FINITE FIELD MODELS IN ARITHMETIC COMBINATORICS – TEN YEARS ON

J. WOLF

Abstract. It has been close to ten years since the publication of Green's influential survey *Finite field models in additive combinatorics* [28], in which the author championed the use of high-dimensional vector spaces over finite fields as a toy model for tackling additive problems concerning the integers. The path laid out by Green has proven to be a very successful one to follow. In the present article we survey the highlights of the past decade and outline the challenges for the years to come.

Contents

1. Introduction

Green's 2004 survey *Finite field models in additive combinatorics* [28] must be counted amongst the most influential and widely cited papers in arithmetic combinatorics. By spotlighting an accessible toy model for a number of notoriously difficult problems in additive number theory, it inspired countless proof ideas in subsequent years and served as introductory reading material for many a graduate student.

Its main tenet was the idea that many of the problems traditionally of interest in additive number theory can be rephrased in the context of high-dimensional vector spaces over finite fields of fixed characteristic. For example, instead of counting the number of 3-term arithmetic progressions in a (finite) subset A of the integers, we may choose to count the

number of such progressions inside a subset $A \subseteq \mathbb{F}_3^n$, where the dimension n is to be thought of as very large. The advantage of working in the latter additive group instead of the former is that it has a plentiful supply of non-trivial additively closed subsets, namely the vector subspaces of $\mathbb{F}_3^n$. These closely resemble the original space itself, making it possible to run arguments locally and facilitating especially those relying on iteration. In addition, we have the notion of orthogonality and linear independence at our disposal, which instantly adds a host of techniques from linear algebra to our toolbox.

Working in $\mathbb{F}_p^n$ for some fixed small prime p thus often simplifies the problem at hand, but it does so in such a way that the principal features of the problem are preserved, meaning that solving the toy problem constitutes a significant step towards solving the problem in the integers. Because of their exact algebraic nature, arguments for the toy problem are often more accessible, highlighting the core idea which in the integers tends to be obscured by technical details. Moreover, in certain cases there is a (by now well-established if often technically challenging) procedure for transferring a proof from the model setting to the integers. Although these remarks may appear senseless if not mystifying at this point, we hope that their meaning will become more transparent to the reader through the examples provided in Chapter 3.

The finite field model $\mathbb{F}_p^n$, and in particular the case of characteristic $p = 2$, is also of unparalleled importance in theoretical computer science, by virtue of representing the set of strings of 0s and 1s of length n under the operation of addition modulo 2. It is nigh on impossible to do justice to the multitude of applications which the ideas sketched out here have had in theoretical computer science, covering areas as diverse as property testing and decoding of Reed-Muller codes [1, 2, 55], probabilistically checkable proofs [56], communication complexity [5, 74], pseudorandom generators [10] and hardness amplification [73]. Nor can we hope to pay adequate tribute to the contributions made to arithmetic combinatorics by that community. We shall therefore confine ourselves to the occasional reference in the context of a particular result or technique, in the hope that the interested reader will delve further into the literature.

Having explained what we mean by the 'finite field model', we should point out that it is something of a misnomer, which we hardly stand any chance of correcting as the term is widely used in the field. But for the sake of clarity we feel compelled to discuss it early on. Indeed, many additive problems concerning finite subsets of the integers are traditionally tackled in the group $\mathbb{Z}/p\mathbb{Z}$, with p a prime chosen large enough so that the relevant segment of the integers can be embedded in it without disturbing the additive structure in question by its cyclic nature. Of course, the ring $\mathbb{Z}/p\mathbb{Z}$ is canonically isomorphic to the finite field with p elements, but from our point of view it behaves more like the integers in the sense that it has no non-trivial additive subgroups. So we would argue that the group $\mathbb{F}_p^n$ should more accurately be called a 'vector space model', as the vector space structure is really its determining feature.[1]

[1]As it stands, the present article is almost entirely disjoint from one that recently appeared under a confusingly similar title in this very same journal [64]. There the focus was on sum-product-type problems and incidence geometry over finite fields. These are central areas of arithmetic combinatorics which we shall omit from this survey as they are very different in flavour: the fields involved tend to be large without

There are a number of reasons for writing a follow-up to [28] at this point in time. First and foremost, significant progress has been made on all of the central problems in arithmetic combinatorics since 2004. This has resulted in a better qualitative and quantitative understanding of additive structures in dense subsets of the integers through the introduction of new analytic and combinatorial techniques. Furthermore, interesting new problems have been posed and new applications of classical tools have been discovered. Finally, since the writing of [28] this still relatively young field of mathematics has settled on certain notational and normalisation conventions, which the present article reflects.

We have attempted to write this paper in such a way that both readers familiar with Green's original survey will find it interesting, as well those who are not. We give proofs whenever these are not overwhelmingly complex, and references to the most relevant literature otherwise.

The structure of the remainder of the text is as follows. In Chapter 2 we review three classical problems that fall within the scope of the model: counting 3-term arithmetic progressions in dense sets, identifying arithmetic structure in sum sets and dealing with longer arithmetic progressions. These topics are also discussed in Green's original paper, but we repeat them here to settle the notation and because they are fundamental to the narrative of the rest of the article. In Chapter 3 we explain in more detail in what sense the vector space $\mathbb{F}_p^n$ can be considered a toy model for the integers. Chapter 4 is devoted to some of the spectacular recent developments surrounding the three problems discussed in Chapter 2. Finally, in Chapter 5 we describe some of the limitations of the vector space model as well as avenues of future research to be explored.

Acknowledgements. The author is deeply indebted to Ben Green and Tom Sanders for many useful discussions over the years as well as helpful comments on the manuscript. She would also like to thank Trevor Wooley for pointers and references regarding the material in Chapter 5, and the anonymous referees for a careful reading of the article.

1.1. **Preliminaries.** Throughout this article, $G := \mathbb{F}_p^n$ denotes a vector space of dimension n over a finite field of characteristic p, where p is a prime. Fields of prime-power characteristic are only of interest to us in special circumstances as they frequently introduce distracting divisibility restrictions, which we shall try to avoid here. In the 'finite field model' as we understand it here, it is important that p be thought of as small and fixed (we shall see $p = 2, 3$ and 5 most frequently in the sequel), while the dimension n is to be thought of as arbitrarily large (tending to ∞). Asymptotic results are thus always asymptotic in n.

A character on the group G takes the form $\gamma : G \to \mathbb{C}$, sending $x \in G$ to $\omega^{x\cdot t}$ for some $t \in G$, where $\omega = e^{2\pi i/p}$ is a primitive pth root of unity and $x \cdot t$ is the standard scalar product on G. Abusing notation, we may identify the character γ with the corresponding $t \in G$. The group of characters $\widehat{G}$, also known as the *dual* of G, is thus in our case isomorphic to $G = \mathbb{F}_p^n$ itself, but we shall nevertheless write $\widehat{G}$ for emphasis.

any ambient vector space structure, making the problems closer in nature to what we shall call the 'integer case' in this article.

The *Fourier transform* $\widehat{f} : \widehat{G} \to \mathbb{C}$ of a function $f : G \to \mathbb{C}$ is then defined, for each $t \in \widehat{G}$, by the formula

$$\widehat{f}(t) := \mathbb{E}_{x \in G} f(x) \omega^{x \cdot t},$$

where the expectation operator $\mathbb{E}_{x \in G}$ stands for the normalised (finite) sum over all elements $x \in G$. More generally, for any subset $B \subseteq G$ and any function $g : B \to \mathbb{C}$, we let

$$\mathbb{E}_{x \in B} g(x) := \frac{1}{|B|} \sum_{x \in B} g(x).$$

The *inversion formula* states that with the above definition of the Fourier transform, we can recover f from its Fourier coefficients via the sum

$$f(x) = \sum_{t \in \widehat{G}} \widehat{f}(t) \omega^{-x \cdot t}.$$

Plancherel's identity asserts that

$$\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle,$$

where

$$\langle f, g \rangle := \mathbb{E}_x f(x) \overline{g(x)} \text{ and } \langle \widehat{f}, \widehat{g} \rangle := \sum_t \widehat{f}(t) \overline{\widehat{g}(t)},$$

and we shall refer to it as *Parseval's identity* whenever f and g are equal. Note that inner products on physical space are normalised, while those on frequency space are not. We shall adopt the same convention for the $L^p(G)$ and $\ell^p(\widehat{G})$ norms, respectively. Finally, we define the *convolution* $f * g : G \to \mathbb{C}$ of two functions $f, g : G \to \mathbb{C}$ by

$$f * g(x) := \mathbb{E}_{y \in G} f(y) g(x - y).$$

In additive number theory, which deals with the structure of *sum sets* $A + B := \{a + b : a \in A, b \in B\}$ for finite sets A, B, the convolution operator is an exceedingly useful tool since $A + B = \mathrm{supp}(1_A * 1_B)$, where 1_A and 1_B denote the characteristic functions of the sets A and B, respectively. The convolution thus opens up an analytic approach to the study of sum sets. The crucial fact driving the utility of the Fourier transform in this context is that it diagonalizes the convolution operator, in other words, for all $t \in \widehat{G}$,

$$\widehat{f * g}(t) = \widehat{f}(t) \cdot \widehat{g}(t).$$

We leave the verification of this identity as an exercise to the reader.

In order to be able to refer to the set of large Fourier coefficients of a function, we introduce the following definition.

Definition 1.1. *Let $\epsilon > 0$, and let $f : G \to \mathbb{C}$. We define the* ϵ-large spectrum *of f to be the set*

$$\mathrm{Spec}_\epsilon(f) := \{t \in \widehat{G} : |\widehat{f}(t)| \geq \epsilon \|f\|_1\}.$$

Notice that the ϵ-large spectrum of a bounded function cannot contain too many elements. Indeed, by Parseval's identity we have

$$\|f\|_2^2 = \|\widehat{f}\|_2^2 \geq \sum_{t \in \mathrm{Spec}_\epsilon(f)} |\widehat{f}(t)|^2 \geq \epsilon^2 \|f\|_1^2 |\mathrm{Spec}_\epsilon(f)|,$$

and hence $|\mathrm{Spec}_\epsilon(f)| \leq \epsilon^{-2}\|f\|_2^2/\|f\|_1^2$. In particular, when $f = 1_A$ is the characteristic function of a subset $A \subseteq G$ of density $\alpha := |A|/|G|$, then $|\mathrm{Spec}_\epsilon(1_A)| \leq \epsilon^{-2}\alpha^{-1}$. More precise and extremely useful statements can be made about the large spectrum of a function, which we shall barely touch upon in this article. At this point we only mention the following theorem due to Chang [15].

Theorem 1.2. *Let $\epsilon > 0$ and suppose that $A \subseteq G$ is a subset of density α. Then the spectrum $\mathrm{Spec}_\epsilon(1_A)$ is contained in a subspace of dimension at most*

$$8\epsilon^{-2}\log(\alpha^{-1}).$$

For a proof we refer the interested reader to [69] or the alternative approach in [43].

Throughout this text, we employ Vinogradov's notation $f \ll g$ to mean "f is bounded above by a constant times g", or $f = O(g)$. We shall denote by $o(1)$ a quantity that goes to 0 as $|G|$ (and in particular n) goes to infinity.

1.2. **Additive structure in iterated sum sets.** As a warm-up exercise we shall use the Fourier transform to show that iterated sum sets contain a surprising amount of additive structure. For any subset $A \subseteq G = \mathbb{F}_p^n$, let $2A-2A := \{a_1+a_2-a_3-a_4 : a_1, a_2, a_3, a_4 \in A\}$. A rather simple argument, which can be traced back to [11], shows that such an iterated sum set always contains a rather large subspace.

Proposition 1.3. *Let $A \subseteq G$ be a subset of density α. Then there exists a subspace $V \leqslant G$ of codimension at most $2\alpha^{-2}$ such that $2A - 2A \supseteq V$.*

Proof: Observe that the iterated sum set $2A - 2A$ is the support of the function $g := 1_A * 1_A * 1_{-A} * 1_{-A}$, so it suffices to exhibit a large subspace V with the property that $g(x) > 0$ for all $x \in V$. For reasons which shall become clear in a moment, we set $K := \mathrm{Spec}_\rho(1_A)$ for some parameter $\rho > 0$ to be chosen later, and let $V :=< K >^\perp$ be the orthogonal complement of the vector subspace of G spanned by K. By Parseval's identity the codimension of V, which is at most the size of K, is bounded above by $\rho^{-2}\alpha^{-1}$. Now consider

(1.1)

$$g(x) = 1_A * 1_A * 1_{-A} * 1_{-A}(x) = \sum_t |\widehat{1}_A(t)|^4 \omega^{-x\cdot t} = \sum_{t \in K} |\widehat{1}_A(t)|^4 \omega^{-x\cdot t} + \sum_{t \notin K} |\widehat{1}_A(t)|^4 \omega^{-x\cdot t}.$$

When $x \in V =< K >^\perp$, the first sum is by positivity at least as large as the contribution made by $t = 0$, that is, bounded below by $|\widehat{1}_A(0)|^4 = \alpha^4$. On the other hand, the second sum is in absolute value at most

$$\sup_{t \notin K} |\widehat{1}_A(t)|^2 \sum_{t \notin K} |\widehat{1}_A(t)|^2 \leq \alpha^2\rho^2 \cdot \alpha = \alpha^3\rho^2.$$

Setting $\rho := \sqrt{\alpha/2}$, we find that the second sum in (1.1) is at most half the size of the first, and therefore $g(x) \geq \alpha^4/2 > 0$ for all $x \in V$. Finally, with this choice of ρ the codimension of V is bounded above by $2\alpha^{-2}$ as claimed. □

The same is not necessarily true of the 2-fold difference set $A - A$, in a rather strong sense. Consider the group $\mathbb{F}_2^n$, where addition and subtraction conveniently coincide. Green [28] adapted an example of Ruzsa [53] to show that there exist sets of density $1/4$ such that $A + A$ does not contain the coset of any subspace of codimension $\sqrt{n}$. An example of a set with this remarkable property is the set consisting of all vectors $x \in \mathbb{F}_2^n$ with at least $n/2 + \sqrt{n}/2$ 1s with respect to the standard basis. We refer the reader to [28] for a proof and to [57] for further interesting work on this problem.

2. The classics

We begin our exposition with a finite field analogue of a classical result of Roth [52], who proved in 1953 that any subset of the first N integers which does not contain a non-trivial 3-term arithmetic progression, i.e. a configuration of the form $x, x + d, x + 2d$ for some $x, d \in \mathbb{Z}$, must be of vanishingly small size if we think of N as tending to infinity (see Theorem 3.4 in Chapter 3). The corresponding question in $\mathbb{F}_3^n$ was studied by Meshulam [48] in 1995, whose argument is a prime example of both the power of the Fourier transform and the advantages of the finite field model. For these reasons we present it in full detail in Section 2.1. We continue in Section 2.2 with a discussion of Freiman's theorem, which informs us about the structure of sets with small sum set and whose finite field version is due to Ruzsa [54]. Finally, in Section 2.3 we introduce the Gowers uniformity norms, which have been fundamental to the quantitative study of arithmetic patterns of higher complexity, leading up to a sketch of Szemerédi's celebrated theorem on the existence of long arithmetic progressions in dense sets [66].

