Introduction to additive combinatorics

E. Kowalski

Version of January 8, 2024
kowalski@math.ethz.ch

Задача оказалась глубокой, видимость простоты была обманчивой.
Zadacha okazalas' glubokoy, vidimost' prostoty byla obmanchivoy.
(The task turned out to be deep, the appearance of simplicity was deceptive.)

A. Khintchine, ***Tri zhemchuzhiny teorii chisel***

(*"Three pearls of number theory"*)

Contents

Preface

These are lecture notes for an introduction to additive combinatorics and to some of its applications. It focuses on a few specific topics, chosen partly for their importance in the development of the subject, and partly from a definite bias concerning subjects of interest to the author (this bias is not necessarily positive – there will very little discussion of arithmetic aspects of additive combinatorics,[1] for instance).

Zürich, January 8, 2024

Acknowledgments. The first draft of these notes was prepared for a course "Introduction to additive combinatorics" that I taught at ETH Zürich during the Fall Semester 2023. Thanks to the students of the course for their interest and corrections, and to C. Bortolotto for organizing and preparing the exercise sessions. Thanks also to J. Fresán for his comments, and to B. Green for sending me his own notes for a similar course and his write-up of Schoen's argument for the Balog–Szemerédi–Gowers Theorem.

[1] Such as the use of Schnirelman's ideas to study sums of primes.

CHAPTER 1

Introduction

1.1. What is additive combinatorics?

Like many mathematical terminology, "additive combinatorics" is both perfectly accurate and deeply misleading.[1] It is accurate because its meaning is clear to the mathematical community, and reflects well the early history of the subject; it is misleading because it hides the breadth and importance this topic has acquired in recent years. For instance, the adjective "additive" hints at abelian groups, whereas some of the most striking applications of "additive" combinatorics lie squarely in fundamental problems related to non-abelian groups (including simple finite groups).

The best known early result of additive combinatorics, proven well before the name was introduced, is van der Waerden's Theorem [86].[2] This has been an extraordinarily influential statement.

THEOREM 1.1.1 (Van der Waerden, 1928). *For every positive integers l and k, there exists an integer n, such if the set $[n] = \{1, \dots, n\}$ is partitioned in k disjoint subsets, then one at least of these subsets contains l integers in arithmetic progression.*

In other words, if

$$\{1, \dots, n\} = \mathrm{C}_1 \cup \cdots \cup \mathrm{C}_k$$

with $\mathrm{C}_i \cap \mathrm{C}_j = \emptyset$ for $i \neq j$, then there exists an integer $i \in \{1, \dots, k\}$, $a \geqslant 1$ and $q \geqslant 1$ such that

$$a, \quad a+q, \quad a+2q, \ \dots \ , a+(l-1)q$$

all belong to C_i.

This statement immediately conveys the flavor of additive combinatorics: finite partitions of finite sets are clearly natural objects of combinatorics,[3] but the notion of arithmetic progressions involves the *algebraic* operation of addition on integers, and does not make sense in an arbitrary finite set.

One can move the birth date of additive combinatorics significantly; for instance, another prototypical result that fully belongs to the subject was already proved by Cauchy [16, Th. VII], before being rediscovered independently by Davenport [19]:

THEOREM 1.1.2 (Cauchy, 1813). *Let p be a prime number. If $\mathrm{A} \subset \mathbf{Z}/p\mathbf{Z}$ and $\mathrm{B} \subset \mathbf{Z}/p\mathbf{Z}$ are arbitrary non-empty sets, then we have*

$$|\mathrm{A} + \mathrm{B}| \geqslant \min(p, |\mathrm{A}| + |\mathrm{B}| - 1),$$

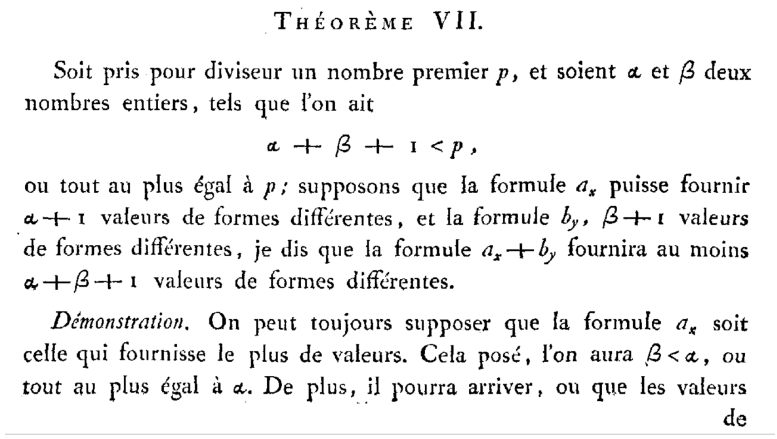
where $\mathrm{A} + \mathrm{B}$ is the set of all elements of the form $a + b$ with $(a, b) \in \mathrm{A} \times \mathrm{B}$.

---

[1] Another example of this phenomenon is the "large sieve".

[2] The mathematician P. Baudet, who formulated the corresponding conjecture (in the case $k = 2$) died young in 1921, and never knew the importance of his question.

[3] Although defining "combinatorics" concisely and precisely is certainly a very difficult endeavour.



THÉORÈME VII.

Soit pris pour diviseur un nombre premier p, et soient α et β deux nombres entiers, tels que l'on ait

$$\alpha + \beta + 1 < p,$$

ou tout au plus égal à p; supposons que la formule a_x puisse fournir $\alpha + 1$ valeurs de formes différentes, et la formule b_y, $\beta + 1$ valeurs de formes différentes, je dis que la formule $a_x + b_y$ fournira au moins $\alpha + \beta + 1$ valeurs de formes différentes.

Démonstration. On peut toujours supposer que la formule a_x soit celle qui fournisse le plus de valeurs. Cela posé, l'on aura $\beta < \alpha$, ou tout au plus égal à α. De plus, il pourra arriver, ou que les valeurs de

FIGURE 1.1. Cauchy's Theorem.

And finally, here is a third basic example: the "sum-product phenomenon", in the first version of Erdős and Szemerédi [28]. This has been the motivation for many of the developments in additive combinatorics.

THEOREM 1.1.3 (Erdős–Szemerédi, 1983). *There exists real numbers $\delta > 0$ and $c > 0$ such that for any finite set A of positive integers, we have*

$$\max(|\mathrm{A} + \mathrm{A}|, |\mathrm{A} \cdot \mathrm{A}|) \geqslant c|\mathrm{A}|^{1+\delta},$$

where $\mathrm{A} + \mathrm{A}$ and $\mathrm{A} \cdot \mathrm{A}$ are the sets of sums $a + b$ or products ab, respectively, with $(a, b) \in \mathrm{A}^2$.

In more intuitive words, this states that one cannot find a large finite set of integers which behave "as if it were a ring", i.e., as if it were both "almost" stable by addition and by multiplication. We will discuss this in Chapter 3.

1.2. What is special with additive combinatorics?

What one thinks of the theorems stated above is partly a question of taste, certainly. In Khintchine's account of van der Waerden's Theorem (see [55, Ch. I]), he mentions that he heard of the problem just when it had been solved, and[4] that "Nearly all mathematicians whom I met told me about it with enthusiasm". Since Göttingen was, at that time, most likely the leading mathematicial center in the world, this proves how immediately appealing the result was. Khintchine adds: "All to whom this question was put regarded the problem at first sight as quite simple; its solution in the affirmative appeared to be almost self-evident. The first attempts to solve it, however, led to nought", and concludes the introduction with: "The task turned out to be deep, the appearance of simplicity was deceptive".

Nevertheless, it seems fair to state that for a long time, the type of results that now are parts of additive combinatorics were not in the forefront of the mainstream of mathematics. Even in number theory, which is a closely related subject, much of the focus lies on "more structured" questions (for instance, in algebraic number theory, arithmetic geometry, and the theory of automorphic forms). From this point of view, the problems and the phenomena of additive combinatorics are viewed as being (maybe) cute and attractive, but not really "deep".

[4] In translation from the Russian.

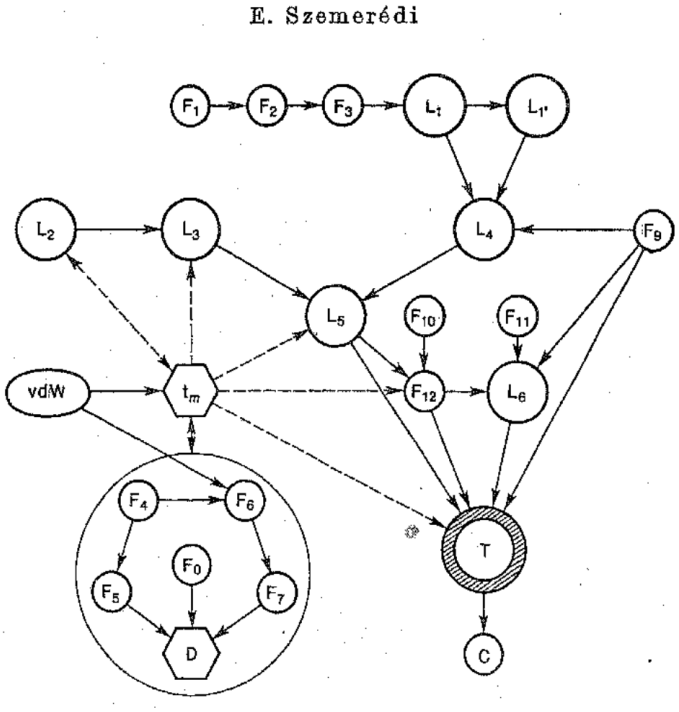


FIGURE 1.2. The logical structure of Szemerédi's proof.

Judgements of this type have now mostly disappeared. They were already challenged by any serious look at some of the works expanding, say, on van der Waerden's Theorem. For instance, among its direct "descendants", one find Szemerédi's Theorem [80], which strengthens Theorem 1.1.1 by finding arithmetic progressions in any "sufficiently dense" subset:

THEOREM 1.2.1 (Szemerédi). *Let $k \geqslant 1$ be an integer and let $\alpha > 0$ be a real number. There exists an integer $\mathrm{N}_k \geqslant 1$ such that for any integer $\mathrm{N} \geqslant \mathrm{N}_k$ and any subset A of $[\mathrm{N}]$, the condition $|\mathrm{A}| \geqslant \alpha \mathrm{N}$ implies that A contains a k-term arithmetic progression.*

The proof of this result is an extraordinary achievement, as the graph (included in Szemerédi's paper) displaying its logical structure already suggests.

Moreover, through a number of remarkable works, it has turned out that these types of results could be used in apparently completely unrelated areas of mathematics, and that ideas from additive combinatorics could both solve important open problems which didn't seem to have anything to do with it, as well as bring new light and insight apparently far from their simple-looking statements.

To give some examples:

(1) Furstenberg [37] realized early the connection between generalizations of van der Waerden's Theorem, especially Szeremédi's Theorem, and ergodic theory.

(2) Green [44] (for 3-term progressions) and then Green–Tao, in their celebrated work [45] proving the existence of arbitrarily long arithmetic progressions in the primes, showed how additive combinatorics led to various breakthroughs in the study of prime numbers.

(3) Bourgain, in a variety of contexts, exploited various aspects of additive combinatorics in different areas of analysis (Kakeya sets, applications to harmonic analysis, etc...)

(4) Building on the deep Inverse Theorem for Gowers norms of Green, Tao and Ziegler [46], a number of researchers have made significant progress in various problems of arithmetic geometry related to point-counting on algebraic varieties (see for instance the paper [14] of Browning, Matthiesen and Skorobogatov).

(5) Starting from a sensational contribution of Helfgott [49], combined with ideas of Bourgain and Gamburd [9], completely new phenomena related to quantitative properties of finite simple groups have been discovered since the mid 2000's, leading to close connections with geometric group theory, the theory of expander graphs, and to extremely surprising applications, including in number theory (see for instance the book [59]).

1.3. A soft beginning

We begin the formal mathematical part of these notes by discussing Cauchy's Theorem and giving some applications. We first recall the statement:

THEOREM 1.3.1. *For any two non-empty subsets* A *and* B *of the group* $\mathbf{Z}/p\mathbf{Z}$*, where* p *is a prime number, we have* $|A+B| \geqslant \min(p, |A|+|B|-1)$.

We begin with some remarks to show the result in some context. The first is that the question of estimating the size of the *sumset* $A+B$ of two finite sets can be asked in any abelian group (in fact, in any group, but the addition is taken to refer to a commutative operation; examples of this question with non-abelian will appear later).

It is natural then to look first at the case of $\mathbf{Z}$, which intuitively might be simpler than the case of $\mathbf{Z}/p\mathbf{Z}$ (because "wrap-around" does not happen in the addition process). In fact, for (non-empty) finite subsets of $\mathbf{Z}$ or $\mathbf{R}$, we easily obtain

$$|A+B| \geqslant |A|+|B|-1,$$

and we also easily see that this cannot be improved (without extra assumptions or information on A and B).

Indeed, for the former, it suffices to enumerate the elements of A and B in order, say

$$a_1 < a_2 < \cdots < a_{|A|}, \qquad b_1 < b_2 < \cdots < b_{|B|}.$$

Then note that $A+B$ contains the following ordered sequence

$$a_1+b_1 < a_2+b_1 < \cdots < a_{|A|}+b_1 < a_{|A|}+b_2 < \cdots < a_{|A|}+b_{|B|}$$

of $|A|+|B|-1$ elements. Because the ordering is strict, this gives the result. The optimality of the bound is demonstrated by the simple example $A = B = [n]$, for which we have $A+B = \{2, 3, \ldots, 2n\}$, with size $2n-1$.

These simple remarks put Cauchy's Theorem in context. Indeed, it leads us to think that $|A|+|B|-1$ is a natural quantity, and a reasonable goal to ask about (while the fact that $|A+B| \leqslant p$ is an obvious upper-bound for subsets of $\mathbf{Z}/p\mathbf{Z}$). Moreover, it also gives the optimality: although we cannot order $\mathbf{Z}/p\mathbf{Z}$ in a way which respects addition in general, we can do so for subsets that are not too large if we perform "not too many" additions. Precisely, the example $A = B = [n]$ in $\mathbf{Z}$ can be transplanted to $\mathbf{Z}/p\mathbf{Z}$ for any prime such that $p > 2n$: if we take for A and B the set of residue classes modulo p of the integers from 1 to n, then the subset $A+B$ of $\mathbf{Z}/p\mathbf{Z}$ coincides with the set of residue classes of integers from 2 to $2n$, and in particular contains $2n-1$ elements. (Soon, we will formalize such a property as the fact the reduction modulo p defines "Freiman homomorphisms" on suitable subsets of $\mathbf{Z}$.)

Moreover, thinking in general terms, it is immediately apparent that there must be some interplay between the group structure and the behavior of sumsets: for instance, if A and B are both contained in a proper subgroup $H \subset G$, then $A+B$ is also contained in H, and this gives an a-priori upper-bound for $|A+B|$ which might be smaller than the "expected" quantity $|A|+|B|-1$. Indeed, if $A = B = H$, then we have $|A+B| = |H| = |A|$. Note that $|A+B| \geqslant \max(|A|, |B|)$ in all cases (because A and B are not empty, so we can fix some element of B, say b, and then $A+B$ contains $A+b$).

The case of Cauchy's Theorem (and of $\mathbf{Z}$ or $\mathbf{R}$) are therefore simplified by the fact that $\mathbf{Z}/p\mathbf{Z}$ is a simple abelian group: it has no proper subgroup except for $\{0\}$ (and if $A = B = \{0\}$ then $|A| + |B| - 1 = |A|$...)

But now let us finally prove the result... (The reader is encouraged, before going on, to devote some time to try to find a proof.)

Proof of Cauchy's Theorem. It is natural to argue by induction on the size of one of the two sets, say B, since the nature of the desired lower-bound (the size of B plus some quantity independent of B) is such that one only needs to get one more element at each stage of the induction.

We therefore proceed this way. If $|B| = 1$, then $A + B$ is just an additive translate of A, so $|A + B| = |A| = |A| + |B| - 1$, which gives the desired result.

Suppose now that $|B| \geqslant 2$ and that the result holds for $|A + B'|$ for any set B' with $1 \leqslant |B'| < |B|$.

It is somewhat convenient to replace B by a translate to ensure that $0 \in B$, so that $A \subset A + B$. We can moreover assume of course that A is not the whole group $\mathbf{Z}/p\mathbf{Z}$.

We then first observe (this may be considered as trying to first handle the case where $|B| = 2$) that for any fixed $b_0 \in B - \{0\}$, there exists some $a_0 \in A$ such that $a_0 + b_0 \notin A$. Indeed, otherwise the set A would be invariant by a non-trivial translation, which is impossible since A is neither empty nor the whole group (concretely, for a given $a_0 \in A$, note that this implies that $a_0 + kb_0 \in A$ for all $k \in \mathbf{Z}$, but the set of values $a_0 + kb_0$ is all of $\mathbf{Z}/p\mathbf{Z}$ when $b_0 \neq 0$, so this is impossible; abstractly, we would have $A + b_0 = A$, so the stabilizer of A is non-trivial, hence is equal to $\mathbf{Z}/p\mathbf{Z}$ since this is a simple group, and we would again have $\mathbf{Z}/p\mathbf{Z} = A$).

Now here comes the trick. Given $b_0 \in B - \{0\}$ and an element $a_0 \in A$ with $a_0 + b_0 \notin A$, define

$$B' = \{b \in B \mid a_0 + b \in A\}, \qquad B'' = B - B' = \{b \in B \mid a_0 + b \notin A\}.$$

Since $b_0 \in B''$, this set is not empty, and hence $|B'| < |B|$. Moreover, $0 \in B'$, so B' is also not empty. Furthermore, let $A' = A \cup (a_0 + B'')$.

The key point of these definitions is that $A + B$ contains $A' + B'$. Indeed, the only issue is whether $a_0 + b'' + b' \in A + B$ for $b' \in B'$ and $b'' \in B''$, but this is so because $a_0 + b'' + b' = (a_0 + b') + b''$, and $a_0 + b' \in A$ by definition of B'.

Using the induction hypothesis for the sets A' (which is not empty since it contains A) and B', we conclude that

$$|A + B| \geqslant |A' + B'| \geqslant \min(p, |A'| + |B'| - 1),$$

and finally, we note that $|A'| = |A| + |B''|$ (because $a + B''$ and A are disjoint) so that $|A'| + |B'| = |A| + |B''| + |B'| = |A| + |B|$. □

Here is a simple but useful corollary of Cauchy's Theorem.

COROLLARY 1.3.2. *Let p be a prime number and $\alpha \geqslant 2/p$ a real number. Let* A *be a subset of $\mathbf{F}_p$ such that $|\mathrm{A}| \geqslant \alpha p$. There exists an integer $k \leqslant \lceil 2\alpha^{-1} \rceil$ such that every element of $\mathbf{F}_p$ is the sum of at most k elements of* A.

PROOF. Let $j \geqslant 1$ be an integer and A_j the set of elements which can be represented as the sum of j elements of A, so that $\mathrm{A}_2 = \mathrm{A} + \mathrm{A}$. We deduce by induction from the Cauchy–Davenport inequality that

$$|\mathrm{A}_j| \geqslant \min(p, j|\mathrm{A}| - j + 1)) = \min(p, j\alpha p - j + 1),$$

and hence A_j must be equal to all of $\mathbf{F}_p$ when $j(\alpha p - 1) \geqslant p$. This inequality holds if $j \geqslant 2\alpha^{-1}$ (because of the assumption $\alpha \geqslant 2/p$), hence the result. □

REMARK 1.3.3. This study of finite sumsets might have reminded those readers familiar with integration theory of a well-known result which is often presented as an elementary surprising application of the convolution operation on integrable functions on $\mathbf{R}$: if A and B are two measurable subsets of $\mathbf{R}$ with *positive* measure, then the sumset $\mathrm{A} + \mathrm{B}$ contains a non-empty open interval. (This is because, if we denote by φ_{A} and φ_{B} the characteristic functions of A and B, respectively, and assume, as we may, that A and B have finite measure, then the convolution $f = \varphi_{\mathrm{A}} * \varphi_{\mathrm{B}}$ defined by

$$f(x) = \int_{\mathbf{R}} \varphi_{\mathrm{A}}(y)\varphi_{\mathrm{B}}(x - y)dy$$

is known to be a continuous function on $\mathbf{R}$; it is not identically zero, because its integral is equal to the product of the integrals of A and B, i.e., of the measures of the two sets, which are positive by assumption; and if $f(x_0) > 0$ for some x_0, then by continuity we get $f(x) > 0$ for all x in some non-empty open interval containing x_0, and for any such x, we must have $x \in \mathrm{A} + \mathrm{B}$ since otherwise the integral definition would give $f(x) = 0$.)

Here is a first "concrete" application of Cauchy's Theorem (which indeed is closely related to his motivation; such results had been proved earlier by Lagrange, using more algebraic methods).

COROLLARY 1.3.4. *Let p be a prime number. Let α and β be non-zero elements of $\mathbf{Z}/p\mathbf{Z}$. For any $x \in \mathbf{Z}/p\mathbf{Z}$, there exist y and z in $\mathbf{Z}/p\mathbf{Z}$ such that $x = \alpha y^2 + \beta z^2$.*

PROOF. We apply Cauchy's Theorem with

$$\mathrm{A} = \alpha \cdot \mathrm{Q}, \qquad \mathrm{B} = \beta \cdot \mathrm{Q},$$

where Q is the set of squares in $\mathbf{Z}/p\mathbf{Z}$. We have $|\mathrm{Q}| = 2$ if $p = 2$, and if p is odd, then it is elementary that $|\mathrm{Q}| = (p + 1)/2$ (this can be proved elementarily, but algebraically, this comes from the fact that the set of *non-zero* squares is the image of the group homomorphism $f\colon x \mapsto x^2$ from $(\mathbf{Z}/p\mathbf{Z})^\times$ to itself, and since the kernel of this homomorphism is $\{-1, 1\}$, of size 2, the image has size $(p - 1)/2$ by the isomorphism $\mathrm{Im}(f) \simeq (\mathbf{Z}/p\mathbf{Z})^\times / \ker(f)$; adding the element $0 \in \mathrm{Q}$, we find $|\mathrm{Q}| = (p + 1)/2$). Since α and β are non-zero, this implies that $|\mathrm{A}| = |\mathrm{B}| = |\mathrm{Q}|$, and then $|\mathrm{A}| + |\mathrm{B}| - 1$ is either equal to 3 (if $p = 2$) or to p if p is odd. In any case, Cauchy's Theorem gives $|\mathrm{A} + \mathrm{B}| \geqslant p$, and the result follows. □

EXERCISE 1.3.5. Prove directly the case of Cauchy's Theorem when $|\mathrm{A}| + |\mathrm{B}| \geqslant p + 1$.

EXERCISE 1.3.6. (1) Show by example that Cauchy's Theorem does not hold for subsets of $\mathbf{Z}/q\mathbf{Z}$ in general, if $q \geqslant 1$ is arbitrary.

(2) Show that if $q \geqslant 1$ is an arbitrary positive integer, if $\mathrm{A} \subset \mathbf{Z}/q\mathbf{Z}$ is arbitrary and if $\mathrm{B} \subset \mathbf{Z}/q\mathbf{Z}$ is such that $0 \in \mathrm{B}$ then

$$|\mathrm{A}+\mathrm{B}| \geqslant \min(q, |\mathrm{A}| + |\mathrm{B}^\times| - 1),$$

where $\mathrm{B}^\times$ is the set of elements of B which are invertible in $\mathbf{Z}/q\mathbf{Z}$ (i.e., those $b \in \mathrm{B}$ which are residue classes of integers coprime to q).

There are (at least) three important lessons that can be learnt by considering the most elementary properties of sumsets.

An inverse theorem. We go back to the case of sumsets in $\mathbf{Z}$ (or $\mathbf{R}$). Since we exhibited easily $|\mathrm{A}| + |\mathrm{B}| - 1$ distinct elements of $\mathrm{A} + \mathrm{B}$, it seems tempting to think that one could understand *in which cases exactly the bound is sharp*. This is indeed possible, and is the first instance of a frequent phenomenon in (additive) combinatorics, where "extremal" examples for certain problems can be characterized quite precisely.

PROPOSITION 1.3.7. *Let* A *and* B *be non-empty finite subsets of* $\mathbf{R}$. *We have* $|\mathrm{A}+\mathrm{B}| = |\mathrm{A}| + |\mathrm{B}| - 1$ *if and only if one of the following conditions is valid:*

(1) *We have* $|\mathrm{A}| = 1$.

(2) *We have* $|\mathrm{B}| = 1$.

(3) *The sets* A *and* B *are arithmetic progressions with the same common difference.*

PROOF. We denote $k = |\mathrm{A}|$ and $l = |\mathrm{B}|$. If either k or l is equal to 1, then there is equality $|\mathrm{A} + \mathrm{B}| = |\mathrm{A}| + |\mathrm{B}| - 1$. If A and B are arithmetic progressions with the same common difference $h \neq 0$, so

$$\mathrm{A} = \{a, a+h, \dots, a+(k-1)h\}, \qquad \mathrm{B} = \{b, b+h, \dots, b+(l-1)h\}$$

for some real numbers a and b, then we see easily that

$$\mathrm{A} + \mathrm{B} = \{a+b, a+b+h, \dots, a+b+(k+l-2)h\},$$

which has $k + l - 1$ elements.

The question is therefore to prove the converse. For this, we assume that A and B have at least two elements, and we then prove that they are arithmetic progressions with the same common difference.

We enumerate the elements of A and B in order, as before:

$$a_1 < \cdots < a_k, \qquad b_1 < \cdots < b_l.$$

Note that $l \geqslant 2$, so that b_2 exists. The assumption $|\mathrm{A} + \mathrm{B}| = k + l - 1$ implies that $\mathrm{A} + \mathrm{B}$, in order, consists of the elements

$$a_1 + b_1 < a_2 + b_1 < \cdots < a_k + b_1 < a_k + b_2 < \cdots < a_k + b_l.$$

Observe now that

$$a_1 + b_1 < a_1 + b_2 < \cdots < a_{k-1} + b_2 < a_k + b_2,$$

which, by comparing, means that the $k - 1$ tuples of ordered integers

$$a_2 + b_1 < \cdots < a_k + b_1$$
$$a_1 + b_2 < \cdots < a_{k-1} + b_2$$

have to be identical. This means that $a_j + b_1 = a_{j-1} + b_2$ for all $2 \leqslant j \leqslant k$. Thus A is an arithmetic progression with common difference $b_2 - b_1$. By exchanging the role of A

and B, it follows that B is also an arithmetic progression, with common difference $a_2 - a_1$. But since the above (with $j = 2$) shows that

$$a_2 + b_1 = a_1 + b_2,$$

these arithmetic progressions have the same common difference. □

As we will see later in this book, the real depth of such questions arises rather when one attempts to classify examples which are "almost" extremal – in that case, examples where $|A + B|$ is just a bit larger than $|A| + |B| - 1$.

REMARK 1.3.8. This proof used the order structure of the real numbers. In view of Cauchy's Theorem, it is natural to ask whether a similar statement holds for subsets of $\mathbf{Z}/p\mathbf{Z}$. This is indeed the case, and is due to Vosper (see, e.g., [84, Th. 5.9]). Precisely, if p is a prime number, A and B are subsets of $\mathbf{Z}/p\mathbf{Z}$ with $|A + B| = |A| + |B| - 1$, and in addition if $|A| + |B| \leqslant p - 2$, then A and B are arithmetic progressions with the same common difference.

Randomness helps. The inverse theorem above should only confirm a definite intuition that the Cauchy–Davenport lower bound, even if is true that it sometimes cannot be improved, is "usually" far from the truth. In other words, if one takes A and B among "typical" subsets of $\mathbf{Z}/p\mathbf{Z}$, the sumset $A + B$ should be quite a bit bigger than $|A| + |B| - 1$ (assuming the sets are not so big that the sumset is all of $\mathbf{Z}/p\mathbf{Z}$). There are many different ways to try to make this precise, and these can be very useful in applications. One of the simplest statements exploits the fact that we can also *multiply* elements of $\mathbf{Z}/p\mathbf{Z}$, which is a finite field; it turns out that if we are allowed to replace one of the sets by a dilate, then we get a much stronger lower bound.

PROPOSITION 1.3.9. *Let p be a prime number and let* A *and* B *be non-empty subsets of $\mathbf{Z}/p\mathbf{Z}$. There exists $x \in (\mathbf{Z}/p\mathbf{Z})^\times$ such that*

$$|A + xB| \geqslant \min\Big(\frac{|A||B|}{2}, \frac{p}{6}\Big),$$

where here $xB = \{xb \mid b \in B\}$.

Note that the constant $1/6$ here is not particularly important, and arises when dealing with large sets. One should think of a set of size $p/6$ as simply "a good chunk" of $\mathbf{Z}/p\mathbf{Z}$.

PROOF. We may assume that $\min(|A|, |B|) \leqslant p/6$, since otherwise just taking $x = 1$ gives $|A + B| \geqslant p/6$.

The idea is to take x "at random", and in this case, this means uniformly among elements of $(\mathbf{Z}/p\mathbf{Z})^\times$: we look at the function f from $(\mathbf{Z}/p\mathbf{Z})^\times$ to $\mathbf{R}$ which sends x to $|A + xB|$, and show that its average is quite large – since some value of f is at least as large as the average, the result will follow.[5]

We claim that

$$f(x) \geqslant |A||B| - \frac{1}{2}g(x) \tag{1.1}$$

where $g(x)$ is the number of $(a_1, a_2, b_1, b_2) \in A^2 \times B^2$ such that $x = (a_1 - a_2)/(b_2 - b_1)$. To see this, we write $A + xB$ as the union of the set $a + xB$ for $a \in A$; this union is of

[5] In other cases, for instance in the proof of Theorem 2.7.1, it may be better to select an element "at random" with a different distribution.

course not disjoint in general, but a truncated form of the inclusion exclusion formula (see Example A.3.3) implies that

$$|\mathrm{A}+x\mathrm{B}| = \Big|\bigcup_{a\in\mathrm{A}}(a+x\mathrm{B})\Big| \geqslant \sum_{a\in\mathrm{A}}|a+x\mathrm{B}| - \frac{1}{2}\sum_{\substack{a_1,a_2\in\mathrm{A}\\ a_1\neq a_2}}|(a_1+x\mathrm{B})\cap(a_2+x\mathrm{B})|.$$

The first term on the right-hand side is $|\mathrm{A}||\mathrm{B}|$. As for the other term, for given $a_1 \neq a_2$ in A, there is a bijection between $(a_1+x\mathrm{B})\cap(a_2+x\mathrm{B})$ and the set of $(b_1,b_2)\in\mathrm{B}^2$, with $b_1\neq b_2$, such that $x=(a_1-a_2)/(b_2-b_1)$, defined by mapping y to $((y-a_1)/x,(y-a_2)/x)$, with reciprocal bijection mapping (b_1,b_2) to $a_1+xb_1=a_2+xb_2$. Hence the two sets have the same size, and we find that

$$\sum_{\substack{a_1,a_2\in\mathrm{A}\\ a_1\neq a_2}}|(a_1+x\mathrm{B})\cap(a_2+x\mathrm{B})| = g(x),$$

by definition of g, which concludes the proof of (1.1).

This lower bound implies that

$$\frac{1}{p-1}\sum_{x\in(\mathbf{Z}/p\mathbf{Z})^\times} f(x) \geqslant |\mathrm{A}||\mathrm{B}| - \frac{1}{2(p-1)}\sum_{x\in(\mathbf{Z}/p\mathbf{Z})^\times} g(x) = |\mathrm{A}||\mathrm{B}| - \frac{|\mathrm{A}|^2|\mathrm{B}|^2}{2(p-1)},$$

since the sum of $g(x)$ over all x is just the number of quadruples $(a_1,a_2,b_1,b_2)\in\mathrm{A}^2\times\mathrm{B}^2$.

If $|\mathrm{A}||\mathrm{B}|\leqslant p/2$, then it follows that

$$\frac{1}{p-1}\sum_{x\in(\mathbf{Z}/p\mathbf{Z})^\times} f(x) \geqslant \frac{|\mathrm{A}||\mathrm{B}|}{2},$$

from which the existence of a suitable x follows.

The remaining case where $|\mathrm{A}||\mathrm{B}|>p/2$ is handled by a simple trick: we observe that we can find subsets $\mathrm{A}'\subset\mathrm{A}$ and $\mathrm{B}'\subset\mathrm{B}$ such that $p/3\leqslant|\mathrm{A}'||\mathrm{B}'|\leqslant p/2$; by the first case, we then find some $x\in(\mathbf{Z}/p\mathbf{Z})^\times$ for which the lower bound

$$|\mathrm{A}+x\mathrm{B}| \geqslant |\mathrm{A}'+x\mathrm{B}'| \geqslant \frac{|\mathrm{A}'||\mathrm{B}'|}{2} \geqslant \frac{p}{6}$$

holds. To check that A' and B' exist, we may assume that $|\mathrm{A}|\leqslant|\mathrm{B}|$ (up to exchanging the two sets). Recall that we assumed that $|\mathrm{A}|\leqslant p/6$; it follows that there exists an integer j such that $0\leqslant j<|\mathrm{B}|$ and

$$p\Big(\frac{1}{2}-\frac{1}{6}\Big)\leqslant j|\mathrm{A}|\leqslant\frac{p}{2},$$

and we can then take $\mathrm{A}'=\mathrm{A}$ and define B' to be any subset of B with j elements. □

Harmonic analysis. The concrete application of the Cauchy–Davenport inequality that we have discussed has an arithmetic flavor, and gives a first idea of the links between number theory and additive combinatorics. There is one particular type of methods in (analytic) number theory that is extremely powerful in many counting problems, and which is related to the ideas of harmonic analysis. It is natural to look back at Cauchy's Theorem from this point of view, as this allows us to get some intuition on the differences between these two circles of ideas (and their respective strengths and weaknesses).

Quite generally, starting from an abelian group G and subsets A and B of G, we want to study $A + B$ by looking at the representation function $r\colon G \to \mathbf{R}$ which sends $x \in G$ to the *number of representations* of x as a sum $x = a + b$ with $(a, b) \in A \times B$: formally,

$$r(x) = \sum_{\substack{a,b \\ a+b=x}} 1 \geqslant 0.$$

The sumset $A + B$ is then recovered as the *support* of r, i.e., the set of $x \in G$ such that $r(x) \neq 0$, but the function r contains more information – it can distinguish between elements which are sums in $A + B$ in just one way, and those which are in many ways. Also, r gives us, for instance, access to the "average" number of representations: this is

$$\frac{1}{|G|}\sum_{x\in G} r(x) = \frac{1}{|G|}\sum_{x\in G}\sum_{\substack{(a,b)\in A\times B \\ a+b=x}} 1 = \frac{1}{|G|}\sum_{a\in A}\sum_{b\in B}\sum_{\substack{x\in G \\ a+b=x}} 1 = \frac{|A||B|}{|G|}. \tag{1.2}$$

The harmonic analysis approach to the function r then consists in representing this function in a well-chosen basis of the space $C(G)$ of functions from G to $\mathbf{C}$, so that this average appears naturally in the decomposition. The last requirement amounts to asking that the constant function appears in the chosen basis, but there remain many potential choices of bases. A hint of a useful approach is given by interpreting the average of a function as the *inner product* of this function with the constant function; thus one should make use of the natural inner product

$$\langle f, g\rangle = \frac{1}{|G|}\sum_{x\in G} f(x)\overline{g(x)}$$

on the complex vector space $C(G)$, and one should look for an orthonormal basis for this inner product, which contains the constant function 1 (which is normalized: it satisfies $\|1\| = 1$).

There are still many choices of bases, but now we use the fact that G is a group, and not just a finite set. There is then a distinguished orthonormal basis, namely that of *characters* of G. These are the functions $\chi\colon G \to \mathbf{C}$ which are in fact group morphisms from G to $\mathbf{C}^\times$. The basic properties of these are summarized (with proofs) in Section A.7, and this precise fact is Theorem A.7.3. This allows to write

$$r(x) = \frac{|A||B|}{|G|} + \sum_{\chi\neq 1}\langle r, \chi\rangle\chi(x)$$

for any $x \in G$, where the coefficients are

$$\langle r, \chi\rangle = \frac{1}{|G|}\sum_{x\in G} r(x)\overline{\chi(x)}.$$

In fact, using the definition of r, these coefficients can be transformed into

$$\langle r, \chi\rangle = \frac{1}{|G|}\sum_{x\in G}\Big(\sum_{\substack{(a,b)\in A\times B \\ a+b=x}} 1\Big)\overline{\chi(x)} = \frac{1}{|G|}\sum_{a\in A}\sum_{b\in B}\overline{\chi(a+b)} = \frac{1}{|G|}\Big(\sum_{a\in A}\overline{\chi(a)}\Big)\Big(\sum_{b\in B}\overline{\chi(b)}\Big).$$

We note here that is a special case of the interaction of the discrete Fourier transform with the convolution: indeed, we are using the fact that $r = \varphi_A * \varphi_B$, and that, up to normalizing constants, the Fourier transform of a convolution is the product of the Fourier transforms of the factors (see Proposition A.7.7 and the remarks following).

