ADDITIVE COMBINATORICS AND DIOPHANTINE PROBLEMS

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Science and Engineering

2021

Jonathan C. Chapman
Department of Mathematics
School of Natural Sciences

Contents

Word count 24044

The University of Manchester

Jonathan C. Chapman
Doctor of Philosophy
Additive Combinatorics and Diophantine Problems
August 25, 2021

This thesis presents new results on the topics of partition regularity and density regularity. The first chapter provides an introduction to these subjects and an overview of the main results of this thesis.

In Chapter 2, we study the connections between partition regularity and multiplicatively syndetic sets. In particular, we prove that a dilation invariant system of polynomial equations is partition regular if and only if it has a solution inside every multiplicatively syndetic set. We also adapt methods of Green-Tao and Chow-Lindqvist-Prendiville to develop a syndetic version of Roth's density increment strategy. This argument is then used to obtain bounds on the Rado-Ramsey numbers of configurations of the form $\{x, d, x + d, x + 2d\}$.

In Chapter 3, we establish new partition and density regularity results for systems of diagonal equations in kth powers. Our main result shows that if the coefficient matrix of such a system is sufficiently non-singular, then the system is partition regular if and only if it satisfies Rado's columns condition. Furthermore, if the system also admits constant solutions, then we prove that the system has non-trivial solutions over every set of integers of positive upper density.

In Chapter 4, we obtain a double exponential bound in Brauer's generalisation of van der Waerden's theorem on arithmetic progressions with the same colour as their common difference. Using Gowers' local inverse theorem, we obtain a bound which is quintuple exponential in the length of the progression. We refine this bound in the colour aspect for three-term progressions, and combine our arguments with an insight of Lefmann to obtain analogous bounds for the Rado-Ramsey numbers of certain non-linear quadratic equations. The content of this chapter is joint work with Sean Prendiville.

Finally, in Chapter 5, we conclude with a summary of this thesis and describe possible directions for future research.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

First and foremost, I would like to thank Sean Prendiville for taking me on as a PhD student and for starting me on this journey. His endless enthusiasm and encouragement have been invaluable to my mathematical development and to the creation of this thesis. I am grateful for the many constructive discussions and collaborations we have had and that I hope will continue.

I am grateful to Christopher Frei for supervising me during my third year and for providing me with the opportunity to explore research in a new area of mathematics. Our regular meetings have been particularly insightful and helpful in completing my research projects. I am thankful to Donald Robertson for his supervision during the final year of my studies, and for his continual support and encouragement.

I have benefited from the wisdom and support of many mathematicians. In particular, I would like to thank Tom Sanders, Joel Moreira, Trevor Wooley, and Borys Kuca for enlightening conversations. I am indebted to Tom Sanders and Trevor Wooley for providing academic references for my job applications, and to Hung Bui for providing a reference for my application for associate fellowship of the Higher Education Academy.

I extend my sincerest gratitude to the University of Manchester for funding my PhD studies, providing numerous opportunities to develop my academic and professional skills, and for providing a nurturing research environment.

I am grateful for the love and support of my friends and family. Finally, I would like to thank my mother, to whom I owe all that I have achieved and all that I will do.

Publications

- Chapter 2 and Appendix A are based on the pre-print: J. Chapman, *Partition regularity and multiplicatively syndetic sets*, arXiv:1902.01149v3. This article has been published by Acta Arithmetica as:

  J. Chapman, *Partition regularity and multiplicatively syndetic sets*, Acta Arith. **196** (2020), 109–138. DOI: 10.4064/aa190421-11-3. MR4146371.

- Chapter 3 is based on the pre-print: J. Chapman, *Partition regularity for systems of diagonal equations*, arXiv:2003.10977v1. A revised version of this pre-print has been accepted (06 April 2021) and published online under "Advance articles" (11 May 2021) by International Mathematics Research Notices (IMRN). DOI: 10.1093/imrn/rnab100.

- Chapter 4 is based on the pre-print: J. Chapman and S. Prendiville, *On the Ramsey number of the Brauer configuration*, arXiv:1904.07567v2. This article has been published by Bulletin of the London Mathematical Society as:

  J. Chapman and S. Prendiville, *On the Ramsey number of the Brauer configuration*, Bull. Lond. Math. Soc. **52** (2020), no. 2, 316–334. DOI: 10.1112/blms.12327. MR4171368.

Chapter 1

Introduction

1.1 Background

The central topic of this thesis concerns the combinatorial properties of integer solutions to systems of integer polynomial equations. Such equations are known as *Diophantine equations* in honour of the 3rd century Hellenistic mathematician Diophantus of Alexandria and his influential work *Arithmetica* (see [Bas97]). The two types of combinatorial Diophantine problems we are primarily interested in are *density problems* and *colouring problems*. These concern, respectively, the properties of *density regularity* and *partition regularity*.

1.1.1 Density problems

A density problem asks how large a set of integers can be which lacks solutions to a given system of Diophantine equations. One way of measuring the largeness of an infinite set of positive integers A is by considering its *upper density*, which is defined as the quantity

$$\limsup_{N\to\infty} \frac{|A \cap \{1, \ldots, N\}|}{N}.$$

We call a system of Diophantine equations *density regular* if the system has non-constant solutions over any set of positive integers with positive upper density. Here, a *non-constant solution* is a solution $(x_1, \ldots, x_s)$ for which the x_i are not all equal.

The foundational results of this subject originate with a question posed by Erdős and Turán [ET36]: how large can a subset of $\{1, \ldots, N\}$ be which does not contain a k-term arithmetic progression? Here, a *k-term arithmetic progression* is a configuration of the form

$$\{x, x+d, x+2d, \ldots, x+(k-1)d\} \quad (x, d \in \mathbb{N}).$$

More precisely, one can speculate on the asymptotic behaviour of the function $r_k(N)$, which is defined to be the largest $M \in \{1, \ldots, N\}$ such that there exists a set $A \subseteq \{1, \ldots, N\}$ with $|A| = M$ which does not contain any k-term arithmetic progressions. Proving density regularity for k-term arithmetic progressions, meaning density regularity for the system of $k-2$ equations

$$\begin{aligned} x_1 + x_3 &= 2x_2; \\ x_2 + x_4 &= 2x_3; \\ &\vdots \\ x_{k-2} + x_k &= 2x_{k-1}, \end{aligned}$$

is therefore equivalent to showing that $r_k(N)/N \to 0$ as $N \to \infty$. For $k = 3$, this was achieved by Roth [Rot53] using Fourier analytic techniques. The situation for general k was resolved by Szemerédi [Sze75] using an intricate combinatorial argument. Consequently, the statement that k-term arithmetic progressions are density regular is known today as *Szemerédi's theorem*. Moreover, it follows from Szemerédi's theorem that the linear homogeneous system of integer equations

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,s}x_s &= 0; \\ &\vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,s}x_s &= 0 \end{aligned}$$

is density regular if and only if the system admits at least one non-constant solution and $\sum_{j=1}^{s} a_{i,j} = 0$ for all $i \in \{1, \ldots, n\}$ (see [FGR88, Fact 4]).

Since Szemerédi's seminal work, many other proofs of Szemerédi's theorem have been found. Shortly following the publication of [Sze75], Furstenberg discovered an ergodic theoretic proof [Fur77] which led to the birth of ergodic Ramsey theory. Much later, Gowers [Gow01] generalised the techniques of Roth to obtain a new quantitative proof of Szemerédi's theorem. This work of Gowers is especially significant for this thesis, as it is in [Gow01] that Gowers introduced higher order Fourier analysis. These techniques continue to deliver the best known quantitative estimates for $r_k(N)$ and for analogous quantities arising in density problems for other arithmetic configurations [Pre17, Pel20]. Of particular note is the use of Fourier analysis and its higher order analogue by Green [Gre05A] and, subsequently, Green and Tao [GT08] to extend the theorems of Roth and Szemerédi respectively to the primes. For further information on the numerous proofs of Szemerédi's theorem and the interplay between them, we refer the reader to the survey [Tao07] and the references therein.

1.1.2 Partition regularity

Colouring problems (also known as the *partition problems* or *Ramsey problems*) ask if, for a given system of Diophantine equations, there exists a finite colouring of the positive integers with no monochromatic solutions to the system. A system of Diophantine equations

$$P_1(x_1, \ldots, x_s) = P_2(x_1, \ldots, x_s) = \cdots = P_n(x_1, \ldots, x_s) = 0$$

is called *partition regular* if, for any finite partition of the set of positive integers $\mathbb{N} = C_1 \cup \cdots \cup C_r$, we can find some $i \in \{1, \ldots, r\}$ such that there is a solution $x_1, \ldots, x_s \in C_i$ to the system. In analogy with graph Ramsey theory, we usually refer to a finite partition of this form as a *finite colouring*, and the cells of the partition C_i as the *colours*. A set $S \subseteq \mathbb{N}$ is said to be *monochromatic* with respect to this colouring if $S \subseteq C_i$ for some i.

The earliest result in partition regularity appears in a paper of Hilbert [Hil92] in which it is shown that every finite colouring of the positive integers admits arbitrarily large monochromatic structures of the form

$$\{x + \varepsilon_1 v_1 + \cdots + \varepsilon_n v_n : \varepsilon_1, \ldots, \varepsilon_n \in \{0, 1\}\} \quad (x, v_1, \ldots, v_n \in \mathbb{N}).$$

Two foundational colouring results are the theorems of Schur [Sch16] and van der Waerden [Wae27]. Schur's theorem asserts that $x + y = z$ is partition regular, whilst van der Waerden's theorem states that every finite colouring of the positive integers produces arbitrarily long monochromatic arithmetic progressions. A common generalisation of these theorems was obtained by Brauer [Bra28], who showed that every finite colouring of the positive integers produces arbitrarily long monochromatic configurations of the form $\{x, d, x+d, x+2d, \dots, x+(k-1)d\}$, with $d, k, x \in \mathbb{N}$.

Observe that partition regularity is a weaker property than density regularity. This is because for every finite partition of the positive integers $\mathbb{N} = C_1 \cup \cdots \cup C_r$, one of the C_i must have positive upper density and therefore contains non-constant solutions to every density regular system of equations. However, the converse is false: there exist systems of Diophantine equations which are partition regular but not density regular. For example, Schur's theorem informs us that $x + y = z$ is partition regular, but there are no solutions to this equation in odd numbers. The relationship between density regularity and partition regularity is further investigated in Chapter 2.

The first major classification of partition regularity for families of Diophantine equations was undertaken by Rado. In his PhD thesis [Rad33], Rado established necessary and sufficient conditions for a finite system of linear equations to be partition regular.

Theorem 1.1.1 ([Rad33, Satz IV]). *Let $\mathbf{M} = (a_{i,j})$ denote an $n \times s$ integer matrix of rank n with no zero columns. For each $j \in \{1, \dots, s\}$, let $\mathbf{c}^{(j)}$ denote the jth column of $\mathbf{M}$. The system of equations*

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,s}x_s &= 0; \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,s}x_s &= 0; \\ &\vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,s}x_s &= 0. \end{aligned} \tag{1.1}$$

is partition regular if and only if there exists a partition $\{1, 2, \dots, s\} = J_1 \cup \cdots \cup J_k$ such that $\sum_{j \in J_1} \mathbf{c}^{(j)} = \mathbf{0}$, and, for each $1 < t \leqslant k$,

$$\sum_{j \in J_t} \mathbf{c}^{(j)} \in \langle \mathbf{c}^{(r)} : r \in J_1 \cup \cdots \cup J_{t-1} \rangle_{\mathbb{Q}}.$$

Here, $\langle V \rangle_{\mathbb{Q}}$ denotes the $\mathbb{Q}$-linear span of a set of vectors V with rational entries. Matrices $\mathbf{M}$ which possess this latter property are said to obey the columns condition.

Since the work of Rado, there has been great interest in classifying partition regularity for other Diophantine systems, such as systems of infinitely many linear equations [HLS03, BHLS15], or non-linear systems [EG80, CGS12, DL18]. In this thesis, we are particularly interested in the kth power analogue of the above system:

$$\begin{aligned} a_{1,1}x_1^k + a_{1,2}x_2^k + \cdots + a_{1,s}x_s^k &= 0; \\ &\vdots \\ a_{n,1}x_1^k + a_{n,2}x_2^k + \cdots + a_{n,s}x_s^k &= 0. \end{aligned} \tag{1.2}$$

For the case of single equations ($n = 1$) there are numerous open problems, the most tantalising of which is the conjecture of Erdős and Graham [EG80] that $x^2 + y^2 = z^2$ is partition regular. Currently, this conjecture is only known to be true for colourings which use at most 2 colours, as shown in a computer-assisted proof by Heule, Kullman, and Marek [HKM16].

An immediate consequence of Rado's Theorem is that (1.2) is partition regular only if the coefficient matrix $\mathbf{M} = (a_{i,j})$ obeys the columns condition. This can be seen by noting that (1.2) is partition regular if and only if the corresponding linear system (1.1) admits monochromatic solutions with respect to any finite colouring of the kth powers $\{1, 2^k, 3^k, \ldots\}$ (see [Lef91, Theorem 2.1] for further details). Using non-standard analysis, Di Nasso and Luperi Baglini [DL18, Theorem 3.10] generalised this result by obtaining necessary conditions for partition regularity for polynomial equations $P(x_1, \ldots, x_s) = 0$, where P is a non-zero integer polynomial in s variables such that every monomial appearing in P contains exactly one variable. Barrett, Lupini, and Moreira [BLM21] have recently generalised these results further by establishing necessary conditions for partition regularity for general polynomial equations.

Obtaining corresponding sufficient conditions for systems of the form (1.2) to be partition regular is a far more delicate issue; one has to contend with additional number theoretic and geometric obstructions to the existence of solutions. For example, the system of equations

$$\begin{aligned} x^2 + z^2 &= 2y^2; \\ x^2 + d^2 &= y^2 \end{aligned}$$

obeys the columns condition, and yet, by Fermat's right triangle theorem (see [Con08]), it has no solutions over $\mathbb{N}$ and is therefore not partition regular. It was speculated in [DL18, Open Problem 1] that, for $n = 1$, in addition to the columns condition, the existence of positive integer solutions should be sufficient for (1.2) to be partition regular. However, this is insufficient, as demonstrated by the equation

$$x^5 - y^5 = 211z^5. \tag{1.3}$$

This equation obeys the columns condition and has solutions $(x, y, z) = (3\lambda, 2\lambda, \lambda)$ for any positive integer λ. However, Faltings's theorem from arithmetic geometry [Fal83] implies that there are only finitely many primitive solutions to (1.3), where a solution (x, y, z) is primitive if the greatest common divisor of x, y, z is 1. It follows that there exists a finite collection of positive rational numbers $q_1, \ldots, q_t > 1$ such that $\{x/y, x/z, y/x, y/z, z/x, z/y\} \cap \{q_1, \ldots, q_t\} \neq \emptyset$ holds for every positive integer solution (x, y, z) to (1.3). As a consequence of Rado's theorem, for each $i \in \{1, \ldots, t\}$, there exists a finite colouring of the positive integers with no monochromatic solution (u, v) to $u = q_i v$. Combining all of these finite colourings proves that (1.3) is not partition regular.

If we make additional non-singularity assumptions on the system (1.2), then we can use methods from analytic number theory to count solutions. By combining the Hardy-Littlewood circle method, Green's transference techniques to prove Roth's theorem in the primes [Gre05A], and restriction theory, Chow, Lindqvist, and Prendiville proved that, for $n = 1$, if s is sufficiently large in terms of k, then (1.2) is partition regular if and only if it obeys the columns condition. In Chapter 3, we develop these techniques further to obtain sufficient conditions for partition and density regularity for (1.2) for general n under certain non-singularity conditions.

Before concluding this section, we note that one can consider a quantitative analogue of partition regularity. When establishing partition regularity for a given system of Diophantine equations $\mathcal{E}$, one often finds that, for each positive integer r, there exists a positive integer $N = N_{\mathcal{E}}(r)$ such that every r-colouring $\{1, \ldots, N\} = C_1 \cup \cdots \cup C_r$ yields a monochromatic solution to $\mathcal{E}$. Indeed, the original works of Hilbert [Hil92], Schur [Sch16], and van der Waerden [Wae27] all deliver quantitative bounds of this type[1]. The minimum such $N_{\mathcal{E}}(r)$ is called the *r-colour Rado-Ramsey number for $\mathcal{E}$*. For further information on these numbers and their relationship with Ramsey numbers in graph theory, we refer the reader to [GRS90, §4].

[1] In fact, if $\mathcal{E}$ is partition regular, then the existence of such a $N_{\mathcal{E}}(r)$ is guaranteed by the *compactness principle*, which follows from the axiom of choice (see [FGR88, §1.5, Theorem 4]).

Using our earlier observation that density regularity implies partition regularity, we remark that quantitative bounds for Rado-Ramsey numbers can often be deduced from quantitative bounds for the corresponding density problems. For example, Gowers' bounds in Szemerédi's theorem [Gow01] imply that

$$\mathrm{W}(r,k) \leqslant 2^{2^{r^{2^{2^{k+9}}}}},$$

where $\mathrm{W}(r,k)$ (known as the *r-colour van der Waerden number*) is the r-colour Rado-Ramsey number for k-term arithmetic progressions. However, as observed previously, there does not always exist a corresponding density result for a given partition regular system. Consequently, one often needs to incorporate additional apparatus in order to apply density methods, such as higher order Fourier analysis, to quantitative colouring problems. A major theme of this thesis, particularly in Chapters 2 and 4, is the development of such apparatuses.

1.2 Results

1.2.1 Partition regularity and multiplicatively syndetic sets

We begin in Chapter 2 by elucidating the relationship between partition and density regularity through the mechanism of multiplicatively syndetic sets.

Definition (Multiplicatively syndetic set)**.** Let $S \subseteq \mathbb{N}$. Let $F \subset \mathbb{N}$ be a non-empty finite set. We say that S is *multiplicatively F-syndetic* if, for each $n \in \mathbb{N}$, we can find some $t \in F$ such that $nt \in S$. Equivalently, for every $n \in \mathbb{N}$, we have $(n \cdot F) \cap S \neq \emptyset$. We call $S \subseteq \mathbb{N}$ *multiplicatively syndetic* if S is multiplicatively F-syndetic for some non-empty finite set $F \subset \mathbb{N}$.

In Chapter 2, we investigate the connections between multiplicatively syndetic sets and colouring problems for systems of dilation invariant Diophantine equations. Here, a system of equations is *dilation invariant* if $(\lambda x_1, \ldots, \lambda x_s)$ is a solution to the system for every $\lambda \in \mathbb{Q}$ and solution $(x_1, \ldots, x_s)$.

Theorem 2.1.1. *Let $\mathcal{E}$ be a dilation invariant finite system of equations. Then $\mathcal{E}$ is partition regular if and only if $\mathcal{E}$ has a non-constant solution inside every multiplicatively syndetic set.*

As an immediate corollary, we show that every finite colouring of the positive integers admits monochromatic solutions to all dilation invariant partition regular systems which are all of the same colour.

Corollary 2.1.2. *Let $\mathcal{E}_1, \ldots, \mathcal{E}_s$ be s dilation invariant partition regular finite systems of equations. Then in any finite colouring $\mathbb{N} = C_1 \cup \cdots \cup C_r$ there exists a colour class C_t such that each $\mathcal{E}_i$ has a solution inside C_t.*

We then proceed to demonstrate that multiplicatively syndetic sets provide a framework for the adaptation of density methods to address colouring problems. The main result of Chapter 2 utilises quadratic Fourier analysis to obtain a 'multiplicatively syndetic' version of Brauer's theorem [Bra28] on monochromatic progressions with the same colour as their common difference.

Theorem 2.1.3. *There exists a positive absolute constant $c > 0$ such that the following is true. Let $S \subseteq \mathbb{N}$ be a multiplicatively F-syndetic set, for some non-empty finite set $F \subset \mathbb{N}$. Let M denote the largest element of F. If $N \geqslant 3$ satisfies*

$$M \leqslant \exp\left(c\sqrt{\log\log N}\right),$$

then there exists $d, x \in \mathbb{N}$ such that $\{x, d, x+d, x+2d\} \subseteq S \cap \{1, 2, \ldots, N\}$.

This theorem is an adaptation of a density result of Green and Tao [GT09]; if $A \subseteq \{1, 2, \ldots, N\}$ does not contain a non-trivial 4-term arithmetic progression, then

$$|A| \leqslant N \exp\left(-c\sqrt{\log\log N}\right).$$

Finally, we show that this quantitative syndeticity result may be iteratively applied to obtain a tower-type quantitative bound in Brauer's theorem for progressions of length 3. Let $\mathrm{B}(r, k)$ denote the minimum N (which exists by Brauer's theorem) such that every colouring of $\{1, \ldots, N\}$ with r colours produces a monochromatic configuration of the form $\{x, d, x+d, x+2d, \ldots, x+(k-1)d\}$. As we are interested in the case $k = 3$, we write $\mathrm{B}(r) := \mathrm{B}(r, 3)$ for all r. To state our result, we set $\mathrm{tow}(1) := 2$, and inductively define $\mathrm{tow}(n+1) := 2^{\mathrm{tow}(n)}$ for all $n \in \mathbb{N}$.

Theorem 2.1.4. *For each $r \in \mathbb{N}$,*

$$\mathrm{B}(r) \leqslant \mathrm{tow}\left((1 + o(1))r\right).$$

1.2.2 Partition regularity for systems of diagonal equations

In Chapter 3, we investigate the qualitative problem of finding necessary and sufficient conditions for a system of Diophantine equations to be partition or density regular. Our main result generalises the work of Chow, Lindqvist, and Prendiville [CLP21] to sufficiently non-singular systems of equations in kth powers.

Theorem 3.1.2. *Let $k, n, s \in \mathbb{N}$, with $k \geqslant 2$, and let $\mathbf{M} = (a_{i,j})$ be an $n \times s$ matrix with integer entries. Suppose that the following condition holds:*

(I) *for every non-empty set $\{\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)}\} \subseteq \mathbb{Q}^s$ of linearly independent non-zero vectors in the row space of $\mathbf{M}$, we have*

$$\left| \bigcup_{i=1}^{d} \operatorname{supp}\left(\mathbf{v}^{(i)}\right) \right| \geqslant dk^2 + 1.$$

Then the system of equations

$$\begin{aligned} a_{1,1}x_1^k + \cdots a_{1,s}x_s^k &= 0; \\ &\vdots \\ a_{n,1}x_1^k + \cdots a_{n,s}x_s^k &= 0 \end{aligned}$$

is non-trivially partition regular if and only if $\mathbf{M}$ obeys the columns condition.

We similarly classify density regularity for such systems.

Theorem 3.1.4. *Let $k, n, s \in \mathbb{N}$, with $k \geqslant 2$, and let $\mathbf{M} = (a_{i,j})$ be an $n \times s$ matrix with integer entries and no zero columns. Let $\delta > 0$. If $\mathbf{M}$ satisfies condition (I) of Theorem 3.1.2 and the columns of $\mathbf{M}$ sum to $\mathbf{0}$, then there exists a constant $c_1 = c_1(\delta, k, \mathbf{M}) > 0$ and a positive integer $N_1 = N_1(\delta, k, \mathbf{M}) \in \mathbb{N}$ such that the following is true. If $N \geqslant N_1$ and $A \subseteq \{1, 2, \ldots, N\}$ satisfies $|A| \geqslant \delta N$, then there are at least $c_1 N^{s-kn}$ non-trivial solutions $\mathbf{x} = (x_1, \ldots, x_s) \in A^s$ to (3.2).*

These theorems are proven by combining the argument of Chow, Lindqvist, and Prendiville with an arithmetic regularity lemma of Green [Gre05B]. The key new idea in our work is a decomposition lemma for systems of kth power equations satisfying condition (I). To state our decomposition lemma we require a few definitions. Let $\mathbb{Q}^{n \times s}$ and $\mathbb{Z}^{n \times s}$ denote the sets of $n \times s$ matrices with rational and integer entries respectively. To quantify the non-singularity properties of matrices, we introduce the μ and q functions.

Definition (μ and q functions)**.** Let $\mathbf{M} \in \mathbb{Q}^{n\times s}$ and $0 \leqslant d \leqslant n$. We write $\mu(d; \mathbf{M})$ to denote the largest number of columns of $\mathbf{M}$ whose $\mathbb{Q}$-linear span has dimension at most d. If $d \leqslant \mathrm{rank}(\mathbf{M})$, then we define

$$\mathrm{q}(d; \mathbf{M}) := \min \left| \bigcup_{i=1}^{d} \mathrm{supp}\left(\mathbf{v}^{(i)}\right) \right|,$$

where the minimum is taken over all collections of d linearly independent vectors $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(d)}$ in the row space of $\mathbf{M}$. By convention $\mathrm{q}(0; \mathbf{M}) = |\emptyset| = 0$.

Note that condition (I) from Theorem 3.1.4 is equivalent to the assertion that $\mathrm{q}(d; \mathbf{M}) > dk^2$ for all $1 \leqslant d \leqslant n$.

Given matrices $\mathbf{M}_i \in \mathbb{Q}^{n_i \times s_i}$ for $1 \leqslant i \leqslant r$, we say that a matrix $\mathbf{M} \in \mathbb{Q}^{n\times s}$ is *equivalent to a block upper triangular matrix with diagonal* $(\mathbf{M}_1, \dots, \mathbf{M}_r)$ if we can perform column permutations and elementary row operation to $\mathbf{M}$ to obtain a matrix of the form

$$\begin{pmatrix} \mathbf{M}_1 & \mathbf{A}^{(1,2)} & \dots & \mathbf{A}^{(1,r)} \\ \mathbf{0} & \mathbf{M}_2 & \dots & \mathbf{A}^{(2,r)} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{M}_r \end{pmatrix},$$

where $\mathbf{A}^{(i,j)} \in \mathbb{Q}^{n_i \times s_j}$ for all $1 \leqslant i < j \leqslant r$. Our decomposition lemma states that all matrices satisfying condition (I) are equivalent to a block upper triangular matrix with diagonal $(\mathbf{M}_1, \dots, \mathbf{M}_r)$ for some particularly non-singular matrices $\mathbf{M}_i$ which we have termed *quasi-partitionable*.

Definition (Quasi-partitionable matrix)**.** A non-empty $n \times s$ matrix $\mathbf{M}$ with rational entries is called *quasi-q-partitionable* if $s \geqslant nq$ and $\mu(d; \mathbf{M}) \leqslant dq$ holds for all $0 \leqslant d < n$.

Lemma 3.3.3. *Let $n, q, s \in \mathbb{N}$, and let $\mathbf{M} \in \mathbb{Q}^{n\times s}$ be a matrix of rank n with no zero columns. If $\mathrm{q}(d; \mathbf{M}) > dq$ for all $1 \leqslant d \leqslant n$, then $\mathbf{M}$ is equivalent to a block upper triangular matrix with diagonal $(\mathbf{M}_1, \dots, \mathbf{M}_r)$, where each $\mathbf{M}_i \in \mathbb{Z}^{n_i \times s_i}$ is quasi-q-partitionable and $s_i > n_i q$.*

The inspiration for this lemma comes from an article of Brüdern and Cook [BC92] in which it is noted that the circle method can be used to count solutions to a system of kth power equations whose coefficient matrix has the shape described in the conclusion of the above lemma. They did not use the terminology of quasi-partitionability, which is our own invention. This observation of Brüdern and Cook was previously noted by Low, Pitman, and Wolff [LPW88] who connected this structural property of matrices with results in matroid theory. Our work therefore demonstrates the efficacy of incorporating ideas from analytic number to study partition regularity for systems of Diophantine equations.

1.2.3 Ramsey numbers of Brauer configurations

Chapter 4, which is joint work with Sean Prendiville, returns to the problem of determining quantitative bounds for the Brauer numbers previously studied in Chapter 2. For each pair of positive integers k and r, recall that $\mathrm{B}(r,k)$ denotes the minimum positive integer N such that every colouring of $\{1,\ldots,N\}$ with r colours produces a monochromatic configuration of the form $\{x,d,x+d,x+2d,\ldots,x+(k-1)d\}$. Configurations of this form are called *Brauer configurations* in reference to the work of Brauer [Bra28] who established the existence of $\mathrm{B}(r,k)$ for every k and r. The main theorem of Chapter 4 proves that one can obtain an upper bound for $\mathrm{B}(r,k)$ which is double exponential in a power of the number of colours r, and quintuple exponential in the length of the progression k.

Theorem 4.1.1. *There exists an absolute constant $C=C(k)$ such that if $r\geqslant 2$ and $N\geqslant \exp\exp(r^C)$, then any r-colouring of $\{1,2,...,N\}$ yields a monochromatic k-term progression which is the same colour as its common difference. Moreover, it suffices to assume that*

$$N\geqslant 2^{2^{r^{2^{2^{k+10}}}}}.$$

This theorem is proven by modifying the higher order Fourier analytic techniques devised by Gowers [Gow01]. Gowers' main result shows that every set $A\subseteq\{1,\ldots,N\}$ with $|A|\geqslant \delta N$ contains a k-term arithmetic progression provided that

$$N\geqslant 2^{2^{(1/\delta)^{2^{2^{k+9}}}}}.$$

This implies a corresponding colouring result by taking $\delta = 1/r$. As noted earlier, density methods are not immediately applicable to configurations which are not translation-invariant, such as Brauer configurations. The main contribution of our work is the development of strategies to remove translation invariant assumptions from these methods.

Independently of our work, Sanders [San20] has obtained a generalisation of the above theorem to general linear configurations. Sanders proves that, for a given partition regular linear system of equations $\mathcal{E}$, there exists a constant $C_{\mathcal{E}} > 0$ such that the following is true. If $N \geqslant \exp\exp(r^{C_{\mathcal{E}}})$, then every r-colouring of $\{1, \ldots, N\}$ produces a monochromatic non-constant solution to $\mathcal{E}$. This generalises the double exponential colour bound in Theorem 4.1.1. As in our work, Sanders develops a novel modification of Gowers' density increment argument.

We observed in Chapter 2 that one could combine quadratic Fourier analysis with properties of multiplicatively syndetic sets to obtain an exponential tower-type bound for $\mathrm{B}(3, r)$. By incorporating these same quadratic Fourier analytic techniques, which come from work of Green and Tao [GT09] on arithmetic progressions of length 4, into our work, we improve upon the bound given in Theorem 4.1.1 for $\mathrm{B}(r, 3)$.

Theorem 4.1.2. *There exists an absolute constant C such that if $N \geqslant \exp\exp(Cr\log^2 r)$, then in any r-colouring of $\{1, 2, \ldots, N\}$ there exists a monochromatic three-term progression with the same colour as its common difference.*

Finally, we make use of an observation of Lefmann [Lef91] to obtain new quantitative bounds for the Rado-Ramsey numbers for certain non-linear configurations. Lefmann noted that certain diagonal quadric equations, which we term *Lefmann quadrics*, can be solved over sets of any set of the form $\{x, \lambda d, x + d, x + 2d, \ldots, x + (k-1)d\}$, for certain positive integers k and λ. Our work in Chapter 4 then leads to the following quantitative colouring result for Lefmann quadrics.

Theorem 4.1.3. *Let $a_1, \ldots, a_s \in \mathbb{Z} \setminus \{0\}$ satisfy the following:*

(i) *there exists a non-empty set $I \subset [s]$ such that $\sum_{i \in I} a_i = 0$;*

(ii) *the system*

$$x_0^2 \sum_{i \notin I} a_i + \sum_{i \in I} a_i x_i^2 = \sum_{i \in I} a_i x_i = 0.$$

has a rational solution with $x_0 \neq 0$.

Then there exists an absolute constant $C = C(a_1, \ldots, a_s)$ *such that for* $r \geqslant 2$ *and* $N \geqslant \exp\exp(r^C)$, *any* r*-colouring of* $\{1, 2, \ldots, N\}$ *yields a monochromatic solution to the diagonal quadric*

$$a_1x_1^2 + \cdots + a_sx_s^2 = 0.$$

1.3 Thesis format

This thesis is presented in the *Journal Format*; Chapters 2, 3 and 4 are based on pre-prints of papers that I have authored or co-authored which have subsequently been published (see the 'Publications' section above for further details). Each chapter contains its own introduction, bibliography, and notation sections. Therefore, each of these three chapters may be regarded as self-contained and be read independently of any other chapter of this thesis. I have used the Journal Format as it most clearly presents the contributions I have made to additive combinatorics and arithmetic Ramsey theory over the course of my PhD studies.

The versions of my papers incorporated into this thesis are the arXiv electronic pre-prints. In the case of Chapters 2 and 4, these pre-print versions are identical in mathematical content to the published papers, with only minor stylistic format adjustments made and updates to bibliographic references. The exposition in Chapter 3 differs slightly from that presented in the published paper; the pre-print version on which Chapter 3 is based is the "Author's Original Version". The mathematical content of the two versions do not differ significantly; the published paper (the "Version of Record") incorporates referee suggestions by including additional detail in certain proofs and a slightly extended introduction.

The content of Chapter 4 is joint work with Sean Prendiville. Both authors have agreed that our contribution to this chapter is equal. All other chapters are solely authored by myself.

There are two appendix chapters, Appendix A and Appendix B, which are both addenda to Chapter 2. Appendix A previously appeared in the paper on which Chapter 2 is based, whilst Appendix B previously appeared in an early pre-print version of that paper[2]. The purpose of Appendix A is to show how the tower-type bound in Theorem 2.1.4 may be extracted from the recursive bound given in Theorem 2.4.2. Appendix B recalls some of the technical definitions from [GT09] which are used throughout §2.4.3.

[2] Accessible at arXiv:1902.01149v1.

Bibliography

[BHLS15] B. Barber, N. Hindman, I. Leader, and D. Strauss, *Partition regularity without the columns property*, Proc. Amer. Math. Soc. **143** (2015), 3387–3399.

[BLM21] J. M. Barrett, M. Lupini, and J. Moreira, *On Rado conditions for nonlinear Diophantine equations*, European J. Combin. **94** (2021), no. 103277, 20 pp.

[Bas97] I. G. Bashmakova, *Diophantus and Diophantine equations*. Translated from the 1972 Russian original by Abe Shenitzer and updated by Joseph Silverman. The Dolciani Mathematical Expositions, **20**. Mathematical Association of America, Washington, DC, 1997. xiv+90 pp.

[Bra28] A. Brauer, *Über Sequenzen von Potenzresten*, Sitzungsber. Preuss. Akad. Wiss. (1928), 9–16.

[BC92] J. Brüdern and R. J. Cook, *On simultaneous diagonal equations and inequalities*, Acta Arith. **62** (1992), 125–149.

[CLP21] S. Chow, S. Lindqvist, and S. Prendiville, *Rado's criterion over squares and higher powers*, J. Eur. Math. Soc. (JEMS) **23** (2021), no. 6, 1925–1997.

[Con08] K. Conrad, *The congruent number problem*, Harvard College Mathematics Review, **2** (2008), 58–74.

[CGS12] P. Csikvári, K. Gyarmati, and A. Sárközy, *Density and Ramsey type results on algebraic equations with restricted solution sets*, Combinatorica **32** (2012), 425–449.

[DL18] M. Di Nasso and L. Luperi Baglini, *Ramsey properties of nonlinear Diophantine equations*, Adv. Math. **324** (2018), 84–117.

[EG80] P. Erdős and R. L. Graham, *Old and new problems and results in combinatorial number theory*, Vol. 28, Monographies de L'Enseignement Mathématique [Monographs of L'Enseignement Mathématique]. Université de Genève, L'Enseignement Mathématique, Geneva, 1980. 128 pp.

[ET36] P. Erdős and P. Turán, *On Some Sequences of Integers*, J. London Math. Soc. **11** (1936), no. 4, 261–264.

[Fal83] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), 349–366. Erratum in: Invent. Math. **75** (1984), 381.

[FGR88] P. Frankl, R. L. Graham, and V. Rödl, *Quantitative theorems for regular systems of equations*, J. Combin. Theory Ser. A **47** (1988), 246–261.

[Fur77] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.

[Gow01] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465–588.

[GRS90] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey Theory*, Second Edition, Wiley, New York, 1990. xii+196 pp.

[Gre05A] B J. Green, *Roth's theorem in the primes*, Ann. of Math. **161** (2005), 1609–1636.

[Gre05B] B. J. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), no. 2, 340–376.

[GT08] B. J. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), no. 2, 481–547.

[GT09] B. J. Green and T. Tao, *New bounds for Szemerédi's theorem. II. A new bound for* $r_4(N)$, Analytic number theory, Cambridge Univ. Press, Cambridge, 2009, pp. 180–204.

[HKM16] M. J. H. Heule, O. Kullman, and V. W. Marek, *Solving and verifying the Boolean Pythagorean triples problem via cube-and-conquer*, in: Theory and applications of satisfiability testing - SAT 2016, 228–245, Lecture Notes in Comput. Sci. **9710**, Springer, Cham, 2016.

[HLS03] N. Hindman, I. Leader, and D. Strauss, *Open problems in partition regularity*, Combin. Probab. Comput. **12** (2003), no. 5–6, 571–583.

[Hil92] D. Hilbert, *Über die Irreduzibilatatät ganzer rationaler Funktionen mit ganzzahligen Koeffizienten (On the irreducibility of entire rational functions with integer coefficients)*, J. Reine Angew. Math. **110** (1892), 104–129.

[Lef91] H. Lefmann, *On partition regular systems of equations*, J. Combin. Theory Ser. A **58** (1991), 35–53.

[LPW88] L. Low, J. Pitman, and A. Wolff, *Simultaneous diagonal congruences*, J. Number Theory **29** (1988), no. 1, 31–59.

[Pel20] S. Peluse, *Bounds for sets with no polynomial progressions*, Forum Math. Pi **8** (2020), e16, 55 pp.

[Pre17] S. Prendiville, *Quantitative bounds in the polynomial Szemerdi theorem: the homogeneous case*, Discrete Anal. (2017), Paper No. 5, 34 pp.

[Rad33] R. Rado, *Studien zur Kombinatorik*, Math. Z. **36** (1933), 242–280.

[Rot53] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 245–252.

[San20] T. Sanders, *Bootstrapping partition regularity of linear systems*, Proc. Edinb. Math. Soc. (2) **63** (2020), no. 3, 630–653.

[Sch16] I. Schur, *Uber die Kongruenz $x^m+y^m = z^m \pmod p$*, Jahresber. Dtsch. Math. Ver. **25** (1916), 114–117.

[Sze75] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.

[Tao07] T. Tao, *The dichotomy between structure and randomness, arithmetic progressions, and the primes*, in: International Congress of Mathematicians. Vol. I, 581–608, Eur. Math. Soc., Zürich, 2007.

[Wae27] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wisk. **15** (1927), 212–216.

Chapter 2

Partition regularity and multiplicatively syndetic sets

Abstract. We show how multiplicatively syndetic sets can be used in the study of partition regularity of dilation invariant systems of polynomial equations. In particular, we prove that a dilation invariant system of polynomial equations is partition regular if and only if it has a solution inside every multiplicatively syndetic set. We also adapt the methods of Green-Tao and Chow-Lindqvist-Prendiville to develop a syndetic version of Roth's density increment strategy. This argument is then used to obtain bounds on the Rado numbers of configurations of the form $\{x, d, x+d, x+2d\}$.

2.1 Introduction

A system of equations is called *partition regular* if, in any finite colouring of the positive integers $\mathbb{N} = C_1 \cup \cdots \cup C_r$, there exists a non-constant monochromatic solution $\mathbf{x} = (x_1, \ldots, x_s)$, meaning that $\mathbf{x} \in C_k^s$ for some k, and $x_i \neq x_j$ for some $i \neq j$. The foundational results in the study of partition regularity are the theorems of Schur [Sch16] and van der Waerden [Wae27]. Schur's theorem states that the equation $x+y = z$ is partition regular, whilst van der Waerden's theorem shows that any finite colouring of $\mathbb{N}$ yields arbitrarily long monochromatic (non-trivial) arithmetic progressions.

The theorems of Schur and van der Waerden are both examples of partition regularity being exhibited by certain linear systems of equations. In particular, these systems are *dilation invariant*, meaning that if $\mathbf{x} = (x_1, \ldots, x_s)$ is a solution, then so is $\lambda\mathbf{x} = (\lambda x_1, \ldots, \lambda x_s)$ for any $\lambda \in \mathbb{Q}$. In this chapter we study the properties of general dilation invariant systems of equations, not just those which are linear. We show that the regularity of such systems is inexorably connected with a special class of sets known as *multiplicatively syndetic sets*.

2.1.1 Syndeticity

Syndetic sets originate from the study of topological dynamics of semigroups (see [EEN00, HS12]). Given a semigroup $(G, \cdot)$, a set $S \subseteq G$ is called *(left)-syndetic* if there exists a finite set $F \subseteq G$ such that, for each $g \in G$, we have $S \cap (g \cdot F) \neq \emptyset$. Here $g \cdot F := \{gt : t \in F\}$.

The most familiar notion of syndeticity arises in the additive setting where $(G, \cdot) = (\mathbb{N}, +)$. In this case a syndetic subset S is called *additively syndetic* and is just an infinite set with 'bounded gaps'. That is, S is additively syndetic if and only if $S = \{a_1, a_2, \ldots\}$ for some infinite sequence $a_1 < a_2 < \ldots$ such that the gaps $|a_{n+1} - a_n|$ are uniformly bounded.

In this chapter , we study syndetic sets in the multiplicative semigroup $(\mathbb{N}, \cdot)$.

Definition (Multiplicatively syndetic set)**.** Let $F \subset \mathbb{N}$ be a non-empty finite set. We say that $S \subseteq \mathbb{N}$ is a *multiplicatively F-syndetic set* if, for every $a \in \mathbb{N}$, we have $S \cap (a \cdot F) \neq \emptyset$.