2.1. 3-term arithmetic progressions in dense sets.

In this section we study a first application of the discrete Fourier transform to a question in additive number theory: how large can a subset of the integers $\{1, \dots, N\}$ be before it must contain a 3-term arithmetic progression? It is intuitively obvious that the larger a subset of the first N integers is, the harder it will be for it to avoid arithmetic structures such as 3-term progressions. It turns out that establishing the exact threshold at which a set is guaranteed to contain a certain arithmetic structure on density grounds alone is a very difficult question, which has kept mathematicians occupied for almost a century and remains unsolved to this day. Successive attempts at improving the bounds for this problem (in particular the upper bound) have led to important developments in discrete harmonic analysis that are of interest in their own right.

Before we prove the first main result in this direction in the finite field model, let us give a brief indication of why the Fourier transform might be useful for counting 3-term arithmetic progressions. Looking back at the definition, it is not difficult to see that if a subset of $G = \mathbb{F}_p^n$ (or rather, its indicator function) has no large Fourier coefficients besides the trivial one at zero, then we would expect it to be quite uniformly distributed across

the entire space. In particular, it turns out to be rather straightforward to count 3-term arithmetic progressions in such a set.

Our first lemma below states that the number of progressions in a set whose non-trivial Fourier coefficients are small coincides with the number one would expect to find in a set whose elements are chosen from G with probability $\alpha = |A|/|G|$. In what follows we fix $p = 3$ for simplicity, although it can easily be taken to be a larger prime.[2]

Lemma 2.1. *Let $A \subseteq G = \mathbb{F}_3^n$ be a subset of density α satisfying $\sup_{t\neq 0} |\widehat{1}_A(t)| = o(1)$. Then A contains $(\alpha^3 + o(1))|G|^2$ 3-term arithmetic progressions.*

PROOF: Let $2 \cdot A = \{2a : a \in A\}$, not to be confused with $2A := A + A$ used earlier. The normalised count of 3-term progressions in A can be written as

$$T_3(1_A, 1_A, 1_A) := \mathbb{E}_{x,d} 1_A(x)1_A(x+d)1_A(x+2d) = \langle 1_A * 1_A, 1_{2\cdot A}\rangle,$$

where the latter equality comes from a change of variable $x + d \mapsto y$ and the definition of convolution. But by Parseval's identity,

$$T_3(1_A, 1_A, 1_A) = \langle \widehat{1}_A^2, \widehat{1}_{2\cdot A}\rangle = \alpha^3 + \sum_{t\neq 0} \widehat{1}_A(t)^2 \widehat{1_A}(-2t) = \alpha^3 + \sum_{t\neq 0} \widehat{1}_A(t)^3,$$

where we have used the fact that $-2 \equiv 1 \bmod 3$. Now the latter sum is bounded above by

$$\sup_{t\neq 0} |\widehat{1}_A(t)| \sum_t |\widehat{1}_A(t)|^2 = \sup_{t\neq 0} |\widehat{1}_A(t)| \cdot \mathbb{E}_x |1_A(x)|^2 = \sup_{t\neq 0} |\widehat{1}_A(t)| \cdot \alpha$$

by Plancherel, from which we obtain

$$T_3(1_A, 1_A, 1_A) = \alpha^3 + o(1)$$

as claimed. □

We shall see in a moment that this relatively easy argument forms a fruitful starting point for the proof of Meshulam's theorem [48], which states that any sufficiently dense set must contain a non-trivial 3-term progression.

Theorem 2.2. *Let $A \subseteq G = \mathbb{F}_3^n$. Suppose that A contains no non-trivial 3-term progressions. Then*

$$|A| \ll \frac{|G|}{\log |G|}.$$

The dichotomy lying at the heart of the proof of Theorem 2.2 is as follows. Either a set A is uniform in the sense that its non-trivial Fourier coefficients are small, in which case Lemma 2.1 tells us that it contains lots of 3-term arithmetic progressions; or it is *not* uniform, in which case we stand to gain some additional structural information about the set from its large Fourier coefficients, which in turn also implies, as we shall see, that it contains non-trivial progressions.[3]

[2]Note, however, that the problem of counting a 3-term arithmetic progressions becomes trivial in characteristic 2, as it amounts to computing the number of triples $(x, y, z) \in A^3$ satisfying $x + y = 2z \equiv 0 \bmod 2$.

[3]Note that while any 3-term progression in $\mathbb{F}_3^n$ forms a complete line, the proof makes no use of this special geometric feature.

To underscore this dichotomy we shall split the proof into several parts. Let us first prove a lemma to the effect that if a set is progression-free, then its characteristic function must possess a large Fourier coefficient, complementing the introductory Lemma 2.1 above.

Lemma 2.3. *Let $A \subseteq G$ be a subset of density α, with $|G| \geq 2\alpha^{-2}$. Suppose that A contains no non-trivial 3-term progressions. Then there exists $t \neq 0$ such that*

$$|\widehat{1}_A(t)| \geq \frac{1}{2}\alpha^2.$$

PROOF: As before, the normalised number of 3-term progressions in A can be written as

$$T_3(1_A, 1_A, 1_A) = \mathbb{E}_{x,d} 1_A(x)1_A(x+d)1_A(x+2d) = \langle 1_A * 1_A, 1_{2\cdot A}\rangle = \alpha^3 + \sum_{t\neq 0} \widehat{1}_A(t)^3.$$

Assuming that A contains no non-trivial 3-term progressions, we have that $T_3(1_A, 1_A, 1_A) = \alpha/|G|$, and therefore by the hypothesis on $|G|$ that

$$\frac{1}{2}\alpha^3 \leq |\sum_{t\neq 0} \widehat{1}_A(t)^3| \leq \sup_{t\neq 0} |\widehat{1_A}(t)| \sum_t |\widehat{1}_A(t)|^2.$$

The latter sum is as usual equal to α by Parseval's identity, and it follows that there exists a $t \neq 0$ with $|\widehat{1_A}(t)| \geq \frac{1}{2}\alpha^2$. □

We shall next show that the existence of a large Fourier coefficient implies that A exhibits significant bias towards an affine subspace of $\mathbb{F}_3^n$ of codimension 1.

Lemma 2.4. *Let $A \subseteq G$ be a subset of density α. Suppose $t \neq 0$ is such that $|\widehat{1_A}(t)| \geq \alpha^2/2$. Then there exists a subspace $V \leqslant \mathbb{F}_3^n$ of codimension 1 on some translate of which A has density at least $\alpha(1 + \alpha/4)$.*

PROOF: Write $V := < t >^\perp$ for the orthogonal complement of the span of the vector t, and let $v_j + V$, $j = 1, 2, 3$ be the complete set of cosets of V. Write also $f_A := 1_A - \alpha$ for the so-called *balanced function* of A. Then

$$\widehat{1}_A(t) = \widehat{f}_A(t) = \frac{1}{|G|}\sum_{j=1}^3 \sum_{x\in v_j+V} (1_A(x) - \alpha)\omega^{tx} = \sum_{j=1}^3 \frac{1}{|G|} \sum_{x\in v_j+V} (1_A(x) - \alpha)\omega^j = \sum_{j=1}^3 a_j\omega^j,$$

where $a_j = (|A\cap(v_j+V)| - \alpha|V|)/|G|$. By taking absolute values and applying the triangle inequality we find that $\sum_j |a_j| \geq \alpha^2/2$. However, $\sum_j a_j = 0$, so $\sum_j |a_j| + a_j \geq \alpha^2/2$, and thus by the pigeonhole principle there exists at least one value of j for which $|a_j|+a_j \geq \alpha^2/6$. It follows that $a_j \geq \alpha^2/12$, or in other words, as $|G| = 3|V|$, that

$$|A \cap (v_j + V)| \geq \alpha(1 + \frac{1}{4}\alpha)|V|.$$

□

We are now in a position to complete the proof of Meshulam's theorem by iteratively applying Lemmas 2.3 and 2.4.

PROOF OF THEOREM 2.2: Suppose A is of density α in G and contains no 3-term arithmetic progressions. Then by Lemma 2.3, it has a non-trivial Fourier coefficient of size at least $\alpha^2/2$, and consequently by Lemma 2.4 there exists a subspace V of codimension 1 and a vector v_j such that A has density at least $\alpha+\alpha^2/4$ on V. Denote the set $(A-v_j)\cap V$ by A_1. Note that since A was progression free, so is A_1 as 3-term arithmetic progressions are translation invariant. From now on we focus on the set $A_1 \subseteq V \cong \mathbb{F}_3^{n-1}$ of density $\alpha+\alpha^2/4$, and repeat the above procedure with $G = \mathbb{F}_3^{n-1}$.

Clearly, our iteration will come to an end if the density of the set A_k obtained at the kth step were to exceed 1. How soon is this likely to happend? In the above argument the density increases from α to 2α in at most $\alpha/(\alpha^2/4) = 4\alpha^{-1}$ steps, from 2α to 4α in at most $2\alpha/((2\alpha)^2/4) = (4\alpha^{-1})/2$ steps and so on, reaching the critical density 1 in at most $(1+1/2+1/4+\dots)4\alpha^{-1} = 8\alpha^{-1}$ steps. This means that the size of G after the iteration ends is at least $3^{n-8\alpha^{-1}}$. But at the final step we must have had $|G| < 2\alpha^{-2}$ because otherwise we could have applied Lemma 2.3 one more time. Solving $3^{n-8\alpha^{-1}} \leq 2\alpha^{-2}$ now gives the desired bound. □

It is not difficult to see that this proof strategy extends to other (translation-invariant) linear patterns defined by one equation (note that a 3-term progression can also be thought of as a triple (x, y, z) satisfying $x + z = 2y$), for which we shall describe exciting recent progress in Section 4.3. However, the Fourier transform technique breaks down already for arithmetic progressions of length 4, which are defined by two equations in four variables. In Section 2.3 we shall see in detail why this is the case, and show how to overcome this problem.

2.2. **The structure of sets with small sum set.** The study of sets with small sum set in additive combinatorics has a long history. In particular, there is a host of important results concerning finite subsets of the integers and other abelian groups which we are unable to treat in satisfactory detail in this survey. For a comprehensive account of this classical theory we therefore refer the reader to [49].

In the finite field model $\mathbb{F}_p^n$, it is not hard to see that given a subset $A \subseteq \mathbb{F}_p^n$, we have $A + A = A$, and therefore $|A + A| = |A|$, if and only if A is a (possibly affine) subspace of $\mathbb{F}_p^n$. A slightly more involved exercise shows that even if we relax the condition on the size of the sum set somewhat, say to $|A + A| \leq K|A|$ with $K < 3/2$, then A still retains a significant amount of subspace structure in the sense that it is contained in a subspace $V \leqslant \mathbb{F}_p^n$ such that $|A| \geq 2|V|/3$.

We call the least constant K with the property that $|A+A| \leq K|A|$ the *doubling constant* of A. In what follows we shall consider the regime where the doubling constant is potentially large but remains constant (compared with the size of the set A and the group G, which we think of as tending to infinity). The main theme of this section will be that if a set has small doubling in this sense, then this must be because it has some underlying algebraic structure. Specifically, we shall show that if $A \subseteq \mathbb{F}_p^n$ is a subset satisfying $|A + A| \leq K|A|$, then it is efficiently contained in a subspace of $\mathbb{F}_p^n$ whose relative size is bounded only in terms of K.

Of course, when a set is efficiently contained in a subspace its growth under the operation of repeatedly taking sum sets is naturally limited – the iterated sum set will eventually fill up the subspace but cannot grow further. Hence we should expect to be able to prove a weaker statement of the following type: if a set has small doubling, then the growth of its iterated sum sets must remain bounded. The extent to which this is the case was quantified by Plünnecke [51] in the following theorem, known as *Plünnecke's inequality*. Recently, a simple and elegant proof was discovered by Petridis [50], which we are fortunate to be able to present here.[4]

Theorem 2.5. *Let $A, B \subseteq G$ be such that $|A+B| \leq K|A|$. Then for all integers $k, m \geq 1$*

$$|kB - mB| \leq K^{k+m}|A|.$$

PROOF: Without loss of generality assume that $|A+B| = K|A|$. Choose a non-empty subset $A' \subseteq A$ such that the ratio $|A'+B|/|A'|$ is minimized, and call this ratio K'. Observe that $K' \leq K$, and that $|A'+B| = K'|A'|$ as well as $|A''+B| \geq K'|A''|$ for all $A'' \subseteq A$.

Claim 2.6. *Let A', B, K' be as above. Then for every set C, $|A'+B+C| \leq K'|A'+C|$.*

Let us complete the proof of the theorem assuming the claim. We first show that for all $m \in \mathbb{N}$,

$$|A'+mB| \leq K'^m|A'|.$$

Indeed, for $m = 1$ the inequality is true by assumption. For $m > 1$ we assume the inequality holds for $m-1$ and set $C = (m-1)B$ in Claim 2.6 to get $|A'+mB| \leq K'|A'+(m-1)B|$. By the inductive hypothesis, the right-hand side is bounded above by $K'^m|A'|$.

The full result now follows from a simple fact known as *Ruzsa's triangle inequality*, which is the statement that for any sets U, V and W, one has $|U||V-W| \leq |U+V||U+W|$. (This is easily proved by defining a map $\phi : U \times (V-W) \to (U+V) \times (U+W)$, which takes $(u, x = v-w)$ to $(u+v, u+w)$, where for each $x \in V-W$ we fix one representation $v-w$, and checking that this map is injective.) Indeed, with this inequality we have

$$|A'||kB - mB| \leq |A'+kB||A'+mB| \leq K'^k|A'| \cdot K'^m|A'| \leq K^{k+m}|A'|^2,$$

which immediately yields $|kB - mB| \leq K^{k+m}|A'| \leq K^{k+m}|A|$ as desired. □

PROOF OF CLAIM 2.6: We shall prove this by induction on the size of the set C. When $|C| = 1$, the claim is trivially true by assumption. Suppose now that the result is true for C, and consider $C' = C \cup \{x\}$. We observe that

$$A'+B+C' = (A'+B+C) \cup [(A'+B+x) \setminus (D+B+x)],$$

where D is the set $D := \{a \in A' : a+B+x \subseteq A'+B+C\}$. But by the defining property of the constant K', we have $|D+B| \geq K'|D|$, so that

$$|A'+B+C'| \leq |A'+B+C| + |A'+B| - |D+B| \leq K'(|A'+C| + |A'| - |D|). \tag{2.1}$$

We shall apply a similar argument a second time, writing

$$A'+C' = (A'+C) \cup [(A'+x) \setminus (E+x)],$$

[4]The argument that followed is valid in significant generality, although we shall continue to think of G as being $\mathbb{F}_p^n$.

where E is the set $E := \{a \in A' : a + x \in A' + C\}$ and the union is disjoint. We conclude that, as $E \subseteq D$,

$$|A' + C'| = |A' + C| + |A'| - |E| \geq |A' + C| + |A'| - |D|,$$

which together with (2.1) implies the claim. □

Equipped with Plünnecke's inequality we are now in a position to prove the promised structural result about sets with small sum set, which serves as another advertisement for the elegance and power of the finite field model. Theorem 2.7 below is a beautiful theorem of Ruzsa [54], who adapted an earlier result of Freiman in the integers [21] to the group $\mathbb{F}_p^n$. It is therefore often referred to as the *Freiman-Ruzsa theorem.*

Theorem 2.7. *Let $A \subseteq G = \mathbb{F}_p^n$ be a subset satisfying $|A + A| \leq K|A|$. Then A is contained in a subspace $H \leqslant \mathbb{F}_p^n$ of size at most $K^2 p^{K^4}|A|$.*

PROOF: The crucial idea that gets us started is to choose a subset $X \subseteq 2A - A$ which is maximal with respect to the property that the translates $x + A$ for $x \in X$ are disjoint. We first show that such a set X cannot be too large. Indeed, we clearly have $X + A \subseteq 3A - A$, and by Plünnecke's inequality (Theorem 2.5) we know that $|3A - A| \leq K^4|A|$. Since the sets $x + A$ are also disjoint and each of size $|A|$, we therefore have

$$K^4|A| \geq |3A - A| \geq |X + A| = \sum_{x \in X} |x + A| = |X||A|,$$

and thus $|X| \leq K^4$.