If we specialize further (to gain concreteness) to $\mathrm{G} = \mathbf{Z}/p\mathbf{Z}$, then the characters are parameterized by an element $h \in \mathbf{Z}/p\mathbf{Z}$, and are given by $x \mapsto \exp(2i\pi hx/p)$ (which is well-defined because the function on $\mathbf{R}$ given by sending t to $\exp(2i\pi t/p)$ is periodic with period p; see Example A.7.2), which we denote more concisely by $e(hx/p)$. Then we get, for this case, the formula

$$r(x) = \frac{|\mathrm{A}||\mathrm{B}|}{p} + \frac{1}{p}\sum_{\substack{h\in\mathbf{Z}/p\mathbf{Z}\\ h\neq 0}}\Big(\sum_{a\in\mathrm{A}} e\Big(-\frac{ha}{p}\Big)\Big)\Big(\sum_{b\in\mathrm{B}} e\Big(-\frac{hb}{p}\Big)\Big)e\Big(\frac{hx}{p}\Big).$$

At least if $|\mathrm{A}||\mathrm{B}|$ is large, one can hope to show that the first term dominates (so that $r(x)$ is "close" to its average) by exploiting the fact that for each $h \neq 0$, the corresponding expression on the right-hand side involves *oscillating sums*

$$\sum_{a\in\mathrm{A}} e\Big(-\frac{ha}{p}\Big), \qquad \sum_{b\in\mathrm{B}} e\Big(-\frac{hb}{p}\Big),$$

of complex numbers of modulus 1, which should therefore have smaller modulus than the number of terms, due to "cancellations" arising from summing complex numbers with varying arguments.

Even without any "fine" control on these sums, there are in fact other fairly elementary ways to exploit this basic idea of Fourier analysis, which turn out to be quite powerful – instances include Bogolyubov's Lemma (see Proposition 2.9.2), and (in the slightly more sophisticated setting of non-abelian groups) Gowers's product theorem (see 2.8.4).

Exercise 1.3.10. Let $q \geqslant 1$ and $k \geqslant 1$ be integers. Let $\mathrm{A} \subset \mathbf{Z}/q\mathbf{Z}$ be a non-empty set, and let $\mathrm{A}^{(k)} = \mathrm{A} + \cdots + \mathrm{A}$ (with k-summands) be the set of elements of the form $a_1 + \cdots + a_k$ with $a_i \in \mathrm{A}$. For $x \in \mathbf{Z}/q\mathbf{Z}$, define

$$r_k(x) = |\{(a_1, \ldots, a_k) \in \mathrm{A}^k \mid a_1 + \cdots + a_k = x\}|.$$

(1) Show that

$$r_k(x) = \frac{|\mathrm{A}|^k}{q} + \frac{1}{q}\sum_{1\leqslant h<q} \mathrm{W}_{\mathrm{A}}(h)^k e\Big(\frac{hx}{q}\Big)$$

where

$$\mathrm{W}_{\mathrm{A}}(h) = \sum_{a\in\mathrm{A}} e\Big(-\frac{ah}{q}\Big).$$

(2) For $k \geqslant 2$, deduce that

$$r_k(x) \geqslant \frac{|\mathrm{A}|^k}{q} - |\mathrm{A}| \sup_{h\neq 0} |\mathrm{W}_{\mathrm{A}}(h)|^{k-2}.$$

(3) Let $\delta > 0$ be such that $|\mathrm{W}_{\mathrm{A}}(h)| \leqslant q^{1-\delta}$ for all h. Assuming $k \geqslant 2$, show that $\mathrm{A}^{(k)} = \mathbf{Z}/q\mathbf{Z}$ if

$$|\mathrm{A}| > q^{\delta + (1-\delta)/(k-1)}.$$

Exercise 1.3.11. The goal of this exercise is to show that there are natural subsets of $\mathbf{Z}/p\mathbf{Z}$, for p prime, which satisfy the assumption of the last question of the previous exercise, with $\delta = 1/2$.

Let p be an odd prime number, and let Q be the set of non-zero squares in $\mathbf{Z}/p\mathbf{Z}$. It has $(p-1)/2$ elements.

(1) If $p \equiv 3 \bmod 4$, show that $\mathrm{Q} + \mathrm{Q} \neq \mathbf{Z}/p\mathbf{Z}$.

For $h \in \mathbf{Z}/p\mathbf{Z}$, denote

$$\mathrm{W}(h) = \sum_{x \in \mathrm{Q}} e\Big(\frac{hx}{p}\Big).$$

(2) Show that

$$\mathrm{W}(h) = \frac{1}{2} \sum_{x \in \mathbf{Z}/p\mathbf{Z}} e\Big(\frac{hx^2}{p}\Big) - \frac{1}{2}.$$

(3) Show that

$$\Big| \sum_{x \in \mathbf{Z}/p\mathbf{Z}} e\Big(\frac{hx^2}{p}\Big)\Big|^2 = \sum_{u \in \mathbf{Z}/p\mathbf{Z}} \sum_{v \in \mathbf{Z}/p\mathbf{Z}} e\Big(\frac{huv}{p}\Big).$$

(4) Deduce that

$$|\mathrm{W}(h)| \leqslant \frac{1}{2}(\sqrt{p}+1) \leqslant \sqrt{p}$$

for all $h \neq 0$.

Exercise 1.3.12. Let $q \geqslant 1$ be an integer and let $\alpha \in]0,1[$ be a real numnber. We define a *random subset* A of $\mathbf{Z}/q\mathbf{Z}$ by the condition that each $x \in \mathbf{Z}/q\mathbf{Z}$ (independently) belongs to A with probability α.

(1) For any subset $\mathrm{X} \subset \mathbf{Z}/p\mathbf{Z}$, show that

$$\mathbf{P}(\mathrm{A} = \mathrm{X}) = \alpha^{|\mathrm{X}|}(1-\alpha)^{q-|\mathrm{X}|}.$$

(2) Show that the average of the size of A is equal to αq, or in other words

$$\mathbf{E}(|\mathrm{A}|) = \alpha q.$$

(3) For any non-zero $h \in \mathbf{Z}/q\mathbf{Z}$, show that

$$\mathbf{E}\Big(\Big|\sum_{x \in \mathrm{A}} e\Big(\frac{hx}{q}\Big)\Big|^2\Big) = \alpha(1-\alpha)q.$$

Concretely, this result indicates that for a "random" subset of $\mathbf{Z}/q\mathbf{Z}$ of size proportional to q, the coefficients $\mathrm{W}_{\mathrm{A}}(h)$ appearing in Exercise 1.3.10 are usually of size $\sqrt{q}$.

1.4. Outline of the book

Here is now a quick outline of the main topics that will appear in the text. For detailed statements, we refer to the introductory sections of the corresponding chapters.

(1) Approximate subgroups, Sidon sets, and other types of structured (or non-structured) sets in groups (see Chapter 2).

(2) Quasi-randomness in the sense of Gowers and applications to product-free subsets of finite groups (Section 2.8).

(3) The sum-product phenomenon and its generalizations (Chapter 3).

(4) Discussion of arithmetic progressions in subsets of abelian groups, including Roth's Theorem concerning existence and density of 3-term arithmetic progressions in $\mathbf{Z}$ (Chapter 4).

(5) The "structure vs randomness" dichotomy, and the theory of Gowers norms for arithmetic progressions of length 4 and more (also in Chapter 4).

Finally, an Appendix at the end of the notes summarizes some of the background material that is used at some point in the text, e.g. concerning harmonic analysis on finite abelian groups.

It is obvious from this list that this book is just an introduction; it will also be clear from the way we discuss them that we do not intend or attempt in any way to cover any subject in real depth.

For further and deeper information, besides the original papers, the standard reference is the book of Tao and Vu [84].

1.5. Some general remarks

We collect here some notes about certain aspects of additive combinatorics, and how we handle them.

(1) There is a "hard analysis" aspect to many ideas and proofs, in the sense of finitary, quantitative assumptions and conclusions. This often means that either the statements or the proofs (often both) involve a number of parameters with quantitative restrictions on them (often in relation between each other). These can be sometimes hard to digest at first, and moreover there are often unspecified constants arising in various inequalities which are in principle "effective", in the sense that one could write an explicit upper-bound in terms of the parameters, but which are very rarely specified this way.

In this text, we try to be as specific as possible when this is doable without much work, but we also try to point out which finicky part of certain steps of certain proofs are just there to handle this type of goal (see, e.g,, the proof of the sum-product theorem over finite fields due to Bourgain, Katz and Tao, Theorem 3.2.1).

(2) We often go back to the original papers for the proofs, when these are accessible enough; even if slicker arguments sometimes exist, the fact is that the lessons of (additive) combinatorics – like those of analytic number theory – are often best thought of as a collection of methods and techniques with wide applicability, and from this point a view, even a proof that has been superseded in certain respects may still be very important.

1.6. Some basic facts

The arguments in additive combinatorics repeatedly make use of a number of simple techniques and inequalities. These are often simple enough to invoke without reference or to re-implement in each argument (avoiding the necessity to compare notation between a general lemma and a specific application), and we will do this also in this book. However, especially when learning the material, it helps to be aware of the general nature of these steps.

One of the most basic tools is the Cauchy–Schwarz inequality. In its simplest form, it is an upper-bound

$$\Big|\sum_{n=1}^{N}\alpha_n\beta_n\Big| \leqslant \Big(\sum_{n=1}^{N}|\alpha_n|^2\Big)^{1/2}\Big(\sum_{n=1}^{N}|\beta_n|^2\Big)^{1/2},$$

or a "weighted" upper-bound

$$\Big|\sum_{n=1}^{\mathrm{N}} r_n\alpha_n\beta_n\Big| \leqslant \Big(\sum_{n=1}^{\mathrm{N}} r_n|\alpha_n|^2\Big)^{1/2}\Big(\sum_{n=1}^{\mathrm{N}} r_n|\beta_n|^2\Big)^{1/2},$$

both for arbitrary complex numbers α_n and β_n, and the second with $r_n \geqslant 0$. But it is also used frequently in "reversed" forms, to prove lower-bounds, such as

$$\sum_{n=1}^{\mathrm{N}} \alpha_n \geqslant \frac{1}{\mathrm{N}}\Big(\sum_{n=1}^{\mathrm{N}} \alpha_n^{1/2}\Big)^2, \qquad \sum_{n=1}^{\mathrm{N}} \alpha_n^2 \geqslant \frac{\Big(\displaystyle\sum_{n=1}^{\mathrm{N}} r_n\alpha_n\Big)^2}{\displaystyle\sum_{n=1}^{\mathrm{N}} r_n^2},$$

assuming that $\alpha_n \in \mathbf{R}$. Moreover, the focus can be on the set of n where $\alpha_n \neq 0$, in which case we have another pair of variants, namely

$$\sum_{\substack{1\leqslant n\leqslant \mathrm{N}\\ \alpha_n\neq 0}} 1 \geqslant \frac{\Big(\displaystyle\sum_{n=1}^{\mathrm{N}} \alpha_n\Big)^2}{\displaystyle\sum_{n=1}^{\mathrm{N}} \alpha_n^2}, \qquad \sum_{\substack{1\leqslant n\leqslant \mathrm{N}\\ \alpha_n\neq 0}} r_n \geqslant \frac{\Big(\displaystyle\sum_{n=1}^{\mathrm{N}} r_n\alpha_n\Big)^2}{\displaystyle\sum_{n=1}^{\mathrm{N}} r_n\alpha_n^2}.$$

Another basic tool is a lemma concerning the evaluation of "bilinear forms", which occur very frequently in analytic number theory, analysis and combinatorics.

LEMMA 1.6.1. *Let $\mathrm{N} \geqslant 1$ be an integer, and let $(\alpha_{m,n})_{m,n\in[\mathrm{N}]}$ be complex numbers. For any complex numbers $(\beta_m)_{m\in[\mathrm{N}]}$ and $(\gamma_n)_{n\in[\mathrm{N}]}$, we have*

$$\Big|\sum_{1\leqslant m,n\leqslant \mathrm{N}} \alpha_{m,n}\beta_m\gamma_n\Big|^2 \leqslant \sum_{m\leqslant \mathrm{N}} |\beta_m|^2 \times \sum_{n_1,n_2\leqslant \mathrm{N}} \overline{\gamma_{n_1}}\gamma_{n_2}\Delta(n_1,n_2),$$

where

$$\Delta(n_1,n_2) = \sum_{1\leqslant m\leqslant \mathrm{N}} \overline{\alpha_{m,n_1}}\alpha_{m,n_2}.$$

In particular, we have

$$\Big|\sum_{1\leqslant m,n\leqslant \mathrm{N}} \alpha_{m,n}\beta_m\gamma_n\Big|^2 \leqslant \sum_{m\leqslant \mathrm{N}} |\beta_m|^2 \times \sum_{n\leqslant \mathrm{N}} |\gamma_n|^2 \times \max_{1\leqslant n_1\leqslant \mathrm{N}} \sum_{n_2\leqslant \mathrm{N}} |\Delta(n_1,n_2)|.$$

PROOF. We write

$$\sum_{1\leqslant m,n\leqslant \mathrm{N}} \alpha_{m,n}\beta_m\gamma_n = \sum_{m\leqslant \mathrm{N}} \beta_m \sum_{n\leqslant \mathrm{N}} \alpha_{m,n}\gamma_n,$$

and apply the Cauchy–Schwarz inequality to get

$$\begin{aligned}\Big|\sum_{1\leqslant m,n\leqslant \mathrm{N}} \alpha_{m,n}\beta_m\gamma_n\Big|^2 &\leqslant \Big(\sum_{m\leqslant \mathrm{N}} |\beta_m|^2\Big) \times \Big(\sum_{n_1,n_2\leqslant \mathrm{N}} \overline{\gamma_{n_1}}\gamma_{n_2} \sum_{m\leqslant \mathrm{N}} \overline{\alpha_{m,n_1}}\alpha_{m,n_2}\Big)\\ &= \Big(\sum_{m\leqslant \mathrm{N}} |\beta_m|^2\Big) \times \Big(\sum_{n_1,n_2\leqslant \mathrm{N}} \overline{\gamma_{n_1}}\gamma_{n_2}\Delta(n_1,n_2)\Big),\end{aligned}$$

which is the first part of the lemma.

For the second, note that

$$\begin{aligned}\Big|\sum_{n_1,n_2\leqslant \mathrm{N}} \overline{\gamma_{n_1}}\gamma_{n_2}\Delta(n_1,n_2)\Big| &\leqslant \sum_{n_1,n_2\leqslant \mathrm{N}} |\gamma_{n_1}\gamma_{n_2}||\Delta(n_1,n_2)| \\ &\leqslant \frac{1}{2}\sum_{n_1,n_2\leqslant \mathrm{N}} (|\gamma_{n_1}|^2+|\gamma_{n_2}|^2)|\Delta(n_1,n_2)| \\ &\leqslant \sum_{n_1} |\gamma_{n_1}|^2 \sum_{n_2\leqslant \mathrm{N}} |\Delta(n_1,n_2)|,\end{aligned}$$

which gives the second inequality. □

REMARK 1.6.2. A key interest of this lemma is that it "splits" the problem of estimating the sums involving the various quantities $\alpha_{m,n}$, β_m and γ_n in different problems, namely that of computing the norms of the β's and γ's, and that of understanding the "correlations" in $\Delta(n_1,n_2)$. (In terms of matrices, note that $\Delta(n_1,n_2)$ is the standard inner-product of two of the columns of the matrix $(\alpha_{m,n})$).

Prerequisites and notation

The basic requirements for most of this text are standard introductory graduate courses in algebra and real analysis. Some knowledge of other topics (especially Lebesgue integration, probability theory and the basic theory of finite fields) is useful, but most of what is needed is very elementary and will be described from scratch.

Our conventions for discrete Fourier analysis are described in Section A.7; in particular, we note that the Fourier transform will always be normalized so that it is a unitary operator, and that we use the *normalized* convolution, defined by

$$(f * g)(x) = \frac{1}{|\mathrm{G}|} \sum_{y \in \mathrm{G}} f(y) g(y^{-1}x)$$

for two complex-valued functions f and g on a finite group G.

We will use the following notation:

(1) For any integer $n \geqslant 0$, we denote $[n] = \{1, \dots, n\}$; if $n = 0$, this is the empty set. More generally, we write $[n; m] = \{n, n+1, \dots, m\}$ for any integers $n \leqslant m$ in $\mathbf{Z}$.

(2) For subsets Y_1 and Y_2 of an arbitrary set X, we denote by $\mathrm{Y}_1 - \mathrm{Y}_2$ the difference set, i.e., the set of elements $x \in \mathrm{Y}_1$ such that $x \notin \mathrm{Y}_2$.

(3) For a set X, $|\mathrm{X}| \in [0, +\infty]$ denotes its cardinal, with $|\mathrm{X}| = \infty$ if X is infinite. There is no distinction in this text between the various infinite cardinals.

(4) If X is a set and f, g two complex-valued functions on X, then we write synonymously $f = \mathrm{O}(g)$ or $f \ll g$ to say that there exists a constant $\mathrm{C} \geqslant 0$ (sometimes called an "implied constant") such that $|f(x)| \leqslant \mathrm{C}g(x)$ for all $x \in \mathrm{X}$. Note that this implies that in fact $g \geqslant 0$. We also write $f \asymp g$ to indicate that $f \ll g$ and $g \ll f$.

(5) We write $a \mid b$ for the divisibility relation "a divides b"; we denote by (a, b) the gcd of two integers a and b, and by $[a, b]$ their lcm.

(6) We denote by $\mathbf{F}_p$ the finite field $\mathbf{Z}/p\mathbf{Z}$, for p prime, and more generally by $\mathbf{F}_q$ a finite field with q elements, where $q = p^n$, $n \geqslant 1$, is a power of p. The basic theory is reviewed in Section A.6 in the Appendix.

(7) For a complex number z, we write $e(z) = e^{2i\pi z}$. If $q \geqslant 1$ and $x \in \mathbf{Z}/q\mathbf{Z}$, then $e(x/q)$ is then well-defined by taking any representative of x in $\mathbf{Z}$ to compute the exponential.

(8) If $q \geqslant 1$ and $x \in \mathbf{Z}$ (or $x \in \mathbf{Z}/q\mathbf{Z}$) is an integer which is coprime to q (or a residue class invertible modulo q), we sometimes denote by $\bar{q}$ the inverse class such that $x\bar{x} = 1$ in $\mathbf{Z}/q\mathbf{Z}$. This will always be done in such a way that the modulus q is clear from context, in the case where x is an integer.

(9) Given a probability space $(\Omega, \Sigma, \mathbf{P})$, we denote by $\mathbf{E}(\cdot)$ (resp. $\mathbf{V}(\cdot)$) the expectation (resp. the variance) computed with respect to $\mathbf{P}$. If X is a non-empty finite set, it is often viewed as a probability space with the uniform measure without this being specifically mentioned; we then sometimes write for instance

$$\mathop{\mathbf{E}}_{x\in \mathrm{X}} f(x) = \frac{1}{|\mathrm{X}|}\sum_{x\in \mathrm{X}} f(x)$$

for the corresponding averages, or even $\mathbf{E}_x f(x)$ if the set X is clear in context.

(10) Given a measure space (Ω, Σ, μ) (not necessarily a probability space), a set Y with a σ-algebra Σ' and a measurable map $f : \Omega \longrightarrow \mathrm{Y}$, we denote by $f_*(\mu)$ (or sometimes $f(\mu)$) the image measure on Y; in the case of a probability space, so that f is seen as a random variable on Ω, this is the probability law of f seen as a "random Y-valued element". If the set Y is given without specifiying a σ-algebra, we will view it usually as given with the σ-algebra generated by sets $\mathrm{Z} \subset \mathrm{Y}$ such that $f^{-1}(\mathrm{Z})$ belongs to Σ.

(11) Let G be an abelian group, with additive notation. An *arithmetic progression* $\mathrm{A} \subset \mathrm{G}$ is a set of the form

$$\mathrm{A} = a_0 + \mathrm{I}a = \{a_0 + ka \mid k \in \mathrm{I}\}$$

for some elements a_0 and a of G and some interval I of $\mathbf{Z}$. The element a is called the common difference of A; it is unique if A is not empty. If A is finite, then the size of A is also often called the *length* of the arithmetic progression.

Let further $d \geqslant 1$ be an integer. A d-dimensional *generalized arithmetic progression* is a set of the form

$$\mathrm{A} = \{a_0 + k_1a_1 + \cdots + k_da_d \mid k_i \in \mathrm{I}_i\}$$

where $a_i \in \mathrm{G}$ and I_i are intervals in $\mathbf{Z}$. This is called a *proper generalized arithmetic progression* if the equation

$$a_0 + k_1a_1 + \cdots + k_da_d = a_0 + m_1a_1 + \cdots + m_da_d$$

with (k_i) and (l_i) in $\mathrm{I}_1 \times \cdots \times \mathrm{I}_d$ implies $k_i = l_i$ for all i; if A is finite, this is equivalent to the fact that $|\mathrm{A}| = |\mathrm{I}_1| \cdots |\mathrm{I}_d|$.

CHAPTER 2

Product sets

2.1. Definition and notation

This chapter is devoted to the basic theory of "sumsets", which we will rather call "product sets", since we consider general groups, and not only commutative ones. The focus is to a large extent on sets with extremal behavior. There are two very different aspects: (1) sets with *very large* (even maximal) product sets, which are called *Sidon sets*; (2) sets with small subsets. The latter are more interesting: they appear in many applications, and moreover some classification can be attempted in very general groups, with important consequences. Sidon sets, by contrast, are extremely elusive.

We first introduce some basic definitions and notation.

Definition 2.1.1 (Product sets). Let G be a group, with group law written multiplicatively.

(1) For any subsets A and B of G, we denote by $A \cdot B$ or by AB the *product set*

$$A \cdot B = \{x \in G \mid \text{there exist } (a,b) \in A \times B \text{ such that } ab = x\}.$$

If A contains a single element a, then we write $a \cdot B$ or aB for $\{a\} \cdot B$.

(2) For any subset A of G and integers $k \in \mathbf{Z}$, we denote $A^{(k)}$ the k-fold product set, defined inductively by $A^{(0)} = \{1\}$, $A^{(1)} = A$, and

$$A^{(k+1)} = A^{(k)} \cdot A$$

for $k \geqslant 0$, while $A^{(-1)} = \{a \in G \mid a^{-1} \in A\}$ and

$$A^{(-k-1)} = A^{(-k)} \cdot A^{(-1)}$$

for $k \geqslant 0$.

We also write A^{-1} for $A^{(-1)}$, since there is then no ambiguity.

If G is commutative, we sometimes write kA instead of $A^{(k)}$ and $-A$ instead of A^{-1}.

Remark 2.1.2. In the context of an abelian group G with additive notation, and a subset A of G, one must be careful to distinguish between two potential meanings of 2A: it may denote either $A + A$, or the set of elements of the form $2a$ for $a \in A$.

The definitions and the associativity of the group operation give the basic relations

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C$$

for arbitary subsets of G. If m and n are both non-negative or both non-positive, then we also have

$$A^{(m+n)} = A^{(m)} \cdot A^{(n)},$$

but this is false in general if m and n have different signs. For instance, note that

$$A^{(1)} \cdot A^{(-1)} = \{ab^{-1} \mid (a,b) \in A \times A\},$$

and this is almost never the same as $A^{(0)} = \{1\}$.

If G is commutative, we have furthermore

$$A \cdot B = B \cdot A.$$

There are other completely elementary properties of these definitions which we just state, and usually use later on without specific mention:

(1) If $A \subset A_1$, then $A \cdot B \subset A_1 \cdot B$, and similarly if $B \subset B_1$, then $A \cdot B \subset A \cdot B_1$.

(2) For x and $y \in G$, we have $x \in y \cdot A$ if and only if $y \in x \cdot A^{-1}$.

REMARK 2.1.3. If the group G is non-commutative, there are certain elementary facts, valid for abelian groups, which fail. One of the simplest is that it is not true then that $A \cdot B$ and $B \cdot A$ have the same size. Indeed, this will hold, in any group G, for sets of the form

$$A = \{a, b\}, \qquad B = \{1, a^{-1}b\},$$

provided $a^2 \neq b^2$ (in particular $a \neq b$) and $ab \neq ba$: indeed, we have

$$A \cdot B = \{a, b, ba^{-1}b\}, \qquad B \cdot A = \{a, b, a^{-1}ba, a^{-1}b^2\},$$

and it is elementary that the last set has four elements under the stated conditions.

For example, in the smallest non-abelian group, symmetric group S_3: for instance, we can take

$$A = \{(1,3), (1,2,3)\}, \qquad B = \{1, (1,3) \circ (1,2,3)\} = \{1, (1,2)\}.$$

In the study of product sets, the following extra conditions on A tend to simplify the arguments.

DEFINITION 2.1.4 (Symmetric, neutral sets). Let G be a group. A subset $A \subset G$ is *symmetric* if it is stable under inversion, i.e., if $A = A^{-1}$. It is called *neutral*[1] if $1 \in A$.

If A is symmetric, then $A^{(k)} = A^{(-k)}$ for all integers k. If A is neutral, then

$$A^{(k)} \subset A^{(k+1)}$$

for all integers $k \geqslant 0$.

It is often neutral symmetric subsets which are the best behaved. Note that A has this property if and only if

$$A = A^{(-1)} \cup A^{(0)} \cup A^{(1)}.$$

One can sometimes reduce problems to this situation by replacing A by $A^{(-1)} \cup A^{(0)} \cup A^{(1)}$ (i.e., adding all inverses of elements of A, and the neutral element). However, another natural neutral symmetric set related to A is the product set $A \cdot A{-}1$ (assuming that A is not empty).

REMARK 2.1.5. Let $\langle A \rangle$ denote the subgroup generated by a subset $A \subset G$. We have the equality

$$\bigcup_{k \in \mathbf{Z}} (A \cup A^{-1})^{(k)} = \langle A \rangle$$

essentially by definition. In particular, a symmetric set A generates G if and only if any $x \in G$ belongs to *some* product set $A^{(k)}$ with $k \geqslant 0$.

[1] This is not a standard notation.

As suggested by the discussion at the end of Section 1.3, it is also useful to take into account the multiplicity of representations of an element as a product in a product set. For this, we assume that the sets A and B are finite, and then define the *representation function for* $A \cdot B$ by

$$r_{A,B}(x) = |\{(a,b) \in A \times B \mid ab = x\}|,$$

which is a function $G \to \mathbf{R}$ which is non-zero exactly on $A \cdot B$. On occasion, we also argue with the set of representations itself, i.e, with the set

$$R_{A,B}(x) = \{(a,b) \in A \times B \mid ab = x\},$$

which we call the *representation set* for $A \cdot B$.

If the group G itself is finite, then if we denote by φ_A and φ_B the characteristic functions of A and B, respectively, we see that the formula

$$r_{A,B}(x) = \sum_{\substack{(a,b)\in G^2 \\ ab=x}} \varphi_A(a)\varphi_B(b) = |G|\varphi_A * \varphi_B,$$

holds, where the right-hand side involves the *convolution* of the characteristic functions of A and B (see Definition A.7.8).

As a consequence, or by direct computation, we have

$$\frac{1}{|G|}\sum_{x\in G} r_{A,B}(x) = \frac{|A||B|}{|G|}, \tag{2.1}$$

generalizing (1.2).

Remark 2.1.6. Depending on the situation, it may be convenient to consider $\widetilde{r}_{A,B} = \varphi_A * \varphi_B$ instead of $r_{A,B}$. Adjusting the constants, this satisfies

$$\frac{1}{|G|}\sum_{x\in G} \widetilde{r}_{A,B}(x) = \frac{|A||B|}{|G|^2},$$

which is interpreted as the product of the "densities" of A and B in G.

If G is abelian, then the properties of the convolution imply that the Fourier transform of $r_{A,B}$ is proportional to the (pointwise) product of the Fourier transforms of φ_A and φ_B (see Proposition A.7.7). Similar formulas (with repeated convolutions) apply to sets like $A^{(k)}$ for $k \in \mathbf{Z}$.

Coming back to an arbitrary group G, we note also that if A and B are not empty, then we have an inequality

$$|A \cdot B| \geqslant \frac{|A||B|}{\max_{x\in G} |r_{A,B}(x)|},$$

which can sometimes be useful; this follows from the computation

$$|A||B| = \sum_{x\in G} r_{A,B}(x) = \sum_{x\in A\cdot B} r_{A,B}(x) \leqslant |A \cdot B| \max_{x\in G} r_{A,B}(x),$$

and the fact that $r_{A,B}$ is not identically zero.

Remark 2.1.7. (1) There are important links between iterated product sets $A^{(k)}$ and properties (especially connectedness) of the *Cayley graph* associated to G and A. More or less equivalently, there are links with the random walks on G defined with steps given by random variables with values in A. We will discuss briefly some of these in Section 2.6.

(2) The definition of the product set $A \cdot B$ does not require any property of the product on a group, simply that it exists (as a map $G \times G \to G$). We will sometimes use the notation in this more general context, as it already appeared for instance (in the case of the multiplication on $\mathbf{Z}$) in the statement of Theorem 1.1.3.

There are a few "obvious" statements which hold for the sizes of product sets. The fact that $|a \cdot B| = |B|$ for any set B and any element $a \in G$ is used constantly (note that it reflects the fact that we work with subsets of a *group*, so that multiplication by a fixed element is a bijection; in the case of multiplication in $\mathbf{Z}$, or any other ring, this property fails when $a = 0$). The other general bounds are the following:

(1) If A and B are non-empty, then we have
$$|A \cdot B| \geqslant \max(|A|, |B|),$$
since $|a \cdot B| = |B|$ and $a \cdot B \subset A \cdot B$ for any $a \in A$, and similarly for the bound $|A \cdot B| \geqslant |A|$.

(2) We have, for A, B arbitrary, the bound
$$|A \cdot B| \leqslant |A||B|.$$

(3) If G is commutative and $A = B$, then we have the better bound
$$|A^{(2)}| \leqslant \frac{|A|(|A|+1)}{2},$$
since the matrix $(a_1 + a_2)_{(a_1,a_2)\in A\times A}$ is then symmetric.

None of these bounds can be improved in general. The basic theory of product sets can then be described as attempts to classify first the extremal cases, and then "almost" extremal cases. This leads to completely different outcomes for the upper and the lower bounds. We will first discuss sets where $A^{(2)}$ is maximal in Section 2.3, simply because this is more straightforward (partly in the absence of really definitive results, or even conjectural statements), but the reader can very well skip immediately to Section 2.4 for the discussion of the "approximate subgroups", which are related to the extrema of the lower bounds.

2.2. Freiman homomorphisms

In general, if we want to compare the algebraic properties of two groups G and H, then we use group homomorphisms from G to H (or from H to G). In particular, two groups are considered to be identical when there is an isomorphism between them, and then any group-theoretical property of one of these groups has a perfect translation in a property of the other.

When dealing with product sets, which involve often relatively small subsets of the ambient groups, Freiman realized that some weaker notion could be used to go back and forth between non-isomorphic groups.

Definition 2.2.1. Let G and H be arbitrary groups and let $A \subset G$ and $B \subset H$ be subsets of these. Let $k \geqslant 0$ be an integer.

A map $f\colon A \to B$ is a *Freiman k-morphism* if and only if, for any tuple $(a_1, \dots, a_{2k})$ of elements of A, the condition
$$a_1 \cdots a_k = a_{k+1} \cdots a_{2k} \tag{2.2}$$
implies
$$f(a_1) \cdots f(a_k) = f(a_{k+1}) \cdots f(a_{2k}). \tag{2.3}$$

If f is bijective and its inverse is a Freiman k-morphism from B to A, then f is said to be a *Freiman k-isomorphism*.

Note that any homomorphism $f\colon \mathrm{G} \to \mathrm{H}$ induces by restriction a Freiman k-morphism $\mathrm{A} \to \mathrm{B}$ for any $k \geqslant 0$ and any subsets $\mathrm{A} \subset \mathrm{G}$ and $\mathrm{B} \supset f(\mathrm{A})$. If f is an isomorphism, then it defines a Freiman k-isomorphism from A to $f(\mathrm{A})$ for any k. In particular, the identity map from a set to itself is always a Freiman k-isomorphism.

It is also straightforward that if $f\colon \mathrm{A} \to \mathrm{B}$ and $g\colon \mathrm{B} \to \mathrm{C}$ are Freiman k-morphisms, then $g \circ f$ is also one, and similarly for k-isomorphisms.

Any Freiman k-morphism from A to B is also an l-morphism for any $l \leqslant k$: indeed, this is trivial if A is empty (in which case the unique map $\mathrm{A} \to \mathrm{B}$ is a k-morphism for any k), and otherwise given a relation

$$a_1 \cdots a_l = a_{l+1} \cdots a_{2l},$$

we obtain an equation between k-tuples by multiplying by $k - l$ elements equal to some fixed element α_0 of A on both sides, and then we cancel out the corresponding factors $f(\alpha_0)$ after applying f and the Freiman property for k factors.

Although the definition of a Freiman k-morphism only involves sums, it does have good properties also for differences. Indeed, suppose $f\colon \mathrm{A} \to \mathrm{B}$ is a Freiman 2-morphism. If $a_1, \ldots, a_4$ in A satisfy $a_1 - a_2 = a_3 - a_4$, then we get $a_1 + a_4 = a_3 + a_2$, hence $f(a_1) + f(a_4) = f(a_3) + f(a_2)$, and then $f(a_1) - f(a_2) = f(a_3) - f(a_4)$.

Suppose furthermore that $\mathrm{A} \subset \mathrm{G}$ is symmetric and that $f(1_\mathrm{G}) = 1_\mathrm{H}$. We then have $f(a^{-1}) = f(a)^{-1}$ for all $a \in \mathrm{A}$ (indeed, from $a \cdot a^{-1} = 1_\mathrm{G} \cdot 1_\mathrm{G}$ we deduce $f(a)f(a^{-1}) = 1_\mathrm{H}$ from the assumption). Note that this is not true without some condition: for instance, if $f\colon \mathrm{A} \to \mathrm{B}$ is a Freiman k-morphism and $b \in \mathrm{B}$, then $f + b\colon x \mapsto f(x) + b$ is also a Freiman k-morphism.

The following proposition is elementary but quite useful.

Proposition 2.2.2. *Let $f\colon \mathrm{A} \to \mathrm{B}$ be a Freiman k-morphism for some $k \geqslant 1$. Then f is a k-isomorphism if and only if f is surjective and if the equations*

$$a_1 \cdots a_k = a_{k+1} \cdots a_{2k}$$

and

$$f(a_1) \cdots f(a_k) = f(a_{k+1}) \cdots f(a_{2k})$$

are equivalent for all $(a_1, \ldots, a_{2k}) \in \mathrm{A}^{2k}$.

Proof. Taking $k = 1$ (as we can by the remarks above), we see that $f(a_1) = f(a_2)$ if and only if $a_1 = a_2$, so that f is injective, hence bijective by the assumptions. Then it follows also that the inverse f^{-1} is a Freiman k-morphism. □

Corollary 2.2.3. *Let* G *and* H *be arbitrary groups and let $f\colon \mathrm{G} \to \mathrm{H}$ be a group morphism. Let $k \geqslant 1$ be an integer. If $\mathrm{A} \subset \mathrm{G}$ is any subset such that the restriction of f to $\mathrm{A}^{(k)}$ is injective, then f defines by restriction a Freiman k-isomorphism from* A *to $f(\mathrm{A})$.*

Proof. Let $\mathrm{B} = f(\mathrm{A})$. Then we have by restriction a map $f\colon \mathrm{A} \to \mathrm{B}$ which is surjective and is a Freiman k-morphism. Since f is morphism of groups, the condition (2.2) implies (2.3). Conversely, suppose that $(a_i) \in \mathrm{A}^{2k}$ is such that

$$f(a_1) \cdots f(a_k) = f(a_{k+1}) \cdots f(a_{2k}).$$

The elements

$$x = a_1 \cdots a_k, \qquad y = a_{k+1} \cdots a_{2k}$$

of $A^{(k)}$ then satisfy $f(x) = f(y)$, and hence $x = y$ by assumption, which proves that the criterion of Proposition 2.2.2 is satisfied. □

The basic property of Freiman morphisms is the following lemma:

LEMMA 2.2.4. *Let $k \geqslant 2$ and let $f\colon A \to B$ be a Freiman k-morphism. For any subsets A_1 and A_2 of A, we have*

$$|f(A_1) \cdot f(A_2)| \leqslant |A_1 \cdot A_2|$$

with equality if f is a 2-isomorphism. In particular, we have $|f(A)^{(2)}| \leqslant |A^{(2)}|$.

PROOF. Partition the *cartesian product* set $A_1 \times A_2$ into sets formed with elements (a_1, a_2) such that the product $a_1 \cdot a_2$ is the same, and similarly for $f(A_1) \times f(A_2)$; call X and Y the sets of the resulting subsets, and observe that the set X has cardinality $|A_1 \cdot A_2|$, while Y has cardinality $|f(A_1) \cdot f(A_2)|$ (in other words, the elements of X are the equivalence classes for the equivalence relation on $A_1 \times A_2$ defined by $(a_1, a_2) \sim (b_1, b_2)$ if and only if $a_1 a_2 = b_1 b_2$).

