures & Algorithms*, 46(1):1–25, 2015.

[97] Sergei V. Konyagin and Vsevolod F. Lev. Combinatorics and linear algebra of Freiman's isomorphism. *Mathematika*, 47(1-2):39–51, 2000.

[98] Daniel Král', Oriol Serra, and Lluís Vena. A removal lemma for systems of linear equations over finite fields. *Israel J. Math.*, 187(1):193–207, 2012.

[99] Michael Krivelevich. The critical bias for the Hamiltonicity game is $(1 + o(1))n/\ln(n)$. *Journal of the American Mathematical Society*, 24(1):125–131, 2011.

[100] Fritz Krückeberg. B_2-Folgen und verwandte Zahlenfolgen. *Journal für die reine und angewandte Mathematik*, 206:53–60, 1961.

[101] Christopher Kusch, Juanjo Rué, Christoph Spiegel, and Tibor Szabó. On the optimality of the uniform random strategy. *Random Structures & Algorithms*, 55(2):371–401, 2019.

[102] Izabella Łaba and Michael T. Lacey. On sets of integers not containing long arithmetic progressions. 2011. arXiv preprint math/0108155.

[103] Vsevolod F. Lev. Distribution of points on arcs. *INTEGERS: Electronic Journal of Combinatorial Number Theory*, 5(2):A11, 2005.

[104] Vsevolod F. Lev and Pavel Y. Smeliansky. On addition of two distinct sets of integers. *Acta Arithmetica*, 70(1):85–91, 1995.

[105] Neil Lyall. A new proof of Sárközy's theorem. *Proceedings of the American Mathematical Society*, 141(7):2253–2264, 2013.

[106] Ian G. Macdonald. The volume of a lattice polyhedron. *Mathematical Proceedings of the Cambridge Philosophical Society*, 59(4):719–726, 1963.

[107] Abdul Majid Mian and S Chowla. On the b_2 sequences of Sidon. *Proceedings of the National Academy of Sciences, India Section A*, 14(3–4), 1944.

[108] H. L. Montgomery and R. C. Vaughan. On the Erdős–Fuchs theorems. In *A tribute to Paul Erdős*, pages 331–338. Cambridge University Press, 1990.

[109] Leo Moser. An application of generating series. *Mathematics Magazine*, 35(1):37–38, 1962.

[110] Rajko Nenadov and Angelika Steger. A short proof of the random Ramsey theorem. *Combinatorics, Probability and Computing*, 25(1):130–144, 2016.

[111] Kevin O'Bryant. Sets of integers that do not contain long arithmetic progressions. *Electronic Journal of Combinatorics*, 18(1):P59, 2011.

[112] Richard Rado. Studien zur Kombinatorik. *Mathematische Zeitschrift*, 36(1):424–470, 1933.

[113] Robert Alexander Rankin. Xxiv.—sets of integers containing not more than a given number of terms in arithmetical progression. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 65(4):332–344, 1961.

[114] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness.* Methods of Modern Mathematical Physics. Academic Press, 1975.

[115] Vojtech Rödl and Andrzej Ruciński. Rado partition theorem for random subsets of integers. *Proc. London Math. Soc.*, s3-74(3):481–502, 1997.

[116] Øystein J. Rødseth. On Freiman's 2.4-theorem. *Skrifter. Det Kongelige Norske Videnskabers Selskab*, 4:11–18, 2006.

[117] Klaus F. Roth. On certain sets of integers. *Journal of the London Mathematical Society*, s1-28(1):104–109, 1953.

[118] Eszter Rozgonyi and Csaba Sándor. A converse to an extension of a theorem of Erdős and Fuchs. *Journal of Combinatorics and Number Theory*, 5(3):151–163, 2013.

[119] Andrzej Ruciński. Small subgraphs of random graphs – a survey. In *Random Graphs '87*, pages 283–303. John Wiley & Sons, 1990.

[120] Juanjo Rué. On polynomial representation functions for multilinear forms. *European Journal of Combinatorics*, 34(8):1429–1435, 2013.

[121] Juanjo Rué and Christoph Spiegel. On a problem of Sárközy and Sós on multivariate linear forms. *Revista Iberoamericana*, to appear.

[122] Juanjo Rué, Christoph Spiegel, and Ana Zumalacárregui. Threshold functions and Poisson convergence for systems of equations in random sets. *Mathematische Zeitschrift*, 288(1-2):333–360, 2018.

[123] Imre Z. Ruzsa. Difference sets without squares. *Periodica Mathematica Hungarica*, 15(3):205–209, 1984.

[124] Imre Z. Ruzsa. Solving a linear equation in a set of integers I. *Acta Arith.*, 65(3):259–282, 1993.

[125] Imre Z. Ruzsa. Generalized arithmetical progressions and sumsets. *Acta Mathematica Hungarica*, 65(4):379–388, 1994.

[126] Imre Z. Ruzsa. Solving a linear equation in a set of integers II. *Acta Arith.*, 72(4):385–397, 1995.

[127] Imre Z. Ruzsa. A converse to a theorem of Erdős and Fuchs. *Journal of Number Theory*, 62(2):397–402, 1997.