Multiplicatively syndetic sets possess a number of interesting properties. Graham, Spencer, and Witsenhausen [GSW77] observed that multiplicatively syndetic sets have positive density. Much later, Bergelson [Ber10, Lemma 5.11] used methods from ultrafilter theory to show that multiplicatively syndetic sets are *additively central*[1] (which implies that they have positive density).

The fact that multiplicatively syndetic sets have positive density plays a significant role in the work of Chow, Lindqvist and Prendiville [CLP21]. They demonstrate how multiplicatively syndetic sets can be used to obtain partition regularity results for non-linear equations via an "induction on colours" argument. Their work shows that a sufficient condition for a dilation invariant equation to be partition regular is that it

[1] A subset of $\mathbb{N}$ is called *additively central* if it is a member of a minimal idempotent ultrafilter on $(\mathbb{N}, +)$ (see [Ber10, Definition 5.8]).

has a solution inside[2] every multiplicatively syndetic set. Our first main theorem is a converse of this result.

Theorem 2.1.1 (Partition regularity is equivalent to syndetic solubility)**.** *Let $\mathcal{E}$ be a dilation invariant finite system of equations. Then $\mathcal{E}$ is partition regular if and only if $\mathcal{E}$ has a non-constant solution inside every multiplicatively syndetic set.*

As an immediate corollary to this theorem, we obtain the following *dilation invariant consistency theorem.*

Corollary 2.1.2 (Dilation invariant consistency theorem)**.** *Let $\mathcal{E}_1, \ldots, \mathcal{E}_s$ be s dilation invariant partition regular finite systems of equations. Then in any finite colouring $\mathbb{N} = C_1 \cup \cdots \cup C_r$ there exists a colour class C_t such that each $\mathcal{E}_i$ has a solution inside C_t.*

2.1.2 Brauer configurations

Van der Waerden [Wae27] proved that, for all $r, k \in \mathbb{N}$, there exists a (minimal) positive integer $W(r,k) \in \mathbb{N}$ such that, in any r-colouring of the set $\{1, \ldots, W(r,k)\}$, there exists a monochromatic arithmetic progression of length k. Obtaining good bounds for $W(r,k)$ is a notoriously difficult problem. Over 60 years after van der Waerden's original paper, Shelah [She88] obtained the first primitive recursive bounds. The best bounds currently known are due to Gowers [Gow01] who obtained the bound

$$\mathrm{W}(r,k) \leqslant 2^{2^{r^{2^{2^{k+9}}}}}. \tag{2.1}$$

In §2.4 we consider a variation of van der Waerden's theorem concerning configurations of the form

$$\{x, d, x+d, x+2d\}.$$

These are arithmetic progressions of length 3 along with their common difference. Brauer [Bra28] was the first to establish the partition regularity of these configurations, and so we refer to them as *Brauer configurations (of length 3).* We also call the corresponding Rado numbers the *(r-colour) Brauer numbers.* Specifically, we define

[2] A system $\mathcal{E}$ is said to have a solution in a set S if there exists a solution $\mathbf{x}$ to $\mathcal{E}$ with each entry of $\mathbf{x}$ lying in S.

$\mathrm{B}(r) \in \mathbb{N}$ to the the smallest positive integer such that every r-colouring of the interval $\{1, \ldots, \mathrm{B}(r)\}$ yields a monochromatic set of the form $\{x, d, x+d, x+2d\}$.

To show that Brauer configurations (of length 3) are partition regular, Theorem 2.1.1 informs us that it is sufficient to prove that all multiplicatively syndetic sets contain such configurations. Our next result establishes a quantitative version of Brauer's theorem for multiplicatively syndetic sets.

Theorem 2.1.3. *There exists a positive absolute constant $c > 0$ such that the following is true. Let $S \subseteq \mathbb{N}$ be a multiplicatively F-syndetic set, for some non-empty finite set $F \subset \mathbb{N}$. Let M denote the largest element of F. If $N \geqslant 3$ satisfies*

$$M \leqslant \exp\left(c\sqrt{\log\log N}\right),$$

then there exists $d, x \in \mathbb{N}$ such that $\{x, d, x+d, x+2d\} \subseteq S \cap \{1, 2, \ldots, N\}$.

This theorem is analogous to Green and Tao's result [GT09, Theorem 1.1] that sets $A \subseteq \{1, 2, \ldots, N\}$ which lack 4-term arithmetic progressions have size

$$|A| \leqslant N \exp\left(-c\sqrt{\log\log N}\right).$$

We deduce Theorem 2.1.3 from a more general density result (Theorem 2.4.1), which concerns dense sets $A \subseteq \{1, 2, \ldots, N\}$ lacking arithmetic progressions of length 3 with common difference lying in a given multiplicatively syndetic set S. This density result is proven in §2.4 by combining the methods of Green and Tao [GT09] with a 'multiplicatively syndetic induction on colours' argument of Chow, Lindqvist, and Prendiville [CLP21].

Brauer's theorem may be proved by iteratively applying van der Waerden's theorem. As indicated by Cwalina and Schoen [CS17], the best bounds one can obtain for the Brauer numbers by incorporating Gowers' bound (2.1) into this argument are of the form

$$\mathrm{B}(r) \leqslant \mathrm{tow}\left((5 + o(1))r\right).$$

Here $\mathrm{tow}(n)$ denotes an exponential tower of 2's of height n. Explicitly, we take $\mathrm{tow}(1) := 2$ and for all $n \geqslant 2$ define

$$\mathrm{tow}(n) := 2^{\mathrm{tow}(n-1)}.$$

By incorporating Theorem 2.1.3 into an induction on colours argument, we obtain an asymptotic improvement on this bound.

Theorem 2.1.4 (Tower bound for $\mathrm{B}(r)$)**.** *For each $r \in \mathbb{N}$,*

$$\mathrm{B}(r) \leqslant \mathrm{tow}\left((1+o(1))r\right). \tag{2.2}$$

In general, for a partition regular system of equations $\mathcal{E}$, one can define the *r-colour Rado number* $\mathrm{R}_{\mathcal{E}}(r)$ to be the smallest $N \in \mathbb{N}$ such that every r-colouring of the interval $\{1, \dots, N\}$ yields a monochromatic solution to $\mathcal{E}$. Cwalina and Schoen [CS17, Theorem 1.5] proved that if $\mathcal{E}$ is a partition regular homogeneous linear equation of the form

$$a_1x_1 + \cdots + a_sx_s = 0,$$

where $a_1, \dots, a_s \in \mathbb{Z} \setminus \{0\}$, then

$$\mathrm{R}_{\mathcal{E}}(r) \ll_{\mathcal{E}} 2^{O_{\mathcal{E}}(r^4 \log r)}.$$

The improvements obtained by Cwalina and Schoen for single equations ultimately derive from the fact that single linear equations are controlled by the U^2 norm (see §2.4 for a definition of the U^s norms), and so they can be analysed with (linear) Fourier analysis. However, Brauer configurations of length 3 are controlled by the U^3 norm and therefore require methods from quadratic Fourier analysis. In [CP20] we use higher order Fourier analysis to improve on Theorem 2.1.4 by obtaining a double exponential bound of the form

$$\mathrm{B}(r) \leqslant \exp\exp(r^C).$$

More generally, we show that a bound of the above form holds for Brauer configurations of any length k (with constant C depending on the length k).

Notation

The positive integers are denoted by $\mathbb{N}$. Given $X \geqslant 1$, we let $[X] := \{n \in \mathbb{N} : 1 \leqslant n \leqslant X\} = \{1, 2, \dots, \lfloor X \rfloor\}$.

Let f and g be positively valued functions. We write $f \ll g$, or $g \gg f$, or $f = O(g)$ if there exists a positive constant C such that $f(x) \leqslant Cg(x)$ for all x. If we require the constant C to depend on some parameters $\lambda_1, \dots, \lambda_k$, then we write $f \ll_{\lambda_1,\dots,\lambda_k} g$ or $f = O_{\lambda_1,\dots,\lambda_k}(g)$.

The letters c and C are typically used to denote absolute constants, whose values may change from line to line. We usually write c to denote a small constant $0 < c < 1$, whereas C usually denotes a large constant $C > 1$.

Acknowledgements

The author would like to thank Sean Prendiville for his constant support and encouragement, and for his helpful comments on an earlier draft of this paper. We also thank the anonymous referee for their comments and suggestions on an earlier version of the paper on which this chapter is based.

2.2 Multiplicative syndeticity and partition regularity

We begin by formally introducing the concepts of partition regularity and multiplicative syndeticity mentioned in the introduction. After establishing the basic properties of multiplicatively syndetic sets, we prove Theorem 2.1.1 and Corollary 2.1.2.

2.2.1 Systems of equations

We consider finite systems of polynomial equations $\mathcal{E}$ in $s \in \mathbb{N}$ variables of the form

$$\begin{aligned} p_1(t_1, t_2, \dots, t_s) &= 0; \\ p_2(t_1, t_2, \dots, t_s) &= 0; \\ &\vdots \\ p_k(t_1, t_2, \dots, t_s) &= 0, \end{aligned} \tag{2.3}$$

where each $p_i \in \mathbb{Q}[t_1, t_2, \dots, t_s]$ is a polynomial in the variables $\{t_i\}_{i=1}^{s}$. In this chapter we only consider systems of finitely many equations, each with finitely many variables. For related results concerning the regularity of infinite systems, see [BHLS15, HLS03].

We usually refer to such a system of polynomial equations $\mathcal{E}$ simply as a *system of equations*. We call $\mathbf{x} \in \mathbb{N}^s$ a *solution* to the system $\mathcal{E}$ if $p_i(\mathbf{x}) = 0$ for all i, meaning that $\mathbf{x}$ is a solution to all of the equations in $\mathcal{E}$ simultaneously. A solution $\mathbf{x} = (x_1, \dots, x_s)$ is called a *non-constant solution* if the entries of $\mathbf{x}$ are not all equal, meaning that

$x_i \neq x_j$ for some $i \neq j$. Given a set $S \subseteq \mathbb{Q}$, we say that $\mathcal{E}$ has a (non-constant) solution in S if there is a (non-constant) solution $\mathbf{x} = (x_1, \ldots, x_s) \in \mathbb{N}^s$ to $\mathcal{E}$ such that $x_i \in S$ for all i.

A system of equations $\mathcal{E}$ is called *dilation invariant* if the following is true. If $\mathbf{x} = (x_1, \ldots, x_s)$ is a solution to $\mathcal{E}$, then $\lambda\mathbf{x} = (\lambda x_1, \ldots, \lambda x_s)$ is also a solution for every $\lambda \in \mathbb{Q}$. For the majority of this chapter , we restrict our attention to dilation invariant systems of polynomial equations. However it should be noted that most of the results we prove in this section apply to any dilation invariant system of equations and not just those consisting of polynomial equations.

2.2.2 Partition regularity

As mentioned in the introduction, the partition regularity of equations is a well-studied topic in Ramsey theory. Recall that an *r-colouring* of a set X is a partition $X = C_1 \cup \cdots \cup C_r$ of X into r *colour classes* C_i. Equivalently, an r-colouring can be defined by a function $\chi : X \to A$, for some set $A = \{a_1, \ldots, a_r\}$ with $|A| = r$ (usually we take $A = [r]$). These two characterisations can be seen to be equivalent by taking $\chi^{-1}(a_i) = C_i$. A subset $Y \subseteq X$ is called (χ)-*monochromatic* if χ is constant on Y, or equivalently that $Y \subseteq C_i$ for some colour class C_i.

Definition (Partition regularity)**.** Let $S \subseteq \mathbb{Q}$ be a non-empty set and let $\mathcal{E}$ be a system of equations with coefficients in $\mathbb{Q}$. Let $r \in \mathbb{N}$. We say that $\mathcal{E}$ is *(kernel) r-regular over S* if, for each r-colouring $\chi : \mathbb{Q} \to [r]$, there exists a χ-monochromatic non-constant solution $\mathbf{x}$ to $\mathcal{E}$ with entries in S. We call such an $\mathbf{x}$ a (χ-)*monochromatic (non-constant) solution* to A. We say that $\mathcal{E}$ is *(kernel) partition regular over S* if $\mathcal{E}$ is r-regular over S for every $r \in \mathbb{N}$.

In practice, when one shows that a given system of equations $\mathcal{E}$ is r-regular, the proof actually yields a number $\mathrm{R}_{\mathcal{E}}(r)$ (known as the r *colour Rado number* for $\mathcal{E}$) such that $\mathcal{E}$ is r-regular over the finite interval $[\mathrm{R}_{\mathcal{E}}(r)]$. This is certainly the case whenever one obtains a quantative regularity result, such as in [CS17, Gow01, Sch16, Wae27]. By assuming (some form of) the axiom of choice, one can show that if $\mathcal{E}$ is r-regular, then such an $\mathrm{R}_{\mathcal{E}}(r)$ necessarily exists. This result is known as the compactness principle.

Compactness Principle. *Let $\mathcal{E}$ be a finite system of equations in finitely many* *variables. Let $A \subseteq \mathbb{N}$ and let $r \in \mathbb{N}$. Then $\mathcal{E}$ is r-regular over A if and only if there exists a finite set $F \subseteq A$ such that $\mathcal{E}$ is r-regular over F.*

Proof. See [GRS90, Theorem 4]. □

Remark. For the rest of this section, we assume (some form of) the axiom of choice in order to make use of the compactness principle. This assumption is not required for any of the remaining sections.

2.2.3 Multiplicatively thick sets

The compactness principle informs us that a system of equations $\mathcal{E}$ is partition regular if and only if, for each $r \in \mathbb{N}$, we can find a finite set $F_r \subset N$ such that $\mathcal{E}$ is r-regular over F_r. Thus, a sufficient condition for $\mathcal{E}$ to be partition regular over a set $A \subseteq \mathbb{N}$ would be that $F_r \subseteq A$ for all $r \in \mathbb{N}$. The problem with this condition is that it is quite possible that the only set which could satisfy this property is $A = \mathbb{N}$. If $\mathcal{E}$ is a dilation invariant system of equations, then we can relax this condition to the requirement that, for each $r \in \mathbb{N}$, we can find $t_r \in \mathbb{N}$ such that $t_r \cdot F_r \subseteq A$. This motivates the following definition.

Definition (Multiplicatively thick set)**.** Let $T \subseteq \mathbb{N}$. We say that T is a *multiplicatively thick set* if, for each finite set $F \subset \mathbb{N}$, there exists $t \in \mathbb{N}$ such that $t \cdot F \subseteq T$.

Proposition 2.2.1 (Regularity over thick sets)**.** *Let $\mathcal{E}$ be a dilation invariant system of equations. Let $r \in \mathbb{N}$. Then the following are all equivalent:*

(I) *$\mathcal{E}$ is r-regular;*

(II) *$\mathcal{E}$ is r-regular over every multiplicatively thick set;*

(III) *$\mathcal{E}$ is r-regular over some multiplicatively thick set T.*

Proof. The implications (II)⇒(III) and (III)⇒(I) are immediate. It only remains to establish (I)⇒(II).

Suppose $\mathcal{E}$ is r-regular. By compactness, we can find a finite set $F \subset \mathbb{N}$ such that $\mathcal{E}$ is r-regular over F. Now let $T \subseteq \mathbb{N}$ be a multiplicatively thick set. We can then find $t \in \mathbb{N}$ such that $t \cdot F \subseteq T$. Now suppose $\chi : T \to [r]$ is an r-colouring of T. Define a new r-colouring $\tilde{\chi} : F \to [r]$ of F by $\tilde{\chi}(x) = \chi(tx)$. Since $\mathcal{E}$ is r-regular over F, we can find a $\tilde{\chi}$-monochromatic solution $\mathbf{x}$ to $\mathcal{E}$ in F. By dilation invariance, we deduce that $t\mathbf{x}$ is a χ-monochromatic solution to $\mathcal{E}$ in T. Thus $\mathcal{E}$ is r-regular over T. $\square$

2.2.4 Multiplicatively syndetic sets

We have now reduced regularity over $\mathbb{N}$ to regularity over a multiplicatively thick set. The utility of Proposition 2.2.1 is demonstrated in the following argument. Suppose that we have a dilation invariant system of equations $\mathcal{E}$ and an integer $r > 1$ such that $\mathcal{E}$ is $(r-1)$-regular. We would like to use this to test whether $\mathcal{E}$ is r-regular. Suppose that we have an r-colouring $\mathbb{N} = C_1 \cup \cdots \cup C_r$. Informally, if we know that one of the colour classes C_j is 'small', then we would expect, by $(r-1)$-regularity, to find a monochromatic solution in a colour class C_i with $i \neq j$.

To make this rigorous, suppose that we have a colour class C_j such that the complement $\mathbb{N} \setminus C_j$ is multiplicatively thick. The remaining $(r-1)$ colour classes induce an $(r-1)$-colouring on $\mathbb{N} \setminus C_j$. By Proposition 2.2.1, since $\mathcal{E}$ is $(r-1)$-regular, we can find a monochromatic solution to $\mathcal{E}$ inside $\mathbb{N} \setminus C_j$.

This shows that if the dilation invariant system $\mathcal{E}$ is $(r-1)$-regular but not r-regular, then there is an r-colouring $\mathbb{N} = C_1 \cup \cdots C_r$ without non-constant monochromatic solutions to $\mathcal{E}$ such that each complement $\mathbb{N} \setminus C_i$ is not multiplicatively thick. Observe that $\mathbb{N} \setminus C_i$ is not multiplicatively thick if and only if there exists a finite set $F \subset \mathbb{N}$ such that, for every $n \in \mathbb{N}$, we have $(n \cdot F) \cap C_i \neq \emptyset$. This motivates the following definition.

Definition (Multiplicatively syndetic set)**.** Let $S \subseteq \mathbb{N}$. Let $F \subset \mathbb{N}$ be a non-empty finite set. We say that S is *multiplicatively F-syndetic* if, for each $n \in \mathbb{N}$, we can find some $t \in F$ such that $nt \in S$. Equivalently, for every $n \in \mathbb{N}$, we have $(n \cdot F) \cap S \neq \emptyset$. We call $S \subseteq \mathbb{N}$ *multiplicatively syndetic* if S is multiplicatively F-syndetic for some non-empty finite set $F \subset \mathbb{N}$.

Remark. Chow, Lindqvist, and Prendiville [CLP21] define an *M-homogeneous set* to be a set which intersects every homogeneous arithmetic progression $x \cdot [M]$ of length M for every $x \in \mathbb{N}$. We therefore observe that an M-homogeneous set is exactly the same as a multiplicatively $[M]$-syndetic set.

As mentioned previously, multiplicatively syndetic sets can be equivalently defined in terms of multiplicatively thick sets.

Proposition 2.2.2. *Let $S \subseteq \mathbb{N}$. Then the following are all equivalent:*

(I) *S is multiplicatively syndetic;*

(II) *for every multiplicatively thick set T, we have $S \cap T \neq \emptyset$;*

(III) *$\mathbb{N} \setminus S$ is not multiplicatively thick.*

Proof.

(I)⇒(II): Suppose S is multiplicatively F-syndetic for some $F \subset \mathbb{N}$, and suppose $T \subseteq \mathbb{N}$ is a multiplicatively thick set. This means that we can find $t_T \in \mathbb{N}$ such that $t_T \cdot F \subseteq T$. Since S is multiplicatively F-syndetic, we have $(t_T \cdot F) \cap S \neq \emptyset$. In particular, $S \cap T \neq \emptyset$.

(II)⇒(III): Follows from the fact that S and $\mathbb{N} \setminus S$ are disjoint.

(III)⇒(I): Since $\mathbb{N} \setminus S$ is not multiplicatively thick, we can find a non-empty finite set $F \subset \mathbb{N}$ such that $t \cdot F \nsubseteq \mathbb{N} \setminus S$ for every $t \in \mathbb{N}$. This implies that S is multiplicatively F-syndetic. □

In Proposition 2.2.1 we showed that a dilation invariant system of equations is r-regular if and only if it is r-regular over all multiplicatively thick sets. This is a consequence of the 'largeness' of multiplicatively thick sets. We now prove a similar result for multiplicatively syndetic sets. To do this, we identify multiplicatively syndetic sets with finite colourings in the following manner.

Definition (Encoding function)**.** Let $S \subseteq \mathbb{N}$ be a multiplicatively F-syndetic set, for some non-empty finite $F \subset \mathbb{N}$. The *encoding function for* (S, F) is the function $\tau_{S;F} : \mathbb{N} \to F$ defined by

$$\tau_{S;F}(n) := \min\{t \in F : nt \in S\}.$$

Note that the assertion that $\tau_{S;F}$ is a well-defined total function is equivalent to the statement that S is multiplicatively F-syndetic.

The encoding function $\tau_{S;F}$ defines a finite colouring of $\mathbb{N}$. Moreover, if a set A is monochromatic with respect to this colouring, then there exists some $t \in F$ such that $t \cdot A \subseteq S$. This observation leads to the following result.

Proposition 2.2.3 (Syndetic sets contain PR configurations)**.** *Let $A \subseteq \mathbb{N}$. Let $S \subseteq \mathbb{N}$ be a multiplicatively F-syndetic set, for some non-empty finite set $F \subset \mathbb{N}$. Let $\mathcal{E}$ be a dilation invariant system of equations, and let $r \in \mathbb{N}$. If $\mathcal{E}$ is $(|F| \cdot r)$-regular over A, then $\mathcal{E}$ is r-regular over $S \cap (F \cdot A)$.*

Proof. Suppose $\chi : S \cap (F \cdot A) \to [r]$ is an r-colouring. Let $\tau = \tau_{S;F}$. Now let $\tilde{\chi} : A \to F \times [r]$ be the product colouring given by

$$\tilde{\chi}(n) := (\tau(n), \chi(n\tau(n))).$$

Since $\mathcal{E}$ is $(|F| \cdot r)$-regular over A, we can find a $\tilde{\chi}$-monochromatic solution $\mathbf{a}$ to $\mathcal{E}$ whose entries a_i all lie in A. From the definition of $\tilde{\chi}$, we can find $t \in F$ such that $\tau(a_i) = t$ for each entry a_i. From the dilation invariance of $\mathcal{E}$, we deduce that $\mathbf{x} := t\mathbf{a}$ is a χ-monochromatic solution to $\mathcal{E}$ whose entries $x_i = ta_i$ all lie in $S \cap (F \cdot A)$. □

This proposition immediately gives the following corollary.

Corollary 2.2.4. *Let $\mathcal{E}$ be a dilation invariant system of equations. Then the following are all equivalent:*

(I) *$\mathcal{E}$ is partition regular (over $\mathbb{N}$);*

(II) *$\mathcal{E}$ is partition regular over every multiplicatively syndetic set;*

(III) *$\mathcal{E}$ is partition regular over some multiplicatively syndetic set $S \subseteq \mathbb{N}$.*

We have thus shown that partition regularity over $\mathbb{N}$ is equivalent to partition regularity over a particular multiplicatively syndetic set. Our goal now is to prove Theorem 2.1.1 and therefore show that partition regularity over $\mathbb{N}$ is actually equivalent to 1-regularity over every multiplicatively syndetic set.

Recall that our motivation for introducing multiplicatively syndetic sets came from considering colourings in which some of the colour classes were not multiplicatively thick. This leads to the following induction argument first developed in [CLP21] to establish partition regularity of certain non-linear dilation invariant equations.

Lemma 2.2.5 (Induction on colours schema)**.** *Let $\mathcal{E}$ be a dilation invariant system of equations. If $\mathcal{E}$ is r-regular (for some $r \in \mathbb{N}$), then there exists a finite set $F =$*

$F(\mathcal{E}, r) \subset \mathbb{N}$ so that the following holds. If $\mathbb{N} = C_1 \cup \cdots \cup C_{r+1}$ is an $(r+1)$-colouring which lacks monochromatic solutions to $\mathcal{E}$, then each colour class C_i must be a multiplicatively F-syndetic set.

Proof. Since $\mathcal{E}$ is r-regular, the compactness principle allows us to find a non-empty finite set $F = F(\mathcal{E}, r) \subset \mathbb{N}$ such that $\mathcal{E}$ is r-regular over F. By dilation invariance, in any colouring χ of $\mathbb{N}$, if there exists a set of the form $x \cdot F$ (with $x \in \mathbb{N}$) which receives at most r distinct colours, then there exists a χ-monochromatic solution to $\mathcal{E}$ in $x \cdot F$. By contraposition we deduce that if $\mathbb{N} = C_1 \cup \cdots \cup C_{r+1}$ is an $(r+1)$-colouring which lacks monochromatic solutions to $\mathcal{E}$, then each colour class C_i is a multiplicatively F-syndetic set. $\square$

This lemma shows that when we are trying to prove that a given dilation invariant system $\mathcal{E}$ is partition regular, we only need to consider colourings in which all of the colour classes are multiplicatively syndetic. Combining this with Corollary 2.2.4 allows us to prove Theorem 2.1.1.

Proof of Theorem 2.1.1. If $\mathcal{E}$ is partition regular, then Corollary 2.2.4 implies that $\mathcal{E}$ is partition regular over every multiplicatively syndetic set. In particular, $\mathcal{E}$ has a non-constant solution inside every multiplicatively syndetic set.

Conversely, suppose $\mathcal{E}$ is not partition regular. If $\mathcal{E}$ is not 1-regular, then $\mathcal{E}$ has no non-constant solutions in the multiplicatively syndetic set $\mathbb{N}$. Suppose then that $\mathcal{E}$ is 1-regular. By Lemma 2.2.5, there exists a finite colouring of $\mathbb{N}$ with no monochromatic non-constant solutions to $\mathcal{E}$ and with each colour class being a multiplicatively syndetic set. Therefore each colour class is a multiplicatively syndetic set which has no non-constant solutions to $\mathcal{E}$. $\square$

This result therefore reduces the task of establishing r-regularity over $\mathbb{N}$ for every $r \in \mathbb{N}$ to establishing solubility in every multiplicatively syndetic set. Whilst this may not immediately appear to be helpful, we can obtain Corollary 2.1.2 very easily from this new approach.

Proof of Corollary 2.1.2. For each $k \in [s]$, let m_k denote the number of variables appearing in the equations defining the system $\mathcal{E}_k$. We can therefore define a dilation invariant system $\mathcal{E}$ in $m = m_1 + \cdots + m_s$ variables whose solutions are precisely tuples of the form $(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(s)})$, where $\mathbf{x}^{(k)} \in \mathbb{Q}^{m_k}$ is a solution to the system $\mathcal{E}_k$. Since each $\mathcal{E}_i$ is partition regular, it follows from Theorem 2.1.1 that $\mathcal{E}$ is partition regular. This implies the desired result. □

Remark. In the case that each $\mathcal{E}_i$ is a partition regular linear homogeneous equation, the above result is an immediate consequence of Rado's Criterion [Rad33, Satz IV]. Our proof shows that it is not necessary to utilise such a strong result.

2.2.5 Multiplicatively piecewise syndetic sets

By using encoding functions, one can show that for a non-empty finite set $F \subset \mathbb{N}$, a set $S \subseteq \mathbb{N}$ is multiplicatively F-syndetic if and only if

$$\mathbb{N} = \bigcup_{t \in F} t^{-1} S,$$

where $t^{-1}S := \{n \in \mathbb{N} : nt \in S\}$. Our proof of Proposition 2.2.3 used this fact to show that regularity over $\mathbb{N}$ can be 'lifted' to regularity over a multiplicatively syndetic set. However, we proved in Proposition 2.2.1 that regularity over $\mathbb{N}$ is equivalent to regularity over a multiplicatively thick set. This motivates the introduction of the following weaker form of syndeticity.

Definition (Multiplicatively piecewise syndetic set)**.** Let $F \subset \mathbb{N}$ be a non-empty finite set, and let $S \subseteq \mathbb{N}$. We say that S is *multiplicatively piecewise F-syndetic* if the set $\cup_{t \in F} (t^{-1}S)$ is a multiplicatively thick set.
We call $S \subseteq \mathbb{N}$ *(multiplicatively) piecewise syndetic* if S is multiplicatively piecewise F-syndetic for some $F \subset \mathbb{N}$.

Another way to view multiplicatively piecewise syndetic sets is through the following 'partial encoding' formulation. Given a non-empty finite set $F \subset \mathbb{N}$ and a set $S \subseteq \mathbb{N}$, define a partial function[3] $\tau_{S;F} : \mathbb{N} \nrightarrow F$ by

$$\tau_{S;F}(n) := \min\{t \in F : nt \in S\}, \tag{2.4}$$

for all $n \in \mathbb{N}$ for which the above quantity is defined. We refer to this partial function as the *(partial) encoding function for* (S, F). By the *domain of* $\tau_{S;F}$ we mean the set of all $n \in \mathbb{N}$ for which (2.4) is defined.

[3] We use the partial function notation $f : A \nrightarrow B$ to mean that f only defines a function on a (possibly empty) subset of A.

We remarked earlier that S is multiplicatively F-syndetic if and only if $\tau_{S;F}$ is a total function, meaning that $\tau_{S;F}(n)$ is defined for all $n \in \mathbb{N}$. Similarly, we see that S is multiplicatively piecewise F-syndetic if and only if the domain of $\tau_{S;F}$ is multiplicatively thick.

This technique of identifying a multiplicatively syndetic set with its encoding function was the key idea in the proof of Proposition 2.2.3. A similar argument can be used to obtain the following analogous result.

Proposition 2.2.6 (Piecewise syndetic sets contain PR configurations)**.** *Let $S \subseteq \mathbb{N}$ be a multiplicatively piecewise F-syndetic set, for some non-empty finite set $F \subset \mathbb{N}$. Let $\mathcal{E}$ be a dilation invariant system of equations, and let $r \in \mathbb{N}$. If $\mathcal{E}$ is $(|F| \cdot r)$-regular over $\mathbb{N}$, then $\mathcal{E}$ is r-regular over S.*

Proof. Suppose $\chi : S \to [r]$ is an r-colouring of S. Since S is multiplicatively piecewise syndetic, the set $T := \cup_{t \in F} (t^{-1} \cdot S)$ is multiplicatively thick. Let $\tau : T \to F$ denote the encoding function given by

$$\tau(n) := \min\{t \in F : nt \in S\}.$$

Now let $\tilde{\chi} : T \to F \times [r]$ be the product colouring given by

$$\tilde{\chi}(n) := (\tau(n), \chi(n\tau(n))).$$

By Proposition 2.2.1, we know that $\mathcal{E}$ is $(|F| \cdot r)$-regular over T. Thus, we can find a $\tilde{\chi}$-monochromatic solution $\mathbf{x}$ to $\mathcal{E}$ such that every entry of $\mathbf{x}$ is an element of T. From the definition of $\tilde{\chi}$, we can find $t \in F$ such that $\tau(x_i) = t$ for every entry x_i of $\mathbf{x}$. The dilation invariance of $\mathcal{E}$ then shows that $\mathbf{y} := t\mathbf{x}$ is a χ-monochromatic solution to $\mathcal{E}$ whose entries $y_i = tx_i$ all lie in S. $\square$

We end this section by synthesising all of our results relating partition regularity with solubility in multiplicatively syndetic sets into the following theorem.

Theorem 2.2.7 (Summary of results)**.** *Suppose $\mathcal{E}$ is a dilation invariant system of equations. Then the following are all equivalent:*

(I) *$\mathcal{E}$ is partition regular (over $\mathbb{N}$);*

(II) *$\mathcal{E}$ is partition regular over every multiplicatively thick set;*

(III) $\mathcal{E}$ *is partition regular over every multiplicatively piecewise syndetic set;*

(IV) $\mathcal{E}$ *is partition regular over every multiplicatively syndetic set;*

(V) $\mathcal{E}$ *has a solution in every multiplicatively piecewise syndetic set;*

(VI) $\mathcal{E}$ *has a solution in every multiplicatively syndetic set.*

Proof. Proposition 2.2.1 establishes the equivalence (I)⇔(II). Proposition 2.2.3 and Theorem 2.1.1 together show that (I)⇔(IV)⇔(VI). Similarly, we deduce from Proposition 2.2.6 that (I)⇔(III). Since syndeticity implies piecewise syndeticity, we see that (V)⇒(VI). Finally, since solubility is equivalent to 1-regularity, we observe that (III)⇒(V). □

2.3 Density of multiplicatively syndetic sets

In Ramsey theory, there are multiple concepts of 'largeness'. The most familiar of these is the notion of (asymptotic) density.

Definition (Asymptotic Density)**.** Let $A \subseteq \mathbb{N}$. The *upper (asymptotic) density* $\bar{d}(A)$ of A is defined by

$$\bar{d}(A) := \limsup_{N\to\infty} \frac{|A \cap [N]|}{N}.$$

The *lower (asymptotic) density* $\underline{d}(A)$ of A is defined by

$$\underline{d}(A) := \liminf_{N\to\infty} \frac{|A \cap [N]|}{N}.$$

The *natural (asymptotic) density* $d(A)$ of A is defined by

$$d(A) := \lim_{N\to\infty} \frac{|A \cap [N]|}{N},$$

whenever the above limit exists, which occurs if and only if $\bar{d}(A) = \underline{d}(A)$.

One can generalise this definition to obtain a notion of asymptotic density for general cancellative, left amenable semigroups (see [BG18] for further details). The above definition comes from the case where the semigroup in question is $(\mathbb{N}, +)$. As such, asymptotic density is a form of 'additive largeness'.

Recent research has led to the surprising discovery that multiplicatively large sets, such as multiplicatively syndetic sets, are additively large. Bergelson [Ber10, Lemma

5.11] proved that multiplicatively syndetic sets are additively central, which implies that they have positive upper asymptotic density. However, due to the infinitary nature of central sets, no explicit bounds on the density of multiplicatively syndetic sets can be extracted from this result.

In their work on the partition regularity of non-linear equations, Chow, Lindqvist, and Prendiville [CLP21, Lemma 4.2] independently proved that multiplicatively syndetic sets have positive (lower) asymptotic density. They obtained the following quantitative result (for the case $F = [M]$).

Lemma 2.3.1. *Let $F \subset \mathbb{N}$ be a non-empty finite set, and let M denote the largest element of F. Then for any $N \in \mathbb{N}$ and any multiplicatively F-syndetic set $S \subseteq [N]$, we have*

$$|S \cap [N]| \geqslant \frac{1}{|F|} \left\lfloor \frac{N}{M} \right\rfloor . \tag{2.5}$$

Proof. Define an encoding function $\tau : [N/M] \to F$ for S by

$$\tau(x) := \min\{t \in F : tx \in S\}.$$

By the pigeonhole principle, there exists $t \in F$ such that $|\tau^{-1}(t)| \geqslant \frac{1}{|F|}|[N/M]|$. Thus,

$$|S| \geqslant |\{tx : x \in \tau^{-1}(t)\}| \geqslant \frac{1}{|F|} \left\lfloor \frac{N}{M} \right\rfloor .$$

□

In fact, the density of multiplicatively syndetic sets had been studied much earlier. In 1977 Graham, Spencer, and Witsenhausen [GSW77] determined the maximum asymptotic density for sets lacking linear configurations of the form $\{a_1x, a_2x, \ldots, a_sx\}$. Taking complements enables one to determine the minimum density of a multiplicatively F-syndetic set for $F = \{a_1, \ldots, a_s\}$. After performing this reformulation, their result is as follows.

Theorem 2.3.2 ([GSW77, Theorem 2])**.** *Let $F \subset \mathbb{N}$ be a non-empty finite set. Let P_F be the set of primes dividing elements of F. Let $\mathrm{S}(P_F)$ denote the set of all P_F-*smooth numbers*, meaning that $x \in \mathrm{S}(P_F)$ if and only if every prime factor of x lies in P_F. We write $\mathrm{S}(P_F) = \{d_1 < d_2 < d_3 < \ldots\}$, where d_k is the kth smallest element of $\mathrm{S}(P_F)$. For each $k \in \mathbb{N}$, let*

$$g_F(k) := \min\{|X \cap \{d_1, \ldots, d_k\}| : X \subseteq \mathbb{N} \textit{ is multiplicatively } F\textit{-syndetic}\}$$

Finally, let $\mathrm{K}(F) = \{k \in \mathbb{N} : g_F(k) = g_F(k-1)\}$*. Then for any multiplicatively* F*-syndetic set* $S \subseteq \mathbb{N}$*, we have the sharp bound*

$$\underline{d}(S) \geqslant \delta_{min}(F) := 1 - \prod_{p \in \mathrm{S}(P_F)} (1 - p^{-1}) \sum_{k \in \mathrm{K}(F)} d_k^{-1}.$$

Graham, Spencer, and Witsenhausen remark that there are difficulties in evaluating $\delta_{\min}(F)$ due to the complicated nature of the set $\mathrm{K}(F)$. In particular, they could not obtain an explicit evaluation for $\delta_{\min}(F)$ in the case where $F = \{1, 2, 3\}$. Erdős and Graham [EG80] subsequently conjectured that $\delta_{\min}(\{1, 2, 3\})$ is irrational. This conjecture remains open (see [CEG02] for further details and developments related to this problem).

In the case where $F = \{1, p, p^2, \ldots, p^{k-1}\}$ for some prime p and some $k \in \mathbb{N}$, Graham, Spencer, and Witsenhausen observed that $\delta_{\min}(F) = \frac{p+1}{p^k - 1}$. We now consider $F = \{1, a, a^2, \ldots, a^{k-1}\}$, where $a \in \mathbb{N}$ need not be prime, and explicitly construct a multiplicatively F-syndetic set of minimum density.

Definition (Multiplicity function)**.** Let $a \geqslant 2$ be a positive integer. Define the *a-multiplicity function* $\nu_a : \mathbb{N} \to \mathbb{N} \cup \{0\}$ by

$$\nu_a(n) = \max\{k \in \mathbb{N} \cup \{0\} : a^k | n\}.$$

Lemma 2.3.3 (The set $S(a, k)$)**.** *Let* $a, k \in \mathbb{N} \setminus \{1\}$*, and let* $F = \{1, a, a^2, \ldots, a^{k-1}\}$*. Let* $S(a, k) \subseteq \mathbb{N}$ *be defined by*

$$S(a, k) := \{n \in \mathbb{N} : \nu_a(n) \equiv k - 1 \pmod{k}\}. \tag{2.6}$$

Then $S(a, k)$ *is a multiplicatively* F*-syndetic set and has natural density*

$$d\left(S(a, k)\right) = \frac{a - 1}{a^k - 1}.$$

Proof. By noting that $\nu_a(an) = \nu_a(n) + 1$, we see that, for any $n \in \mathbb{N}$, the set $\nu_a(n \cdot F)$ is a complete residue system modulo k. Thus, $S(a, k)$ is a multiplicatively F-syndetic set.

It only remains to check that $S(a, k)$ achieves the required density bound. For each $m \in \mathbb{N}$, let $A_m = \{n \in \mathbb{N} : \nu_a(n) = km - 1\}$. Hence,

$$S(a, k) = \bigcup_{m \in \mathbb{N}} A_m.$$

Observe that $n \in A_m$ holds if and only if $n \equiv a^{km-1}b \pmod{a^{km}}$ for some $b \in \{1, 2, \ldots, a-1\}$. We therefore deduce that A_m has natural density

$$d(A_m) = \frac{a-1}{a^{km}}.$$

For each $r \in \mathbb{N}$, let $B_r = \cup_{m=1}^{r} A_m$. Since A_i and A_j are disjoint for all $i \neq j$, we deduce from the finite additivity of natural density that

$$d(B_r) = \sum_{m=1}^{r} d(A_m) = (a-1) \sum_{m=1}^{r} a^{-km} = (a-1) \frac{1 - a^{-rk}}{a^k - 1}.$$

By noting that $B_r \subseteq S(a,k)$, we deduce that $\underline{d}\,(S(a,k)) \geqslant d(B_r)$ for all $r \in \mathbb{N}$. Taking $r \to \infty$ gives the lower bound $\underline{d}(S(a,k)) \geqslant (a-1)/(a^k - 1)$.

We now compute an upper bound. First observe that

$$S(a,k) \setminus B_r \subseteq a^{kr-1} \cdot \mathbb{N}$$

for all $r \in \mathbb{N}$. Thus,

$$\bar{d}(S(a,k)) \leqslant \bar{d}(B_r) + \bar{d}(a^{kr-1} \cdot \mathbb{N}) = d(B_r) + \frac{1}{a^{kr-1}}$$

holds for all $r \in \mathbb{N}$. Taking $r \to \infty$ gives the desired upper bound. $\square$

We now show that $S(a,k)$ has minimal density.

Theorem 2.3.4 (Minimal $\{1, a, \ldots, a^{k-1}\}$-syndetic set)**.** *Let $a, k \in \mathbb{N} \setminus \{1\}$, and let $F = \{1, a, a^2, \ldots, a^{k-1}\}$. Let $S(a,k)$ be the set defined in (2.6). Let $N \in \mathbb{N}$. If $X \subseteq \mathbb{N}$ is a multiplicatively F-syndetic set, then*

$$|X \cap [N]| \geqslant |S(a,k) \cap [N]|.$$

Proof. We may assume that $N \geqslant a^{k-1}$, since otherwise $S(a,k) \cap [N] = \emptyset$ and the result is vacuously true. Let $X \subseteq \mathbb{N}$ be a multiplicatively F-syndetic set. Let $m \in S(a,k) \cap [N]$. Since every element of $S(a,k)$ is divisible by a^{k-1}, we deduce that $(a^{-(k-1)}m) \cdot F \subseteq [N]$. As X is multiplicatively F-syndetic, we can find some $t(m) \in F$ such that $a^{-(k-1)}mt(m) \in X$. We can therefore define a function $g : S(a,k) \cap [N] \to X$ by $g(n) = a^{-(k-1)}nt(n)$. To complete the proof it is sufficient to show that g is an injective function.