Observe now that the assumption that f is a Freiman 2-morphism means that there is a well-defined map $\widetilde{f}\colon X \to Y$ which maps the set of elements with $a_1 a_2$ equal to a given value to the set in Y with the common value of $f(a_1)f(a_2)$. By construction, this map $\widetilde{f}$ is surjective, and therefore $|Y| \leqslant |X|$, which gives the stated inequality. Applying it also to f^{-1} if f is a 2-isomorphism, we get equality in that case. □

EXAMPLE 2.2.5. (1) Suppose that H is commutative. If $f\colon A \to B$ is a Freiman k-morphism, then for any $h_0 \in H$, the map $\widetilde{f}\colon A \to h_0 B$ defined by $\widetilde{f}(x) = h_0 f(x)$ is a Freiman k-morphism (and is a k-isomorphism if f is one). Indeed, from (2.2) and (2.3), and from the commutativity of H, we deduce

$$\widetilde{f}(a_1) \cdots \widetilde{f}(a_k) = h_0^k f(a_{k+1}) \cdots f(a_{2k}) = h_0^k f(a_{k+1}) \cdots f(a_{2k}) = \widetilde{f}(a_{k+1}) \cdots \widetilde{f}(a_{2k}).$$

(2) The following two examples give possibly the two most important examples of Freiman isomorphism which do not arise from a group isomorphism.

Let $q \geqslant 1$ be an integer and let $f\colon \mathbf{Z} \to \mathbf{Z}/q\mathbf{Z}$ be the reduction map. This is a group morphism, hence a Freiman k-morphism for all $k \geqslant 0$. Suppose that $k \geqslant 2$ is fixed, and that n is an integer such that $0 \leqslant kn \leqslant q$. We claim that the map induced from f by restriction to the interval $A = \{0, \dots, n-1\}$ and to $B = f(A) \subset \mathbf{Z}/q\mathbf{Z}$, is a Freiman k-isomorphism.

Indeed, the map $f\colon A \to B$ induced by restriction is of course surjective. Moreover, we have the inclusion

$$A^{(k)} \subset \{0, \dots, k(n-1)\},$$

and since $k(n-1) < kn \leqslant q$, it follows that the restriction of f to $A^{(k)}$ is injective. Thus the result follows directly from Corollary 2.2.3.

This example demonstrates the essential fact that, from the point of view of Freiman k-isomorphisms, two subsets of very different groups can "look the same" (in particular, we can pass from an infinite torsion-free group like $\mathbf{Z}$ to a finite cyclic group). From Lemma 2.2.4, it follows that one can often study product sets by moving to a different group. In particular, one can move some questions from $\mathbf{Z}$ to finite abelian groups, one advantage of which is that Fourier analysis is often much simpler in that context.

(3) The other fundametnal standard example shows that Freiman isomorphisms can "alter the apparent dimension" of a set. Let $d \geqslant 1$ be an integer and let $I_1, \dots, I_d$ be finite intervals in $\mathbf{Z}$, of length $|I_i| = n_i \geqslant 1$. Consider the subset $A = I_1 \times \cdots \times I_d$ of $\mathbf{Z}^d$.

Now pick an integer $k \geqslant 1$ and an integer $m > k\max(n_i) \geqslant 2$. Define $f\colon \mathbf{Z}^d \to \mathbf{Z}$ by

$$f(x_1, \dots, x_d) = y_1 + my_2 + m^2 y_3 + \cdots + m^{d-1} y_d, \qquad y_i = x_i - \min(\mathrm{I}_i).$$

The map f is a (surjective) affine map (i.e., a group homomorphism composed with a translation), hence is a Freiman k-isomorphism for any k. We claim that it is injective on $\mathrm{A}^{(k)}$; by Corollary 2.2.3 (adapted to affine maps), this implies that A is Freiman k-isomorphic to $f(\mathrm{A})$. Thus, a seemingly d-dimensional parallelepided is "the same" as a certain subset of "line" $\mathbf{Z}$.

The injectivity assertion is a consequence of the fact that any $x = (x_i) \in \mathrm{A}^{(k)}$ satisfies $0 \leqslant x_i < kn_i \leqslant n$, hence $f(x)$ is an integer whose expansion in base n has "digits" precisely given by $x_1, \dots, x_d$. This means we can recover x from $f(x)$, proving the desired injectivity.

For later purposes, as it can be important in applications, we note also that this construction gives a good control of the size of an interval containing the image of f: indeed, suppose for simplicity that $n_i = n$ for all i, so that we can take $m = kn + 1$. We then get $f(\mathrm{A}) \subset [\mathrm{N}]$ with $\mathrm{N} \leqslant m^d = (kn+1)^d$.

(4) Freiman morphisms can be used to characterize certain additive structures. For instance, let G be an abelian group, and let a_0 and a be elements of G and I an interval in $\mathbf{Z}$. The map $f\colon \mathrm{I} \to \mathrm{G}$ defined by $f(i) = a_0 + ia$ is then a Freiman k-morphism for any $k \geqslant 1$, with image equal to the arithmetic progression $a_0 + [\mathrm{I}]a$. Now, conversely, suppose $g\colon \mathrm{I} \to \mathrm{G}$ is a Freiman 2-morphism; we claim that $g(\mathrm{I})$ is an arithmetic progression, so that arithmetic progressions are characterized as images of Freiman 2-morphisms defined on intervals in $\mathbf{Z}$.

To check the claim, notice that we may assume that I has at least two elements, since otherwise the result is clear. Pick then some fixed $j \in \mathrm{I}$ such that $j+1 \in \mathrm{I}$, and define $b = g(j+1) - g(j) \in \mathrm{G}$. Whenever $i \in \mathrm{I}$ is such that $i+1 \in \mathrm{I}$, the equation $(i+1) - i = (j+1) - j$ implies that $g(i+1) - g(i) = b$. By induction on $i \in \mathrm{I}$, we conclude that $g(i) = g(j) + (j-i)b$ (the induction is done separately over $i \geqslant j$ in I and over $i \leqslant j$ in I; if I has a minimum or a maximum, one of these inductions can be avoided by picking j to be the minimum or maximum of I).

2.3. Sidon sets

If G is non-commutative, then the upper-bound $|\mathrm{A}\cdot\mathrm{B}| = |\mathrm{A}||\mathrm{B}|$ is relatively frequently an equality. For a group like $\mathrm{GL}_n(\mathbf{R})$ or $\mathrm{GL}_n(\mathbf{Z})$, there is no hope at all of saying anything new or interesting about sets which achieve this bound.

However, for abelian groups, there is definitely some interest in the study of sets with $\mathrm{A}^{(2)}$ as large as possible.

Throughout this section, all groups are assumed to be commutative unless otherwise specified. We use additive notation unless we deal with subgroups of multiplicative groups.

Definition 2.3.1. Let G be an abelian group. A subset $\mathrm{A} \subset \mathrm{G}$ is called a *Sidon set* if the Sidon equation

$$a + b = c + d$$

with $(a,b,c,d) \in \mathrm{A}^4$ implies $a \in \{c,d\}$. Equivalently, the equation

$$a - b = c - d$$

with $(a,b,c,d) \in \mathrm{A}^4$ such that $a \neq b$ implies that $a = c$ and $b = d$.

REMARK 2.3.2. (1) If $A \subset G$ is finite, then A is a Sidon set if and only if $|A^{(2)}| = |A|(|A|+1)/2$, but as we will see, there are some fairly interesting infinite examples.

(2) If $A \subset G$ is a Sidon set and $B \subset H$ is a subset of another abelian group, and if $f\colon A \to B$ is a Freiman 2-morphism, then the image $f(A) \subset B$ is also a Sidon set, by Lemma 2.2.4.

There is a slightly different but slightly stronger statement: if $f\colon G \to H$ is a group homomorphism, and if $A \subset G$ is a subset such that the restriction of f to A is injective, and such that $f(A)$ is a Sidon set in H, then A is a Sidon set. Indeed, if $(a,b,c,d) \in A^4$ satisfy $a+b=c+d$, then we get $f(a)+f(b)=f(c)+f(d)$, an equation with all terms in $f(A)$, so that $f(a) \in \{f(c), f(d)\}$, which means that $a \in \{c,d\}$ by injectivity of f on A.

(3) An equivalent form of the definition (for A finite) is that $r_{A,A}(x) \leqslant 2$ for all x.

(4) Certain authors (see, e.g., the paper [1] of Babai and Sós) define a Sidon set $A \subset G$ to be a subset such that whenever (a,b,c,d) are elements of A with three at least of them distinct, we have $a+b \neq c+d$. This coincides with the definition above if G has no 2-torsion, but otherwise it allows A to contains elements $a \neq b$ such that $2a = 2b$, which Definition 2.3.1 excludes in a Sidon set.

In particular, if G is an abelian group where $2x = 0$ for all x, then G contains no Sidon set of size $\geqslant 2$, since we have equations $x+x = y+y$ with $x \neq y$.

EXAMPLE 2.3.3. To say that $A \subset G$ is a Sidon set is a form of restricted "linear independance" of the elements of A. Thus there are immediate "tautological" examples: for any set X, we obtain a Sidon set A in bijection with X in the free abelian group $G = \mathbf{Z}^{(X)}$ by taking A to be the canonical basis of G, since the exact linear independance of the elements of A imply *a fortiori* what is required to have a Sidon set.

Furthermore (and this maybe explains why there is no real "theory" of Sidon sets, but only examples and applications of the defining property) *all* Sidon sets of a given (finite) size $n \geqslant 1$ are Freiman 2-isomorphic to the canonical basis of $\mathbf{Z}^n$.

Indeed, given an abelian group G and a Sidon set $A \subset G$ with $|A| = n$, there exists (by the "universal property" of free abelian groups) a group morphism $f\colon \mathbf{Z}^A \to G$ such that $f(e_a) = a$ for any $a \in A$, where $(e_a)_{a\in A}$ is the canonical basis of $\mathbf{Z}^A$ (concretely, we have

$$f\Big(\sum_{a\in A} n_a e_a\Big) = \sum_{a\in A} n_a a$$

for any integers $n_a \in \mathbf{Z}$, with the right-hand sum computed in G). This morphism is of course injective on the set $B = \{e_a\}$ of the elements of the canonical basis, so that it defines by restriction a Freiman 2-isomorphism from B to $f(B) = A$ by Corollary 2.2.3. Since $\mathbf{Z}^A$ is isomorphic (as a group) to $\mathbf{Z}^n$, this gives the claim.

EXAMPLE 2.3.4. The set of prime numbers is a Sidon set in the group $\mathbf{Q}^\times$ of invertible rational numbers; equivalently, the set of numbers $\log p$, for p prime, is a Sidon set in $\mathbf{R}$.

EXAMPLE 2.3.5. Let $r \geqslant 2$ be a real number. Any geometric progression with common ratio r is a Sidon set in $\mathbf{Z}$, since the equation $r^i + r^j = r^k + r^l$ with i, j, k, l non-negative integers has only the solutions $(i,j) = (k,l)$ and $(i,j) = (l,k)$.

More generally, let $A \subset \mathbf{R}_+^\times$ be any set of positive real numbers such that, whenever $a < b$ are elements of A, we have $b \geqslant ra$. Then A is a Sidon set in the additive group $\mathbf{R}$. Indeed, assume that $(a,b,c,d) \in A^4$ satisfy $a+b=c+d$. If a is the largest of the four, then we have

$$a+b > a \geqslant 2\max(c,d) \geqslant c+d,$$

hence the result.

This result does not always hold for geometric progressions whose common ratio is not a real number $\geqslant 2$. For instance, one checks easily that the equation

$$1 + r^4 = r^2 + r^3$$

has a solution $r = 1.32471... > 1$. But note that if $r \in \mathbf{C}$ is *transcendental*, the set of its powers is a Sidon set in $\mathbf{C}$.

Example 2.3.6. There are many (often "folklore") conjectures concerning the fact that "concrete" sets should be Sidon sets. Here are two:

- Is the set of fifth powers of positive integers a Sidon set in $\mathbf{Z}$?
- Is the set of positive ordinates of zeros of the Riemann zeta function a Sidon set in $\mathbf{R}$?

Note that the first question involves the smallest degree where the answer could be Yes: the formulas

$$1^2 + 7^2 = 5^2 + 5^2 = 50,$$
$$10^3 + 9^3 = 1^3 + 12^3 = 1729,$$
$$1584^4 + 594^4 = 1344^4 + 1334^4 = 635318657 \qquad \text{(Euler)}$$

show that the sets of squares, cubes or fourth powers of positive integers are *not* Sidon sets in $\mathbf{Z}$.

Although we motivated the definition of Sidon sets using their extremal sumset property, the main applications of Sidon set belong to harmonic analysis and related areas, and can be quite surprising. The key fact in the original study by Sidon [75] is the following simple property of trigonometric polynomials supported on a Sidon subset of $\mathbf{Z}$.

Proposition 2.3.7. *Let $\mathrm{A} \subset \mathbf{Z}$ be a finite Sidon set. For any family of complex coefficients $(\lambda_a)_{a\in\mathrm{A}}$, we have*

$$\int_0^1 \Big|\sum_{a\in\mathrm{A}} \lambda_a e(at)\Big|^4 dt = 2\Big(\sum_{a\in\mathrm{A}} |\lambda_a|^2\Big)^2 - \sum_{a\in\mathrm{A}} |\lambda_a|^4$$
$$= 2\Big(\int_0^1 \Big|\sum_{a\in\mathrm{A}} \lambda_a e(at)\Big|^2 dt\Big)^2 - \sum_{a\in\mathrm{A}} |\lambda_a|^4.$$

In particular, we have

$$\int_0^1 \Big|\sum_{a\in\mathrm{A}} \lambda_a e(at)\Big|^4 dt \leqslant 2\Big(\int_0^1 \Big|\sum_{a\in\mathrm{A}} \lambda_a e(at)\Big|^2 dt\Big)^2.$$

The point is that, in general, the left-hand side (often called the *fourth moment* of the trigonometric polynomial) only satisfies much weaker bounds in terms of the second moment. Sidon used this fact (among other tools) to prove a result concerning the Fourier coefficients of continuous periodic functions (roughly speaking, if $\mathrm{A} \subset \mathbf{Z}$ is an infinite Sidon set, and if $(\lambda_a)_{a\in\mathrm{A}}$ are complex numbers such that

$$\sum_{a\in\mathrm{A}} |\lambda_a|^2 < +\infty,$$

then there exists a 1-periodic continuous function $\varphi\colon \mathbf{R} \to \mathbf{C}$ such that

$$\lambda_a = \int_0^1 \varphi(t) e(-at) dt$$

for all $a \in \mathrm{A}$, i.e., the λ_a's are the Fourier coefficients of φ for these frequencies).

PROOF. By writing $|z|^4 = z \cdot z \cdot \bar{z} \cdot \bar{z}$ and expanding the resulting sum when z is the value at t of the trigonometric polynomial, we get

$$\Big|\sum_{a\in\mathrm{A}} \lambda_a e(at)\Big|^4 = \sum_{a,b,c,d\in\mathrm{A}} \lambda_a\lambda_b\overline{\lambda_c\lambda_d}\; e\big((a+b-c-d)t\big)$$

for $t \in \mathbf{R}$. Integrating over t using the orthogonality relation

$$\int_0^1 e(ht)dt = \begin{cases} 1 & \text{if } h = 0, \\ 0 & \text{if } h \neq 0, \end{cases}$$

for $h \in \mathbf{Z}$, we deduce

$$\int_0^1 \Big|\sum_{a\in\mathrm{A}} \lambda_a e(at)\Big|^4 dt = \sum_{\substack{a,b,c,d\in\mathrm{A}\\ a+b=c+d}} \lambda_a\lambda_b\overline{\lambda_c\lambda_d}.$$

The usefulness of the Sidon condition is now obvious, as it allows us to fully parameterize the solutions of the equation $a + b = c + d$. One must simply be a bit careful of possible multiplicity. Precisely, we will sum over the possible values of (c, d) for each value of $(a, b) \in \mathrm{A}^2$.

If $a = b$, the equation for (c, d) is $a + a = c + d$, which admits the unique solution $c = d = a$ by the Sidon condition. Hence the corresponding contribution is

$$\sum_{a\in\mathrm{A}} \lambda_a\lambda_a\overline{\lambda_a\lambda_a} = \sum_{a\in\mathrm{A}} |\lambda_a|^4.$$

If $a \neq b$, the equation $a + b = c + d$ admits the two solutions $(c, d) = (a, b)$ and $(c, d) = (b, a)$ by the Sidon condition, and the corresponding contribution is

$$\sum_{\substack{a,b\in\mathrm{A}\\ a\neq b}} \Big(\lambda_a\lambda_b\overline{\lambda_a\lambda_b} + \lambda_a\lambda_b\overline{\lambda_b\lambda_a}\Big) = 2\sum_{\substack{a,b\in\mathrm{A}\\ a\neq b}} |\lambda_a|^2|\lambda_b|^2.$$

Adding these two contributions, we obtain

$$\int_0^1 \Big|\sum_{a\in\mathrm{A}} \lambda_a e(at)\Big|^4 dt = 2\sum_{\substack{a,b\in\mathrm{A}\\ a\neq b}} |\lambda_a|^2|\lambda_b|^2 + \sum_{a\in\mathrm{A}} |\lambda_a|^4.$$

Noting the relation

$$\sum_{\substack{a,b\in\mathrm{A}\\ a\neq b}} |\lambda_a|^2|\lambda_b|^2 + \sum_{a\in\mathrm{A}} |\lambda_a|^4 = \sum_{a,b\in\mathrm{A}} |\lambda_a|^2|\lambda_b|^2 = \Big(\sum_{a\in\mathrm{A}} |\lambda_a|^2\Big)^2,$$

we see that this gives the first stated formula. The second just follows from the Parseval identity

$$\sum_{a\in\mathrm{A}} |\lambda_a|^2 = \int_0^1 \Big|\sum_{a\in\mathrm{A}} \lambda_a e(at)\Big|^2 dt$$

(which is straightforward here since A is finite). □

The proposition above admits the following immediate variant for any finite abelian group (the proof is left as an exercise; there is another variant for any compact abelian group, the case of Fourier series corresponding to $\mathbf{R}/\mathbf{Z}$).

PROPOSITION 2.3.8. *Let* G *be a finite abelian group with dual group* $\widehat{\mathrm{G}}$. *Let* $\mathrm{A} \subset \widehat{\mathrm{G}}$ *be a Sidon set. For any family of complex coefficients* $(\lambda_\chi)_{\chi\in\mathrm{A}}$, *we have*

$$\mathop{\mathbf{E}}_{x\in\mathrm{G}}\Big(\Big|\sum_{\chi\in\mathrm{A}}\lambda_\chi\chi(x)\Big|^4\Big) = 2\Big(\sum_{\chi\in\mathrm{A}}|\lambda_\chi|^2\Big)^2 - \sum_{\chi\in\mathrm{A}}|\lambda_\chi|^4.$$

We refer to the work of Forey, Fresán and Kowalski [32, § 8.4] for a discussion of applications of certain Sidon sets in arithmetic geometry.

The most natural questions about Sidon sets (both from the point of view of intellectual curiosity as from that of the type of applications discussed by Sidon) are then about constructing Sidon sets in a given abelian group G which are "as large as possible". This can be expressed in different variants, among which the following are the most popular:

– How "dense" can a Sidon set $\mathrm{A} \subset \mathbf{Z}$, in the sense of "maximizing" the function

$$\mathrm{N} \mapsto \frac{1}{\mathrm{N}}|\mathrm{A} \cap \{-\mathrm{N}, \dots, 0, \dots, \mathrm{N}\},$$

(which can take different meanings, depending on how we want to compare functions) and similarly for a Sidon set among the non-negative integers?

– Given an integer $\mathrm{N} \geqslant 1$, how dense can a finite Sidon set $\mathrm{A} \subset [\mathrm{N}]$ be? Of course, any infinite Sidon set in $\mathbf{Z}$ gives one in $[\mathrm{N}]$ by intersection, but it could be that there are "better" sets for a finite integer N that do not extend to very large infinite Sidon sets.

– Given a finite abelian group G, how large can a Sidon set $\mathrm{A} \subset \mathrm{G}$ be?

We will give some of the basic results concerning all of these questions. First, we present what are essentially all the "densest" known Sidon sets in finite abelian groups: these are five infinite families, indexed by powers of prime numbers (or equivalently by finite fields), of pairs (A, G) where $|\mathrm{A}|$ is very close to $|\mathrm{G}|^{1/2}$. Note that $|\mathrm{G}|^{1/2}$ is the best possible one can hope for: if $\mathrm{A} \subset \mathrm{G}$ is a Sidon subset of a finite group, then the differences $a - b$ for $a \neq b$ are all distinct in G, so that $|\mathrm{A}|(|\mathrm{A}| - 1) \leqslant |\mathrm{G}|$.

We present these dense Sidon sets twice: first, as separate examples, as they arose historically (sometimes being discovered independently more than once), then through a recent uniform construction of Eberhard and Manners [24] (there is another, very different, uniform construction due to Forey, Fresán and Kowalski [33], but we won't discuss it).

EXAMPLE 2.3.9. (1) Let E be a field (possibly infinite). Consider the abelian group $\mathrm{G} = \mathrm{E} \times \mathrm{E}^\times$. Define then

$$\mathrm{A} = \{(x, x) \in \mathrm{G} \mid x \in \mathrm{E}^\times\},$$

the "diagonal" in G.

The set A is a Sidon set. Indeed, given elements a, b, c, d in $\mathrm{E}^\times$, the equation

$$(a, a) \cdot (b, b) = (c, c) \cdot (d, d)$$

in the group G is equivalent to the *system* of equations

$$\begin{cases} a + b = c + d \\ ab = cd \end{cases} \tag{2.4}$$

in E. Since these equations imply that $\{a, b\}$ and $\{c, d\}$ are both the solution sets to the quadratic equation

$$\mathrm{X}^2 - (a + b)\mathrm{X} + ab = \mathrm{X}^2 - (c + d)\mathrm{X} + cd = 0,$$

we have $\{a,b\} = \{c,d\}$, hence $a \in \{c,d\}$.

If we want to have a finite Sidon set, we take E to be a finite field, of size q say (for instance, $\mathrm{E} = \mathbf{Z}/p\mathbf{Z}$ with p prime); then G has size $q(q-1)$ and A has size $q-1$, which is extremely close to $\sqrt{|\mathrm{G}|}$. The structure of G can in fact be determined exactly: if $q = p^\nu$ for p prime and $\nu \geqslant 1$, then it is well-known that E is isomorphic to $(\mathbf{Z}/p\mathbf{Z})^\nu$ as abelian group, whereas $\mathrm{E}^\times$ is cyclic of order $q-1$, so G is isomorphic to $(\mathbf{Z}/p\mathbf{Z})^\nu \times \mathbf{Z}/(q-1)\mathbf{Z}$. In the case where $\mathrm{E} = \mathbf{Z}/p\mathbf{Z}$ (namely, when $\nu = 1$), the Chinese Remainder Theorem shows that G is isomorphic to the cyclic group $\mathbf{Z}/p(p-1)\mathbf{Z}$, since p and $p-1$ are coprime.

This example is due independently to Spence (cited by Ganley in [38, p. 328]) and to Rusza [70, Th. 4.4].

(2) Again, let E be an arbitrary field. Let $\mathrm{G} = \mathrm{E}^\times \times \mathrm{E}^\times$, and define

$$\mathrm{A} = \{(x, 1-x) \in \mathrm{G} \mid x \in \mathrm{E}^\times - \{1\}\}.$$

This is again a Sidon set, and the reason is very similar to the first example: for a, b, c, d in $\mathrm{E}^\times - \{1\}$, the equation

$$(a, 1-a) \cdot (b, 1-b) = (c, 1-c) \cdot (d, 1-d)$$

in the group G is equivalent to the *system* of equations

$$\begin{cases} ab = cd \\ (1-a)(1-b) = (1-c)(1-d) \end{cases} \quad \text{or} \quad \begin{cases} ab = cd \\ 1-(a+b)+ab = 1-(c+d)+cd \end{cases}$$

in E, which is equivalent to the system (2.4) above, hence has only the solutions where $a \in \{c,d\}$.

When E is finite of size q, then G has size $(q-1)^2$ (more precisely, the group G is isomorphic to $(\mathbf{Z}/(q-1)\mathbf{Z})^2$) and A has size $q-2 = \sqrt{|\mathrm{G}|} - 1$; this example is due to Hughes [52] and Cilleruelo [17, Ex. 3].

(3) Let now E be a field of characteristic different from 2. Take $\mathrm{G} = \mathrm{E} \times \mathrm{E}$, and consider

$$\mathrm{A} = \{(x, x^2) \in \mathrm{G} \mid x \in \mathrm{E}\}$$

(in other words, the "parabola" which is the graph of the squaring function). It is again a Sidon set, for pretty much the same reasons as before: if (a,b,c,d) are in E then

$$(a, a^2) + (b, b^2) = (c, c^2) + (d, d^2)$$

if and only if

$$\begin{cases} a+b = c+d \\ a^2+b^2 = c^2+d^2. \end{cases}$$

The standard identities

$$ab = \frac{(a+b)^2 - a^2 - b^2}{2}, \qquad a^2+b^2 = (a+b)^2 - 2ab,$$

valid since E has characteristic different from 2, show that the system is also equivalent to (2.4), so that the solutions satisfy $a \in \{c,d\}$.

When E is finite of size $q = p^\nu$ with p prime, then G has size q^2 and is isomorphic to $(\mathbf{Z}/p\mathbf{Z})^{2\nu}$, while A has size $q = \sqrt{|\mathrm{G}|}$. This example was in fact the first of those discussed above, and is due to Erdős and Turán [30].

(4) The next two examples seem to be based on slightly different principles, but in fact turn out to be again ultimately similar. Here E is again a field, but we now also consider an extension F/E of degree 2. We fix an element $\varepsilon \in \mathrm{F}$ which is not in E (so

that $\mathrm{F} = \mathrm{E}(\alpha)$ necessarily). We consider the group $\mathrm{G} = \mathrm{F}^\times$ and set $\mathrm{A} = \alpha + \mathrm{E} \subset \mathrm{F}$. Note that $\mathrm{A} \subset \mathrm{G}$; we claim that this is again a Sidon set.

Suppose indeed that a, b, c, d are elements of E and that

$$(\alpha + a)(\alpha + b) = (\alpha + c)(\alpha + d).$$

This means that $(a + b)\alpha + ab = (c + d)\alpha + cd$, and since $(1, \alpha)$ is basis of F as an E-vector space, we obtain once more the system (2.4), so that $a \in \{c, d\}$ as before.

When E is finite of size $q = p^\nu$ with p prime, then there is a unique choice of F (up to isomorphism of fields); the group G has size $q^2 - 1$ and is isomorphic to $\mathbf{Z}/(q^2 - 1)\mathbf{Z}$, while A has size q. This example is due to Bose [7].

(5) Finally, for the last of the classical examples, we take a field E and an extension F/E of degree 3. Pick $\alpha \in \mathrm{F}$ such that $\mathrm{F} = \mathrm{K}(\alpha)$, so that $(1, \alpha, \alpha^2)$ is a basis of F as an E vector space. Now let $\mathrm{G} = \mathrm{F}^\times/\mathrm{E}^\times$, and let $\mathrm{A} = ((\mathrm{E} + \alpha\mathrm{E}) - \{0\})/\mathrm{K}^\times$, which is a subset of G. This is the fifth Sidon set.

Indeed, note that any element of A has a unique representative in F which is either α or of the form $1 + \alpha a$ for some $a \in \mathrm{E}$. Suppose then that (a, b, c, d) are elements of E of E such that

$$(1 + \alpha a)(1 + \alpha b) = (1 + \alpha c)(1 + \alpha d)$$

in G. Since $(1, \alpha, \alpha^2)$ is a basis of E, this relation is once more equivalent to the system (2.4). This already shows that $\mathrm{A} - \{\alpha\mathrm{E}^\times\}$ is a Sidon set. It is then elementary to check that adding α preserves the Sidon property (for instance, no equation

$$\alpha(1 + b\alpha) = (1 + c\alpha)(1 + d\alpha)$$

holds, etc).

When E is finite of size $q = p^\nu$ with p prime, then there is a unique choice of F (up to isomorphism of fields); the group G has size $(q^3 - 1)/(q - 1) = q^2 + q + 1$ and is cyclic (since it is a quotient of a cyclic group), while A has size $(q^2 - 1)/(q - 1) = q + 1$. This example is due to Singer [76]; it is interesting to note that it is in fact chronologically the first of those we have described, although it might look more complicated than the previous ones.

These examples have clearly a similar flavor, but it's not obvious at first if one can describe them as special cases of a uniform construction. This is however possible, as shown by Eberhard and Manners. Abstractly, they construct Sidon sets out of plane projective geometry over a field. To state their result, we recall some basic notation from projective geometry.

For a field E and an integer $d \geqslant 1$, the d-dimensional projective space over E is the set $\mathbf{P}^d(\mathrm{E})$ of lines in the vector space E^{d+1}; it can (and will) be identified with the quotient set

$$(\mathrm{E}^{d+1} - \{0\})/\mathrm{E}^\times,$$

where the group $\mathrm{E}^\times$ acts on E^{d+1} by scalar multiplication, i.e., we have $t \cdot (x_1, \dots, x_{d+1}) = (tx_1, \dots, tx_{d+1})$.

The group $\mathrm{GL}_{d+1}(\mathrm{E})$ acts on E^{d+1} by the usual multiplication of matrices and vectors, and this respects the action by scalar multiplication, so we obtain an action on the projective space $\mathbf{P}^d(\mathrm{E})$ (by homogenenous linear change of variable). The group of scalar matrices (isomorphic to $\mathrm{E}^\times$) acts trivially, hence we obtain an action of the quotient group $\mathrm{PGL}_{d+1}(\mathrm{E}) = \mathrm{GL}_{d+1}(\mathrm{E})/\mathrm{E}^\times$ on the projective space.

A *line* in $\mathbf{P}^d(\mathrm{E})$ is defined to be the image by the canonical projection $(\mathrm{E}^{d+1} - \{0\}) \to \mathbf{P}^d(\mathrm{E})$ of the set of non-zero elements in a two-dimensional vector subspace contained

in E^{d+1}. The usual action of $\mathrm{GL}_{d+1}(\mathrm{E})$ on two-dimensional subspaces induces an action of $\mathrm{PGL}_{d+1}(\mathrm{E})$ on the set of lines in $\mathbf{P}^d(\mathrm{E})$. This is compatible with the action on points, in the sense that if $x \in \mathbf{P}^d(\mathrm{E})$ and if $\ell \subset \mathbf{P}^d(\mathrm{E})$ is a line, then for all $g \in \mathrm{PGL}_{d+1}(\mathrm{E})$, we have $x \in \ell$ if and only if $g \cdot x \in g \cdot \ell$.

THEOREM 2.3.10 (Eberhard–Manners). *Let* E *be a field. Let* G *be an abelian subgroup of the group* $\mathrm{PGL}_3(\mathrm{E})$. *Fix a point* $p \in \mathbf{P}^2(\mathrm{E})$ *and a line* $\ell \subset \mathbf{P}^2(\mathrm{E})$. *Assume that the stabilizers of* p *and* ℓ *in* G *are trivial, i.e., that*

$$\{g \in \mathrm{G} \mid g \cdot p = p\} = \{g \in \mathrm{G} \mid g \cdot \ell = \ell\} = \{1\}.$$

Then the set

$$\mathrm{A} = \{g \in \mathrm{G} \mid g \cdot p \in \ell\}$$

is a Sidon set in G.

PROOF. Let $(a, b, c, d) \in \mathrm{A}^4$ be elements such that $a \neq b$ and $ab^{-1} = cd^{-1}$. We need to prove that $a = c$. Let $r = ab^{-1} \cdot p$, which is also equal to $cd^{-1} \cdot p$.

We observe that the points p and r both belong to the intersection of the lines $b^{-1} \cdot \ell$ and $d^{-1} \cdot \ell$. Indeed, this holds by definition for p, and for r, we have $b \cdot r = b(ab^{-1}) \cdot p = a \cdot p \in \ell$ (using the commutativity of G) and similarly $d \cdot r = c \cdot p \in \ell$.

In the *projective* plane $\mathbf{P}^2(\mathrm{E})$, two lines are either equal or intersect in a single point (which may be "at infinity" from the perspective of affine geometry). Hence we have either $b^{-1} \cdot \ell = d^{-1} \cdot \ell$ or $p = r$. The first possibility implies that $b = d$ by the assumption on the stabilizer of ℓ is trivial, and the second is excluded from the assumption that the stabilizer of p is trivial, since $r = ab^{-1} \cdot p$ and $a \neq b$. □

Using the classification of (maximal) commutative subgroups of $\mathrm{PGL}_3(\mathrm{E})$, Eberhard and Manners recover the classical constructions of Sidon sets. We illustrate this in one case.

EXAMPLE 2.3.11. Let G be the subgroup of diagonal matrices, modulo scalars:

$$\mathrm{G} = \left\{ \begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix} \mid x, y, z \in \mathrm{E}^\times \right\} / \mathrm{E}^\times \subset \mathrm{PGL}_3(\mathrm{E}).$$

This is a commutative subgroup of $\mathrm{PGL}_3(\mathrm{E})$, which is isomorphic to $(\mathrm{E}^\times)^2$ by mapping a diagonal matrix as above to $(x/z, y/z)$.

Define $p \in \mathbf{P}^2(\mathrm{E})$ to be the class of $(1, 1, 1)$; since

$$\begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

we see that the elements of the stabilizer of p in G must satisfy $x = y = z$, i.e., are classes of scalar matrices, hence the stabilizer of p in G is trivial.

Now let $\ell \subset \mathbf{P}^2(\mathrm{E})$ be the line defined to be the image in $\mathbf{P}^2(\mathrm{E})$ of the non-zero elements of the plane $\mathrm{L} \subset \mathrm{E}^3$ defined by the equation

$$a + b = c$$

for $(a, b, c) \in \mathrm{E}^3$.

For $g \in \mathrm{G}$ with diagonal coefficients x, y, z, the image $g \cdot \ell$ is the line defined similarly from the equation

$$x^{-1}a + y^{-1}b - z^{-1}c = 0,$$

and this coincides with ℓ if and only if the corresponding planes are the same, which means that the matrix

$$\begin{pmatrix} 1 & 1 & -1 \\ x^{-1} & y^{-1} & -z^{-1} \end{pmatrix}$$

must have rank 1. This only occurs if $x = y = z$, so we deduce that the stabilizer of ℓ in G is also trivial.

We can therefore apply Theorem 2.3.10. The corresponding Sidon set is the set of (classes modulo $\mathrm{E}^\times$ of) matrices

$$g = \begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix}$$

such that

$$\begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathrm{L},$$

which translates to

$$x + y = z.$$

Using the isomorphism $\mathrm{G} \to \mathrm{E}^\times \times \mathrm{E}^\times$ described above, a matrix of this type maps to $(x/z, y/z)$ with $x/z + y/z = 1$. The image in $\mathrm{E}^\times \times \mathrm{E}^\times$ of A is then the set of elements of the form $(t, 1-t)$ with $t \in \mathrm{E}^\times - \{1\}$, which is precisely Example 2.3.9, (2) above.

Exercise 2.3.12. Let $\mathrm{Z} \simeq \mathrm{E}^\times$ denote the subgroup of diagonal matrices in $\mathrm{GL}_3(\mathrm{E})$.

(1) Show that the following are commutative subgroups of $\mathrm{PGL}_3(\mathrm{E})$, and that the Sidon set they define can recover two of the four other classical examples for suitable choices of points and lines:

$$\mathrm{G} = \left\{ \begin{pmatrix} x & y & 0 \\ 0 & x & 0 \\ 0 & 0 & 1 \end{pmatrix} \mid x \in \mathrm{E}^\times,\ y \in \mathrm{E} \right\} \mathrm{Z}/\mathrm{Z},$$

$$\mathrm{G} = \left\{ \begin{pmatrix} 1 & x & y \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix} \mid x, y \in \mathrm{E} \right\} \mathrm{Z}/\mathrm{Z}.$$

(2) Let E_3 (resp. E_2) be a cubic (resp. quadratic extension of) E. Show that there exists a subgroup G_3 (resp. G_2) of $\mathrm{GL}_3(\mathrm{E})$ isomorphic to $\mathrm{E}_3^\times$ (resp. to $\mathrm{E}_2^\times \times \mathrm{E}^\times$) so that the subgroup $\mathrm{G}_3\mathrm{Z}/\mathrm{Z}$ (resp. $\mathrm{G}_2\mathrm{Z}/\mathrm{Z}$) can be used to recover the last two constructions.

We now address the question of the density of Sidon sets in $\mathbf{Z}$. Of course, in order to count these, we either look at finite sets, or consider the intersection with finite intervals. The simplest bound follows by noting that if A is a Sidon set contained in an interval of length N, then the differences $b - a$ for $\{a, b\} \subset \mathrm{A}$ are pairwise distinct, and all lie in $[\mathrm{N}]$, so that

$$\frac{|\mathrm{A}|(|\mathrm{A}| - 1)}{2} \leqslant \mathrm{N},$$

which gives $|\mathrm{A}| \leqslant \sqrt{2\mathrm{N}} + o(\mathrm{N}^{1/2})$.

The simplest general results are due to Erdős and Turán. We summarize them as follows.

Theorem 2.3.13 (Erdős; Turán). *The following statements hold:*

(1) *Any interval of* $\mathbf{Z}$ *of length* $N \geqslant 1$ *contains a Sidon set of size* $\geqslant \frac{1}{2}|N|^{1/2}$.