Next we show that

$$2A - A \subseteq X + (A - A).$$

To see this, observe that if $y \in 2A - A$, then $y + A \cap x + A \neq \emptyset$ for some $x \in X$: if $y \in X$ the statement is trivial, and if $y \notin X$ it follows from the assumption that X was chosen maximally. In either case it follows that $y \in X + (A - A)$.

Adding A repeatedly to both sides of the preceding inclusion leads to the statement that

$$kA - A \subseteq (k - 1)X + (A - A) \tag{2.2}$$

for all $k \geq 2$. This is encouraging as we seem to be able to compress more and more sums of A into very few translates of $A - A$ (remember that X is of constant size).

Writing H for the subgroup of $\mathbb{F}_p^n$ generated by A and Y for the subgroup generated by X, we infer from (2.2) that

$$H = \bigcup_{k \geq 1} (kA - A) \subseteq Y + (A - A).$$

But every element of Y can be written as the sum of at most $|X|$ elements with coefficients between 1 and p, so $|Y| \leq p^{|X|} \leq p^{K^4}$. The observation that

$$|H| \leq |Y||A - A| \leq K^2 p^{K^4}|A|$$

concludes the proof. □

The dependence of the size of H on the doubling constant K can be improved by more careful arguments (in particular using combinatorial compressions, see [33, 19, 20]). However, the following example shows that it must be exponential.

Example 2.8. *Consider the set A consisting of the union of a very large subspace H with $K-1$ randomly chosen elements (not in H). Then A has doubling about K, but any (affine) subspace H' containing A must have size at least $p^{K-2}|A|$.*

Note, however, that the above example is still highly structured: apart from a constant number of points, the set A really is contained in a subspace of size at most $|A|$ (namely H). This leads to the following natural reformulation of the Freiman-Ruzsa theorem.

Theorem 2.9. *Let $A \subseteq G = \mathbb{F}_p^n$ be a subset satisfying $|A+A| \leq K|A|$. Then there exists a subspace $H \leqslant \mathbb{F}_p^n$ of size at most $C_1(K)|A|$ such that for some $x \in G$,*

$$|A \cap (x+H)| \geq \frac{|A|}{C_2(K)},$$

where $C_1(K)$ and $C_2(K)$ are constants depending only on K.

We are now able to state the *Polynomial Freiman-Ruzsa Conjecture* (often referred to by the shorthand 'PFR'), which, over one decade after its popularisation by Green, is still one of the most central open problems in additive combinatorics (see Section 10 of [28] for various alternative formulations and a detailed history of the conjecture).

Conjecture 2.10. *The constants $C_1(K)$ and $C_2(K)$ in Theorem 2.9 above can be taken to be polynomial in K.*

Recent work of Sanders [60] comes very close to resolving this conjecture; we shall discuss it in Section 4.2. In the past decade there have also been numerous generalisations of Theorem 2.7 to other groups, including non-abelian ones. The most general result in this direction to date is [14], and some excellent recent surveys on this topic are [30, 61, 41].

2.3. **Quadratic uniformity and Szemerédi's theorem.** In Section 2.1 we saw how the Fourier transform was useful for establishing the existence of certain arithmetic structures in dense subsets of $\mathbb{F}_p^n$. The crucial starting point of the proof of Theorem 2.2 was the Fourier identity

$$T_3(A,A,A) = \mathbb{E}_{x,d} 1_A(x) 1_A(x+d) 1_A(x+2d) = \sum_t \widehat{1}_A(t)^3,$$

which related the number of 3-term progressions in A to a sum of Fourier coefficients that was easily estimated. It is natural to ask whether a similar identity can be formulated for 4-term progressions, and the answer turns out to be negative. (The sum on the right-hand side requires two distinct parameters, prohibiting any straightforward estimation.)

This fact alone, however, would not be a good reason to rule out the use of Fourier analysis in order to establish a version of Meshulam's theorem for 4-term progressions. But there is a more serious obstacle: the analogue of Lemma 2.1, which constitutes the fundamental premise of the proof of Meshulam's theorem, breaks down for 4-term progressions. Recall

that this lemma said that if a set is uniform in the sense that all its non-trivial Fourier coefficients are small, then it contains the number of 3-term progressions expected in the random case.

Theorem 2.11 below exhibits an example of a set which is uniform in the Fourier sense just described, but which contains significantly more than the expected number of 4-term progressions. This was well known in ergodic theory before it was rediscovered by Gowers in [23]. We give a finite field version of his argument here.

Theorem 2.11. *Let $p > 4$ be a prime. There exists $\epsilon > 0$ such that for every $\delta > 0$ there exists n and a set $A \subseteq G = \mathbb{F}_p^n$ of density α with the following properties.*

1. *The set A is uniform in the sense that $\sup_{t \neq 0} |\widehat{1}_A(t)| \leq \delta$.*
2. *The set A contains significantly more than the expected number of 4-term progressions, namely at least a proportion $\alpha^4 + \epsilon$.*

PROOF: Luckily such a set is very easy to write down in $\mathbb{F}_p^n$: we can take, for example,

$$A := \{x \in \mathbb{F}_p^n : x \cdot x = 0\}.$$

Then the characteristic function of A can be written as $1_A(x) = \mathbb{E}_u \omega^{u(x \cdot x)}$, where the expectation in u is taken over $\mathbb{F}_p$. Using a standard Gauss sum estimate for the quadratic exponential sum, it is straightforward to estimate the Fourier coefficient of 1_A at $t \neq 0$ in absolute value by

$$|\widehat{1}_A(t)| = |\mathbb{E}_{x \in \mathbb{F}_p^n, u \in \mathbb{F}_p} \omega^{ux \cdot x + x \cdot t}| \leq \mathbb{E}_{u \neq 0} |\mathbb{E}_{x \in \mathbb{F}_p^n} \omega^{q_u(x)}| \leq p^{-n/2},$$

where we have written $q_u(x)$ for the quadratic form $ux \cdot x + x \cdot t$, whose rank is n whenever u is non-zero. We conclude that property (1) holds for any δ provided n is taken large enough. Similarly, we find that $|\mathbb{E}_x 1_A(x) - p^{-1}| \leq p^{-n/2}$ (as $u = 0$ with probability p^{-1}, and the rest of the time the exponential sum is tiny), and so the density of A is approximately p^{-1} for large n.

The proof of property (2) relies crucially on the following elementary identity

$$x \cdot x - 3(x + d) \cdot (x + d) + 3(x + 2d) \cdot (x + 2d) - (x + 3d) \cdot (x + 3d) = 0, \tag{2.3}$$

which is valid for all x and d in G. Recall that our aim is to count the number of 4-term progressions in A, that is, the number of $(x, d) \in (\mathbb{F}_p^n)^2$ for which $L_i(x, d) \cdot L_i(x, d)$ are simultaneously zero for $i = 1, \ldots, 4$, where $L_i(x, d) = x + (i - 1)d$. Observe that by (2.3) $L_4 \cdot L_4$ is automatically zero whenever $L_1 \cdot L_1, L_2 \cdot L_2, L_3 \cdot L_3$ are zero, so it suffices to count the number of 3-term progressions in A. But Lemma 2.1 told us that if the non-trivial Fourier coefficients of A are small (as they are in this case), then A contains the number of 3-term progressions expected in the random case. It therefore follows that, up to a negligible error, the set A contains a proportion of at least $(p^{-1})^3$ 4-term progressions, which is bounded below by $\alpha^4 + \epsilon$ for some $\epsilon > 0$. □

If the Fourier transform is not sufficient for counting 4-term progressions, a new tool is needed. One of Gowers's crucial achievements in [24] was the introduction of a new sequence of norms, usually referred to as the *uniformity, Gowers* or *U^k norms*, and the realisation that these norms play an important role in counting long arithmetic progressions. These

norms can be defined in a variety of contexts, including the integers, but for simplicity we shall continue to think of G as a vector space over a finite field.

Definition 2.12. *Let $k \geq 2$ be an integer. The* U^k norm *of a function $f : G \to \mathbb{C}$ is defined by the formula*

$$\|f\|_{U^k}^{2^k} = \mathbb{E}_{x,h_1,\dots,h_k \in G} \Delta_{h_1} \Delta_{h_2} \dots \Delta_{h_k} f(x),$$

where $\Delta_h f(x) := f(x)\overline{f(x+h)}$ is to be thought of as a discrete phase derivative.

It is not too difficult to see, but certainly not obvious, that $\|\cdot\|_{U^k}$ is indeed a norm, and that these norms are nested in the sense that for any integer $k \geq 2$,

$$\|f\|_{U^2} \leq \|f\|_{U^3} \leq \dots \leq \|f\|_{U^k} \leq \|f\|_\infty.$$

We leave the details to the keen reader, who may wish to refer to the book [69] or the lecture notes [29].

Let us first examine the case $k = 2$ in more detail. Unravelling the definition, we see that

$$\|f\|_{U^2}^4 = \mathbb{E}_{x,a,b} f(x)\overline{f(x+a)f(x+b)}f(x+a+b),$$

This expression counts the number of so-called *additive quadruples* of f, which are closely related to the size of the sum set studied in the preceding section. Indeed, it is not difficult to show that if A has small doubling, then it must contain many quadruples of the form $(a_1, a_2, a_3, a_4) \in A^4$ satisfying $a_1 + a_2 = a_3 + a_4$, which can be reparameterised as $(x, x+a+b, x+a, x+b)$. This is a first indication that the results obtained in Section 2.2 may be of use to us beyond their immediate combinatorial applications.

Observe that we can also write

$$\|f\|_{U^2}^4 = \mathbb{E}_x |f * f(x)|^2 = \sum_t |\widehat{f}(t)|^4 = \|\widehat{f}\|_4^4,$$

where the penultimate equality is again just Parseval's identity, linking the U^2 norm of f to the ℓ^4 norm of its Fourier transform. In fact, we can even relate it to the ℓ^∞ norm of $\widehat{f}$, which we employed to such great effect in Section 2.1, by observing that

$$\sup_t |\widehat{f}(t)|^4 \leq \sum_t |\widehat{f}(t)|^4 \leq \sup_t |\widehat{f}(t)|^2 \sum_t |\widehat{f}(t)|^2$$

and thus

$$\|\widehat{f}\|_\infty \leq \|\widehat{f}\|_4 = \|f\|_{U^2} \leq \|\widehat{f}\|_\infty^{1/2} \|f\|_2^{1/2}.$$

This means that for L^2 bounded functions, the U^2 norm and the ℓ^∞ norm on Fourier space are (at least in a qualitative sense) equivalent. This raises the hope that arguments using the Fourier transform might be rewritten in terms of the U^2 norm, and that those new proofs might then be more amenable to generalisation.

Indeed, in many cases such a strategy turns out to be feasible and fruitful. An example is the following lemma, whose content we have already established in Lemma 2.1: it says that the U^2 norm controls the count of 3-term arithmetic progressions. As an immediate corollary, we note that whenever the balanced function $f_A = 1_A - \alpha$ of a set $A \subseteq G$ of

density α has small U^2 norm, then the set A contains roughly the expected number $\alpha^3|G|^2$ of arithmetic progressions of length 3.

Lemma 2.13. *Let $p \geq 3$ be a prime. Let $f : G = \mathbb{F}_p^n \to \mathbb{C}$ be a function satisfying $\|f\|_\infty \leq 1$. Then*

$$|\mathbb{E}_{x,d} f(x)f(x+d)f(x+2d)| \leq \|f\|_{U^2}.$$

PROOF: Set $u = x + d$ and write

$$|\mathbb{E}_{x,d} f(x)f(x+d)f(x+2d)|^2 = |\mathbb{E}_{x,u} f(x)f(u)f(2u-x)|^2,$$

which is bounded above by

$$\mathbb{E}_u |\mathbb{E}_x f(x)f(2u-x)|^2 = \mathbb{E}_u \mathbb{E}_{x,x'} f(x)f(2u-x)\overline{f(x')f(2u-x')}.$$

Finally, reparameterising with $x' = x + a$, $2u - x' = x + b$ gives[5]

$$\mathbb{E}_{x,a,b} f(x)\overline{f(x+a)f(x+b)}f(x+a+b) = \|f\|_{U^2}^4.$$

□

The usefulness of the U^3 norm for our immediate purpose, namely establishing the existence of 4-term progressions in dense subsets of G, lies in the following proposition from [23]. It states that the U^3 norm controls the 4-term progression count in f, and as such constitutes a generalisation of Lemma 2.13 above.[6]

Proposition 2.14. *Let $p > 4$ be a prime. Let $f : G \to \mathbb{C}$ be a function satisfying $\|f\|_\infty \leq 1$. Then*

$$|\mathbb{E}_{x,d} f(x)f(x+d)f(x+2d)f(x+3d)| \leq \|f\|_{U^3}.$$

In other words, if the balanced function $f_A = 1_A - \alpha$ of a set $A \subseteq G$ of density α has small U^3 norm (in which case we call the set A *quadratically uniform*), then the set A contains roughly the expected number $\alpha^4|G|^2$ of arithmetic progressions of length 4. We omit the proof of Proposition 2.14 as it is a straightforward (if slightly tedious) generalisation of the Cauchy-Schwarz argument used to prove Lemma 2.13 above.