Suppose that $n, n' \in S(a,k)$ satisfy $g(n) = g(n')$. Thus, $nt(n) = n't(n')$. Applying ν_a to this equation and then reducing modulo k gives

$$\nu_a(t(n)) \equiv \nu_a(t(n')) \pmod{k}.$$

Now observe that the function which maps $t \in F$ to the residue class of $\nu_a(t)$ modulo k is injective. This shows that $t(n) = t(n')$, which implies that $n = n'$. □

2.4 A syndetic density increment strategy

Brauer [Bra28] established the following common generalisation of Schur's theorem and van der Waerden's theorem.

Brauer's Theorem: *For all $k, r \in \mathbb{N}$, there exists $N_0 = N_0(k,r) \in \mathbb{N}$ such that, in any r-colouring $[N_0] = C_1 \cup \cdots \cup C_r$ of $[N_0] = \{1, 2, \ldots, N_0\}$, there is a colour class C_i containing a set of the form*

$$\{x, d, x+d, x+2d, \ldots, x+(k-1)d\}. \tag{2.7}$$

In this section we study the case where $k = 3$. This corresponds to configurations of the form

$$\{x, d, x+d, x+2d\}. \tag{2.8}$$

We refer to such configurations as *Brauer configurations (of length* 3*)*. These are three term arithmetic progressions along with their common difference. Alternatively, one can view Brauer configurations as being solutions $\{x, y, z, d\}$ to the following dilation invariant system of equations:

$$x - 2y + z = 0;$$
$$x - y + d = 0.$$

As shown by Theorem 2.1.1, Brauer's theorem is equivalent to the assertion that all multiplicatively syndetic sets contain Brauer configurations. In this section, we derive quantitative bounds on the minimal $N \in \mathbb{N}$ for which the set $S \cap [N]$ must contain a Brauer configuration, for a given multiplicatively syndetic set S. The main theorem of this section is as follows.

Theorem 2.4.1 (Syndetic Brauer). *Let $M \geqslant 2$ be a positive integer. Let $A \subseteq [N]$ be such that $|A| \geqslant \delta N$ for some $0 < \delta \leqslant 1/2$. Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set. If there does not exist a 3-term arithmetic progression in A with common difference in S, then*

$$\log\log N \ll \log(\delta^{-1})\log(M/\delta). \tag{2.9}$$

Remark. We impose the restriction $\delta \leqslant 1/2$ to ensure that δ is bounded away from 1. One could replace $1/2$ by any positive quantity strictly less than 1 at the cost of increasing the implicit constant in (2.9).

In [GT09], Green and Tao used a density increment strategy over quadratic factors to obtain new bounds for the sizes of subsets of $[N]$ lacking 4-term arithmetic progressions. In this section we combine their methods with the induction on colours argument (Lemma 2.2.5) of Chow, Lindqvist, and Prendiville [CLP21] to prove Theorem 2.4.1.

By incorporating the density bounds obtained in §2.3, we can use Theorem 2.4.1 to prove Theorem 2.1.3.

Proof of Theorem 2.1.3 given Theorem 2.4.1. Let $c > 0$ be a small positive constant to be specified later, and assume that $M, N \in \mathbb{N}$ satisfy

$$2 \leqslant M \leqslant \exp\left(c\sqrt{\log\log N}\right).$$

Let $S \subseteq \mathbb{N}$ be a multiplicatively F-syndetic set, for some non-empty $F \subseteq [M]$. By taking c sufficiently small, we may assume that $N \geqslant M$. This implies that the density of $S \cap [N]$ in $[N]$ is positive. Moreover, by Lemma 2.3.1, we can choose a subset $S' \subseteq S$ such that the density δ of $S' \cap [N]$ in $[N]$ satisfies both of the bounds $0 < \delta \leqslant 1/2$ and $\delta^{-1} \ll M^2$. This gives

$$\log(\delta^{-1})\log(M/\delta) \ll \log^2 M.$$

Hence, by choosing c to be sufficiently small, we can ensure that (2.9) does not hold. Thus, by taking $A = S' \cap [N]$, we conclude from Theorem 2.4.1 that $S \cap [N]$ contains a Brauer configuration. □

For each $r \in \mathbb{N}$, define the *r colour Brauer number* $\mathrm{B}(r)$ to be the minimum $N \in \mathbb{N}$ such that any r-colouring of $[N]$ yields a monochromatic Brauer configuration. In other words, $\mathrm{B}(r)$ is the minimal value that $N_0(3,r)$ may take in the statement of Brauer's theorem. We can use Theorem 2.4.1 to obtain the following recursive bound for these numbers.

Theorem 2.4.2 (Recursive bound for $\mathrm{B}(r)$)**.** *For each $r \in \mathbb{N}$, we have*

$$\mathrm{B}(r+1) \leqslant 2^{\mathrm{B}(r)^{O(\log(r+1))}}. \tag{2.10}$$

By some computation (which is given in Appendix A), this theorem leads to the tower type bound in Theorem 2.1.4.

Proof of Theorem 2.4.2 given Theorem 2.4.1. Let $r \in \mathbb{N}$. Let $M := \mathrm{B}(r)$, and let $\delta := (r+1)^{-1}$. Note that $M \geqslant \mathrm{B}(1) = 3$. Suppose $N \in \mathbb{N}$ is such that $N < \mathrm{B}(r+1)$. Therefore, we have an $(r+1)$-colouring $[N] = C_1 \cup \cdots \cup C_{r+1}$ with no monochromatic sets of the form (2.8). By dilation invariance and our choice of M, it follows that there cannot exist a set of the form $x \cdot [M] \subseteq [N]$ which is r-coloured. This implies that $C_i \cup (\mathbb{N} \setminus [N])$ is multiplicatively $[M]$-syndetic for all $i \in [r]$.

Without loss of generality, assume that C_1 is the largest colour class. By the pigeonhole principle, we observe that $|C_1| \geqslant \delta N$. Hence, by taking $A = C_1$ and $S = C_1 \cup (\mathbb{N} \setminus [N])$ in the statement of Theorem 2.4.1, we deduce that

$$\log\log \mathrm{B}(r+1) \ll \log(r+1) \log\left((r+1)\mathrm{B}(r)\right).$$

By noting that $\mathrm{B}(r) \geqslant r+1$ (since one can r-colour $[r]$ so that each element has a unique colour), this gives

$$\log\log \mathrm{B}(r+1) \ll \log(r+1)\log(\mathrm{B}(r)).$$

Exponentiating twice then gives (2.10). $\square$

2.4.1 Norms

Given a non-empty finite set A and a function $f : A \to \mathbb{C}$, we define the expectation $\mathbb{E}_A(f)$ of f over A by

$$\mathbb{E}_A(f) = \mathbb{E}_{x \in A}(f) := \frac{1}{|A|}\sum_{x \in A} f(x).$$

The functions we encounter in this section are usually defined on $[N]$ or $\mathbb{Z}/p\mathbb{Z}$, where $N \in \mathbb{N}$, and $p \in \mathbb{N}$ is a prime. When $p > N$, it is convenient to consider $[N]$ as a subset of the field $\mathbb{Z}/p\mathbb{Z}$ by reducing modulo p. Given a function $f : \mathbb{Z} \to \mathbb{C}$ which is supported on $[N]$, we can then consider f as a function defined on $\mathbb{Z}/p\mathbb{Z}$ by taking $f(x) = 0$ for all $x \in (\mathbb{Z}/p\mathbb{Z}) \setminus [N]$.

We make use of two different types of norms. The standard L^p norms are used to measure the overall size of a function, whilst the Gowers uniformity U^s norms (introduced in [Gow01, Lemma 3.9]) measure the degree to which a function exhibits non-uniformity.

Definition (L^p norms)**.** Let A be a set and let $f : A \to \mathbb{C}$ be a finitely supported function. The L^1 *norm* $\|f\|_{L^1(A)}$ of f is defined by

$$\|f\|_{L^1(A)} := \mathbb{E}_{x\in A}|f(x)|.$$

The L^∞ *norm* $\|f\|_{L^\infty(A)}$ of f is defined by

$$\|f\|_{L^\infty(A)} := \max_{x\in A} |f(x)|.$$

We say that f is 1-*bounded* (on A) if $\|f\|_{L^\infty(A)} \leqslant 1$.

Definition (U^s norms)**.** Let $f : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$. For each $s \geqslant 1$, the U^s *norm*[4] $\|f\|_{U^s(\mathbb{Z}/p\mathbb{Z})}$ of f is defined by

$$\|f\|_{U^s(\mathbb{Z}/p\mathbb{Z})} := \left(\mathbb{E}_{x,h_1,h_2,\ldots,h_s\in\mathbb{Z}/p\mathbb{Z}}\, \Delta_{h_1,\ldots,h_s} f(x)\right)^{1/2^s}, \tag{2.11}$$

where the *difference operators* $\Delta_{h_1,\ldots,h_s}$ are defined by

$$\Delta_h f(x) := f(x)\overline{f(x+h)}$$

and

$$\Delta_{h_1,\ldots,h_s} f := \Delta_{h_1}\Delta_{h_2}\cdots\Delta_{h_s} f.$$

The Gowers uniformity norms can also be defined recursively. By expanding and rearranging (2.11), we observe that

$$\|f\|_{U^{s+1}}^{2^{s+1}} = \mathbb{E}_{h\in\mathbb{Z}/p\mathbb{Z}}\|\Delta_h f\|_{U^s}^{2^s}. \tag{2.12}$$

[4]For $s \geqslant 2$ the U^s norms are indeed norms (see [TV06, Chapter 11]), however the U^1 'norm' is only a seminorm.

2.4.2 Counting Brauer configurations

To prove Theorem 2.4.2, we use an induction on colours argument similar to [CLP21, §4].

For the remainder of this section, we let N denote a positive integer and consider Brauer configurations in the interval $[N]$. It is useful to embed $[N]$ in an abelian group which is not much larger than $[N]$. We therefore let p denote a prime[5] satisfying $3N < p < 6N$, and embed $[N]$ in the group $\mathbb{Z}/p\mathbb{Z}$ by reducing modulo p. We also let $M \in \mathbb{N}$ be a positive integer with $M > 1$ so that we may consider multiplicatively $[M]$-syndetic sets.

Let $S \subseteq \mathbb{Z}/p\mathbb{Z}$ and let $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$. The counting functional Λ_S we use is defined by

$$\Lambda_S(f_1, f_2, f_3) := \mathop{\mathbb{E}}_{x\in\mathbb{Z}/p\mathbb{Z}} \mathop{\mathbb{E}}_{d\in S} f_1(x)f_2(x+d)f_3(x+2d).$$

For brevity, we write $\Lambda_S(f) := \Lambda_S(f, f, f)$.

Lemma 2.4.3 (Counting Brauer configurations with common difference in S)**.** *Let $N \in \mathbb{N}$ with $N \geqslant 3$. If $S \subseteq [N/3]$ is non-empty, then*

$$\Lambda_S(1_{[N]}) > \frac{1}{18}.$$

Proof. Since $S \subseteq [N/3]$, for all $d \in S$ we have $N - 2d \geqslant N/3$. Combining this with the bound $p < 6N$ gives

$$\Lambda_S(1_{[N]}) = \frac{1}{p} \mathop{\mathbb{E}}_{d\in S}(N-2d) \geqslant \frac{N}{3p} > \frac{1}{18}.$$

□

We now use the two different types of norms introduced earlier to control the size of Λ_S. The simplest way to bound Λ_S is by using the L^1 norm.

Lemma 2.4.4 (L^1 control for Λ_S)**.** *Let $S \subseteq [N]$ and let $f, g : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ be 1-bounded functions. Then we have*

$$|\Lambda_S(f) - \Lambda_S(g)| \leqslant 3\|f - g\|_{L^1(\mathbb{Z}/p\mathbb{Z})}. \tag{2.13}$$

[5] Such a prime exists by Bertrand's postulate.

Proof. First let $k \in \{1,2,3\}$ and let $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ be functions such that f_i is 1-bounded for all $i \neq k$. By a change of variables, we see that

$$\begin{aligned} |\Lambda_S(f_1, f_2, f_3)| &\leqslant \mathop{\mathbb{E}}_{d \in S} \mathop{\mathbb{E}}_{x \in \mathbb{Z}/p\mathbb{Z}} |f_1(x) f_2(x+d) f_3(x+2d)| \\ &\leqslant \mathop{\mathbb{E}}_{d \in S} \mathop{\mathbb{E}}_{x \in \mathbb{Z}/p\mathbb{Z}} |f_k(x+(k-1)d) 1_S(d)| \\ &= \left(\mathop{\mathbb{E}}_{y \in \mathbb{Z}/p\mathbb{Z}} |f_k(y)| \right) \left(\mathop{\mathbb{E}}_{d \in S} 1_S(d) \right). \end{aligned}$$

We therefore deduce that

$$|\Lambda_S(f_1, f_2, f_3)| \leqslant \|f_k\|_{L^1(\mathbb{Z}/p\mathbb{Z})} \tag{2.14}$$

holds for all $k \in \{1,2,3\}$.

Now observe that, by multilinearity, we have the telescoping identity

$$\Lambda_S(f) - \Lambda_S(g) = \Lambda_S(f-g, f, f) + \Lambda_S(g, f-g, f) + \Lambda_S(g, g, f-g). \tag{2.15}$$

Applying the triangle inequality to this identity and using (2.14) gives (2.13). □

In addition to Λ_S, for functions $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$, we introduce the auxiliary counting functional AP_3 given by

$$AP_3(f_1, f_2, f_3) := \mathop{\mathbb{E}}_{x, d \in \mathbb{Z}/p\mathbb{Z}} f_1(x) f_2(x+d) f_3(x+2d).$$

Since Brauer configurations contain three term arithmetic progressions, it is perhaps unsurprising that the uniformity of Λ_S is related to the uniformity of AP_3. Indeed, the original motivation for the introduction of the U^s norms in [Gow01, §3] was the observation that they control counting functionals for arithmetic progressions. This result is referred to in the literature as a *generalised von Neumann theorem*. In the case of three term arithmetic progressions, the result is as follows.

Lemma 2.4.5 (Generalised von Neumann theorem)**.** *Let $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ be 1-bounded functions. Then we have*

$$|AP_3(f_1, f_2, f_3)| \leqslant \min_{1 \leqslant k \leqslant 3} \|f_k\|_{U^2(\mathbb{Z}/p\mathbb{Z})}. \tag{2.16}$$

Proof. This follows from two applications of the Cauchy-Schwarz inequality. For the full details, see [TV06, Lemma 11.4]. □

We now prove an analogous result for the Λ_S functional.

Lemma 2.4.6 (Generalised von Neumann theorem for Λ_S)**.** *Let $S \subseteq [N]$ be a non-empty set, and let $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ be 1-bounded functions. Then*

$$|\Lambda_S(f_1, f_2, f_3)| \leqslant \frac{p^{1/2}}{|S|^{1/2}} \min_{1\leqslant k\leqslant 3} \|f_k\|_{U^3(\mathbb{Z}/p\mathbb{Z})}.$$

Proof. Observe that we can rewrite $\Lambda_S(f_1, f_2, f_3)$ as

$$\Lambda_S(f_1, f_2, f_3) = \frac{p}{|S|} \mathop{\mathbb{E}}_{x,d\in\mathbb{Z}/p\mathbb{Z}} f_1(x)f_2(x+d)f_3(x+2d)1_S(d).$$

By applying the Cauchy-Schwarz inequality (with respect to the d variable), we see that the quantity $|\Lambda_S(f_1, f_2, f_3)|^2$ is bounded above by

$$\begin{aligned}
&\left(\mathop{\mathbb{E}}_{d\in\mathbb{Z}/p\mathbb{Z}} \frac{p^2}{|S|^2} 1_S(d)\right)\left(\mathop{\mathbb{E}}_{d\in\mathbb{Z}/p\mathbb{Z}} \left|\mathop{\mathbb{E}}_{x,d\in\mathbb{Z}/p\mathbb{Z}} f_1(x)f_2(x+d)f_3(x+2d)\right|^2\right)\\
&= \frac{p}{|S|} \mathop{\mathbb{E}}_{d,x,x'\in\mathbb{Z}/p\mathbb{Z}} f_1(x)f_1(x')f_2(x+d)f_2(x'+d)f_3(x+2d)f_3(x'+2d)\\
&= \frac{p}{|S|} \mathop{\mathbb{E}}_{d,h,x\in\mathbb{Z}/p\mathbb{Z}} \Delta_h f_1(x)\Delta_h f_2(x+d)\Delta_h f_3(x+2d)\\
&= \frac{p}{|S|} \mathop{\mathbb{E}}_{h\in\mathbb{Z}/p\mathbb{Z}} AP_3(\Delta_h f_1, \Delta_h f_2, \Delta_h f_3).
\end{aligned}$$

Note that the penultimate equality above follows from a changed of variables of the form $h = x' - x$. Now let $k \in \{1, 2, 3\}$. Using Lemma 2.4.5 and (2.12) along with an application of Hölder's inequality gives

$$\begin{aligned}
|\Lambda_S(f_1, f_2, f_3)|^2 &\leqslant \frac{p}{|S|} \mathop{\mathbb{E}}_{h\in\mathbb{Z}/p\mathbb{Z}} \|\Delta_h f_k\|_{U^2(\mathbb{Z}/p\mathbb{Z})}\\
&\leqslant \frac{p}{|S|} \left(\mathop{\mathbb{E}}_{h\in\mathbb{Z}/p\mathbb{Z}} 1^{4/3}\right)^{3/4} \left(\mathop{\mathbb{E}}_{h\in\mathbb{Z}/p\mathbb{Z}} \|\Delta_h f_k\|^4_{U^2(\mathbb{Z}/p\mathbb{Z})}\right)^{1/4}\\
&= \frac{p}{|S|} \|f_k\|^2_{U^3(\mathbb{Z}/p\mathbb{Z})}.
\end{aligned}$$

□

Lemma 2.4.7 (U^3 controls Λ_S)**.** *Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set, and let $f, g : \mathbb{Z}/p\mathbb{Z} \to [0, 1]$. If $N \geqslant 18M^2$, then*

$$|\Lambda_{S\cap[N/3]}(f) - \Lambda_{S\cap[N/3]}(g)| \leqslant 18M\|f - g\|_{U^3(\mathbb{Z}/p\mathbb{Z})}. \tag{2.17}$$

Proof. Note that as f and g are non-negative, the difference $f - g$ is a 1-bounded function. Applying the previous lemma and the triangle inequality to the telescoping identity (2.15) then gives

$$|\Lambda_{S\cap[N/3]}(f) - \Lambda_{S\cap[N/3]}(g)| \leqslant \frac{3p^{1/2}}{|S\cap[N/3]|^{1/2}}\|f-g\|_{U^3(\mathbb{Z}/p\mathbb{Z})}.$$

Combining Lemma 2.3.1 with the assumption $N \geqslant 18M^2$ gives

$$|S\cap[N/3]| \geqslant \frac{N}{6M^2} > \frac{p}{36M^2}.$$

This implies the desired bound (2.17). □

We now proceed to prove Theorem 2.4.1 using a density increment strategy. This argument combines Green and Tao's quadratic Fourier analytic methods [GT09] for finding sets lacking arithmetic progressions of length 4 with the techniques used by Chow, Lindqvist, and Prendiville [CLP21, Lemma 7.1] to obtain a 'homogeneous' generalisation of Sárközy's theorem [Sár78].

The original density increment strategy of Roth [Rot53] was used to show that subsets of $[N]$ which lack arithmetic progressions of length 3 have size $o(N)$. This method was subsequently modified by Gowers to prove an analogous result for arithmetic progressions of length 4 [Gow98], and then further generalised for progressions of arbitrary length [Gow01]. The argument proceeds as follows. Let $\delta_0 > 0$. Suppose $A \subseteq [N]$ lacks arithmetic progressions of length 3 and satisfies $|A| = \alpha N$ for some $\alpha \geqslant \delta_0$. Then provided that N is 'not too small', meaning that $N > C(\delta_0)$ for some positive constant $C(\delta_0)$ depending only on δ_0, we can find an arithmetic progression $P \subseteq [N]$ of length $N' := |P| \geqslant F(N, \delta_0)$ on which A has a density increment

$$\alpha' := \frac{|A\cap P|}{|P|} \geqslant \alpha + c(\delta_0). \tag{2.18}$$

Here $c(\delta_0) > 0$ is a positive constant depending only on δ_0, and F is an explicit positive function such that, for any fixed $\delta > 0$, $F(N,\delta) \to \infty$ as $N \to \infty$.

We can then apply an affine transformation of the form $x \mapsto ax + b$ to injectively map $A \cap P$ into $[N']$ with image $A' \subseteq [N']$. Since arithmetic progressions are translation-dilation invariant, we deduce that A' also lacks arithmetic progressions of length 3 and satisfies $|A'| = \alpha' N' > \delta_0 N'$. We can then iterate this argument. Since the density is increasing by at least $c(\delta_0)$ after each iteration, this process must eventually terminate. We can then procure an upper bound for the size of the original N in terms of $C(\delta_0), c(\delta_0)$ and $F(\cdot\,, \delta_0)$.

An important aspect of this method is that it uses the translation-dilation invariance of arithmetic progressions. However, more general configurations, such as Brauer configurations, are not translation invariant. This is emphasised by the fact that the odd numbers have density $1/2$ and yet they do not contain any Brauer configurations. A density analogue of Brauer's theorem is therefore impossible, and so this argument cannot help us prove Brauer's theorem.

The major insight of Chow, Lindqvist, and Prendiville [CLP21, §7] is that one can separate such configurations into a 'translation invariant part' and a 'non-translation invariant part'. For instance, observe that if $\{x, d, x+d, x+2d\}$ is a Brauer configuration, then the set $\{x+h, d, (x+d)+h, (x+2d)+h\}$ is also a Brauer configuration for any $h \in \mathbb{N}$. Brauer configurations therefore consist of a translation invariant part $\{x, x+d, x+2d\}$ and a non-translation invariant part $\{d\}$.

This allows us to modify the density increment strategy of Roth and Gowers to prove Brauer's theorem. Instead of studying a single set A lacking Brauer configurations, we study a pair of sets A and S with the following properties.

(i) (Density). $A \subseteq [N]$ satisfies $|A| \geqslant \delta N$.

(ii) (Syndeticity). $S \subseteq \mathbb{N}$ is a multiplicatively $[M]$-syndetic set.

(iii) (Brauer free). There does not exist an arithmetic progression of length 3 in A with common difference in $S \cap [N/3]$.

As in the original density increment argument, we show that, provided N is 'not too small', we can find a long arithmetic progression $P \subseteq [N]$ on which we have a density increment of the form (2.18). As before, we can apply an affine transformation to obtain a new set $A' \subseteq [N']$ with increased density. We can also obtain a new multiplicatively $[M]$-syndetic set $S' = d^{-1}S$, where d is the common difference of the progression P.

Recall that the translation invariant part $\{x, x+d, x+2d\}$ of a Brauer configuration is required to come from the dense set A, whilst the non-translation invariant part $\{d\}$ comes from the multiplicatively syndetic set S. Thus we have obtained new sets $A', S' \subseteq [N']$ satisfying (i)-(iii). Iterating this procedure as in the Gowers-Roth argument allows us to prove Theorem 2.4.1.

Theorem 2.4.8 (Density increment for Brauer configurations). *There exists a constant $C_0 > 1$ such that the following is true. Let $A \subseteq [N]$ be such that $|A| \geqslant \delta N$, for some $\delta > 0$, and let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set. Suppose that there do not exist $x \in A$ and $d \in S \cap [N/3]$ such that $\{x, x+d, x+2d\} \subseteq A$. If N satisfies*

$$N > \exp(C_0 \delta^{-C_0} M^{C_0}), \tag{2.19}$$

then there exists positive constants $C, c > 0$ and an arithmetic progression P in $[N]$ satisfying

$$|P| \gg N^{c\delta^C M^{-C}}$$

such that we have the density increment

$$\frac{|A \cap P|}{|P|} \geqslant (1 + c)\,\delta.$$

Remark. The bound (2.19) is needed to ensure that N is not too small to satisfy the conclusion of the theorem. Moreover, by taking $C_0 \geqslant 2$, we can assume that $N \geqslant 18M^2$. This allows us to make use of Lemma 2.4.7.

Proof of Theorem 2.4.1 given Theorem 2.4.8. Let $C_1 \geqslant C_0$ be a large positive parameter (which does not depend on M or δ) to be specified later. We use the following iteration algorithm. After the ith iteration, we have a positive integer $N_i \in \mathbb{N}$, a positive real number $\delta_i \geqslant \delta > 0$, an infinite set $S_i \subseteq \mathbb{N}$, and a finite set $A_i \subseteq [N_i]$ satisfying the following three properties:

(I) $|A_i| \geqslant \delta_i |N_i|$;

(II) S_i is a multiplicatively $[M]$-syndetic set;

(III) there does not exist a 3-term arithmetic progression in A_i with common difference in $S_i \cap [N_i/3]$.

We begin by defining the initial variables $N_0 := N$, $A_0 := A$, $S_0 := S$, $\delta_0 := \delta$. The iteration step of the algorithm proceeds as follows. If after the ith iteration we have

$$N_i \leqslant \exp(C_1 \delta^{-C_1} M^{C_1}), \tag{2.20}$$

then the algorithm terminates. If not, then we can apply Theorem 2.4.8 with $N = N_i$ and $f = 1_{A_i}$ to obtain an arithmetic progression $P_i \subseteq [N_i]$ of the form

$$P_i = \{a_i, a_i + d_i, \ldots, a_i + (|P_i| - 1)d_i\}$$

which satisfies the length bound

$$|P_i| \gg N_i^{c\delta^C M^{-C}}, \tag{2.21}$$

and provides the density increment

$$\frac{|A_i \cap P_i|}{|P_i|} \geqslant (1+c)\,\delta_i. \tag{2.22}$$

Moreover, by partitioning P_i into two shorter progressions if necessary, we can ensure that $d_i|P_i| \leqslant N_i$, provided that C_1 is sufficiently large. We then take

$$\begin{aligned} N_{i+1} &:= |P_i|; \\ A_{i+1} &:= \{x \in [N_{i+1}] : a_i + (x-1)d_i \in A_i \cap P_i\}; \\ S_{i+1} &:= d_i^{-1} S_i = \{x \in \mathbb{N} : d_i x \in S_i\}; \\ \delta_{i+1} &:= \frac{|A_{i+1}|}{|N_{i+1}|}. \end{aligned}$$

We now claim that $(N_{i+1}, A_{i+1}, S_{i+1}, \delta_{i+1})$ satisfy properties (I), (II), and (III). Property (I) follows immediately from our choice of δ_{i+1}. Property (II) follows from the fact that $x \cdot [M]$ intersects $d_i^{-1}S_i$ if and only if $(d_i x) \cdot [M]$ intersects S_i. Finally, notice that if A_{i+1} contains a 3-term arithmetic progression with common difference q, then A_i contains a 3-term arithmetic progression with common difference $d_i q$. Since $d_i|P_i| \leqslant N_i$, we conclude that $(N_{i+1}, A_{i+1}, S_{i+1}, \delta_{i+1})$ satisfies property (III).

We have therefore shown that our algorithm may continue with the new variables $(N_{i+1}, A_{i+1}, S_{i+1}, \delta_{i+1})$. Moreover, after applying the iteration process t times, we see from (2.22) that the density δ_t satisfies $\delta_t \geqslant (1+c)^t \delta$. Since $\delta_t \leqslant 1$ for all t, we conclude that the algorithm must terminate after T steps for some $T \ll \log(\delta^{-1})$.

We therefore deduce that (2.20) must hold for $i = T$. By (2.21), we see that

$$N_T \geqslant N^{O(c^T \delta^{CT} M^{-CT})}.$$

Combining these two bounds for N_T gives

$$c^T \delta^{CT} M^{-CT} \log N \ll C_1 \delta^{-C_1} M^{C_1}.$$

Rearranging and taking logarithms gives

$$\log\log N \ll \log C_1 + T\log(c^{-1}) + (C_1 + CT)\log\left(M/\delta\right) \ll \log(\delta^{-1})\log\left(M/\delta\right).$$

We have therefore established (2.9), as required. □

2.4.3 Quadratic Fourier analysis for Brauer configurations

The goal of the rest of this section is to prove Theorem 2.4.8. We achieve this by adapting the methods used by Green and Tao [GT09] to study subsets of $[N]$ which lack arithmetic progressions of length 4.

The objective of their argument is to show that if a subset $A \subseteq [N]$ of density α lacks arithmetic progressions of length 4, then there exists a long arithmetic progression $P \subseteq [N]$ upon which A achieves a density increment $|A \cap P| \geqslant (\alpha + c(\alpha))|P|$. Gowers' argument yields an increment of the form $c(\alpha) \gg \alpha^C$. The key insight of Green and Tao is that one can obtain a larger density increment if one first shows that A has a density increment on a 'quadratic Bohr set'. A linearisation procedure can then be applied to obtain a long arithmetic progression P' which provides a density increment $c(\alpha) \gg \alpha$.

Remark. In this subsection we make use of several results from [GT09]. The statements of these theorems contain a number of technical terms from quadratic Fourier analysis. Definitions of all the relevant terms and notions are given in Appendix B.

The first theorem we need is [GT09, Theorem 5.6].

Theorem 2.4.9 (Quadratic Koopman-von Neumann theorem)**.** *Let $\varepsilon > 0$ and let $f : \mathbb{Z}/p\mathbb{Z} \to [-1, 1]$. Suppose $K \in \mathbb{N}$ satisfies $K \geqslant C\varepsilon^{-C}$ for some absolute constant $C > 0$. Then there exists a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\varepsilon^{-C}), O(\varepsilon^{-C}))$ and resolution K such that*

$$\|f - \mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{triv})\|_{U^3(\mathbb{Z}/p\mathbb{Z})} \leqslant \varepsilon. \tag{2.23}$$

This theorem allows us to approximate (in the U^3 norm) a 1-bounded function f with a more 'structured' function $g := \mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{triv})$. In the proof of Theorem 2.4.8 we take $f = 1_A$, where $A \subseteq [N]$ is a dense subset of $[N]$ which lacks arithmetic progressions of length 3 with common difference in $S \cap [N/3]$, for a given multiplicatively syndetic set S. Our goal is to obtain a density increment on a quadratic factor for this f. To do this, we first show that it is sufficient to obtain a density increment with respect to the approximation g.

Corollary 2.4.10 (Brauer configurations on a quadratic factor)**.** *Let $f : \mathbb{Z}/p\mathbb{Z} \to [0, 1]$ be a 1-bounded non-negative function which is supported on $[N]$. Let $\delta > 0$. Suppose that $N \geqslant 18M^2$ and*

$$|\Lambda_{S\cap[N/3]}(f) - \Lambda_{S\cap[N/3]}(\delta 1_{[N]})| \geqslant \delta^3/18. \tag{2.24}$$

Then there exists a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of resolution $O(\delta^{-C}M^C)$ and complexity at most $(O(\delta^{-C}M^C), O(\delta^{-C}M^C))$ such that

$$|\Lambda_{S\cap[N/3]}(g) - \Lambda_{S\cap[N/3]}(\delta 1_{[N]})| \geqslant \delta^3/36, \tag{2.25}$$

where $g := \mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{triv})$.

Proof. Let $\varepsilon = (\delta^3 M^{-1})/648$. By Theorem 2.4.9, for some absolute constant $C > 0$, we have a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\delta^{-C}M^C), O(\delta^{-C}M^C))$ and resolution $O(\delta^{-C}M^C)$ such that (2.23) holds.

From Lemma 2.4.7 and (2.23) we deduce

$$|\Lambda_{S\cap[N/3]}(f) - \Lambda_{S\cap[N/3]}(g)| \leqslant 18M\varepsilon = \delta^3/36.$$

An application of the triangle inequality to (2.24) then gives (2.25). □

We now follow the approach of Green and Tao [GT09, Corollary 5.8] by replacing f with $\mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{triv})$ to obtain a density increment on a quadratic factor.

Corollary 2.4.11 (Density increment on a quadratic Bohr set)**.** *There exists a constant $\tilde{C}_0 \geqslant 2$ such that the following is true. Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set, and let $f : \mathbb{Z}/p\mathbb{Z} \to [0, 1]$ be supported on $[N]$. Suppose $\mathbb{E}_{[N]}(f) \geqslant \delta$, for some $\delta > 0$. Suppose further that conditions (2.19) and (2.24) both hold, for some $C_0 \geqslant \tilde{C}_0$. Then there exists a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\delta^{-C}M^C), O(\delta^{-C}M^C))$ and resolution $O(\delta^{-C}M^C)$, and an atom B of the factor $\mathcal{B}_2 \vee \mathcal{B}_{triv}$ with density $\frac{|B|}{p} \gg \exp(-O(\delta^{-C}M^C))$ which is contained in $[N]$ and is such that*

$$\mathbb{E}_B(f) \geqslant (1 + c)\,\delta, \tag{2.26}$$

for some absolute constant $c > 0$.

Proof. Let $(\mathcal{B}_1, \mathcal{B}_2)$ be the quadratic factor obtained from Corollary 2.4.10. Let $g = \mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{triv})$. Note that $[N] \in \mathcal{B}_{triv}$, and so $[N] \in \mathcal{B}_2 \vee \mathcal{B}_{triv}$. Since g is constant on atoms of $\mathcal{B}_2 \vee \mathcal{B}_{triv}$ and f is supported on $[N]$, we see that g is also supported on $[N]$. This implies that $\mathbb{E}_B(f) = \mathbb{E}_B(g)$ holds for any $B \in \mathcal{B}_2 \vee \mathcal{B}_{triv}$. Thus, it is sufficient to establish (2.26) with g in place of f.

Let $\eta > 0$ be a small constant (to be chosen later) and define

$$\Omega := \{x \in [N] : g(x) \geqslant (1+\eta)\,\delta\}\,.$$

Since g is constant on atoms of $\mathcal{B}_2 \vee \mathcal{B}_{triv}$, we deduce that Ω can be partitioned into atoms of $\mathcal{B}_2 \vee \mathcal{B}_{triv}$. We can therefore finish the proof if we can show that one of these atoms B satisfies

$$|B| \gg \exp(-O(\delta^{-C}M^C))p. \tag{2.27}$$

Now recall that the factor $\mathcal{B}_2 \vee \mathcal{B}_{triv}$ has complexity and resolution $O(\delta^{-C}M^C)$, and so contains at most $\exp(O(\delta^{-C}M^C))$ atoms. Thus the pigeonhole principle implies that (2.27) holds if we can obtain a bound of the form

$$|\Omega| \gg \delta^3 p. \tag{2.28}$$

Define the function $h : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ by

$$h(x) = 1_{[N]\setminus\Omega}(x)g(x).$$

Thus $\|h\|_{L^\infty(\mathbb{Z}/p\mathbb{Z})} < (1+\eta)\,\delta$. Taking $\eta \leqslant 1$ and applying Lemma 2.4.4 gives

$$|\Lambda_{S\cap[N/3]}(h) - \Lambda_{S\cap[N/3]}(\delta 1_{[N]})| \leqslant 12\delta^2 \|h - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})}.$$

From the fact that g is a 1-bounded function, we have

$$\|h - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} \leqslant \|g - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} + \frac{|\Omega|}{p}.$$

Since h is a 1-bounded function, Lemma 2.4.4 also gives

$$|\Lambda_{S\cap[N/3]}(g) - \Lambda_{S\cap[N/3]}(h)| \leqslant \frac{3|\Omega|}{p}.$$

Combining these three bounds and using the triangle inequality in (2.25) gives

$$\delta^2 \|g - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} + \frac{|\Omega|}{p} \gg \delta^3. \tag{2.29}$$

Recall that $\mathbb{E}_{[N]}(g) = \mathbb{E}_{[N]}(f) \geqslant \delta$. Thus,

$$\begin{aligned}\|g - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} &\leqslant \|g - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} + \mathbb{E}_{\mathbb{Z}/p\mathbb{Z}}(g - \delta 1_{[N]}) \\ &= 2\|(g - \delta 1_{[N]})_+\|_{L^1(\mathbb{Z}/p\mathbb{Z})} \\ &\ll \frac{|\Omega|}{p} + \delta\eta.\end{aligned}$$

If η is sufficiently small (relative to the implicit constants), we can then substitute this bound into (2.29) to obtain the desired result (2.28). □

To complete the proof of Theorem 2.4.8 it only remains to convert this density increment on a quadratic Bohr set into a density increment on an arithmetic progression. This is accomplished by implementing the following 'linearisation' procedure of Green and Tao [GT09, Proposition 6.2].

Theorem 2.4.12 (Linearisation of quadratic Bohr sets)**.** *Suppose $(\mathcal{B}_1, \mathcal{B}_2)$ is a quadratic factor in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most (d_1, d_2) and resolution K, for some $K \in \mathbb{N}$. Let B_2 be an atom of $\mathcal{B}_2$. Then for all $N \in \mathbb{N}$, there is a partition of $B_2 \cap [N]$ as a union of $\ll d_2^{O(d_2)} N^{1-c/(d_1+1)(d_2+1)^3}$ disjoint arithmetic progressions in $\mathbb{Z}/p\mathbb{Z}$.*

Proof of Theorem 2.4.8. Let $\tilde{C}_0 \geqslant 2$ be the positive constant appearing in the statement of Corollary 2.4.11. Let A and S be as defined in the statement of Theorem 2.4.8. Suppose N satisfies (2.19), for some $C_0 \geqslant \tilde{C}_0$. Let $f = 1_A$. By Corollary 2.4.11 there is a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\delta^{-C}M^C), O(\delta^{-C}M^C))$ and resolution $O(\delta^{-C}M^C)$, and an atom $B \subseteq [N]$ of $\mathcal{B}_2 \vee \mathcal{B}_{triv}$ with density $\frac{|B|}{p} \geqslant \exp(-O(\delta^{-C}M^C))$ such that $E_B(f) \geqslant (1 + c_0)\,\delta$. By Theorem 2.4.12, we may write B as a union of $\exp(O(\delta^{-C}M^C))N^{1-c\delta^C M^{-C}}$ arithmetic progressions. By an application of the pigeonhole principle (see [GT09, Lemma 6.1]), we deduce that one of these progressions P satisfies

$$|P| \geqslant \exp(-O(\delta^{-C}M^C))N^{c\delta^C M^{-C}}$$

and

$$\frac{|A \cap P|}{|P|} = \mathbb{E}_P f \geqslant \left(1 + \frac{c_0}{2}\right)\delta.$$

Notice that the only property of the parameter C_0 appearing in (2.19) that we have used is that $C_0 \geqslant \tilde{C}_0$. We may therefore take C_0 to be sufficiently large so that

$|P| \gg N^{c'\delta^{C'} M^{-C'}}$ holds for some absolute constants $C', c' > 0$. This completes the proof. $\square$

Bibliography

[BHLS15] B. Barber, N. Hindman, I. Leader, and D. Strauss, *Partition regularity without the columns property*, Proc. Amer. Math. Soc. **143** (2015), 3387–3399.

[Ber10] V. Bergelson, *Ultrafilters, IP sets, dynamics, and combinatorial number theory*, Ultrafilters across mathematics, Contemp. Math., vol. 530, Amer. Math. Soc., Providence, RI, 2010, pp. 23–47.

[BG18] V. Bergelson and D. Glasscock, *Multiplicative richness of additively large sets in* $\mathbb{Z}^d$, J. Algebra **503** (2018), 67–103.

[Bra28] A. Brauer, *Über Sequenzen von Potenzresten*, Sitzungsber. Preuss. Akad. Wiss. (1928), 9–16.

[CP20] J. Chapman and S. Prendiville, *On the Ramsey number of the Brauer configuration*, Bull. Lond. Math. Soc. **52** (2020), no. 2, 316–334.

[CLP21] S. Chow, S. Lindqvist, and S. Prendiville, *Rado's criterion over squares and higher powers*, J. Eur. Math. Soc. (JEMS) **23** (2021), no. 6, 1925–1997.

[CEG02] F. Chung, P. Erdős, and R. Graham, *On sparse sets hitting linear forms*, Number Theory for the Millennium, I (Urbana, IL, 2000), A K Peters, Natick, MA, 2002, pp. 257–272.

[CS17] K. Cwalina and T. Schoen, *Tight bounds on additive Ramsey-type numbers*, J. London Math. Soc., **96** (2017) 601–620.

[EEN00] D. B. Ellis, R. Ellis, and M. Nerurkar, *The topological dynamics of semigroup actions*, Trans. Amer. Math. Soc. **353** (2000), 1279–1320.

[EG80] P. Erdős and R. Graham, *Old and new problems and results in combinatorial number theory*, Monographies de L'Enseignement Mathématique vol. **28**, Université de Genève, Geneva, 1980.

[Gow98] W. T. Gowers, *A new proof of Szemerédi's theorem for progressions of length four*, Geom. Funct. Anal. **8** (1998), 529–551.

[Gow01] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465–588.

[GRS90] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey Theory*, Second Edition, Wiley, New York, 1990. xii+196 pp.

[GSW77] R. L. Graham, J. H. Spencer, and H. S. Witsenhausen, *On extremal density theorems for linear forms*, Number theory and algebra, Academic Press, 1977, pp. 103–107.

[GT09] B. J. Green and T. Tao, *New bounds for Szemerédi's theorem. II. A new bound for $r_4(N)$*, Analytic number theory, Cambridge Univ. Press, Cambridge, 2009, pp. 180–204.

[HLS03] N. Hindman, I. Leader, and D. Strauss, *Open problems in partition regularity*, Comb. Prob. and Comp. **12** (2003), 571–583.

[HS12] N. Hindman and D. Strauss, *Algebra in the Stone-Čech compactification: theory and applications*, Second Edition, Walter de Gruyter & Co., Berlin, 1998.

[Rad33] R. Rado, *Studien zur Kombinatorik*, Math. Z. **36** (1933), 242–280.

[Rot53] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.