(2) *There exists a Sidon set* $A \subset \mathbf{Z}$ *such that*

$$|A \cap [N]| \geqslant N^{1/3}$$

for all integers $N \geqslant 1$.

(3) *There exists a constant* $c > 0$ *sucht that, for any infinite Sidon set* A *of positive integers, we have*

$$\liminf_{N \to +\infty} \frac{(\log N)^{1/2}}{N^{1/2}} |A \cap [N]| \leqslant c, \tag{2.5}$$

or equivalently, such that

$$|A \cap [N]| \leqslant \frac{cN^{1/2}}{(\log N)^{1/2}}$$

for arbitrarily large values of N.

(4) *For any Sidon subset* $A \subset [N]$, *we have*

$$|A| \leqslant N^{1/2} + N^{1/4} + 6. \tag{2.6}$$

Among these, the contrast between the first and third results might be the most surprising: it shows that although the first N integers always contain pretty big Sidon sets, these are not "compatible" as N varies, and any attempt to construct an infinite Sidon set will lead to a set which has rather smaller density in a subsequence of intervals.

Proof of (1). This follows from the existence of dense Sidon subsets in suitable finite abelian groups (as given by any of the "classical" examples), with some added basic information on the distribution of primes.

To be precise, we first observe that if $N = p^2 - 1$ for some prime number p, then we have the Sidon set $\bar{A}$ of Bose in $\mathbf{Z}/(p^2-1)\mathbf{Z}$ (Example 2.3.9, (4)), that the set A of integers n in $[p^2]$ such that $n \bmod p \in \bar{A}$ is again a Sidon set (see Remark 2.3.2, (2)) of size $p \geqslant \frac{1}{2}\sqrt{p^2-1}$.

Now let $N \geqslant 4$ be arbitrary. By the so-called Bertrand Postulate (proved by Chebychev; see, e.g. [78, § 2.2] for a very accessible account), there exists a prime number p such that

$$\tfrac{1}{2}\sqrt{N} \leqslant p \leqslant \sqrt{N}$$

hence $\frac{1}{4}N \leqslant p^2 \leqslant N$. A Sidon set A of size p in $[p^2-1]$ is also a Sidon set in $[N]$, and we have

$$|A| = p \geqslant \tfrac{1}{2}\sqrt{N},$$

To conclude, we observe that any interval of length N in $\mathbf{Z}$ is of the form $j + [N]$, and contains the Sidon set $j + A$. □

Proof of (2). The idea here is to use a "greedy" construction: starting with $a_1 = 1$, we construct by induction a strictly increasing sequence of integers by defining inductively a_{k+1} to be the smallest positive integer, distinct from $a_1, \ldots, a_k$, such that the set $\{a_1, \ldots, a_{k+1}\}$ is a Sidon set.

The existence of this sequence is elementary: when $a_1, \ldots, a_k$ have been defined, any integer which is not among the integers $a + b - c$, with a, b, c taken among $\{a_1, \ldots, a_k\}$, is such that $\{a_1, \ldots, a_k, a\}$ is a Sidon set, and this a is not one of the a_i's with $i \leqslant k$ (since $a_i = a_i + a_i - a_i$). Thus the set of possible integers a to chose from is infinite, and we take the smallest possible to get a_{k+1}.

Now observe more precisely that, given k, there are $\leqslant k^3$ integers of the form $a+b-c$ with a, b, c in $\{a_1, \dots, a_k\}$, and thus the smallest positive integer which is *not* of this form is $\leqslant k^3+1$. Thus $a_{k+1} \leqslant k^3+1$, and it follows that the interval $[\mathrm{N}]$ always contains at least $\mathrm{N}^{1/3}$ elements of this sequence. □

Remark 2.3.14. The sequence (a_k) used above is known as the *Mian–Chowla sequence*. It appears as sequence A005282 in the *Online Encyclopedia of Integer Sequences* (oeis.org), and its first few terms are:

$$1, 2, 4, 8, 13, 21, 31, 45, 66, 81, 97, 123, 148, 182, 204, 252, 290, 361, 401, 475, 565, 593,$$
$$662, 775, 822, 916, 970, 1016, 1159, 1312, 1395, 1523, 1572, 1821, 1896, 2029, 2254, 2379,$$
$$2510, 2780, 2925, 3155, 3354, 3591, 3797, 3998, 4297, 4433, 4779, 4851.$$

We now prove the statements (3) and (4). They are based on similar ideas, so we begin with (3) and indicate the variant used in the proof of (4) later on. The key property of Sidon sets used in (3) is the following, which was in fact already mentioned.

Lemma 2.3.15. *Let* A *be a Sidon set of positive integers and let* $\mathrm{M} \geqslant 1$ *be an integer. The number of subsets* $\{a, b\}$ *of elements of* A *of size* 2 *with* $|b-a| \leqslant \mathrm{M}$ *is at most* $\mathrm{M}-1$.

Proof. By assumption, we have a well-defined map

$$\{a, b\} \mapsto \max(a, b) - \min(a, b) = |b-a|$$

from the set of pairs of elements in $\mathrm{A} \cap \mathrm{I}$ to $[\mathrm{M}-1]$. The fact that A is a Sidon set implies that this map is injective, and hence the result follows. □

We will prove that if A is a set of positive integers which satisfies the condition of Lemma 2.3.15, then it also satisfies the conclusion of the third part of Theorem 2.3.13. In fact, we will show the following inequality, which turns out to imply the statement.

Proposition 2.3.16. *Let* A *be a Sidon set of positive integers. For any integer* $\mathrm{N} \geqslant 1$, *define the intervals*

$$\mathrm{I}_j = (j-1)\mathrm{N} + [\mathrm{N}] = \{(j-1)\mathrm{N}+1, \dots, j\mathrm{N}\}$$

for $1 \leqslant j \leqslant \mathrm{N}$. *For any family* $(\lambda_j)_{1 \leqslant j \leqslant \mathrm{N}}$ *of complex numbers, the inequality*

$$\Big|\sum_{j=1}^{\mathrm{N}} \lambda_j |\mathrm{A} \cap \mathrm{I}_j|\Big| \leqslant 5^{1/2}\mathrm{N}^{1/2}\Big(\sum_{j=1}^{\mathrm{N}} |\lambda_j|^2\Big)^{1/2}$$

holds.

Proof. There is no choice here but to use the Cauchy–Schwarz inequality: we have

$$\Big|\sum_{j=1}^{\mathrm{N}} \lambda_j |\mathrm{A} \cap \mathrm{I}_j|\Big| \leqslant \Big(\sum_{j=1}^{\mathrm{N}} |\mathrm{A} \cap \mathrm{I}_j|^2\Big) \sum_{j=1}^{\mathrm{N}} |\lambda_j|^2 \tag{2.7}$$

for any family (λ_j). So the question is to bound the sum

$$\sum_{j=1}^{\mathrm{N}} |\mathrm{A} \cap \mathrm{I}_j|^2$$

from above (in fact, the proposition is equivalent to proving that this sum is $\leqslant 5\mathrm{N}$). We can do this because the problem is closely related to counting pairs of elements from

$\mathrm{A} \cap \mathrm{I}_j$. Precisely, noting the inequality $x^2 \leqslant 1 + 2x(x-1)$ for $x \in \mathbf{R}$, we obtain the estimate

$$\sum_{j=1}^{\mathrm{N}} |\mathrm{A} \cap \mathrm{I}_j|^2 \leqslant \mathrm{N} + 4 \sum_{j=1}^{\mathrm{N}} \binom{|\mathrm{A} \cap \mathrm{I}_j|}{2}.$$

But $\binom{|\mathrm{A}\cap \mathrm{I}_j|}{2}$ is the number of pairs $\{a, b\}$ in $\mathrm{A} \cap \mathrm{I}_j$; any such pair, for any j, satisfies $1 \leqslant |b-a| \leqslant \mathrm{N}-1$, so the *total number*, allowing j to range over all integers, is at most N, by Lemma 2.3.15. This means that

$$\sum_{j=1}^{\mathrm{N}} \binom{|\mathrm{A} \cap \mathrm{I}_j|}{2} \leqslant \mathrm{N},$$

and hence

$$\sum_{j=1}^{\mathrm{N}} |\mathrm{A} \cap \mathrm{I}_j|^2 < 5\mathrm{N},$$

which gives the result when combined with (2.7). □

We apply the general inequality with $\lambda_j = j^{-1/2}$ (we encourage the reader to try other choices also to get a feeling for this inequality; for instance, check that $\lambda_j = 1$ only recovers the "trivial" bound $|\mathrm{A} \cap [\mathrm{N}^2]| \ll \mathrm{N}$). We thus obtain

$$\sum_{j=1}^{\mathrm{N}} \frac{|\mathrm{A} \cap \mathrm{I}_j|}{\sqrt{j}} \ll \sqrt{\mathrm{N} \log \mathrm{N}},$$

for all integers $\mathrm{N} \geqslant 2$. We then transform the sum on the right-hand side using summation by parts (see Lemma A.2.1 in the Appendix if this is not familiar). Noting that

$$\sum_{j=1}^{k} |\mathrm{A} \cap \mathrm{I}_j| = |\mathrm{A} \cap [k\mathrm{N}]|$$

for any integer $k \geqslant 1$, we get

$$\sum_{j=1}^{\mathrm{N}} \frac{|\mathrm{A} \cap \mathrm{I}_j|}{\sqrt{j}} = \sum_{j=1}^{\mathrm{N}-1} |\mathrm{A} \cap [j\mathrm{N}]| \Big(\frac{1}{\sqrt{j}} - \frac{1}{\sqrt{j+1}} \Big) + \frac{|\mathrm{A} \cap [\mathrm{N}^2]|}{\sqrt{\mathrm{N}}}.$$

Since we want a lower bound, we can ignore the last term, and using elementary estimates, we get

$$\sum_{j=1}^{\mathrm{N}-1} \frac{|\mathrm{A} \cap \mathrm{I}_j|}{\sqrt{j}} \gg \sum_{j=1}^{\mathrm{N}-1} \frac{|\mathrm{A} \cap [j\mathrm{N}]|}{j^{3/2}}.$$

Combined with the previous result, this gives the estimate

$$\sum_{j=1}^{\mathrm{N}} \frac{|\mathrm{A} \cap [j\mathrm{N}]|}{j^{3/2}} \ll \sqrt{\mathrm{N}} (\log \mathrm{N})^{1/2}$$

for all $\mathrm{N} \geqslant 2$.

The $\sqrt{\log \mathrm{N}}$ on the right hand side shows that it is not possible that $|\mathrm{A} \cap [j\mathrm{N}]| \gg \sqrt{\mathrm{N}}$ for all j, since the left-hand side would then be of size $\sqrt{\mathrm{N}} \log \mathrm{N}$. More precisely, to deduce (2.5), we may define

$$e(\mathrm{N}) = \inf_{n \geqslant \mathrm{N}} \frac{|\mathrm{A} \cap [n]|}{(\log n/n)^{1/2}}$$

for $N \geqslant 2$. Then the left-hand side of the last inequality is at least

$$\geqslant e(N) \sum_{j=1}^{N} \sqrt{\frac{jN}{\log jN}} \frac{1}{j^{3/2}} \gg e(N)\sqrt{N(\log N)}$$

and hence we deduce that the function $e(N)$ is bounded for $N \geqslant 2$. This amounts to proving (2.5).

We now explain how to prove the bound for Sidon sets contained in $[N]$. We pick a parameter $M \geqslant 1$, to be determined later, with $M < N$. We consider the intervals $J_j = j + [M]$ for $j \in \mathbf{Z}$. As in Proposition 2.3.16, we use Cauchy's inequality to obtain the bound

$$\Big|\sum_{j\in\mathbf{Z}} \lambda_j |A \cap J_j|\Big|^2 \leqslant \Big(\sum_{j\in\mathbf{Z}} |A \cap J_j|^2\Big) \Big(\sum_{j\in\mathbf{Z}}^{N} |\lambda_j|^2\Big) \tag{2.8}$$

for any sequence (λ_j), say with finite support so that the last sum is finite. We now simply take $\lambda_j = 1$ if $[N] \cap J_j \neq \emptyset$, and $\lambda_j = 0$ otherwise. Then it follows that

$$\sum_{j\in\mathbf{Z}}^{N} |\lambda_j|^2 = |\{j \in \mathbf{Z} \mid [N] \cap (j + [M]) \neq \emptyset\}| = N + M - 1.$$

On the other hand, the sum of $|A \cap J_j|$ is easily determined:

$$\sum_{j\in\mathbf{Z}} |A \cap J_j| = \sum_{j\in\mathbf{Z}} \sum_{a\in A\cap J_j} 1 = \sum_{a\in A} \sum_{\substack{j\in\mathbf{Z} \\ a\in j+[M]}} 1 = M|A|.$$

Finally, the sum of $|A \cap J_j|^2$ is handled using the Sidon property: first, note that

$$\sum_{j\in\mathbf{Z}} |A \cap J_j|^2 = \sum_{a,b\in A} \sum_{\substack{j\in\mathbf{Z} \\ \{a,b\}\subset J_j}} 1.$$

For $a = b$, we have the same sum equal to $M|A|$ as before. On the other hand, when $a \neq b$, the Sidon property implies that $\{a, b\}$ is determined by the value of $\delta = |b-a|$, which is an integer in $[M - 1]$. Moreover, for any given δ, there are $M - \delta$ intervals J_j in which the pair $\{a, b\}$ is contained (namely, assuming $a < b$, those of the form $j + [M]$ with $b - M \leqslant j < a$). Thus

$$\sum_{a,b\in A} \sum_{\substack{j\in\mathbf{Z} \\ \{a,b\}\subset J_j}} 1 \leqslant 2 \sum_{\delta=1}^{M-1} (M - \delta) = M(M - 1),$$

and altogether

$$\sum_{j\in\mathbf{Z}} |A \cap J_j|^2 \leqslant M|A| + M(M - 1) = M(|A| + M - 1).$$

Using (2.8) and dividing by M, Cauchy's inequality therefore gives

$$M|A|^2 \leqslant (N + M - 1)(M + |A| - 1). \tag{2.9}$$

The final bound results from some optimization of the parameter M. For instance, note that we can assume that $N \geqslant 2$. If $|A| < N^{1/2} + N^{1/4}$, then we are done. Otherwise,

let $M = \lceil N^{3/4} \rceil$; we then get

$$\frac{N + M - 1}{|A|} \leqslant \frac{N + N^{3/4}}{N^{1/2} + N^{1/4}} = N^{1/2},$$

and the inequality (2.9) implies

$$M|A| \leqslant N^{1/2}(M + |A| - 1),$$

hence

$$|A| \leqslant \frac{N^{1/2}(M-1)}{M - N^{1/2}} \leqslant \frac{N^{3/4}}{N^{1/4} - 1} = N^{1/2} + N^{1/4} + 1 + \frac{1}{N^{1/4} - 1},$$

which concludes the proof since the last term is $\leqslant 1/(2^{1/4} - 1) \leqslant 6$. (With more care, one can show that the last term can in fact be omitted.)

Remark 2.3.17. The best known infinite Sidon sets in positive integers satisfy

$$|A \cap [N]| \gg N^{\sqrt{2}-1+\varepsilon}$$

for all N, where $\varepsilon > 0$ is a fixed arbitrarily small number. The first examples were constructed by Ruzsa [72], using a spectacular construction starting with the fact that the numbers $\log p$, for p prime, form a Sidon set in $\mathbf{R}$. A different construction of Cilleruelo [18] gives the same exponent.

On the other hand, the bound (2.6) has been improved recently by Balogh, Füredi and Roy [3] to

$$|A| \leqslant \sqrt{N} + 0.998N^{1/4},$$

for N large enough.

Exercise 2.3.18. Show that if $A_1 \subset G_1$ and $A_2 \subset G_2$ are Sidon sets with $|A_i| \geqslant 2$, then $A_1 \times A_2$ is not a Sidon set in $G_1 \times G_2$.

Exercise 2.3.19. Let G be a finite abelian group. Let $\alpha \in G$ be a fixed element. A subset $A \subset G$ is called a *symmetric Sidon set* with center α if $A = \alpha - A$ (i.e., for any x in A, the element $\alpha - x$ is also in A) and if the equation

$$a + b = c + d$$

with $(a, b, c, d) \in A^4$ implies that $a \in \{c, d\}$ or $a + b = \alpha$.

(1) Let E be a field with characteristic different from 3. Prove that the set

$$A = \{(x, x^3) \mid x \in E\} \subset E \times E$$

is a symmetric Sidon set with center 0.

(2) Prove that if G is a finite group without 2-torsion (so that $2x = 0$ if and only if $x = 0$), then any symmetric Sidon set $A \subset G$ contains a subset $A' \subset A$ with $|A'| \geqslant (|A| - 1)/2$ such that A' is a Sidon set.

(3) Let G be a finite abelian group and $A \subset \widehat{G}$ a finite set of characters of G. If A is a symmetric Sidon set with center α, prove that

$$\frac{1}{|G|} \sum_{x \in G} \Big| \sum_{\chi \in A} \lambda_\chi \chi(x) \Big|^4 \leqslant 3 \Big(\sum_{\chi \in A} |\lambda_\chi|^2 \Big)^2.$$

Exercise 2.3.20. Let G be an abelian group, denoted additively. For a finite subset $A \subset G$, we denote by $E(A)$ the number of quadruples $(a, b, c, d) \in A^4$ such that $a + b = c + d$ (this quantity occurs again later, and will be called the *additive energy* of A).

(1) Show that A is a Sidon set in G if and only if $\mathrm{E}(\mathrm{A}) = 2|\mathrm{A}|^2 - |\mathrm{A}|$.

The goal of the remainder of the exercise shows that a finite set A may satisfy $\mathrm{E}(\mathrm{A}) = 2|\mathrm{A}|^2 + \mathrm{O}(|\mathrm{A}|)$, but not contain any Sidon subset of size $\sim |\mathrm{A}|$.

We take $\mathrm{G} = \mathbf{Z}$.

(1) Show that for all large integers N, there exists a Sidon set $\mathrm{A} \subset \{1, \dots, \mathrm{N}\} \cap 2\mathbf{Z}$ with $|\mathrm{A}| \to +\infty$ as $\mathrm{N} \to +\infty$.

(2) Consider a Sidon set $\mathrm{A} \subset \{1, \dots, \mathrm{N}\} \cap 2\mathbf{Z}$. Define

$$\mathrm{A}' = \mathrm{A} \cup \{a + \mathrm{N}2^{a+1} \mid a \in \mathrm{A}\} \cup \{a - \mathrm{N}2^{a+1} \mid a \in \mathrm{A}\} \subset \mathbf{Z}.$$

(3) Show that if $\mathrm{A}'' \subset \mathrm{A}'$ is a Sidon set, we have $|\mathrm{A}''| \leqslant \frac{2}{3}|\mathrm{A}'|$.

(4) Let

$$x_1 + x_2 = x_3 + x_4,$$

with

$$x_i = a_i + \varepsilon_i \mathrm{N}2^{a_i+1}, \quad a_i \in \mathrm{A}, \quad \varepsilon_i \in \{-1, 0, 1\},$$

Show that $a_1 + a_2 = a_3 + a_4$.

(5) Suppose that $a_1 = a_3$, hence $a_2 = a_4$. Show that

$$(\varepsilon_1 - \varepsilon_3)2^{a_1} = (\varepsilon_4 - \varepsilon_2)2^{a_2}.$$

(6) Deduce that $x_1 = x_3$ if $\varepsilon_1 = \varepsilon_3$ or $\varepsilon_2 = \varepsilon_4$.

(7) Suppose further that $\varepsilon_1 \neq \varepsilon_3$ and $\varepsilon_2 \neq \varepsilon_4$. Show that $a_1 = a_2 = a_3 = a_4$ and $\varepsilon_1 + \varepsilon_2 = \varepsilon_3 + \varepsilon_4$.

(8) Conclude that if $x_1 \notin \{x_3, x_4\}$, then (x_1, x_2, x_3, x_4) has one of the forms

$$(a + \mathrm{N}2^{a+1}, a - \mathrm{N}2^{a+1}, a, a), \qquad (a - \mathrm{N}2^{a+1}, a + \mathrm{N}2^{a+1}, a, a),$$
$$(a, a, a - \mathrm{N}2^{a+1}, a + \mathrm{N}2^{a+1}), \qquad (a, a, a + \mathrm{N}2^{a+1}, a - \mathrm{N}2^{a+1}),$$

for some $a \in \mathrm{A}$. (Hint: consider the various possibilities for $(\varepsilon_1, \dots, \varepsilon_4)$ for given $(\varepsilon_3, \varepsilon_4)$.)

(9) Deduce that

$$\mathrm{E}(\mathrm{A}'') = 2|\mathrm{A}''|^2 + \mathrm{O}(|\mathrm{A}''|).$$

Exercise 2.3.21. Let G be a finite abelian group, with additive notation.

(1) With the usual notation for representation functions, show that for any subsets A and B of G, we have

$$\sum_{x \in \mathrm{G}} r_{\mathrm{A},-\mathrm{B}}(x)^2 = \sum_{x \in \mathrm{G}} r_{\mathrm{A},-\mathrm{A}}(x) r_{\mathrm{B},-\mathrm{B}}(x).$$

(2) We assume for the remainder of the exercise that A is a Sidon set in G. Prove that

$$\sum_{x \in \mathrm{G}} r_{\mathrm{A},-\mathrm{A}}(x) r_{\mathrm{B},-\mathrm{B}}(x) \leqslant |\mathrm{A}||\mathrm{B}| + |\mathrm{B}|^2 - |\mathrm{B}|.$$

(3) Deduce from the previous questions that

$$\sum_{x \in \mathrm{G}} \Big(r_{\mathrm{A},-\mathrm{B}}(x) - \frac{|\mathrm{A}||\mathrm{B}|}{|\mathrm{G}|} \Big)^2 \leqslant |\mathrm{B}|(|\mathrm{A}| - 1) + \frac{|\mathrm{B}|^2(|\mathrm{G}| - |\mathrm{A}|^2}{|\mathrm{G}|}.$$

(4) Let also C be a subset of G and define

$$\mathrm{N} = |\{(b,c) \in \mathrm{B} \times \mathrm{C} \mid b + c \in \mathrm{A}\}|.$$

Show that

$$\mathrm{N} - \frac{|\mathrm{A}||\mathrm{B}||\mathrm{C}|}{|\mathrm{G}|} = \sum_{c \in \mathrm{C}} \Big(r_{\mathrm{A},-\mathrm{B}}(c) - \frac{|\mathrm{A}||\mathrm{B}|}{|\mathrm{G}|} \Big).$$

(5) Deduce that

$$\mathrm{N} - \frac{|\mathrm{A}||\mathrm{B}||\mathrm{C}|}{|\mathrm{G}|} \leqslant |\mathrm{C}|^{1/2} \Big(|\mathrm{B}|(|\mathrm{A}|-1) + \frac{|\mathrm{B}|^2(|\mathrm{G}|-|\mathrm{A}|^2}{|\mathrm{G}|} \Big)^{1/2}.$$

(6) Define $\delta \in \mathbf{Z}$ by $|\mathrm{A}| = |\mathrm{G}|^{1/2} - \delta$. Show that

$$\mathrm{N} = \frac{|\mathrm{A}||\mathrm{B}||\mathrm{C}|}{|\mathrm{G}|} + \theta(|\mathrm{B}||\mathrm{C}|\sqrt{|\mathrm{G}|})^{1/2},$$

where

$$\theta \leqslant 1 + \frac{|\mathrm{B}|}{|\mathrm{G}|}\max(0,\delta), \qquad \theta \leqslant 1 + \frac{|\mathrm{C}|}{|\mathrm{G}|}\max(0,\delta).$$

(7) Show that

$$|\mathrm{C}|\,|\mathrm{A} \cap \mathrm{B}| \leqslant |\{(x,y) \in -\mathrm{C} \times (\mathrm{B}+\mathrm{C}) \mid x + y \in \mathrm{A}\}|.$$

(8) Deduce that

$$|\mathrm{A} \cap \mathrm{B}| \leqslant \frac{|\mathrm{B}+\mathrm{C}||\mathrm{A}|}{|\mathrm{G}|} + \theta\Big(\frac{|\mathrm{B}+\mathrm{C}|}{|\mathrm{C}|}\Big)^{1/2}|\mathrm{G}|^{1/4},$$

with the same bounds as before for θ.

(The results of this exercise are due to Cilleruelo [17, Th. 2.1, Lemma 3.1].)

2.4. Approximate subgroups

The previous section was concerned with product sets of size very close to the maximal value. The other extreme is that of sets A and B for which $\mathrm{A} \cdot \mathrm{B}$ is closed to the minimal value. In general, this minimum is $\max(|\mathrm{A}|, |\mathrm{B}|)$, assuming that A and B are not empty. In constrast to Sidon sets, the extremal cases are here easily characterized, and it becomes possible (and, it turns out, important) to investigate sets which are close to achieving this – these will be called *approximate subgroups*.

We begin with looking at the smallest possible size of product sets, and obtain a relatively nice algebraic characterization.

Proposition 2.4.1. *Let* G *be a finite group, and let* A, B *be non-empty subsets of* G *such that* $|\mathrm{A} \cdot \mathrm{B}| = |\mathrm{B}|$. *Define*

$$\mathrm{H} = \{x \in \mathrm{G} \mid x\mathrm{B} = \mathrm{B}\},$$

the left *stabilizer of* B *in* G *for the action of* G *on itself by left multiplication. This is a subgroup of* G, *and for any* $a_0 \in \mathrm{A}$, *there exists a subset* $\mathrm{X} \subset \mathrm{G}$ *such that we have*

$$\mathrm{A} \subset a_0\mathrm{H}, \qquad \mathrm{B} = \mathrm{H}\mathrm{X}.$$

In particular, if $|\mathrm{A} \cdot \mathrm{A}| = |\mathrm{A}|$ *and* A *is neutral, the set* A *is equal to* H *and is a subgroup of* G.

In other words, this states that the product set $A \cdot B$ is minimal when one of the sets is the union of *left* cosets of a certain subgroup, and the other is contained in one of its *right* cosets.

PROOF. Since $|a_0^{-1}A \cdot B| = |A \cdot B| = |B|$, and $a_0^{-1}A$ is neutral, so that $a_0^{-1}A \cdot B$ contains B, the assumption implies that

$$a_0^{-1}A \cdot B = B.$$

For any $a \in A$, we have $a_0^{-1}aB \subset a_0^{-1}A \cdot B = B$ and $|a_0^{-1}aB| = |B|$, which means that $a_0^{-1}aB = B$. This implies that $a_0^{-1}A$ is contained in H, i.e., that $A \subset a_0H$.

The fact that B is a union of *right* cosets of its *left* stabilizer is a completely general fact: if $b \in B$, then by definition of H, we have $Hb \subset B$, so

$$B = \bigcup_{b \in B} Hb.$$

Finally, in the case where $A = B$ and A is neutral, the fact that $|A| = |B|$ implies that A is a full coset of H, which must be equal to H since $1 \in A$. □

This suggests to look at sets with $|A \cdot A|$ close to A as being close to subgroups. It turns out however that this does not lead to a really successful theory, and this also has the defect of being restricted to finite sets. As in the case of Sidon sets, there is a suitable definition which does apply to the infinite case. Although it seems a bit stronger than looking at sizes of product sets, the first key result will be that, up to manageable changes of quantitative parameters, it is equivalent to such a quantitative condition, in the case of finite groups.

DEFINITION 2.4.2. Let G be an arbitrary group. Let $\alpha \geqslant 1$ be a real number. A subset $H \subset G$ is called an *α-approximate subgroup* if H is neutral and symmetric and if there exists a symmetric subset X of G with $|X| \leqslant \alpha$ such that

$$H \cdot H \subset X \cdot H.$$

REMARK 2.4.3. (1) If $\alpha = 1$, then we obtain $|H \cdot H| = |H|$, so H is an actual subgroup of G by Proposition 2.4.1.

(2) If G is finite, note that one can always take $G = X$, so that any subset which is neutral and symmetric is a $|G|$-approximate subgroup. In this context, this means that the notion is only of interest if α is rather smaller than $|G|$. Usually, this translates to the fact that we consider a sequence of finite groups G_n of increasing size and subsets A_n which are α-approximate subgroups for some α independent of n, or growing slowler than the size of A_n.

(3) Let H be an α-approximate subgroup and X a subset of size at most α satisfying the property above. Let h_1 and h_2 be elements of H. Since H is symmetric, there exists $x \in X$ and $h_3 \in H$ such that $h_2^{-1}h_1^{-1} = xh_3$; taking inverse, we obtain

$$h_1h_2 = h_3^{-1}x^{-1} \in H \cdot X$$

(but there is no reason in general to expect that $H \cdot X = X \cdot H$, although both sets contain $H \cdot H$).

The following consequence of the definition is immediate by induction:

LEMMA 2.4.4. *Let* G *be a group, and let* $H \subset G$ *be a finite α-approximate subgroup for some $\alpha \geqslant 1$. For any integer $n \geqslant 0$, we have*

$$|A^{(n)}| \leqslant \alpha^{n-1}|A|.$$

EXAMPLE 2.4.5. Let $G = \mathbf{Z}$ (with additive notation). For any integer $N \geqslant 1$, the subset

$$H = [-N; N]$$

is a 2-approximate subgroup of $\mathbf{Z}$. Indeed, we have

$$H + H = [-2N; 2N] = (-N + H) \cup (N + H),$$

so we can take $X = \{-N, N\}$ in the definition.

We now begin to study the relationship between the notion of an approximate subgroup and that of sizes of product sets. The first goal will be to prove the following quite striking result:

THEOREM 2.4.6. *For any finite group G, for any real number $\alpha \geqslant 1$ and for any neutral symmetric subset $A \subset G$ such that $|A^{(3)}| \leqslant \alpha|A|$, the set $H = A^{(3)}$ is a β-approximate subgroup, with $\beta \leqslant 2\alpha^5$.*

In this result, as well as in similar results later, the value $2\alpha^5$ is not of essential importance (and is unlikely to be sharp); it is however essential that it has "polynomial" dependency on α: a choice of β with (say) $\beta = 2^\alpha$ would usually be useless for applications.

In fact, we will also prove later another characterization of a similar "numerical" flavor which is also important in some applications, and is significantly more involved.

Theorem 2.4.6 is based essentially on ideas of Ruzsa.

DEFINITION 2.4.7 (Ruzsa distance). Let G be a group. For any two non-empty finite subsets A and B of G, we define the *Ruzsa distance* between A and B to be

$$d(A, B) = \log\Big(\frac{|A \cdot B^{-1}|}{\sqrt{|A||B|}}\Big).$$

Concretely, we often use $d(A, B)$ in the estimate

$$|A \cdot B^{-1}| = \sqrt{|A||B|}\exp(d(A.B)),$$

and in particular, for A symmetric, we get

$$|A^{(2)}| = |A|\exp(d(A, A)). \tag{2.10}$$

The Ruzsa distance is not quite a distance in the usual sense of the word, since for instance $d(A, A) = \log(|A \cdot A^{-1}|/|A|)$ is rarely equal to zero. However, it satisfies the other two properties of distances.

PROPOSITION 2.4.8. *The Ruzsa distance is symmetric and satisfies the triangle inequality. In other words, for any non-empty finite subsets A, B and C in G, we have*

$$d(A, B) = d(B, A)$$
$$d(A, B) \leqslant d(A, C) + d(C, B).$$

PROOF. Since $|A| = |A^{-1}|$ for any subset of G, the symmetry amounts to the equality

$$|A \cdot B^{-1}| = |B \cdot A^{-1}|$$

for A and B non-empty subsets of G, and follows from the fact that $x \mapsto x^{-1}$ is a bijection from $A \cdot B^{-1}$ to $B \cdot A^{-1}$.

The triangle inequality $d(A, C) \leqslant d(A, B) + d(B, C)$, on the other hand, translates to the inequality

$$|A \cdot C^{-1}| \leqslant \frac{|A \cdot B^{-1}||B \cdot C^{-1}|}{|B|},$$

which we prove by exhibiting an injective map

$$f\colon \mathrm{B}\times(\mathrm{A}\cdot\mathrm{C}^{-1})\to(\mathrm{A}\cdot\mathrm{B}^{-1})\times(\mathrm{B}\times\mathrm{C}^{-1}),$$

in more or less the way that seems the most obvious: given $x\in\mathrm{A}\cdot\mathrm{C}^{-1}$, we pick $a(x)\in\mathrm{A}$ and $c(x)\in\mathrm{C}$ such that $x=ab^{-1}$ (arbitrarily), and we define

$$f(b,x)=(a(x)b^{-1},bc(x)),$$

for $b\in\mathrm{B}$ and $x\in\mathrm{A}\cdot\mathrm{C}^{-1}$. To see that this map is injective, note that when composed with the (restriction of the) multiplication map, we obtain the map $(b,x)\mapsto a(x)c(x)=x$. This shows that knowing $f(b,x)=(y,z)$ allows us to recover uniquely $x=yz$, and then we recover b uniquely by $b=zc(x)^{-1}$. □

Remark 2.4.9. The relation (2.10) shows that if A is symmetric, then $d(\mathrm{A},\mathrm{A})=0$ if and only if $|\mathrm{A}^{(2)}|=|\mathrm{A}|$. If A is also neutral, then this is equivalent to saying that A is a subgroup of G. Indeed, since $\mathrm{A}\subset\mathrm{A}^{(2)}$ in that case, the fact that $|\mathrm{A}|=|\mathrm{A}^{(2)}|$ implies $\mathrm{A}=\mathrm{A}^{(2)}$, i.e., that A is stable by product.

This simple property has remarkable consequences, which show that the sizes of multiple product sets cannot grow in arbitrary fashion. The simplest version is the following:

Proposition 2.4.10 (Ruzsa's Lemma). *Let* G *be a group. Let* $\mathrm{A}\subset\mathrm{G}$ *be a non-empty finite subset which is neutral and symmetric. Defining*

$$\alpha=\frac{|\mathrm{A}^{(3)}|}{|\mathrm{A}|},$$

we have the inequality

$$\frac{|\mathrm{A}^{(n)}|}{|\mathrm{A}|}\leqslant\alpha^{n-2}$$

for all integers $n\geqslant 3$.

Proof. We argue by induction on $n\geqslant 3$, with the case $n=3$ being valid by definition. Assume now that $n\geqslant 4$ and that the statement holds for $\mathrm{A}^{(n-1)}$. We have $\mathrm{A}^{(n+1)}=\mathrm{A}^{(n-1)}\cdot\mathrm{A}^{(2)}$, and hence

$$\frac{|\mathrm{A}^{(n+1)}|}{|\mathrm{A}|}=\frac{|\mathrm{A}^{(n-1)}\cdot\mathrm{A}^{(2)}|}{|\mathrm{A}|}=\frac{\sqrt{|\mathrm{A}^{(n-1)}||\mathrm{A}^{(2)}|}}{|\mathrm{A}|}\exp(d(\mathrm{A}^{(n-1)},\mathrm{A}^{(2)})).$$

We apply the Ruzsa triangle inequality to deduce

$$\begin{aligned}\frac{|\mathrm{A}^{(n+1)}|}{|\mathrm{A}|}&\leqslant\frac{\sqrt{|\mathrm{A}^{(n-1)}||\mathrm{A}^{(2)}|}}{|\mathrm{A}|}\exp(d(\mathrm{A}^{(n-1)},\mathrm{A})+d(\mathrm{A},\mathrm{A}^{(2)}))\\&=\frac{\sqrt{|\mathrm{A}^{(n-1)}||\mathrm{A}^{(2)}|}}{|\mathrm{A}|}\cdot\frac{|\mathrm{A}^{(n)}|}{\sqrt{|\mathrm{A}^{(n-1)}||\mathrm{A}|}}\cdot\frac{|\mathrm{A}^{(3)}|}{\sqrt{|\mathrm{A}||\mathrm{A}^{(2)}|}}=\frac{|\mathrm{A}^{(n)}|}{|\mathrm{A}|}\cdot\frac{|\mathrm{A}^{(3)}|}{|\mathrm{A}|},\end{aligned}$$

and this is $\leqslant\alpha^{n-2}\cdot\alpha$ by the induction hypothesis and the definition of α. This concludes the induction. □

Remark 2.4.11. There is no general analogue of Proposition 2.4.10 where $\alpha=|\mathrm{A}^{(3)}|/|\mathrm{A}|$ is replaced by $\beta=|\mathrm{A}^{(2)}|/|\mathrm{A}|$. Indeed, let $\mathrm{G}=\mathrm{SL}_2(\mathbf{F}_p)$ for a prime p (two by two matrices with coefficients in $\mathbf{F}_p$ with determinant 1). This is a group of size $p(p-1)(p+1)$, which is about p^3 for large p.