At this point it is critical to recall that the proof of Meshulam's theorem (Theorem 2.2) proceeded via a dichotomy: either the non-trivial Fourier coefficients of 1_A were small, in which case we were able to count the number of 3-term progressions in A rather precisely, or there existed at least one large Fourier coefficient, which led to a density increase on a hyperplane. Reformulated in terms of the U^2 norm, the first part of the dichotomy now states that if $\|f_A\|_{U^2}$ is small, where $f_A = 1_A - \alpha$ is the balanced function of A, then A contains approximately the expected number of 3-term progressions. The second part of the dichotomy corresponds to the following U^2 *inverse theorem,* which describes the structure of functions whose U^2 norm is large.

[5]The attentive reader may have noticed that, in fact, we have proved a stronger result than stated. The more general version $|\mathbb{E}_{x,d} f_1(x)f_2(x+d)f_3(x+2d)| \leq \min_{i=1,2,3} \|f_i\|_{U^2}$, however, which is used in many applications, requires one further application of Cauchy-Schwarz.

[6]Green and Tao refer to Proposition 2.14 as a *von Neumann theorem.*

Theorem 2.15. *Let $\delta > 0$ and let $f : \mathbb{F}_p^n \to \mathbb{C}$ be a function satisfying $\|f\|_\infty \leq 1$ and $\|f\|_{U^2} \geq \delta$. Then there exists $b \in \mathbb{F}_p^n$ such that*

$$|\mathbb{E}_x f(x)\omega^{x\cdot b}| \geq \delta^2.$$

In other words, f correlates with a linear phase function of the form $\phi(x) = \omega^{x\cdot b}$.

PROOF: The proof follows straight from the above-mentioned equivalence between the U^2 norm and the ℓ^∞ norm of the Fourier transform of f. Indeed, following Definition 2.12 we saw that $\|f\|_{U^2}^2 \leq \|\widehat{f}\|_\infty \|f\|_2$, so the hypotheses imply that $\|\widehat{f}\|_\infty \geq \delta^2$. The result now follows from the definition of the Fourier transform. □

In order to prove a version of Theorem 2.2 for 4-term progressions, we must therefore generalise Theorem 2.15 to the U^3 norm. In other words, we must provide an answer to the following question: when is a function large in the U^3 norm? A first observation (and easy exercise) in this direction is that if a function $f : G \to \mathbb{C}$ is such that $\|f\|_{U^3} = 1$, then f must be of the form $f(x) = \omega^{q(x)}$ for some quadratic form q on $\mathbb{F}_p^n$, where ω as usual is a primitive pth root of unity. This suggests that functions with large U^3 norm are of a somewhat quadratic nature. In view of Theorem 2.15 it is therefore not unreasonable to conjecture that if f has non-negligible U^3 norm, then it must correlate with a *quadratic phase function*, by which we shall mean a function of the form $\omega^{q(x)}$ for some quadratic form q. Theorem 2.16 to this effect, due to Green and Tao ($p \neq 2$) [31] and Samorodnitsky ($p = 2$) [55], is known as the U^3 *inverse theorem* over finite fields. Its proof is largely based on groundbreaking work of Gowers [23], who had previously provided a (slightly weaker) statement for functions defined on the integers.

Theorem 2.16. *Let $\delta > 0$, and let $f : \mathbb{F}_p^n \to \mathbb{C}$ be a function satisfying $\|f\|_\infty \leq 1$ and $\|f\|_{U^3} \geq \delta$. Then there exists an $n \times n$ matrix M and $b \in \mathbb{F}_p^n$ such that*

$$|\mathbb{E}_x f(x)\omega^{xMx+b\cdot x}| \geq c(\delta),$$

for some constant $c(\delta)$ going to 0 as δ tends to 0.[7]

The proof of Theorem 2.16 is deep and combines Fourier analysis and additive combinatorics in an ingenious way. We shall be unable to do it justice in this brief account and only provide the vaguest of sketches. For full details see Green's lecture notes [29].

The starting point of the proof is the observation that the U^3 norm satisfies the inductive formula

$$\|f\|_{U^3}^8 = \mathbb{E}_h \|\Delta_h f\|_{U^2}^4,$$

where we remind the reader that we had defined $\Delta_h f(x) = f(x)\overline{f(x+h)}$. We now immediately see that if $\|f\|_{U^3}$ is large, then for many values of h, $\|\Delta_h f\|_{U^2}$ is large, which by the inverse theorem for the U^2 norm (Theorem 2.15) means that each such derivative $\Delta_h f$ correlates with a linear phase function. Now the intuition is that if the derivative of f correlates with a linear phase function, then f itself must correlate with a quadratic phase function. The problem with making this heuristic rigorous is that for each h we obtain a potentially different linear phase, and it is not a priori clear that these will line up to

[7] We shall discuss the quantitative dependence of $c(\delta)$ on δ below.

allow us to 'integrate' the linear structure of $\Delta_h f$ to the desired statement about f. It is here that various results from additive combinatorics, including the Freiman-Ruzsa theorem (Theorem 2.7), come into play: they provide us with additional structural information which enables us to 'glue' the different linear phases together.

The use of the Freiman-Ruzsa theorem is so crucial in the proof of Theorem 2.16 that the dependence of $c(\delta)$ on δ depends entirely on the bound in the Freiman-Ruzsa theorem. In point of fact, Green and Tao [35] and independently Lovett [46] have shown that the Polynomial Freiman-Ruzsa Conjecture (Conjecture 2.10) is equivalent to polynomial bounds in Theorem 2.16 (a conjecture which they termed the *Polynomial Gowers Inverse Conjecture*, or PGI for short).

Combining Proposition 2.14 with Theorem 2.16 in the case $p = 5$, it is not too difficult to deduce a generalisation of Meshulam's theorem to progressions of length 4. Such a generalisation is better known as *Szemerédi's theorem for finite fields*, and indeed it was Szemerédi [66] who first proved a qualitative statement to this effect for the integers in 1975, using an exceedingly clever combination of purely combinatorial techniques (amongst them his celebrated regularity lemma). Its quantitative finite field analogue is due to Green and Tao [31].

Theorem 2.17. *Let $A \subseteq G = \mathbb{F}_5^n$ be a set containing no (non-trivial) 4-term arithmetic progressions. Then*

$$|A| \ll \frac{|G|}{(\log\log|G|)^c},$$

where $c = 2^{-21}$.

This bound is of the same shape as that obtained by Gowers [23] in the integers, and indeed the proof follows his argument very closely. Not surprisingly, it proceeds via a density increment strategy similar to that used in the proof of Theorem 2.2. Instead of obtaining a density increment on a hyperplane, however, Theorem 2.16 implies a density increment on the zero-set of a quadratic form, but such sets can (with some effort) be efficiently linearised.[8]

3. Passing from the model setting to the integers

Even though historically many of the results we have presented so far were known in the integers (sometimes decades) before the finite field analogue was proved, the real strength of the finite field model lies in the possibility of reversing this order: one hopes to be able to attack a problem in the simplified model setting of $\mathbb{F}_p^n$ and then to transfer the main features of the argument to the technically more demanding context of the integers. As we mentioned in the introduction, in many cases this can be achieved in a now more or less standardised fashion. Because the method is highly technical, however, we shall only be able to explain the general principle here. Unfortunately, for many applications its implementation remains a case of 'learning by doing'.

[8] A bound of $|G|/(\log|G|)^c$ in the above theorem, obtained by refinements of the basic iteration argument, is now known [36] and far exceeds anything proved in the integer case.

In fact, instead of dealing with the integers directly (which, from the point of view of a discrete harmonic analyst, are a somewhat uncomfortable infinite group with a continuous dual), one can consider many questions of interest in arithmetic combinatorics in a large cyclic group $\mathbb{Z}/N\mathbb{Z}$ with N a prime, which we shall abbreviate as $\mathbb{Z}_N$ in the sequel (not to be confused with the N-adic integers). Specifically, it is often convenient to embed the initial segment of integers $\{1,\dots,M\}$, in which the problem is originally posed, into a cyclic group $\mathbb{Z}_N$ of prime order N, which is chosen sufficiently large so that no 'wrap-around' issues arise. We shall gloss over this point repeatedly below by interchanging $\{1,\dots,N\}$ and $\mathbb{Z}_N$ whenever it is convenient.

Before moving on, let us clarify what we mean by the Fourier transform on $\mathbb{Z}_N$. For a function $f:\mathbb{Z}_N\to\mathbb{C}$ we let $\widehat{f}:\widehat{\mathbb{Z}_N}\to\mathbb{C}$ be the function which is defined, for each $t\in\widehat{\mathbb{Z}_N}$, by the formula

$$\widehat{f}(t)=\mathbb{E}_{x\in\mathbb{Z}_N}f(x)\omega^{xt},$$

where ω is an Nth root of unity. (Again, the Pontryagin dual of $\mathbb{Z}_N$ is isomorphic to $\mathbb{Z}_N$ itself, but we still write $\widehat{\mathbb{Z}_N}$ for emphasis.) The inversion formula and Parseval's identity are virtually identical to those in $\mathbb{F}_p^n$, following the same normalisation conventions as before.

In a such group $\mathbb{Z}_N$ then we clearly have no non-trivial algebraic substructures that could play the role of the subspaces of $\mathbb{F}_p^n$, which we made incessant use of in Section 2. Considering a dense subspace $V\leqslant\mathbb{F}_p^n$ as being defined by the bounded number of vectors that span its orthogonal complement, however, we arrive at the following natural definition that will serve as a substitute for a subspace in $\mathbb{Z}_N$.

Definition 3.1. *Let $K\subseteq\widehat{\mathbb{Z}_N}$ be a set of frequencies and let $\rho>0$. The* Bohr set *or* Bohr neighbourhood *$B(K,\rho)$ is defined as*

$$B(K,\rho):=\{x\in\mathbb{Z}_N:\sup_{t\in K}\|xt/N\|<\rho\},$$

where $\|\beta\|$ denotes the distance of β from the nearest integer.[9]

The parameter ρ is called the *width* of $B(K,\rho)$, while the size of K is often referred to as the *dimension* of the Bohr set (even though strictly speaking this parameter corresponds to the *co*dimension of the subspace of which it is supposed to be an analogue).

Without further ado, let us give the $\mathbb{Z}_N$-analogue of Bogolyubov's lemma (Proposition 1.3). The reader will notice that the two arguments are nearly identical, modulo the fact that we have traded in the notion of exact orthogonality for an approximate one.

Proposition 3.2. *Let $A\subseteq\mathbb{Z}_N$ be a subset of density α. Then there exists a subset $K\subseteq\widehat{\mathbb{Z}_N}$ of size at most $2\alpha^{-2}$ such that $A+A-A-A\supseteq B(K,1/4)$.*

Proof: As in the finite field version, we begin by writing, for each $x\in\mathbb{Z}_N$,

$$g(x)=1_A*1_A*1_{-A}*1_{-A}(x)=\sum_{t\in\widehat{\mathbb{Z}_N}}|\widehat{1}_A(t)|^4\omega^{xt}.$$

[9]In the setting of a general group H, we define the Bohr set $B(K,\rho)$ for $K\subseteq\widehat{H}$ and $\rho>0$ by $B(K,\rho):=\{x\in H:\sup_{\gamma\in K}|\gamma(x)-1|\leq\rho\}$. When $H=\mathbb{Z}_N$ this is equivalent to Definition 3.1 (up to a constant factor), and we stick with the latter throughout this section for reasons of notational simplicity.

Set $K := \mathrm{Spec}_\gamma(1_A)$ for some parameter γ to be chosen later. We shall show that for $x \in B(K, 1/4)$, the real part of the expression on the right-hand side is strictly positive, and thus any $x \in B(K, 1/4)$ also lies in $2A - 2A$.

First, observe that for $x \in B(K, 1/4)$ and $t \in K$, we have $\cos(2\pi xt/N) > 0$, and hence

$$\begin{aligned}\Re \sum_t |\widehat{1}_A(t)|^4 \omega^{xt} &= \sum_{t\in K} |\widehat{1}_A(t)|^4 \cos(2\pi xt/N) + \sum_{t\notin K} |\widehat{1}_A(t)|^4 \cos(2\pi xt/N) \\ &\geq |\widehat{1}_A(0)|^4 - \sum_{t\notin K} |\widehat{1}_A(t)|^4.\end{aligned}$$

We easily compute $|\widehat{1}_A(0)|^4 = \alpha^4$ and $\sum_{t\notin K} |\widehat{1}_A(t)|^4 \leq \gamma^2\alpha^3$ by Parseval as before. It follows that for $x \in B(K, 1/4)$, $g(x)$ is strictly positive if we set $\gamma := \sqrt{\alpha/2}$. □

In practice we will often need to extract more rigid structural information from a Bohr set. In particular, for an argument such as the one used to prove Meshulam's theorem to be applied in iteration, we will need the structure obtained upon a first application of the Fourier transform to resemble the original space $\{1, \dots, N\}$ itself, in other words, we would like the Bohr set to somehow look like an arithmetic progression. The following folklore lemma is an easy first step in this direction.

Lemma 3.3. *Let K be a non-empty subset of $\widehat{\mathbb{Z}_N}$ and let $\rho > 0$. Then the Bohr set $B(K, \rho)$ contains an arithmetic progression of size at least $\rho N^{1/|K|}$ centered at 0.*

Proof: Consider the $|K|$-dimensional torus $\mathbb{T}^{|K|} := (\mathbb{R}/\mathbb{Z})^{|K|}$, and split it into $|K|$-dimensional subcubes of side length $N^{-1/|K|}$ (as equitably as possible). There are at most $\lceil N^{1/|K|} \rceil$ such subcubes. Now consider the map $\phi : \mathbb{Z}_N \to \mathbb{T}^{|K|}$ which sends $x \in \mathbb{Z}_N$ to the vector $(xt/N)_{t\in K} \in \mathbb{T}^{|K|}$. By the pigeonhole principle, there must be two elements $x, x' \in \mathbb{Z}_N$ such that $\phi(x)$ and $\phi(x')$ lie in the same subcube, that is, $\|(x-x')t/N\| \leq N^{-1/|K|}$ for all $t \in K$. Let $z := x - x'$. Then for all $s \in \mathbb{Z}_N$ such that $-\rho N^{1/|K|}/2 \leq s \leq \rho N^{1/|K|}/2$, we have $\|szt/N\| < \rho$ for all $t \in K$, which implies that the arithmetic progression

$$\{sz : -\rho N^{1/|K|}/2 \leq s \leq \rho N^{1/|K|}/2\}$$

is contained in $B(K, \rho)$. □

Indeed, Roth's original argument [52] essentially proceeded via a density increment strategy as in Theorem 2.2 but relative to a long arithmetic progression, resulting in the following theorem.