[Sár78] A. Sárközy, *On difference sets of sequences of integers I*, Acta Math. Acad. Sci. Hungar. **31** (1978), 125–149.

[Sch16] I. Schur, *Uber die Kongruenz $x^m + y^m = z^m \pmod{p}$*, Jahresber. Dtsch. Math. Ver. **25** (1916), 114–117.

[She88] S. Shelah, *Primitive recursive bounds for van der Waerden numbers*, J. Amer. Math. Soc. **1** (1988), 683–697.

[TV06] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Univ. Press, Cambridge, 2006.

[Wae27] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wisk. **15** (1927), 212–216.

Chapter 3

Partition regularity for systems of diagonal equations

Abstract. We consider systems of n diagonal equations in kth powers. Our main result shows that if the coefficient matrix of such a system is sufficiently non-singular, then the system is partition regular if and only if it satisfies Rado's columns condition. Furthermore, if the system also admits constant solutions, then we prove that the system has non-trivial solutions over every set of integers of positive upper density.

3.1 Introduction

3.1.1 Partition regularity

A system of equations is said to be *(non-trivially) partition regular* over a set S if, whenever we finitely colour S, we can always find a (non-trivial) solution to our system of equations over S such that each variable has the same colour. Here, we refer to a solution $\mathbf{x} = (x_1, \ldots, x_s)$ as being *non-trivial* if $x_i \neq x_j$ for all $i \neq j$. When S is the set of positive integers $\mathbb{N}$, we typically omit S and refer to the system of equations as being (non-trivially) partition regular.

A foundational result in the field of arithmetic Ramsey theory is *Rado's criterion* [Rad33, Satz IV], which classifies all (finite) systems of homogeneous linear equations that are partition regular over $\mathbb{N}$.

Definition (Columns condition)**.** Let $\mathbf{M}$ be an $n \times s$ matrix with rational entries. Let

$\mathbf{c}^{(1)}, \ldots, \mathbf{c}^{(s)}$ denote the columns of $\mathbf{M}$. We say that the matrix $\mathbf{M}$ obeys the *columns condition* if there exists a partition $\{1, 2, \ldots, s\} = J_1 \cup \cdots \cup J_k$ such that $\sum_{j \in J_1} \mathbf{c}^{(j)} = \mathbf{0}$, and, for each $1 < t \leqslant k$,

$$\sum_{j \in J_t} \mathbf{c}^{(j)} \in \langle \mathbf{c}^{(r)} : r \in J_1 \cup \cdots \cup J_{t-1} \rangle_{\mathbb{Q}}.$$

Here $\langle V \rangle_{\mathbb{Q}}$ denotes the $\mathbb{Q}$-linear span of a set of vectors V with rational entries. By convention, $\langle \emptyset \rangle_{\mathbb{Q}} := \{\mathbf{0}\}$.

Rado's criterion. *Let $\mathbf{M}$ be an $n \times s$ non-empty matrix of rank n (over $\mathbb{Q}$) with integer entries. The system of equations $\mathbf{Mx} = \mathbf{0}$ is partition regular if and only if $\mathbf{M}$ obeys the columns condition.*

Recent developments in the study of partition regularity are motivated by the desire to extend Rado's classification to systems of non-linear equations. Chow, Lindqvist, and Prendiville [CLP21] recently established the following generalisation of Rado's criterion for diagonal polynomial equations in sufficiently many variables.

Theorem 3.1.1 ([CLP21, Theorem 1.3]). *For each $k \in \mathbb{N}$, there exists $s_0(k) \in \mathbb{N}$ such that the following is true. If $s \geqslant s_0(k)$ and $a_1, \ldots, a_s \in \mathbb{Z} \setminus \{0\}$, then the equation*

$$a_1 x_1^k + \cdots + a_s x_s^k = 0 \tag{3.1}$$

is non-trivially partition regular over $\mathbb{N}$ if and only if there exists a non-empty set $I \subseteq [s]$ such that $\sum_{i \in I} a_i = 0$. Moreover, one may take $s_0(2) = 5$, $s_0(3) = 8$, and $s_0(k) = k(\log k + \log\log k + 2 + O(\log\log k/\log k))$.

The main purpose of this chapter is to generalise this theorem to suitably non-singular systems of equations. To state our result, we require the following notation. Given a vector $\mathbf{v} = (v_1, \ldots, v_s) \in \mathbb{Q}^s$, we define the *support* of $\mathbf{v}$ to be the set $\operatorname{supp}(\mathbf{v}) := \{j \in \{1, \ldots, s\} : v_j \neq 0\}$.

Our main result is the following.

Theorem 3.1.2 (Partition regularity for diagonal polynomial systems). *Let $k, n, s \in \mathbb{N}$, with $k \geqslant 2$, and let $\mathbf{M} = (a_{i,j})$ be an $n \times s$ matrix with integer entries. Suppose that the following condition holds:*

(I) *for every non-empty set* $\{\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(d)}\} \subseteq \mathbb{Q}^s$ *of linearly independent non-zero vectors in the row space of* $\mathbf{M}$*, we have*

$$\left| \bigcup_{i=1}^{d} \operatorname{supp}\left(\mathbf{v}^{(i)}\right) \right| \geqslant dk^2 + 1.$$

Then the system of equations

$$\begin{aligned} a_{1,1}x_1^k + \cdots a_{1,s}x_s^k &= 0; \\ &\vdots \\ a_{n,1}x_1^k + \cdots a_{n,s}x_s^k &= 0 \end{aligned} \tag{3.2}$$

is non-trivially partition regular if and only if $\mathbf{M}$ *obeys the columns condition.*

The necessity of the columns condition in Theorem 3.1.2 was originally established by Lefmann [Lef91, Theorem 2.1]. In fact, Lefmann observed that (3.2) is (non-trivially) partition regular over $\mathbb{N}$ if and only if the linear system $\mathbf{Mx} = \mathbf{0}$ is (non-trivially) partition regular over the set $\{n^k : n \in \mathbb{N}\}$.

We remark that whilst it may be possible to weaken condition (I) of Theorem 3.1.2, it cannot be removed entirely. This is because the columns condition is not sufficient to ensure that (3.2) has non-zero integral solutions. To see this, consider the matrix

$$\mathbf{M} = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}.$$

Since the first three columns of $\mathbf{M}$ span $\mathbb{Q}^2$ and sum to $\mathbf{0}$, we see that the columns condition holds. Observe that if $(x, y, z, d) \in \ker \mathbf{M}$, then $\{x, y, z\}$ is an arithmetic progression with common difference d. However, Fermat's right triangle theorem implies that there are no non-trivial arithmetic progressions of length 3 in the squares whose common difference is also a square (see [Con08] for further details). Thus, the corresponding system of quadric equations

$$\begin{aligned} x^2 + z^2 &= 2y^2; \\ x^2 + d^2 &= y^2 \end{aligned}$$

has no solutions over $\mathbb{N}$, and is therefore not partition regular.

In general, it is a very difficult problem in number theory and algebraic geometry to determine whether an arbitrary system of Diophantine equations has any non-trivial solutions. For this reason, we are required to impose some form of non-singularity condition on our system, such as (I), so that we may use the Hardy-Littlewood circle method to count solutions. This allows us to develop the tools introduced by Chow, Lindqvist, and Prendiville [CLP21] so that we may establish partition regularity for systems of equations.

In the case when $k = n = 2$ and the rows of $\mathbf{M}$ are linearly independent, condition (I) is equivalent to the assertion that $s \geqslant 9$ and every non-zero vector in the row space of $\mathbf{M}$ has at least 5 non-zero entries. This observation shows that Theorem 3.1.2 confirms a conjecture of Chow, Lindqvist, and Prendiville [CLP21, Conjecture 3.1] concerning pairs of quadric equations in 9 variables.

Corollary 3.1.3 (Partition regularity for pairs of quadrics)**.** *Let $s \in \mathbb{N}$, and let $\mathbf{M} = (a_{i,j})$ be a $2 \times s$ integer matrix with no zero columns. Suppose that $s \geqslant 9$, and that every non-zero vector in the row space of $\mathbf{M}$ has at least 5 non-zero entries. Then the system of equations*

$$\begin{aligned} a_{1,1}x_1^2 + \cdots + a_{1,s}x_s^2 &= 0; \\ a_{2,1}x_1^2 + \cdots + a_{2,s}x_s^2 &= 0 \end{aligned}$$

is non-trivially partition regular if and only if $\mathbf{M}$ obeys the columns condition.

3.1.2 Density regularity

A set of positive integers A is said to have *positive upper density* if

$$\limsup_{N\to\infty} \frac{|A \cap \{1, \ldots, N\}|}{N} > 0.$$

We call a system of equations *(non-trivially) density regular* (over $\mathbb{N}$) if the system has a (non-trivial) solution over every set of positive integers with positive upper density. It follows from Szemerédi's theorem [Sze75] that the linear homogeneous system $\mathbf{Ax} = \mathbf{0}$ is density regular if and only if the columns of $\mathbf{A}$ sum to zero (see [FGR88, Fact 4]).

More recent work on density regularity has focused on non-linear configurations. Browning and Prendiville [BP17] obtained quantitative bounds for the largest subset of $\{1, \ldots, N\}$ lacking non-trivial solutions to (3.1) for $k = 2$. In particular, for $k = 2$ and $s \geqslant 5$, they prove that (3.1) is non-trivially density regular if and only if $a_1 + \cdots + a_s = 0$.

For $k \geqslant 3$ and $s \geqslant k^2 + 1$, Chow [Cho18] showed that (3.1) has non-trivial solutions over every relatively dense subset of the primes.

Our second main theorem classifies, in a quantitative manner, density regular systems of diagonal equations satisfying condition (I) of Theorem 3.1.2.

Theorem 3.1.4 (Density regularity for polynomial systems)**.** *Let $k, n, s \in \mathbb{N}$, with $k \geqslant 2$, and let $\mathbf{M} = (a_{i,j})$ be an $n \times s$ matrix with integer entries and no zero columns. Let $\delta > 0$. If $\mathbf{M}$ satisfies condition (I) of Theorem 3.1.2 and the columns of $\mathbf{M}$ sum to $\mathbf{0}$, then there exists a constant $c_1 = c_1(\delta, k, \mathbf{M}) > 0$ and a positive integer $N_1 = N_1(\delta, k, \mathbf{M}) \in \mathbb{N}$ such that the following is true. If $N \geqslant N_1$ and $A \subseteq \{1, 2, \ldots, N\}$ satisfies $|A| \geqslant \delta N$, then there are at least $c_1 N^{s-kn}$ non-trivial solutions $\mathbf{x} = (x_1, \ldots, x_s) \in A^s$ to (3.2).*

3.1.3 Structure of the chapter

We begin in §3.2 by introducing all the general notation and conventions used throughout this chapter. We also state the restriction estimates that are needed to control the counting operators studied in subsequent sections.

In §3.3 we investigate the structure and properties of matrices obeying condition (I) of Theorem 3.1.2. This allows us to establish a relative generalised von Neumann theorem for counting operators which count solutions to (3.2). This is the key result of this chapter, as it is integral to the arguments developed in subsequent sections to prove Theorem 3.1.2 and Theorem 3.1.4. Furthermore, we prove that condition (I) implies that the set of trivial solutions to (3.2) is sparse in the set of all solutions.

In §3.4 we introduce the notion of multiplicative syndeticity, and recall the induction on colours argument from [CLP21]. This allows us to reduce both Theorem 3.1.2 and Theorem 3.1.4 to a result, Theorem 3.4.5, concerning solutions of (3.2) over dense and multiplicatively syndetic sets.

In §3.5 we use the linearisation and W-trick procedures from [CLP21] and [Lin19] to reduce Theorem 3.4.5 to a 'linearised' version (Theorem 3.5.1). We also note that only a special case of Theorem 3.5.1 is needed to prove Theorem 3.1.4, and that this case follows from existing results (such as [FGR88, Theorem 2]).

Finally, in §3.6 we use the arithmetic regularity lemma of Green [Gre05B] to prove

Theorem 3.5.1, and thereby prove Theorem 3.1.2 and Theorem 3.1.4.

Acknowledgements

The author would like to thank Sean Prendiville for his continual support and encouragement, and for providing informative discussions regarding the use of the arithmetic regularity lemma. We also thank Christopher Frei for his support and helpful comments on an earlier draft of the paper on which this chapter is based. We are grateful to Trevor Wooley for his comments on Lemma 3.2.2.

3.2 Notation and preliminaries

3.2.1 Set notation

The set of positive integers is denoted by $\mathbb{N}$. The set of non-negative integers is denoted by $\mathbb{Z}_{\geqslant 0}$. The set of non-negative real numbers is denoted by $\mathbb{R}_{\geqslant 0}$. Given $X \in \mathbb{R}$, we write $[X] := \{n \in \mathbb{N} : n \leqslant X\}$. The indicator function of a set A is denoted by 1_A. We follow the convention that, for non-empty sets A and B, if $t = 0$, then $A \times B^t = A$.

We write $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$ to denote the d-dimensional torus, which we often identify with $[0,1)^d$ in the usual way. We define a metric $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \mapsto \|\boldsymbol{\alpha} - \boldsymbol{\beta}\|$ on $\mathbb{T}^d$ by

$$\|\boldsymbol{\theta}\| := \min_{\mathbf{n} \in \mathbb{Z}^d} \left(\sum_{i=1}^{d} |\theta_i - n_i| \right) = \sum_{i=1}^{d} \left(\min_{n \in \mathbb{Z}} |\theta_i - n| \right). \tag{3.3}$$

3.2.2 Asymptotic notation

Let $f : A \to \mathbb{C}$ and $g : A \to \mathbb{R}_{\geqslant 0}$ be functions defined on some set A, and let $\lambda_1, \ldots, \lambda_s$ be some parameters. We write $f \ll_{\lambda_1,\ldots,\lambda_s} g$ if there exists a positive constant $C = C(\lambda_1, ..., \lambda_s)$ depending only on the parameters λ_i such that $|f(x)| \leqslant Cg(x)$ for all $x \in A$. We also write $g \gg_{\lambda_1,\ldots,\lambda_s} f$ or $f = O_{\lambda_1,\ldots,\lambda_s}(g)$ to denote this same property.

3.2.3 Linear Algebra

The set of $n \times s$ matrices with entries in a given set S is denoted by $S^{n \times s}$. We allow n or s to be zero, in which case an $n \times s$ matrix is an *empty matrix*. The *row space* of $\mathbf{M} \in \mathbb{Q}^{n\times s}$ is the $\mathbb{Q}$-linear subspace of $\mathbb{Q}^s$ spanned by the rows of $\mathbf{M}$. Given $k \in \mathbb{N}$ and $\mathbf{x} = (x_1, \dots, x_s) \in \mathbb{Q}^s$, we write $\mathbf{x}^{\otimes k} := (x_1^k, \dots, x_s^k)$.

Given matrices $\mathbf{M}_i \in \mathbb{Q}^{n_i \times s_i}$ for $1 \leqslant i \leqslant r$, a *block upper triangular matrix with diagonal* $(\mathbf{M}_1, \dots, \mathbf{M}_r)$ is a matrix of the form

$$\begin{pmatrix} \mathbf{M}_1 & \mathbf{A}^{(1,2)} & \dots & \mathbf{A}^{(1,r)} \\ \mathbf{0} & \mathbf{M}_2 & \dots & \mathbf{A}^{(2,r)} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{M}_r \end{pmatrix},$$

where $\mathbf{A}^{(i,j)} \in \mathbb{Q}^{n_i \times s_j}$ for all $1 \leqslant i < j \leqslant r$.

Two matrices $\mathbf{M}, \mathbf{M}' \in \mathbb{Q}^{n\times s}$ are said to be *equivalent* if one can be obtained from the other by performing column permutations and elementary row operations. Observe that if $\mathbf{M}'$ can be obtained from $\mathbf{M}$ by performing elementary row operations, then $\ker(\mathbf{M}') = \ker(\mathbf{M})$. If $\mathbf{M}'$ can be obtained by applying the permutation $\sigma : [s] \to [s]$ to the columns of $\mathbf{M}$, then there is a linear isomorphism between $\ker(\mathbf{M})$ and $\ker(\mathbf{M}')$ given by $(x_1, \dots, x_s) \mapsto (x_{\sigma(1)}, \dots, x_{\sigma(s)})$. We therefore see that properties such as having (non-trivial) solutions over a set A or being partition regular are preserved under coefficient matrix equivalence. Throughout this chapter we frequently make use of this fact to pass to equivalent matrices when necessary.

3.2.4 Norms

Given a bounded function $f : A \to \mathbb{C}$ on a set A, we write $\|f\|_\infty := \sup_{x\in A} |f(x)|$. Let $d \in \mathbb{N}$ and $p \in \mathbb{R}$ with $p \geqslant 1$. For any function $\psi : \mathbb{T}^d \to \mathbb{C}$, we define the L^p norm of ψ by

$$\|\psi\|_{L^p(\mathbb{T}^d)} := \left(\int_{\mathbb{T}^d} |\psi(\boldsymbol{\alpha})|^p d\boldsymbol{\alpha} \right)^{\frac{1}{p}},$$

whenever the above integral exists and is finite. For any function $g : \mathbb{Z}^d \to \mathbb{C}$ of finite support, we define the L^p norm of g by

$$\|g\|_{L^p(\mathbb{Z}^d)} := \left(\sum_{x_1,\dots,x_d} |g(\mathbf{x})|^p \right)^{\frac{1}{p}}.$$

If the domain of a complex-valued function f is understood to be either $\mathbb{T}^d$ or $\mathbb{Z}^d$ for some $d \in \mathbb{N}$, then we write $\|f\|_p$ to denote the L^p norm of f. We say that f is *integrable* if the L^p norm of f exists and is finite for all $p \in [1, \infty]$.

Let $f_1, \dots, f_s$ be complex-valued integrable functions defined on a set A, where $A = \mathbb{T}^d$ or $A = \mathbb{Z}^d$ for some $d \in \mathbb{N}$. If $A = \mathbb{Z}^d$, then assume further that each f_i has finite support. Given $p_1, \dots, p_s \in [1, \infty]$ such that $p_1^{-1} + \cdots + p_s^{-1} = 1$ (with the convention that $\infty^{-1} := 0$), we have *Hölder's inequality*:

$$\|f_1 \cdots f_s\|_1 \leqslant \|f_1\|_{p_1} \cdots \|f_s\|_{p_s}. \tag{3.4}$$

Taking $s = 2$ and $p_1 = p_2 = 2$, we recover the *Cauchy-Schwarz inequality*:

$$\|f_1 \cdot f_2\|_1 \leqslant \|f_1\|_2 \|f_2\|_2.$$

3.2.5 Fourier analysis

Given $\alpha \in \mathbb{T}$, we write $e(\alpha) := \exp(2\pi i \alpha)$, where α has been identified with an element of $[0, 1)$ in the usual way. Given $\boldsymbol{\alpha} \in \mathbb{T}^d$ and $\mathbf{x} \in \mathbb{Q}^d$, we write $\boldsymbol{\alpha} \cdot \mathbf{x} := \alpha_1 x_1 + \cdots + \alpha_d x_d$.

Let $f : \mathbb{Z} \to \mathbb{C}$ be a function with finite support. The *(linear) Fourier transform of f* is the function $\hat{f} : \mathbb{T} \to \mathbb{C}$ defined by

$$\hat{f}(\alpha) := \sum_x f(x) e(\alpha x).$$

Let $k \in \mathbb{N}$. The *degree k Fourier transform of f* is defined by

$$\mathfrak{F}_k[f](\alpha) := \sum_x f(x) e(\alpha x^k).$$

Observe that if f is supported on $[N]$ for some $N \in \mathbb{N}$, then $\mathfrak{F}_k[f] = \hat{F}$, where $F : \mathbb{Z} \to \mathbb{C}$ is defined by

$$F(x) = \begin{cases} f(y), & \text{if } x = y^k \text{ for some } y \in [N]; \\ 0, & \text{otherwise.} \end{cases}$$

Given $d \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{Z}^d$, we have the *orthogonality relations*:

$$\int_{\mathbb{T}^d} e(\boldsymbol{\alpha} \cdot \mathbf{x}) d\boldsymbol{\alpha} = \begin{cases} 1, & \text{if } \mathbf{x} = \mathbf{0}; \\ 0, & \text{otherwise.} \end{cases}$$

For $x \in \mathbb{Z}$ and a function $f : \mathbb{Z} \to \mathbb{C}$ supported on a finite subset of $\mathbb{Z}_{\geqslant 0}$, these relations provide us with the *Fourier inversion formulae*:

$$f(x) = \int_{\mathbb{T}} \hat{f}(\alpha) e(-\alpha x) d\alpha = \int_{\mathbb{T}} \mathfrak{F}_k[f](\alpha) e(-\alpha x^k) d\alpha,$$

and *Parseval's identity*: $\|f\|_{L^2(\mathbb{Z})} = \|\hat{f}\|_{L^2(\mathbb{T})}$.

Our applications of Fourier analysis frequently make use of *restriction estimates*, which are uniform bounds for exponential sums over arithmetic sets. The precise definition we need is as given in [CLP21, Definition 5.4].

Definition (Restriction estimate)**.** Let $N \in \mathbb{N}$, $K > 0$, and $p \geqslant 1$. A function $\nu : \{1, \dots, N\} \to \mathbb{R}_{\geqslant 0}$ is said to satisfy a *p-restriction estimate with constant* K if, for every function $f : \{1, \dots, N\} \to \mathbb{C}$ satisfying $|f| \leqslant \nu$, we have

$$\|\hat{f}\|_{L^p(\mathbb{T})}^p = \int_{\mathbb{T}} |\hat{f}(\alpha)|^p d\alpha \leqslant K\|\nu\|_1^p N^{-1}.$$

As an immediate consequence of Parseval's identity, we have the following restriction estimates for the function $1_{[N]}$.

Lemma 3.2.1 (Linear restriction estimate)**.** *Let $N \in \mathbb{N}$, and let $f : [N] \to \mathbb{C}$ be such that $\|f\|_\infty \leqslant 1$. If $p \in \mathbb{R}$ with $p \geqslant 2$, then*[1]

$$\|\hat{f}\|_{L^p(\mathbb{T})}^p = \int_{\mathbb{T}} |\hat{f}(\alpha)|^p d\alpha \leqslant N\|\hat{f}\|_\infty^{p-2}.$$

Proof. Parseval's identity gives $\|\hat{f}\|_{L^2(\mathbb{T})}^2 = \|f\|_{L^2(\mathbb{Z})}^2 \leqslant N$, which implies that

$$\int_{\mathbb{T}} |\hat{f}(\alpha)|^p d\alpha \leqslant \|\hat{f}\|_\infty^{p-2}\|\hat{f}\|_{L^2(\mathbb{T})}^2 \leqslant N\|\hat{f}\|_\infty^{p-2}.$$

□

Observe that this lemma immediately implies that $1_{[N]}$ satisfies a p-restriction estimate with constant 1 for every $p \geqslant 2$. We now seek an analogue of this lemma for the degree k Fourier transform. For this we require the following auxiliary result.

Lemma 3.2.2. *Let $k, N, t \in \mathbb{N}$, and let $\mathcal{N}(k, t, N)$ denote the number of solutions $(\mathbf{x}, \mathbf{y}) \in [N]^{2t}$ to the equation $x_1^k + \cdots + x_t^k = y_1^k + \cdots + y_t^k$. Let $t_k := \lfloor k^2/2 \rfloor$. If $k \geqslant 4$, then $\mathcal{N}(k, t_k, N) \ll_k N^{2t_k - k}$.*

Proof. Let $k, N \in \mathbb{N}$ be fixed, with $k \geqslant 4$. For each (measurable) $A \subseteq \mathbb{T}$, let

$$I_s(A) := \int_A |\mathfrak{F}_k[1_{[N]}](\alpha)|^s d\alpha.$$

Let $\mathfrak{m} = \mathfrak{m}_k$ be as defined in [Woo12, §1]. From [Vau89, Lemma 5.1], it follows that if $s \geqslant k + 2$, then $I_s(\mathbb{T} \setminus \mathfrak{m}) \ll_{s,k} N^{s-k}$. Combining the mean value estimate [BDG16,

[1] *Here we use the convention* $\|0\|_\infty^0 = 1$.

Theorem 1.1] with the bound [Woo12, Theorem 2.1], we deduce that $I_{k(k+1)}(\mathfrak{m}) \ll_{\varepsilon,k} N^{k^2-1+\varepsilon}$ holds for all $\varepsilon > 0$. Following the proof of [Woo12, Theorem 3.1], we can apply Hölder's inequality and Hua's lemma to this bound to deduce that $I_s(\mathfrak{m}) \ll_{s,k} N^{s-k-\delta}$ holds for all $s \geqslant k^2 - 1$, for some $\delta = \delta(k) > 0$. The desired result now follows by observing that $\mathcal{N}(k,s,N) = I_{2s}(\mathfrak{m}) + I_{2s}(\mathbb{T} \setminus \mathfrak{m})$ and $2t_k \geqslant k^2 - 1$. □

We are now ready to state our analogue of Lemma 3.2.1 for higher degree Fourier transforms. The following previously appeared in [Lin19, Lemma D.4].

Lemma 3.2.3 (Polynomial restriction estimates)**.** *Let $k \in \mathbb{N} \setminus \{1\}$. Let $N \in \mathbb{N}$, and let $f : [N] \to \mathbb{C}$ be such that $\|f\|_\infty \leqslant 1$. If $p \in \mathbb{R}$ with $p > k^2$, then*[2]

$$\|\mathfrak{F}_k[f]\|_{L^p(\mathbb{T})}^p = \int_{\mathbb{T}} |\mathfrak{F}_k[f](\alpha)|^p d\alpha \ll_p N^{p-k}.$$

Proof. The case $k = 2$ is due to Bourgain [Bou89, Eqn. (4.1)]. Now suppose that $k \geqslant 3$. Let $t \in \mathbb{N}$, and let $\mathcal{N}(k,t,N)$ be as defined in Lemma 3.2.2. Let $p \in \mathbb{R}$ with $p > k^2$, and let $f : [N] \to \mathbb{C}$ be such that $\|f\|_\infty \leqslant 1$. If $2t \leqslant p$, then the orthogonality relations imply that

$$\||\mathfrak{F}_k[f]\|_{L^p(\mathbb{T})}^p \leqslant \|\mathfrak{F}_k[f]\|_\infty^{p-2t} \int_{\mathbb{T}} |\mathfrak{F}_k[f](\alpha)|^{2t} d\alpha \leqslant N^{p-2t}\mathcal{N}(k,t,N).$$

By the work of Vaughan [Vau86, Theorem 2], we have $\mathcal{N}(3,4,N) \ll N^5$. This proves the lemma for $k = 3$. For $k \geqslant 4$, the result now follows immediately from Lemma 3.2.2. □

3.3 Quasi-partitionable matrices

In this section we investigate the structure and properties of diagonal systems (3.2) whose coefficient matrices obey condition (I) of Theorem 3.1.2. We then establish a generalised von Neumann theorem for such systems, and also show that the set of trivial solutions for these systems is sparse in the set of all solutions.

As intimated in the introduction, we need to impose some non-singularity conditions on the coefficient matrix $\mathbf{M} = (a_{i,j})$ in order to count solutions to the system (3.2). Stating these conditions requires the following notation.

[2] *Since there are only finitely many k satisfying $k^2 < p$, the dependence on k of the implicit constant can be removed.*

Definition (μ and q functions)**.** Let $\mathbf{M} \in \mathbb{Q}^{n\times s}$ and $0 \leqslant d \leqslant n$. We write $\mu(d;\mathbf{M})$ to denote the largest number of columns of $\mathbf{M}$ whose $\mathbb{Q}$-linear span has dimension at most d. If $d \leqslant \mathrm{rank}(\mathbf{M})$, then we define

$$\mathrm{q}(d;\mathbf{M}) := \min \left| \bigcup_{i=1}^{d} \mathrm{supp}\left(\mathbf{v}^{(i)}\right) \right|,$$

where the minimum is taken over all collections of d linearly independent vectors $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)}$ in the row space of $\mathbf{M}$. By convention $\mathrm{q}(0;\mathbf{M}) = |\emptyset| = 0$.

Remark. Condition (I) of Theorem 3.1.2 can now be seen to be equivalent to the statement that $\mathrm{q}(d;\mathbf{M}) > dk^2$ holds for all $1 \leqslant d \leqslant \mathrm{rank}(\mathbf{M})$.

The study of systems of diagonal polynomial equations (3.2) was initiated in the seminal work of Davenport and Lewis [DL69]. They established that such systems possess non-zero integer solutions provided they admit non-singular real solutions and that $\mathrm{q}(d;(a_{i,j}))$ is sufficiently large in terms of d, k, and n for all $1 \leqslant d \leqslant n$.

In their work on simultaneous diagonal congruences, Low, Pitman, and Wolff [LPW88] discovered that these non-singularity conditions could be better understood by being put in the context of matroid theory. Most notably, they observed that a suitable alternative non-singularity condition could be formulated using a theorem known as *Aigner's criterion*[3] [Aig79, Proposition 6.45].

Definition (Partitionable matrix)**.** Let $\mathbf{M} \in \mathbb{Q}^{n\times s}$, for some $n, s \in \mathbb{N}$. Let $k \in \mathbb{N}$. We say that $\mathbf{M}$ is *k-partitionable* if $s = kn$ and the columns of $\mathbf{M}$ can be partitioned into k disjoint blocks of size n such that each block forms an $n \times n$ non-singular submatrix.

The following is a special case of Aigner's criterion when the matroid under consideration is a vector matroid.

Lemma 3.3.1 (Aigner's criterion)**.** *Let $\mathbf{M} \in \mathbb{Q}^{n\times kn}$, for some $k, n \in \mathbb{N}$. Then $\mathbf{M}$ is k-partitionable if and only if $\mu(d;\mathbf{M}) \leqslant dk$ holds for all $0 \leqslant d \leqslant n$.*

Proof. See [Aig79, Proposition 6.47] or [LPW88, Lemma 1]. □

Brüdern and Cook [BC92, Theorem 1] subsequently combined Aigner's criterion with the circle method to count the number of solutions to (3.2). They consider

[3] Originally proved by Edmonds [Edm65].

systems (3.2) whose coefficient matrices $\mathbf{M} = (a_{i,j}) \in \mathbb{Z}^{n \times s}$ are such that there exists an $n \times (ns_1 + 1)$ submatrix $\mathbf{M}'$ of $\mathbf{M}$ such that $\mu(d; \mathbf{M}') \leqslant ds_1$ holds for all $0 \leqslant d < n$. Provided $s_1 = s_1(k)$ is sufficiently large, and that (3.2) has non-singular real and p-adic solutions for every prime p, they prove that there are $\gg_{\mathbf{M}} N^{s-kn}$ solutions $\mathbf{x} \in [N]^s$ to (3.2).

For $n = 2$ and $s = 9$, such a result had previously been obtained by Cook [Coo71]. In this case, as in Corollary 3.1.3, the condition on $\mathbf{M}$ may be replaced with the condition that every non-zero vector in the row space of $\mathbf{M}$ has at least 5 non-zero entries.

From Aigner's criterion, we see that this condition on $\mathbf{M}$ implies that $\mathbf{M}$ contains an $n \times (ns_1 + 1)$ submatrix $\mathbf{M}'$ such that every $n \times ns_1$ submatrix of $\mathbf{M}'$ is partitionable. This leads us to consider what we have termed *quasi-partitionable matrices.*

Definition (Quasi-partitionable matrix)**.** A non-empty matrix $\mathbf{M} \in \mathbb{Q}^{n \times s}$ is called *quasi-q-partitionable* if $s \geqslant nq$ and $\mu(d; \mathbf{M}) \leqslant dq$ holds for all $0 \leqslant d < n$. We say that $\mathbf{M}$ is *quasi-partitionable* if $\mathbf{M} \in \mathbb{Q}^{n \times s}$ is quasi-q-partitionable for some $q \in \mathbb{N}$.

Corollary 3.3.2. *Let $q \in \mathbb{N}$, and $\mathbf{M} \in \mathbb{Q}^{n \times s}$. If $\mathbf{M}$ is quasi-q-partitionable, then $\mathbf{M}$ has full rank and every $n \times nq$ submatrix of $\mathbf{M}$ is q-partitionable.*

Proof. If $\mathbf{M}$ is quasi-q-partitionable, then $s > (n-1)q \geqslant \mu(n-1; \mathbf{M})$, which implies that $\mathbf{M}$ has rank n. The rest of the corollary follows immediately from Aigner's criterion. □

We now show that all of the systems (3.2) we consider, namely those obeying condition (I) of Theorem 3.1.2, may be 'semi-decomposed' into systems whose coefficient matrices are quasi-partitionable.

Lemma 3.3.3 (Decomposition lemma)**.** *Let $n, q, s \in \mathbb{N}$, and let $\mathbf{M} \in \mathbb{Q}^{n \times s}$ be a matrix of rank n with no zero columns. If $\mathrm{q}(d; \mathbf{M}) > dq$ for all $1 \leqslant d \leqslant n$, then $\mathbf{M}$ is equivalent to a block upper triangular matrix with diagonal $(\mathbf{M}_1, \dots, \mathbf{M}_r)$, where each $\mathbf{M}_i \in \mathbb{Z}^{n_i \times s_i}$ is quasi-q-partitionable and $s_i > n_i q$.*

Proof. We proceed by induction on n. If $n = 1$, then the hypotheses on $\mathbf{M}$ imply that $\mathbf{M}$ is quasi-q-partitionable. Suppose then that $n \geqslant 2$, and assume the induction hypothesis that if $1 \leqslant n' < n$ and $\mathbf{M}' \in \mathbb{Q}^{n' \times s'}$ is a full rank matrix with no zero columns which satisfies $\mathrm{q}(d;\mathbf{M}) > dq$ for all $1 \leqslant d \leqslant n'$, then $\mathbf{M}'$ satisfies the conclusion of the theorem.

We may henceforth assume that $\mathbf{M}$ is not quasi-q-partitionable, since otherwise we are done. From the bound $s = \mathrm{q}(n;\mathbf{M}) > nq$, we deduce that there exists a minimal $d_0 \in [n-1]$ such that $s_0 := \mu(d_0;\mathbf{M}) > d_0 q$. Hence, $\mathbf{M}$ is equivalent to a block upper triangular matrix with diagonal $(\mathbf{M}_0, \mathbf{M}')$, for some full rank matrices $\mathbf{M}_0 \in \mathbb{Q}^{d_0 \times s_0}$ and $\mathbf{M}' \in \mathbb{Q}^{n' \times s'}$ without zero columns.

Note that $\mu(d;\mathbf{M}_0) \leqslant dq$ holds for all $0 \leqslant d < d_0$ by the minimality of d_0. Since $s_0 > d_0 q$, we therefore deduce that $\mathbf{M}_0$ is quasi-q-partitionable. We also note that $\mathrm{q}(d;\mathbf{M}') \geqslant \mathrm{q}(d;\mathbf{M}) > dq$ holds for all $0 \leqslant d \leqslant n'$. Thus, by the induction hypothesis, $\mathbf{M}'$ satisfies the conclusion of the lemma, and therefore so does $\mathbf{M}$. □

3.3.1 Fourier control

We now use the properties of quasi-partitionable matrices to control counting operators for diagonal polynomial systems (3.2). More precisely, given a matrix $\mathbf{M} \in \mathbb{Z}^{n \times s}$ satisfying condition (I) of Theorem 3.1.2, we consider sums of the form

$$\Lambda_{\mathbf{M}}(f_1, \ldots, f_s) = \sum_{\mathbf{M}\mathbf{x}=\mathbf{0}} f_1(x_1) \cdots f_s(x_s), \tag{3.5}$$

where $f_1, \ldots, f_s : [N] \to \mathbb{C}$. Such counting operators are frequently treated in the arithmetic combinatorics literature, see [Cha20, CLP21, GS16, GT10, Tao12]. The primary method of understanding counting operators is via a *generalised von Neumann theorem*. This term is used to describe any result which asserts that $\Lambda_{\mathbf{M}}(f_1, \ldots, f_s) \approx \Lambda_{\mathbf{M}}(g_1, \ldots, g_s)$ holds whenever $\|f_i - g_i\|$ is 'small' for all i with respect to some (semi-)norm $\|\cdot\|$. Determining precisely which (semi-)norms are admissible for a given family of counting operators is also a widely studied problem in its own right, see [GW10] for further details.

In this subsection, we establish a generalised von Neumann theorem using the norm given by $f \mapsto \|\hat{f}\|_\infty$. We begin with the observation that, by orthogonality, the sum (3.5) is equal to the integral

$$\int_{\mathbb{T}^n} \prod_{j=1}^{s} \hat{f}_j(\boldsymbol{\alpha} \cdot \mathbf{c}^{(j)}) d\boldsymbol{\alpha}.$$

Here, we have written $\mathbf{c}^{(j)} \in \mathbb{Z}^n$ to denote the jth column of $\mathbf{M}$. Our goal is to use the L^p norms of the $\hat{f}_j$ to bound this integral.

Let $\mathbf{M}_i \in \mathbb{Q}^{n_i \times s_i}$ for $1 \leqslant i \leqslant r$, and consider a block upper triangular matrix $\mathbf{M}$ with diagonal $(\mathbf{M}_1, \ldots, \mathbf{M}_r)$. Typically, the jth column of $\mathbf{M}$ is denoted by $\mathbf{c}^{(j)}$. However, as we are interested in utilising the properties of the $\mathbf{M}_i$, it is convenient for us to reindex the columns of $\mathbf{M}$ using (i, j) where $1 \leqslant i \leqslant r$ and $j \in [s_i]$. Hence, we write $\mathbf{c}^{(i,j)}$ to denote the column of $\mathbf{M}$ which intersects the jth column of $\mathbf{M}_i$. Explicitly, $\mathbf{c}^{(i,j)}$ is the $(j + \sum_{1 \leqslant t < i} s_t)$th column of $\mathbf{M}$.

Theorem 3.3.4 (L^p control for integral operators)**.** *Let $n, q, r, s \in \mathbb{N}$, and let $\mathbf{M} \in \mathbb{Z}^{n \times s}$ be a block upper triangular matrix with diagonal $(\mathbf{M}_1, \ldots, \mathbf{M}_r)$, for some non-empty $\mathbf{M}_i \in \mathbb{Z}^{n_i \times s_i}$. As described above, denote the columns of $\mathbf{M}$ by $\mathbf{c}^{(i,j)} \in \mathbb{Z}^n$ for $i \in [r]$ and $j \in [s_i]$. Let $p_i := s_i/n_i$ for each $i \in [r]$. If every $\mathbf{M}_i$ is quasi-q-partitionable, then for any collection of integrable functions $\psi_{i,j} : \mathbb{T} \to \mathbb{C}$ for $i \in [r]$ and $j \in [s_i]$, we have*

$$\int_{\mathbb{T}^n} \prod_{i=1}^{r} \prod_{j=1}^{s_i} |\psi_{i,j}(\boldsymbol{\alpha} \cdot \mathbf{c}^{(i,j)})| d\boldsymbol{\alpha} \leqslant \prod_{i=1}^{r} \prod_{j=1}^{s_i} \|\psi_{i,j}\|_{L^{p_i}(\mathbb{T})}. \tag{3.6}$$

Proof. Let $\mathcal{B} := \{\mathbf{J} = (J_1, \ldots, J_r) : J_i \subseteq [s_i], |J_i| = n_i q\}$. Note that

$$|\mathcal{B}| = \prod_{i=1}^{r} \binom{s_i}{n_i q}.$$

For each $i \in [r]$ and $j \in [s_i]$, let $a_{i,j}$ be an arbitrary non-negative real number, and let

$$m_i := \binom{s_i - 1}{n_i q - 1} \prod_{\substack{t=1 \\ t \neq i}}^{r} \binom{s_t}{n_t q} = \frac{q|\mathcal{B}|}{p_i}.$$

Observe that, for any $j \in [s_i]$, the number of $\mathbf{J} = (J_1, \ldots, J_r) \in \mathcal{B}$ such that $j \in J_i$ is given by m_i. This provides us with the identity

$$\prod_{i=1}^{r} \prod_{j=1}^{s_i} a_{i,j} = \prod_{\mathbf{J} \in \mathcal{B}} \prod_{i=1}^{r} \prod_{j \in J_i} a_{i,j}^{1/m_i}. \tag{3.7}$$

Let $i \in [r]$. By Corollary 3.3.2, any set $J \subseteq [s_i]$ with $|J| = n_i q$ admits a partition $J = I_1^{(i,J)} \cup \cdots \cup I_q^{(i,J)}$ such that, for each $1 \leqslant t \leqslant q$, the columns of $\mathbf{M}_i$ indexed by $I_t^{(i,J)}$ form a non-singular $n_i \times n_i$ matrix. Thus, given any $\mathbf{J} \in \mathcal{B}$, we obtain the identity

$$\prod_{\mathbf{u} \in [q]^r} \prod_{j \in I_{u_i}^{(i,J_i)}} a_{i,j} = \prod_{t=1}^{q} \prod_{\substack{\mathbf{u} \in [q]^r \\ u_i = t}} \prod_{j \in I_t^{(i,J_i)}} a_{i,j} = \prod_{j \in J_i} a_{i,j}^{q^{r-1}}. \tag{3.8}$$

Combining this identity with (3.7) and applying Hölder's inequality shows that the left-hand side of (3.6) is bounded above by

$$\prod_{\mathbf{J}\in\mathcal{B}}\prod_{\mathbf{u}\in[q]^r}\left(\int_{\mathbb{T}^n}\prod_{i=1}^{r}\prod_{j\in I_{u_i}^{(i,J_i)}}|\psi_{i,j}(\boldsymbol{\alpha}\cdot\mathbf{c}^{(i,j)})|^{p_i}d\boldsymbol{\alpha}\right)^{(q^r|\mathcal{B}|)^{-1}}.$$

Now fix some $\mathbf{u}\in[q]^r$ and $\mathbf{J}\in\mathcal{B}$. For each $i\in[r]$, let $\mathbf{A}_i$ be the $n_i\times n_i$ submatrix of $\mathbf{M}_i$ formed by the columns of $\mathbf{M}_i$ indexed by $I_{u_i}^{(i,J_i)}$. Our choice of the $I_{u_i}^{(i,J_i)}$ ensures that each $\mathbf{A}_i$ is non-singular. Hence, every $n\times n$ block upper triangular matrix with diagonal $(\mathbf{A}_1,\dots,\mathbf{A}_r)$ is non-singular (this can be seen directly from the structure, or by noting that the determinant of such a matrix is given by the product of the determinants of the $\mathbf{A}_i$). Thus, the set of vectors $\{\mathbf{c}^{(i,j)}: i\in[r], j\in I_{u_i}^{(i,J_i)}\}$ is linearly independent. We may therefore perform a change of variables to deduce that

$$\left(\int_{\mathbb{T}^n}\prod_{i=1}^{r}\prod_{j\in I_{u_i}^{(i,J_i)}}|\psi_{i,j}(\boldsymbol{\alpha}\cdot\mathbf{c}^{(i,j)})|^{p_i}d\boldsymbol{\alpha}\right)^{(q^r|\mathcal{B}|)^{-1}}=\prod_{i=1}^{r}\prod_{j\in I_{u_i}^{(i,J_i)}}\|\psi_{i,j}\|_{L^{p_i}(\mathbb{T})}^{(q^{r-1}m_i)^{-1}}.$$

The theorem now follows from (3.7) and (3.8). $\square$

We now return to the problem of bounding the counting operators (3.5). For the applications in §3.5, we require control for counting operators with weights f_i which may be unbounded as $N\to\infty$. Such a result is given for single equations (3.1) in [CLP21, Lemma C.2], and we now provide a generalisation for systems (3.2).