To construct a set for which β is small, we take a subgroup (which is stable by product), and add a few extra elements (to "break" the stability). For instance, let

$$\mathrm{B} = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \mathrm{G} \right\}, \qquad w = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and let

$$\mathrm{A} = \mathrm{B} \cup \{w, w^{-1}\}.$$

The set A is neutral and symmetric in G, and $|\mathrm{A}| = |\mathrm{B}| + 2 = p(p-1) + 2$. Since B is a subgroup of G and $w^2 = -\mathrm{Id} \in \mathrm{B}$, the product set $\mathrm{B} \cdot \mathrm{B}$ is the union of the cosets B, $w\mathrm{B}$, $w^{-1}\mathrm{B}$, $\mathrm{B}w$ and $\mathrm{B}w^{-1}$, which implies that $\beta \leqslant 5$ (and, in fact, $w\mathrm{B} = w^{-1}\mathrm{B}$ and $\mathrm{B}w = \mathrm{B}w^{-1}$, since $w^{-1} = -w$ and $-\mathrm{Id} \in \mathrm{B}$, so $\beta \leqslant 3$). On the other hand, the computation

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} w \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix} = \begin{pmatrix} -ba' & -bb' + ac' \\ -da' & -db' \end{pmatrix}$$

implies easily that $\mathrm{B}w\mathrm{B} \subset \mathrm{A}^{(3)}$ contains all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in G with $c \neq 0$. The number of these is $|\mathrm{G}| - |\mathrm{B}| = p(p-1)(p-2)$, so we have

$$\alpha = \frac{|\mathrm{A}^{(3)}|}{|\mathrm{A}|} \geqslant \frac{|\mathrm{B}w\mathrm{B}|}{|\mathrm{A}|} = \frac{p(p-1)(p-2)}{p(p-1)+2}$$

which is of size approximately p. In particular, there is no fixed constant $\mathrm{C} \geqslant 0$ such that $\beta \leqslant \alpha^{\mathrm{C}}$ for all primes p.

It is clear that a key feature of this computation has been the non-commutativity of G, which made $\mathrm{B}w\mathrm{B}$ much bigger than $\mathrm{B}w$ or $w\mathrm{B}$. This may suggest the possibility of having a statement like Ruzsa's lemma using $|\mathrm{A}^{(2)}|/|\mathrm{A}|$ if G is assumed to be commutative, and this is indeed possible. This was first established by Plünnecke [65], and we present one version of such a result (due to Petridis [64]) in Theorem 2.4.13 below.

Before proving Theorem 2.4.6, we need one last lemma, again due to Ruzsa. It moves from purely numerical constraints to more structural ones.

Lemma 2.4.12 (Ruzsa's covering lemma). *Let* G *be a group, let* A *and* B *be non-empty finite subsets of* G*, and let* $\alpha \geqslant 1$ *be such that* $|\mathrm{A} \cdot \mathrm{B}| \leqslant \alpha|\mathrm{A}|$*. There exists a subset* $\mathrm{X} \subset \mathrm{B} \subset \mathrm{G}$ *such that*

$$|\mathrm{X}| \leqslant \alpha, \qquad \mathrm{B} \subset \mathrm{A}^{-1} \cdot \mathrm{A} \cdot \mathrm{X}.$$

Proof. We consider in B a set X of elements x such that the subsets $\mathrm{A}x$ are *disjoint* for $x \in \mathrm{X}$, and which is maximal for inclusion (such a set exists, because any singleton has the disjointness property, and we can then take a set having this property with the largest size since B is finite). Since, by construction, we have

$$|\mathrm{A}||\mathrm{X}| = |\mathrm{A} \cdot \mathrm{X}| \leqslant |\mathrm{A} \cdot \mathrm{B}| \leqslant \alpha|\mathrm{A}|,$$

we must have $|\mathrm{X}| \leqslant \alpha$. Furthermore, for $b \in \mathrm{B}$, we note that $\mathrm{A}b$ and $\mathrm{A}\mathrm{X}$ cannot be disjoint: either $b \in \mathrm{X}$ and then $\mathrm{A}b \subset \mathrm{A}\mathrm{X}$, or otherwise the set $\mathrm{X} \cup \{b\}$ would contradict the maximality of X; thus there exist $(a_1, a_2, x) \in \mathrm{A}^2 \times \mathrm{X}$ such that $a_1 b = a_2 x$, and thence $b = a_1^{-1} a_2 x \in \mathrm{A}^{-1} \cdot \mathrm{A} \cdot \mathrm{X}$. □

We now prove Theorem 2.4.6.

PROOF. Let $H = A^{(3)}$. Since A is neutral and symmetric, Ruzsa's Lemma gives $|A \cdot H^{(2)}| = |A^{(7)}| \leqslant \alpha^5|A|$. Ruzsa's Covering Lemma applied to A and $H^{(2)}$ gives a subset X_0 of $H^{(2)}$, of size $\leqslant \alpha^5$, with the property that

$$H^{(2)} \subset A^{(2)} \cdot X_0.$$

A fortiori, the set $H^{(2)}$ is contained in $A^{(3)} \cdot X_0$, which is equal to $H \cdot X_0$. This is almost the conclusion we want: the only issue is that X_0 might not be symmetric. However, if we put $X = X_0 \cup X_0^{-1}$, then have still $H^{(2)} \subset H \cdot X$, and X is symmetric of size $\leqslant 2\alpha^5$. □

We conclude with the result of Plünnecke which, in particular, improves Proposition 2.4.10 for abelian groups.

THEOREM 2.4.13 (Plünnecke). *Let* G *be an abelian group, with additive notation. Let* A *and* B *be non-empty finite subsets of* G*, and let* $\alpha \geqslant 1$ *be such that*

$$|A + B| \leqslant \alpha|A|.$$

For any non-negative integers n *and* m*, we have*

$$|mB - nB| \leqslant \alpha^{m+n}|A|.$$

In particular, we get $|nA - mA| \leqslant \alpha^{m+n}|A|$ for any $n \geqslant 1$ if $|2A| \leqslant \alpha|A|$, which strenghtens Ruzsa's Lemma for commutative groups (and does not require A to be symmetric).

PROOF. We give the proof of Petridis [64].

The key claim is the following: if we denote by β the minimum of $|A' + B|/|A'|$ for $A' \subset A$ not empty, and by M a subset of A such that $|M + B| = \beta|M|$, then for any non-empty subset $C \subset G$, we have

$$|M + B + C| \leqslant \beta|M + C|. \tag{2.11}$$

Assuming this, we first deduce by induction that

$$|M + nB| \leqslant \beta^n|M|$$

for any integer $n \geqslant 0$. Pick then non-negative integers m and n; we have

$$|mB - nB| = \sqrt{|mB||nB|}\exp(d(mB, nB)), \tag{2.12}$$

and Ruzsa's triangle inequality gives

$$d(mB, nB) \leqslant d(mB, -M) + d(-M, nB).$$

Since

$$d(mB, -M) = \log\Big(\frac{|mB + M|}{\sqrt{|mB||M|}}\Big), \quad d(-M, nB) = \log\Big(\frac{|nB + M|}{\sqrt{|nB||M|}}\Big),$$

we get

$$\exp(d(mB, nB)) \leqslant \frac{|mB + M||nB + M|}{|M|\sqrt{|mB||nB|}} \leqslant \frac{\beta^{m+n}|M|}{\sqrt{|mB||nB|}}$$

and this, combined with (2.12), gives

$$|mB - nB| \leqslant \beta^{m+n}|M| \leqslant \alpha^{m+n}|A|,$$

since $\beta \leqslant \alpha$ by definition and $M \subset A$; this concludes the proof.

We now prove the key claim. This is done by induction on $|C|$. If C has a single element, then $|M + B + C| = |M + B| = \beta|M| = \beta|M + C|$. We now assume that $C = C' \cup \{x\}$ with $x \notin C'$, and that the statement holds for C'.

We can express

$$\mathrm{M}+\mathrm{B}+\mathrm{C}=(\mathrm{M}+\mathrm{B}+\mathrm{C}')\cup((\mathrm{M}+\mathrm{B}+x)-(\mathrm{M}_0+\mathrm{B}+x)),$$

with

$$\mathrm{M}_0=\{a\in\mathrm{M}\mid a+\mathrm{B}+x\subset\mathrm{M}+\mathrm{B}+\mathrm{C}'\}.$$

The key observation is then that $|\mathrm{M}+\mathrm{M}_0|\geqslant\beta|\mathrm{M}_0|$ by definition of M. Hence, using induction for C' and the fact that $\mathrm{M}_0+\mathrm{B}+x\subset\mathrm{M}+\mathrm{B}+x$, we get

$$|\mathrm{M}+\mathrm{B}+\mathrm{C}|\leqslant|\mathrm{M}+\mathrm{B}+\mathrm{C}'|+|\mathrm{M}+\mathrm{B}|-|\mathrm{M}_0+\mathrm{B}|\leqslant\beta(|\mathrm{M}+\mathrm{C}'|+|\mathrm{M}|-|\mathrm{M}_0|). \tag{2.13}$$

Furthermore, we have similarly

$$\mathrm{M}+\mathrm{C}=(\mathrm{M}+\mathrm{C}')\cup((\mathrm{M}+x)-(\mathrm{M}_1+x))$$

with $\mathrm{M}_1=\{a\in\mathrm{M}\mid a+x\in\mathrm{M}+\mathrm{C}'\}$, so that $|\mathrm{M}+\mathrm{C}|=|\mathrm{M}+\mathrm{C}'|+|\mathrm{M}|-|\mathrm{M}_1|$. By definition, $\mathrm{M}_1\subset\mathrm{M}$, hence

$$|\mathrm{M}+\mathrm{C}|=|\mathrm{M}+\mathrm{C}'|+|\mathrm{M}|-|\mathrm{M}_1|\geqslant|\mathrm{M}+\mathrm{C}'|+|\mathrm{M}|-|\mathrm{M}_0|,$$

and inputing this in (2.13) leads to the inequality

$$|\mathrm{M}+\mathrm{B}+\mathrm{C}|\leqslant\beta|\mathrm{M}+\mathrm{C}|,$$

which establishes (2.11) for C. □

For later purposes, we reformulate some of the last results using a useful notation.

Definition 2.4.14 (Approximate inclusion). Let G be a group. For subsets A and B of G, and $\alpha\geqslant 1$, we say that A is *α-contained* in B, denoted $\mathrm{A}\sqsubset_\alpha\mathrm{B}$, if there exists a subset $\mathrm{Y}\subset\mathrm{G}$ with $|\mathrm{Y}|\leqslant\alpha$ such that $\mathrm{A}\subset\mathrm{X}\cdot\mathrm{B}$.

Example 2.4.15. (1) In this language, an α-approximate subgroup of G is therefore (essentially) a neutral symmetric set H such that $\mathrm{H}\cdot\mathrm{H}$ is α-contained in H.
(2) A is 1-contained in B if and only if A is contained in a (left) translate of B.

Remark 2.4.16. The following properties are formal consequences of the definition, and of the upper-bound $|\mathrm{Y}\cdot\mathrm{Z}|\leqslant|\mathrm{Y}||\mathrm{Z}|$ for subsets of G:

(1) If $\mathrm{A}\sqsubset_\alpha\mathrm{B}$ and $\mathrm{B}\sqsubset_\beta\mathrm{C}$, then $\mathrm{A}\sqsubset_{\alpha\beta}\mathrm{C}$.

(2) If G is abelian and if $\mathrm{A}\sqsubset_\alpha\mathrm{B}$ and $\mathrm{A}'\sqsubset_\beta\mathrm{B}'$, then $\mathrm{A}\cdot\mathrm{A}'\sqsubset_{\alpha\beta}\mathrm{B}\cdot\mathrm{B}'$.

Furthermore, Ruzsa's Covering Lemma means that for any non-empty finite subsets A and B of a group G, the condition $|\mathrm{A}\cdot\mathrm{B}|\leqslant\alpha|\mathrm{A}|$ implies that $\mathrm{B}\sqsubset_\alpha\mathrm{A}^{-1}\cdot\mathrm{A}$. Similarly $|\mathrm{A}\cdot\mathrm{B}^{-1}|\leqslant\alpha|\mathrm{A}|$ implies that $\mathrm{B}\sqsubset_\alpha\mathrm{A}^{-1}\cdot\mathrm{A}$.

Plünnecke's Theorem admits the following corollary.

Proposition 2.4.17. *Let* G *be an abelian group. Let* $\mathrm{A}\subset\mathrm{G}$ *be a non-empty finite subset, and let* $\alpha\geqslant 1$ *be such that* $|\mathrm{A}-\mathrm{A}|\leqslant\alpha\mathrm{A}$. *For any non-negative integers* k *and* l, *we have*

$$m\mathrm{A}-n\mathrm{A}\sqsubset_{\alpha^{m+n+1}}\mathrm{A}-\mathrm{A}.$$

Proof. By Plünnecke's Theorem, we have

$$|\mathrm{A}+m\mathrm{A}-n\mathrm{A}|\leqslant\alpha^{m+n+1}|\mathrm{A}|,$$

so Ruzsa's Covering Lemma implies $m\mathrm{A}-n\mathrm{A}\sqsubset_{\alpha^{m+n+1}}\mathrm{A}-\mathrm{A}$. □

EXERCISE 2.4.18. (1) For any integer $N \geqslant 1$, find examples of sets A and B of positive integers such that $|A| = |B| = N$ and

$$\frac{|2A|}{A} \leqslant 2, \qquad \frac{|2B|}{|B|} \leqslant 2,$$

but

$$\frac{|2(A \cup B)|}{|A \cup B|} \geqslant \frac{N}{2}.$$

(2) For any integer $N \geqslant 1$, find examples of sets A and B of positive integers such that $|A| = |B| = N$ and

$$\frac{|2A|}{A} \leqslant 10, \qquad \frac{|2B|}{|B|} \leqslant 10,$$

but

$$\frac{|2(A \cap B)|}{|A \cap B|} \geqslant \frac{N^{1/2}}{10}.$$

EXERCISE 2.4.19. Let A_1, A_2, A_3 be non-empty finite subsets of some group G. If $\alpha \geqslant 1$ is such that

$$|A_j \cap A_3| \geqslant \frac{|A_j|}{\alpha}, \qquad |A_j \cdot A_j| \leqslant \alpha |A_j|$$

for $1 \leqslant j \leqslant 3$, then show that

$$|A_1 \cdot A_2| \leqslant \alpha^6 |A_3|.$$

(Hint: use the Ruzsa triangle inequality suitably)

EXERCISE 2.4.20. Let G be a finite *abelian* group and A, B non-empty subsets of G. Let

$$r(x) = \sum_{\substack{(a,b) \in A \times B \\ a+b=x}} 1$$

be the representation function for $A + B$.

(1) Show that $r(x) = |A \cap (x - B)|$.

(2) Show that

$$E(A, B) = \sum_{x \in (A-A) \cap (B-B)} |A \cap (x + A)|\, |B \cap (x + B)|.$$

EXERCISE 2.4.21. Let G be a finite group and A, B non-empty subsets of G.

(1) Let $x_0 \in A \cdot B$. Prove that

$$|\{(a, b) \in A \times B \mid ab = x_0\}| \times |(B \cdot A)| \leqslant |B \cdot A^{-1}|\, |B^{-1} \cdot A|.$$

(Hint: construct an injective map from the left-hand set to the cartesian product $B \cdot A^{-1} \times B^{-1} \cdot A$.)

(2) Deduce that if $x \in A \cdot B$, then

$$|A \cap xB^{-1}| \leqslant \frac{|B \cdot A^{-1}||B^{-1} \cdot A|}{|B \cdot A|}.$$

(3) If G is abelian, deduce that

$$|A \cap xB^{-1}| \leqslant \frac{|B \cdot A^{-1}|^2}{|A \cdot B|}.$$

EXERCISE 2.4.22. Let G be a finite abelian group.

(1) If H_1 and H_2 are subgroups of G, then show that the Ruzsa distance $d(H_1, H_2)$ satisfies

$$d(H_1, H_2) = \log\Big(\frac{\sqrt{|H_1||H_2|}}{|H_1 \cap H_2}\Big).$$

(2) Show that

$$d(H_1, H_2) = d(H_1, H_1 + H_2) + d(H_1 + H_2, H_2) = d(H_1, H_1 \cap H_2) + d(H_1 \cap H_2, H_2).$$

EXERCISE 2.4.23. Let G be a finite abelian group and $A \subset G$ a non-empty subset such that

$$|2A - 2A| < 2|A|.$$

(1) Show that there exists $x_0 \in G$ such that

$$A - 2A \subset A - A + x_0.$$

(2) Deduce that $A - A$ is a subgroup of G.

2.5. Multiplicative energy

A "dual" perspective on approximate subgroups arises from the following observation: if $H \subset G$ is a finite subgroup of a group, then the equation

$$ab = cd$$

with $(a, b, c, d) \in H^4$ has the largest possible numbers of solutions, namely $|H|^3$: for any given choice of (a, b, c), the value of d is uniquely determined and belongs to H because it is a subgroup. (This is a different way of saying that subgroups are "completely opposite" to Sidon sets.) This fact, once again, has a simple exact inverse property. To state it, we introduce formally the quantity that was just discussed for a subgroup.

DEFINITION 2.5.1 (Multiplicative energy). Let G be a group and let A, B be finite subsets of G. The *multiplicative energy* of (A, B) is the quantity

$$E(A, B) = |\{(a_1, a_2, b_1, b_2) \in A^2 \times B^2 \mid a_1 b_1 = a_2 b_2\}|.$$

If $A = B$, we call $E(A) = E(A, B)$ the multiplicative energy of A.

If A and B are not empty, then the *normalized multiplicative energy* is

$$e(A, B) = \frac{E(A, B)}{(|A|\,|B|)^{3/2}}.$$

REMARK 2.5.2. (1) Solutions of $a_1 b_1 = a_2 b_2$ are often called *multiplicative quadruples*, or *additive quadruples* when G is abelian.

(2) Part of the reason the energy is extremely useful is that, by counting solutions of equation, it takes into accoung multiplicity. This feature often means that a quantity has a nice analytic interpretation. Here, if we denote

$$r(x) = \sum_{\substack{(a,b)\in A\times B\\ ab=x}} 1$$

the representation function for $A \cdot B$, then we have

$$E(A, B) = \sum_{x\in G} r(x)^2. \tag{2.14}$$

Indeed, by expanding the square, we get

$$\sum_{x\in \mathrm{G}} r(x)^2 = \sum_{x\in \mathrm{G}} \Big(\sum_{ab=x} 1\Big)^2 = \sum_{(a_1,a_2,b_2,b_2)\in \mathrm{A}^2\times \mathrm{B}^2} \sum_{\substack{x\in \mathrm{G}\\ a_1b_1=x=a_2b_2}} 1$$
$$= \sum_{\substack{(a_1,a_2,b_2,b_2)\in \mathrm{A}^2\times \mathrm{B}^2\\ a_1b_1=a_2b_2}} 1 = \mathrm{E}(\mathrm{A},\mathrm{B}).$$

As an example of application, we note that

$$\sum_{x\in \mathrm{G}} r(x) \leqslant \sqrt{|\mathrm{A}\cdot \mathrm{B}|}\Big(\sum_{x\in \mathrm{G}} r(x)^2\Big)^{1/2}$$

by the Cauchy–Schwarz inequality, and therefore

$$\mathrm{E}(\mathrm{A},\mathrm{B}) \geqslant \frac{|\mathrm{A}|^2|\mathrm{B}|^2}{|\mathrm{A}\cdot \mathrm{B}|}. \tag{2.15}$$

If G is finite and commutative, the discrete Plancherel formula (see (A.9), noting that $r \geqslant 0$ so that $r^2 = |r|^2$) shows that we also have

$$\mathrm{E}(\mathrm{A},\mathrm{B}) = \sum_{\xi\in \widehat{\mathrm{G}}} |\widehat{r}(\xi)|^2,$$

where

$$\widehat{r}(\xi) = \frac{1}{\sqrt{|\mathrm{G}|}} \sum_{x\in \mathrm{G}} r(x)\overline{\xi(x)}$$

is the discrete Fourier transform of r.

Note in passing that the contribution of the trivial character to $\mathrm{E}(\mathrm{A},\mathrm{B})$ is $|\mathrm{A}|^2|\mathrm{B}|^2/|\mathrm{G}|$; this has a probabilistic interpretation: think that $(a_1,a_2,b_1,b_2)\in \mathrm{A}^2\times \mathrm{B}^2$ are picked at random, and that their product is "uniformly distributed", so that their is a chance about $1/|\mathrm{G}|$ that $a_2^{-1}a_1^{-1}b_1b_2 = 1$, which is the condition to have a multiplicative quadruple.

(3) Various identities hold for the multiplicative energy. For instance, if G is abelian, then the equation $ab = cd$ is equivalent to $ad^{-1} = cb^{-1}$, and therefore we get $\mathrm{E}(\mathrm{A},\mathrm{A}) = \mathrm{E}(\mathrm{A},\mathrm{A}^{-1})$.

Exercise 2.5.3. Let A be a non-empty finite subset of a group G. Prove that if $\alpha \geqslant 1$ is such that $|\mathrm{A}^{(2)}| \leqslant \alpha|\mathrm{A}|$, then

$$e(\mathrm{A},\mathrm{A}) \geqslant \frac{1}{\alpha}.$$

Before stating the inverse result for energy, we explain why the normalization of $e(\mathrm{A},\mathrm{B})$ is natural. The simplest upper-bound for $\mathrm{E}(\mathrm{A},\mathrm{B})$ in general is

$$\mathrm{E}(\mathrm{A},\mathrm{B}) \leqslant \min(|\mathrm{A}|^2|\mathrm{B}|, |\mathrm{A}||\mathrm{B}|^2),$$

since fixing three of the variables in the equation $a_1b_1 = a_2b_2$ determines uniquely the remaining one. We have

$$\min(|\mathrm{A}|^2|\mathrm{B}|, |\mathrm{A}||\mathrm{B}|^2) = |\mathrm{A}||\mathrm{B}|\min(|\mathrm{A}|,|\mathrm{B}|) \leqslant |\mathrm{A}||\mathrm{B}|\sqrt{|\mathrm{A}||\mathrm{B}|},$$

(where the last step amounts to saying that the minimum of two numbers is at most their geometric mean). This gives

$$\mathrm{E}(\mathrm{A},\mathrm{B}) \leqslant \sqrt{|\mathrm{A}|^3|\mathrm{B}|^3},$$

which means that the normalized energy $e(\mathrm{A},\mathrm{B})$ is always at most 1.

We can also deduce from this the following estimate: if $\alpha \geqslant 1$ is such that $e(A, B) \geqslant \alpha^{-1}$, then

$$\frac{1}{\alpha^2}|A| \leqslant |B| \leqslant \alpha^2|A|. \tag{2.16}$$

Indeed, assume for instance that $|A| \leqslant |B|$, the other case being similar. Then

$$\frac{|A||B|^{3/2}}{\alpha} \leqslant E(A, B) \leqslant \min(|A||B|^2, |A|^2|B|) = |A|^2|B|,$$

which gives

$$\frac{1}{\alpha^2}|A| \leqslant |A| \leqslant |B| \leqslant \alpha^2|A|.$$

We now state the inverse result; note that it is a rather "cleaner" statement than Proposition 2.4.1.

PROPOSITION 2.5.4. *Let* G *be a group and let* A, B *be non-empty finite subsets of* G. *We have* $e(A, B) \leqslant 1$, *and* $e(A, B) = 1$ *if and only if there exist a subgroup* $H \subset G$ *and* $(x, y) \in G^2$ *such that* $A = xH$ *and* $B = Hy$.

In fact, H *is the stabilizer of* A *for the action of* G *on itself by right multiplication, or equivalently the stabilizer of* B *for the action by left multiplication, and one can take* x *and* y *to be any element of* A *or* B, *respectively.*

In particular, this result says that $|A| = |B|$ if $e(A, B) = 1$ (and A, B are not empty).

PROOF. We have seen that $e(A, B) \leqslant 1$; moreover, checking how this was done, we see that equality can only occur if $\min(|A|, |B|) = \sqrt{|A||B|}$, which is only true (for non-empty sets) if $|A| = |B|$.

Now define H as first indicated, namely $H = \{x \in G \mid Ax = A\}$.

Since $|A| = |B|$, the condition $e(A, B) = 1$ can be written $E(A, B) = |A||B|^2$, and implies that for any triple $(a_1, b_1, b_2) \in A \times B^2$ the unique element $a_2 = a_1 b_1 b_2^{-1}$ in G such that $a_1 b_1 = a_2 b_2$ is in A. More concisely, this is equivalent to $A \cdot B \cdot B^{-1} \subset A$, and can also be translated as the inclusion

$$B \cdot B^{-1} \subset H.$$

It follows that $|A| = |B| \leqslant |B \cdot B^{-1}| \leqslant |H|$. However, taking any $a_0 \in A$, we have $a_0 H \subset A$, so that $|H| \leqslant |A|$ (which can also be deduced from the observation, used in the proof of Proposition 2.4.1, that A is a union of cosets of H). Hence $|A| = |H|$, and in fact $A = a_0 H$.

Writing dually $E(A, B) = |A|^2|B|$, we see also that for any $(a_1, a_2, b_1) \in A^2 \times B$, the element $a_2^{-1} a_1 b_1$ is in B. This means that $A^{-1} \cdot A \cdot B \subset B$; since $A^{-1} \cdot A = H^{-1} \cdot a_0^{-1} a_0 \cdot H = H$, it follows that H is contained in the stabilizer of B for the *left* multiplication action. From $|H| = |B|$, we conclude as before that B must be a left coset of H, i.e., that $B = Hb$ for any $b \in B$. □

The following result is a version of this proposition for approximate subgroups. As we will see, it is rather deeper than Ruzsa's Theorem (but will use it in the proof); the first version was proved by Balog and Szemerédi [2], and strong quantitative versions were then given by Gowers [41], in the course of his proof of Szemerédi's Theorem.

THEOREM 2.5.5 (Balog–Szemerédi; Gowers). *Let* G *be a finite group and let* A *and* B *be non-empty finite subsets of* G. *Let* $\alpha \geqslant 1$ *be a real number such that* $e(A, B) \geqslant \alpha^{-1}$. *There exists a real number*

$$\beta \leqslant 2^{3000} \alpha^{1000},$$

a β-approximate subgroup $H \subset G$ *and elements* x, y *in* G, *such that*

$$|H| \leqslant \beta|A|, \qquad |A| \leqslant \beta|A \cap xH|, \qquad |B| \leqslant \beta|B \cap Hy|.$$

The possible value of β is only indicated to give an explicit result. It can be improved (see for instance [59, Th. A.3.7] for one version), but in most applications, it is rather irrelevant which numbers appear – what matters is that $\beta \leqslant C\alpha^d$ for some *fixed* real numbers $C \geqslant 0$ and $d \geqslant 0$ (even any other dependency of β on α may be useful, as in the original work of Balog and Szemerédi). In fact, we will only prove the theorem in this form since it simplifies notation (and makes computational mistakes less likely).

Example 2.5.6. Suppose that $A = B$ and is neutral and symmetric. One can wonder if it would be possible to take simply $H = A$ in Theorem 2.5.5, or in other words, if A itself would have to be an approximate subgroup. This is not the case, and this fact explains to some extent why the result is rather subtle.

Here is a very simple kind of example to illustrate what can happen. Consider a finite abelian group G, a subgroup H of G and a Sidon set $S \subset G$. Assume that H and S have approximately the same size (in particular, they are of size $\ll |G|^{1/2}$), e.g. $\frac{1}{2}|S| \leqslant |H| \leqslant 2|S|$. Define $A = H \cup S$. Since H is a subgroup, we have $E(A, A) \geqslant |H|^3$, hence $e(A, A) \geqslant \frac{1}{3^3}|A|^3$ under our assumption on the size of H. On the other hand, $A + A$ contains $S + S$, which has size $\gg |S|^2 \gg |A|^2$, so that $|2A|/|A|$ is very large, and in particular is not bounded independently of the size of A.

On the other hand, the conclusion of Theorem 2.5.5 clearly holds in this case, with H itself a possible choice of approximate subgroup.

To apply the theorem, it is useful to have various criteria that imply lower bounds for the multiplicative energy. Here are some "deterministic" versions; in Section 3.5, we will see variants involving random variables.

Proposition 2.5.7. *Let* G *be a finite group and let* A *and* B *be non-empty finite subsets of* G.

(1) *For any subset* C *of* $A \times B$, *we have*

$$E(A, B) \geqslant \frac{|C|^2}{|A \times_C B|},$$

where

$$A \times_C B = \{x \in G \mid x = ab \text{ for some } (a, b) \in C\}.$$

(2) *Let* $C \subset G$ *and* $\alpha \geqslant 1$ *be such that* $r_{A\cdot B}(x) \geqslant \alpha^{-1}|A|$ *for all* $x \in C$. *We have*

$$E(A, B) \geqslant \alpha^{-2}|A|^2|C|.$$

Proof. For (1), we define the "relative" representation function $r^C_{A\cdot B} \colon G \to \mathbf{R}$ by

$$r^C_{A\cdot B}(x) = \sum_{\substack{(a,b)\in C \\ ab=x}} 1.$$

Since $r^C_{A\cdot B} \leqslant r_{A\cdot B}$ we have

$$E(A, B) = \sum_{x\in G} r_{A\cdot B}(x)^2 \geqslant \sum_{x\in G} r^C_{A\cdot B}(x)^2,$$

and noting that $r^C_{A\cdot B}(x) = 0$ unless $x \in A \times_C B$, the Cauchy–Schwarz inequality gives

$$E(A, B) \geqslant \frac{1}{|A \times_C B|}\Big(\sum_{x\in G} r^C_{A\cdot B}(x)\Big)^2.$$

Thus the desired lower-bound follows from the formula

$$\sum_{x\in G} r^C_{A\cdot B}(x) = \sum_{x\in G}\sum_{\substack{(a,b)\in C\\ ab=x}} 1 = \sum_{(a,b)\in C}\sum_{x=ab} 1 = |C|.$$

For (2), we just note that

$$E(A,B) = \sum_{x\in G} r_{A\cdot B}(x)^2 \geqslant \alpha^{-2}|A|^2|C|$$

by assumption. □

Before embarking on the proof of Theorem 2.5.5, which is much more involved than all those we have seen before, we will present a sketch of an application of this theorem, which illustrates *how* approximate subgroups can arise in practice in apparently unrelated problems.

2.6. Sketch of application

We will discuss parts of the ideas of Bourgain and Gamburd, which they used in [9] to solve an important problem at the boundary of graph theory and geometric group theory. Readers wishing to look immediately at the proof of Theorem 2.5.5 may safely skip this section.

THEOREM 2.6.1 (Bourgain–Gamburd). *Let* S *be a finite symmetric subset of* $\mathrm{SL}_2(\mathbf{Z})$. *Let* Γ *be the subgroup generated by* S. *Assume that for all primes* p *large enough, the reduction modulo* p *of* Γ *is equal to* $\mathrm{SL}_2(\mathbf{F}_p)$. *Then for all such primes, the* Cayley graphs *of* $\mathrm{SL}_2(\mathbf{F}_p)$ *with respect to the reduction of* S *modulo* p *form an* expander family.

We will not give a full proof, but only explain the link between this question and Theorem 2.5.5. A complete account,[2] using only elementary tools, can be found in [59, Ch. 6]; this involves other deep tools, in particular the work of Helfgott [49], and Gowers's notion of quasirandom groups, both of which will be discussed later. And before starting, let us mention that the assumption on S is no onerous, and is often valid, and often easy to check in practice (abstractly, it is equivalent to asking that the group Γ is what is known as "Zariski-dense" in the group $\mathrm{SL}_2(\mathbf{C})$, which is a very mild condition).

EXAMPLE 2.6.2. A good example to keep in mind is

$$S = \left\{ \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} \right\}. \tag{2.17}$$

In this case, it is quite elementary that Γ reduces to all of $\mathrm{SL}_2(\mathbf{F}_p)$ for all primes $p \neq 3$, because for such a prime we get

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma \bmod p, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in \Gamma \bmod p,$$

by looking at the k-th power of the generators, where k is the inverse of 3 modulo p, and it is an elementary fact that these two matrices generate $\mathrm{SL}_2(\mathbf{F}_p)$.

(It is also true that the integral matrices $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ generate the group $\mathrm{SL}_2(\mathbf{Z})$, but note that no power of the elements in S are equal to one of these; in fact, one can show that the group Γ has *infinite* index in $\mathrm{SL}_2(\mathbf{Z})$, see for instance [59, Prop. B.1.3].)

[2] Up to the use of the so-called Tits alternative at some point, which can be avoided however in many cases, including the example (2.17).

To discuss the theorem of Bourgain and Gamburd, we first explain briefly what it is meaning of the conclusion. The involves Cayley graphs and expander graphs, and we first spell out our convention concerning graphs (see also Section A.5).

DEFINITION 2.6.3 (Graph). A *graph* is a pair (V, E) of sets, called *vertices* and *edges*, respectively, such that E is a set of subsets of V with $|e| = 2$ for any $e \in \mathrm{E}$.

A graph is said to be finite if V is finite, in which case E is also finite.

Given a vertex $x \in \mathrm{V}$, the *neighbours* of x are the $y \in \mathrm{V}$ such that $\{x, y\}$ is an edge of the graph.

REMARK 2.6.4. For certain purposes, this definition is not the best, since it excludes the possibility of having multiple edges between vertices or loops,[3] and always codes edges by subsets of size 2 of the set of vertices, which can be awkward.[4] However, this will be sufficient for us.

Cayley graphs are defined using a group G and a symmetric subset S of G: the corresponding Cayley Graph $\mathscr{C}(\mathrm{G}, \mathrm{S})$ has vertex set G and edges of the form $\{g, gs\}$ for $g \in \mathrm{G}$ and $s \in \mathrm{S} - \{1\}$. For instance, if $\mathrm{G} = \mathbf{Z}^2$ and $\mathrm{S} = \{(0,1), (0,-1), (1,0), (-1,0)\}$, then $\mathscr{C}(\mathrm{G}, \mathrm{S})$ can be viewed as an infinite square grid.

Expander graphs, on the other hand, are certain *infinite* families of finite graphs which are simultaneously highly connected and rather sparse. To give the precise combinatorial definition, we introduce the notation

$$\partial \mathrm{X} = \{y \in \mathrm{V} - \mathrm{X} \mid \{x, y\} \in \mathrm{E}\}$$

for the *boundary* of a set X of vertices of G: the set of edges with one vertex in X and one outside X (figuratively speaking, the edges that join X to the rest of the graph).

DEFINITION 2.6.5 (Expander graphs). For any finite non-empty graph $\gamma = (\mathrm{V}, \mathrm{E})$, the quantity

$$h(\gamma) = \min_{\substack{\mathrm{X} \subset \mathrm{V} \\ |\mathrm{X}| \leqslant |\mathrm{V}|/2}} \frac{|\partial \mathrm{X}|}{\min(|\mathrm{X}|, |\mathrm{V} - \mathrm{X}|)},$$

is called the discrete *Cheeger constant* of the graph.

A sequence $(\gamma_n)_{n \geqslant 1}$ of finite non-empty graphs $\gamma_n = (\mathrm{V}_n, \mathrm{E}_n)$ is called an *expander family* if

(1) the size of V_n tends to infinity as $n \to +\infty$;

(2) there exists a real number $\mathrm{C} \geqslant 0$ such that for all $n \geqslant 1$ and all $x \in \mathrm{V}_n$, the number of neighbours of x is $\leqslant \mathrm{C}$;

(3) there exists $\delta > 0$, independent of n, such that $h(\gamma_n) \geqslant \delta$ for all n.

There is much to say about expander graphs (see, for instance, the survey [51] of Hoory, Linial and Wigderson, or the book [59]), but for the present discussion, two points are especially relevant: (1) these graphs have remarkable properties; (2) their existence is not obvious at all, and checking if "explicit" families of graphs form an expander family can be very difficult. These facts explain in part the interest of Theorem 2.6.1.

The proof of this theorem is quite involved. Here are the key steps.

Step 1. (*Reduction to return probability*) One of the basic properties of expander graphs is that the combinatorial definition, for a family (γ_n), is equivalent to a spectral

[3] Which means one cannot say, e.g., that an infinite regular tree is the universal covering of a *bouquet* of loops.

[4] It may require non-natural identifications.

property of the discrete Laplace operator of the graphs (the existence of a uniform spectral gap), and is implied by[5] uniform exponential convergence to "equilibritum" of a lazy random walk on γ_n. In the case of a family $(\mathscr{C}(\mathrm{G}_n, \mathrm{S}_n))_{n\geqslant 1}$ of Cayley graphs, with S_n a neutral symmetric set of size bounded independently of n and the size of G_n going to infinity, this boils down to the following concrete statement: there exists a positive real number $\delta < 1$, independent of n, such that

$$\max_{g\in \mathrm{G}_n} \left| \frac{1}{|\mathrm{G}_n|} - \frac{1}{|\mathrm{S}_n|^k} |(s_1, \dots, s_k) \in \mathrm{S}_n^k \mid s_1 \cdots s_k = g\}| \right| \leqslant \delta^k$$

for all $k \geqslant 0$ and $n \geqslant 1$. (In other words: a "long" product of elements of S_n, taken arbitrarily, has about the same chance to be equal to any given element of G_n, and the difference between the chances for different elements goes to 0 at an exponential rate independent of n and of the element which is targeted.)

If the sets S_n generate G_n, a relatively elementary argument (which amounts to the basic theory of finite Markov chains) shows that the property above at least holds with δ replaced by some $\delta_n < 1$, which may depend on n: for fixed n, there exists $\delta_n < 1$ such that

$$\left| \frac{1}{|\mathrm{G}_n|} - \frac{1}{|\mathrm{S}_n|^k} |(s_1, \dots, s_k) \in \mathrm{S}_n^k \mid s_1 \cdots s_k = g\}| \right| \leqslant \delta_n^k$$

for all $g \in \mathrm{G}_n$ and $k \geqslant 0$.