Theorem 3.4. *Let $A \subseteq \{1, \dots, N\}$. Suppose that A contains no non-trivial 3-term progressions. Then*

$$|A| \ll \frac{N}{\log\log N}.$$

Notice that the bound here is much weaker than the one we obtained in the finite field case. But the attentive reader may already have been struck by the fact that in passing to an arithmetic progression inside the Bohr set we appear to have discarded its potentially

useful multidimensional structure. Using some arguments from the geometry of numbers, it is instead possible to prove the following stronger statement (see [69]).[10]

Proposition 3.5. *Let K be a non-empty subset of $\widehat{\mathbb{Z}_N}$ and let $\rho > 0$. Then the Bohr set $B(K, \rho)$ contains a proper multidimensional arithmetic progression of dimension $|K|$ and size at least $(\rho/|K|)^{|K|}N$. In other words,*

$$B(K,\rho) \supseteq \{x \in \mathbb{Z}_N : x = \sum_{i=1}^{|K|} m_i x_i : m_i \in \mathbb{Z}, |m_i| \leq l_i\} \tag{3.1}$$

for some $x_1, \dots, x_{|K|} \in \mathbb{Z}_N$ and some integers l_i such that $\prod_{i=1}^{|K|} l_i \geq (\rho/|K|)^{|K|}N$.

Reassuringly, these multidimensional arithmetic progressions also crop up as analogues of subspaces in less analytic arguments. For example, the integer equivalent of the Freiman-Ruzsa theorem (Theorem 2.9), due to Freiman [21], states that a finite subset of the integers whose sum set is small is efficiently contained in a multidimensional arithmetic progression.

Theorem 3.6. *Let $A \subseteq \mathbb{Z}$ be a finite subset of the integers satisfying $|A + A| \leq C|A|$ for some constant C. Then A is contained in a translate of a multidimensional arithmetic progression of the form (3.1) of dimension at most C_1 and size at most $C_2|A|$, where C_1 and C_2 are constants depending only on C.*

A conjecture analogous to Conjecture 2.10 applies in this context.

It turns out that for many purposes, a Bohr set $B(K, \rho)$ can also usefully be thought of as a metric ball of radius ρ and dimension $|K|$. Indeed, it displays a similar behaviour as far as size is concerned. The proof of the following lemma is elementary and omitted (see [69] for details).

Lemma 3.7. *For any frequency set $K \subseteq \widehat{\mathbb{Z}_N}$ and width parameter $\rho > 0$, we have*

$$|B(K,\rho)| \geq \rho^{|K|}N \quad \text{and} \quad |B(K,2\rho)| \leq 4^{|K|}|B(K,\rho)|.$$

Whether we think of Bohr sets as balls or (multidimensional) arithmetic progressions, what we have lost by passing from our model setting to the integers is the very useful property of additive closure that a subspace of $\mathbb{F}_p^n$ possesses: it is no longer true that when two elements $x, y \in \mathbb{Z}_N$ belong to a Bohr set B, then $x + y$ also belongs to B. However, such a closure property can be replicated in an approximate sense by considering pairs of Bohr sets (B, B'). It turns out that, with an appropriate choice of parameters, these can be made to behave like *approximate subgroups* in the sense that $B + B' \approx B$. We illustrate this idea with the following simple lemma, which says that for a pair of Bohr sets (B, B') for which $B + B'$ is not much larger than B itself, the number of 3-term progressions in B is almost what one would expect to find in the case of two subspaces.

[10]Note that the size of the arithmetic progression obtained here is linear in N, instead of a small power of N as in Lemma 3.3 above.

Lemma 3.8. *Let $\epsilon > 0$, let $B \subseteq \mathbb{Z}_N$ and let $B' \subseteq \mathbb{Z}_N$ be a symmetric set containing 0 such that $|B + B'| \leq (1+\epsilon)|B|$. Then*

$$T_3(B, B, B) \geq (1 - 2\epsilon)\beta\beta',$$

where β and β' denote the densities of B and B' in $\mathbb{Z}_N$, respectively.

PROOF: For each fixed $b \in B'$ we have $B \cup (B+b) \subseteq B+B'$, and therefore by hypothesis $|B \cup (B+b)| \leq (1+\epsilon)|B|$. The same bound holds by a similar argument for $|B \cup (B-b)|$. Now for fixed $b \in B'$, we have by elementary considerations that

$$|B \cap (B+b) \cap (B-b)| = 3|B| - |B \cup (B+b)| - |B \cup (B-b)| + |B \cup (B+b) \cup (B-b)| - |(B+b) \cup (B-b)|,$$

which is bounded below by $3|B| - 2(1+\epsilon)|B| + 0 = (1-2\epsilon)|B|$. It follows that there are at least $(1-2\epsilon)|B||B'|$ 3-term arithmetic progressions of the form $x-b, x, x+b$ in B, giving the desired result upon normalisation. □

This notion of approximate additive closure, which is due to Bourgain [12], is formalised in the following definition.

Definition 3.9. *A Bohr set $B := B(K, \rho)$ is said to be* regular *if, for every $0 < \epsilon < (100|K|)^{-1}$,*

$$|B(K, \rho(1+\epsilon))| \leq |B|(1 + 100|K|\epsilon) \text{ and } |B(K, \rho(1-\epsilon))| \geq |B|(1 - 100|K|\epsilon).$$

Now since $B(K, \sigma) + B(K, \tau) \subseteq B(K, \sigma+\tau)$, we see that a regular Bohr set $B := B(K, \rho)$ together with its scaled-down version $B' := B(K, \rho\epsilon)$ provides a pair (B, B') with the property that $B + B' \approx B$. The following lemma, again due to Bourgain [12], tells us that there are a fair number of choices for the width parameter ρ that make a Bohr set with given frequency set $K \subseteq \widehat{\mathbb{Z}_N}$ regular.

Lemma 3.10. *Let K be a subset of $\widehat{\mathbb{Z}_N}$ and let $\rho_0 > 0$. Then there exists $\rho \in [\rho_0, 2\rho_0]$ such that the Bohr set $B(K, \rho)$ is regular.*

In applications, one often requires nested sequences of regular Bohr sets with carefully chosen width parameters, making the resulting arguments technically rather challenging. Increasingly sophisticated refinements of this technology have led to several quantitative improvements in Theorem 3.4 above, starting with Bourgain [12], who proved a bound of the form $|A| \ll N \cdot (\log\log N / \log N)^{1/2}$ and culminating in recent work of Sanders, who showed that $|A| \ll N \cdot (\log\log N)^6 / \log N$ (see also Section 4.3).

To conclude this section, let us briefly remark on the analogue of quadratic Fourier analysis in $\mathbb{Z}_N$, which turns out to be vastly more complex than its finite field counterpart. Indeed, it is not true that any function f of large U^3 norm correlates with a quadratic phase function of the form ω^q for some quadratic function q defined on $\mathbb{Z}_N$. Instead, such a statement is only true locally, that is, with the quadratic restricted to a Bohr set. To make this notion precise, we define a *local quadratic phase* on a Bohr set B to be a function $\gamma : B \to \mathbb{T}$ such that

$$\gamma(x)\gamma(x+a)^{-1}\gamma(x+b)^{-1}\gamma(x+c)^{-1}\gamma(x+a+b)\gamma(x+a+c)\gamma(x+b+c)\gamma(x+a+b+c)^{-1} = 1 \quad (3.2)$$

whenever all of $x, x+a, x+b, \dots x+a+b+c$ lie in B. To justify this rather bizarre-looking definition, notice that (3.2) is satisfied by a 'global' quadratic phase of the form ω^q, where q is a quadratic function on $\mathbb{Z}_N$ (and conversely any function satisfying (3.2) globally is of the form ω^q). The following inverse theorem for the U^3 norm in $\mathbb{Z}_N$ is due to Green and Tao [31], based on prior work of Gowers [23].

Theorem 3.11. *Let $\delta > 0$, let $C = 2^{24}$ and let $f : \mathbb{Z}_N \to \mathbb{C}$ be a function such that $\|f\|_\infty \leq 1$ and $\|f\|_{U^3} \geq \delta$. Then there exists an element $y \in \mathbb{Z}_N$, a regular Bohr set $B = B(K, \rho)$ with $|K| \leq (2/\delta)^C$ and $\rho \geq (\delta/2)^C$ as well as a quadratic phase function γ defined on $y + B$ such that*

$$\mathbb{E}_{x \in y+B} f(x)\gamma(x) \geq (\delta/2)^C.$$

For the higher-order Gowers norms the situation is even more complicated, and the language of *nilsequences,* inspired by parallel developments in ergodic theory, has become indispensable for describing obstructions to Gowers uniformity in the integers (see [37]).

4. Recent highlights

After this brief foray into the world of the integers we return to studying arithmetic structure in vector spaces over finite fields. In this chapter we shall give an overview of the most consequential developments in recent years, most of which have also had, as predicted by Green in [28], important ramifications in the integer setting.

As we hope to have demonstrated in Chapter 2, algebraic, analytic and combinatorial methods are inextricably intertwined in this subject. Our outline of recent highlights begins with a new technique at the interface of combinatorics and analysis, giving rise to so-called *almost periodicity* results for sum sets, which has driven many of the recent improvements to classical results. We shall discuss its applications to the theory of set addition and the study of linear patterns in dense sets in Sections 4.2 and 4.3, respectively. It has also had quantitative implications for higher-order Fourier analysis, whose recent progress we review in Section 4.4.

4.1. Croot-Sisask almost periodicity.

In 2009 Croot and Sisask introduced a "probabilistic technique for finding almost periods of convolutions" [16], which turned out to have several far-reaching implications in arithmetic combinatorics. Roughly speaking, their main result says that if $A \subseteq \mathbb{F}_p^n$ is a set whose sum set $A + A$ is small, then there exists a dense set X such that for all $x \in X$, the convolution $1_A * 1_A(\cdot)$ and its translate $1_A * 1_A(\cdot + x)$ are almost indistinguishable in the L^2 (or any higher L^p) norm. For obvious reasons, we refer to such a set X as the set of *almost periods* of $A + A$.[11]

In our exposition we shall follow a slight variant [6] of Croot and Sisask's original argument [16]. It proceeds in a more combinatorial fashion without reference to L^p norms, instead relying on a standard Chernoff-type tail estimate, which we now state (see, for example, [69]).

[11] In fact, their result is much more general and even applies to non-abelian groups, but for ease of exposition we shall concentrate on $\mathbb{F}_p^n$ here.

Lemma 4.1. *Let Y be a random variable with $|Y| \leq 1$, and let $\widehat{Y}$ be the empirical average of Y obtained from t samples. Then for any $\gamma > 0$, we have*

$$\mathbb{P}[|\mathbb{E}Y - \widehat{Y}| > \gamma] \leq \exp(-2\gamma^2 t).$$

For simplicity we shall restrict our attention to finite fields of characteristic $p = 2$ in this section, although the more general case only requires minimal modifications. Given subsets $A, B \subseteq G = \mathbb{F}_2^n$ then, we shall be interested in the measure $\rho_{A\to B} : \mathbb{F}_2^n \to [0,1]$ defined by

$$\rho_{A\to B}(y) := \mathbb{P}_{a\in A}\left[y + a \in B\right] = \frac{|(y+A)\cap B|}{|A|} = \mu_A * 1_B(y)$$

whenever $y \in \mathbb{F}_2^n$, where $\mu_A := \frac{|G|}{|A|}1_A$ denotes the so-called *characteristic measure* of the subset A. Notice that $\rho_{A\to B}(y) = 1$ whenever $y + A \subseteq B$, and $\rho_{A\to B}(y) = 0$ whenever $(y + A) \cap B = \emptyset$. We shall also write $a \approx_\epsilon b$ whenever a and b are two real numbers satisfying $|a - b| \leq \epsilon$. With this notation, Theorem 4.2 below is an almost-periodicity result in the spirit of Croot and Sisask [6].

Theorem 4.2. *Let $\epsilon > 0$ and let $A \subseteq \mathbb{F}_2^n$ be a subset satisfying $|A + A| \leq K|A|$. Then for every $t \in \mathbb{N}$ and every subset $B \subseteq \mathbb{F}_2^n$, there exists a set $X \subseteq \mathbb{F}_2^n$ with the following properties.*

1. *The set X is contained in an affine shift of A.*
2. *The size of X is at least $|A|/(2K^{t-1})$.*
3. *For all $x \in X$ and for all subsets $S \subseteq \mathbb{F}_2^n$,*

$$\mathbb{P}_{y\in S}\left[\rho_{A\to B}(y) \approx_{2\epsilon} \rho_{A\to B}(y+x)\right] \geq 1 - 8\frac{|A+B|}{|S|}\cdot\exp\left(-2\epsilon^2 t\right).$$

Proof: Fix $t \in \mathbb{N}$, and for simplicity of notation set $\rho(y) := \rho_{A\to B}(y)$. For any t-tuple $a = (a_1, \ldots, a_t) \in (\mathbb{F}_2^n)^t$, define the *$a$-estimator of ρ* to be the function $\widehat{\rho}_a : \mathbb{F}_2^n \to [0,1]$ which, for each $y \in \mathbb{F}_2^n$, takes the value

$$\widehat{\rho}_a(y) := \frac{|\{y + a_i \in B : i = 1, \ldots, t\}|}{t}.$$

We say that $a \in (\mathbb{F}_2^n)^t$ is an *ϵ-good estimator for y* if $\rho(y) \approx_\epsilon \widehat{\rho}_a(y)$. We call $a \in (\mathbb{F}_2^n)^t$ an *(ϵ,δ)-good estimator* if a is an ϵ-good estimator for all but a δ-fraction of $A + B$, i.e.,

$$\mathbb{P}_{y\in A+B}\left[\rho(y) \not\approx_\epsilon \widehat{\rho}_a(y)\right] < \delta.$$

Fix $y \in \mathbb{F}_2^n$. The first step towards constructing the set X of almost periods is to show that most sample vectors $a \in (\mathbb{F}_2^n)^t$ are ϵ-good for a given y, provided that t, the sample size, is large enough with respect to $1/\epsilon$. To see this, let Y_i be the indicator random variable for the event $y + a_i \in B$ where a_i is chosen uniformly at random from A. Then $\widehat{\rho}_a(y) = \frac{1}{t}\sum_{i=1}^t Y_i$ is the average of t i.i.d. indicator random variables each of mean $\rho(y)$, so Lemma 4.1 implies that for each $y \in \mathbb{F}_2^n$,

$$\mathbb{P}_{a\in A^t}\left[\rho(y) \not\approx_\epsilon \widehat{\rho}_a(y)\right] \leq 2\exp\left(-2\epsilon^2 t\right). \tag{4.1}$$

Setting $\delta := 2\exp(-2\epsilon^2 t)$ and letting Z_a be the random variable measuring the fraction of $y \in A+B$ for which a is an ϵ-good estimator, that is,

$$Z_a := \mathbb{P}_{y\in A+B}\left[\rho(y) \approx_\epsilon \widehat{\rho}_a(y)\right],$$

we conclude from (4.1) via linearity of expectation that

$$\mathbb{E}_{a\in A^t}\left[Z_a\right] \geq 1-\delta.$$

It now follows from Markov's inequality that at least half the sample vectors $a \in A^t$ are ϵ-good estimators for all but a 2δ-fraction of $y \in A+B$. In other words, the set $G_{\epsilon,2\delta} \subseteq A^t$ of $(\epsilon, 2\delta)$-good estimators for ρ, defined by

$$G_{\epsilon,2\delta} := \{a \in A^t : \mathbb{P}_{y\in A+B}\left[\rho(y) \approx_\epsilon \widehat{\rho}_a(y)\right] \geq 1-2\delta\},$$

satisfies $|G_{\epsilon,2\delta}| \geq |A|^t/2$.