Lemma 3.3.5 (Relative Fourier control)**.** *Let $k,n,s\in\mathbb{N}$ with $k\geqslant 2$. Let $p:=k^2+\frac{1}{2n}$, and let $\eta:=(2k^2n+2)^{-1}$. Let $N\in\mathbb{N}$, and suppose that $\nu_1,\dots,\nu_s:[N]\to\mathbb{R}_{\geqslant 0}$ are non-zero functions which each satisfy a p-restriction estimate with constant K. Let $\mathbf{M}\in\mathbb{Z}^{n\times s}$ be a matrix of rank n with no zero columns. If $\mathbf{M}$ satisfies condition (I) of Theorem 3.1.2, then, for any functions $f_1,\dots,f_s:[N]\to\mathbb{C}$ such that $|f_i|\leqslant\nu_i$, we have*

$$\left|\sum_{\mathbf{Mx}=\mathbf{0}}\prod_{i=1}^{s}\frac{f_i(x_i)}{\|\nu_i\|_1}\right|\leqslant K^nN^{-n}\prod_{i=1}^{s}\left(\frac{\|\hat{f}_i\|_\infty}{\|\nu_i\|_1}\right)^{\eta}.$$

Proof. By applying Lemma 3.3.3, and relabelling the f_i if necessary, we may assume that $\mathbf{M}$ is a block upper triangular matrix with diagonal $(\mathbf{M}_1,\dots,\mathbf{M}_r)$, where each $\mathbf{M}_i\in\mathbb{Z}^{n_i\times s_i}$ is quasi-partitionable and satisfies $s_i>k^2n_i$. For each $i\in[r]$ and $j\in[s_i]$, let $\psi_{i,j} : \mathbb{T} \to \mathbb{C}$ be defined by $\psi_{i,j} := \|\nu_l\|_1^{-1}\hat{f}_l$, where $l \in [s]$ is such that the lth column of $\mathbf{M}$ intersects the jth column of $\mathbf{M}_i$.

Let $p_i = s_i/n_i$ for each $i \in [r]$, and note that $p_i > p$. From orthogonality and Theorem 3.3.4 we obtain the bound

$$\left|\sum_{\mathbf{Mx}=\mathbf{0}} \prod_{i=1}^{s} \frac{f_i(x_i)}{\|\nu_i\|_1}\right| \leqslant \prod_{i=1}^{r}\prod_{j=1}^{s_i} \|\psi_{i,j}\|_{p_i} \leqslant \prod_{i=1}^{r}\prod_{j=1}^{s_i} \|\psi_{i,j}\|_p^{\frac{p}{p_i}} \|\psi_{i,j}\|_\infty^{1-\frac{p}{p_i}}. \tag{3.9}$$

From the restriction estimates for the ν_i, we deduce that

$$\prod_{i=1}^{r}\prod_{j=1}^{s_i} \|\psi_{i,j}\|_p^{\frac{p}{p_i}} \leqslant \prod_{i=1}^{r}\prod_{j=1}^{s_i} \left(\frac{K}{N}\right)^{\frac{1}{p_i}} = \left(\frac{K}{N}\right)^{n}. \tag{3.10}$$

Observe that our choice of η gives $\eta p_i \leqslant p_i - p$. Hence, for each $i \in [r]$ and $j \in [s_i]$, the hypothesis $|f_i| \leqslant \nu_i$ implies that $\|\psi_{i,j}\|_\infty^{1-\frac{p}{p_i}} \leqslant \|\psi_{i,j}\|_\infty^{\eta} \leqslant 1$. The lemma may now be deduced from (3.9) by invoking (3.10). □

Finally, by applying a telescoping identity, we obtain the desired generalised von Neumann theorem.

Theorem 3.3.6 (Relative generalised von Neumann)**.** *Let $k, n, s \in \mathbb{N}$ with $k \geqslant 2$. Let $p := k^2 + \frac{1}{2n}$, and let $\eta := (2k^2n+2)^{-1}$. Let $N \in \mathbb{N}$, and suppose that $\nu_1, \ldots, \nu_s, \mu_1, \ldots, \mu_s : [N] \to \mathbb{R}_{\geqslant 0}$ are non-zero functions which each satisfy a p-restriction estimate with constant K. Let $\mathbf{M} \in \mathbb{Z}^{n\times s}$ be a matrix of rank n with no zero columns. If $\mathbf{M}$ satisfies condition (I) of Theorem 3.1.2, then, for any functions $f_1, \ldots, f_s, g_1, \ldots, g_s : [N] \to \mathbb{C}$ such that $|f_i| \leqslant \nu_i$ and $|g_i| \leqslant \mu_i$, we have*

$$\begin{aligned}\left|\sum_{\mathbf{Mx}=\mathbf{0}} \left(\frac{f_1(x_1)}{\|\nu_1\|_1}\cdots\frac{f_s(x_s)}{\|\nu_s\|_1} - \frac{g_1(x_1)}{\|\mu_1\|_1}\cdots\frac{g_s(x_s)}{\|\mu_s\|_1}\right)\right| \\ \leqslant 2sK^nN^{-n}\max_{1\leqslant i\leqslant s}\left\|\frac{\hat{f}_i}{\|\nu_i\|_1} - \frac{\hat{g}_i}{\|\mu_i\|_1}\right\|_\infty^{\eta}.\end{aligned}$$

Proof. For each $i \in [s]$, define functions $h_i : [N] \to \mathbb{C}$ and $\tau_i : [N] \to \mathbb{R}_{\geqslant 0}$ by

$$h_i := \frac{f_i}{\|\nu_i\|_1} - \frac{g_i}{\|\mu_i\|_1} \quad ; \quad \tau_i := \frac{\nu_i}{\|\nu_i\|_1} + \frac{\mu_i}{\|\mu_i\|_1}.$$

By [CLP21, Lemma C.1], the functions $\tau_1, \ldots, \tau_s$ each satisfy a p-restriction estimate with constant K. Note that $|h_i| \leqslant \tau_i$ and $\|\tau_i\|_1 = 2$ for all $i \in [s]$. The result now follows by applying the triangle inequality and Lemma 3.3.5 to the telescoping identity

$$\sum_{\mathbf{Mx}=\mathbf{0}} \left(\frac{f_1(x_1)}{\|\nu_1\|_1} \cdots \frac{f_s(x_s)}{\|\nu_s\|_1} - \frac{g_1(x_1)}{\|\mu_1\|_1} \cdots \frac{g_s(x_s)}{\|\mu_s\|_1} \right)$$
$$= 2\sum_{r=1}^{s} \sum_{\mathbf{Mx}=\mathbf{0}} \left(\prod_{i=1}^{r-1} \frac{f_i(x_i)}{\|\nu_i\|_1} \right) \frac{h_r(x_r)}{\|\tau_r\|_1} \left(\prod_{i=r+1}^{s} \frac{g_i(x_i)}{\|\mu_i\|_1} \right).$$

□

3.3.2 Counting trivial solutions

We close this section by exhibiting a second consequence of Theorem 3.3.4: we can bound the number of trivial solutions to (3.1.1). In particular, we show that the set of trivial solutions is sparse in the set of all solutions. From this it follows that one can obtain non-trivial solutions to (3.1.1) over a set S by taking N sufficiently large and showing that a positive proportion of all solutions $\mathbf{x} \in [N]^s$ lie in $(S \cap [N])^s$.

Theorem 3.3.7. *Let $k, n, s \in \mathbb{N}$ with $k \geqslant 2$. Let $\mathbf{M} = (a_{i,j}) \in \mathbb{Z}^{n\times s}$ be a matrix of rank n with no zero columns. Let $\delta = \delta(k, n) = 2kn(k^2n+1)^{-1}$. If $\mathbf{M}$ satisfies condition (I) of Theorem 3.1.2, then there are $O_{n,s}(N^{s-kn+\delta-1})$ trivial solutions $\mathbf{x} \in [N]^s$ to (3.2).*

Proof. By the union bound, it suffices to show that $\mathcal{N}_{\mathbf{M}}(u, v; N) \ll_{\mathbf{M}} N^{s-kn+\delta-1}$ holds for all $u, v \in [s]$ with $u < v$, provided $N \in \mathbb{N}$ is sufficiently large. Here, $\mathcal{N}_{\mathbf{M}}(u, v; N)$ denotes the number of solutions $\mathbf{x} \in [N]^s$ to (3.2) with $x_u = x_v$.

By Lemma 3.3.3, we may assume that $\mathbf{M}$ is a block upper triangular matrix with diagonal $(\mathbf{M}_1, \dots, \mathbf{M}_r)$, where each $\mathbf{M}_i \in \mathbb{Z}^{n_i\times s_i}$ is quasi-k^2-partitionable and satisfies $s_i > k^2 n_i$. As in the statement of Theorem 3.3.4, denote the columns of $\mathbf{M}$ by $\mathbf{c}^{(i,j)}$. Hence, we can find indices $i_u, i_v \in [r]$, $j_u \in [s_{i_u}]$, and $j_v \in [s_{i_v}]$ such that $\mathbf{c}^{(i_u,j_u)}$ and $\mathbf{c}^{(i_v,j_v)}$ correspond to the uth and vth columns of $\mathbf{M}$ respectively. By orthogonality, we observe that

$$\mathcal{N}_{\mathbf{M}}(u, v; N) = \int_{\mathbb{T}^n} \mathfrak{F}_k[1_{[N]}](\boldsymbol{\alpha} \cdot (\mathbf{c}^{(i_u,j_u)} + \mathbf{c}^{(i_v,j_v)})) \prod_{i=1}^{r} \prod_{j=1}^{s_i} \psi_{i,j}(\boldsymbol{\alpha} \cdot \mathbf{c}^{(i,j)}) d\boldsymbol{\alpha},$$

where, for all $\alpha \in \mathbb{T}$,

$$\psi_{i,j}(\alpha) := \begin{cases} 1, & \text{if } (i, j) \in \{(i_u, j_u), (i_v, j_v)\}; \\ \mathfrak{F}_k[1_{[N]}](\alpha), & \text{otherwise.} \end{cases}$$

Thus, by setting $p_i := s_i/n_i$, Theorem 3.3.4 provides us with the upper bound

$$\mathcal{N}_{\mathbf{M}}(u,v;N) \leqslant \|\mathfrak{F}_k[1_{[N]}]\|_\infty \prod_{i=1}^{r}\prod_{j=1}^{s_i}\|\psi_{i,j}\|_{L^{p_i}(\mathbb{T})}.$$

For each $i \in [r]$ and $j \in [s_i]$ with $(i,j) \notin \{(i_u,j_u),(i_v,j_v)\}$, Lemma 3.2.3 gives

$$\|\psi_{i,j}\|_{L^{p_i}(\mathbb{T})} \ll_{n,s} N^{1-\frac{k}{p_i}} = N^{1-\frac{kn_i}{s_i}}.$$

Hence, on noting that $\|\mathfrak{F}_k[1_{[N]}]\|_\infty = N$, we find that

$$\mathcal{N}_{\mathbf{M}}(u,v;N) \ll_{n,s} N^{s-kn+1} \cdot N^{-2+\frac{k}{p_{i_u}}+\frac{k}{p_{i_v}}}.$$

The result now follows on noting that $p_i \geqslant k^2 + \frac{1}{n_i} \geqslant k^2 + \frac{1}{n}$. □

3.4 Induction on colours

The goal of this section is to show that the task of establishing partition regularity for the system (3.2) can be accomplished by counting solutions over dense sets and multiplicatively syndetic sets. We begin by recalling the definition of a multiplicatively syndetic set.

Definition (Multiplicatively syndetic sets)**.** Let $M \in \mathbb{N}$. A set $S \subseteq \mathbb{N}$ is called *multiplicatively* $[M]$*-syndetic* if $S \cap \{x, 2x, \dots, Mx\} \neq \emptyset$ for every $x \in \mathbb{N}$. We say that S is a *multiplicatively syndetic set* if S is multiplicatively $[M]$-syndetic for some $M \in \mathbb{N}$.

Remark. In [CLP21], multiplicatively $[M]$-syndetic set are also called M*-homogeneous sets.*

Chow, Lindqvist, and Prendiville [CLP21] observed that a homogeneous system of equations such as (3.2) is partition regular if it has a solution over every multiplicatively syndetic set. In fact, one can show that the converse statement is also true, see [Cha20]. This argument enables us to reduce Theorem 3.1.2 to the following.

Theorem 3.4.1. *Let $M \in \mathbb{N}$, and let $\mathbf{M} = (a_{i,j})$ be a $n \times s$ integer matrix of rank n. Let $k \geqslant 2$. If $\mathbf{M}$ satisfies the columns condition and condition* (I) *of Theorem 3.1.2, then there exist positive constants $N_0 = N_0(k,M;\mathbf{M}) \in \mathbb{N}$ and $c_0 = c_0(k,M;\mathbf{M}) > 0$ such that the following is true. If $S \subseteq \mathbb{N}$ is a multiplicatively $[M]$-syndetic set, and $N \geqslant N_0$, then there are at least c_0N^{s-kn} non-trivial solutions $\mathbf{x} \in (S \cap [N])^s$ to (3.2).*

Proof of Theorem 3.1.2 given Theorem 3.4.1. Note that we may assume $\mathbf{M}$ has rank n by deleting any linearly dependent rows. The result now follows by the induction on colours argument given in [CLP21, §4.2] and [CLP21, §13.2]. □

Observe that any diagonal polynomial equation of degree k which satisfies the columns condition can be written in the form

$$\sum_{i=1}^{s} a_i x_i^k = \sum_{j=1}^{t} b_j y_j^k, \tag{3.11}$$

for some integers $s \in \mathbb{N}$, $t \geqslant 0$, and $a_1, \dots, a_s, b_1, \dots, b_t \in \mathbb{Z} \setminus \{0\}$ are such that $a_1 + \dots + a_s = 0$. For equations of this form, Chow, Lindqvist, and Prendiville [CLP21] establish partition regularity (Theorem 3.1.1) by showing that one can find (many) solutions to (3.11) with x_i lying in a dense set of smooth numbers and y_j lying in a multiplicatively syndetic set.

We seek a similar reformulation for the systems considered in Theorem 3.1.2. We start by listing a number of equivalent definitions of the columns condition.

Proposition 3.4.2. *Let $\mathbf{M} \in \mathbb{Z}^{h \times k}$ be a matrix of rank h. Then the following are all equivalent.*

(i) *$\mathbf{M}$ obeys the columns condition;*

(ii) *the system of equations $\mathbf{M}\mathbf{x} = \mathbf{0}$ is partition regular;*

(iii) *for every $\mathbf{v} = (v_1, \dots, v_k) \in \mathbb{Q}^k \setminus \{\mathbf{0}\}$ in the row space of $\mathbf{M}$, there exists a non-empty set $J \subseteq [k]$ such that $v_j \neq 0$ for all $j \in J$, and $\sum_{j \in J} v_j = 0$;*

(iv) *there exist positive integers $n, s \in \mathbb{N}$ and non-negative integers $m, t \in \mathbb{Z}_{\geqslant 0}$ with $h = n + m$ and $k = s + t$ such that the following is true. The matrix $\mathbf{M}$ is equivalent to a matrix of the form*

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}, \tag{3.12}$$

for some matrices $\mathbf{A} \in \mathbb{Z}^{n \times s}, \mathbf{B} \in \mathbb{Z}^{n \times t}, \mathbf{C} \in \mathbb{Z}^{m \times t}$ such that $\mathbf{A}$ is a matrix of rank n whose columns sum to $\mathbf{0}$, and $\mathbf{C}$ obeys the columns condition.[4]

[4] *Note that $\mathbf{C}$ is allowed to contain zero columns.*

Proof. The equivalence of (i) and (ii) is provided by [Rad33, Satz IV], whilst the equivalence of (ii) and (iii) follows from [Rad43, Lemma 4].

Suppose that $\mathbf{M}$ satisfies (iv), and let $\mathbf{M}'$ be the matrix given in (3.12). Let $\mathbf{v} = (v_1, \ldots, v_k) \in \mathbb{Q}^k$ be a non-zero vector in the row space of $\mathbf{M}'$. Since the columns of $\mathbf{A}$ sum to $\mathbf{0}$, we see that $v_1 + \ldots + v_s = 0$. Thus, if $v_1, \ldots, v_s$ are not all zero, then we may take $J = \{i \in [s] : v_i \neq 0\} \neq \emptyset$. If, on the other hand, we have $v_1 = \ldots = v_s = 0$, then the hypothesis that $\mathbf{A}$ has full rank implies that $(v_{s+1}, \ldots, v_k) \in \mathbb{Q}^t$ is a non-zero element of the row space of $\mathbf{C}$. Since property (iii) holds for $\mathbf{C}$, we deduce that there exists $J \subseteq [k] \setminus [s]$ such that $\sum_{j \in J} v_j = 0$ and $v_j \neq 0$ for all $j \in J$. We therefore find that (iii) holds for $\mathbf{M}'$. Since vectors in the row space of $\mathbf{M}$ are just permutations of vectors in the row space of $\mathbf{M}'$, we conclude that (iii) holds for $\mathbf{M}$.

Finally, suppose that (i) and (iii) both hold for $\mathbf{M}$. Let $I = J_1$ and $J = [k] \setminus I$, where $[k] = J_1 \cup \cdots \cup J_p$ is the partition provided by the columns condition. By permuting the columns of $\mathbf{M}$, we may assume that $I = [s]$ for some $0 < s \leqslant k$. Let $\mathbf{M}'$ be the $h \times s$ submatrix of $\mathbf{M}$ formed from the columns of $\mathbf{M}$ indexed by I. By performing elementary row operations, we may assume that the bottom $(h - n)$ rows of $\mathbf{M}'$ are identically zero, where n is the rank of $\mathbf{M}'$. By performing these same operations to $\mathbf{M}$, we see that $\mathbf{M}$ is equivalent to a matrix of the form (3.12). Our choice of I ensures that $\mathbf{A}$ has rank n and that the columns of $\mathbf{A}$ sum to $\mathbf{0}$. By considering only linear combinations of the bottom $h - n$ rows of the matrix (3.12), we see that $\mathbf{C}$ satisfies property (iii). Hence, by the equivalence of (i) and (iii), we deduce that $\mathbf{C}$ obeys the columns condition. We have therefore shown that (iv) holds. □

This proposition therefore shows that to prove Theorem 3.1.2 we need only consider systems of equations of the form

$$\begin{aligned} \mathbf{A}\mathbf{x}^{\otimes k} &= \mathbf{B}\mathbf{y}^{\otimes k}; \\ \mathbf{0} &= \mathbf{C}\mathbf{y}^{\otimes k}. \end{aligned} \tag{3.13}$$

where $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are as described in (iv) of Proposition 3.4.2. Here we have also recalled the notation $(x_1, \ldots, x_s)^{\otimes k} := (x_1^k, \ldots, x_s^k)$. Thus, Theorem 3.4.1 may be rewritten to the following.

Theorem 3.4.3. *Let $k \in \mathbb{N} \setminus \{1\}$. Let $\mathbf{A} \in \mathbb{Z}^{n \times s}, \mathbf{B} \in \mathbb{Z}^{n \times t}$, and $\mathbf{C} \in \mathbb{Z}^{m \times t}$, for some $n, s \in \mathbb{N}$ and $m, t \in \mathbb{Z}_{\geqslant 0}$. Let $\mathbf{M}$ be the matrix given by (3.12). Suppose that $\mathbf{A}$ is a* *matrix of rank n whose columns sum to $\mathbf{0}$, and that $\mathbf{C}$ obeys the columns condition. If $\mathbf{M}$ satisfies condition (I) of Theorem 3.1.2, then for every $M \in \mathbb{N}$, there exists $N_0 = N_0(k, M; \mathbf{M}) \in \mathbb{N}$ and $c_0 = c_0(k, M; \mathbf{M}) > 0$ such that the following is true. If $S \subseteq \mathbb{N}$ is a multiplicatively $[M]$-syndetic set, and $N \geqslant N_0$, then there are at least $c_0 N^{s+t-k(m+n)}$ non-trivial solutions $(\mathbf{x}, \mathbf{y}) \in (S \cap [N])^{s+t}$ to (3.13).*

Proof of Theorem 3.4.1 given Theorem 3.4.3. This is an immediate consequence of the equivalence between (i) and (iv) in Proposition 3.4.2. □

To prove Theorem 3.4.3, we follow the approach of Chow, Lindqvist, and Prendiville by seeking solutions to (3.13) with the x_i lying in a dense set, and the y_j lying in a multiplicatively syndetic set. Such an approach has the advantage that it allows us to address both Theorem 3.1.2 and Theorem 3.1.4 simultaneously by proving a stronger result (Theorem 3.4.5).

To elucidate this argument further, we first recall the fact that multiplicatively syndetic sets have positive (lower) density.

Lemma 3.4.4 (Density of multiplicatively syndetic sets)**.** *Let $M, N \in \mathbb{N}$. If $S \subseteq \mathbb{N}$ is multiplicatively $[M]$-syndetic, then*

$$|S \cap [N]| \geqslant \frac{1}{M} \left\lfloor \frac{N}{M} \right\rfloor .$$

Proof. See [CLP21, Lemma 4.2] or [Cha20, Lemma 3.1]. □

Our earlier remarks therefore show that partition regularity is a special case of density regularity, namely the case where the dense set we seek solutions over is multiplicatively syndetic. Combining Proposition 3.4.2 with these observations allows us to generalise [CLP21, Theorem 12.1] to systems of equations. We also take this opportunity to impose a number of helpful properties on the matrices $\mathbf{M}$ in order to simplify our arguments in §3.6.

Theorem 3.4.5 (Dense-syndetic regularity)**.** *Let $k \in \mathbb{N} \setminus \{1\}$. Let $n, s \in \mathbb{N}$, and $m, t \in \mathbb{Z}_{\geqslant 0}$. Let $\mathbf{A} \in \mathbb{Z}^{n \times s}$, $\mathbf{B} \in \mathbb{Z}^{n \times t}$, and $\mathbf{C} \in \mathbb{Z}^{m \times t}$. Let $\delta > 0$, and $M \in \mathbb{N}$. Let $\mathbf{M}$ be the matrix defined by (3.12). Suppose that the following conditions all hold.*

(i) *the matrix $\mathbf{A}$ has rank n and no zero columns;*

(ii) *the columns of* $\mathbf{A}$ *sum to* $\mathbf{0}$*;*

(iii) *if* $m, t > 0$*, then* $\mathbf{C}$ *obeys the columns condition;*

(iv) *the matrix* $\mathbf{M}$ *satisfies condition* (I) *of Theorem 3.1.2;*

(v) *the matrix* $\mathbf{A}$ *contains a non-singular* $n \times n$ *diagonal submatrix;*

(vi) *for each* $i \in [n]$*, the entries in the* i*th row of* $\mathbf{A}$ *are coprime;*

(vii) *every entry of* $\mathbf{B}$ *is divisible by every non-zero entry of* $\mathbf{A}$*;*

(viii) *if* $m, t > 0$*, then, for any* $M' \in \mathbb{N}$ *and multiplicatively* $[M']$*-syndetic set* $S \subseteq \mathbb{N}$*, there are* $\gg_{M',\mathbf{C}} N^{t-km}$ *solutions* $\mathbf{y} \in (S \cap [N])^t$ *to* $\mathbf{C}\mathbf{y}^{\otimes k} = \mathbf{0}$*, provided that* N *is sufficiently large relative to* M' *and* $\mathbf{C}$*.*

Then there exists a positive integer $N_0 = N_0(\delta, k, M, \mathbf{M})$ *and a positive constant* $c_0 = c_0(\delta, k, M, \mathbf{M}) > 0$ *such that the following is true. Let* $S \subseteq \mathbb{N}$ *be a multiplicatively* $[M]$*-syndetic set. Let* $N \in \mathbb{N}$*, and* $A \subseteq [N]$ *be such that* $|A| \geqslant \delta N$*. If* $N \geqslant N_0$*, then there are at least* $c_0 N^{s+t-k(n+m)}$ *solutions* $(\mathbf{x}, \mathbf{y}) \in A^s \times (S \cap [N])^t$ *to (3.13).*

Remark. Condition (viii) simply asserts that Theorem 3.4.1 is true if "$\mathbf{M}$" and "(3.2)" are replaced by "$\mathbf{C}$" and "$\mathbf{C}\mathbf{y}^{\otimes k} = \mathbf{0}$" respectively.

Proof of Theorem 3.4.3 given Theorem 3.4.5. Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{M}$ be as given in the statement of Theorem 3.4.3. Observe that the hypotheses of Theorem 3.4.3 guarantee that $\mathbf{A}, \mathbf{C}$, and $\mathbf{M}$ obey conditions (i)-(iv) of Theorem 3.4.5. By performing elementary row operations, we can ensure that condition (v) is also satisfied. We henceforth assume that these five conditions all hold.

Note that, by taking N sufficiently large, Theorem 3.3.7 implies that if we can find $cN^{s+t-k(n+m)}$ solutions $(\mathbf{x}, \mathbf{y}) \in A^s \times (S \cap [N])^t$ to (3.13), then we can obtain $(c_0/2)N^{s+t-k(n+m)}$ non-trivial solutions over $A^s \times (S \cap [N])^t$. Thus, we can remove from the conclusion of Theorem 3.4.3 the condition that the solutions we find are non-trivial.

We now seek to show that, without loss of generality, we may assume that conditions (vi) and (vii) of Theorem 3.4.5 hold. Let $K \in \mathbb{N}$ denote the absolute value of the product of all the non-zero entries of $\mathbf{A}$ (counting multiplicity). Let $\mathbf{B}' := K^{k^2}\mathbf{B}$, $\mathbf{C}' := K^{k^2}\mathbf{C}$, and

$$\mathbf{M}' := \begin{pmatrix} \mathbf{A} & \mathbf{B}' \\ \mathbf{0} & \mathbf{C}' \end{pmatrix}.$$

In other words, $\mathbf{M}'$ is the result of multiplying the last t columns of $\mathbf{M}$ by K^{k^2}. Observe that $\mathbf{A}, \mathbf{B}', \mathbf{C}'$, and $\mathbf{M}'$ all obey conditions (i)-(v) of Theorem 3.4.5. Let $M \in \mathbb{N}$ and let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set. Observe that the set $S' := \{x \in \mathbb{N} : K^k x \in S\}$ is also multiplicatively $[M]$-syndetic. Moreover, if $(\mathbf{x}, \mathbf{y}) \in [N]^s \times (S' \cap [N])^t$ satisfies both $\mathbf{A}\mathbf{x} = \mathbf{B}'\mathbf{y}$ and $\mathbf{C}'\mathbf{y} = \mathbf{0}$, then $(\mathbf{x}, K^k\mathbf{y}) \in [N]^s \times (S \cap [K^k N])^t$ is a solution to (3.13). Hence, by rescaling, we deduce that the conclusion of Theorem 3.4.3 holds for $\mathbf{A}, \mathbf{B}, \mathbf{C}$ if it holds for $\mathbf{A}, \mathbf{B}', \mathbf{C}'$. Furthermore, since every entry of $\mathbf{B}'$ is a multiple of K^k, for each $i \in [n]$ we can divide the ith row of $\mathbf{M}'$ by the greatest common divisor of the ith row of $\mathbf{A}$. The resulting matrix $\mathbf{M}''$ has integer entries (by choice of K), has the same solution set as $\mathbf{M}'$, and obeys conditions (i)-(vii) of Theorem 3.4.5. Hence, by replacing $\mathbf{M}$ with $\mathbf{M}''$, we may henceforth assume without loss of generality that $\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$ satisfy conditions (i)-(vii).

We now proceed by induction on $n + m$ to show that the conclusion of Theorem 3.4.3 holds for $\mathbf{A}, \mathbf{B}, \mathbf{C}$. Note that if $m = 0$ or $t = 0$, then condition (viii) holds vacuously, and so the result follows from Theorem 3.4.5 and Lemma 3.4.4 (by taking $A = S \cap [N]$). In particular, this proves the result for $n + m = 1$.

Now suppose that $n + m \geqslant 2$ and $m, t \geqslant 1$. Assume the induction hypothesis that the conclusion of Theorem 3.4.3 holds for any $\tilde{\mathbf{A}} \in \mathbb{Z}^{n' \times s'}, \tilde{\mathbf{B}} \in \mathbb{Z}^{n' \times t'}$, $\tilde{\mathbf{C}} \in \mathbb{Z}^{m' \times t'}$ with $1 \leqslant n' + m' < n + m$ which satisfy conditions (i)-(v) of Theorem 3.4.5. Consider the system of equations $\mathbf{C}\mathbf{y}^{\otimes k} = \mathbf{0}$. Since $m, t \geqslant 1$, condition (iii) implies that $\mathbf{C}$ obeys the columns condition. By considering linear combinations of the bottom m rows of $\mathbf{M}$, we also find that $\mathbf{C}$ obeys condition (iv) of Theorem 3.4.5. Thus, the induction hypothesis and Proposition 3.4.2 imply that condition (viii) holds. We therefore deduce from Theorem 3.4.5 and Lemma 3.4.4 that the conclusion of Theorem 3.4.3 holds for $\mathbf{A}, \mathbf{B}, \mathbf{C}$. □

Proof of Theorem 3.1.4 given Theorem 3.4.5. By performing elementary row operations and deleting linearly dependent rows, we may assume that $\mathbf{M}$ has rank n, each row of $\mathbf{M}$ has coprime entries, and that $\mathbf{M}$ contains a non-singular square diagonal matrix with the same number of rows as $\mathbf{M}$. By taking $\mathbf{B}$ and $\mathbf{C}$ to be empty matrices and putting $\mathbf{A} = \mathbf{M}$, the result follows immediately from Theorem 3.4.5. □

3.5 Linearisation and the W-trick

The purpose of this section is to obtain a lower bound for the number of solutions to (3.13) in terms of solutions to the 'linearised' system

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{B}\mathbf{y}^{\otimes k}; \\ \mathbf{0} &= \mathbf{C}\mathbf{y}^{\otimes k}. \end{aligned} \tag{3.14}$$

We accomplish this by using the linearisation procedure developed by Chow, Lindqvist, and Prendiville [CLP21, §12]. For $k \geqslant 3$, we avoid using smooth numbers and instead use the linearisation procedure detailed in Lindqvist's thesis [Lin19, §6.5]. This version follows the same outline as [CLP21, §12], but with the smoothness parameter η set to 1 (as in the quadratic case) and using the restriction estimates given in Lemma 3.2.3.

The upshot of applying these methods is that we are able to prove that Theorem 3.4.5 follows from the following linearised version.

Theorem 3.5.1 (Dense-syndetic regularity for linearised systems)**.** *Let $\delta > 0$, $k \in \mathbb{N} \setminus \{1\}$, and $M \in \mathbb{N}$. Let $n, s \in \mathbb{N}$, and $m, t \in \mathbb{Z}_{\geqslant 0}$. Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{M}$ be as defined in Theorem 3.4.5, and suppose that they satisfy conditions (i)-(viii). Then there exists a positive integer $N_1 = N_1(\delta, k, M, \mathbf{M}) \in \mathbb{N}$ and a positive constant $c_1 = c_1(\delta, k, M, \mathbf{M}) > 0$ such that the following is true. Let $N \in \mathbb{N}$, and suppose $A \subseteq [N]$ is such that $|A| \geqslant \delta N$. Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set. If $N \geqslant N_1$, then there are at least $c_1 N^{s+\frac{t}{k}-(n+m)}$ solutions $(\mathbf{x}, \mathbf{y}) \in A^s \times (S \cap [N^{1/k}])^t$ to (3.14).*

For the rest of this section we fix a choice of $k \in \mathbb{N} \setminus \{1\}$, and matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{M}$ satisfying conditions (i)-(viii) of Theorem 3.4.5. For each $r \in \mathbb{N}$ and for finitely supported functions $f_1, \ldots, f_s, g_1, \ldots, g_s : \mathbb{Z} \to \mathbb{C}$, we define the counting operators

$$\Lambda_r(f_1, \ldots, f_s; g_1, \ldots, g_s) := \sum_{\mathbf{C}\mathbf{y}^{\otimes k} = \mathbf{0}} \sum_{\mathbf{A}\mathbf{x}^{\otimes r} = \mathbf{B}\mathbf{y}^{\otimes k}} f_1(x_1) \cdots f_s(x_s) g_1(y_1) \cdots g_t(y_t).$$

For clarity, the outer sum is taken over all $\mathbf{y} \in \mathbb{Z}^t$ satisfying $\mathbf{C}\mathbf{y}^{\otimes k} = \mathbf{0}$, whilst the inner sum is over all $\mathbf{x} \in \mathbb{Z}^s$ such that $\mathbf{A}\mathbf{x}^{\otimes r} = \mathbf{B}\mathbf{y}^{\otimes k}$. In the case where

$\mathbf{C}$ is empty, the outer summation is omitted and Λ_r is defined by the inner sum only. If $\mathbf{C}$ and $\mathbf{B}$ are both empty, then the inner sum is taken over $\mathbf{A}\mathbf{x}^{\otimes r} = \mathbf{0}$. For brevity, we write $\Lambda_r(f_1, \ldots, f_s; g) := \Lambda_r(f_1, \ldots, f_s; g, \ldots, g)$, and similarly write $\Lambda_r(f; g) := \Lambda_r(f, \ldots, f; g)$. For sets A, B, we use the abbreviation $\Lambda_r(A; g_1, \ldots, g_t) := \Lambda_r(1_A; g_1, \ldots, g_t)$, and similarly for the quantities $\Lambda_r(f_1, \ldots, f_s; B)$ and $\Lambda_r(A; B)$.

In §3.3, we obtained a generalised von Neumann theorem (Theorem 3.3.6) for counting operators with weights satisfying p-restriction estimates. We now specialise this result to Λ_1.

Lemma 3.5.2 (Generalised von Neumann for Λ_1)**.** *Let $p := k^2 + \frac{1}{2(n+m)}$, and let $\eta := (2k^2(n+m)+2)^{-1}$. Let $N \in \mathbb{N}$, and suppose that $\nu, \mu : [N] \to \mathbb{R}_{\geqslant 0}$ each satisfy a p-restriction estimate with constant K. Let $f, g : [N] \to \mathbb{C}$, and let $D \subseteq [N^{1/k}]$. If $|f| \leqslant \nu$ and $|g| \leqslant \mu$, then*

$$\left|\Lambda_1\left(\|\nu\|_1^{-1} f; D\right) - \Lambda_1\left(\|\mu\|_1^{-1} g; D\right)\right| \ll_{K,\mathbf{M}} N^{\frac{t}{k}-(n+m)} \left\| \frac{\hat{f}}{\|\nu\|_1} - \frac{\hat{g}}{\|\mu\|_1} \right\|_\infty^\eta .$$

Proof. For each $\varphi : [N^{1/k}] \to \mathbb{C}$, let $Q_\varphi : \mathbb{Z} \to \mathbb{C}$ be the function given by

$$Q_\varphi(x) = \begin{cases} \varphi(y), & \text{if } x = y^k \text{ for some } y \in [N^{1/k}]; \\ 0, & \text{otherwise.} \end{cases}$$

If $B \subseteq [N^{1/k}]$, then we write $Q_B := Q_{1_B}$. Hence, if $h : \mathbb{Z} \to \mathbb{C}$ has finite support, then we may write

$$\Lambda_1(h; D) = \sum_{\mathbf{M}\mathbf{x}=\mathbf{0}} h(x_1) \cdots h(x_s) Q_D(x_{s+1}) \cdots Q_D(x_{s+t}).$$

Let $Q := Q_{[N^{1/k}]}$. Note that $\|Q_B\|_1 = |B|$ and $Q_B \leqslant Q$ for all $B \subseteq [N^{1/k}]$. Moreover, if $h : \mathbb{Z} \to \mathbb{C}$ satisfies $|h| \leqslant Q$, then $h = Q_H$, where $H : [N^{1/k}] \to \mathbb{C}$ is defined by $H(x) = h(x^k)$. Thus, by recalling from §3.2.5 that $\hat{Q}_\varphi = \mathfrak{F}_k[\varphi]$, Lemma 3.2.3 shows that Q satisfies a p-restriction estimate with constant $O_p(1)$ for all $p > k^2$. Now let $\nu_i = \nu$ and $\mu_i = \mu$ for all $i \in [s]$, and let $\nu_j = \mu_j = Q$ for all $s < j \leqslant s+t$. The result now follows from Theorem 3.3.6. □

3.5.1 The W-trick

Observe that one could attempt to linearise Theorem 3.4.5 by equating solutions $(\mathbf{x}, \mathbf{y}) \in A^s \times S^t$ to (3.13) with solutions $(\mathbf{x}^{\otimes k}, \mathbf{y}) \in B^s \times S^t$ to (3.14), where $B =$

$\{x^k : x \in A\}$. The most immediate problem with this approach is that the number of solutions sought are different; we need to find $\gg_{M,\mathbf{M}} N^{s+t-k(n+m)}$ solutions $(\mathbf{x}, \mathbf{y}) \in (A^s \times S^t) \cap [N]^{s+t}$ to (3.13), and $\gg_{M,\mathbf{M}} N^{s+\frac{t}{k}-(n+m)}$ solutions $(\mathbf{x}, \mathbf{y}) \in A^s \times (S \cap [N^{1/k}])^t$ to (3.14). This issue can be resolved by taking a weighted count of solutions to (3.14). We return to this topic in the next subsection.

The second problem that arises comes from the fact that, in general, the kth powers are not uniformly distributed modulo p for all primes p. This means that the Fourier coefficients of the indicator function of the kth powers in $[N]$ differ significantly from the Fourier coefficients of a weighted indicator function of $[N]$. This prevents us from making use of the Generalised von Neumann theorem to compare solutions of (3.13) with those of (3.14). The principle used to fix this problem is known as the *W-trick*. Originally introduced by Green [Gre05A] to solve equations in primes, a W-trick for squares was subsequently developed by Browning and Prendiville [BP17] and later generalised to squares and (smooth) higher powers by Chow, Lindqvist, and Prendiville [CLP21, §12].

We now describe the steps of the W-trick. Given $w \in \mathbb{N}$, define $W \in \mathbb{N}$ by

$$W = W(k, w) := k^{k-1} \prod_{p \leqslant w} p^k, \tag{3.15}$$

where the product is taken over all primes which do not exceed w. To transfer from a dense subset of $[N]$ to a subprogression, we require the following technical lemma.

Lemma 3.5.3 ([CLP21, Lemma A.4])**.** *Let $\delta > 0$, $w \in \mathbb{N}$, and let $W \in \mathbb{N}$ be defined by (3.15). Let $N \in \mathbb{N}$, and let $A \subseteq [N]$ be such that $|A| \geqslant \delta N$. There exist positive integers $\xi, \zeta \in \mathbb{N}$ (which depend on A) with $\xi \in [W]$ and $\zeta \ll_{\delta,w} 1$ which satisfy the following three properties.*

- *ξ and W are coprime;*
- *there does not exist a prime $p > w$ which divides ζ;*
- *$|\{x \in \mathbb{Z} : \zeta(\xi + Wx) \in A\}| \geqslant \frac{1}{2}\delta |\{x \in \mathbb{Z} : \zeta(\xi + Wx) \in [N]\}|$.*

Let $\delta > 0$ and $N, w \in \mathbb{N}$. Suppose that we are given sets $A \subseteq [N]$ and $S \subseteq \mathbb{N}$ such that $|A| \geqslant \delta N$ and S is multiplicatively $[M]$-syndetic. Let ζ and ξ be the positive integers provided by the above lemma. Define sets $A_1, S_1 \subseteq \mathbb{N}$ by

$$A_1 := \left\{ \frac{(Wz+\xi)^k - \xi^k}{kW} \in \mathbb{N} : \zeta(Wz+\xi) \in A \setminus \{\zeta\xi\} \right\};$$
$$S_1 := \{y \in \mathbb{N} : \zeta(kW)^{1/k} y \in S\}.$$

Observe that $A_1 \subseteq [X]$, where X is the positive rational number defined by

$$X = X(k, N, w, \zeta) := \frac{N^k}{kW\zeta^k}. \tag{3.16}$$

Lemma 3.5.3 implies that if $N \geqslant 2\zeta\xi$, then

$$|A_1| \geqslant \frac{\delta}{2}\left(\frac{N}{\zeta W} - \frac{\xi}{W}\right) \geqslant \frac{\delta N}{4\zeta W}.$$

Our aim is to count solutions to (3.13) over $A^s \times (S \cap [N])^t$ by counting solutions to (3.14) over $A_1^s \times (S_1 \cap [X^{1/k}])^t$. This leads to the following result.