In the case of interest, the sequence of groups are the groups $\mathrm{SL}_2(\mathbf{F}_p)$, indexed by the primes, and S_p is the image of S modulo p, which by assumption generates $\mathrm{SL}_2(\mathbf{F}_p)$ for all primes p large enough. Then, an argument related to the "quasirandomness" of these groups (in the sense of Gowers, see Section 2.8), leads to a bound of the form

$$\delta_n \leqslant \left(\frac{2|\mathrm{SL}_2(\mathbf{F}_p)|}{p-1} \frac{1}{|\mathrm{S}_p|^{2k}} \left| \{(s_1, \dots, s_{2k}) \in \mathrm{S}_p^{2k} \mid s_1 \cdots s_{2k} = 1\} \right| \right)^{1/(2k)}$$

for any integer $k \geqslant 1$. A simple computation shows then that one will have a uniform $\delta < 1$ such that $\delta_n \leqslant \delta$ for all n *if* one can show that there exists $c > 0$, independent of p, such that

$$\frac{1}{|\mathrm{SL}_2(\mathbf{F}_p)|} |\{(s_1, \dots, s_{2k}) \in \mathrm{S}_p^{2k} \mid s_1 \cdots s_{2k} = 1\}| \leqslant \frac{1}{|\mathrm{SL}_2(\mathbf{F}_p)|^{5/6}}, \tag{2.18}$$

(say), for *some* integer k with $k \leqslant c \log p$. The left-hand side is a "return probability": intuitively, it is the probability of coming back to 1 if one multiples $2k$ matrices in S_p chosen uniformly and independently at random.[6]

Remark 2.6.6. This reduction is reasonable because the basic properties of expander graphs imply that this will be true if the Cayley graphs $\mathscr{C}(\mathrm{SL}_2(\mathbf{F}_p), \mathrm{S}_p)$ form an expander.

Step 2. (L^2-*flattening and approximate subgroups*) For $k \geqslant 0$, let

$$\varrho_k = \frac{1}{|\mathrm{SL}_2(\mathbf{F}_p)|} |\{(s_1, \dots, s_{2k}) \in \mathrm{S}_p^{2k} \mid s_1 \cdots s_{2k} = 1\}|,$$

which we want to show is small for suitable k.

Extremely roughly, computing ϱ_{2k} means counting products of length $4k$ of elements of S_p that are equal to 1, i.e., counting solutions of

$$x_1 x_2 x_3 x_4 = 1$$

[5] But not quite equivalent with.

[6] The exponent 5/6 is not crucial, but it must be $> 2/3$.

where $x_i \in \mathrm{S}_p^{(k)}$. Since $\mathrm{S}_p = \mathrm{S}_p^{-1}$, this suggests that ϱ_{2k} can only be large if S_k has "large" multiplicative energy, which means that some approximate subgroup contains a significant part of a translate of S_k.

This already shows the relevance of approximate subgroups to the problem, but the actual argument is more involved. Indeed, in order to compute ϱ_k, the multiplicative energy of $\mathrm{S}_p^{(k)}$ must be combined with the number of ways to write, say, $x_1 = s_1 \dots s_k$, which is not the same for all $x_1 \in \mathrm{S}_p^{(k)}$.

The actual approach of Bourgain and Gamburd to the proof of (2.18) is through an iterative process which demonstrates that ϱ_{2k} is in fact significantly smaller than ϱ_k, in the range of k of interest, unless certain conditions are met.

More precisely, they (essentially) prove that there exists a real number $c > 0$ (independent of any data involved here, including p), such that we have

$$\varrho_{2k} \leqslant c\left(\frac{1}{|\mathrm{SL}_2(\mathbf{F}_p)|^{5/6}} + \varrho_k e(\mathrm{A},\mathrm{B})\right) \tag{2.19}$$

for *some* subsets A and B of $\mathrm{S}_p^{(k)}$ of controlled size, in particular with

$$\max(|\mathrm{A}|,|\mathrm{B}|) \leqslant |\mathrm{SL}_2(\mathbf{F}_p)|^{1-\gamma}$$

for suitably small $\gamma > 0$. The idea (and this explains the name "L^2-flattening lemma" given to this argument) is that we have

$$\varrho_k = \sum_{g\in\mathrm{SL}_2(\mathbf{F}_p)} \mu_k(g)^2,$$

where

$$\mu_k(g) = \frac{1}{|\mathrm{SL}_2(\mathbf{F}_p)|}|\{(s_1,\dots,s_k) \in \mathrm{S}_p^k \mid s_1\cdots s_k = g\}|,$$

and one combines this expression (in the case of ϱ_{2k}) with various ingredients, such as partitioning "dyadically" the sum over $g \in \mathrm{SL}_2(\mathbf{F}_p)$ according to the number of products of k elements of S_p which are equal to g).

From (2.19), it is not difficult to deduce that the result will be reached if one can show that there is a real number $\kappa > 0$, independent of p and k, such that the subsets which occur in that inequality satisfy

$$e(\mathrm{A},\mathrm{B}) \leqslant \frac{1}{p^\kappa}. \tag{2.20}$$

Step 3. (*Conclusion*) By contraposition, the previous steps leads to the appearance of some approximate subgroups of $\mathrm{SL}_2(\mathbf{F}_p)$, through the Balog–Szemerédi–Gowers Theorem: if the goal (2.20) *fails*, this means that we have A, B with

$$e(\mathrm{A},\mathrm{B}) \geqslant \frac{1}{p^\kappa},$$

hence A and B are both related by Theorem 2.5.5 to a β-approximate subgroup with β of size at most $p^{1000\kappa}$.

This would be a dead-end if one didn't have some knowledge about β-approximate subgroups of $\mathrm{SL}_2(\mathbf{F}_p)$ for such values of β, namely a small but *fixed* positive power of p. And indeed, the impetus for the work of Bourgain and Gamburd was precisely that, at that time, Helfgott had proved an extremely strong statement concerning them. This takes the following form:

THEOREM 2.6.7 (Helfgott). *There exists a real number $\delta > 0$ with the following property: for any prime number p and for any neutral symmetric subset H of $\mathrm{SL}_2(\mathbf{F}_p)$, we have*

$$|\mathrm{H}^{(3)}| \geqslant \min(|\mathrm{SL}_2(\mathbf{F}_p)|, |\mathrm{H}|^{1+\delta}),$$

unless A is contained in a proper subgroup of $\mathrm{SL}_2(\mathbf{F}_p)$.

In particular, any β-approximate subgroup H of $\mathrm{SL}_2(\mathbf{F}_p)$ such that $\beta \leqslant |\mathrm{H}|^{\delta/2}$ is either contained in a proper subgroup of $\mathrm{SL}_2(\mathbf{F}_p)$ or satisfies $|\mathrm{H}| \geqslant |\mathrm{G}|/\beta^2$.

Bourgain and Gamburd proved that, in their situation, one could control the sizes of the sets A and B sufficiently to ensure that $|\mathrm{A}|$ is not "too big", in particular so that the conclusion $\mathrm{A}^{(3)} = \mathrm{SL}_2(\mathbf{F}_p)$ can be excluded, and not "too small" (of size at least a small positive power of A). Thus only approximate subgroups contained in a (relatively large) proper subgroup could possibly thwart the implementation of the L^2-flattening strategy. But one knows that $\mathrm{A} \subset \mathrm{S}_p^{(k)}$ and that S_p generates $\mathrm{SL}_2(\mathbf{F}_p)$, and it is relatively easy (using the classification of proper subgroups of $\mathrm{SL}_2(\mathbf{F}_p)$) to show that this can not happen for k in the required range.[7]

REMARK 2.6.8. (1) Helfgott did not compute a value of the constant δ in Theorem 2.6.7, although this was clearly doable from his argument. It was proved in [57], by following through the simplest case of the proof of Pyber and Szabó of a considerable generalization of the theorem (see the introduction of [66]), that $\delta = 1/3024$ is possible. Rudnev and Shkredov [69] have improved this much further, obtaining the constant $\delta = 1/20$, up to an absolute multiplicative factor (i.e., they show that $|\mathrm{A}^{(3)}| \geqslant c|\mathrm{A}|^{1+1/20}$ for some absolute constant $c > 0$).

(2) One can compare this statement with the Cauchy–Davenport Theorem, for a subset $\mathrm{A} \subset \mathbf{Z}/p\mathbf{Z}$, and we see that it is incredibly more powerful: instead of $|2\mathrm{A}| \geqslant \min(p, 2|\mathrm{A}| - 1)$, which only gives "growth" by a constant factor, we have growth by a factor $|\mathrm{A}|^{\delta}$.

The following exercise shows one of the easiest consequence of the expansion property.

EXERCISE 2.6.9. (1) Show that if $(\gamma_n)_{n\geqslant 1}$ is an expander family, then the diameter of γ_n grows logarithmically: there exists $c > 0$ such that $\mathrm{diam}(\gamma_n) \leqslant c \log |\mathrm{V}_n|$ for all $n \geqslant 1$, where V_n is the set of vertices of γ_n. (The *diameter* of a graph is the supremum of the integers k such that for any two vertices x and y, there are vertices

$$x_0 = x, \quad x_1, \quad \dots, \quad x_{k-1}, \quad x_k = y$$

such that $\{x_i, x_{i+1}\}$ is an edge for $0 \leqslant i \leqslant k-1$.)

In particular, if $\gamma_n = \mathscr{C}(\mathrm{G}_n, \mathrm{S}_n)$ for some finite groups G_n and generating sets S_n, there exists $k_n \leqslant c \log |\mathrm{G}_n|$ such that $\mathrm{S}_n^{(k_n)} = \mathrm{G}_n$.

(2) For comparison, show that Theorem 2.6.7 implies that there exists real numbers c_1, $c_2 \geqslant 0$ such that, for any symmetric generating set S of $\mathrm{SL}_2(\mathbf{F}_p)$, the Cayley graph $\mathscr{C}(\mathrm{SL}_2(\mathbf{F}_p), \mathrm{S})$ has diameter $\leqslant c_1(\log p)^{c_2}$.

Here is however an open problem: given a "concrete" family of expanding Cayley graphs (say those from Theorem 2.6.1), find an efficient algorithm which, given $x \in \mathrm{G}_n$,

[7] E.g., one can show that if $p \geqslant 7$, the set $\mathrm{S}_p^{(3)}$ must contain a matrix which is diagonalizable with distinct eigenvalues.

expresses it as a short product of the generators. Already if we take the case of (2.17), the question is not fully solved for

$$x = \begin{pmatrix} 1 & \frac{1}{2}(p-1) \bmod p \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbf{F}_p).$$

Exercise 2.6.10. Let G be a group and H a subgroup of G. Let $x \in \mathrm{G}$, and define $\mathrm{I} = \mathrm{H} \cap x^{-1}\mathrm{H}x$; this is a subgroup of H.

(1) For h_1 and $h_2 \in \mathrm{H}$, show that

$$\mathrm{H}xh_1 \cap \mathrm{H}xh_2 = \emptyset$$

unless if $h_1^{-1}h_2 \in \mathrm{I}$.

(2) If $h_1^{-1}h_2 \in \mathrm{I}$, on the other hand, show that

$$\mathrm{H}xh_1 = \mathrm{H}xh_2.$$

(3) Deduce that the product set $\mathrm{H}x\mathrm{H}$ (known as a *double coset* of H) is the disjoint union of $\mathrm{H}xy$ for y running over a set of representatives of the cosets $h\mathrm{I}$ of I in H. In particular, if H is finite, deduce that

$$|\mathrm{H}x\mathrm{H}| = [\mathrm{H} : \mathrm{I}]\ |\mathrm{H}|.$$

Exercise 2.6.11. Let p be a prime number and let

$$\mathrm{U} = \left\{ \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mid t \in \mathbf{F}_p \right\}, \qquad \mathrm{B} = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mid a, b, d \in \mathbf{F}_p,\ ad = 1 \right\}.$$

Set $\mathrm{U}^* = \mathrm{U} - \{1\}$.

(1) Show that U and B are subgroups of $\mathrm{SL}_2(\mathbf{F}_p)$ with $|\mathrm{U}| = p$ and $|\mathrm{B}| = p(p-1)$.

(2) Let $x \in \mathrm{SL}_2(\mathbf{F}_p) - \mathrm{B}$. Show that the map

$$\begin{cases} \mathrm{U}^* \times \mathrm{U}^* \times \mathrm{U}^* & \to & \mathrm{SL}_2(\mathbf{F}_p) \\ (u, v, w) & \mapsto & uxvx^{-1}w \end{cases}$$

is injective.

(3) Let A be a symmetric subset of $\mathrm{SL}_2(\mathbf{F}_p)$. Show that either $\mathrm{A} \subset \mathrm{B}$ or

$$|\mathrm{U}^* \cap \mathrm{A}|^3 \leqslant |\mathrm{A}^{(5)}|.$$

(This is a very special case of what are called *Larsen–Pink non-concentration inequalities*.)

(4) Let $x \in \mathrm{SL}_2(\mathbf{F}_p) - \mathrm{B}$. Let $\mathrm{A} = \mathrm{U} \cup \{x, x^{-1}\}$. Show that there exists $c > 0$ and $\delta > 0$, independent of p and x, such that

$$|\mathrm{A}^{(3)}| \geqslant c|\mathrm{A}|^{1+\delta}.$$

How large can you get δ to be?

Exercise 2.6.12. Let p be an odd prime number. With the same notation as in the previous exercise, consider

$$x = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \in \mathrm{SL}_2(\mathbf{F}_p).$$

Let K be a subgroup of B such that $x^2 \in \mathrm{K}$. Let $\mathrm{A} = \mathrm{K} \cup \{x, x^{-1}\}$.

(1) Show that

$$\mathrm{A}^{(3)} = \mathrm{K} \cup \mathrm{K}x\mathrm{K} \cup x^{-1}\mathrm{K}x.$$

(2) Deduce that
$$|\mathrm{A}^{(3)}| \leqslant (2+c)|\mathrm{K}|,$$
where c is the index of $\mathrm{K} \cap x^{-1}\mathrm{K}x$ in K. (Hint: use the first exercise.)

(3) Assume that -1 is a square modulo p (which means that p is congruent to 1 modulo 4). Let K be the subgroup of B of the form
$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$
where a is a square modulo p. Show that $x^2 \in \mathrm{K}$ and
$$[\mathrm{K} : \mathrm{K} \cap x^{-1}\mathrm{K}x] = p.$$

(4) Under the same assumption, show that $\mathrm{A}^{(3)} \neq \mathrm{SL}_2(\mathbf{F}_p)$, and
$$|\mathrm{A}^{(3)}| \leqslant c'|\mathrm{A}|^{3/2}$$
for some constant $c' \geqslant 0$; you may use without proof the fact that $|\mathrm{SL}_2(\mathbf{F}_p)| = p(p^2-1)$ for all p odd.

(One can show that A is a generating set of $\mathrm{SL}_2(\mathbf{F}_p)$, so this example shows that the best exponent in Theorem 2.6.7 cannot be larger than $1/2$.)

2.7. Proof of Theorem 2.5.5

We first give the proof of a weaker statement, which is however sufficient for many applications, and is both simpler and involves better constants.

THEOREM 2.7.1. *Let* G *be a group and* $\mathrm{A} \subset \mathrm{G}$ *a non-empty finite subset. Let* $\alpha \geqslant 1$ *be such that* $e(\mathrm{A}) \geqslant \alpha^{-1}$. *There exists a subset* $\mathrm{B} \subset \mathrm{A}$ *such that*
$$|\mathrm{B}| \geqslant \frac{|\mathrm{A}|}{4\alpha}, \qquad |\mathrm{B}\cdot\mathrm{B}^{-1}| \ll \alpha^5|\mathrm{A}| \ll \alpha^6|\mathrm{B}|, \tag{2.21}$$
where the implied constant is absolute. In particular, we have also
$$\exp d(\mathrm{B},\mathrm{B}) \ll \alpha^7. \tag{2.22}$$
Similarly, there exists a subset $\mathrm{B}' \subset \mathrm{A}$ *such that*
$$|\mathrm{B}'| \geqslant \frac{|\mathrm{A}|}{4\alpha}, \qquad |\mathrm{B}'\cdot\mathrm{B}'| \ll \alpha^6|\mathrm{A}| \ll \alpha^7|\mathrm{B}'|, \tag{2.23}$$
where the implied constant is absolute.

The proof follows a write-up by B. Green of the original argument of Schoen [74]. The key step is to find a large subset X of A such that the elements of $\mathrm{X}\cdot\mathrm{X}^{-1}$ have a large number of representations as elements of $\mathrm{A}\cdot\mathrm{A}^{-1}$. The precise statement is the following:

PROPOSITION 2.7.2. *Let* G *be a group and* $\mathrm{A} \subset \mathrm{G}$ *a non-empty finite subset. Let* $\alpha \geqslant 1$ *be such that* $e(\mathrm{A}) \geqslant \alpha^{-1}$. *Fix a real number* δ *such that* $0 < \delta < 1$.

There exists $z \in \mathrm{G}$ *such that*
$$|\mathrm{A} \cap \mathrm{A}\cdot z| \geqslant \frac{|\mathrm{A}|}{2\alpha} \tag{2.24}$$
and
$$\left|\left\{(x,y) \in (\mathrm{A}\cap\mathrm{A}\cdot z)^2 \mid r(xy^{-1}) \geqslant \frac{\delta|\mathrm{A}|}{2\alpha^2}\right\}\right| \geqslant (1-\delta)|\mathrm{A}\cap\mathrm{A}\cdot z|^2, \tag{2.25}$$
where r *is the representation function for* $\mathrm{A}\cdot\mathrm{A}^{-1}$.

Note that this will be applied for a fixed δ (for instance, $\delta = 1/10$), so the proof can be read with such a value in mind.

PROOF. The key idea of Schoen is to take z "at random", but not according to the uniform probability measure on G. Rather, we pick a given element z with probability proportional to $r(z)$. Since

$$\sum_{x\in \mathrm{G}} r(x) = |\mathrm{A}||\mathrm{A}^{-1}| = |\mathrm{A}|^2,$$

(see (2.1)), this means that we have

$$\mathbf{P}(z = x) = \frac{1}{|\mathrm{A}|^2} r(x)$$

for any $x \in \mathrm{G}$.

Thus z is a G-valued random variable with this distribution. We further denote $\mathrm{B} = \mathrm{A} \cap \mathrm{A} \cdot z$, which is a random subset of G (contained in A).

Let $\gamma > 0$ be a parameter, to be chosen later. We define

$$\mathrm{Y} = \{(a, b) \in \mathrm{A} \times \mathrm{A} \mid r(ab^{-1}) < \gamma|\mathrm{A}|\}.$$

We will show that for $\gamma = \delta/(2\alpha^2)$, the inequality

$$\mathbf{E}\Big(|\mathrm{B}|^2 - \delta^{-1}|(\mathrm{B} \times \mathrm{B}) \cap \mathrm{Y}|\Big) \geqslant \frac{|\mathrm{A}|^2}{2\alpha^2} \tag{2.26}$$

holds. It implies the existence of some element $z \in \mathrm{G}$ such that

$$|\mathrm{A} \cap \mathrm{A} \cdot z|^2 - \delta^{-1}|(\mathrm{A} \cap \mathrm{A} \cdot z)^2 \cap \mathrm{Y}| \geqslant \frac{|\mathrm{A}|^2}{2\alpha^2},$$

and from this we deduce, on the one hand, that $|\mathrm{A} \cap \mathrm{A} \cdot z|^2 \geqslant |\mathrm{A}|^2/(2\alpha^2)$, which implies (2.24), and on the other hand that

$$|(\mathrm{A} \cap \mathrm{A} \cdot z)^2 \cap \mathrm{Y}| \leqslant \delta|\mathrm{A} \cap \mathrm{A} \cdot z|^2,$$

which is equivalent to (2.25).

To prove (2.26), we first find a lower-bound for $\mathbf{E}(|\mathrm{B}|^2)$. By the Cauchy–Schwarz inequality, we have

$$\mathbf{E}(|\mathrm{B}|^2) \geqslant \mathbf{E}(|\mathrm{B}|)^2,$$

and the expectation of the size of B is

$$\mathbf{E}(|\mathrm{B}|) = \sum_{a\in \mathrm{A}} \mathbf{P}(a \in \mathrm{A} \cdot z) = \sum_{a\in \mathrm{A}} \sum_{b\in \mathrm{A}} \mathbf{P}(z = b^{-1}a) = \frac{1}{|\mathrm{A}|^2} \sum_{a\in \mathrm{A}} \sum_{b\in \mathrm{A}} r(b^{-1}a).$$

But, by replacing $r(b^{-1}a)$ by its definition, we compute

$$\frac{1}{|\mathrm{A}|^2} \sum_{a\in \mathrm{A}} \sum_{b\in \mathrm{A}} r(b^{-1}a) = \frac{1}{|\mathrm{A}|^2} \sum_{a\in \mathrm{A}} \sum_{b\in \mathrm{A}} \sum_{\substack{(x,y)\in \mathrm{A}^2 \\ xy^{-1}=b^{-1}a}} 1 = |\mathrm{A}|e(\mathrm{A}).$$

Using the assumption, we get from this the lower bound

$$\mathbf{E}(|\mathrm{B}|^2) \geqslant \frac{|\mathrm{A}|^2}{\alpha^2}.$$

We now handle separately an upper bound for the expectation of $(\mathrm{B} \times \mathrm{B}) \cap \mathrm{Y}$. We simply write

$$\mathbf{E}(|(\mathrm{B} \times \mathrm{B}) \cap \mathrm{Y}|) \leqslant |\mathrm{A}|^2 \max_{(a,b)\in \mathrm{Y}} \mathbf{P}(\{a, b\} \subset \mathrm{B}),$$

and estimate the probability that $\{a,b\} \subset \mathrm{B}$ for each $(a,b) \in \mathrm{Y}$ separately. Since $\mathrm{Y} \subset \mathrm{A}^2$, this is

$$\mathbf{P}(a \in \mathrm{B} \text{ and } b \in \mathrm{B}) = \mathbf{P}(a \in \mathrm{A}\cdot z \text{ and } b \in \mathrm{A}\cdot z) = \mathbf{P}(z \in \mathrm{A}^{-1}\cdot a \cap \mathrm{A}^{-1}\cdot b).$$

From the crude bound $r(x) \leqslant |\mathrm{A}|$, it follows that $\mathbf{P}(z = x) \leqslant 1/|\mathrm{A}|$ for any $x \in \mathrm{G}$, and we deduce that

$$\mathbf{P}(z \in \mathrm{A}^{-1}\cdot a \cap \mathrm{A}^{-1}\cdot b) \leqslant \frac{1}{|\mathrm{A}|}|\mathrm{A}^{-1}\cdot a \cap \mathrm{A}^{-1}\cdot b|.$$

Note that $\mathrm{A}^{-1}\cdot a \cap \mathrm{A}^{-1}\cdot b$ is in bijection with the set of pairs $(x,y) \in \mathrm{A}^2$ such that $xy^{-1} = ab^{-1}$, by means of the map f which sends an element w of the intersection to (aw^{-1}, bw^{-1}), with inverse $(x,y) \mapsto a^{-1}x = b^{-1}y$. Thus we get

$$\mathbf{P}(a \in \mathrm{B} \text{ and } b \in \mathrm{B}) \leqslant \frac{1}{|\mathrm{A}|}\sum_{\substack{(x,y)\in \mathrm{A}^2\\ xy^{-1}=ab^{-1}}} 1 = \frac{r(ab^{-1})}{|\mathrm{A}|},$$

and by definition of Y, this is $< \gamma|\mathrm{A}|$. Thus we have

$$\mathbf{E}\Big(|\mathrm{B}|^2 - \delta^{-1}|(\mathrm{B}\times\mathrm{B})\cap \mathrm{Y}|\Big) \geqslant \frac{|\mathrm{A}|^2}{\alpha^2} - \frac{\gamma|\mathrm{A}|^2}{\delta},$$

and this is $\geqslant |\mathrm{A}|^2/(2\alpha^2)$ if we take $\gamma = \delta/(2\alpha^2)$, as claimed. □

Proof of Theorem 2.7.1. We first observe that the last conclusion (2.22) follows from (2.21) and the definition of the Ruzsa distance: we have

$$\exp d(\mathrm{B},\mathrm{B}) = \frac{|\mathrm{B}\cdot\mathrm{B}^{-1}|}{|\mathrm{B}|} \ll \frac{\alpha^6|\mathrm{A}|}{|\mathrm{B}|} \ll \alpha^7.$$

We will prove the existence of B, and leave the similar proof of the existence of the set B′ to the reader.

We apply the proposition with $\delta = 1/10$, and denote by C the set $\mathrm{A}\cap \mathrm{A}\cdot z$ which it provides, and by X the set of $g \in \mathrm{G}$ with $r(g) \geqslant \delta|\mathrm{A}|/(2\alpha^2) = |\mathrm{A}|/(20\alpha^2)$, where r is the representation function for $\mathrm{A}\cdot\mathrm{A}^{-1}$. We note that since the sum of $r(g)$ over all g is equal to $|\mathrm{A}|^2$, we have

$$|\mathrm{X}| \leqslant 20\alpha^2|\mathrm{A}|. \tag{2.27}$$

Further, for any element $a \in \mathrm{A}$, we let $\mathrm{N}(a)$ denote the set of $b \in \mathrm{C}$ such that $ab^{-1} \in \mathrm{X}$. We have $0 \leqslant |\mathrm{N}(c)| \leqslant |\mathrm{C}|$ for any $c \in \mathrm{C}$; moreover, by (2.25), we have

$$\sum_{c\in\mathrm{C}} |\mathrm{N}(c)| \geqslant (1-\delta)|\mathrm{C}|^2,$$

and this implies that $\mathrm{N}(c)$ must often be quite close to its maximal value. Precisely, for any $\gamma > 0$, we have[8]

$$\begin{aligned}\frac{1}{|\mathrm{C}|}|\{c\in\mathrm{C} \mid |\mathrm{N}(c)| < (1-\gamma)|\mathrm{C}|\}| &= \frac{1}{|\mathrm{C}|}\sum_{|\mathrm{C}|-|\mathrm{N}(c)|>\gamma|\mathrm{C}|} 1\\ &\leqslant \frac{1}{\gamma|\mathrm{C}|}\times\frac{1}{|\mathrm{C}|}\sum_{c\in\mathrm{C}}(|\mathrm{C}|-|\mathrm{N}(c)|) \leqslant \frac{\delta}{\gamma},\end{aligned}$$

and taking $\gamma = \sqrt{\delta}$, we find that there are at least $(1-\sqrt{\delta})|\mathrm{C}|$ elements of C such that $|\mathrm{N}(c)| \geqslant (1-\sqrt{\delta})|\mathrm{C}|$.

[8] This is really the Chebychev inequality implemented "inline".

Let B be the subset of C (hence of A) defined by this condition on $\mathrm{N}(c)$; since the proposition implies that $|\mathrm{C}| \geqslant |\mathrm{A}|/(2\alpha)$, we already get

$$|\mathrm{B}| \geqslant (1-\sqrt{\delta})|\mathrm{C}| \geqslant \frac{|\mathrm{C}|}{2} \geqslant \frac{|\mathrm{A}|}{4\alpha}.$$

To conclude the proof, we claim that

$$\mathrm{B}\cdot\mathrm{B}^{-1} \subset \left\{x \in \mathrm{G} \mid s(x) \geqslant \frac{|\mathrm{C}|}{3}\right\}, \tag{2.28}$$

where s is the representation function for $\mathrm{X}\cdot\mathrm{X}^{-1}$. Assuming this, we observe that the right-hand set satisfies

$$\left|\left\{x \in \mathrm{G} \mid s(x) \geqslant \frac{|\mathrm{C}|}{3}\right\}\right| \leqslant \frac{3|\mathrm{X}|^2}{|\mathrm{C}|}$$

(as before, since the sum of all $s(x)$ is $|\mathrm{X}|^2$). Using $|\mathrm{C}| \geqslant |\mathrm{A}|/(2\alpha)$ together with (2.27), we deduce

$$|\mathrm{B}\cdot\mathrm{B}^{-1}| \leqslant \frac{3|\mathrm{X}|^2}{|\mathrm{C}|} \leqslant 6\cdot 20^2\cdot\alpha^5|\mathrm{A}|,$$

which finishes the proof of the theorem.

To prove (2.28), pick any a and b in B; we need a lower bound for $s(ab^{-1})$, or in other words for the size of the set

$$\{(u,v) \in \mathrm{X}\times\mathrm{X} \mid uv^{-1} = ab^{-1}\}.$$

There is an injective map

$$\mathrm{N}(a)\cap\mathrm{N}(b) \to \{(u,v) \in \mathrm{X}\times\mathrm{X} \mid uv^{-1} = ab^{-1}\}$$

defined by $f(z) = (az^{-1}, bz^{-1})$ (the crucial point here is that this map is well-defined: we have $(az^{-1}, bz^{-1}) \in \mathrm{X}\times\mathrm{X}$ by definition of $\mathrm{N}(a)$ and $\mathrm{N}(b)$). Hence $s(ab^{-1}) \geqslant |\mathrm{N}(a)\cap\mathrm{N}(b)|$. But, by definition, $|\mathrm{N}(a)|$ and $|\mathrm{N}(b)|$ are very large, and so is their intersection. In fact, we get

$$|\mathrm{N}(a)\cap\mathrm{N}(b)| \geqslant (1-2\sqrt{\delta})|\mathrm{C}| \geqslant \frac{|\mathrm{C}|}{3},$$

(recall that $\delta = 1/10$), so that $s(ab^{-1}) \geqslant |\mathrm{C}|/3$, as desired. □

The proof of the full version of Theorem 2.5.5 proceeds in a few steps, which we summarize as follows, using c_i to denote some numerical constants, which can all be made explicit:

(1) A result in graph theory, of independent interest, states that if $\gamma = (\mathrm{V},\mathrm{E})$ is a finite bipartite graph, with $\mathrm{V} = \mathrm{V}_1 \sqcup \mathrm{V}_2$ its bipartite decomposition, such that $|\mathrm{E}| \geqslant \alpha^{-1}|\mathrm{V}_1||\mathrm{V}_2$ for some $\alpha \geqslant 1$, then one can find subsets $\mathrm{U}_i \subset \mathrm{V}_i$ with $|\mathrm{V}_i| \leqslant c_1\alpha|\mathrm{U}_i|$, such that any pair $(u_1,u_2) \in \mathrm{U}_1\times\mathrm{U}_2$ is joined by at $\geqslant c_2\alpha^{-4}|\mathrm{V}_1||\mathrm{V}_2|$ paths of length 3.

(2) Given A, B with $e(\mathrm{A},\mathrm{B}) \geqslant \alpha^{-1}$, one uses this result to deduce that there exist $\mathrm{A}_1 \subset \mathrm{A}$ and $\mathrm{B}_1 \subset \mathrm{B}$ such that $|\mathrm{A}| \leqslant 2c_1\alpha|\mathrm{A}_1|$, $|\mathrm{B}| \leqslant 2c_2\alpha|\mathrm{B}_1|$, and

$$\log d(\mathrm{A}_1,\mathrm{B}_1) \leqslant c_3\alpha^9.$$

The graph which is used has vertex set the disjoint union $\mathrm{A}\sqcup\mathrm{B}$, and edges those $\{a,b\}$ (with $a \in \mathrm{A}$ and $b \in \mathrm{B}$ in the respective parts of the disjoint union) such that ab has $\geqslant \frac{1}{2}\alpha^{-1}\sqrt{|\mathrm{A}||\mathrm{B}|}$ representations in $\mathrm{A}\cdot\mathrm{B}$.

Remark 2.7.3. Theorem 2.7.1 is really a variant (and its proof a simplification) of the combination of (1) and (2), in the case of a single set.

(3) Given now sets A, B with $\log d(A,B) \leqslant \alpha$ for some $\alpha \geqslant 1$, one proves that there exists a $c_4\alpha^{80}$-approximate subgroup H and a set X such that

$$|X| \leqslant c_5\alpha^{104}, \qquad |H| \leqslant c_6\alpha^{14}|A|$$
$$A \subset XH, \qquad B \subset HX.$$

One can then a simple positivity argument see that there exists elements $x \in X$ and $y \in X$ such that

$$|A \cap xH| \geqslant \frac{|A|}{|X|}, \qquad |B \cap Hy| \geqslant \frac{|A|}{|X|}$$

(because

$$|A| = \sum_{x\in X} |A \cap xH| \leqslant |X| \max_{x\in X} |A \cap xH|$$

for instance; this argument is often called the "pigeonhole principle").

All together, this combines to yield Theorem 2.5.5.

2.8. Quasirandom groups and product-free sets

Gowers [42] introduced a notion of "quasi-random" groups, which is a fairly simple group-theoretic condition which implies that certain associated graphs behave, in certain respects, like random graphs. We will present this relatively briefly, since accounts of this theory already appear in a number of sources. It allows us however to also mention another important topic in additive combinatorics: product-free sets.

Definition 2.8.1. Let $\alpha \geqslant 1$ be a real number. A finite group G is called α-*quasirandom*, or α-quasirandom, if any group morphism $\varrho\colon G \to \mathrm{GL}_n(\mathbf{C})$ with $n < \alpha$ is trivial, in the sense that $\varrho(g) = \mathrm{Id}$ for all $g \in G$.

In the language of representation theory, this is equivalent to saying that G is α-quasirandom if and only if any non-trivial irreducible representation of G has dimension $\geqslant \alpha$.

Example 2.8.2. (1) Any non-trivial finite group G is 1-quasirandom. If G is a abelian, then this is best possible: G is not α-quasirandom for any $\alpha > 1$ (taking $g \in G$ different from 1, Lemma A.7.4 shows that there exists a group morphism $\varrho\colon G \to \mathbf{C}^\times = \mathrm{GL}_1(\mathbf{C})$ such that $\varrho(g) \neq 1$).

More generally, any finite group which is α-quasirandom for some $\alpha > 1$ must be a *perfect group*, meaning that there is no non-trivial group morphism from G to a finite abelian group (given a morphism $f\colon G \to A$ where A is abelian, we would obtain a non-trivial morphism $f\colon G \to \mathbf{C}^\times$ by composing f with a character of A which is non-trivial on some element of the image of f).

(2) On the other hand, if p is an odd prime number, the group $\mathrm{SL}_2(\mathbf{F}_p)$ is α-quasirandom with $\alpha = \frac{1}{2}(p-1)$. This result goes back to Frobenius, who classified the irreducible representations of $\mathrm{SL}_2(\mathbf{F}_p)$; inspecting the dimensions of these, the result immediately follows.

There is also however an elegant elementary argument which we now present. Let $\varrho\colon \mathrm{SL}_2(\mathbf{F}_p) \to \mathrm{GL}_n(\mathbf{C})$ be a group morphism which is not trivial, where $n \geqslant 1$ is an integer. Consider the matrices

$$u_+ = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad u_- = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

which are known to generate $\mathrm{SL}_2(\mathbf{F}_p)$. One at least is therefore not in the kernel of ϱ, say u_+ (the other case is treated similarly). There is then in E a non-zero eigenvector x of the matrix $\varrho(u_+)$ for some eigenvalue $\xi \neq 1$, which must be a primitive p-root of unity.

Let $k \geqslant 1$ be an integer coprime to p. Writing k^{-1} for the inverse of the class of k modulo p, let

$$a = \begin{pmatrix} k & 0 \\ 0 & k^{-1} \end{pmatrix} \in \mathrm{SL}_2(\mathbf{F}_p).$$

A direct computation gives the matrix identity $au_+a^{-1} = u_+^{k^2}$, and the standard computation

$$\varrho(u_+)\varrho(a^{-1})x = \varrho(a^{-1})\varrho(au_+a^{-1})x = \varrho(a^{-1})\varrho(u_+)^{k^2}x = \xi^{k^2}\varrho(a)x,$$

i.e., the vector $\varrho(a)x$ is an eigenvector of $\varrho(u_+)$ with eigenvalue ξ^{k^2}.

As k runs over integers between 1 and $p-1$, the square $k^2 \pmod p$ runs over $(p-1)/2$ distinct elements in $\mathbf{F}_p^\times$, and the corresponding roots of unity ξ^{k^2} are also distinct; thus the matrix $\varrho(u_+)$ on $\mathbf{C}^n$ has at least $(p-1)/2$ distinct eigenvalues, which implies that $n \geqslant (p-1)/2$.

(3) Let $n \geqslant 2$ be an integer. The symmetric group S_n is only 1-quasirandom since the signature is a non-trivial homomorphism $\mathrm{S}_n \to \{-1, 1\} \subset \mathbf{C}^\times$. If $n \geqslant 6$, it is known that the alternating group A_n is $(n-1)$-quasirandom, and that this is sharp.[9] The former result can be deduced from a theorem of Burnside [15, p. 468], and can nowadays also be proved using formulas for the dimension of the irreducible representations of S_n; to see that the bound is sharp, note that there is an injective (hence non-trivial) group morphism $\mathrm{A}_n \to \mathrm{GL}_n(\mathbf{C})$, mapping each permutation to the corresponding permutation matrix, and observe that each permutation matrix leaves invariant the $(n-1)$-dimensional subspace of $\mathbf{C}^n$ defined by the condition

$$x_1 + \cdots + x_n = 0, \qquad \text{for } (x_i) \in \mathbf{C}^n.$$

Picking a basis of this subspace, the resulting morphism $\mathrm{A}_n \to \mathrm{GL}_{n-1}(\mathbf{C})$ is still non-trivial.