To obtain X we partition $G_{\epsilon,2\delta}$ as follows. Define a map $\phi : A^t \to \{0\}\times(A+A)^{t-1}$ by shifting the sequence $a = (a_1,\ldots,a_t) \in A^t$ by the first element a_1, that is, let

$$\phi(a) := (a_1+a_1, a_2+a_1, \ldots, a_t+a_1).$$

Since ϕ maps the set $G_{\epsilon,2\delta}$ of size at least $|A|^t/2$ into a set of size $|A+A|^{t-1} \leq (K|A|)^{t-1}$, there exists a subset $G^*_{\epsilon,2\delta} \subseteq G_{\epsilon,2\delta}$ of size

$$|G^*_{\epsilon,2\delta}| \geq \frac{|A|^t}{2K^{t-1}|A|^{t-1}} = \frac{|A|}{2K^{t-1}}, \tag{4.2}$$

all of whose elements map to the same value $b \in \{0\}\times(2A)^{t-1}$ under ϕ. Finally, fix an arbitrary $a' = (a'_1,\ldots,a'_t) \in G^*_{\epsilon,2\delta}$ and set

$$X := \{a'_1 + a_1 : (a_1,\ldots,a_t) \in G^*_{\epsilon,2\delta}\}.$$

To complete the proof we need to show that X has the three properties listed in the statement of the theorem. First note that by definition, $X \subseteq a'_1 + A$, so X is indeed contained in an affine shift of A. Secondly, observe that by construction of the set $G^*_{\epsilon,2\delta}$ the map $G^*_{\epsilon,2\delta} \to X$ given by $(a_1,\ldots,a_t) \mapsto a'_1+a_1$ is invertible, and hence $|X| = |G_{\epsilon,2\delta^*}|$ is bounded below by (4.2) as desired.

Last but not least, suppose that we have $x = a'_1 + a_1 \in X$, where a_1 is the first element of an $(\epsilon,2\delta)$-good estimator $a = (a_1,\ldots,a_t) \in G^*_{\epsilon,2\delta}$. We claim that $a'+x = a$. Indeed, the definition of $G^*_{\epsilon,2\delta}$ implies that $\phi(a') = \phi(a)$, so that for all $i = 1,\ldots,t$, we have $a_i + a_1 = a'_i + a'_1$, or, in other words, $a' + x = a' + (a'_1 + a_1) = a$.

Now recalling that a' is an $(\epsilon,2\delta)$-good estimator, we know that for all but a 2δ-fraction of $y \in A+B$,

$$\rho(y) \approx_\epsilon \widehat{\rho}_{a'}(y). \tag{4.3}$$

This inequality also holds for all $y \notin A+B$ since in this case both sides are trivially zero. Hence we have that (4.3) holds for all but a $(2\delta|A+B|/|S|)$-fraction of $y \in S$. Similarly, since a is an $(\epsilon,2\delta)$-good estimator, we have that for any x,

$$\rho(y+x) \approx_\epsilon \widehat{\rho}_a(y+x) \tag{4.4}$$

for all but a $(2\delta|A+B|/|x+S|)$-fraction of $y \in S$. Using a union bound and the fact that $|S+x| = |S|$, we find that for all but a $(4\delta|A+B|/|S|)$-fraction of $y \in S$ both (4.3) and (4.4) hold. For such y we have

$$\rho(y) \approx_\epsilon \widehat{\rho}_{a'}(y) = \widehat{\rho}_{a'+x}(y+x) = \widehat{\rho}_a(y+x) \approx_\epsilon \rho(y+x),$$

completing the proof of the theorem. □

4.2. **A quasipolynomial Bogolyubov lemma.** Croot and Sisask's original applications of Theorem 4.2 included non-commutative analogues of the Freiman-Ruzsa theory of set addition (cf. Section 2.2), a low-density version of a structure theorem on long arithmetic progressions in sum sets inside $\mathbb{Z}_N$ by Bourgain [13] and Green [27], as well as a new probabilistic proof of Roth's theorem (cf. Theorem 3.4).

Subsequently, Sanders quickly realised and ingeniously exploited the fact that almost periodicity could be used to make a substantial leap in the bounds of Bogolyubov's lemma (Proposition 1.3). In [60] he proved (a much more general version of) the following result.

Theorem 4.3. *Let $A \subseteq \mathbb{F}_2^n$ be a subset of density α. Then $2A - 2A$ contains a subspace V of codimension at most $O(\log^4 \alpha^{-1})$.*

In the remainder of this section we give a sketch of the argument, again following the approach in [6]. First, note that by an inductive application of Theorem 4.2, using nothing more than the triangle inequality, one can prove the following iterated version.

Corollary 4.4. *Let $\epsilon > 0$ and let $A \subseteq \mathbb{F}_2^n$ be a subset satisfying $|A+A| \leq K|A|$. Then for every $t \in \mathbb{N}$ and every subset $B \subseteq \mathbb{F}_2^n$, there exists a set $X \subseteq \mathbb{F}_2^n$ with the following properties.*

1. *The set X is contained in an affine shift of A.*
2. *The size of X is at least $|A|/(2K^{t-1})$.*
3. *For any $\ell \in \mathbb{N}$, all $x_1, \dots, x_\ell \in X$ and all subsets $S \subseteq \mathbb{F}_2^n$,*

$$\mathbb{P}_{y \in S}\left[\rho_{A\to B}(y) \approx_{2\epsilon\ell} \rho_{A\to B}(y + x_1 + \ldots + x_\ell)\right] \geq 1 - 8\ell\frac{|A+B|}{|S|} \cdot \exp\left(-2\epsilon^2 t\right).$$

But we have already seen that iterated sum sets are highly structured, so one might hope to be able to convert the periodicity statement above from one involving iterated sum sets into one in which the set of periods is in fact a subspace. Moreover, observe that any subset $A \subseteq \mathbb{F}_2^n$ of density α trivially has doubling at most α^{-1}, so one can use the above corollary with α^{-1} in place of K, omitting the small sum set condition from the statement altogether.

Corollary 4.5. *Let $A \subseteq \mathbb{F}_2^n$ be a subset of density α. Then for every integer t and set $B \subseteq \mathbb{F}_2^n$ there exists a subspace V of*

$$\operatorname{codim}(V) \leq 32\log(2/\alpha^t)$$

with the property that for all $v \in V$, for all subsets $S \subseteq \mathbb{F}_2^n$, for every $\epsilon, \eta > 0$ and for every integer ℓ,

$$\mathbb{P}_{y\in S}\left[\rho_{A\to B}(y) \approx_{\epsilon'} \rho_{A\to B}(y+v)\right] \geq 1 - 16\frac{\ell}{\eta}\frac{|A+B|}{|S|}\cdot \exp\left(-2\epsilon^2 t\right), \tag{4.5}$$

where $\epsilon' := 4\epsilon\ell + 2\eta + 2^{-\ell}(|B|/|A|)^{1/2}$.

In order to prove (4.5) one first shows, using Corollary 4.4, a simple averaging argument and the triangle inequality, that there exists a set X of density at least $\alpha/(2K^{t-1}) \geq \alpha^t/2$ such that for most $y \in S$,[12]

$$\mathbb{E}_{x_1,\ldots,x_\ell\in X}\rho_{A\to B}(y + x_1 + \ldots + x_\ell) \approx_{2\epsilon\ell+\eta} \rho_{A\to B}(y), \tag{4.6}$$

and similarly that for any $v \in \mathbb{F}_2^n$ and most $y \in S$,

$$\mathbb{E}_{x_1,\ldots,x_\ell\in X}\rho_{A\to B}(y + v + x_1 + \ldots + x_\ell) \approx_{2\epsilon\ell+\eta} \rho_{A\to B}(y+v). \tag{4.7}$$

In order to complete the proof of Corollary 4.5, we now only need to connect the left-hand sides of (4.6) and (4.7), respectively, for elements v of a suitable subspace $V \leqslant \mathbb{F}_2^n$. This is achieved by Lemma 4.6 below, whose proof is strongly reminiscent of that of Proposition 1.3 in the sense that the desired subspace is taken to be the orthogonal complement of the set of large Fourier coefficients of X.

Lemma 4.6. *Let $X \subseteq \mathbb{F}_2^n$ be as above, and set $V := \mathrm{Spec}_{1/2}(X)^\perp$. Then*

$$\mathrm{codim}(V) \leq 32\log(2/\alpha^t),$$

and for all $y \in \mathbb{F}_2^n$ and $v \in V$, we have

$$\mathbb{E}_{x_1,\ldots,x_\ell\in X}\rho_{A\to B}(y + x_1 + \ldots + x_\ell) \approx_{\epsilon''} \mathbb{E}_{x_1,\ldots,x_\ell\in X}\rho_{A\to B}(y + v + x_1 + \ldots + x_\ell), \tag{4.8}$$

where $\epsilon'' := 2^{-\ell}(|B|/|A|)^{1/2}$.

Proof: Since the density of X is at least $\alpha^t/2$, Chang's theorem (Theorem 1.2) implies that

$$\mathrm{codim}(V) \leq 8\cdot(1/2)^{-2}\log(\alpha^t/2)^{-1} = 32\log(2/\alpha^t).$$

In order to prove (4.8), note that we can write the difference between the two sides as

$$(\mu_X)^{*\ell} * \mu_A * 1_B(y) - (\mu_X)^{*\ell} * \mu_A * 1_B(y+v),$$

which in terms of the Fourier transform equals

$$\sum_{t\in\mathbb{F}_2^n} \widehat{\mu_A}(t)\cdot\widehat{\mu_X}(t)^\ell\cdot\widehat{1_B}(t)\cdot\left((-1)^{y\cdot t} - (-1)^{(y+v)\cdot t}\right).$$

This expression in turn is bounded above in absolute value by

$$\sum_{t\notin V^\perp} |\widehat{\mu_A}(t)|\cdot|\widehat{\mu_X}(t)|^\ell\cdot|\widehat{1_B}(t)|\cdot\left|1-(-1)^{v\cdot t}\right| \leq 2^{-\ell}\sum_{t\notin V^\perp}|\widehat{\mu_A}(t)|\cdot|\widehat{1_B}(t)|.$$

The Cauchy-Schwarz inequality followed by an application of Parseval's identity yields the final bound of $2^{-\ell}(|B|/|A|)^{1/2}$. □

[12]The proportion of such $y \in S$ depends inversely on η. For details see [6].

Equipped with Corollary 4.5 we are now in a position to complete the proof of Sanders's quasipolynomial Bogolyubov bound.

Proof of Theorem 4.3: On application of Corollary 4.5 with

$$B := A + A,\ S := A,\ \ell := \log(30^2/\alpha)/2,\ \eta := 1/60,\ \epsilon := 1/(120\ell) \text{ and } t := O(\log^3(1/\alpha))$$

we obtain a subspace $V \leqslant \mathbb{F}_2^n$ of $\mathrm{codim}(V) = O(\log^4(1/\alpha))$ which has the property that for all $v \in V$,

$$\mathbb{P}_{a\in A}\left[\rho_{A\to 2A}(a) \approx_{\epsilon'} \rho_{A\to 2A}(a+v)\right] \geq 1 - 16\frac{\ell}{\eta}\frac{|3A|}{|A|}\cdot \exp\left(-2\epsilon^2 t\right) \geq 0.9,$$

where $\epsilon' = 4\epsilon\ell + 2\eta + 2^{-\ell}(|A+A|/|A|)^{1/2} \leq 1/30 + 1/30 + (\alpha^{1/2}/30)\cdot\alpha^{-1/2} \leq 1/10$. Now since $\rho_{A\to 2A}(a) = 1$ for all $a \in A$, this implies that for all $v \in V$,

$$\mathbb{P}_{a\in A}\left[\rho_{A\to 2A}(a+v) \geq 0.9\right] \geq 0.9.$$

Recalling the definition of $\rho_{A\to B}$, it follows that for all $v \in V$,

$$\mathbb{P}_{a,a'\in A}[a + a' + v \in 2A] \geq 0.9^2 = 0.81.$$

By averaging there therefore exists a pair $a, a' \in A$ such that

$$\mathbb{P}_{v\in V}[a + a' + v \in 2A] \geq 0.81,$$

or equivalently, a choice of $a, a' \in A$ for which $|V \cap (a + a' + 2A)\ | \geq 0.81|V|$. But if a subset $E \subseteq \mathbb{F}_2^n$ is such that $|V \cap E| > \frac{1}{2}|V|$, then $E + E \supseteq V$, so we conclude that $V \subseteq 2(a + a' + A + A) \subseteq 4A$, which terminates the proof. □

Using Theorem 4.3, Sanders was additionally able to make substantial progress towards the Polynomial Freiman-Ruzsa Conjecture (Conjecture 2.10), by showing that the constants $C_1(K)$ and $C_2(K)$ defined there can be taken to be quasipolynomial and polynomial in K, respectively. Moreover, these results hold in a much more general setting, for details of which we refer the reader to the survey [59].

4.3. **Solutions to translation-invariant equations.** In addition to the applications already mentioned, the techniques discussed in this chapter have also transformed the counting of solutions to translation-invariant equations. The most striking instance is another result of Sanders [58] in the integers, which we made brief reference to in Chapter 3 and which improves the bound for 3-term progression free subsets of $\{1, \dots, N\}$ to

$$|A| \ll \frac{(\log\log N)^6 N}{\log N}.$$

To fully appreciate the strength of this result, observe that it comes very close to establishing a Roth-type theorem for dense subsets of the primes, which have density $1/\log N$.[13] However, it is still a long way from matching the best known lower bound, which is of the form $\exp(-c(\log N)^{1/2})$ for some constant $c > 0$, proved by Behrend [4, 18, 39] nearly 70 years ago.

[13]Very recently, Bloom [8] managed to improve this result by a factor of $(\log\log N)^2$ using a completely different technique, namely a strengthening of Chang's theorem.

In the finite field model setting, sadly, these rather sophisticated methods do not lead to improvements upon the bound presented in Theorem 2.2. It turns out that in all its simplicity, Meshulam's argument is already surprisingly powerful. However, by a very careful and highly complex analysis of the large spectrum of a set, Bateman and Katz [3] were recently able to gain a small but positive exponent in Theorem 2.2.