Proposition 3.5.4. *Let $M, N, w \in \mathbb{N}$, and let $\delta > 0$. Let $A \subseteq [N]$ be such that $|A| \geqslant \delta N$. Let $W, \xi, \zeta \in \mathbb{N}$ be as given in Lemma 3.5.3, and let X be defined by (3.16). Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set. Let $A_1, S_1 \subseteq \mathbb{N}$ be defined as above. Then*

$$\Lambda_1\left(A_1; S_1 \cap [X^{1/k}]\right) \leqslant \Lambda_k\left(A; S \cap [N]\right).$$

Proof. Suppose that $\mathbf{x} \in A_1^s$ and $\mathbf{y} \in (S_1 \cap [X^{1/k}])^t$ are solutions to (3.14). We can therefore find $z_i \in \mathbb{N}$ for each $i \in [s]$ such that $kWx_i = (Wz_i + \xi)^k - \xi^k$. Let $u_i = \zeta(Wz_i+\xi)$ for each $i \in [s]$, and $v_j = \zeta(kW)^{1/k} y_j$ for each $j \in [t]$. Our construction of A_1 and S_1 shows that $\mathbf{u} \in A^s$ and $\mathbf{v} \in S^t$. Furthermore, we see from (3.16) that $\mathbf{v} \in [N]^t$.

Since $\mathbf{v}$ is a scalar multiple of $\mathbf{y}$, we find that $\mathbf{C}\mathbf{v}^{\otimes k} = \mathbf{0}$. Now let $i \in [n]$. Since $a_{i,1} + \cdots + a_{i,s} = 0$, we deduce that

$$\begin{aligned}\sum_{j=1}^{s} a_{i,j} u_j^k &= \sum_{j=1}^{s} a_{i,j}\zeta^k\left((Wz_j+\xi)^k - \xi^k\right) + (\zeta\xi)^k \sum_{j=1}^{s} a_{i,j} \\ &= k\zeta^k W \sum_{j=1}^{s} a_{i,j} x_j \\ &= \sum_{j=1}^{s} b_{i,j} v_j^k.\end{aligned}$$

We therefore conclude that $\mathbf{A}\mathbf{u}^{\otimes k} = \mathbf{B}\mathbf{v}^{\otimes k}$. Since the map $(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{u}, \mathbf{v})$ described above is injective, the desired result may now be obtained from a change of variables.

□

3.5.2 Weighted solutions

We now turn our attention to the problem of counting weighted solutions to (3.14). We begin by considering bounded weights. In this case, such sums can be handled by performing a minor modification to Theorem 3.5.1.

Lemma 3.5.5 (Functional Theorem 3.5.1)**.** *Let $\delta > 0$ and $M \in \mathbb{N}$. If Theorem 3.5.1 is true, then there exist constants $N_0(\delta, k, M, \mathbf{M}) \in \mathbb{N}$ and $c_0(\delta, k, M, \mathbf{M}) > 0$ such that the following is true. Let $f : [N] \to [0,1]$, and let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set. If $N \geqslant N_0(\delta, k, M, \mathbf{M})$ and $\|f\|_1 \geqslant \delta N$, then*

$$\Lambda_1\left(f; S \cap [N^{1/k}]\right) \geqslant c_0(\delta, k, M, \mathbf{M}) N^{s+\frac{t}{k}-(n+m)}.$$

Proof. This result follows from the same argument used to prove both [CLP21, Lemma 5.2] and [CLP21, Lemma 11.3]. □

Let $N, w \in \mathbb{N}$, and $\delta > 0$. Let $W, \xi, \zeta \in \mathbb{N}$ be as given in Proposition 3.5.4, and let X be given by (3.16). The weight function $\nu : [X] \to \mathbb{R}_{\geqslant 0}$ is defined by

$$\nu(n) := \begin{cases} x^{k-1}, & \text{if } n = \frac{x^k - \xi^k}{kW} \text{ for some } x \in [N/\zeta] \text{ with } x \equiv \xi \bmod W; \\ 0, & \text{otherwise.} \end{cases} \tag{3.17}$$

We now record some of the pseudorandomness properties possessed by the weight ν, as given in [Lin19, §6.5] (see also [CLP21, §6] and [CLP21, §12]).

Proposition 3.5.6. *Let $N, w \in \mathbb{N}$, and $\delta > 0$. Let $W, \xi, \zeta \in \mathbb{N}$ and $A_1 \subseteq \mathbb{N}$ be as given in Proposition 3.5.4. Let $X \in \mathbb{R}_{\geqslant 0}$ and $\nu : [X] \to \mathbb{R}_{\geqslant 0}$ be defined by (3.16) and (3.17) respectively. If N is sufficiently large with respect to w and δ, then the following properties all hold.*

- *(Density transfer).*
$$\sum_{n \in A_1} \nu(n) \gg_k \delta^k \|\nu\|_{L^1(\mathbb{Z})};$$
- *(Fourier decay).*
$$\left\|\hat{\nu} - \hat{1}_{[X]}\right\|_\infty \ll_k X w^{-1/k};$$
- *(Restriction estimate). For any $f : \mathbb{Z} \to \mathbb{C}$ and $p \in \mathbb{R}$ such that $p > k^2$ and $|f| \leqslant \nu$, we have*
$$\|\hat{f}\|_p^p = \int_{\mathbb{T}} \left|\hat{f}(\alpha)\right|^p d\alpha \ll_p X^{p-1}.$$

Proof. The above properties are given by the results [Lin19, Lemma 6.5.2], [Lin19, Lemma 6.5.3], and [Lin19, Lemma 6.5.4] respectively. □

These properties allow us to approximate functions majorised by ν with functions majorised by $1_{[X]}$. This enables us to transfer solutions from the linearised setting (3.14) to the squares (3.13). To achieve this, we require the following technical lemma.

Lemma 3.5.7 (Dense model lemma). *Let $N, w \in \mathbb{N}\backslash\{1\}$, and $\delta > 0$. Let $X \in \mathbb{R}_{\geqslant 0}$ and $\nu : [X] \to \mathbb{R}_{\geqslant 0}$ be as given in Proposition 3.5.6. Then for any function $f : [X] \to \mathbb{R}_{\geqslant 0}$ with $f \leqslant \nu$, there exists a function $g : [X] \to \mathbb{R}_{\geqslant 0}$ with $\|g\|_\infty \leqslant 1$ such that*

$$\|\hat{f} - \hat{g}\|_\infty \ll_k X(\log w)^{-3/2}.$$

Proof. Using the Fourier decay estimate given in Proposition 3.5.6, the result follows by applying [Pre17, Theorem 5.1] to ν. □

This lemma allows us to consider unbounded weights f for our counting operators by replacing them with bounded weights g and applying Lemma 3.5.5. Following the strategy used in the proof of [CLP21, Theorem 5.5], we can now show that Theorem 3.5.1 implies Theorem 3.4.5.

Proof of Theorem 3.4.5 given Theorem 3.5.1. Given $\delta > 0$ and $M \in \mathbb{N}$, we choose $w = w(\delta, k, M, \mathbf{M}) \in \mathbb{N}$ to be sufficiently large. By fixing this choice of w, the assumption $N \gg_{\delta,k,M,\mathbf{M}} 1$ allows us to ensure that N and X are sufficiently large relative to $\delta, M, \mathbf{M}$ and w.

Proposition 3.5.4 implies that

$$\|\nu\|_\infty^s \Lambda_k(A; S \cap [N]) \geqslant \Lambda_1(\nu 1_{A_1}; S_1 \cap [X^{1/k}]). \tag{3.18}$$

Recall from (3.16) and (3.17) respectively that $X \gg_{\delta,k,M} N^k$ and $\|\nu\|_\infty \leqslant N^{k-1}$. We may therefore deduce Theorem 3.4.5 from (3.18) if we can prove that

$$\Lambda_1(\nu 1_{A_1}; S_1 \cap [X^{1/k}]) \gg_{\delta,k,M,\mathbf{M}} X^{s+\frac{t}{k}-(n+m)}. \tag{3.19}$$

Let $f := \nu 1_{A_1}$. By choosing N sufficiently large with respect to w and δ, the density transfer estimate in Proposition 3.5.6 implies that $\|f\|_1 \gg \delta^k \|\nu\|_1$. Lemma 3.5.7 provides us with a function $g : [X] \to [0, 1]$ such that

$$\|\hat{f} - \hat{g}\|_\infty \ll_k X(\log w)^{-3/2}.$$

Note that the Fourier decay estimate given in Proposition 3.5.6 implies that

$$\left| \|\nu\|_1 - X \right| \leqslant \left\| \hat{\nu} - \hat{1}_{[X]} \right\|_\infty \ll_k X w^{-1/k}. \tag{3.20}$$

Thus, if w is sufficiently large, then

$$\left\| \frac{\hat{f}}{\|\nu\|_1} - \frac{\hat{g}}{X} \right\|_\infty \ll_k (\log w)^{-3/2}.$$

By considering the Fourier coefficients at 0, if w is sufficiently large, then the above inequality implies that $\|g\|_1 \gg \delta^k X$. Hence, from Lemma 3.5.5 it follows that

$$\Lambda_1(g; S \cap [X^{1/k}]) \gg_{\delta,k,M,\mathbf{M}} X^{s+\frac{t}{k}-(n+m)}. \tag{3.21}$$

Taking $p = k^2 + \frac{1}{2(n+m)}$ and $D = S \cap [X^{1/k}]$, Lemma 3.5.2 gives the bound

$$\left| \Lambda_1(\|\nu\|_1^{-1} f; D) - \Lambda_1(X^{-1} g; D) \right| \ll_{\delta,k,M,\mathbf{M}} X^{\frac{t}{k}-(n+m)} (\log w)^{-\eta},$$

where $\eta = \frac{3}{4}(k^2(n+m)+1)^{-1} \in (0,1)$. We may therefore deduce (3.19) from (3.20) and (3.21) upon choosing w to be sufficiently large relative to the implicit constants appearing in these two inequalities and the above. $\square$

It should be emphasised that the above proof shows that if the conclusion of Theorem 3.5.1 holds for a given matrix $\mathbf{M}$, then the conclusion of Theorem 3.4.5 holds for the same matrix $\mathbf{M}$. In particular, if $\mathbf{B}$ and $\mathbf{C}$ are empty ($m = t = 0$), then the above proof provides us with a means to establish density regularity for $\mathbf{A}\mathbf{x}^{\otimes k} = \mathbf{0}$. This method gives an alternative proof of Theorem 3.1.4 which does not require us to first prove Theorem 3.5.1 in full generality. All that we require is the following theorem of Frankl, Graham, and Rödl [FGR88].

Theorem 3.5.8 ([FGR88, Theorem 2]). *Let $\mathbf{M} \in \mathbb{Z}^{n\times s}$ be an integer matrix whose columns sum to $\mathbf{0}$. If there exists at least one non-trivial solution $\mathbf{y} \in \mathbb{N}^s$ to the system $\mathbf{M}\mathbf{y} = \mathbf{0}$, then there exist constants $N_1 = N_1(\delta, \mathbf{M}) \in \mathbb{N}$ and $c_1 = c_1(\delta, \mathbf{M}) > 0$ such that the following is true. If $N \geqslant N_1$, then for any $A \subseteq [N]$ such that $|A| \geqslant \delta N$, there are at least $c_0 N^{s-n}$ non-trivial solutions $\mathbf{x} \in A^s$ to the system of equations $\mathbf{M}\mathbf{x} = \mathbf{0}$.*

Proof of Theorem 3.1.4. As shown at the end of §3.4, Theorem 3.1.4 follows from the case $m = t = 0$ of Theorem 3.4.5. By the argument above that Theorem 3.5.1 implies Theorem 3.4.5, it only remains to prove Theorem 3.5.1 in the case where $m = t = 0$. This result follows immediately from Theorem 3.5.8. $\square$

3.6 Arithmetic regularity

The objective of this final section is to use the arithmetic regularity lemma to prove Theorem 3.5.1. For the rest of this section, we fix a choice of matrices $\mathbf{A} \in \mathbb{Z}^{n\times s}$, $\mathbf{B} \in \mathbb{Z}^{n\times t}$, $\mathbf{C} \in \mathbb{Z}^{m\times t}$, $\mathbf{M} \in \mathbb{Z}^{(n+m)\times(s+t)}$ as given in Theorem 3.4.5, and assume that they satisfy conditions (i)-(viii). In particular, by permuting columns, we may assume that the first n columns of $\mathbf{A}$ form a non-singular diagonal $n \times n$ matrix.

Our general strategy is similar to the method used to prove Roth's theorem in [Tao12, §1.2]. The regularity lemma enables us to decompose the indicator function of our dense set A into more manageable functions. To describe these functions, we require the following definition.

Definition (Lipschitz function)**.** A function $F : \mathbb{T}^d \to [0,1]$ is called an *L-Lipschitz function* if, for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{T}^d$, we have

$$|F(\boldsymbol{\alpha}) - F(\boldsymbol{\beta})| \leqslant L\|\boldsymbol{\alpha} - \boldsymbol{\beta}\|.$$

For our work, we only require the 'abelian' arithmetic regularity lemma. This result, originally introduced by Green [Gre05B], is a special case of the general arithmetic regularity lemma (see for instance [GT10]). The following version is a combination of [Ebe16, Theorem 5] and [GL19, Proposition 4.2] (see also [Tao12, Theorem 1.2.11]).

Lemma 3.6.1 (Abelian arithmetic regularity lemma)**.** *Let $\mathcal{F} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ be a monotone increasing function, and let $\varepsilon > 0$. Then there exists a positive integer $L_0(\varepsilon, \mathcal{F}) \in \mathbb{N}$ such that the following is true. If $f : [N] \to [0,1]$ for some $N \in \mathbb{N}$, then there is a positive integer $L \leqslant L_0(\varepsilon, \mathcal{F})$ and a decomposition*

$$f = f_{\mathrm{str}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$$

of f into functions $f_{\mathrm{str}}, f_{\mathrm{sml}}, f_{\mathrm{unf}} : [N] \to [-1,1]$ such that:

(I) *the functions f_{str} and $f_{\mathrm{str}} + f_{\mathrm{sml}}$ take values in $[0,1]$;*

(II) *the function f_{sml} obeys the bound $\|f_{\mathrm{sml}}\|_{L^2(\mathbb{Z})} \leqslant \varepsilon \|1_{[N]}\|_{L^2(\mathbb{Z})}$;*

(III) *the function f_{unf} obeys the bound $\|\hat{f}_{\mathrm{unf}}\|_\infty \leqslant \|\hat{1}_{[N]}\|_\infty / \mathcal{F}(L)$;*

(IV) *there exists a positive integer $d \leqslant L$, a phase $\boldsymbol{\theta} \in \mathbb{T}^d$, and an L-Lipschitz function $F : \mathbb{T}^d \to [0,1]$ such that $F(x\boldsymbol{\theta}) = f_{\mathrm{str}}(x)$ for all $x \in [N]$.*

We now employ the arithmetic regularity lemma to count solutions to (3.14). Recall from §3.5 the counting operator

$$\Lambda_1(f_1,\ldots,f_s;g_1,\ldots,g_t) := \sum_{\mathbf{C}\mathbf{y}^{\otimes k}=\mathbf{0}} \sum_{\mathbf{A}\mathbf{x}=\mathbf{B}\mathbf{y}^{\otimes k}} f_1(x_1)\cdots f_s(x_s)g_1(y_1)\cdots g_t(y_t).$$

We are interested in the quantity $\Lambda_1(A; S\cap[N^{1/k}])$, where $A\subseteq[N]$ satisfies $|A|\geqslant\delta N$, and $S\subseteq\mathbb{N}$ is multiplicatively $[M]$-syndetic. Applying the arithmetic regularity lemma with $f = 1_A$ (for a choice of parameters $\varepsilon,\mathcal{F}$ to be specified later), we obtain a decomposition $1_A = f_{\mathrm{str}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$.

The first step towards the proof of Theorem 3.5.1 involves removing the uniform part f_{unf} via the Generalised von Neumann theorem (Theorem 3.3.6).

Lemma 3.6.2 (Removing f_{unf}). *Let $N\in\mathbb{N}$, and let $A\subseteq[N]$. Let $\varepsilon>0$, and let $\mathcal{F}:\mathbb{R}_{\geqslant 0}\to\mathbb{R}_{\geqslant 0}$ be a monotone increasing function. Let $f_{\mathrm{str}}, f_{\mathrm{sml}}, f_{\mathrm{unf}}$ be the functions provided by applying Lemma 3.6.1 to $f = 1_A$. Then for any $D\subseteq[N^{1/k}]$, we have*

$$|\Lambda_1(A;D) - \Lambda_1(f_{\mathrm{str}}+f_{\mathrm{sml}};D)| \ll_{\mathbf{M}} N^{s+\frac{t}{k}-(n+m)}\mathcal{F}(L)^{-(2k^2(n+m)+2)^{-1}}.$$

Proof. Let $\nu=\mu=1_{[N]}$, and note that $\|1_{[N]}\|_1 = N$. Furthermore, property (I) of Lemma 3.6.1 implies that $0\leqslant(f_{\mathrm{str}}+f_{\mathrm{sml}}), 1_A\leqslant 1_{[N]}$. Thus, the result follows from Lemma 3.5.2. □

3.6.1 Auxiliary counting operators

Lemma 3.6.2 shows that the main contribution to $\Lambda_1(A;D)$ comes from $\Lambda_1(f_{\mathrm{str}} + f_{\mathrm{sml}};D)$. We therefore focus our attention on finding lower bounds for this latter quantity. Rather than count all solutions to (3.14), it is convenient for us to restrict our attention to an explicit dense subcollection of solutions. We then show that the counting operators associated with these subcollections obey a generalised von Neumann theorem with respect to the L^2 norm. This allows us to remove the f_{sml} function from our analysis.

To make this precise, let $q := s-n-1$, and let $\{\mathbf{u}^{(0)},\ldots,\mathbf{u}^{(q)}\}\subseteq\mathbb{Z}^s$ denote a $\mathbb{Q}$-basis for $\ker(\mathbf{A})$. Note that condition (ii) of Theorem 3.4.5 implies that $q\geqslant 0$. Let $u_{i,j}$ denote the jth entry of $\mathbf{u}^{(i)}$. Since the columns of $\mathbf{A}$ sum to zero, we may take $\mathbf{u}^{(0)} := (1,1,\ldots,1)$ and assume that $u_{i,1}=0$ for all $i\in[q]$. Such a basis is not uniquely defined, however we can insist that $|u_{i,j}|\ll_{\mathbf{M}} 1$ for all i and j.

Observe that the each row of $\mathbf{B}\mathbf{y}^{\otimes k}$ defines an integer homogeneous diagonal polynomial of degree k in the variables $y_1, \ldots, y_t$. Thus, by condition (vii), we may define for each $i \in [n]$ an integer polynomial $P_i \in \mathbb{Z}[y_1, \ldots, y_t]$ by dividing the ith row of $\mathbf{B}\mathbf{y}^{\otimes k}$ by the ith entry of the ith row of $\mathbf{A}$. We also set $P_j = 0$ for all $n < j \leqslant s$. Now note that every solution to $\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{y}^{\otimes k}$ takes the form $\mathbf{x} = (z_1 + P_1(\mathbf{y}), \ldots, z_s + P_s(\mathbf{y}))$, where $(z_1, \ldots, z_s) \in \ker(\mathbf{A})$.

We are now ready to define our auxiliary counting operators. Let $\mathbf{y} \in \mathbb{Z}^t$, and let $B \subseteq \mathbb{Z}$ be a finite set. For each $\mathbf{d} \in B^q$ and $j \in [s]$, define

$$Q_j(\mathbf{d}, \mathbf{y}) := u_{1,j}d_1 + \cdots + u_{q,j}d_q + P_j(\mathbf{y}). \tag{3.22}$$

Given functions $f_1, \ldots, f_s : \mathbb{Z} \to \mathbb{C}$ with finite support, we define

$$\Psi_{B,\mathbf{y}}(f_1, \ldots, f_s) := \sum_{x \in \mathbb{Z}} \sum_{\mathbf{d} \in B^q} \prod_{j=1}^{s} f_j(x + Q_j(\mathbf{d}, \mathbf{y})). \tag{3.23}$$

For brevity, we write $\Psi_{B,\mathbf{y}}(f) := \Psi_{B,\mathbf{y}}(f, \ldots, f)$. Hence, since $\{\mathbf{u}^{(0)}, \ldots, \mathbf{u}^{(q)}\}$ is a $\mathbb{Q}$-basis for $\ker(\mathbf{A})$, if $f_1, \ldots, f_s : [N] \to [0, 1]$, then

$$\Lambda_1(f_1, \ldots, f_s; S) \geqslant \sum_{\mathbf{C}\mathbf{y}^{\otimes k} = \mathbf{0}} \Psi_{B,\mathbf{y}}(f_1, \ldots, f_s) 1_S(y_1) \cdots 1_S(y_t). \tag{3.24}$$

We may therefore use Ψ to obtain a lower bound for $\Lambda_1(f_{\text{str}} + f_{\text{sml}}; D)$. In fact, by exploiting the fact that $\|f_{\text{sml}}\|_2$ is 'small', we show that it is sufficient to obtain a lower bound for $\Lambda_1(f_{\text{str}}; D)$. To proceed in this way, we utilise the following L^2 generalised von Neumann theorem for Ψ.

Lemma 3.6.3 (Generalised von Neumann for Ψ)**.** *Let $N \in \mathbb{N}$, $\mathbf{y} \in \mathbb{Z}^t$, and let $B \subseteq \mathbb{Z}$ be a finite set. Let Ψ be defined by (3.23). If $f, g : [N] \to [0, 1]$, then*

$$|\Psi_{B,\mathbf{y}}(f) - \Psi_{B,\mathbf{y}}(g)| \leqslant s|B|^{s-n-1} N^{1/2} \|f - g\|_2.$$

Proof. Let $q := s - n - 1$. We show that

$$|\Psi_{B,\mathbf{y}}(f_1, \ldots, f_s)| \leqslant |B|^q N^{1/2} \|f_i\|_2 \tag{3.25}$$

holds for all $i \in [s]$ and all $f_1, \ldots, f_s : [N] \to [-1, 1]$. The lemma then follows by applying this bound to the telescoping identity

$$\Psi_{B,\mathbf{y}}(f) - \Psi_{B,\mathbf{y}}(g) = \sum_{r=1}^{s} \Psi_{B,\mathbf{y}}(g_1, \ldots, g_{r-1}, f_r - g_r, f_{r+1}, \ldots, f_s),$$

where $f_j = f$ and $g_j = g$ for all $j \in [s]$.

Let $i \in [s]$. Since the function f_i is supported on $[N]$, the change of variables $z = x + Q_i(\mathbf{d}, \mathbf{y})$ yields

$$\Psi_{B,\mathbf{y}}(f_1, \ldots, f_s) = \sum_{z\in[N]} \sum_{\mathbf{d}\in B^q} \prod_{j=1}^{s} f_j(z + Q_j(\mathbf{d}, \mathbf{y}) - Q_i(\mathbf{d}, \mathbf{y})).$$

Applying the Cauchy-Schwarz inequality with respect to z gives

$$|\Psi_{B,\mathbf{y}}(f_1, \ldots, f_s)|^2 \leqslant \|f_i\|_2^2 \sum_{z\in[N]} \left| \sum_{\mathbf{d}\in B^q} \prod_{\substack{j=1 \\ j\neq i}}^{n} f_j(z + Q_j(\mathbf{d}, \mathbf{y}) - Q_i(\mathbf{d}, \mathbf{y})) \right|^2 .$$

We may therefore obtain (3.25) from the bound $\|f_j\|_\infty \leqslant 1$. □

Lemma 3.6.4 (Removing f_{sml})**.** *Let the assumptions and definitions be as in Lemma 3.6.2. Let $\mathbf{y} \in \mathbb{Z}^t$, and let $B \subseteq \mathbb{Z}$ be a finite set. Then we have*

$$\Psi_{B,\mathbf{y}}(f_{\mathrm{str}} + f_{\mathrm{sml}}) = \Psi_{B,\mathbf{y}}(f_{\mathrm{str}}) - O_{\mathbf{M}}(|B|^{s-n-1} N\varepsilon).$$

Proof. Recall from Lemma 3.6.1 that f_{str} and $f_{\mathrm{str}} + f_{\mathrm{sml}}$ take values in $[0, 1]$. Thus, the result follows immediately from Lemma 3.6.3. □

3.6.2 Bohr sets

We wish to better understand the behaviour of the structured function f_{str} appearing in Lemma 3.6.1. From the definition of Lipschitz functions, we see that $f_{\mathrm{str}}(x) \approx f_{\mathrm{str}}(x + y)$ holds whenever $\|y\boldsymbol{\theta}\|$ is 'small'. Such y are sometimes referred to as *almost periods* for f_{str}, see [Tao12, Lemma 1.2.13]. This leads us to consider the properties of sets of such y, which are known as *(polynomial) Bohr sets.*

Definition (Polynomial Bohr sets)**.** Let $d, h \in \mathbb{N}$, $\rho > 0$, and let $\boldsymbol{\alpha} \in \mathbb{T}^d$. The (*polynomial*) *Bohr set* $\mathrm{B}_h(\boldsymbol{\alpha}, \rho)$ is the set

$$\mathrm{B}_h(\boldsymbol{\alpha}, \rho) := \bigcap_{i=1}^{d} \{n \in \mathbb{N} : \|n^h \alpha_i\| < \rho\}.$$

A key property of Bohr sets is that they have positive density. Furthermore, the density of $\mathrm{B}_h(\boldsymbol{\alpha}, \rho)$ on $[N]$ (for N suitably large) can be bounded from below by a positive quantity which depends only on d, h and ρ. A crucial aspect of this result is that this uniform lower bound does not depend on $\boldsymbol{\alpha}$.

Using Lemma 3.4.4, we can deduce such a result by first showing that Bohr sets are multiplicatively syndetic. Furthermore, we prove that the intersection of a multiplicatively syndetic set with any finite intersection of non-empty Bohr sets is multiplicatively syndetic. This fact may be of independent interest.

The key tool needed to establish these facts is the following polynomial recurrence result recorded in [Sch77].

Lemma 3.6.5 (Polynomial recurrence)**.** *If $h \in \mathbb{N}$, then there exists a constant $C = C(h) > 0$ such that the following holds. If $\alpha \in \mathbb{R}$ and $N \in \mathbb{N}$, then*

$$\min_{1 \leqslant n \leqslant N} \|n^h \alpha\| \leqslant C N^{-2^{-h}}.$$

Proof. See [Sch77, Theorem 7A]. □

Lemma 3.6.6 (Syndeticity of Bohr sets)**.** *Let $h \in \mathbb{N}$ and $0 < \rho \leqslant 1$. Then there exists a constant $M_0 = M_0(\rho, h) \in \mathbb{N}$ such that the following is true. If $\alpha \in \mathbb{T}$, then the Bohr set $\mathrm{B}_h(\alpha, \rho)$ is multiplicatively M_0-syndetic.*

Proof. Lemma 3.6.5 immediately implies that there exists some $M \ll_h \rho^{-2^h}$ such that $[M]$ intersects $\mathrm{B}_h(\beta, \rho)$ for each $\beta \in \mathbb{R}$. Hence, by replacing β with $m^h\alpha$, we deduce that $\mathrm{B}_h(\alpha, \rho)$ intersects $m \cdot [M]$ for every $m \in \mathbb{N}$. We therefore find that $\mathrm{B}_h(\alpha, \rho)$ is multiplicatively $[M]$-syndetic. □

Corollary 3.6.7 (Syndeticity for Bohr-syndetic intersections)**.** *Let $d, h, M \in \mathbb{N}$, and $0 < \rho \leqslant 1$. There exists a constant $M_1 = M_1(\rho, d, h, M) \in \mathbb{N}$ such that the following is true. Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set. If $\boldsymbol{\alpha} \in \mathbb{T}^d$, then the set $S \cap \mathrm{B}_h(\boldsymbol{\alpha}, \rho)$ is multiplicatively $[M_1]$-syndetic. Moreover, we may assume that, when considered as a function of ρ, d, M, the quantity M_1 satisfies*

$$M_1(\rho, d, h, M) = \max\{M_1(\rho', d', h, M') : \rho' \in [\rho, 1], d' \in [d], M' \in [M]\}. \tag{3.26}$$

Proof. Observe that every multiplicatively $[M]$-syndetic set is multiplicatively $[M+1]$-syndetic, and $\mathrm{B}_h((\alpha_1, \ldots, \alpha_{d'}), \rho) \subseteq \mathrm{B}_h((\alpha_1, \ldots, \alpha_d), \rho')$ for all $\rho' \geqslant \rho$ and $d' \in [d]$. These observations show that we can guarantee (3.26) holds once the rest of the result is proven.

By writing $S \cap \mathrm{B}_h(\boldsymbol{\alpha}, \rho) = \mathrm{B}_h(\alpha_1, \rho) \cap (S \cap \mathrm{B}_h((\alpha_2, \ldots, \alpha_d), \rho))$, we see that the result follows by induction from the case $d = 1$. Thus, it is sufficient to show that there exists some $M_1 = M_1(\rho, h, M) \in \mathbb{N}$ such that $S \cap \mathrm{B}_h(\alpha, \rho)$ is multiplicatively $[M_1]$-syndetic for all $\alpha \in \mathbb{T}$.

Let $\alpha \in \mathbb{T}$, $a \in \mathbb{N}$, and let $\rho' = \rho/M$. Lemma (3.6.6) provides us with some $M' = M'(\rho, h, M) \in \mathbb{N}$ such that $\mathrm{B}_h(\alpha, \rho')$ is multiplicatively $[M']$-syndetic. We can therefore choose some $m \in [M']$ such that $am \in \mathrm{B}_h(\alpha, \rho')$. Moreover, we have $am \cdot [M] \subseteq \mathrm{B}_h(\alpha, \rho)$. Thus, $am \cdot [M]$ intersects $S \cap \mathrm{B}_h(\alpha, \rho)$. We therefore conclude that $S \cap \mathrm{B}_h(\alpha, \rho)$ is multiplicatively $[M \cdot M']$-syndetic. $\square$

Let $\boldsymbol{\theta} \in \mathbb{T}^d$ be as given in Lemma 3.6.1, and let $0 < \rho < 1$. Observe that if $x, r \in [N]$ with $r \in \mathrm{B}_1(\boldsymbol{\theta}, \rho)$ and $x + r \in [N]$, then

$$|f_{\mathrm{str}}(x+r) - f_{\mathrm{str}}(x)| = |F(\boldsymbol{\theta}(x+r)) - F(\boldsymbol{\theta}x)| \leqslant dL\rho \leqslant L^2\rho.$$

Similarly, if $\mathbf{y} \in \mathrm{B}_k(\boldsymbol{\theta}, \rho)^t$ and $j \in [s]$ are such that $x, x + P_j(\mathbf{y}) \in [N]$, then

$$|f_{\mathrm{str}}(x + P_j(\mathbf{y})) - f_{\mathrm{str}}(x)| \ll_{\mathbf{M}} L^2\rho.$$

It is important to note that we must assume $x, x+r, x+P_j(\mathbf{y}) \in [N]$ for the above inequalities to hold. This is because we are using the convention that $f_{\mathrm{str}}(x) = 0$ for all $x \notin [N]$. To circumvent this issue, we use a trick of Tao [Tao12, Lemma 1.2.13]. Rather then extend f_{str} outside of $[N]$ by 0, we instead use the function F and replace $f_{\mathrm{str}}(x)$ by $F(\boldsymbol{\theta}x)$. Since $f_{\mathrm{str}}(x) = F(\boldsymbol{\theta}x)$ only holds for $x \in [N]$, we restrict our choice of $\mathbf{y}$ and $\mathbf{d}$ so that $|Q_j(\mathbf{d}, \mathbf{y})| \leqslant \rho N$ holds for all $j \in [s]$. Proceeding in this way produces the following result.

Lemma 3.6.8 (Lower bound for $\Psi_{B,\mathbf{y}}(f_{\mathrm{str}})$)**.** *Let the assumptions and definitions be as in Lemma 3.6.2. Let $q := s - n - 1$, and let $\mathcal{U} = \{\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(q)}\}$ denote a $\mathbb{Q}$-basis for* $\ker(\mathbf{A})$ *with the properties described in §3.6.1. There exists positive constants $\rho_0 = \rho_0(k, \mathbf{M}, \mathcal{U}) \in (0,1)$ and $N_0 = N_0(k, \mathbf{M}, \mathcal{U}) \in \mathbb{N}$ such that the following is true. Let $\rho \in (0,1)$ and $N \in \mathbb{N}$. Let $B = \mathrm{B}_1(\boldsymbol{\theta}, \rho) \cap [\rho N]$, and let $\mathbf{y} \in (\mathrm{B}_k(\boldsymbol{\theta}, \rho) \cap [(\rho N)^{1/k}])^q$. Let $Q_1, \dots, Q_s$ be given by (3.22), and let $\Psi_{B,\mathbf{y}}$ be as defined in (3.23). If $N \geqslant N_0$ and $0 < \rho \leqslant \rho_0$, then*

$$\Psi_{B,\mathbf{y}}(f_{\mathrm{str}}) \geqslant |B|^q N \left[(\delta - \varepsilon - \mathcal{F}(L)^{-1})^s - O_{\mathbf{M},\mathcal{U}}(L^{2s}\rho^s + L^2\rho) \right].$$

Proof. Observe that our choice of $\mathbf{y}$ and B implies that $|Q_i(\mathbf{d}, \mathbf{y})| \ll_{\mathbf{M},\mathcal{U}} \rho N$ holds for all $i \in [s]$ and $\mathbf{d} \in B^q$. Hence, provided N and ρ^{-1} are sufficiently large in terms of

$k, \mathbf{M}$, and $\mathcal{U}$, there exists some $\rho' > 0$ with $\rho' \ll_{k,\mathbf{M},\mathcal{U}} \rho$ such that the following is true. Let Ω denote the set of $x \in \mathbb{N}$ such that $\rho' N < x < (1-\rho')N$. If $x \in \Omega$, then

$$|f_{\mathrm{str}}(x + Q_i(\mathbf{d},\mathbf{y})) - f_{\mathrm{str}}(x)| = |F((x + Q_i(\mathbf{d},\mathbf{y}))\boldsymbol{\theta}) - F(x\boldsymbol{\theta})| \ll_{\mathbf{M},\mathcal{U}} L^2\rho$$

holds for all $i \in [s]$ and $\mathbf{d} \in B^q$. Since $|[N] \setminus \Omega| \ll_{k,\mathbf{M},\mathcal{U}} \rho N$, we can take ρ sufficiently small to ensure that $\Omega \neq \emptyset$. Thus, the inequality $0 \leqslant f_{\mathrm{str}} \leqslant 1_{[N]}$ implies that

$$\begin{aligned}
\Psi_{B,\mathbf{y}}(f_{\mathrm{str}}) &\geqslant \sum_{x\in\Omega}\sum_{\mathbf{d}\in B^q}\prod_{j=1}^{s} f_{\mathrm{str}}(x + Q_j(\mathbf{d},\mathbf{y})) \\
&\geqslant \sum_{\mathbf{d}\in B^q}\left(\sum_{x=1}^{N}\left(f_{\mathrm{str}}(x)^s - O_{k,\mathbf{M},\mathcal{U}}(L^{2s}\rho^s + L^2\rho)\right) - |[N]\setminus\Omega|\right) \\
&\geqslant |B|^q N\left(N^{-1}\|(f_{\mathrm{str}})^s\|_1 - O_{k,\mathbf{M},\mathcal{U}}(L^{2s}\rho^s + L^2\rho)\right). \qquad (3.27)
\end{aligned}$$

By an application of Hölder's inequality, we find that

$$\|(f_{\mathrm{str}})^s\|_1 \geqslant N^{1-s}\left(\sum_{x=1}^{N}(1_A(x) - f_{\mathrm{sml}}(x) - f_{\mathrm{unf}}(x))\right)^s \geqslant N(\delta - \varepsilon - \mathcal{F}(L)^{-1})^s.$$

Substituting the above into (3.27) completes the proof. □

This final lemma, in combination with the previous results of this section, finally provides us with a means to prove Theorem 3.5.1.

Proof of Theorem 3.5.1. Condition (viii) implies that if $\mathbf{C}$ is non-empty, then there exist functions $\mathcal{M} : \mathbb{N} \to \mathbb{N}$ and $\kappa : \mathbb{N} \to (0,1]$ (which depend on k and $\mathbf{M}$) such that the following is true. Let $M \in \mathbb{N}$, and let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set. If $N \geqslant \mathcal{M}(M)$, then there are at least $\kappa(M)N^{t-km}$ solutions $\mathbf{y} \in (S \cap [N])^t$ to the system $\mathbf{C}\mathbf{y}^{\otimes k} = \mathbf{0}$. Moreover, this condition shows that we may assume that $\mathcal{M}$ is a monotone increasing function satisfying $\mathcal{M}(M) \geqslant M$ for all $M \in \mathbb{N}$, and that κ is monotone decreasing. In the case where $\mathbf{C}$ is empty, we take $\kappa = 1$ and $\mathcal{M}(M) = M$ for all $M \in \mathbb{N}$.

Let $\delta \in (0,1)$ and $M \in \mathbb{N}$ be fixed. Let $c = c(k,\mathbf{M}) > 1$ be a sufficiently small positive constant depending only on k and $\mathbf{M}$. Let $M_1 = M_1(\rho, d, h, M)$ be the quantity given by Corollary 3.6.7, which satisfies (3.26). Let $\varepsilon := c\delta^s$, and let $\mathcal{F} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ be a sufficiently fast growing monotone increasing function (specifically, $\mathcal{F}$ is chosen so that (3.33) holds). For this choice of $\mathcal{F}$ and ε, let $L_0 = L_0(\varepsilon, \mathcal{F})$ be the integer given by Lemma 3.6.1.

Let $N \in \mathbb{N}$ be sufficiently large in terms of all the previously defined parameters (specifically, N is chosen to satisfy (3.29)). Let $A \subseteq [N]$ with $|A| \geqslant \delta N$. Let $f_{\rm str}, f_{\rm sml}, f_{\rm unf}$ be the functions obtained by applying Lemma 3.6.1 with respect to $f = 1_A$. Let $d, L, \boldsymbol{\theta}$ and F be as given in property (IV) of Lemma 3.6.1. Let $\mathcal{U} = \{\mathbf{u}^{(0)}, \ldots, \mathbf{u}^{(s-n-1)}\}$ denote a $\mathbb{Q}$-basis for $\ker(\mathbf{A})$ with the properties described in §3.6.1. Note that, by explicit computation, we can choose such a basis $\mathcal{U}$ such that $|u_{i,j}| \ll_{\mathbf{M}} 1$ for all i and j. Consequently, any emergent quantities which may depend on $\mathcal{U}$ are instead considered to depend on $\mathbf{M}$.

Let $\rho := c\delta^s L^{-2}$, and let $B := \mathrm{B}_1(\boldsymbol{\theta}, \rho) \cap [\rho N]$. By choosing c sufficiently small, an application of Lemma 3.6.8 followed by Lemma 3.6.4 reveals that

$$\Psi_{B,\mathbf{y}}(f_{\rm str} + f_{\rm sml}) \geqslant \tfrac{1}{2}\delta^s |B|^{s-n-1} N$$

holds for all $\mathbf{y} \in (\mathrm{B}_k(\boldsymbol{\theta}, \rho) \cap [(\rho N)^{1/k}])^t$.

Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$-syndetic set, and let

$$\Omega := \left(\mathrm{B}_k(\boldsymbol{\theta}, \rho) \cap S \cap [(\rho N)^{1/k}]\right)^t \cap \{\mathbf{y} \in \mathbb{N}^t : \mathbf{C}\mathbf{y}^{\otimes k} = \mathbf{0}\}.$$

By the non-negativity of $f_{\rm str} + f_{\rm sml}$, we deduce from (3.24) the bound

$$\Lambda_1(f_{\rm str} + f_{\rm sml}; S \cap [N^{1/k}]) \geqslant \tfrac{1}{2}\delta^s |B|^{s-n-1} N |\Omega|. \tag{3.28}$$

By Corollary 3.6.7, the set $\mathrm{B}_1(\boldsymbol{\theta}, \rho)$ is multiplicatively $[M_1(\rho, d, 1, 1)]$-syndetic, whilst $\mathrm{B}_k(\boldsymbol{\theta}, \rho) \cap S$ is multiplicatively $[M_1(\rho, d, k, M)]$-syndetic. Note that Lemma 3.4.4 implies that if $N \geqslant 2\tilde{M}^2$, then any multiplicatively $[\tilde{M}]$-syndetic set has density at least $\frac{1}{2}\tilde{M}^{-2}$ on $[N]$. Thus, if N satisfies

$$cN^{1/k} \geqslant \delta^{-s} L^2 \mathcal{M}\left(M_1(c\delta^s L_0^{-2}, L_0, 1, 1) + M_1(c\delta^s L_0^{-2}, L_0, 2, M)\}\right), \tag{3.29}$$

for c sufficiently small, then we obtain the bounds

$$|B| \geqslant \tfrac{1}{4} c\delta^s L^{-2} M_1(c\delta^s L^{-2}, L, 1, 1)^{-2} N; \tag{3.30}$$

$$|\Omega| \geqslant \kappa\left(M_1(c\delta^s L^{-2}, L, k, M)\right) \cdot \left(\tfrac{1}{4} c\delta^s L^{-2} N\right)^{\frac{t}{k} - m}. \tag{3.31}$$

Lemma 3.6.2 therefore gives

$$\Lambda_1(A; S \cap [N^{1/k}]) \geqslant \left(\gamma(\delta, k, L, M, \mathbf{M}) - O_{k,\mathbf{M}}(\mathcal{F}(L)^{-\eta})\right) N^{s + \frac{t}{k} - (n+m)}, \tag{3.32}$$

where $\eta := (2k^2(n+m)+2)^{-1}$, and $\gamma(\delta, k, L, M, \mathbf{M})$ is the function of δ, k, L, M, and $\mathbf{M}$ obtained by substituting the lower bounds (3.30) and (3.31) into (3.28) (ignoring the factors of N). Using (3.26), we may assume that $\gamma(\delta, k, L, M, \mathbf{M})$ is a decreasing function of L, for fixed $\delta, k, M, \mathbf{M}$. We can therefore construct a monotone increasing function $\mathcal{F}_0 : \mathbb{N} \to \mathbb{R}_{\geqslant 0}$ such that

$$2C\mathcal{F}_0(x)^{-\eta} \leqslant \gamma(\delta, k, x, M, \mathbf{M}) \tag{3.33}$$

holds for all $x \in \mathbb{N}$, where $C = C(k, \mathbf{M}) > 1$ is the implicit positive constant appearing in (3.32) (which can be assumed to be greater than 1). We can then extend $\mathcal{F}_0$ to a monotone increasing function $\mathcal{F} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ by interpolation. With this choice of $\mathcal{F}$, we deduce from (3.32) that

$$\Lambda_1(A; S \cap [N^{1/k}]) \gg_{\delta,k,M,\mathbf{M}} N^{s+\frac{t}{k}-(n+m)},$$

as required. $\square$

Bibliography

[Aig79] M. Aigner, *Combinatorial Theory*, Springer, Berlin, 1979.

[Bou89] J. Bourgain, *On $\Lambda(p)$-subsets of squares*, Israel J. Math. **67** (1989), 291–311.

[BDG16] J. Bourgain, C. Demeter, and L. Guth, *Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three*, Ann. of Math. **184** (2016), no. 2, 633–682.

[BP17] T. D. Browning and S. Prendiville, *A transference approach to a Roth-type theorem in the squares*, IMRN (2017), no. 7, 2219–2248.

[BC92] J. Brüdern and R. J. Cook, *On simultaneous diagonal equations and inequalities*, Acta Arith. **62** (1992), 125–149.

[Cha20] J. Chapman, *Partition regularity and multiplicatively syndetic sets*, Acta Arith. **196** (2020), 109–138.

[Cho18] S. Chow, *Roth–Waring–Goldbach*, IMRN (2018), no. 8, 2341–2374.

[CLP21] S. Chow, S. Lindqvist, and S. Prendiville, *Rado's criterion over squares and higher powers*, J. Eur. Math. Soc. (JEMS) **23** (2021), no. 6, 1925–1997.

[Con08] K. Conrad, *The congruent number problem*, Harvard College Mathematics Review, **2** (2008), 58–74.

[Coo71] R. J. Cook, *Simultaneous quadratic equations*, J. Lond. Math. Soc. (2) **4** (1971), 319–326.

[DL69] H. Davenport and D. J. Lewis, *Simultaneous equations of additive type*, Philos. Trans. Roy. Soc. A **264** (1969), 557–595.

[Ebe16] S. Eberhard, *The abelian arithmetic regularity lemma*, preprint arXiv:1606.09303v1 (2016).

[Edm65] J. Edmonds, *Minimum partition of a matroid into independent subsets*, J. Res. Nat. Bureau Standards **69**B (1965), 67–72.

[FGR88] P. Frankl, R. L. Graham, and V. Rödl, *Quantitative theorems for regular systems of equations*, J. Combin. Theory Ser. A **47** (1988), 246–261.

[GW10] W. T. Gowers and J. Wolf, *The true complexity of a system of linear equations*, Proc. Lond. Math. Soc. **100** (2010), 155–176.

[Gre05A] B. J. Green, *Roth's theorem in the primes*, Ann. of Math. **161** (2005), 1609–1636.

[Gre05B] B. J. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), no. 2, 340–376.

[GL19] B. J. Green and S. Lindqvist, *Monochromatic solutions to $x + y = z^2$*, Canad. J. Math. **71** (3) (2019), 579–605.

[GS16] B. J. Green and T. Sanders, *Monochromatic sums and products*, Discrete Anal. (2016), Paper No. 5, 43 pp.

[GT10] B. J. Green and T. Tao, *An arithmetic regularity lemma, an associated counting lemma, and applications*, An irregular mind, Bolyai, Soc. Math. Stud. **21**, János Bolyai Math. Soc., Budapest, (2010), 261–334.

[Lef91] H. Lefmann, *On partition regular systems of equations*, J. Combin. Theory Ser. A **58** (1991), 35–53.

[Lin19] S. Lindqvist, *Quadratic phenomena in additive combinatorics and number theory*, PhD thesis, University of Oxford, 2019.

[LPW88] L. Low, J. Pitman, and A. Wolff, *Simultaneous diagonal congruences*, J. Number Theory **29** (1988), no. 1, 31–59.

[Pre17] S. Prendiville, *Four variants of the Fourier analytic transference principle*, Online J. Anal. Comb. **12** (2017), 25 pp.

[Rad33] R. Rado, *Studien zur Kombinatorik*, Math. Z. **36** (1933), 242–280.

[Rad43] R. Rado, *Note on combinatorial analysis*, Proc. London Math. Soc. **48** (1943), 122–160.

[Sch77] W. M. Schmidt, *Small fractional parts of polynomials*, CBMS Regional Conference Series in Math. **32**, Amer. Math. Soc., Providence, 1977.

[Sze75] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.

[Tao12] T. Tao, *Higher order Fourier analysis*, Graduate Studies in Mathematics **142**, American Mathematical Society, Providence, RI, 2012.

[Vau86] R. C. Vaughan, *On Warings problem for cubes*, J. Reine Angew. Math. **365** (1986), 122–170.

[Vau89] R. C. Vaughan, *A new iterative method in Warings problem*, Acta Math. **162** (1989), 1–71.

[Woo12] T. D. Wooley, *The asymptotic formula in Waring's problem*, IMRN (2012), no. 7, 1485–1504.

Chapter 4

On the Ramsey number of the Brauer Configuration

Abstract. We obtain a double exponential bound in Brauer's generalisation of van der Waerden's theorem, which concerns progressions with the same colour as their common difference. Such a result has been obtained independently and in much greater generality by Sanders. Using Gowers' local inverse theorem, our bound is quintuple exponential in the length of the progression. We refine this bound in the colour aspect for three-term progressions, and combine our arguments with an insight of Lefmann to obtain analogous bounds for the Ramsey numbers of certain nonlinear quadratic equations.

4.1 Introduction

Schur's theorem states that in any partition of the positive integers into finitely many pieces, at least one part contains a solution to the equation $x + y = z$. By a theorem of van der Waerden, the same is true for the equation of three-term arithmetic progressions $x + y = 2z$. A common generalisation of these theorems due to Brauer states that, in any finite colouring of the positive integers, there is a monochromatic arithmetic progression of length k with the same colour as its common difference.

There is a finitary analogue of these results, asserting that the same holds for colourings of the interval $\{1, 2, \ldots, N\}$, provided that N is sufficiently large in terms of the number of colours. Determining the minimal such number N (the *Ramsey* or

Rado number of the system) has received much attention for arithmetic progressions, and a celebrated breakthrough of Shelah [She88] showed that these numbers (*van der Waerden numbers*) are primitive recursive. One spectacular consequence of Gowers' work on Szemerédi's theorem [Gow01] is a bound on van der Waerden numbers which is quintuple exponential in terms of the length of the progression, and double exponential in terms of the number of pieces of the partition.

Using Gowers' local inverse theorem for the uniformity norms, we obtain a bound for the Ramsey number of Brauer configurations which is comparable to that obtained for arithmetic progressions.

Theorem 4.1.1 (Ramsey bound for Brauer configurations)**.** *There exists an absolute constant $C = C(k)$ such that if $r \geqslant 2$ and $N \geqslant \exp\exp(r^C)$, then any r-colouring of $\{1, 2, ..., N\}$ yields a monochromatic k-term progression which is the same colour as its common difference. Moreover, it suffices to assume that*

$$N \geqslant 2^{2^{r^{2^{2^{k+10}}}}}. \tag{4.1}$$

A double exponential bound has been obtained independently and in maximal generality by Sanders [San20], who bounds the Ramsey number of an *arbitrary* system of linear equations with this colouring property, often termed *partition regularity.*[1] Our results and those of Sanders are the first quantitatively effective bounds for configurations lacking translation invariance and of 'true complexity' greater than one (see [GW10] for further explanation).

Gowers [Gow01] obtains the bound (4.1) for progressions of length $k+1$, this being the appropriate analogue of the $(k+1)$-point Brauer configuration in Theorem 4.1.1. For a four-point Brauer configuration, we improve the exponent of r on combining our method with an energy-increment argument of Green and Tao [GT09].

Theorem 4.1.2 (Improved bound for four-point Brauer configurations)**.** *There exists an absolute constant C such that if $N \geqslant \exp\exp(Cr\log^2 r)$, then in any r-colouring of $\{1, 2, \dots, N\}$ there exists a monochromatic three-term progression with the same colour as its common difference.*

[1] A system of equations is said to be *partition regular* (over $\mathbb{N}$) if any finite colouring of $\mathbb{N}$ yields a monochromatic solution to the system.

Due to an insight of Lefmann [Lef91] we are able to use (a variant of) Theorem 4.1.1 to bound the Ramsey number of certain partition regular *nonlinear* equations.

Theorem 4.1.3. *Let $a_1, \ldots, a_s \in \mathbb{Z} \setminus \{0\}$ satisfy the following:*

(i) *there exists a non-empty set $I \subset [s]$ such that $\sum_{i \in I} a_i = 0$;*

(ii) *the system*
$$x_0^2 \sum_{i \notin I} a_i + \sum_{i \in I} a_i x_i^2 = \sum_{i \in I} a_i x_i = 0.$$
has a rational solution with $x_0 \neq 0$.

Then there exists an absolute constant $C = C(a_1, \ldots, a_s)$ such that for $r \geqslant 2$ and $N \geqslant \exp\exp(r^C)$, any r-colouring of $\{1, 2, \ldots, N\}$ yields a monochromatic solution to the diagonal quadric
$$a_1 x_1^2 + \cdots + a_s x_s^2 = 0.$$

Previous work

Hitherto, little is recorded regarding the Ramsey number of general partition regular systems. Cwalina–Schoen [CS17] observe that one can use Gowers' bounds [Gow01] in Szemerédi's theorem to obtain a bound which is tower in nature, of height proportional to $5r$. Gowers' methods are well suited to delivering double exponential bounds for so-called translation invariant systems (such as arithmetic progressions), but such systems are far from typical. In Cwalina–Schoen [CS17], Fourier-analytic arguments are adapted to give an exponential bound on the Ramsey number of a single partition regular equation. The first author [Cha20] has shown how multiplicatively syndetic sets allow one to reduce the tower height to $(1 + o(1))r$ for the four-point Brauer configuration
$$x, \ x + d, \ x + 2d, \ d. \tag{4.2}$$

Our method

The approach underlying Theorem 4.1.1 generalises that of Roth [Rot53] and Gowers [Gow01]. Given a subset A of $[N] := \{1, 2, \ldots, N\}$ of density δ, Roth uses a density increment procedure to locate a subprogression $P = a + q \cdot [M]$ of length $M \geqslant N^{\exp(-C/\delta)}$

where $A \cap P$ has density at least δ and is 'Fourier uniform', in the sense that all but its trivial Fourier coefficients are small. An application of Fourier analysis (in the form of the circle method) shows that such sets possess of order $\delta^3 M^2$ three-term progressions. This yields a non-trivial three-term progression provided that $N \geqslant \exp\exp(C/\delta)$.

The above application of the circle method relies crucially on the translation-dilation invariance of three-term progressions, so that the number of configurations in P is the same as that in $[M]$. Unfortunately, the Brauer configuration (4.2) is not translation invariant. To overcome the lack of translation-invariance, given a colouring $A_1 \cup \cdots \cup A_r = [N]$, we use Gowers' local inverse theorem for the uniformity norms [Gow01] to run a density increment procedure with respect to the *maximal translate density*

$$\sum_{i=1}^{r} \max_{a} \frac{|A_i \cap (a + q \cdot [M])|}{M}.$$

This outputs (see Lemma 4.3.7) a homogeneous progression $q \cdot [M]$ such that for each colour class A_i there is a translate $a_i + q \cdot [M]$ on which A_i achieves its maximal translate density and on which A_i is suitably uniform. For the four-point Brauer configuration (4.2), the correct notion of uniformity is *quadratic uniformity* (as measured by the Gowers U^3-norm).

Write α_i for the density of A_i on the maximal translate $a_i + q \cdot [M]$ and β_i for its density on the homogeneous progression $q \cdot [M]$. An application of quadratic Fourier analysis shows that the number of four-point Brauer configurations (4.2) satisfying

$$\{x, x+d, x+2d\} \subset A_i \cap (a_i + q \cdot [M]) \quad \text{and} \quad d \in A_i \cap q \cdot [M]$$

is of order $\alpha_i^3 \beta_i M^2$. By the pigeonhole principle, some colour class has $\beta_i \geqslant 1/r$, which also implies that $\alpha_i \geqslant 1/r$, so we deduce that some colour class contains at least $r^{-4} M^2$ Brauer configurations. Unravelling the quantitative dependence in our density increment then yields a double exponential bound on N in terms of r.

In §4.2 we give a more detailed exposition of this method for the model problem of Schur's theorem in the finite vector space $\mathbb{F}_2^n$. In §4.3 we generalise the argument to arbitrarily long Brauer configurations over the integers. We improve our bound for four-point Brauer configurations in §4.4. Finally, in §4.5 we show how our methods give comparable bounds for the Ramsey number of certain quadratic equations.

Notation

The set of positive integers is denoted by $\mathbb{N}$. Given $x \geqslant 1$, we write $[x] := \{1, 2, ..., \lfloor x \rfloor\}$. If f and g are functions, and g takes only positive values, then we use the Vinogradov notation $f \ll g$ if there exists an absolute positive constant C such that $|f(x)| \leqslant Cg(x)$ for all x. We also write $g \gg f$ or $f = O(g)$ to denote this same property. The letters C and c are used to denote absolute constants, whose values may change from line to line. Typically C denotes a large constant $C > 1$, whilst c denotes a small constant $0 < c < 1$.

Acknowledgements

We thank Tom Sanders for alerting us to the existence of [San20], and for his generosity in synchronising release. We also thank Tom Sanders for his thorough and helpful comments. The second author thanks Ben Green for suggesting this problem.

4.2 Schur in the finite field model

We illustrate the key ideas of our approach in proving Schur's theorem over $\mathbb{F}_2^n$. This asserts that, provided the dimension n is sufficiently large relative to the number of colours r, any partition $\mathbb{F}_2^n = A_1 \cup \cdots \cup A_r$ possesses a colour class A_i containing vectors x, y, z with $y \neq 0$ and such that $x + y = z$. The goal of this section is to obtain a quantitative bound on the dimension n in terms of r.

The argument of this section is purely expository, the resulting bound being slightly worse than that given by a standard application of Ramsey's theorem (see [GRS90, §3.1]) or Schur's original argument (see [CS17]). We have since learned that the same ideas are discussed in Shkredov [Shk10, §5].

Theorem 4.2.1 (Schur in the finite field model)**.** *Consider a partition of $\mathbb{F}_2^n$ into r sets $A_1, ..., A_r$. If n satisfies*

$$n > \sqrt{2}\, r^3 + \log_2(2r),$$

then there exists $i \in [r]$ and $x, y, z \in A_i$ with $y \neq 0$ such that $x + y = z$.

Remark (Distinctness of x, y, z)**.** One can guarantee that the $x, y, z \in A_i$ that are obtained are distinct and non-zero by introducing a new partition $\mathbb{F}_2^n = A_0' \cup A_1' \cup \cdots \cup A_r'$ by setting $A_0' = \{0\}$ and $A_i' := A_i \setminus \{0\}$ for all $i \in [r]$. Applying the above theorem (with r replaced by $r+1$) to this new partition gives distinct non-zero $x, y, z \in A_i$ for some $i \in [r]$ satisfying $x + y = z$.

Inspired by Cwalina–Schoen [CS17], we deduce Theorem 4.2.1 from the following dichotomy. This argument is a variant of Sanders's '99% Bogolyubov theorem' [San08], which asserts that the difference set of a dense set contains 99% of a subspace of bounded codimension.

Lemma 4.2.2 (Sparsity–expansion dichotomy)**.** *Let $A_1 \cup \cdots \cup A_r = \mathbb{F}_2^n$ be a partition of $\mathbb{F}_2^n$ into r parts. Then there exists a subspace $H \leqslant \mathbb{F}_2^n$ with $\operatorname{codim}(H) \leqslant \sqrt{2}r^3$ such that for any $i \in [r]$ we have one of the two following possibilities.*

- *(Sparsity).*
$$|A_i \cap H| < \tfrac{1}{r}|H|; \tag{4.3}$$
- *(Expansion).*
$$|(A_i - A_i) \cap H| \geqslant \left(1 - \tfrac{1}{2r}\right)|H|. \tag{4.4}$$

The idea is that, as one of the colour classes A_i is dense, its difference set $A_i - A_i$ must (by Sanders' result) contain 99% of a 'large' subspace H. Were A_i itself to contain more that 1% of this subspace, then we would be done, since then $(A_i - A_i) \cap A_i \neq \emptyset$ and we would obtain the desired Schur triple $x, y, z \in A_i$. Unfortunately, this cannot always be guaranteed: consider the case in which A_i is a non-trivial coset of a subspace of co-dimension 1.

To overcome this, we run Sanders' proof with respect to all of the colour classes simultaneously, constructing a subspace H which is almost covered by $A_i - A_i$ for *all* $i \in [r]$. If such a H were obtainable we would be done as before, since (by the pigeonhole principle) some colour class has large density on H. Again, this is slightly too much to hope for, as sets which are 'hereditarily sparse' cannot be good candidates for a 99% Bogolyubov theorem. Fortunately, such sets can be accounted for in our argument.

Before we proceed to the proof of Lemma 4.2.2, let us use this lemma to prove our finite field model of Schur's theorem.

Proof of Theorem 4.2.1. Let H denote the subspace provided by the dichotomy. By the pigeonhole principle, there exists some A_i satisfying

$$|A_i \cap H| \geqslant \tfrac{1}{r}|H|.$$

Our assumption on the size of n then implies that

$$|A_i \cap H| > \tfrac{1}{2r}|H| + 1.$$

Since A_i is not sparse on H, in the sense of (4.3), it must instead satisfy the expansion property (4.4). By inclusion–exclusion

$$|A_i \cap (A_i - A_i) \cap H| \geqslant |A_i \cap H| + |(A_i - A_i) \cap H| - |H| > 1.$$

In particular, the set $A_i \cap (A_i - A_i)$ contains a non-zero element. □

4.2.1 A maximal translate increment strategy

It remains to prove Lemma 4.2.2. Following Sanders [San08], we accomplish this via density increment. We cannot merely increment the density of each individual colour class on translates of different subspaces, since our final dichotomy involves a single subspace H which is uniform for all A_i. We therefore have to increment a more subtle notion of density, namely the *maximal translate density*

$$\Delta_H = \Delta_H(A_1, \dots, A_r) := \sum_{i=1}^{r} \max_x \frac{|A_i \cap (x + H)|}{|H|}.$$

This is a non-negative quantity bounded above by r. It follows that a procedure passing to subspaces $H_0 \geqslant H_1 \geqslant H_2 \geqslant \dots$, which increments Δ_{H_i} by a constant amount at each iteration, must terminate in a constant number of steps (depending on r).

We first observe that to increment Δ_H it suffices to find a subspace where one of the colour classes increases their maximal translate density. To this end, write

$$\delta_H(A) := \max_x \frac{|A \cap (x + H)|}{|H|}. \tag{4.5}$$

Lemma 4.2.3 (Maximal translate density is preserved on passing to subspaces)**.** *Let $H_1 \geqslant H_2$ be subspaces of $\mathbb{F}_2^n$. Then for any $A \subset \mathbb{F}_2^n$ we have*

$$\delta_{H_2}(A) \geqslant \delta_{H_1}(A).$$

Proof. As $H_2 \leqslant H_1$, we can write H_1 as a disjoint union of cosets of H_2. This means that we can find $V \subset H_1$ such that $H_1 = \sqcup_{y \in V}(y + H_2)$. Hence

$$|A \cap (x + H_1)| = \sum_{y \in V} |A \cap (x + y + H_2)| \leqslant |V| \max_z |A \cap (z + H_2)|.$$

Choosing x so that A has maximal density on $x + H_1$ gives the result. □

We now prove Lemma 4.2.2 using a density increment strategy for the maximal translate density. The argument proceeds by showing that if our claimed dichotomy does not hold, then we may pass to a subspace on which the colour classes have larger maximal translate density.

The process of identifying such a subspace involves the use of Fourier analysis. Given a subspace $H \leqslant \mathbb{F}_2^n$ and a function $f : H \to \mathbb{C}$, we define the *Fourier transform* $\hat{f} : \hat{H} \to \mathbb{C}$ of f by

$$\hat{f}(\gamma) := \sum_{x \in H} f(x)\gamma(x).$$

Here $\hat{H}$ denotes the *dual group* of H, which is the group of homomorphisms $\gamma : H \to \mathbb{C}^{\times}$. Since every element of H has order at most 2, the value $\gamma(x)$ must be ± 1 for all $x \in H$ and $\gamma \in \hat{H}$.

Proof of Lemma 4.2.2. We proceed by an iterative procedure, at each stage of which we have a subspace $H = H^{(m)} \leqslant \mathbb{F}_2^n$ of codimension m satisfying

$$\Delta_{H^{(m)}} \geqslant \frac{m}{\sqrt{2}r^2}.$$

We initiate this procedure on taking $H^{(0)} := \mathbb{F}_2^n$. Since $\Delta_{H^{(m)}} \leqslant r$ this procedure must terminate at some $m \leqslant \sqrt{2}r^3$.

Given $H = H^{(m)}$ we define three types of colour class.

- (Sparse colours). A_i is *sparse* if

$$\delta_H(A_i) < \tfrac{1}{r};$$

- (Dense expanding colours). A_i is *dense expanding* if $\delta_H(A_i) \geqslant \frac{1}{r}$ and we have the expansion estimate

$$|(A_i - A_i) \cap H| > \left(1 - \tfrac{1}{2r}\right)|H|; \tag{4.6}$$

- (Dense non-expanding colours). A_i is *dense non-expanding* if it is neither sparse nor dense expanding.

If there are no dense non-expanding colour classes, then the dichotomy claimed in our lemma is satisfied, and we terminate our procedure. Let us show how the existence of a dense non-expanding colour class A_i allows the iteration to continue.

By the definition of maximal translate density, there exists t such that

$$|A_i \cap (t + H)| = \delta_H(A_i)|H|.$$

We define dense subsets $A, B \subset H$ by taking

$$A := (A_i - t) \cap H \quad \text{and} \quad B := H \setminus \big(A_i - A_i\big). \tag{4.7}$$

Writing α and β for the respective densities of A and B in H, our dense non-expanding assumption implies that $\alpha \geqslant 1/r$ and $\beta \geqslant 1/(2r)$. Moreover, it follows from our construction (4.7) that

$$\sum_{x-x'=y} 1_A(x)1_A(x')1_B(y) = 0.$$

Comparing this to the count

$$\sum_{x-x'=y} \alpha 1_H(x)1_A(x')1_B(y) = \alpha^2\beta|H|^2,$$

we deduce, on writing $f_A := 1_A - \alpha 1_H$, that

$$\left|\mathbb{E}_{\gamma\in\hat{H}}\overline{\hat{f}_A(\gamma)}\hat{1}_A(\gamma)\hat{1}_B(\gamma)\right| = \left|\sum_{x-x'=y} f_A(x)1_A(x')1_B(y)\right| \geqslant \alpha^2\beta|H|^2.$$

By Cauchy–Schwarz and Parseval's identity, there exists $\gamma \neq 1_H$ such that

$$\left|\sum_{x\in H} f_A(x)\gamma(x)\right| \geqslant \frac{\alpha^2\beta}{\sqrt{\alpha\beta}}|H| \geqslant \frac{|H|}{\sqrt{2}r^2}.$$

Partitioning H into level sets of γ, gives

$$\left|\sum_{\gamma(x)=1} f_A(x)\right| + \left|\sum_{\gamma(x)=-1} f_A(x)\right| \geqslant \frac{|H|}{\sqrt{2}r^2}.$$

Since f_A has mean zero, we deduce that the two terms on the left of the above inequality are equal. This implies that there exists $\omega \in \{\pm 1\}$ such that

$$2\sum_{\gamma(x)=\omega} f_A(x) \geqslant \frac{|H|}{\sqrt{2}r^2}.$$

Observe that, since $\gamma \neq 1_H$, the set $H' := \{x \in H : \gamma(x) = 1\}$ is a subspace of H of codimension 1. Hence on choosing $y \in H$ with $\gamma(y) = \omega$ we have

$$\frac{|A \cap (y + H')|}{|H'|} \geqslant \alpha + \frac{1}{\sqrt{2}r^2}.$$

By combining this with Lemma 4.2.3 and our definition (4.7) of A, we deduce that

$$\Delta_{H'} \geqslant \Delta_H + \tfrac{1}{\sqrt{2}r^2} \qquad \text{and} \qquad \operatorname{codim}(H') = \operatorname{codim}(H) + 1.$$

We have therefore established that our iteration may continue, completing the proof of the lemma. □

4.3 Brauer configurations over the integers

In this section we use higher order Fourier analysis to study longer Brauer configurations and prove Theorem 4.1.1. Henceforth, we fix the parameter $k \geqslant 2$ to denote the length of the progression in the Brauer configuration under consideration. We emphasise that this section streamlines substantially if the reader is only interested in a double exponential bound in the colour aspect, as opposed to the more explicit bound (4.1).

Given finitely supported $f_1, f_2, ..., f_k, g : \mathbb{Z} \to \mathbb{R}$ we introduce the counting operator

$$\Lambda(f_1, f_2, ..., f_k; g) := \sum_{d,x \in \mathbb{Z}} f_1(x) f_2(x + d) \cdots f_k(x + (k-1)d) g(d). \tag{4.8}$$

For brevity, write $\Lambda(f; g) := \Lambda(f, f, ..., f; g)$. For given finite sets $A, B \subset \mathbb{N}$, the number of arithmetic progressions of length k in A with common difference in B is given by $\Lambda(1_A; 1_B)$.

Lemma 4.3.1. *Let $M \in \mathbb{N}$ with $M \geqslant k$. If $B \subset [M/(2k-2)]$, then*

$$\Lambda(1_{[M]}; 1_B) \geqslant \tfrac{1}{2}|B|M.$$

Proof. Since $B \subset [M/(2k-2)]$, we have $M - (k-1)d \geqslant M/2$ for all $d \in B$. Thus

$$\Lambda(1_{[M]}; 1_B) = \sum_{d \in B} (M - (k-1)d) \geqslant \tfrac{1}{2}|B|M.$$

□

4.3.1 Gowers norms

Gowers [Gow98] observed that arithmetic progressions of length four or more are not controlled by ordinary (linear) Fourier analysis. Similarly, four-point Brauer configurations (and longer) require higher order notions of uniformity – they have *true complexity* greater than 1 (see [GW10] for further details). To overcome this difficulty, Gowers introduced a sequence of norms which can be used to measure the higher order uniformity of sets and functions.

Definition (U^d norms)**.** Let $f : \mathbb{Z} \to \mathbb{R}$ be a finitely supported function. For each $d \geqslant 2$, the U^d *norm* $\|f\|_{U^d}$ of f is defined by

$$\|f\|_{U^d} := \left(\sum_{x \in \mathbb{Z}} \sum_{\mathbf{h} \in \mathbb{Z}^d} \Delta_{h_1,\ldots,h_d} f(x) \right)^{1/2^d}, \tag{4.9}$$

where the *difference operators* Δ_h are defined inductively by

$$\Delta_h f(x) := f(x) f(x+h)$$

and

$$\Delta_{h_1,\ldots,h_d} f := \Delta_{h_1} \Delta_{h_2} \cdots \Delta_{h_d} f.$$

Remark. In the literature, and in Gowers' original paper, it is common to work with functions $f : \mathbb{Z}/p\mathbb{Z} \to \mathbb{R}$, defining $\|f\|_{U^d(\mathbb{Z}/p\mathbb{Z})}$ by summing over $x \in \mathbb{Z}/p\mathbb{Z}$ and $\mathbf{h} \in (\mathbb{Z}/p\mathbb{Z})^d$ in (4.9). Given a prime $p > N$, one can embed the interval $[N]$ into $\mathbb{Z}/p\mathbb{Z}$ by reduction modulo p. This allows us to identify a function $f : [N] \to \mathbb{R}$ with an extension $\tilde{f} : \mathbb{Z}/p\mathbb{Z} \to \mathbb{R}$ on taking $\tilde{f}(x) = 0$ for all $x \in (\mathbb{Z}/p\mathbb{Z}) \setminus [N]$. One can observe that if $p > 2(d+1)N$, then $\|f\|_{U^d} = \|\tilde{f}\|_{U^d(\mathbb{Z}/p\mathbb{Z})}$. This is due to the fact that the interval $[N] \subset \mathbb{Z}$ and the embedding of $[N]$ into $\mathbb{Z}/p\mathbb{Z}$ are *Freiman isomorphic of order* $d+1$ (see [TV06, §5.3] for further details).

We note that if $f : \mathbb{Z} \to [-1, 1]$ is supported on $[N]$, then

$$\tfrac{1}{2} N^{\frac{d+1}{2^d}} \leqslant \|f\|_{U^d} \leqslant N^{\frac{d+1}{2^d}}. \tag{4.10}$$

The lower bound follows from inductively applying the Cauchy–Schwarz inequality in the form

$$\|f\|_{U^d} = \left(\sum_{h_1,\ldots,h_{d-1}} \left| \sum_x \Delta_{h_1,\ldots,h_{d-1}} f(x) \right|^2 \right)^{1/2^d} \geqslant (2N)^{-\frac{d-1}{2^d}} \|f\|_{U^{d-1}},$$

the factor of 2 resulting from the observation that if $\mathrm{supp}(f) \subset [N]$, then the h_i in (4.9) contribute only if $h_i \in (-N, N)$. The upper bound is a consequence of the fact that $\|f\|_{U^d}^{2^d}$ counts the number of solutions to a system of $2^d - d - 1$ independent linear equations in 2^d variables, each weighted by f.

Definition (Uniform of degree d)**.** We say that $A \subset [N]$ is *ε-uniform of degree d* if

$$\|1_A - \mathbb{E}_{[N]}(1_A)1_{[N]}\|_{U^{d+1}} \leqslant \varepsilon\|1_{[N]}\|_{U^{d+1}},$$

where $\mathbb{E}_{[N]}(1_A) := |A \cap [N]|/N$ denotes the density of A on $[N]$. More generally, given $P \subset [N]$, we say that A is *ε-uniform of degree d on P* if

$$\|1_A - \mathbb{E}_P(1_A)1_P\|_{U^{d+1}} \leqslant \varepsilon\|1_P\|_{U^{d+1}}.$$

Gowers showed that one can study sets which lack arithmetic progressions of length k by considering their uniformity. If a set has density α in $[N]$ and is ε-uniform of degree $k-2$, for some small $\varepsilon = \varepsilon(k, \alpha)$, then A contains a proportion of α^k of the total progressions of length k in the interval $[N]$. Hence the only way a uniform set can lack k-term progressions is if it has few elements.

A similar result holds for Brauer configurations, see for instance [GT10, Appendix C]. In order to avoid the introduction of an (admittedly harmless) absolute constant resulting from the passage to a cyclic group, we give the simple proof.

Lemma 4.3.2 (Generalised von Neumann for Λ)**.** *Let $f_1, ..., f_k, g : [N] \to [-1, 1]$. Then for each $j \in [k]$ we have*

$$|\Lambda(f_1, ..., f_k; g)| \leqslant N^2 \left(\frac{\|f_j\|_{U^k}^{2^k}}{N^{k+1}}\right)^{1/2^k} \left(\frac{\|g\|_{U^k}^{2^k}}{N^{k+1}}\right)^{1/2^k}.$$

Proof. We prove the case where $j = k$. The other cases follow on performing a change of variables $x' = x + id$ preceding each application of the Cauchy-Schwarz inequality.

Applying the Cauchy-Schwarz inequality with respect to the x variable shows that $|\Lambda(f_1, ..., f_k; g)|$ is bounded above by

$$\left(\sum_{x\in\mathbb{Z}} |f_1(x)|^2\right)^{1/2} \left(\sum_{x,d,d'\in\mathbb{Z}} g(d)g(d') \prod_{i=1}^{k-1} f_{i+1}(x + id)f_{i+1}(x + id')\right)^{1/2}.$$

Using the fact that $|f_1(x)| \leqslant 1_{[N]}(x)$ holds for all $x \in \mathbb{Z}$, and by performing a change of variables $d' = d + h$, we deduce that

$$|\Lambda(f_1, ..., f_k; g)|^2 \leqslant N \sum_{x,h\in\mathbb{Z}} \sum_{d\in\mathbb{Z}} \Delta_h g(d) \prod_{i=1}^{k-1} \Delta_{ih} f_{i+1}(x + id).$$

By applying the Cauchy-Schwarz inequality a further $k - 2$ times, each time with respect to all variables except for d, we see that $|\Lambda(f_1, ..., f_k; g)|^{2^{k-1}}$ is bounded above by

$$N^{2^k-k-1} \sum_{\mathbf{h}\in\mathbb{Z}^{k-1}} \sum_{x\in\mathbb{Z}} \Delta_{(k-1)h_1,(k-2)h_2,...,h_{k-1}} f_k(x) \sum_{d\in\mathbb{Z}} \Delta_{h_1,...,h_{k-1}} g(d).$$

By applying Cauchy-Schwarz with respect to the $\mathbf{h}$ variable, the above sum is at most $S^{1/2}\|g\|_{U^k}^{2^{k-1}}$, where S is equal to

$$\sum_{\mathbf{h}\in\mathbb{Z}^{k-1}} \left| \sum_{x\in\mathbb{Z}} \Delta_{(k-1)h_1,(k-2)h_2,...,h_{k-1}} f_k(x) \right|^2 .$$

Since the terms in the sum over $\mathbf{h}$ are non-negative, we can extend the summation from $\mathbb{Z}^{k-1}$ to $(k-1)^{-1} \cdot \mathbb{Z} \times (k-2)^{-1} \cdot \mathbb{Z} \times \cdots \times \mathbb{Z}$, yielding the lemma. □

Corollary 4.3.3 (U^k controls Λ)**.** *Let* $f_1, f_2, g : [N] \to [0, 1]$*. Then*

$$|\Lambda(f_1; g) - \Lambda(f_2; g)| \leqslant kN^2 \frac{\|f_1 - f_2\|_{U^k}}{N^{(k+1)2^{-k}}}.$$

Proof. Observe that $\Lambda(f_1; g) - \Lambda(f_2; g)$ can be written as the sum of k terms

$$\Lambda(f_1 - f_2, f_1, ..., f_1; g) + \Lambda(f_2, f_1 - f_2, f_1, ..., f_1; g) + \cdots + \Lambda(f_2, ..., f_2, f_1 - f_2; g).$$

Recall from (4.10) that $\|g\|_{U^k}^{2^k} \leqslant N^{k+1}$. Since $f_1 - f_2$ takes values in $[-1, 1]$, the result now follows from the triangle inequality and Lemma 4.3.2. □

Lemma 4.3.1 shows us that, for any non-empty $B \subset [N/(2k-2)]$ and $\alpha > 0$, we have

$$\Lambda(\alpha 1_{[N]}; 1_B) \geqslant \tfrac{1}{2}\alpha^k |B| N.$$

Combining this with Corollary 4.3.3 we see that, if $A \subset [N]$ has density $\alpha > 0$ and is ε-uniform of degree $k - 1$ for some 'very small' $\varepsilon > 0$, then the difference

$$|\Lambda(1_A; 1_B) - \Lambda(\alpha 1_{[N]}; 1_B)|$$

is also small. This then implies that A contains an arithmetic progression of length k with common difference in B. Hence sets A lacking such arithmetic progression cannot be uniform. A key observation of Gowers is that this lack of uniformity implies that the set A exhibits significant bias towards a long arithmetic progression inside $[N]$.

Gowers' density increment lemma. *Let $d \geqslant 1$ and $0 < \varepsilon \leqslant \frac{1}{2}$. Suppose that $A \subset [N]$ is not ε-uniform of degree d as in Definition 4.3.1. Then, on setting*

$$\eta = \eta(d, \varepsilon) := \frac{1}{4}\left(\frac{\varepsilon}{8(d+2)}\right)^{2^{d+1+2^{d+10}}}, \tag{4.11}$$

there exists an arithmetic progression $P \subset [N]$ such that

$$|P| \geqslant \eta N^{\eta} \quad \textit{and} \quad \frac{|A \cap P|}{|P|} \geqslant \frac{|A|}{N} + \eta.$$

Proof. Let p be a prime in the interval $2(d+2)N < p \leqslant 4(d+2)N$, so that on setting $f := 1_A - \mathbb{E}_{[N]}(1_A)1_{[N]}$ and viewing this as a function on $\mathbb{Z}/p\mathbb{Z}$ the lower bound in (4.10) gives

$$\sum_{\mathbf{h} \in (\mathbb{Z}/p\mathbb{Z})^d} \left| \sum_{x \in \mathbb{Z}/p\mathbb{Z}} \Delta_h f(x) \right|^2 \geqslant \left(\tfrac{\varepsilon}{8(d+2)}\right)^{2^{d+1}} p^{d+2}.$$

Hence, according to Gowers' [Gow01, p.478] definition of α-uniformity, f is not α-uniform of degree d on $\mathbb{Z}/p\mathbb{Z}$ with

$$\alpha = \left(\tfrac{\varepsilon}{8(d+2)}\right)^{2^{d+1}}.$$

Let $\beta := \alpha^{2^{2^{d+10}}}$. Applying Gowers' local inverse theorem for the U^{d+1}-norm [Gow01, Theorem 18.1] there exists a partition of $\mathbb{Z}/p\mathbb{Z}$ into (integer) arithmetic progressions $P_1, \ldots, P_M$ with $M \leqslant p^{1-\beta}$ and such that

$$\sum_{j=1}^{M} \left| \sum_{x \in P_j} f(x) \right| \geqslant \beta p.$$

Since f is supported on $[N]$, we may assume that $P_j \subset [N]$ for all j. As f has mean zero we may apply (the proof of) [Gow01, Lemma 5.15] to obtain a progression $P \subset [N]$ with $|P| \geqslant \frac{1}{4}\beta p^{\beta}$ which also satisfies

$$\sum_{x \in P} f(x) \geqslant \tfrac{1}{4}\beta |P|.$$

□

4.3.2 Maximal translate density

As in the previous section, we prove Theorem 4.1.1 by a maximal translate density increment argument. For $q, M \in \mathbb{N}$ and $A \subset \mathbb{Z}$, define the *maximal translate density*

$$\delta_{q,M}(A) := \max_{x\in\mathbb{Z}} \frac{|A \cap (x + q \cdot [M])|}{M}.$$

Given a collection of non-empty subsets $A_1, ..., A_r \subset [N]$, we collate their densities into the quantity

$$\Delta(q, M) = \Delta(q, M; \{A_i\}_{i=1}^r) := \sum_{i=1}^r \delta_{q,M}(A_i).$$

We write $\Delta(q, M)$ when it is clear from the context which collection of sets $\{A_i\}_{i=1}^r$ we are working with.

In the previous section, where we worked with subspaces of $\mathbb{F}_2^n$, we showed (Lemma 4.2.3) that the maximal translate density does not decrease when passing to a subspace. This is no longer true when passing to subprogressions in $\mathbb{Z}$. However, we can still increment $\Delta(q, M)$ if the subprogression we pass to is not too long.