[128] Imre Z. Ruzsa. An infinite Sidon sequence. *Journal of Number Theory*, 68(1):63–71, 1998.

[129] Tom Sanders. Appendix to 'Roth's theorem on progressions,' revisited by J. Bourgain. *Journal d'Analyse Mathématique*, 104(1):193–206, 2008.

[130] Tom Sanders. On Roth's theorem on progressions. *Annals of Mathematics*, 174(1):619–636, 2011.

[131] Csaba Sándor. Non-degenerate hilbert cubes in random sets. *Journal de Théorie des Nombres de Bordeaux*, 19(1):249–261, 2007.

[132] Andras Sárközy. On difference sets of sequences of integers. III. *Acta Mathematica Hungaricae*, 31(3-4):355–386, 1978.

[133] András Sárközy. On a theorem of Erdős and Fuchs. *Acta Arithmetica*, 37:333–338, 1980.

[134] András Sárközy and Vera T. Sós. On additive representation functions. In *The mathematics of Paul Erdős I. Algorithms and Combinatorics*, volume 13, pages 129–150. Springer, 1997.

[135] David Saxton and Andrew Thomason. Hypergraph containers. *Inventiones mathematicae*, 201(3):925–992, 2015.

[136] Matthias Schacht. Extremal results for random discrete structures. *Annals of Mathematics*, 184(2):333–365, 2016.

[137] Tomasz Schoen. Near optimal bounds in Freiman's theorem. *Duke Mathematical Journal*, 158(1):1–12, 2011.

[138] Alexander Schrijver. *Theory of Linear and Integer Programming.* Wiley-Interscience series in Discrete Mathematics and Optimization. John Wiley & Sons, 1986.

[139] Issai Schur. über die Kongruenz $x^m + y^m \equiv z^m \mod p$. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 25:114–117, 1917.

[140] Oriol Serra and Gilles Zémor. Large sets with small doubling modulo p are well covered by an arithmetic progression. *Annales de l'institut Fourier*, 59(5):2043–2060, 2009.

[141] Asaf Shapira. Behrend-type constructions for sets of linear equations. *Acta Arithmetica*, 122(1):17–33, 2006.

[142] Asaf Shapira. A proof of Green's conjecture regarding the removal properties of sets of linear equations. *Journal of the London Mathematical Society*, 81(2):355–373, 2010.

[143] James Singer. A theorem in finite projective geometry and some applications to number theory. *Transactions of the American Mathematical Society*, 43(3):377–385, 1938.

[144] Christoph Spiegel. A note on sparse supersaturation and extremal results for linear homogeneous systems. *Electronic Journal of Combinatorics*, 24(3):P3.38, 2017.

[145] Yonutz V. Stanchescu. On the simplest inverse problem for sums of sets in several dimensions. *Combinatorica*, 18(1):139–149, 1998.

[146] Yonutz V. Stanchescu. The structure of d–dimensional sets with small sumset. *Journal of Number Theory*, 130(2):289–303, 2010.

[147] Alfred Stöhr. Gelöste und ungelöste Fragen über Basen der natürlichen Zahlenreihe. I. *Journal für die reine und angewandte Mathematik*, 194:40–65, 1955.

[148] Benny Sudakov and Van H. Vu. Local resilience of graphs. *Random Structures & Algorithms*, 33(4):409–433, 2008.

[149] Zoltán Szabó. An application of Lovász' local lemma – a new lower bound for the van der Waerden number. *Random Structures & Algorithms*, 1(3):343–360, 1990.

[150] Endre Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arithmetica*, 27:199–245, 1975.

[151] Endre Szemerédi. Integer sets containing no arithmetic progressions. *Acta Mathematica Hungaricae*, 56(1-2):155–158, 1990.

[152] Min Tang. On a generalization of a theorem of Erdős and Fuchs. *Discrete Mathematics*, 309(21):6288–6293, 2009.

[153] Terence Tao. Freiman's theorem for solvable groups. *Contributions to Discrete Mathematics*, 5(2):137–184, 2010.

[154] Terence Tao and Van H. Vu. *Additive Combinatorics*, volume 105 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 2006.

[155] Matthew C. H. Tointon. Freiman's theorem in an arbitrary nilpotent group. *Proceedings of the London Mathematical Society*, 109(2):318–352, 2014.

[156] Matthew C. H. Tointon. Polylogarithmic bounds in the nilpotent Freiman theorem. In *Mathematical Proceedings of the Cambridge Philosophical Society*, pages 1–17, 2018.

[157] Bartel L. Van der Waerden. Beweis einer Baudetschen Vermutung. *Nieuw Archief voor Wiskunde*, 15(2):212–216, 1927.

[158] Panayiotis Varnavides. On certain sets of positive density. *Journal of the London Mathematical Society*, s1-34(3):358–360, 1959.

[159] Alan G. Vosper. The critical pairs of subsets of a group of prime order. *Journal of the London Mathematical Society*, s1-31(2):200–205, 1956.

[160] Lutz Warnke. Upper tails for arithmetic progressions in random subsets. *Israel Journal of Mathematics*, 221(1):317–365, 2017.

[161] Günter M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer Science+Business Media, 1995.