Theorem 4.7. *There exists $c > 0$ with the following property. Let $A \subseteq \mathbb{F}_3^n$ be a set containing no non-trivial 3-term arithmetic progressions. Then*

$$|A| \ll \frac{N}{(\log N)^{1+c}}.$$

The constant c can be made explicit, but at the time of writing the argument is not sufficiently well understood (at least by the author of this survey) to warrant an exposition of the material.[14] Again, we must compare this bound to the largest known example of a 3-term progression free subset of $\mathbb{F}_3^n$. The current record in this regard is held by Edel [17], who constructed a set of size $3^{0.725851n}$. Green conjectured ([28], Conjecture 4.3) that this construction is close to best possible in the sense that there exists a constant $\delta > 0$ such that the size of any 3-term progression free subset of $\mathbb{F}_3^n$ is bounded above by $(3-\delta)^n$, but this conjecture remains wide open.

However, it turns out that if one considers translation-invariant equations with only a few more variables, one can obtain a marked improvement on Meshulam's bound using Croot-Sisask almost periodicity in the guise of Sanders's quasipolynomial Bogolyubov lemma (Theorem 4.3). The following is a neat result of Schoen and Shkredov [63], whose proof we are able to give in full detail.

Theorem 4.8. *Let $p > 5$ be a prime, and let $A \subseteq G = \mathbb{F}_p^n$. Suppose that A contains no non-trivial solutions to the equation*

$$5y = x_1 + x_2 + x_3 + x_4 + x_5, \tag{4.9}$$

that is, no solution $(y, x_1, x_2, x_3, x_4, x_5) \in A^6$ such that $y \neq x_i$ for some $i = 1, 2, ..., 5$. Then

$$|A| \leq \exp(-c(\log |G|)^{1/5})|G|.$$

PROOF: As usual, let α denote the density of A in $\mathbb{F}_p^n$. The main idea is to partition the set A into two (disjoint) sets A_1, A_2 with the cardinalities of A_1 and A_2 being as equal as possible. By averaging there exists $z \in \mathbb{F}_p^n$ such that

$$|A_1 \cap (z - A_2)| \geq \alpha^2 |G|/4.$$

Set $B := A_1 \cap (z - A_2)$, which by the above is of density $\gg \alpha^2$. By Theorem 4.3 therefore, there exists a subspace V of codimension at most $O(\log^4(\alpha^{-1}))$ with the property that $2B - 2B \supseteq V$, and hence

$$2z + V \subseteq 2z + 2B - 2B \subseteq 2A_1 + 2A_2.$$

[14]The reader may be interested in exploring the aforementioned very recent work of Bloom [8], which exploits ideas along similar lines.

We claim that as a consequence,

$$(5 \cdot A - A) \cap (2z + V) = \emptyset. \tag{4.10}$$

For suppose that there existed $x, y \in A$ such that $5y - x \in 2z + V$. Then we would be able to write $5y - x$ as a sum of two elements in A_1 and two elements in A_2, which (since these two sets are disjoint) would yield a non-trivial solution to (4.9).

It follows from (4.10) that for any fixed $w \in \mathbb{F}_p^n$, at most one of the sets $A \cap (w + V)$ and $5 \cdot A \cap (w + 2z + V)$ can be non-empty, from which we deduce that

$$2|A| = \sum_{w \in V^\perp} |A \cap (w + V)| + |5 \cdot A \cap (w + 2z + V)| \leq |V^\perp| \sup_w |A \cap (w + V)|.$$

We have thus obtained a considerable density increase of A on some coset of V, i.e. an element $w \in \mathbb{F}_p^n$ such that

$$|A \cap (w + V)| \geq 2\alpha |V|.$$

Let $A_1 := (A - w) \cap V$, and observe that this set of density at least 2α inside $V_1 := V$ still contains no non-trivial solutions to (4.9). After t iterations of this argument, we obtain a subspace V_t of codimension at most $O(t \log^4(\alpha^{-1}))$ satisfying

$$|(A - w_t) \cap V_t| \geq 2^t \alpha |V_t|$$

for some $w_t \in \mathbb{F}_p^n$. Arguing as in the proof of Theorem 2.2 yields the stated bound on the density α. □

By replacing subspaces with Bohr sets (as described in Chapter 3) in the above argument, Schoen and Shkredov gave a similar bound in the integers. They thereby obtained, for the first time, an upper bound on the number of solutions to a translation-invariant equation that lies within arm's reach of the lower bound. Further progress in this direction in the not too distant future seems likely.

4.4. **Progress on higher-order polynomial structure.** At the time of Green's survey [28] the field of higher-order Fourier analysis was still in its very early infancy, the foundations having been laid by Gowers [23, 24] only a few years prior. Since then significant progress has been made, ranging from quantitative improvements to quadratic structure theorems, to inverse theorem for higher U^k norms, through to steps towards a more general and abstract theory of higher-order Fourier analysis. In this section we shall only give some pointers to the most important developments in the context of the finite field model.

As discussed in Section 2.3, the Freiman-Ruzsa theorem (Theorem 2.7) and the inverse theorem for the U^3 norm (Theorem 2.16) are intimately connected, and have in fact been shown to be quantitatively equivalent [35]. As a result of Sanders's improvement to the bounds in the Freiman-Ruzsa theorem (see Section 4.2) we therefore now have a quasipolynomial dependence of the constant $c(\delta)$ on the uniformity parameter δ in Theorem 2.16.

The situation is much less satisfactory for a higher-order inverse theory, which has been developed from various viewpoints over the past decade. Notable contributors include Host and Kra [42] and Ziegler [75] from the ergodic theoretic perspective, Szegedy [65] with a novel non-standard analysis approach as well as Bergelson, Green, Tao and Ziegler

[76, 70, 71, 38]. Since it is most relevant for our purposes, we shall focus on the latter body of work in our discussion below.

Conjecturally, an inverse theorem for the U^k norms for $k > 3$ seems straightforward at first sight: loosely speaking, functions with large U^k norm ought to correlate with polynomial phase functions of degree $k-1$.[15] However, even in the supposedly simple finite field model, extending Theorem 2.16 to higher U^k norms turned out to be less than plain sailing. Somewhat surprisingly, although well-understood with hindsight, the naive conjecture for the U^4 norm, namely that every function with large U^4 norm correlates with a cubic phase polynomial, was shown to be false in characteristic 2 [47, 34].

Theorem 4.9. *Let S_4 be the quartic symmetric polynomial on $\mathbb{F}_2^n$, that is,*

$$S_4(x_1, \dots, x_n) := \sum_{1 \le i_1 < \dots < i_4 \le n} x_{i_1} x_{i_2} x_{i_3} x_{i_4}.$$

Then the function defined by $f(x) = (-1)^{S_4(x)}$ satisfies $\|f\|_{U^4}^{16} = \frac{1}{8} + O(2^{-n/2})$, but there exists an absolute constant $c > 0$ such that

$$|\mathbb{E}_x f(x)(-1)^{q(x)}| \ll 2^{-cn}$$

for any cubic polynomial q.

Doubts as to the veracity of a higher-order inverse theory ensued, but it quickly emerged that its principal proposition could nevertheless be salvaged. The key was to adjust the notion of a "phase polynomial" in such a way that Theorem 4.9 no longer presented a counterexample.

In the original inverse conjecture for the U^k norm over $\mathbb{F}_p^n$ (see for example [67]) the correlating phase polynomial $\phi : \mathbb{F}_p^n \to \mathbb{T}$ was assumed to be of the form $\phi = \omega^P$, where ω is a pth root of unity and P is a polynomial of degree at most $k-1$ in the sense that it is a function $P : \mathbb{F}_p^n \to \mathbb{F}_p$ satisfying

$$\Delta_{h_1}^+ \Delta_{h_2}^+ \dots \Delta_{h_k}^+ P(x) = 0 \tag{4.11}$$

for all $h_1, \dots, h_k, x \in \mathbb{F}_p^n$, where $\Delta_h^+ f(x) = f(x+h) - f(x)$. Instead, it was observed by Tao and Ziegler [70] that one needs to consider a wider class of polynomial phases of degree $k-1$, namely functions $\psi : \mathbb{F}_p^n \to \mathbb{C}$ satisfying the multiplicative derivative condition

$$\Delta_{h_1} \Delta_{h_2} \dots \Delta_{h_k} \psi(x) = 1 \tag{4.12}$$

for all $h_1, \dots, h_k, x \in \mathbb{F}_p^n$, where $\Delta_h \psi(x) = \psi(x+h)\overline{\psi(x)}$. The crux is to observe that while for $p \ge k$ any phase polynomial ψ satisfying (4.12) is in fact (up to a constant phase) of the form ω^P for a function $P : \mathbb{F}_p^n \to \mathbb{F}_p$ satisfying (4.11), this is not so when $p < k$. Rather, it can be shown ([71], Lemma 1.7) that ψ must be of the form $e^{2\pi i P'}$ for some function

[15] We already mentioned at the end of Chapter 3 that in the integers this is a problematic statement even for $k = 3$, and that any correlation has to be defined locally with respect to a Bohr set (or in the language of nilsequences). For higher values of k this situation only worsens. For example, when $k = 4$, we would expect to obtain a cubic phase defined on the approximate simultaneous level set of a bunch of 'quadratic characters'. Despite this added complexity, very general results to this effect are now known [40, 37, 38].

$P' : \mathbb{F}_p^n \to \mathbb{T}$ satisfying (4.11) which is no longer restricted to taking values in the group of pth roots of unity $\frac{1}{p}\mathbb{Z}/\mathbb{Z}$, but rather turns out to be confined to a coset of the slightly larger

$$\frac{1}{p^{\lfloor \frac{k-2}{p-1} \rfloor + 1}}\mathbb{Z}/\mathbb{Z}.$$

As an example, consider the function $P' : \mathbb{F}_2^n \to \mathbb{T}$ defined by $P'(x_1, \dots, x_n) = |x|/8 \mod 1$, where $|x| = \sum_{j=1}^n x_j$ denotes the number of components equal to 1, which is a (non-classical) cubic polynomial satisfying (4.11). Moreover, it can be checked that $\psi = e^{2\pi i P'}$ has non-trivial correlation with the function $f = (-1)^{S_4}$ defined in Theorem 4.9.

With this new definition in mind, Bergelson, Tao and Ziegler [76, 70, 71] were able to prove the following *inverse theorem for the U^k norm.*

Theorem 4.10. *Let $k \geq 2$ be an integer and let p be a prime. Then for every $\delta > 0$ there exists $c = c(\delta, k, p)$ such that the following holds. Suppose that $f : \mathbb{F}_p^n \to \mathbb{C}$ is a function satisfying $\|f\|_\infty \leq 1$ and $\|f\|_{U^k} \geq \delta$. Then f has correlation at least c with a non-classical phase polynomial of degree at most $k-1$, in other words, there exists $P' : \mathbb{F}_p^n \to \mathbb{T}$ satisfying (4.11) such that*

$$\mathbb{E}_x f(x) e^{2\pi i P'(x)} \geq c.$$

Frustratingly, despite the statement of the theorem being purely combinatorial, the proof proceeds via an infinitary recurrence result in ergodic theory [76], and then uses a so-called transference principle (in the spirit of Furstenberg) to pass to the finite field model setting [70]. The ergodic-theoretic arguments used are by their very nature not quantitative (and while some bounds could in principle, with some effort, be extracted, they would be unimaginably poor). Providing a direct analytic (combinatorial) proof of Theorem 4.10 remains an central open problem, whose successful resolution would have a number of additional benefits in the form of a quantitative Szemerédi theorem for long progressions in finite fields, as well as applications to error-correcting codes and other areas of theoretical computer science.

Moreover, in spite of the general form of Theorem 4.10, the case of low characteristic remains poorly understood. Indeed, the proof in the case $p \leq k$ [71] is a 67-page tour de force, bootstrapping a weaker version of the result from an earlier paper of the authors, in which the degree of the correlating polynomial was not necessarily $k-1$ but rather bounded above by a function of k. New ideas are needed to successfully tackle these questions.

In the past decade a significant amount of work has also been done towards justifying the labels *quadratic* and *higher-order Fourier analysis*. The strength of the Fourier transform lies in the fact that we can decompose any function into a weighted sum of characters (or linear phase functions). Those phases with small coefficients can often (but not always) be neglected in applications. A quadratic analogue of such a decomposition would be the statement that any bounded function can be written as a weighted sum of *quadratic phase functions*, plus an error term which is *quadratically uniform* (i.e. small in the U^3 norm). Since the set of quadratic phase functions does not form an orthonormal basis, however, this is a highly non-trivial proposition. One of the most concrete and accessible

statements to this effect is the following result [26].[16] In applications, much more sophisticated decompositions and higher-order versions are often needed, see [26, 36] and [25], respectively.

Theorem 4.11. *Let $f : \mathbb{F}_p^n \to \mathbb{C}$ be a function such that $\|f\|_2 \leq 1$. Then for every $\delta > 0$ and $\eta > 0$ there exists $M = M(\eta, \delta, p)$ such that f has a decomposition of the form*

$$f(x) = \sum_i \lambda_i \omega^{q_i(x)} + g(x) + h(x), \tag{4.13}$$

where the q_i are quadratic forms on $\mathbb{F}_p^n$, and

$$\eta^{-1}\|g\|_1 + \delta^{-1}\|h\|_{U^3} + M^{-1}\sum_i |\lambda_i| \leq 1.$$

Finally, if quadratic Fourier analysis is to take its rightful place alongside traditional (linear) Fourier analysis, we need to be able to answer another natural question: can we efficiently compute the 'quadratic Fourier coefficients' of a given function? In joint work Tulsiani and the author [72] gave a probabilistic algorithm that computes the coefficients λ_i corresponding to the large quadratic characters ω^{q_i} in the decomposition (4.13) above. This has applications to list decoding and self-correction procedures for Reed-Muller codes.

5. Limitations of the finite field model and new horizons

As we hope to have demonstrated in this article, the finite field model has proven extremely powerful in attacking a range of number-theoretic and combinatorial problems in the integers. Many striking successes of the recent past have had to be omitted from this survey, which would have further emphasised this point. Because of its undeniable elegance, the finite field model now attracts much interest in its own right, as well as being of enduring importance to computer scientists. And it remains the point of entry to an oftentimes technical subject for many a graduate student.

However, as the appreciation for the model's strengths has grown in recent years, so has our understanding of its weaknesses and limitations. We have hinted at some of these throughout the article, but let us summarise the most significant ones here.

First, while arguments often do transfer from the finite field model to the integers, when it comes to pushing the quantitative bounds to the limit of our capabilities in either setting, the two often differ substantially from each other. As we have seen, this is the case for Roth's and Meshulam's theorem, for example, where the ultimate answers (and the techniques used to obtain them) may well turn out to be fundamentally different. Another case in point is the development of higher-order Fourier analysis, where there is some evidence that a U^4 inverse theorem may actually be strictly harder in the finite field model than in the integer case (see Chapter 1.5 of [68]).

A second issue that makes the finite field model challenging to work with, in a way that the integers are not, is that additional complications arise when the characteristic of the

[16] Alternative formulations, borrowing the language of *factors* from ergodic theory, can be found in [29] (Proposition 3.7), for example.

underlying field is very small. Specifically, we have seen that the problem of determining the structure of functions on $\mathbb{F}_p^n$ with large U^k norm has thrown up unexpected difficulties when $p < k$. Small characteristic also means that we lose access to certain types of geometric arguments, as is the case, for example, in the lower bounds for Meshulam's theorem.