Lemma 4.3.4 (Approximately preserving max translate density)**.** *Given positive integers M, M_1, q, q_1 and a finite set $A \subset \mathbb{Z}$, we have*

$$\delta_{qq_1,M_1}(A) \geqslant \delta_{q,M}(A)\left(1 - \tfrac{q_1 M_1}{M}\right).$$

Proof. By definition of $\delta_{q,M}$, we can find $t \in \mathbb{Z}$ such that

$$\delta_{q,M}(A)M = |A \cap (t + q \cdot [M])|.$$

Let $\tilde{A} := A \cap (t + q \cdot [M])$. Note that

$$\delta_{q,M}(\tilde{A})M = |\tilde{A}| = \delta_{q,M}(A)M.$$

Now observe that the collection of translates $\{x + qq_1 \cdot [M_1] : x \in \mathbb{Z}\}$ covers $\mathbb{Z}$, and each integer $m \in \mathbb{Z}$ lies in exactly M_1 such translates. This gives

$$\sum_{x\in\mathbb{Z}} |\tilde{A} \cap (x + qq_1 \cdot [M_1])| = |\tilde{A}|M_1 = \delta_{q,M}(A)MM_1. \tag{4.12}$$

Let Ω be given by

$$\Omega := \{x \in \mathbb{Z} : \tilde{A} \cap (x + qq_1 \cdot [M_1]) \neq \emptyset\}.$$

Now suppose $x \in \Omega$. Since $\tilde{A} \subset t + q \cdot [M]$, we can find $u \in [M]$ and $u_1 \in [M_1]$ such that $x - t = q(u - q_1 u_1)$. From this we see that

$$(x - t) \in [q(1 - q_1 M_1), q(M - q_1)] \cap (q \cdot \mathbb{Z}).$$

We therefore deduce that $|\Omega| \leqslant M + q_1 M_1$. Applying the pigeonhole principle to (4.12), we conclude that

$$\delta_{qq_1, M_1}(A) \geqslant \delta_{qq_1, M_1}(\tilde{A}) \geqslant \delta_{q,M}(A) \tfrac{M}{M + q_1 M_1}.$$

This implies the desired bound. □

Corollary 4.3.5 (Subprogression density increment)**.** *Let $M, M_1, q, q_1 \in \mathbb{N}$, and let $A_1, ..., A_r \subset [N]$ be non-empty sets. If $\delta_{qq_1, M_1}(A_i) \geqslant \delta_{q,M}(A_i) + \eta$ for some $i \in [r]$ and some $\eta > 0$, then*

$$\Delta(qq_1, M_1) \geqslant \Delta(q, M) - \tfrac{q_1 M_1}{M} r + \eta.$$

The following lemma allows us to pass to a subprogression whose common difference and length are sufficiently small to allow for an effective employment of Corollary 4.3.5.

Lemma 4.3.6. *Let q and M be positive integers with $M \leqslant 2 \lfloor N/q \rfloor$. For any $A \subset [N]$ there exists an arithmetic progression P of common difference q and length $|P| \in [\frac{1}{2}M, M]$ such that*

$$\frac{|A \cap P|}{|P|} \geqslant \frac{|A \cap [N]|}{N}.$$

Proof. Let us first give the argument for $q = 1$. We partition $[N]$ into the intervals

$$(0, \lceil M/2 \rceil] \cup (\lceil M/2 \rceil, 2\lceil M/2 \rceil] \cup \cdots \cup (m \lceil M/2 \rceil, N],$$

for some m with $N - m\lceil M/2 \rceil \leqslant \lceil M/2 \rceil$. If $N - m\lceil M/2 \rceil = \lceil M/2 \rceil$, then we obtain the result on applying the pigeonhole principle. So we may suppose that $N - m\lceil M/2 \rceil \leqslant \lfloor M/2 \rfloor$. The pigeonhole principle again gives the result on partitioning similarly, but with the final interval equal to $((m-1)\lceil M/2 \rceil, N]$.

We generalise to $q > 1$ by first partitioning $[N]$ into congruence classes mod q. Each such congruence class takes the form $-a + q \cdot [N_a]$ where $0 \leqslant a < q$ and $N_a = \lfloor (N + a)/q \rfloor$. Since $N_a \geqslant M/2$ we can use our previous argument to partition $[N_a]$ into intervals, each with length in $[\frac{1}{2}M, M]$. The result follows once again from the pigeonhole principle. □

4.3.3 Uniform translates

We have shown that a highly uniform set contains many Brauer configurations. In general, one cannot guarantee that one of the colour classes in a finite colouring of $[N]$ is uniform. However, we can use Gowers' density increment lemma to show that there exists a long arithmetic progression $q\cdot[M]\subset[N]$ such that each colour class is uniform on a translate of $q\cdot[M]$, and on the same translate its density is not diminished.

Lemma 4.3.7 (Uniform maximal translates)**.** *Given $0<\varepsilon\leqslant\frac{1}{2}$ and $d\geqslant 1$, let $\eta=\eta(d,\varepsilon)$ denote the constant (4.11) appearing in Gowers' density increment lemma. Suppose that*

$$N\geqslant\exp\exp(3r\eta^{-2}). \tag{4.13}$$

Then for any sets $A_1,\ldots,A_r\subset[N]$ there exists a homogeneous progression $q\cdot[M]\subset[N]$ with $M\geqslant N^{\exp(-3r\eta^{-2})}$ such that the following is true. For each $i\in[r]$, there exists a translate $a_i+q\cdot[M]$ on which A_i achieves its maximal translate density and on which A_i is ε-uniform of degree d.

Proof. We give an iterative procedure, at each stage of which we have positive integers q_n and M_n satisfying

$$M_n\geqslant(\eta^2/5r)^{1+\eta+\cdots+\eta^{n-1}}N^{\eta^n}\quad\text{and}\quad\Delta(q_n,M_n)\geqslant n\eta/2. \tag{4.14}$$

We initiate this on taking $q_0:=1$ and $M_0:=N$ (the common difference and length of $[N]$). Since $\Delta(q,M)\leqslant r$ this procedure must terminate at some $n\leqslant 2r\eta^{-1}$.

Suppose that we have iterated n times to give $q=q_n$ and $M=M_n$. If, for each A_i, there is a translate $a_i+q\cdot[M]$ on which A_i is ε-uniform and achieves its maximal translate density, then we terminate our procedure. Suppose then that we can find A_j which does not have this property. We now give the iteration step of our algorithm.

By the definition of maximal translate density, there exists $t\in\mathbb{Z}$ such that

$$|A_j\cap(t+q\cdot[M])|=\delta_{q,M}(A_j)M.$$

Let $A:=\{y\in[M]:t+qy\in A_j\}$. Since A_j is not ε-uniform of degree d on $t+q\cdot[M]$, we see that A is not ε-uniform of degree d on $[M]$. By Gowers' density increment lemma, we deduce the existence of a progression $P\subset[M]$ of length $|P|\geqslant\eta M^{\eta}$ such that

$$|A\cap P|\geqslant(\delta_{q,M}(A_j)+\eta)\,|P|.$$

We would like the length and common difference of P to be sufficiently small to allow for the effective employment of Corollary 4.3.5. Using (4.13) and (4.14) one can verify that $\eta M^{\eta} \geqslant 2r\eta^{-1}$, so that the integer $\lfloor \eta|P|/(2r)\rfloor$ is positive. Lemma 4.3.6 then gives a subprogression $x + q' \cdot [M'] \subset P$, of the same common difference as P, such that $\frac{1}{2}\lfloor \eta|P|/(2r)\rfloor \leqslant M' \leqslant \lfloor \eta|P|/(2r)\rfloor$ and for some x we have

$$\frac{|A \cap (x + q' \cdot [M'])|}{M'} \geqslant \frac{|A \cap P|}{|P|}.$$

Note that, since $P \subset [M]$ has common difference q' we have $q'|P| \leqslant M$ and so $q'M' \leqslant q'|P|\eta/(2r) \leqslant M\eta/(2r)$. Hence by Corollary 4.3.5 we obtain

$$\Delta(q'q, M') \geqslant \Delta(q, M) + \tfrac{1}{2}\eta.$$

Again using (4.14) and (4.13) one can check that $M' \geqslant (\eta^2/5r)M^{\eta}$, so we obtain (4.14) with $(q_{n+1}, M_{n+1}) := (q', M')$, and our iteration can continue. Taking this iteration through to completion gives the lemma. □

We are now in a position to derive our main theorem.

Proof of Theorem 4.1.1. Let $\eta = \eta(k-1, \varepsilon)$ be the quantity given by (4.11) in Gowers' density increment lemma, with ε to be determined, and suppose that $N \geqslant \exp\exp(4r\eta^{-2})$. Let M, q be the positive integers obtained by applying Lemma 4.3.7 to the partition $[N] = A_1 \cup \cdots \cup A_r$. By the pigeonhole principle, there exists $j \in [r]$ such that

$$|A_j \cap q \cdot [M/(2k-2)]| \geqslant \frac{1}{r}\left\lfloor \frac{M}{2(k-1)} \right\rfloor > \frac{M}{4(k-1)r}, \tag{4.15}$$

the latter following from the fact that

$$M \geqslant \exp\big(\exp(-3r\eta^{-2})\log N\big) \geqslant \exp\exp(r\eta^{-2}).$$

Let $t \in \mathbb{Z}$ be such that

$$|A_j \cap (t + q \cdot [M])| = \delta_{q,M}(A_j)M.$$

We now construct sets $A \subset [M]$ and $B \subset [M/(2k-2)]$ by taking

$$A := \{y \in [M] : t + qy \in A_j\} \quad \text{and} \quad B := \{d \in [M/(2k-2)] : qd \in A_j\}.$$

Our goal is to show that $\Lambda(1_A; 1_B) > 0$. If this is the case, then there exists an arithmetic progression of length k in A whose common difference lies in B. We can then infer from our construction of A and B the existence of a $(k+1)$-point Brauer configuration in A_j.

Let α and β denote the respective densities of A and B in $[M]$. From the bound (4.15) it follows that $\alpha, \beta > (4r(k-1))^{-1}$. Combining this with Lemma 4.3.1 gives

$$\Lambda(\alpha 1_{[M]}; 1_B) \geqslant \tfrac{1}{2}\alpha^k \beta M^2 > \frac{M^2}{2^{2k+3}(k-1)^{k+1} r^{k+1}}.$$

Recall that the conclusion of Lemma 4.3.7 guarantees that A is ε-uniform of degree $(k-1)$ (as a subset of $[M]$). Applying Corollary 4.3.3 and the upper bound in (4.10) gives

$$|\Lambda(1_A; 1_B) - \Lambda(\alpha 1_{[M]}; 1_B)| \leqslant kM^2\varepsilon.$$

We therefore obtain $\Lambda(1_A; 1_B) > 0$, and hence the theorem, on taking

$$\varepsilon^{-1} := k2^{2k+3}(k-1)^{k+1} r^{k+1}. \tag{4.16}$$

Since we are assuming that $N \geqslant \exp\exp(4r\eta^{-2})$, this choice of ε gives a double exponential bound in $r^{O_k(1)}$.

Finally, we show that the more precise bound (4.1) suffices. The inequality $\exp\exp(x) \leqslant 2^{2^{2x}}$ is valid for all $x \geqslant 1$, so that

$$\exp\exp(4r\eta^{-2}) \leqslant 2^{2^{8r\eta^{-2}}}.$$

For our choice (4.16) of ε, we have

$$\eta^{-1} \leqslant \left(16(k+1)k(k-1)^{k+1} 2^{2k+3} r^{k+1}\right)^{2^{k+2^{k+9}}}$$

One may check that

$$16(k+1)k(k-1)^{k+1} \leqslant 2^{k^2+4},$$

so that, on using $r \geqslant 2$, we have

$$8r\eta^{-2} \leqslant 8r\left(2^{k^2+2k+7} r^{k+1}\right)^{2^{1+k+2^{k+9}}} \leqslant r^{(k+3)^2 2^{1+k+2^{k+9}}} \leqslant r^{2^{2^{k+10}}},$$

as required. □

4.4 An improved bound for four-point configurations

In this section we focus on the four-point Brauer configuration (4.2), improving our bound on the Ramsey number to $\exp\exp(r^{1+o(1)})$. Instead of finessing the quantitative aspect of Lemma 4.3.7, we opt to mimic the sparsity–expansion dichotomy of §4.2. This requires a higher order analogue of the difference set $A - A$, namely

$$\mathrm{Step}_3(A) := \{d : A \cap (A - d) \cap (A - 2d) \neq \emptyset\}.$$

Lemma 4.4.1 (Sparsity–expansion dichotomy)**.** *There exists an absolute constant C such that for $N \geqslant \exp\exp(Cr\log^2 r)$ the following holds. For any r-colouring $A_1 \cup \cdots \cup A_r = [N]$ there exist positive integers q and M with $qM \leqslant N$ such that for any $i \in [r]$ we have one of the two following possibilities.*

- *(Sparsity).*
$$|A_i \cap q \cdot [3M]| < \tfrac{1}{r} M; \tag{4.17}$$
- *(Expansion).*
$$|\mathrm{Step}_3(A_i) \cap q \cdot [M]| > \left(1 - \tfrac{1}{r}\right) M. \tag{4.18}$$

Before proving this, let us first use it to obtain a bound on the Ramsey number of four-point Brauer configurations.

Proof of Theorem 4.1.2. Let q and M denote the numbers provided by the dichotomy. By the pigeonhole principle, there exists a colour class A_i satisfying

$$|A_i \cap q \cdot [M]| \geqslant \tfrac{1}{r} M.$$

So A_i is not sparse in the sense of (4.17). It follows that A_i must instead satisfy the expansion property (4.18). By inclusion–exclusion

$$|A_i \cap \mathrm{Step}_3(A_i) \cap q \cdot [M]| \geqslant |A_i \cap q \cdot [M]| + |\mathrm{Step}_3(A_i) \cap q \cdot [M]| - M > 0.$$

In particular $A_i \cap \mathrm{Step}_3(A_i)$ contains a non-zero element. □

Proof of Lemma 4.4.1. If $r = 1$, then $[N] = A_1$ and so (4.18) holds with $q = 1$ and $M = N$ for all $N \geqslant \exp\exp(0) > 2$. We may therefore henceforth assume that $r \geqslant 2$. Set $q_0 := 1$ and $N_0 := N$. We proceed by an iterative procedure, at each stage of which we have positive integers q_n, N_n and $M_i = M_i^{(n)}$ $(1 \leqslant i \leqslant r)$ such that for $n \geqslant 1$ we have:

(i) $N_n \geqslant \exp\left(-r^{O(1)}\right) N_{n-1}^{r^{-O(1)}}$;

(ii) $M_i \in [\frac{1}{2}N_n, N_n]$ for all $1 \leqslant i \leqslant r$;

(iii) $\delta_{q_n, M_i^{(n)}}(A_i) \geqslant \delta_{q_{n-1}, M_i^{(n-1)}}(A_i)$ for all $1 \leqslant i \leqslant r$;

(iv) $\delta_{q_n, M_i^{(n)}}(A_i) \geqslant (1+c)\,\delta_{q_{n-1}, M_i^{(n-1)}}(A_i)$ for some i with $\delta_{q_{n-1}, M_i^{(n-1)}}(A_i) \geqslant \frac{1}{7r}$. Here $c = \Omega(1)$ is an absolute constant.

At stage n of the iteration we classify each colour class A_i according to which of the following hold.

- (Sparse colours). These are the colours A_i for which
$$\delta_{q_n, M_i}(A_i) < \tfrac{1}{7r}.$$
- (Dense expanding colours). These are the dense colours $\delta_{q_n, M_i}(A_i) \geqslant \frac{1}{7r}$ for which we have the additional expansion estimate
$$|\mathrm{Step}_3(A_i) \cap q_n \cdot [N_n/6]| > \left(1 - \tfrac{1}{r}\right) \lfloor N_n/6 \rfloor . \tag{4.19}$$
- (Dense non-expanding colours). The class A_i is *dense non-expanding* if it is neither sparse nor dense expanding.

If there are no dense non-expanding colours, then we terminate our procedure. If $N_n < 100$, then we also terminate our procedure. Let us therefore suppose that $N_n \geqslant 100$ and there exists a dense non-expanding colour class A_i. Our aim is to show how, under these circumstances, the iteration may continue.

By the definition of maximal translate density, there exists a such that
$$|A_i \cap (a + q_n \cdot [M_i])| M_i^{-1} = \delta_{q_n, M_i}(A_i) \geqslant \tfrac{1}{7r}. \tag{4.20}$$

Writing $M := M_i$, we define dense subsets $A, B \subset [M]$ by taking
$$A := \{y \in [M] : a + q_n y \in A_i\} \quad \text{and} \quad B := \{d \in [N_n/6] : q_n d \notin \mathrm{Step}_3(A_i)\}. \tag{4.21}$$

We recall that $M/3 \geqslant N_n/6 \geqslant M/6$.

Letting α denote the density of A in $[M]$, we see that α is equal to the left-hand side of (4.20). Our assumption that A_i is dense non-expanding and $N_n \geqslant 100$ together imply that B has size $\gg r^{-1}M$ and that

$$\sum_{x,d} 1_A(x)1_A(x+d)1_A(x+2d)1_B(d) = 0.$$

Using the notation (4.8) and (the proof of) Lemma 4.3.1 we have

$$\left|\Lambda(1_A; 1_B) - \Lambda(\alpha 1_{[M]}; 1_B)\right| \gg \alpha^3 M|B| \gg r^{-4}M^2.$$

From hereon, we assume that the reader is familiar with the notation and terminology of Green and Tao [GT09]. Applying [GT09, Theorem 5.6] in conjunction with Corollary 4.3.3 we obtain a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ of complexity and resolution $\ll r^{O(1)}$ such that the function $f := \mathbb{E}(1_A \mid \mathcal{B}_2)$ satisfies

$$\left|\Lambda(f; 1_B) - \Lambda(\alpha 1_{[M]}; 1_B)\right| \gg \alpha^3 M|B|.$$

Define the $\mathcal{B}_2$-measurable set

$$\Omega := \{x \in [M] : f(x) \geqslant (1+c)\alpha\},$$

where $c > 0$ is small enough to make the following argument valid.[2]

For functions $f_1, f_2, f_3 : [M] \to \mathbb{R}$ we have the bound

$$|\Lambda(f_1, f_2, f_3; 1_B)| \leqslant M|B| \|f_i\|_{L^1([M])} \prod_{j \neq i} \|f_j\|_\infty . \tag{4.22}$$

Invoking the telescoping identity we used to prove Corollary 4.3.3 gives us the bound $|\Lambda(f; 1_B) - \Lambda(f 1_{\Omega^c}; 1_B)| \ll |\Omega||B|$, so that

$$|\Omega||B| + \left|\Lambda(f 1_{\Omega^c}; 1_B) - \Lambda(\alpha 1_{[M]}; 1_B)\right| \gg \alpha^3 M|B|.$$

Another application of the telescoping identity in conjunction with (4.22) gives

$$\begin{aligned}\left|\Lambda(f 1_{\Omega^c}; 1_B) - \Lambda(\alpha 1_{[M]}; 1_B)\right| &\ll \alpha^2 M|B| \left\|f 1_{\Omega^c} - \alpha 1_{[M]}\right\|_{L^1([M])} \\ &\ll \alpha^2 M|B| \left\|f - \alpha 1_{[M]}\right\|_{L^1([M])} + |B||\Omega|,\end{aligned}$$

[2] Specifically, c is chosen to be sufficiently small relative to all the implicit constants appearing in the inequalities preceding (4.23).

so that

$$|\Omega| + \alpha^2 M \left\| f - \alpha 1_{[M]} \right\|_{L^1([M])} \gg \alpha^3 M.$$

Since $f - \alpha 1_{[M]}$ has mean zero, its L^1-norm is equal to twice the mean of its positive part. The function $\left(f - \alpha 1_{[M]}\right)_+$ can only exceed $c\alpha$ on Ω, so taking c small enough gives

$$|\Omega| \gg \alpha^3 M \gg r^{-3} M. \tag{4.23}$$

As $\mathcal{B}_2$ has complexity and resolution $r^{O(1)}$ it contains at most $\exp(r^{O(1)})$ atoms. By [GT09, Proposition 6.2] each such atom can be partitioned into a further

$$\exp(r^{O(1)}) M^{1 - r^{-O(1)}} \tag{4.24}$$

disjoint arithmetic progressions. Hence Ω itself can be partitioned into arithmetic progressions, the number of which is at most (4.24). Combining this with (4.23) and [GT09, Lemma 6.1], we see that there exists an arithmetic progression P of length at least

$$\exp\left(-r^{O(1)}\right) M^{1/r^{O(1)}}$$

on which A has density at least $(1 + \frac{c}{2})\alpha$. By partitioning P into two pieces and applying the pigeon-hole principle, we may further assume that $|P|q \leqslant M$, where q is the common difference of P.

Writing q_{n+1} and $N_{n+1} = M_i^{(n+1)}$ for the common difference and length of P, we see that (i) and (iv) are satisfied. For all $j \in [r] \setminus \{i\}$ we have

$$q_{n+1} N_{n+1} \leqslant M = M_i^{(n)} \leqslant N_n \leqslant 2M_j^{(n)}.$$

Hence we may apply Lemma 4.3.6 to each colour class A_j with $j \neq i$ to obtain a progression P_j of common difference q_n and length $M_j^{(n+1)}$ such that (ii) and (iii) hold. It follows that our iteration may continue.

Our iterative procedure must terminate at stage n for some $n \ll r^2$. To see this, note that the sum of the maximal translate densities $\delta_{q_n, M_i}(A_i)$ is at most r, and this quantity increases by at least $\Omega(1/r)$ at each iteration. Our next task is to improve this upper bound on the number of iterations to $n \ll r \log r$.

Let A_i denote the colour class for which the density increment (iv) occurs most often. By the pigeonhole principle this happens on at least n/r occasions. If the density of A_i increments at least c^{-1} times, then its density doubles. After a further $\frac{1}{2}c^{-1}$ increments the density of A_i quadruples. The density of A_i has therefore increased by a factor of 2^m if the number of iterations is at least

$$\lceil c^{-1} \rceil + \lceil \tfrac{1}{2} c^{-1} \rceil + \cdots + \lceil \tfrac{1}{2^{m-1}} c^{-1} \rceil \leqslant m + 2c^{-1}. \tag{4.25}$$

The first time A_i increments its initial density is at least $1/(7r)$, so if the number of increments experienced by A_i is at least (4.25) then its final density is at least $2^m/(7r)$. If $n/r > 2c^{-1} + \lceil \log_2(7r) \rceil$ then we obtain a density exceeding 1, a contradiction. It follows that the total number of iterations n satisfies $n = O(r \log_2 r)$.

Having shown that our iteration must terminate in $n = O(r \log r)$ steps, let us now ensure that termination results from a lack of dense non-expanding colours. This follows if we can ensure that $N_n \geqslant 100$. Applying the lower bound (i) iteratively we obtain

$$N_n \geqslant \exp\left(-r^{O(1)}\right) N^{r^{-O(n)}}.$$

Using the fact that $n \ll r \log r$, the right-hand side above is at least 100 provided it is not the case that $N \leqslant \exp\exp\left(O(r \log^2 r)\right)$. Given this assumption, we obtain the conclusion of Lemma 4.4.1 on taking $M := \lfloor N_n/6 \rfloor$. □

4.5 Lefmann quadrics

In this section we show how our results can be used to obtain bounds on the Ramsey numbers for quadric equations of the form

$$\sum_{i=1}^{s} a_i x_i^2 = 0. \tag{4.26}$$

Lefmann [Lef91, Fact 2.8] established the following sufficient condition for equations of the above form to be partition regular.

Lefmann's criterion. *Suppose that $a_1, \dots, a_s \in \mathbb{Z} \setminus \{0\}$ satisfy the following two properties:*

(i) *there exists a non-empty set $I \subset [s]$ such that $\sum_{i \in I} a_i = 0$;*

(ii) *there exists a positive integer* λ *such that the system*

$$\lambda^2 \sum_{i \notin I} a_i + \sum_{i \in I} a_i u_i^2 = \sum_{i \in I} a_i u_i = 0.$$

has a solution in integers $(u_i)_{i \in I}$.

Then in any finite partition of the positive integers $\mathbb{N} = A_1 \cup \cdots \cup A_r$, *there exists* $i \in [r]$ *and* $x_1, ..., x_s \in A_i$ *satisfying (4.26).*

Lefmann observed that assumption (i) is necessary for the conclusion of the above theorem to hold. Lefmann then showed that if (i) and (ii) both hold, then (4.26) has a solution over any set of the form $\{x, x+d, ..., x+(k-1)d, \lambda d\}$, where $k = 1 + 2\max_{i \in I} |u_i|$. We therefore obtain our quantitative version of Lefmann's result (Theorem 4.1.3) from the following analogue of Theorem 4.1.1.

Theorem 4.5.1 (Ramsey bound for generalised Brauer configurations)**.** *For positive integers* k, λ *there exists an absolute constant* $C = C(k, \lambda)$ *such that for any* $r \geqslant 2$ *and* $N \geqslant \exp\exp(r^C)$, *if* $[N]$ *is* r*-coloured then there exists a monochromatic configuration of the form* $\{x, x+d, ..., x+(k-1)d, \lambda d\}$.

Proof. This is essentially the same as the proof of Theorem 4.1.1. $\square$

Bibliography

[Cha20] J. Chapman, *Partition regularity and multiplicatively syndetic sets*, Acta Arith. **196** (2020), 109–138.

[CS17] K. Cwalina and T. Schoen, *Tight bounds on additive Ramsey-type numbers*, J. London Math. Soc., **96** (2017), 601–620.

[Gow98] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. **8** (1998), 529–551.

[Gow01] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465–588.

[GW10] W. T. Gowers and J. Wolf, *The true complexity of a system of linear equations*, Proc. Lond. Math. Soc. **100** (2010), 155–176.

[GRS90] R. Graham, B. Rothschild and J. H. Spencer, *Ramsey theory.* Second edition. Wiley-Interscience Series in Discrete Mathematics and Optimization. 1990.

[GT09] B. Green and T. Tao, *New bounds for Szemerédi's theorem, II: A new bound for* $r_4(N)$, Analytic number theory, 180–204, Cambridge Univ. Press, Cambridge, 2009.

[GT10] B. Green and T. Tao,*Linear equations in primes*, Ann. of Math. **171** (2010), 1753–1850.

[Lef91] H. Lefmann, *On partition regular systems of equations*, J. Combin. Theory Ser. A **58** (1991), 35–53.

[Rot53] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.

[San08] T. Sanders, *Additive structures in sumsets*, Math. Proc. Cambridge Philos. Soc. **144** (2008), 289–316.

[San20] T. Sanders, *Bootstrapping partition regularity of linear systems*, Proc. Edinb. Math. Soc. (2) **63** (2020), no. 3, 630–653.

[She88] S. Shelah, *Primitive recursive bounds for van der Waerden numbers*, J. Amer. Math. Soc. **1** (1988), 683–697.

[Shk10] I. D. Shkredov, *Fourier analysis in combinatorial number theory*, Uspekhi Mat. Nauk, **65** (2010), 127–184.

[TV06] T. Tao and V. Vu, *Additive combinatorics*, volume 105 of Cambridge Studies in Advanced Mathematics. Cambridge University Press (2006).

Chapter 5

Conclusions

In this thesis, we have investigated connections between multiplicatively syndetic sets and partition regularity, established new partition and density regularity results for suitably non-singular systems of diagonal polynomial equations, and obtained new quantitative bounds for Ramsey-Rado numbers of non-translation invariant systems.

Our work in Chapter 2 has demonstrated that, when establishing partition regularity, one need only consider finite colourings for which each colour class is multiplicatively syndetic. Further investigation into precisely which colourings are necessary, and into additional properties of multiplicatively sydnetic sets, may be beneficial in the study of partition regularity for systems of non-linear equations.

In Chapter 3, we obtained new partition and density regularity results for sufficiently non-singular systems of kth power equations. Future work on the subject could investigate applications of the analytic methods developed in [CLP21] and Chapter 3 to the study of partition and density regularity for systems of polynomial equations of the form $P_1(x_1) + P_2(x_2) + \cdots + P_s(x_s) = 0$, where $P_1, \ldots, P_s$ are non-zero integer polynomials. In the case where every P_i takes the form $P_i(x_i) = a_i x_i^m$ with $m \in \{1, 2\}$, such an investigation has recently been initiated by Prendiville [Pre20].

Theorem 4.1.1 in Chapter 4 provides a new quantitatively effective version of Brauer's theorem. The shape of our final bound is a direct consequence of Gowers' bound in "Gowers local inverse theorem" [Gow01, Theorem 18.1]. Future research should investigate improvements for the bounds in this local inverse theorem, which, in view of our work and the work of Sanders [San20], would lead to improved upper bounds for Rado-Ramsey numbers of partition regular linear systems.

Bibliography

[CLP21] S. Chow, S. Lindqvist, and S. Prendiville, *Rado's criterion over squares and higher powers*, J. Eur. Math. Soc. (JEMS) **23** (2021), no. 6, 1925–1997.

[Gow01] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465–588.

[Pre20] S. Prendiville, *Counting monochromatic solutions to diagonal Diophantine equations*, preprint arXiv:2003.10161v1 (2020).

[San20] T. Sanders, *Bootstrapping partition regularity of linear systems*, Proc. Edinb. Math. Soc. (2) **63** (2020), no. 3, 630–653.

Appendix A

Obtaining a tower bound

In this appendix to Chapter 2, we show how Theorem 2.4.2 implies Theorem 2.1.4. By Theorem 2.4.2, there is a positive constant $\tilde{C} > 0$ such that

$$\mathrm{B}(r+1) \leqslant 2^{\mathrm{B}(r)^{\tilde{C}\log(r+1)}}$$

holds for all $r \in \mathbb{N}$. Given $n \in \mathbb{N}$ and $a_1, \ldots, a_n \in [2, \infty)$, define the tower function

$$T_n(a_1, a_2, \ldots, a_n) := a_1^{a_2^{\cdot^{\cdot^{\cdot^{a_n}}}}}.$$

Let $K \geqslant 1$ be a large positive constant, to be chosen later. We can now introduce the auxiliary function $F : \mathbb{N} \to \mathbb{R}$ given by

$$F(r) := T_{r+1}(2, 2, \ldots, 2, Kr^2).$$

Thus, we have

$$F(1) = 2^K, \quad F(2) = 2^{2^{4K}}, \quad F(3) = 2^{2^{2^{9K}}}, \quad F(4) = 2^{2^{2^{2^{16K}}}}, \quad \ldots.$$

Our goal is to show that, if K is sufficiently large relative to $\tilde{C}$, then

$$\mathrm{B}(r) \leqslant F(r)$$

holds for all $r \in \mathbb{N}$. To demonstrate why this is enough to prove Theorem 2.1.4, we first need to investigate the growth of tower functions.

Lemma A.0.1 (Towers dominate cubes)**.** *For all $r \geqslant 5$ we have*

$$r^3 \leqslant \mathrm{tow}(r-1). \tag{A.1}$$

Proof. We first observe that (A.1) holds for $r = 5$. Suppose then that $r > 5$ and assume the induction hypothesis that

$$(r-1)^3 \leqslant \mathrm{tow}(r-2).$$

Note that since $r > 5$, we have

$$\frac{r^3}{(r-1)^3} = \left(1 + \frac{1}{r-1}\right)^3 < \frac{125}{64} < 2.$$

This gives

$$r^3 \leqslant 2(r-1)^3 \leqslant 2 \cdot \mathrm{tow}(r-2).$$

Using the elementary fact that $2k \leqslant 2^k$ holds for all $k \in \mathbb{N}$, we deduce

$$r^3 \leqslant 2^{\mathrm{tow}(r-2)} = \mathrm{tow}(r-1).$$

The desired result now follows by induction. □

This lemma enables us to bound F above by a tower function.

Corollary A.0.2 (Tower bound for F)**.** *For all $r \in \mathbb{N}$,*

$$F(r) \leqslant \mathrm{tow}\left((1 + o(1))r\right). \tag{A.2}$$

Proof. Recall that F is an exponential tower of height $r + 1$, with Kr^2 as the 'top' term, and with the remaining terms in the tower equal to 2. By the previous lemma, when r is sufficiently large, we have

$$Kr^2 \leqslant \mathrm{tow}\left(\lceil K^{1/3} r^{2/3} \rceil\right).$$

By adding the heights of the towers, we deduce that

$$F(r) \leqslant \mathrm{tow}\left(r + \lceil K^{1/3} r^{2/3} \rceil\right)$$

holds for all sufficiently large r. This gives (A.2). □

We require the following elementary result concerning manipulations of exponentials.

Lemma A.0.3. *For all $a, b, k \geqslant 2$,*

$$a^b k \leqslant a^{b+k} \leqslant a^{bk}. \tag{A.3}$$

This gives the following bound on the growth of F.

Corollary A.0.4 (Tower growth of F)**.** *For all $r \in \mathbb{N}$,*

$$F(r)^r \leqslant \log_2 F(r+1).$$

Proof. Since $K \geqslant 1$, the case $r = 1$ can be verified by inspection. Suppose then that $r \geqslant 2$. By iteratively applying Lemma A.0.3, we deduce that

$$\begin{aligned}
\log_2 F(r+1) &= T_{r+1}(2, 2, \ldots, 2, 2, K(r+1)^2) \\
&\geqslant T_r(2, 2, \ldots, 2, 2^{Kr^2} \cdot r) \\
&\geqslant T_{r-1}(2, 2, \ldots, 2^{2^{Kr^2}} \cdot r) \\
&\vdots \\
&\geqslant T_2(2, T_r(2, 2, \ldots, 2, Kr^2) \cdot r). \\
&= F(r)^r.
\end{aligned}$$

□

We can now prove Theorem 2.1.4.

Proof of Theorem 2.1.4. By Corollary A.0.2, it is sufficient to show that

$$\mathrm{B}(r) \leqslant F(r) \tag{A.4}$$

holds for all $r \in \mathbb{N}$. Since $\tilde{C}\log(n+1) = o(n)$, we can choose $n_0 \in \mathbb{N}$ with $n_0 \geqslant 5$ such that

$$\tilde{C}\log(n+1) \leqslant n$$

holds for $n \geqslant n_0$. By taking K sufficiently large, we can assume that (A.4) holds for $r \leqslant n_0$. Suppose then that $r > n_0$ and assume the induction hypothesis

$$\mathrm{B}(r-1) \leqslant F(r-1).$$

By Theorem 2.4.2 and the fact that $r > n_0$, we have

$$\log_2 \mathrm{B}(r) \leqslant \mathrm{B}(r-1)^{\tilde{C}\log r} \leqslant \mathrm{B}(r-1)^{r-1}.$$

Now by the induction hypothesis and Corollary A.0.4, we conclude that

$$\log_2 \mathrm{B}(r) \leqslant F(r-1)^{r-1} \leqslant \log_2 F(r).$$

This establishes the induction step and completes the proof. □

Appendix B

Definitions from quadratic Fourier analysis

In this section we explain the various terms from [GT09] which appear in §2.4.3. We begin with the notion of a factor.

Definition (Factors)**.** Let X be a non-empty set. Let $\chi : X \to F$ be a finite colouring of X, where F is some non-empty finite set. The *factor* (or *σ-algebra*) *in* X *induced by* χ is the collection of sets $\mathcal{B}_\chi := \{\chi^{-1}(I) : I \subseteq F\}$. The *atoms* of $\mathcal{B}_\chi$ are the non-empty colour classes $\chi^{-1}(\{i\}) \neq \emptyset$. In general, a collection of sets $\mathcal{B}$ is called a *factor in* X if $\mathcal{B} = \mathcal{B}_{\tilde{\chi}}$ for some finite colouring $\tilde{\chi}$ of X.

Let $\mathcal{B} = \mathcal{B}_\chi$ and $\mathcal{B}' = \mathcal{B}_{\tilde{\chi}}$ be two factors in a set X, which are induced by the finite colourings $\chi : X \to F$ and $\tilde{\chi} : X \to F'$ respectively. We say that $\mathcal{B}'$ *extends* $\mathcal{B}$ if $\mathcal{B} \subseteq \mathcal{B}'$. Equivalently, this means that every χ'-monochromatic subset of X is χ-monochromatic.

The factor $\mathcal{B} \vee \mathcal{B}'$ in X is defined to be the smallest factor in X which extends both $\mathcal{B}$ and $\mathcal{B}'$. The atoms of $\mathcal{B} \vee \mathcal{B}'$ are all of the non-empty sets of the form $A \cap A'$, where A is an atom of $\mathcal{B}$ and A' is an atom of $\mathcal{B}'$. Hence, $\mathcal{B} \vee \mathcal{B}'$ can equivalently be defined by $\mathcal{B} \vee \mathcal{B}' := \mathcal{B}_{\chi \times \chi'}$, where $(\chi \times \chi') : X \to F \times F'$ is the *product colouring* defined by

$$(\chi \times \chi')(x) := (\chi(x), \chi'(x)).$$

Now suppose that X is a non-empty finite set, and let $\mathcal{B}$ be a factor in X. Let $f : X \to \mathbb{C}$ be a function defined on X. The *conditional expectation* $\mathbb{E}(f|\mathcal{B}) : X \to \mathbb{C}$

of f with respect to $\mathcal{B}$ is given by

$$\mathbb{E}(f|\mathcal{B})(x) := \mathbb{E}_{\mathcal{B}(x)} f = \frac{1}{|\mathcal{B}(x)|} \sum_{y \in \mathcal{B}(x)} f(y),$$

where $\mathcal{B}(x)$ denotes the unique atom of $\mathcal{B}$ which contains x. The function $g := \mathbb{E}(f|\mathcal{B})$ can be thought of as an approximation to the function f, which is obtained by taking $g(x)$ to be the average value that f takes on the atom of $\mathcal{B}$ containing x. One reason for working with g rather than f is that g has the useful property that it is $\mathcal{B}$-*measurable*, meaning that g is constant on the atoms of $\mathcal{B}$. In particular, all sets of the form $g^{-1}([\eta, \infty)) = \{x \in X : g(x) \geqslant \eta\}$ are members of $\mathcal{B}$.

Using these properties of g provides us with a way of obtaining a density increment for a function $f : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ supported on $[N]$. We begin with a 'suitably structured' factor $\tilde{\mathcal{B}}$. Define $\mathcal{B}_{triv}$ to be the factor in $\mathbb{Z}/p\mathbb{Z}$ induced by the 2-colouring of $\mathbb{Z}/p\mathbb{Z}$ with colour classes $[N]$ and $(\mathbb{Z}/p\mathbb{Z}) \setminus [N]$. We then take $\mathcal{B} = \tilde{\mathcal{B}} \vee \mathcal{B}_{triv}$ and let $g = \mathbb{E}(f|\mathcal{B})$. We seek a density increment of the form

$$\mathbb{E}_B(f) \geqslant \mathbb{E}_{[N]}(f) + \eta, \tag{B.1}$$

for some $\eta > 0$ and some 'large' atom $B \in \mathcal{B}$. Observe that replacing $\tilde{\mathcal{B}}$ with $\mathcal{B}$ ensures that every atom of $\mathcal{B}$ is either contained in $[N]$ or disjoint from $[N]$. Hence, as f is supported on $[N]$, we see that g is also supported on $[N]$. Moreover, we have $\mathbb{E}_B(f) = \mathbb{E}_B(g)$ for any $B \in \mathcal{B}$. It is therefore sufficient to establish (B.1) with g in place of f.

Before we define the 'structured' factors which are used to approximate f, we first need to introduce *phase functions*.

Definition (Phase functions)**.** Let $X \subseteq \mathbb{Z}/p\mathbb{Z}$ be a non-empty finite set. A *phase function* ϕ on X is a function $\phi : X \to \mathbb{R}/\mathbb{Z}$. A phase function ϕ is called *irrational* if[1] $\phi(x) \notin \mathbb{Q}$ for all $x \in X$. We say that ϕ is *globally linear* if

$$\sum_{\varepsilon_1, \varepsilon_2 \in \{0,1\}} (-1)^{\varepsilon_1 + \varepsilon_2} \phi(x + \varepsilon_1 h_1 + \varepsilon_2 h_2) = 0$$

holds for all $x, h_1, h_2 \in X$. We say that ϕ is *locally quadratic* if

$$\sum_{\varepsilon_1, \varepsilon_2, \varepsilon_3 \in \{0,1\}} (-1)^{\varepsilon_1 + \varepsilon_2 + \varepsilon_3} \phi(x + \varepsilon_1 h_1 + \varepsilon_2 h_2 + \varepsilon_3 h_3) = 0$$

[1] Here we have identified $\mathbb{R}/\mathbb{Z}$ with the interval $[0, 1)$ by reduction modulo 1.

holds whenever $x, h_1, h_2, h_3 \in X$ are such that $x + \varepsilon_1 h_1 + \varepsilon_2 h_2 + \varepsilon_3 h_3 \in X$ for all $\varepsilon_1, \varepsilon_2, \varepsilon_3 \in \{0, 1\}$.

We can now use phase functions to define finite colourings (and hence factors) of a set $X \subseteq \mathbb{Z}/p\mathbb{Z}$. Given an irrational phase function $\phi : X \to \mathbb{R}/\mathbb{Z}$, define a K-colouring $\chi_{\phi,K} : X \to \{0, 1, ..., K-1\}$ by defining $\chi_{\phi,K}(x)$ to be the unique element of $\{0, 1, ..., K-1\}$ satisfying $\|\phi(x) - \chi_{\phi,K}(x)/K\|_{\mathbb{R}/\mathbb{Z}} < 1/2K$. Note that the irrationality of ϕ guarantees that such a $\chi_{\phi,K}(x)$ exists. For brevity, we write $\mathcal{B}_{\phi,K}$ to denote the factor in X induced by $\chi_{\phi,K}$.

Definition (Linear factors)**.** A *linear factor of complexity at most d and resolution K* is any factor $\mathcal{B}$ in $\mathbb{Z}/p\mathbb{Z}$ of the form $\mathcal{B} = \mathcal{B}_{\phi_1,K} \vee \cdots \vee \mathcal{B}_{\phi_{d'},K}$, for some $d' \leqslant d$, where each $\phi_i : \mathbb{Z}/p\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is a globally linear phase function.

Definition (Quadratic factors)**.** A *quadratic factor of complexity at most (d_1, d_2) and resolution K* is any pair of factors $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ with the following properties.

- $\mathcal{B}_1$ is a linear factor of complexity at most d_1 and resolution K;
- $\mathcal{B}_2$ is an extension of $\mathcal{B}_1$;
- for any atom B of the factor $\mathcal{B}_1$, the restriction of $\mathcal{B}_2$ to B can be written as $\mathcal{B}_2|_B = \mathcal{B}_{\phi_1,K} \vee \cdots \vee \mathcal{B}_{\phi_{d'},K}$, where each $\phi_i : B \to \mathbb{R}/\mathbb{Z}$ is a locally quadratic phase function defined on B, and $0 \leqslant d' \leqslant d_2$.