Exercise 2.8.3. This exercise shows that A_n is $\varphi(n)/2$-quasirandom when $n \geqslant 4$ is *odd*, where $\varphi(n)$ is the cardinality of the group of invertible elements in $\mathbf{Z}/n\mathbf{Z}$. In particular, this proves that A_n is $(p-1)/2$-quasirandom when $p \geqslant 5$ is an odd prime. (In general, $\varphi(n)$ is not far from n: it is known that $\varphi(n) \gg n(\log\log n)^{-1}$, see, e.g., [48, Th. 328]).

We recall that A_n is a simple group if $n \geqslant 5$, and we denote by σ an n-cycle in S_n. We always assume here that n is odd.

(1) Show that $\varphi(n)$ is even.

(2) Let $m \geqslant 1$ be an integer, and let $\varrho\colon \mathrm{S}_n \to \mathrm{GL}_m(\mathbf{C})$ be a group morphism which is non-trivial when restricted to A_n. Show that $\varrho(\sigma)$ is not the identity.

(3) Show that if k is an integer coprime to n, then σ^k is an n-cycle.

(4) Deduce that $m \geqslant \varphi(n)$. (Hint: adapt the method used for $\mathrm{SL}_2(\mathbf{F}_p)$.)

(5) Let $m \geqslant 1$ be an integer, and let $\varrho\colon \mathrm{A}_n \to \mathrm{GL}_m(\mathbf{C})$ be a non-trivial group morphism. Show that $m \geqslant \varphi(n)/2$. (Hint: if you know about induced representations, consider the representation $\mathrm{Ind}_{\mathrm{A}_n}^{\mathrm{S}_n}(\varrho)$ of S_n, and apply the previous

[9] One can check numerically that A_5 is only 3-quasirandom.

result; otherwise, show that $\varphi(n)/2$ distinct powers of σ are conjugate to σ in A_n, and argue as before.)

The point of the definition of α-quasirandom groups is that it implies strong properties of product sets if α is relatively large.

THEOREM 2.8.4 (Gowers). *Let $\alpha \geqslant 1$ be a real number and G an α-quasirandom group. Let $k \geqslant 1$ be an integer.*

For any subsets $\mathrm{A}_1, \ldots, \mathrm{A}_k$ of G such that

$$|\mathrm{A}_1| \cdots |\mathrm{A}_k| > \frac{|\mathrm{G}|^k}{\alpha^{k-1}},$$

we have

$$\mathrm{A}_1 \cdots \mathrm{A}_k = \mathrm{G}.$$

In particular, for any subsets A, B and C of G such that $|\mathrm{A}||\mathrm{B}||\mathrm{C}| \geqslant \alpha^{-2}|\mathrm{G}|^3$, we have $\mathrm{A} \cdot \mathrm{B} \cap \mathrm{C} \neq \emptyset$.

PROOF OF THEOREM 2.8.4. We note first that the second statement follows from the first by taking $k = 3$ and $\mathrm{A}_1 = \mathrm{A}$, $\mathrm{A}_2 = \mathrm{B}$, $\mathrm{A}_3 = \mathrm{C}^{-1}$; we deduce that $\mathrm{A} \cdot \mathrm{B} \cdot \mathrm{C}^{-1} = \mathrm{G}$, and in particular writing $abc^{-1} = 1$ for suitable $(a, b, c) \in \mathrm{A} \times \mathrm{B} \times \mathrm{C}$, we get $ab = c \in \mathrm{A} \cdot \mathrm{B} \cap \mathrm{C}$.

The proof of the first statement is a variant of the proof of Gowers, due to Breuillard [13, Lemma 2.2]. It is, in principle, a straightforward implementation of the basic idea of harmonic analysis, already sketched in Section 1.3: we expand in a suitable "basis" the representation function of the product set, and manipulate the resulting expression to show that it is positive. The trick is to see how to use the quasi-randomness assumption.

Let φ_i denote the characteristic function of the set A_i, and let $r \colon \mathrm{G} \to \mathbf{C}$ be the function defined by

$$r = \varphi_1 * \cdots * \varphi_k,$$

which is a positive multiple of the representation function for $\mathrm{A}_1 \cdots \mathrm{A}_k$. It suffices to prove that $r(x) > 0$ for all $x \in \mathrm{G}$. To do this, we use Fourier analysis on G to expand r in the form

$$r = \sum_{\varrho \in \widehat{\mathrm{G}}} r_\varrho$$

as in Section A.8 (see especially Example A.8.7).

Isolating the contribution of the trivial representation, denoted ϱ_0 in Theorem A.8.5, we get

$$r(x) \geqslant \frac{|\mathrm{A}_1| \cdots |\mathrm{A}_k|}{|\mathrm{G}|^k} - \sum_{\varrho \neq 1} |r_\varrho(x)|,$$

since r_{ϱ_0} is a constant function equal to the average of r over G.

The functions r_ϱ can be identified with linear maps on a finite-dimensional Hilbert space E_ϱ, and the space of these linear maps is a Hilbert space, with a norm denoted $u \mapsto \|u\|_\varrho$. We know that for any $x \in \mathrm{G}$, we have

$$|r_\varrho(x)| \leqslant \|r_\varrho\|_\varrho.$$

By the definition of r using convolution and the fact that the ϱ-component of a convolution is given by composition of the corresponding linear maps, we have

$$r_\varrho = \varphi_{1,\varrho} \circ \cdots \circ \varphi_{k,\varrho}$$

for any $\varrho \in \widehat{\mathrm{G}}$. The inequality above becomes

$$r(x) \geqslant \frac{|\mathrm{A}_1| \cdots |\mathrm{A}_k|}{|\mathrm{G}|^k} - \sum_{\varrho \neq 1} \|\varphi_{1,\varrho} \circ \cdots \circ \varphi_{k,\varrho}\|_\varrho. \tag{2.29}$$

The multiplicativity of the norm $\|u\|_\varrho$ (Proposition A.8.9, (2), gives an upper-bound

$$\|\varphi_{1,\varrho} \circ \cdots \circ \varphi_{k,\varrho}\|_\varrho \leqslant \|\varphi_{1,\varrho}\| \cdots \|\varphi_{k-1,\varrho}\| \, \|\varphi_{k,\varrho}\|_\varrho$$

for every ϱ, where $\|u\|$ is the usual norm for linear operators on E_ϱ (note that only one of the functions is measured with the ϱ-norm).

Now quasi-randomness can be used: for $\varrho \neq \varrho_0$ and any linear map u on E_ϱ, we have

$$\|u\| \leqslant \alpha^{-1/2} \|u\|_\varrho,$$

by Proposition A.8.9, (1),[10] and consequently, we get

$$\|\varphi_{1,\varrho} \circ \cdots \circ \varphi_{k,\varrho}\|_\varrho \leqslant \alpha^{-(k-1)/2} \|\varphi_{1,\varrho}\|_\varrho \cdots \|\varphi_{k-1,\varrho}\|_\varrho \, \|\varphi_{k,\varrho}\|_\varrho.$$

If $u = \varphi_{i,\varrho}$, then we argue from the Parseval identity

$$\sum_\varrho \|\varphi_{i,\varrho}\|_\varrho^2 = \|\varphi_i\|^2 = \frac{|\mathrm{A}_i|}{|\mathrm{G}|}$$

and positivity that

$$\|\varphi_{i,\varrho}\|_\varrho \leqslant \Big(\frac{|\mathrm{A}_i|}{|\mathrm{G}|}\Big)^{1/2}.$$

We use this for the first $k-2$ among the terms in the product of ϱ-norms, and obtain

$$\|\varphi_{1,\varrho} \circ \cdots \circ \varphi_{k,\varrho}\|_\varrho \leqslant \alpha^{-(k-1)/2} \Big(\frac{|\mathrm{A}_1| \cdots |\mathrm{A}_{k-2}|}{|\mathrm{G}|^{k-2}}\Big)^{1/2} \|\varphi_{k-1,\varrho}\|_\varrho \, \|\varphi_{k,\varrho}\|_\varrho$$

for $\varrho \neq \varrho_0$.

Furthermore, using the Cauchy–Schwarz inequality and the Parseval identity again, we obtain the estimate

$$\begin{aligned}
\sum_{\varrho \neq \varrho_0} \|\varphi_{1,\varrho} \circ \cdots \circ \varphi_{k,\varrho}\|_\varrho &\leqslant \alpha^{-(k-1)/2} \Big(\frac{|\mathrm{A}_1| \cdots |\mathrm{A}_{k-2}|}{|\mathrm{G}|^{k-2}}\Big)^{1/2} \sum_{\varrho \neq \varrho_0} \|\varphi_{k-1,\varrho}\|_\varrho \, \|\varphi_{k,\varrho}\|_\varrho \\
&\leqslant \alpha^{-(k-1)/2} \Big(\frac{|\mathrm{A}_1| \cdots |\mathrm{A}_{k-2}|}{|\mathrm{G}|^{k-2}}\Big)^{1/2} \|\varphi_{k-1}\| \, \|\varphi_k\| \\
&= \alpha^{-(k-1)/2} \Big(\frac{|\mathrm{A}_1| \cdots |\mathrm{A}_k|}{|\mathrm{G}|^k}\Big)^{1/2}. \qquad (2.30)
\end{aligned}$$

It follows finally by combining (2.30) with (2.29) that $r(x) > 0$ for all x if

$$|\mathrm{A}_1| \cdots |\mathrm{A}_k| > \frac{|\mathrm{G}|^k}{\alpha^{k-1}}.$$

□

Remark 2.8.5. (1) Theorem 2.8.4 is not useful if $\alpha = 1$, since the assumption can then only hold when $\mathrm{A} = \mathrm{B} = \mathrm{C} = \mathrm{G}$, but it becomes very interesting as soon as α grows. In particular, for $\mathrm{SL}_2(\mathbf{F}_p)$ with p an odd prime, Example 2.8.2, (2) combined with Theorem 2.8.4, shows that if a subset A of $\mathrm{SL}_2(\mathbf{F}_p)$ satisfies $|\mathrm{A}| \geqslant (\frac{1}{2}(p-1))^{-1/3} |\mathrm{SL}_2(\mathbf{F}_p)|$,

[10] Here is the only place where we also use the interpretation of the Fourier decomposition in terms of linear actions to say that any $\varrho \neq \varrho_0$ must have degree at least α.

which is of size roughly $2^{1/3}p^{8/9}$, then $A^{(3)} = \mathrm{SL}_2(\mathbf{F}_p)$. The analogue of this fails completely for abelian groups: for instance, if $G = \mathbf{Z}/p\mathbf{Z}$, then the set A which is the image modulo p of an interval of length $< p/3$ in $\mathbf{Z}$ *never* satisfies $A + A + A = \mathbf{Z}/p\mathbf{Z}$.

(2) One can think of this result in analogy with the circle method of diophantine geometry, which is also based on the use of harmonic analysis to compute the number of solutions of various additive equations. It is interesting to note that, just as in that case, the most direct method breaks down for "binary problems", i.e., in that case for a product $A_1 \cdot A_2$: it leads to a condition on the size of A_1 and A_2 which can only be satisfied in the trivial case where each is equal to G.

Gowers deduced from Theorem 2.8.4 and from the existence of α-quasirandom groups with α arbitrarily large the answer to a question of Babai and Sós [1, Problem 7.5]: a finite group G does not always contain a product-free subset of density bounded away from 0 in G. Precisely, a subset A of an arbitrary group G is called *product-free* if $A^{(2)} \cap A$ is empty (i.e., no element of A is also a product of two, possibly equal, elements of A) and the question is whether there exists $c > 0$ such that any finite group G contains a product-free subset of size at least $c|G|$.

The answer is "No": indeed, if G is α-quasirandom, then by taking $A = B = C$ in Theorem 2.8.4, (1), we see that a subset of size at least $\alpha^{-1/3}|G|$ of G is not *product-free*. Taking a sequence of groups G_n which are α_n-quasirandom with $\alpha_n \to +\infty$, the result follows.

There are also some form of converse statements. For instance, the following exercise presents a well-known result of Erdős, according to which any non-empty set A of integers contains a rather large product-free subset. Because of the additive notation, these are also called sum-free sets.

Exercise 2.8.6. (1) Let $p \geqslant 5$ be a prime number. Denote by I the image modulo p of the set of positive integers n such that $p/3 < n \leqslant 2p/3$. Let $A \subset \mathbf{F}_p$ be a finite set. For any $x \in \mathbf{F}_p^\times$, show that the set $B_x = A \cap xI$ is a sum-free set, where xI denotes here the set of elements of the form xy with $y \in I$.

(2) We now consider B_x when x is taken uniformly at random in $\mathbf{F}_p^\times$. Show that $\mathbf{E}(|B_x|) > |A|/3$, i.e., that

$$\frac{1}{p-1}\sum_{x\in\mathbf{F}_p^\times} |B_x| > \frac{|A|}{3}.$$

(3) Deduce that if $A \subset \mathbf{Z}$ is a set of integers, then it contains a sum-free subset B with $|B| > |A|/3$.

Exercise 2.8.7. (1) Let G be a group and H a subgroup of G. For any $x \in G \smallsetminus H$, show that the cosets xH and Hx are product-free sets.

(2) Deduce that the symmetric group S_n contains a product-free set of size $(n-1)!$ for any $n \geqslant 2$.

(3) Try to find a product-free subset of $\mathrm{SL}_2(\mathbf{F}_p)$ which is as large as possible.

2.9. The Freiman–Ruzsa Theorem

The first truly significant study of what are now called approximate subgroups is to be found in the work of Freiman (see his book [35]). In particular, what is now called the Freiman–Green–Ruzsa Theorem provides a basic qualitative description of approximate subgroups in abelian groups. In the case of torsion-free groups (i.e., groups G where the

equation $dx = 0$, with $d \geqslant 1$ and $x \in \mathrm{G}$, implies that $x = 0$), it takes the following form (which was how Ruzsa [71, Th. 1.1] stated his version of the result):

THEOREM 2.9.1 (Freiman, Ruzsa). *Let $\alpha \geqslant 1$ be a real number. Let A be a finite subset of a torsion-free abelian group G such that $|\mathrm{A}+\mathrm{A}| \leqslant \alpha|\mathrm{A}|$. There exist an integer $d \geqslant 1$ and a real number $\beta \geqslant 0$, both depending only on α, such that A is contained in a d-dimensional generalized arithmetic progression B with $|\mathrm{B}| \leqslant \beta|\mathrm{A}|$.*

One of the most challenging questions in additive combinatorics is the *polynomial Freiman–Ruzsa conjecture*, which concerns the dependency of d and β on α in theorems of this type – in particular, d should be bounded by a polynomial in α, see [84, Conj. 5.43] for a precise statement.

We sketch Ruzsa's proof, since it involves important ideas. The key step that extracts some structure from A is the following general result, due to Bogolyubov:

PROPOSITION 2.9.2. *Let G be a finite abelian group. Let A be a non-empty subset of G. There exists an integer $d \leqslant (|\mathrm{G}|/|\mathrm{A}|)^2$ and characters*

$$\chi_1, \dots, \chi_d$$

of G such that the set $2\mathrm{A} - 2\mathrm{A}$ contains the set

$$\mathrm{B} = \{x \in \mathrm{G} \mid \mathrm{Re}(\chi_i(x)) \geqslant 0 \text{ for } 1 \leqslant i \leqslant d\}.$$

PROOF. Let φ_{A} be the characteristic function of A, and $\varphi_{-\mathrm{A}}$ that of $-\mathrm{A}$. By the definition of convolution, the function

$$r_{\mathrm{A}} = \varphi_{\mathrm{A}} * \varphi_{\mathrm{A}} * \varphi_{-\mathrm{A}} * \varphi_{-\mathrm{A}}$$

is proportional to the representation function for the sumset $2\mathrm{A} - 2\mathrm{A}$, and in particular is positive exactly on this subset.

We will find a lower bound for the value of r_{A} by expressing it on the Fourier side. We have $\widehat{r}_{\mathrm{A}} = |\mathrm{G}|^{-3/2}|\widehat{\varphi}_{\mathrm{A}}|^4$ (see Remark A.7.9, (1)) since the Fourier transform of $\varphi_{-\mathrm{A}}$ is $\overline{\widehat{\varphi}}_{\mathrm{A}}$. Hence, by inverse Fourier transform, we know that

$$r_{\mathrm{A}}(x) = \mathrm{Re}(r_{\mathrm{A}}(x)) = \mathrm{Re}\Big(\frac{1}{|\mathrm{G}|^2}\sum_{\chi\in\widehat{\mathrm{G}}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^4\chi(x)\Big).$$

We split the sum according to the size of the Fourier transform of φ_{A}, using a parameter β that we will specify later, with $0 < \beta < 1$. Note that

$$|\widehat{\varphi}_{\mathrm{A}}(\chi)| = \Big|\frac{1}{|\mathrm{G}|^{1/2}}\sum_{x\in\mathrm{A}}\chi(x)\Big| \leqslant \frac{|\mathrm{A}|}{|\mathrm{G}|^{1/2}},$$

and denote by X the set of $\chi \in \widehat{\mathrm{G}}$ such that $|\widehat{\varphi}_{\mathrm{A}}(\chi)| > \beta|\mathrm{A}||\mathrm{G}|^{-1/2}$. We have

$$\sum_{\chi\in\widehat{\mathrm{G}}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^4\chi(x) = \sum_{\chi\in\mathrm{X}} \widehat{\varphi}_{\mathrm{A}}(\chi)|^4\chi(x) + \sum_{\chi\notin\mathrm{X}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^4\chi(x),$$

hence

$$|\mathrm{G}|^2 r_{\mathrm{A}}(x) > \mathrm{Re}\Big(\sum_{\chi\in\mathrm{X}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^4\chi(x)\Big) - \sum_{\chi\notin\mathrm{X}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^4.$$

By definition of X, the second term satisfies the bound

$$\sum_{\chi\notin\mathrm{X}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^4 \leqslant \frac{\beta^2|\mathrm{A}|^2}{|\mathrm{G}|}\sum_{\chi\notin\mathrm{X}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^2 \leqslant \frac{\beta^2|\mathrm{A}|^2}{|\mathrm{G}|}\sum_{\chi\in\widehat{\mathrm{G}}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^2 = \frac{\beta^2|\mathrm{A}|^3}{|\mathrm{G}|},$$

where we used the discrete Plancherel formula in the end.

The set X contains the trivial character, which gives a contribution equal to $|\widehat{\varphi}_{\mathrm{A}}(1)|^4 = |\mathrm{A}|^4/|\mathrm{G}|^2$. Hence, if x satisfies $\mathrm{Re}(\chi(x)) \geqslant 0$ for $\chi \in \mathrm{X} - \{1\}$, then

$$|\mathrm{G}|^2 r_{\mathrm{A}}(x) > \frac{|\mathrm{A}|^4}{|\mathrm{G}|^2} - \frac{\beta^2|\mathrm{A}|^3}{|\mathrm{G}|}.$$

Thus, taking $\beta = (|\mathrm{A}|/|\mathrm{G}|)^{1/2}$, we ensure that the right-hand side is positive, and we conclude that $2\mathrm{A} - 2\mathrm{A}$ contains the set

$$\mathrm{B} = \{x \in \mathrm{G} \mid \mathrm{Re}(\chi(x)) \geqslant 0 \text{ for } \chi \in \mathrm{X} - \{1\}\}.$$

This set is of the form claimed, with

$$d = |\mathrm{X}| - 1 \leqslant \Big(\frac{|\mathrm{G}|}{|\mathrm{A}|}\Big)^2,$$

since

$$|\mathrm{X}|\frac{|\mathrm{A}|^3}{|\mathrm{G}|^2} = |\mathrm{X}|\frac{\beta^2|\mathrm{A}|^2}{|\mathrm{G}|} \leqslant \sum_{\chi\in\mathrm{X}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^2 \leqslant \sum_{\chi\in\widehat{\mathrm{G}}} |\widehat{\varphi}_{\mathrm{A}}(\chi)|^2 = |\mathrm{A}|,$$

by the Plancherel formula once more. □

Sets of the kind described in this lemma, with the conditions $\mathrm{Re}(\chi_i(x)) \geqslant 0$ replaced by the more general conditions that $\chi_i(x)$ should be of the form $e(t) = e^{2i\pi t}$ with $|t| \leqslant \varepsilon_i$ for some parameters $\varepsilon_i > 0$,[11] are called *Bohr sets*, and appear frequently in harmonic analysis as well as additive combinatorics. Using the function $\|x\|$ on the unit circle which gives the angular distance (measured from 0 to 1, i.e., identifying the circle with $\mathbf{R}/\mathbf{Z}$) to the point 1, we can express these sets in the form

$$\{x \in \mathrm{G} \mid \|\chi_i(x)\| < \varepsilon_i \text{ for } 1 \leqslant i \leqslant d\}.$$

Ruzsa combined Proposition 2.9.2 with another structural result about Bohr sets.

PROPOSITION 2.9.3 (Ruzsa). *Let $d \geqslant 1$ be an integer, and let $(\varepsilon_1, \dots, \varepsilon_d)$ be positive real numbers with $\varepsilon_i < \frac{1}{2}$ for all i. Let $\mathrm{N} \geqslant 1$ be an integer, and denote by χ_a the character $x \mapsto e(ax/\mathrm{N})$ of $\mathbf{Z}/\mathrm{N}\mathbf{Z}$.*

For any integers $(a_1, \dots, a_d)$ such that the gcd of $(a_1, \dots, a_d, \mathrm{N})$ is equal to 1, the Bohr set

$$\{x \in \mathbf{Z}/\mathrm{N}\mathbf{Z} \mid \|\chi_{a_i}(x)\| < \varepsilon_i \text{ for } 1 \leqslant i \leqslant d\}$$

in $\mathbf{Z}/\mathrm{N}\mathbf{Z}$ contains a proper generalized arithmetic progression of dimension d and size $> \delta\mathrm{N}$, where

$$\delta = d^{-d}\varepsilon_1 \cdots \varepsilon_d.$$

2.10. Final remarks

We have omitted many topics related to sum (or product) sets. Here are a few examples, with some references:

(1) A substantial part of analytic number theory concerns problems which can be interpreted as asking about properties of sumsets $\mathrm{A}_1 + \cdots + \mathrm{A}_k$ for interesting concrete sets of integers A_i and various values of k. For instance:

[11] The case of Bogolyubov corresponds to taking $\varepsilon_i = \frac{1}{4}$.

– A famous theorem of Lagrange states that $4\square$ is the set of all positive integers, where $\square$ denotes the set of squares of positive integers. Another famous theorem of Gauss identifies exactly the set $3\square$, and shows in particular that it contains an some arithmetic progressions modulo 8.

– Waring's Problem, as solved by Hilbert, can be interpreted as the statement that for any integer $d \geqslant 1$, some iterated sumset of the set of d-th powers of positive integers is equal to the set of positive integers.

– Let $\mathbf{P}$ be the set of primes. Vinogradov proved that $3\mathbf{P}$ contains all sufficiently large odd integers, and it is famously conjectured that $2\mathbf{P}$ contains all even integers $\geqslant 4$.

(2) Remarkably, there is an abstract theory of sumsets of integers which suffices to solve Waring's Problem and to prove that $k\mathbf{P}$ is contains all but finitely many integers for some $k \geqslant 1$: this is the topic of Schnirelman's density (see, e.g., [55, Ch. II]).

CHAPTER 3

The sum-product phenomenon

3.1. Sum-product in integers

The main theme of this chapter is the *sum-product phenomenon*, which was already mentioned in Theorem 1.1.3. Recall the statement: for any finite set A of positive integers, we have

$$\max(|A+A|, |A\cdot A|) \geqslant c|A|^{1+\delta}, \tag{3.1}$$

for some real numbers $c > 0$ and $\delta > 0$, independent of A. Intuitively, this is interpreted as saying that $\mathbf{Z}$ does not contain (finite) subsets which behave like "approximate subrings".

Remark 3.1.1. Note that

$$\begin{aligned}\max(|A+A|, |A\cdot A|) &\leqslant |(A+A)\cup A\cdot A| \\ &\leqslant |A+A|+|A\cdot A| \leqslant 2\max(|A+A|, |A\cdot A|),\end{aligned}$$

hence (up to changing the value of c), the bound (3.1) is equivalent to either

$$|A+A|+|A\cdot A| \geqslant c|A|^{1+\delta}$$

or

$$|(|A+A)\cup A\cdot A| \geqslant c|A|^{1+\delta}. \tag{3.2}$$

To get a first feeling for this result, it is of course useful to check what happens for sets A for which we already know that either $A+A$ or $A\cdot A$ is not much larger than A.

(1) Suppose first that A has minimal growth under multiplication. This means that $|A\cdot A| = 2|A|-1$ (we use here the minor assumption that all elements of A are positive, so we can view A as a subset of the abelian group $\mathbf{R}_+^\times$ of positive real numbers, which is isomorphic to $\mathbf{R}$ by the logarithm and use the very easy form of Cauchy's Theorem in $\mathbf{R}$, see Section 1.3); from the inverse statement (Proposition 1.3.7), it follows that A is a geometric progression. Thus there exists $a_0 \in A$ (in particular non-zero) and $r \geqslant 2$ such that

$$A = \{a_0, a_0 r, \dots, a_0 r^{|A|-1}\}.$$

We recognize that A is a Sidon set in $\mathbf{Z}$, so $A+A$ is of size $|A|^2$, so that (3.1) is confirmed with the best possible value $\delta = 1$ in that case.

(2) The case where A has minimal growth under addition is more curious. By Proposition 1.3.7 again, this means that A is an arithmetic progression; we consider for simplicity the case $A = [N]$ for some integer $N \geqslant 1$. The question of the size of $A\cdot A$ is the "multiplication table problem" of Erdős: how many distinct integers appear in an $N\times N$ multiplication table? The form of the answer is quite surprising: after much earlier work (of Erdős, Hall, Tenenbaum, in particular) it was proved by Ford [31, Cor. 3] that

$$|[N]\cdot[N]| \asymp \frac{N^2}{(\log N)^\beta (\log\log N)^{3/2}}, \qquad \beta = 1 - \frac{1+\log\log 2}{\log 2} = 0.086071\dots$$

for all $N \geqslant 1$.

In particular, although we obtain the bound (3.1) for any $\delta < 1$, it *fails* when $\delta = 1$. One can check these facts more easily than by appealing to the deep result of Ford. For the lower bound, note that $[N]$ contains the set of primes $p \leqslant N$ as a Sidon subset, so that $|[N] \times [N]| \geqslant \frac{1}{2}\pi(N)^2$, where $\pi(N)$ is the number of primes $\leqslant N$. By Chebychev's estimate already (see, e.g., [78, § 2.3] or [48, Th. 414] for a proof), this implies that $|[N] \cdot [N] \gg N^2/(\log N)^2$.

Concerning the upper bound (showing that $\delta = 1$ is not possible, as had already been observed by Erdős and Szemerédit, in stronger form), one can exploit a result of Hardy and Ramanujan, which states that the number of prime factors of a typical integer $n \in [N]$ is usually about $\log\log N$, and hence the number of prime factors of an element of $[N] \cdot [N]$ is about $2\log\log N$, which means that these integers *cannot* be typical as elements of $[N^2]$.

More precisely, we denote by $\Omega(n)$ the number of prime divisors of an integer $n \geqslant 1$, counted with multiplicity; its crucial property is that $\Omega(ab) = \Omega(a)+\Omega(b)$ for any positive integers a and b. Hardy and Ramanujan proved that

$$\sum_{n\in[N]} (\Omega(n) - \log\log N)^2 \ll N\log\log N, \tag{3.3}$$

for $N \geqslant 3$, which is a variance estimate for $\Omega(n)$ (see, e.g., [48, Th. 431]). It follows that

$$|\{n \in [N] \mid |\Omega(n) - \log\log N| \geqslant \tfrac{1}{4}(\log\log N)^{1/2}\}| \ll \frac{N}{\log\log N},$$

(this is Markov's inequality). On the one hand, this implies that

$$|\{n \in [N^2] \mid \Omega(n) \geqslant \tfrac{3}{2}\log\log N\}| \ll \frac{N^2}{\log\log N},$$

giving the typical value of $\Omega(n)$ for $n \in [N^2]$ (noting that $\log(\log(N^2)) = \log(\log N) + \log(\log 2)$).

On the other hand, an element n of $[N]\times[N]$ which does *not* satisfy $\Omega(n) \geqslant \frac{3}{2}\log\log N$ must be of the form $n = ab$ for some $(a, b) \in [N] \times [N]$ with either

$$\Omega(a) < \tfrac{1}{4}\log\log N, \quad \text{or} \quad \Omega(b) < \tfrac{1}{4}\log\log N.$$

The total number of (a, b) with this property is bounded by

$$2N\,|\{k \in [N] \mid \Omega(k) < \tfrac{1}{4}\log\log N\}| \ll \frac{N^2}{\log\log N}.$$

It follows that

$$|[N] \times [N]| \leqslant |\{n \in [N^2] \mid \Omega(n) \geqslant \tfrac{3}{2}\log\log N\}| + O\Big(\frac{N^2}{\log\log N}\Big) \ll \frac{N^2}{\log\log N}.$$

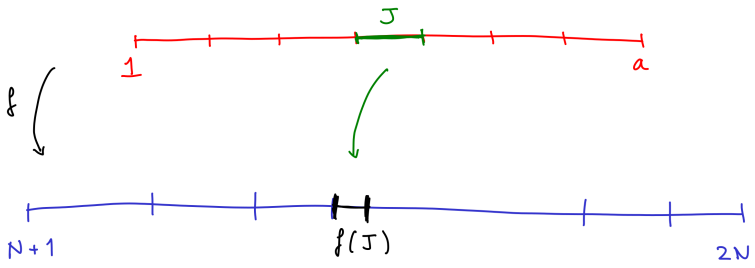
Remark 3.1.2. The inequality (3.3) is only the tip of an iceberg: the Erdős–Kác Theorem (see, e.g., [60, § 2.3] for a proof) states that the normalized quantity

$$\frac{\Omega(n) - \log\log N}{\sqrt{\log\log N}},$$

for $n \in [N]$, becomes distributed, as $N \to +\infty$, like a standard gaussian random variable.

Remark 3.1.3. Erdős and Szemerédi conjectured that the sum-product estimate (3.1) should be valid for any $\delta < 1$.

We will give different proofs of Theorem 1.1.3. We begin by the original argument, which will lead to the respectable value $\delta = \frac{1}{41}$ (without much effort at optimization).



The Erdős–Szemerédi proof. We will prove the statement in the form of the inequality (3.2).

We begin by considering the case where the set A is contained in a dyadic interval $N + [N]$, which we may also assume satisfies $N \geqslant 3$.

The first easy consequence of this reduction is that the sum set $A+A$ and the product set $A \cdot A$ are disjoint (all elements of the former are $\leqslant 4N$, all those of the latter are $\geqslant (N+1)^2 > 4N$).

Let $a = |A|$, and consider the unique strictly increasing function $f\colon [a] \to A$. For a parameter $q < a$ to be chosen later, subdivide $[a]$ in $q+1$ successive intervals $I_1, \dots, I_{q+1}$, all but the latest of the same size $\ell = \lfloor a/q \rfloor$. Finally, let I be one interval among $I_1, \dots, I_q$ chosen so that $\max f(I) - \min f(I)$ is as small as possible. (We will simply ignore the contribution of the elements of A parameterized by the last interval.)

For $1 \leqslant j \leqslant q$, let

$$X_j = (f(I) + f(I_j)) \cup (f(I) \cdot f(I_j)),$$

a subset of $(A + A) \cup (A \cdot A)$.

Step 1. The first observation is that if we only consider integers $j \equiv 1 \pmod 3$, with $I_j \neq I$, then the sets X_j are pairwise disjoint. Indeed, since $(A+A) \cap A \cdot A$ is empty, it suffices to show that (under these conditions on j and k), the equations

$$a + c = b + d$$

and

$$ac = bd$$

have no solutions with $(a, b) \in f(I)^2$ and $(c, d) \in f(I_j) \times f(I_k)$ if $j \neq k$.

For the former, this is because we would deduce that $a - b = d - c$; however, since $|j - k| \geqslant 2$ by assumption, there is at least one full interval, say J, between I_j and I_k. Since f is injective, we get

$$|d - c| > \max(f(J)) - \min(f(J)) \geqslant \max f(I) - \min f(I),$$

by definition of I, whereas $|a - b| \leqslant \max f(I) - \min f(I)$.

For the second equation, we may assume that $j < k$, in particular $c < d$; the equation $ac = bd$ then imposes that $b < a$. Write $b = a - u$ and $d = c + v$ for some positive integers u and v. Observe that since $k \geqslant j + 3$ by assumption, the choice of I implies that $v > 2u$.

The equation $ac = bd$ is equivalent to $bd = (b + u)(d - v)$, i.e., to $uv = du - bv$. But this is impossible, since $uv > 0$ whereas

$$du - bv \leqslant 2bu - bv = b(u - 2v) \leqslant 0$$

(we used the fact that A is contained in a dyadic interval to ensure that $d \leqslant 2b$).

Step 2. Let $\gamma > 0$ be a (small) constant, and let $\mathscr{I}$ be the set of $j \leqslant q$ congruent to 1 modulo 3, with $\mathrm{I}_j \neq \mathrm{I}$, such that $|\mathrm{X}_j| < \ell^{1+\gamma}$. We claim that if $\gamma < 1/3$, then for any $j \in \mathscr{I}$, the system of equations

$$\begin{cases} x_1 + x_3 = x_2 + x_4 \\ x_1 x_5 = x_2 x_6 \end{cases} \tag{3.4}$$

with $(x_1, x_2) \in f(\mathrm{I}_j)^2$ and $(x_3, x_4, x_5, x_6) \in f(\mathrm{I})^4$ has at least one solution which is non-trivial in the sense that $x_3 \neq x_4$.

Indeed, by assumption on j, the set $f(\mathrm{I}_j) \cdot f(\mathrm{I}) \subset \mathrm{X}_j$ has size $< \ell^{1+\gamma}$, and by the pigeonhole principle, there exists $n \in f(\mathrm{I}_j) \cdot f(\mathrm{I})$ such that $|\mathrm{R}(n)| \geqslant \ell^{1-\gamma}$, where $\mathrm{R} \subset f(\mathrm{I}_j) \times f(\mathrm{I})$ is the representation set for $f(\mathrm{I}_j) \cdot f(\mathrm{I})$.

Since we assumed that $\gamma < 1/3$, we have $\ell^{2-2\gamma} > \ell^{1+\gamma}$, and therefore $|\mathrm{R}(n)|^2 > \ell^{1+\gamma} > |\mathrm{X}_j|$; thus the map

$$\begin{cases} \mathrm{R}(n) \times \mathrm{R}(n) \to f(\mathrm{I}_j) + f(\mathrm{I}) \subset \mathrm{X}_j \\ ((a, a'), (b, b')) \mapsto a + b' \end{cases}$$

is not injective, and there must exist elements

$$(a, a'), \quad (c', c), \quad (b, b'), \quad (d', d)$$

of $\mathrm{R}(n)$ such that $a + c = b + d$ and

$$(a, a', c', c) \neq (b, b', d', d).$$

Since furthermore $aa' = bb' = n$, we obtain a solution

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (a, b, c, d, a', b')$$

to the system (3.4), and it does satisfy $x_3 = c \neq d = x_4$, as desired (because otherwise we would get $a = b$, and then $b' = a'$ and $c' = d'$ from $aa' = bb' = n = cc' = dd'$).

Step 3. We can now conclude: still under the condition $\gamma < 1/3$, the previous step allows us to define a map

$$\phi \colon \mathscr{I} \to f(\mathrm{I})^4$$

by sending $j \in \mathscr{I}$ to one quadruple (c, d, e, f) with $c \neq d$ for which the system (3.4) has a solution (a, b, c, d, e, f) in $f(\mathrm{I}_j)^2 \times f(\mathrm{I})^4$.

We claim that the map ϕ is *injective*. Indeed, given any non-zero integers (c, d, e, f) with $c \neq d$, there is at most one pair $(a, b) \in \mathbf{Q}^2$ for which the equation

$$\begin{cases} a + c = b + d \\ ae = bf \end{cases}$$

has a solution, namely $(a, b) = (f(c-d)/(e-f), e(c-d)/(e-f)$ (which is well-defined because the existence of a solution with $c \neq d$ implies that $e \neq f$). In our situation, given $j \in \mathscr{I}$, we can recover j from the tuple $(c, d, e, f) = \phi(j)$ because the corresponding pair (a, b) exists in $f(\mathrm{I}_j)$, and this determines uniquely the value of j.