Finally, a rather important criticism of the model is that additive groups of the form $\mathbb{F}_p^n$ are wholly unsuitable for studying problems involving multiplicative structure, which are ubiquitous in arithmetic combinatorics. Examples include Sarközy's theorem [62], which is a version of Roth's theorem in which the common difference of the arithmetic progression is restricted to the set of squares, or the famously difficult sum-product conjecture of Erdős and Szemerédi [69], which asserts that addition and multiplication are incompatible in a strong sense.

An alternative model that appears to hold promise for tackling such problems is the *function field setting*. Let q be an odd prime power, and consider the set of polynomials over $\mathbb{F}_q$ of degree less than n. This set is easily seen to form a group under addition, and can be endowed with a field structure isomorphic to $\mathbb{F}_{q^n}$ by performing multiplication modulo a fixed irreducible polynomial of degree n. When we speak of the function field model we henceforth refer to $\mathbb{F}_q[t]_{<n} := \mathbb{F}_q[t]/(p(t))$, where p is a fixed irreducible polynomial of degree n. Just like in the traditional finite field (vector space) model $\mathbb{F}_p^n$, the analogues of Bohr sets in the function field model $\mathbb{F}_q[t]_{<n}$ are additively closed, but we now have the operation of multiplication at our disposal. A prime number (in the traditional number-theoretic sense) corresponds to an irreducible polynomial, and problems posed in the initial segment $\{1, \dots, N\}$ of the integers have analogues in $\mathbb{F}_q[t]_{<n}$ where $N \sim q^n$.

This analogy with the integers, which in fact runs much deeper than we have outlined here, has been known in analytic number theory circles for a rather long time (see for example [22]), and initial exploration of the model is starting to take place in arithmetic combinatorics. One of the earliest results is due to Lê Thái [44], who proved a function-field version of the celebrated Green-Tao theorem on long arithmetic progressions in the primes [32] by showing that given any integer $k > 0$, it is possible to find two polynomials $f, g \in \mathbb{F}_q[t]$, $g \neq 0$, such that the polynomials $f + pg$ are all irreducible, where p runs through all polynomials $p \in \mathbb{F}_q[t]$ of degree less than k. In the context of sum-product estimates, Bloom and Jones [9] were able to show that for any subset $A \subseteq \mathbb{F}_q[t]$ and any $\epsilon > 0$, one has

$$\max\{|A + A|, |A \cdot A|\} \gg_{\epsilon,q} |A|^{1+\frac{1}{5}+\epsilon},$$

where the exponent $1/5$ beats what is known in the case of (large) finite fields $\mathbb{F}_p$. Furthermore, very recent work of Bloom [7, 8] provides a strong quantitative analogue of Roth's theorem in function fields. Specifically, he proved that whenever $A \subseteq \mathbb{F}_q[t]_{<n}$ is a set not containing any non-trivial solutions to the equation $c_1x_1 + c_2x_2 + c_3x_3 = 0$, where the coefficients $c_1, c_2, c_3 \in \mathbb{F}_q[t] \setminus \{0\}$ satisfy $c_1 + c_2 + c_3 = 0$, then A must satisfy

$$|A| \ll \frac{(\log n)^2}{n} q^n,$$

which again is stronger than what is known in the integers.

The function field setting also holds hope for perhaps resolving some of the aforementioned issues of small characteristic. The obstacles posed by repeated differencing in small fields have, for instance, been overcome in the context of Waring's problem in the function field setting [45]. It remains to be seen whether similar techniques can be made to bear fruit in arithmetic combinatorics.

References

[1] Noga Alon. Testing subgraphs in large graphs. *Random Structures Algorithms*, 21(3-4):359–370, 2002.

[2] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing Low-Degree Polynomials over GF(2). 2764:1101–1117, 2003.

[3] Michael Bateman and Nets Hawk Katz. New bounds on cap sets. *J. Amer. Math. Soc.*, 25(2):585–613, 2012.

[4] F. A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci. U. S. A.*, 32:331–332, 1946.

[5] Eli Ben-Sasson, Shachar Lovett, and N Ron-Zewi. An Additive Combinatorics Approach Relating Rank to Communication Complexity. *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*, pages 177–186, 2012.

[6] Eli Ben-Sasson, Noga Ron-Zewi, Madhur Tulsiani, and Julia Wolf. Sampling-based proofs of almost-periodicity results and algorithmic applications. *ICALP 2014*, 2014.

[7] Thomas F. Bloom. Translation invariant equations and the method of Sanders. *Bull. Lond. Math. Soc.*, 44(5):1050–1067, October 2012.

[8] Thomas F. Bloom. A quantitative improvement for Roth's theorem on arithmetic progressions. *arXiv*, May 2014.

[9] Thomas F. Bloom and Timothy G F Jones. A Sum–Product Theorem in Function Fields. *Int. Math. Res. Not. IMRN*, page rnt125, June 2013.

[10] A. Bogdanov and Emanuele Viola. Pseudorandom Bits for Polynomials. In *Foundations of Computer Science, 2007. FOCS '07. 48th Annual IEEE Symposium on*, pages 41–51, October 2007.

[11] N. Bogolioùboff. Sur quelques propriétés arithmétiques des presque-périodes. *Ann. Chaire Phys. Math. Kiev*, 4:185–205, 1939.

[12] Jean Bourgain. On triples in arithmetic progression. *Geom. Funct. Anal.*, 9(5):968–984, 1999.

[13] Jean Bourgain. On arithmetic progressions in sums of sets of integers. In A Baker, B Bollobás, and A Hajnal, editors, *A Tribute to Paul Erdős*, pages 105–110. Cambridge University Press, Cambridge, 2009.

[14] Emmanuel Breuillard, Ben J. Green, and Terence Tao. The structure of approximate groups. *Publ.math.IHES*, 116(1):115–221, October 2012.

[15] Mei-Chu Chang. A polynomial bound in Freiman's theorem. *Duke Math. J.*, 113(3):399–419, 2002.

[16] Ernie Croot and Olof Sisask. A probabilistic technique for finding almost-periods of convolutions. *Geom. Funct. Anal.*, 20(6):1367–1396, 2010.

[17] Yves Edel. Extensions of generalized product caps. *Des. Codes Cryptogr.*, 31(1):5–14, 2004.

[18] Michael Elkin. An improved construction of progression-free sets. *Israel J. Math.*, 184:93–128, 2011.

[19] Chaim Even-Zohar. On Sums of Generating Sets in $\mathbb{Z}_2^n$. *Combin. Probab. Comput.*, 21(06):916–941, November 2012.

[20] Chaim Even-Zohar and Shachar Lovett. The Freiman–Ruzsa theorem over finite fields. *Journal of Combinatorial Theory*, 125:333–341, 2014.

[21] G. A. Freiman. *Foundations of a structural theory of set addition*. American Mathematical Society, Providence, R. I., 1973.

[22] D Goss. *Basic Structures of Function Field Arithmetic, 1996*. Ergeb. Math. Grenzgeb.(3), 1996.

[23] W. T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998.
[24] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.
[25] W. T. Gowers and Julia Wolf. Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$. *Geom. Funct. Anal.*, 21(1):36–69, 2011.
[26] W. T. Gowers and Julia Wolf. Linear forms and quadratic uniformity for functions on $\mathbb{F}_p^n$. *Mathematika*, 57(2):215–237, February 2012.
[27] Ben J. Green. Arithmetic progressions in sumsets. *Geom. Funct. Anal.*, 12(3):584–597, 2002.
[28] Ben J. Green. Finite field models in additive combinatorics. In Bridget S Webb, editor, *Surveys in combinatorics 2005*, pages 1–27. Cambridge Univ. Press, Cambridge, Cambridge, 2005.
[29] Ben J. Green. Montréal notes on quadratic Fourier analysis. In *Additive combinatorics*, pages 69–102. Amer. Math. Soc., Providence, RI, 2007.
[30] Ben J. Green. Approximate algebraic structure. *To appear, Proceedings of the ICM*, 2014.
[31] Ben J. Green and Terence Tao. An inverse theorem for the Gowers $U^3(G)$ norm. *Proc. Edinb. Math. Soc. (2)*, 51(1):73–153, 2008.
[32] Ben J. Green and Terence Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008.
[33] Ben J. Green and Terence Tao. Freiman's theorem in finite fields via extremal set theory. *Combin. Probab. Comput.*, 18(3):335–355, 2009.
[34] Ben J. Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms. *Contrib. Discrete Math.*, 4(2):1–36, 2009.
[35] Ben J. Green and Terence Tao. An equivalence between inverse sumset theorems and inverse conjectures for the U^3 norm. *Math. Proc. Cambridge Philos. Soc.*, 149(1):1–19, 2010.
[36] Ben J. Green and Terence Tao. New bounds for Szemeredi's theorem, Ia: Progressions of length 4 in finite field geometries revisited. *arXiv*, May 2012.
[37] Ben J. Green, Terence Tao, and Tamar Ziegler. An inverse theorem for the Gowers $U^{s+1}[N]$-norm. *Electron. Res. Announc. Math. Sci.*, 18(0):69–90, 2011.
[38] Ben J. Green, Terence Tao, and Tamar Ziegler. An inverse theorem for the Gowers $U^{s+1}[N]$-norm. *Ann. of Math. (2)*, 176(2):1231–1372, 2012.
[39] Ben J. Green and Julia Wolf. A note on Elkin's improvement of Behrend's construction. In *Additive number theory*, pages 141–144. Springer, New York, New York, NY, 2010.
[40] Ben J. Green, Tamar Ziegler, and Terence Tao. An inverse theorem for the Gowers U^4-norm. *Glasg. Math. J.*, 53(1):1–50, 2011.
[41] Harald Andres Helfgott. Growth in groups: ideas and perspectives. *arXiv*, March 2013.
[42] Bernard Host and Bryna Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math. (2)*, 161(1):397–488, 2005.
[43] Russell Impagliazzo, Cristopher Moore, and Alexander Russell. An entropic proof of Chang's inequality. *SIAM J. Discrete Math.*, 28(1):173–176, 2014.
[44] Thai Hoang Le. Green-Tao theorem in function fields. *Acta Arith.*, 147(2):129–152, 2011.
[45] Yu-Ru Liu and Trevor D. Wooley. Waring's problem in function fields. *J. Reine Angew. Math.*, 638(638):1–67, 2010.
[46] Shachar Lovett. Equivalence of polynomial conjectures in additive combinatorics. *Combinatorica*, 32(5):607–618, 2012.
[47] Shachar Lovett, Roy Meshulam, and Alex Samorodnitsky. Inverse conjecture for the Gowers norm is false. *Theory Comput.*, 7(1):131–145, 2011.
[48] Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *J. Combin. Theory Ser. A*, 71(1):168–172, 1995.
[49] Melvyn B. Nathanson. *Additive number theory: inverse problems and the geometry of sumsets*, volume 165 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, May 2010.

[50] Giorgis Petridis. New proofs of Plünnecke-type estimates for product sets in groups. *Combinatorica*, pages 1–14, 2012.
[51] Helmut Plünnecke. *Eigenschaften und Abschätzungen von Wirkungsfunktionen*. BMwF-GMD-22. Gesellschaft für Mathematik und Datenverarbeitung, Bonn, 1969.
[52] K. F. Roth. On certain sets of integers. *J. Lond. Math. Soc. (2)*, 28(1):104–109, 1953.
[53] Imre Z. Ruzsa. Arithmetic progressions in sumsets. *Acta Arith.*, 60(2):191–202, 1991.
[54] Imre Z. Ruzsa. An analog of Freiman's theorem in groups. *Astérisque*, (258):xv, 323–326, 1999.
[55] Alex Samorodnitsky. Low-degree tests at large distances. In *STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 506–515. ACM, New York, New York, New York, USA, 2007.
[56] Alex Samorodnitsky and Luca Trevisan. Gowers uniformity, influence of variables, and PCPs. *SIAM J. Comput.*, 39(1):323–360, 2009.
[57] Tom Sanders. Green's sumset problem at density one half. *Acta Arith.*, 146(1):91–101, 2011.
[58] Tom Sanders. On Roth's theorem on progressions. *Ann. of Math. (2)*, 174(1):619–636, 2011.
[59] Tom Sanders. Approximate (Abelian) groups. *arXiv*, December 2012.
[60] Tom Sanders. On the Bogolyubov-Ruzsa lemma. *Anal. PDE*, 5(3):627–655, 2012.
[61] Tom Sanders. The structure theory of set addition revisited. *Bull. Amer. Math. Soc. (N.S.)*, 50(1):93–127, 2013.
[62] Andras Sárközy. On difference sets of sequences of integers. I. *Acta Math. Acad. Sci. Hungar.*, 31(1–2):125–149, 1978.
[63] Tomasz Schoen and Ilya Shkredov. Roth's theorem in many variables. *Israel J. Math.*, 199(1):287–308, 2014.
[64] Igor Shparlinski. Additive combinatorics over finite fields: new results and applications. In *Finite fields and their applications*, pages 233–271. De Gruyter, Berlin, 2013.
[65] Balazs Szegedy. On higher order Fourier analysis. *arXiv*, 2012.
[66] Endre Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.
[67] Terence Tao. Structure and Randomness in Combinatorics. *Foundations of Computer Science, 2007. FOCS '07. 48th Annual IEEE Symposium on*, pages 3–15, 2007.
[68] Terence Tao. *Higher order Fourier analysis*, volume 142 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.
[69] Terence Tao and Van Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
[70] Terence Tao and Tamar Ziegler. The inverse conjecture for the Gowers norm over finite fields via the correspondence principle. *Anal. PDE*, 3(1):1–20, 2010.
[71] Terence Tao and Tamar Ziegler. The inverse conjecture for the Gowers norm over finite fields in low characteristic. *Ann. Comb.*, 16(1):121–188, 2012.
[72] Madhur Tulsiani and Julia Wolf. Quadratic Goldreich-Levin Theorems. *FOCS 2011*, 2011.
[73] Salil Vadhan, Luca Trevisan, and M. Tulsiani. Regularity, boosting, and efficiently simulating every high-entropy distribution. In *Computational Complexity, 2009. CCC'09. 24th Annual IEEE Conference on*, pages 126–136. IEEE, 2009.
[74] Emanuele Viola and Avi Wigderson. Norms, XOR Lemmas, and Lower Bounds for GF(2) Polynomials and Multiparty Protocols. In *Computational Complexity, 2007. CCC '07. Twenty-Second Annual IEEE Conference on*, pages 141–154, June 2007.
[75] Tamar Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.*, 20(1):53–97 (electronic), 2007.
[76] Tamar Ziegler, Vitaly Bergelson, and Terence Tao. An inverse theorem for the uniformity seminorms associated with the action of $\mathbb{F}_p^\infty$. *Geom. Funct. Anal.*, 19(6):1539–1596, 2010.

School of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom
E-mail address: julia.wolf@bristol.ac.uk