Thus, we have $|\mathscr{I}| \leqslant |f(\mathrm{I})|^4 \leqslant \ell^4$. It follows that, as soon as $\ell^4 < \frac{q}{6}$, there are at least $\frac{q}{6}$ integers j with $j \equiv 1 \pmod 3$ and $|\mathrm{X}_j| \geqslant \ell^{1+\gamma}$, and therefore

$$|(\mathrm{A} + \mathrm{A}) \cup \mathrm{A} \cdot \mathrm{A}| \geqslant \sum_{j \equiv 1 \,(\mathrm{mod}\, 3)} |\mathrm{X}_j| \geqslant \frac{1}{2} \times \frac{q}{6} \times \ell^{1+\gamma} \gg a^{1+\gamma} q^{-\gamma},$$

which is clearly suitable provided $q \sim a^\beta$ with $0 < \beta < 1$. To get an explicit exponent, we take for instance $\gamma = 1/4$ and $q = \lfloor 7a^{4/5} \rfloor$, and deduce

$$|(A + A) \cup A \cdot A| \gg |A|^{1+1/20}.$$

This concludes the proof of the result when A is in a dyadic interval. We now establish the general case. For $j \geqslant 0$, let $A_j = A \cap (2^j - 1 + [2^j])$, so that A is the union of the subsets A_j. We can always apply the first case to each subset A_j, but there is a potential difficulty: these sets might be very small, so that the growth given by the dyadic case is not enough to deduce anything significant concerning A. Indeed, A might be so "lacunary" that each A_j contains a single element. This however gives a clue how to proceed: in this extreme case, A would be a Sidon set by Example 2.3.5, so growth under sum or product would arise from all the A_j, instead of from individual pieces.

To make this idea precise, we pick again a parameter $\beta > 0$ to be determined later. Let $\mathscr{J}$ be the set of integers j such that $1 \leqslant |A_j| < |A|^\beta$, and let A' be the union of A_j for $j \in \mathscr{J}$. We have two cases to consider.

First, if $|A'| \leqslant \frac{1}{2}|A|$ (so that many sets A_j are pretty large), then applying the first case of the result to the sets A_j with $j \notin \mathscr{J}$, we obtain the lower bound

$$|(A_j + A_j) \cup A_j \cdot A_j| \gg |A_j|^{1+\gamma} \gg |A|^{\beta\gamma}|A_j|,$$

where $\gamma = 1/20$. Furthermore, the sets A_j are contained in the disjoint dyadic intervals

$$D_j = \{2^j, \dots, 2^{j+1} - 1\}, \qquad j \geqslant 0,$$

which satisfy

$$D_j + D_j = \{2^{j+1}, \dots, 2^{j+2} - 1\}, \quad D_j \cdot D_j \subset \{2^{2j}, \dots, 2^{2j+2} - 1\},$$

and we see that a given integer $n \in (A + A) \cup A \cdot A$ belongs to $(A_j + A_j) \cup A_j \cdot A_j$ for at most three values of j. Therefore

$$|(A + A) \cup A \cdot A| \geqslant \frac{1}{3} \sum_{j \notin \mathscr{J}} |(A_j + A_j) \cup A_j \cdot A_j| \gg |A|^{\beta\gamma} \sum_{j \notin \mathscr{J}} |A_j| = |A|^{1+\beta\gamma}.$$

Finally, if $|A'| > \frac{1}{2}|A|$ (many sets A_j are small), then by construction the size of $\mathscr{J}$ must be $\geqslant \frac{1}{2}|A|^{1-\beta}$. Let $A'' \subset A$ be a set formed by taking one element from each set A_j with $j \in \mathscr{J}$, j even; thus $|A''| \geqslant \frac{1}{4}|A|^{1-\beta}$. Since the A_j's are in disjoint dyadic intervals, it follows that A'' is a Sidon set in $\mathbf{Z}$ (see Example 2.3.5), and thus

$$|(A + A) \cup A \cdot A| \geqslant |A'' + A''| \gg |A|^{2-2\beta}.$$

As soon as $2\beta < 1$, the combination of these two lower bounds implies the sum-product theorem. To get the best possible exponent, we choose β so that the two cases give the same exponent, i.e. so that

$$1 - 2\beta = \beta\gamma, \qquad \gamma = \frac{1}{20},$$

thus obtaining (3.1) with $\delta = \beta\gamma = \gamma/(2 + \gamma) = 1/41$. □

The second proof of the sum-product theorem is due to Elekes [27], and is completely different – it is strikingly elementary, but it relies on another important result of discrete geometry, the Szemerédi–Trotter Theorem (see [81, Th. 2] for this precise statement).

THEOREM 3.1.4 (Szemerédi–Trotter). *Let* P *be a finite subset of* $\mathbf{R}^2$ *and* L *a finite set of affine lines* $\ell \subset \mathbf{R}^2$. *Let* $k \geqslant 2$ *be an integer such that* $k \leqslant |\mathrm{P}|^{1/2}$.

If the lower bound $|\ell \cap \mathrm{P}| \geqslant k$ *holds for all* $\ell \in \mathrm{L}$, *then we have*

$$|\mathrm{L}| \ll |\mathrm{P}|^2 k^{-3},$$

where the implied constant is absolute.

SKETCH OF PROOF. The proof relies on the *crossing number inequality* which we discuss briefly in Section A.5. We apply it to the graph γ with vertices P and edge set E formed by the pairs $\{x, y\} \subset \mathrm{P}$ (with $x \neq y$) such that the line through x and y belongs to L, and there is no element of P other than x and y on the line segment from x to y. This construction gives also automatically a planar realization of γ, in the sense of Definition A.5.2.

Considering each line $\ell \in \mathrm{L}$ separately, we see that the number of edges of γ is then given by

$$|\mathrm{E}| = \sum_{\ell \in \mathrm{L}} (|\ell \cap \mathrm{P}| - 1) = |\mathrm{I}| - |\mathrm{L}|,$$

where $\mathrm{I} = \{(\ell, x) \in \mathrm{L} \times \mathrm{P} \mid x \in \ell\}$ is the *incidence set*.

We now consider the planar realization of γ. Since two distinct lines intersect at most once, the number of crossings of γ is at most $|\mathrm{L}|^2$. Theorem A.5.4 implies that either $|\mathrm{E}| < 4|\mathrm{P}|$ or

$$\frac{|\mathrm{E}|^3}{64|\mathrm{P}|^2} \leqslant |\mathrm{L}|^2.$$

These two cases mean that

$$|\mathrm{I}| < |\mathrm{L}| + 4|\mathrm{P}|, \quad \text{or} \quad |\mathrm{I}| \leqslant |\mathrm{L}| + 4|\mathrm{L}|^{2/3}|\mathrm{P}|^{2/3},$$

which are inequalities of independent interest, summarized for convenience in the single statement

$$|\mathrm{I}| \leqslant |\mathrm{L}| + 4|\mathrm{P}| + 4|\mathrm{L}|^{2/3}|\mathrm{P}|^{2/3}. \tag{3.5}$$

Under the assumptions of the Szemerédi–Trotter Theorem, we have a lower bound

$$|\mathrm{I}| = \sum_{\ell \in \mathrm{L}} |\ell \cap \mathrm{P}| \geqslant k|\mathrm{L}|,$$

hence the inequality

$$|\mathrm{L}| \leqslant \frac{1}{k}(|\mathrm{L}| + 4|\mathrm{P}| + 4|\mathrm{L}|^{2/3}|\mathrm{P}|^{2/3})$$

follows. Using the fact that $k \geqslant 2$, we get

$$|\mathrm{L}| \leqslant \frac{8|\mathrm{P}|}{k} + \frac{8|\mathrm{L}|^{2/3}|\mathrm{P}|^{2/3}}{k} \leqslant 16 \max\Big(\frac{|\mathrm{P}|}{k}, \frac{|\mathrm{L}|^{2/3}|\mathrm{P}|^{2/3}}{k}\Big),$$

or in other words

$$|\mathrm{L}| \leqslant \max\Big(\frac{16|\mathrm{P}|}{k}, \frac{16^3|\mathrm{P}|^2}{k^3}\Big) \leqslant \frac{16^3|\mathrm{P}|^2}{k^3}$$

where the last step uses the fact that $k \leqslant |\mathrm{P}|^{1/2}$. □

We can now prove give the proof of the sum-product inequality following Elekes.

ELEKES'S PROOF. We may assume that $|A| \geqslant 2$. Let $P = (A+A) \times (A \cdot A) \subset \mathbf{R}^2$, and let L be the set of lines in $\mathbf{R}^2$ obtained as graphs of the functions

$$f_{a,b}(x) = a(x-b)$$

for $(a,b) \in A^2$. Observing that

$$f_{a,b}(b+c) = ac \in A \cdot A$$

for all $c \in A$, we see that each line in L contains at least $|A|$ elements $(b+c, ac)$ of P.

We can apply the Szemerédi–Trotter inequality to P and L with $k = |A|$, since $|P| = |A+A||A \cdot A| \geqslant |A|^2 = k^2$. We get

$$|A|^2 = |L| \ll |P|^2 |A|^{-3} = (|A+A|\ |A \cdot A|)^2\ |A|^{-3},$$

i.e., $|A+A|\,|A \cdot A| \gg |A|^{5/2}$. Thus

$$\max(|A+A|, |A \cdot A|) \gg |A|^{5/4},$$

which concludes the proof of (3.1) with $\delta = 1/4$. □

REMARK 3.1.5. (1) To obtain (3.1), for some $\delta > 0$, it would have been enough to have a version of Theorem 3.1.4 with k^3 replaced by $k^{2+\beta}$ for some $\beta > 0$; such a statement was first proved by Beck [4, Th. 1.5].

(2) The Szemerédi–Trotter bound is essentially sharp, as shown by the example of the set $P = [N] \times [N]$ of N^2 points and of the N vertical lines L with equations $x = a$ for $a \in [N]$, where each line contains $k = N = |P|^{1/2}$ points of P, so that $|P|^2 k^{-3} = N$.

(3) We refer to the book of Guth [47, Ch. 7 to 12] for a deep and enlightening discussion of *incidence geometry* and some of its further applications (as well as discussions of open problems).

The third proof we present is due to Solymosi [77], and is also geometric in spirit.

SOLYMOSI'S PROOF. We view A as a subset of the multiplicative group $\mathbf{R}^\times$. The main result is a bound for the *multiplicative energy* of A, in terms of the size of the sumset $A+A$. Precisely, we claim that the inequality

$$E(A,A) \leqslant 3|A+A|^2 \log(4|A|) \tag{3.6}$$

holds. Since

$$E(A,A) \geqslant \frac{|A|^4}{|A \cdot A|},$$

by the Cauchy–Schwarz inequality (see (2.15)), it follows that

$$|A+A|^2\ |A \cdot A| \geqslant \frac{|A|^4}{2\log(4|A|)},$$

and hence, for any $\varepsilon > 0$, the inequality

$$\max(|A+A|, |A \cdot A|) \gg |A|^{4/3-\varepsilon}$$

holds, where the implied constant depends only on ε.

To prove (3.6), the key idea is to use the equality $E(A,A) = E(A, A^{-1})$, and the corresponding formula

$$E(A,A) = \sum_{x \in A \cdot A^{-1}} r(x)^2,$$

in terms of the representation function r for $A \cdot A^{-1}$. This has a geometric interpretation: for any x, the equality

$$r(x) = |\ell_x \cap (A \times A)|$$

holds, where $\ell_x \subset \mathbf{R}^2$ is the half-line in the plane (with coordinates (u, v)) with equation $u = xv$, with $v \geqslant 0$.

There are two features which emerge from this geometric point of view. First, if $x \neq y$ are elements of $\mathrm{A} \cdot \mathrm{A}^{-1}$, then the sumset

$$\mathrm{S}_{x,y} = (\ell_x \cap (\mathrm{A} \times \mathrm{A})) + (\ell_y \cap (\mathrm{A} \times \mathrm{A}))$$

in $\mathbf{R}^2$ satisfies

$$|\mathrm{S}_{x,y}| = |\ell_x \cap (\mathrm{A} \times \mathrm{A})| \, |\ell_y \cap (\mathrm{A} \times \mathrm{A})| = r(x)r(y),$$

simply because the vectors $(1, x)$ and $(1, y)$ are linearly independent in $\mathbf{R}^2$.

Second, the sumset $\mathrm{S}_{x,y}$ is contained in the convex hull of the half-lines ℓ_x and ℓ_y, and even in the interior if $x \neq y$. This set is the (strict) angular sector delimited by the two half-lines ℓ_x and ℓ_y. In particular, this means that if we take three elements of $\mathrm{A} \cdot \mathrm{A}^{-1}$ in increasing order, say $x < y < z$, then we have $\mathrm{S}_{x,y} \cap \mathrm{S}_{y,z} = \emptyset$.

All this suggests the use of a dyadic subdivision of the range of values of $r(x)$, in order to restrict attention to x's where $r(x)$ has roughly the same value, hence $r(x)r(y)$ is close to $r(x)^2$ if $y \neq x$. Doing so, we see that there exists an integer $\mathrm{N} \geqslant 1$ such that

$$\mathrm{E}(\mathrm{A}, \mathrm{A}) \leqslant \log(4|\mathrm{A}|) \sum_{\substack{x \in \mathrm{A} \cdot \mathrm{A}^{-1} \\ \mathrm{N} \leqslant r(x) < 2\mathrm{N}}} r(x)^2.$$

(Indeed, we have

$$\mathrm{E}(\mathrm{A}, \mathrm{A}) = \sum_{j \geqslant 0} \sum_{\substack{x \in \mathrm{A} \cdot \mathrm{A}^{-1} \\ 2^j \leqslant r(x) < 2^{j+1}}} r(x)^2,$$

and the inner sum has no non-zero terms if $2^j > |\mathrm{A}|$, since $s(x) \leqslant |\mathrm{A}|$; defining $\mathrm{N} = 2^{j_0}$ where j_0 is an integer with

$$\sum_{\substack{x \in \mathrm{A} \cdot \mathrm{A}^{-1} \\ 2^{j_0} \leqslant r(x) < 2^{j_0+1}}} r(x)^2$$

maximal, we obtain the result since the number of possibly non-zero terms is $\leqslant 2 + \log(|\mathrm{A}|)/\log(2) \leqslant \log(4|\mathrm{A}|)$.)

For x such that $\mathrm{N} < r(x) \leqslant 2\mathrm{N}$, we denote by $x_+ > x$ the next element with the same property, if it exists. We then have

$$r(x)^2 \leqslant 2r(x)r(x_+) = 2|\mathrm{S}_{x,x_+}| = 2|\mathrm{S}_x|.$$

When x is the largest element (say x_0) such that $\mathrm{N} < r(x) \leqslant 2\mathrm{N}$, the element x_+ does not exist. Still, we obtain first

$$\sum_{\substack{x \in \mathrm{A} \cdot \mathrm{A}^{-1} \\ \mathrm{N} \leqslant r(x) < 2\mathrm{N} \\ x \neq x_0}} r(x)^2 \leqslant 2 \sum_{\substack{x \in \mathrm{A} \cdot \mathrm{A}^{-1} \\ \mathrm{N} \leqslant r(x) < 2\mathrm{N} \\ x \neq x_0}} |\mathrm{S}_{x,x_+}| = 2 \Big| \bigcup_{\substack{x \in \mathrm{A} \cdot \mathrm{A}^{-1} \\ \mathrm{N} \leqslant r(x) < 2\mathrm{N} \\ x \neq x_0}} \mathrm{S}_{x,x_+} \Big|,$$

since by the second observation above, the sets S_{x,x_+} are pairwise disjoint. But, by construction, the sets S_{x,x_+} are also subsets of $(\mathrm{A} + \mathrm{A}) \times (\mathrm{A} + \mathrm{A})$, and thus

$$\sum_{\substack{x \in \mathrm{A} \cdot \mathrm{A}^{-1} \\ \mathrm{N} < r(x) \leqslant 2\mathrm{N} \\ x \neq x_0}} r(x)^2 \leqslant 2|\mathrm{A} + \mathrm{A}|^2.$$

We add $r(x_0)^2$, noting that

$$r(x_0)^2 \leqslant |\mathrm{A}|^2 \leqslant |\mathrm{A} + \mathrm{A}|^2,$$

to get

$$\mathrm{E}(\mathrm{A},\mathrm{A}) \leqslant 3\log(4|\mathrm{A}|)|\mathrm{A}+\mathrm{A}|^2,$$

as claimed. □

REMARK 3.1.6. If one wishes to squeeze the best possible constant, one can use a small trick (as is done by Solymosi): let a_0 be the largest element of A, and O the set of points (y, a_0) for y the abscissa of some element of $\ell_{x_0} \cap (\mathrm{A}\times\mathrm{A})$. Then, defining

$$\mathrm{S}_{x_0} = (\ell_{x_0} \cap (\mathrm{A}\times\mathrm{A})) + \mathrm{O},$$

we can also write

$$r(x_0)^2 = |\ell_{x_0} \cap (\mathrm{A}\times\mathrm{A})||\mathrm{O}| = |\mathrm{S}_{x_0}|,$$

and moreover, it is also elementary geometrically that the set S_{x_0} is disjoint from the subsets S_{y,y_+}. Since $\mathrm{S}_{x_0} \subset (\mathrm{A}+\mathrm{A})\times(\mathrm{A}+\mathrm{A})$ also, we get

$$\sum_{\substack{x\in \mathrm{A}\cdot\mathrm{A}^{-1}\\ \mathrm{N}<r(x)\leqslant 2\mathrm{N}}} r(x)^2 \leqslant 2|\mathrm{A}+\mathrm{A}|^2.$$

REMARK 3.1.7. The basic inequality (3.6) of Solymosi is essentially sharp: if we take $\mathrm{A} = [\mathrm{N}]$ for some positive integer N, then $|\mathrm{A}+\mathrm{A}| = 2\mathrm{N}-1$ and $\mathrm{E}(\mathrm{A},\mathrm{A})$ is certainly at least of size N^2.

3.2. Sum-product in finite fields

The theorem of Erdős and Szemerédi mixes two types of product sets (but note that $\mathbf{Z}$ is not a group for multiplication), and relies on the fact that both addition and multiplication exist on $\mathbf{Z}$, or in other words that it is a ring. In view of Cauchy's Theorem, it is of course natural to also consider the case of finite rings, starting with the finite fields $\mathbf{F}_p$ for p prime. Indeed, we have the following result (see [11]):

THEOREM 3.2.1 (Bourgain–Katz–Tao). *For any $\gamma > 0$, there exists $\delta > 0$ such that for any prime number p and any set $\mathrm{A} \subset \mathbf{F}_p$ such that*

$$p^\gamma \leqslant |\mathrm{A}| \leqslant p^{1-\gamma},$$

we have

$$\max(|\mathrm{A}+\mathrm{A}|, |\mathrm{A}\cdot\mathrm{A}|) \gg |\mathrm{A}|^{1+\delta}, \tag{3.7}$$

where the implied constant depends only on γ.

REMARK 3.2.2. This theorem should remind the reader of the statement of Theorem 2.6.7 (which is, of course, chronologically posterior), and indeed this was one of the inspirations for the work of Helfgott. However, note that the exponent δ here *depends* on γ. This is not an artefact of the proof, but a necessity (see Remark 3.2.11 for a related situation where this is more transparent).

We will again present different proofs. One is the original argument of Bourgain, Katz and Tao, and the second, due to Breuillard [13, § 2.13], is based on an interpretation of the result as a statement about approximate subgroups of the affine-linear group, combined with techniques related to the study of approximate subgroups of non-abelian groups, i.e., to Helfgott's Theorem and its generalizations.

Both proofs will begin with a statement (due to Katz and Tao), which is in spirit something like a version for "approximate rings" of the Balog–Szemerédi–Gowers Theorem 2.5.5.

First comes a translation from a set with *both* sum and product growing slowly, to a single condition.

PROPOSITION 3.2.3. *Let p be a prime number. Let A be a non-empty subset of $\mathbf{F}_p$ and let $\alpha \geqslant 1$ be such that*

$$\max(|\mathrm{A}+\mathrm{A}|, |\mathrm{A}\cdot\mathrm{A}|) \leqslant \alpha|\mathrm{A}|.$$

There exists a subset $\mathrm{B} \subset \mathrm{A}$ such that

$$|\mathrm{A}| \leqslant \beta|\mathrm{B}|, \qquad |\mathrm{B}\cdot\mathrm{B} - \mathrm{B}\cdot\mathrm{B}| \leqslant \beta|\mathrm{B}|. \tag{3.8}$$

for some real number $\beta \leqslant c\alpha^d$, where $c \geqslant 0$ and $d \geqslant 0$ are independent of p and A.

Then we have a kind of ring analogue of Ruzsa's Lemma (Proposition 2.4.10), due to Bourgain, Katz and Tao. We extend here the notation for product sets to handle more complicated sets like $\mathrm{A}\cdot\mathrm{A}-\mathrm{A}\cdot\mathrm{A}$. Precisely, for any polynomial $f \in \mathbf{Z}[\mathrm{X}_1, \ldots, \mathrm{X}_d]$, where $d \geqslant 1$, we define the set $f_*(\mathrm{A})$ as follows: for a monomial

$$f = m\mathrm{X}_1^{n_1}\cdots\mathrm{X}_d^{n_d}$$

we put

$$f_*(\mathrm{A}) = \mathrm{A}^{(n_1+\cdots+n_d)} + \cdots + \mathrm{A}^{(n_1+\cdots+n_d)}$$

with m summands (an m-fold sum of a multiple product set), and for $f = f_1 + f_2$, we define

$$f_*(\mathrm{A}) = f_*(\mathrm{A}_1) + f_*(\mathrm{A}_2).$$

For any subset $\mathrm{B} \subset \mathbf{F}_p$, we will also denote $\mathrm{B}^{-1} = (\mathrm{B} - \{0\})^{-1}$.

EXAMPLE 3.2.4. For instance, if $f = \mathrm{X}_1^2 - \mathrm{X}_2^2$, then $f_*(\mathrm{A}) = \mathrm{A}\cdot\mathrm{A} - \mathrm{A}\cdot\mathrm{A}$. This is also the case for $f = \mathrm{X}_1\mathrm{X}_2 - \mathrm{X}_3\mathrm{X}_4$, among many other choices.

A few minutes's thoughts shows (as a few experiments suggest) that we have the inclusion

$$f_*(g_*(\mathrm{A})) \subset (f\circ g)_*(\mathrm{A}) \tag{3.9}$$

for any polynomials f and g.

PROPOSITION 3.2.5. *Let E be a finite field. Let A be a non-empty subset of E containing 1 and let $\alpha \geqslant 1$ be such that*

$$|\mathrm{A}\cdot\mathrm{A} - \mathrm{A}\cdot\mathrm{A}| \leqslant \alpha|\mathrm{A}|.$$

For any integer $m \geqslant 1$, and for any polynomials $f \in \mathbf{Z}[\mathrm{X}_1, \ldots, \mathrm{X}_m]$, we have

$$|f_*(\mathrm{A})\cdot f_*(\mathrm{A})^{-1}| \leqslant \beta|\mathrm{A}|, \text{ and } |f_*(\mathrm{A})| \leqslant \beta|\mathrm{A}|$$

where $\beta \leqslant c\alpha^d$ for some integers $c \geqslant 0$ and $d \geqslant 0$ depending only on f.

Assuming the previous two results, which will be proved in Section 3.3, we can now prove Theorem 3.2.1 following Bourgain, Katz and Tao.

THE BOURGAIN–KATZ–TAO PROOF. We first observe that it is enough to prove the theorem for sets containing 1, since in any case the case $\mathrm{A}' = \mathrm{A}\cup\{1\}$ satisfies the same size assumptions as A, up to replacing γ by a slightly smaller positive number (any one will do if p is large enough, which we may also assume), and moreover

$$|\mathrm{A}'+\mathrm{A}'| \leqslant |\mathrm{A}+\mathrm{A}| + |\mathrm{A}| + 1 \leqslant 3|\mathrm{A}+\mathrm{A}|, \qquad |\mathrm{A}'\cdot\mathrm{A}'| \leqslant |\mathrm{A}\cdot\mathrm{A}| + |\mathrm{A}| \leqslant 2|\mathrm{A}\cdot\mathrm{A}|.$$

Let $\mathrm{A} \subset \mathbf{F}_p$ be a non-empty set containing 1 such that

$$p^{\gamma} \leqslant |\mathrm{A}| \leqslant p^{1-\gamma}.$$

We let $\alpha \geqslant 1$ be such that

$$|\mathrm{A}\cdot\mathrm{A} - \mathrm{A}\cdot\mathrm{A}| \leqslant \alpha|\mathrm{A}|. \tag{3.10}$$

We will first show that $\alpha \gg |\mathrm{A}|^{\delta}$ for some $\delta > 0$ depending only on γ.

Step 1. There exists an integer k, bounded in terms of γ only, and a linear form $\lambda\colon \mathbf{F}_p^k \to \mathbf{F}_p$ such that the restriction of λ to A^k is surjective.

Indeed, for any integer $n \geqslant 1$, if we iterate n times the variant of Cauchy's Theorem given by Proposition 1.3.9, we find elements $x_1 = 1, \ldots, x_n \in \mathbf{F}_p^{\times}$ such that

$$|x_1\mathrm{A} + \cdots + x_n\mathrm{A}| \geqslant \min\Big(\frac{|\mathrm{A}|^n}{2^n}, \frac{p}{12}\Big),$$

(where here $x\mathrm{A}$ denotes the dilated set $\{xa \mid a \in \mathrm{A}\}$).

Taking $k = \lceil \log(p/12)/\log(|\mathrm{A}|/2)\rceil$, this gives x_i's such that

$$|x_1\mathrm{A} + \cdots + x_n\mathrm{A}| \geqslant \frac{p}{12},$$

and since $(\frac{1}{2}p)^{\gamma} \leqslant \frac{1}{2}p^{\gamma} \leqslant \frac{1}{2}|\mathrm{A}|$, we have $k \leqslant \frac{1}{\gamma} + 1$. Applying then Cauchy's Theorem, in the form of Corollary 1.3.2, we find an integer $k \leqslant 24n \leqslant 24(\gamma^{-1}+1)$ and elements x_1, $\ldots$, x_k in $\mathbf{F}_p^{\times}$ such that $x_1\mathrm{A} + \cdots + x_k\mathrm{A} = \mathbf{F}_p$, hence the result with

$$\lambda(y_1, \ldots, y_k) = x_1y_1 + \cdots + x_ky_k.$$

Step 2. Let B be a subset of $\mathbf{F}_p$. If $k \geqslant 2$ is an integer and λ is a linear form on $\mathbf{F}_p^k$ such that the restriction of λ to B^k is surjective, then there exists a linear form $\widetilde{\lambda}\colon \mathbf{F}_p^{k-1} \to \mathbf{F}_p$ whose restriction to

$$\widetilde{\mathrm{B}} = \mathrm{B}\cdot(\mathrm{B}-\mathrm{B}) + \mathrm{B}\cdot(\mathrm{B}-\mathrm{B})$$

is surjective. (Note that $\widetilde{\mathrm{B}} = f_*(\mathrm{B})$ with $f = \mathrm{X}_1(\mathrm{X}_2 - \mathrm{X}_3) + \mathrm{X}_4(\mathrm{X}_5 - \mathrm{X}_6)$.)

Amusingly, this is a form of Gaussian elimination. Precisely, note that the restriction of λ to B^k is not injective (for the simple reason that this would imply that $p = |\mathbf{F}_p| = |\mathrm{B}|^k$, contradicting the primality of p), hence we find two distinct elements $x_0 \neq y_0$ in B^k such that $\lambda(x_0) = \lambda(y_0)$. Thus $z = x_0 - y_0$ is a non-zero element of $\mathrm{B} - \mathrm{B}$ in the kernel of λ. Assume, for simplicity, that the last coordinate z_k of z is non-zero, and write

$$\lambda(y) = a_1x_1 + \cdots + a_kx_k$$

for all $x \in \mathbf{F}_p^k$. We obtain from $\lambda(z) = 0$ the identity

$$a_kz_k = -\sum_{j=1}^{k-1} a_jz_j.$$

For any $t \in \mathbf{F}_p$, the surjectivity of $\lambda\colon \mathrm{B}^k \to \mathbf{F}_p$ applied to t/z_k gives elements b_1, $\ldots$, b_k of B such that

$$t = a_1z_kb_1 + \cdots + a_kz_kb_k,$$

and we conclude that

$$t = \sum_{j=1}^{k-1} a_j(z_kb_j - z_jb_k),$$

which gives the result with $\widetilde{\lambda}(y) = a_1y_1 + \cdots + a_{k-1}y_{k-1}$, since $z_kb_j - z_jb_k \in \widetilde{\mathrm{B}}$.

Step 3. We apply Step 1 to the set A, and then iterate $k-1$ times the previous step (recalling that k is bounded in terms of γ only, hence is a constant for fixed γ); because

of (3.9), there is a polynomial f (in many variables, but depending only on γ) and a surjective map $f_*(\mathrm{A}) \to \mathbf{F}_p$. It follows that

$$|f_*(\mathrm{A})| \geqslant p.$$

Combining this with Proposition 3.2.5 applied to f, we obtain

$$p \leqslant |f_*(\mathrm{A})| \leqslant c\alpha^d |\mathrm{A}|$$

for some constants c and d. Under the assumption that $|\mathrm{A}| \leqslant p^{1-\gamma}$, this implies that

$$\alpha \gg \Big(\frac{p}{|\mathrm{A}|}\Big)^{1/d} \gg |\mathrm{A}|^{\delta}, \qquad \delta = \frac{\gamma}{d(1-\gamma)}.$$

This estimate was our first objective, but it is not exactly the right kind, since α is defined in (3.10) in terms of the growth of $\mathrm{A}\cdot\mathrm{A} - \mathrm{A}\cdot\mathrm{A}$. However, in general, let $\beta \geqslant 1$ be such that

$$\max(|\mathrm{A}+\mathrm{A}|, |\mathrm{A}\cdot\mathrm{A}|) = \beta|\mathrm{A}|.$$

Proposition 3.2.3 provides a subset $\mathrm{B} \subset \mathrm{A}$ such that

$$|\mathrm{B}\cdot\mathrm{B} - \mathrm{B}\cdot\mathrm{B}| \leqslant \kappa|\mathrm{B}|, \qquad |\mathrm{A}| \leqslant \kappa|\mathrm{B}|,$$

where $\kappa \leqslant c'\beta^{d'}$ for some constants c' and d'. We have of course $|\mathrm{B}| \leqslant |\mathrm{A}| \leqslant p^{1-\gamma}$. On the other hand, from $|\mathrm{B}| \geqslant \kappa^{-1}|\mathrm{A}|$, we see that either $\kappa \geqslant p^{\gamma/2}$, or $|\mathrm{B}| \geqslant p^{\gamma/2}$.

In the first case, we deduce from $\beta \gg \kappa^{1/d'}$ that $\beta \gg p^{\gamma/(2d')}$, and this, in turn, gives

$$\beta \gg |\mathrm{A}|^{\frac{\gamma}{2d'(1-\gamma)}},$$

which is of the desired form.

In the second case, the previous argument can be applied to the set B (with γ replaced by $\gamma/2$). It yields

$$\max(|\mathrm{B}+\mathrm{B}|, |\mathrm{B}\cdot\mathrm{B}|) \gg |\mathrm{B}|^{1+\delta},$$

for some $\delta > 0$ depending only on γ. This implies

$$\beta|\mathrm{A}| = \max(|\mathrm{A}+\mathrm{A}|, |\mathrm{A}\cdot\mathrm{A}|) \gg \Big(\frac{|\mathrm{A}|}{\kappa}\Big)^{1+\delta} \gg \Big(\frac{|\mathrm{A}|}{\beta^{d'}}\Big)^{1+\delta},$$

and this straightens itself into a bound $\beta \gg |\mathrm{A}|^{\delta'}$, where $\delta' = 1 + d'(1+\delta) > 0$, a constant depending again only on γ. □

Remark 3.2.6. As happened for product sets, it is not necessarily the case that a finite subset A of a field E which satisfies

$$\max(|\mathrm{A}+\mathrm{A}|, |\mathrm{A}\cdot\mathrm{A}|) \leqslant \alpha|\mathrm{A}|$$

always satisfies a bound

$$|\mathrm{A}\cdot\mathrm{A} - \mathrm{A}\cdot\mathrm{A}| \leqslant \beta|\mathrm{A}|$$

with β depending polynomially on α (or similar bounds for other sets of the form $f_*(\mathrm{A})$).

However, the easiest examples are $\mathrm{A} = \mathrm{F} \cup \{x\}$ where $\mathrm{F} \subset \mathrm{E}$ is a proper subfield and $x \in \mathrm{E} - \mathrm{F}$ (in which case $|\mathrm{A}\cdot\mathrm{A}| \leqslant 2|\mathrm{A}|$ and $|\mathrm{A}+\mathrm{A}| \leqslant 2|\mathrm{A}|$, as one checks easily, but $|\mathrm{A}\cdot\mathrm{A} - \mathrm{A}\cdot\mathrm{A}| \geqslant |\mathrm{E} + x\mathrm{E}| = (|\mathrm{A}|-1)^2$.

A posteriori, one can see that Theorem 3.2.1 shows this conclusion *does* hold for the field $\mathbf{F}_p$ and for subsets which satisfy the size conditions there, but for the "trivial" reason that α must be at least $|\mathrm{A}|^{\delta} \geqslant p^{\gamma\delta}$, so the trivial bound $|f_*(\mathrm{A})| \leqslant p$ is already polynomial in terms of α.

We will now present a second proof of the sum-product theorem over finite fields, which is due to Breuillard [13], and makes very explicit the connection between this result and the general study of approximate groups. In this case, the question is to classify the approximate subgroups of the *affine linear group* $\mathrm{Aff}(\mathbf{F}_p)$ over $\mathbf{F}_p$, which is the group of transformations $x \mapsto ax+b$ of $\mathbf{F}_p$, where $(a,b) \in \mathbf{F}_p^\times \times \mathbf{F}_p$, with composition as the group law. Although there exist general statements of this kind, goint back to Helfgott [50, Th. 3.6] and Murphy [63] (later improved by Rudnev and Shkredov [69, Th. 6]), we state only the version that leads immediately to the sum-product theorem.

We will identify the group $\mathrm{Aff}(\mathbf{F}_p)$ with the group of invertible matrices of the form

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix},$$

as we may since the composition of affine-linear maps behaves like the product of the corresponding matrices. For a subset A of $\mathbf{F}_p$, we further denote

$$\mathrm{Aff}(\mathrm{A}) = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mid (a,b) \in (\mathrm{A} - \{0\}) \times \mathrm{A} \right\}$$

and $\mathrm{Aff}_s(\mathrm{A}) = \mathrm{Aff}(\mathrm{A}) \cup \mathrm{Aff}(\mathrm{A})^{-1} \cup \{1\}$.

THEOREM 3.2.7 (Partial growth theorem for $\mathrm{Aff}(\mathbf{F}_p)$). *Let p be a prime number and let A be a subset of $\mathbf{F}_p$ containing 0 and 1. The symmetric subset $\mathrm{Aff}_s(\mathrm{A})$ of $\mathrm{Aff}(\mathbf{F}_p)$ satisfies*

$$|\mathrm{Aff}_s(\mathrm{A})^{(3)}| \geqslant \min\Big(p\Big(\frac{|\mathrm{Aff}_s(\mathrm{A})|}{|f_*(\mathrm{A}) \cdot f_*(\mathrm{A})^{-1}|} - 1\Big), |\mathrm{Aff}_s(\mathrm{A})|\Big(\frac{|\mathrm{Aff}_s(\mathrm{A})|}{|f_*(\mathrm{A}) \cdot f_*(\mathrm{A})^{-1}|}\Big)^{1/3}\Big),$$

for some polynomial f independent of p and A.

Using this theorem, we now prove Theorem 3.2.1; in fact, we can now prove it without the assumption that $|\mathrm{A}| \geqslant p^\gamma$ (a version of the sum-product phenomenon over finite fields which did not require this condition was first obtained by Konyagin [56]).

BREUILLARD'S PROOF. We will make the minor assumption that 0 and 1 belong to A.

As in the original argument, we begin with the case where the subset A satisfies the condition $|\mathrm{A} \cdot \mathrm{A} - \mathrm{A} \cdot \mathrm{A}| \leqslant \alpha|\mathrm{A}|$, and show that α must be large.

According to Proposition 3.2.5, the set $\mathrm{B} = f_*(\mathrm{A}) \cdot f_*(\mathrm{A})^{-1}$ appearing in Theorem 3.2.7 satisfies $|\mathrm{B}| \leqslant \beta|\mathrm{A}|$ with $\beta \leqslant c\alpha^d$ for some constants c and α. In particular, we will have $|\mathrm{B}| \leqslant \frac{1}{2}|\mathrm{A}|^2$ unless $\alpha \geqslant |\mathrm{A}|^\delta$ for some constant $\delta > 0$, in which case we are done. When $|\mathrm{B}| \leqslant \frac{1}{2}|\mathrm{A}|^2$, Theorem 3.2.7 yields

$$|\mathrm{Aff}_s(\mathrm{A})^{(3)}| \gg |\mathrm{A}|^2 \min\Big(\frac{p}{|\mathrm{B}|}, \Big(\frac{|\mathrm{Aff}_s(\mathrm{A})|}{|\mathrm{B}|}\Big)^{1/3}\Big) \gg |\mathrm{A}|^2 \min\Big(\frac{p^\gamma}{\beta}, \Big(\frac{|\mathrm{A}|}{\beta}\Big)^{1/3}\Big).$$

On the other hand, from the matrix computations

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} a^{-1} & -a^{-1}b \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} c & d \\ 0 & 1 \end{pmatrix}\begin{pmatrix} e & f \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} ace & acf+ad+b \\ 0 & 1 \end{pmatrix},$$

we deduce that $\mathrm{Aff}_s(\mathrm{A})^{(3)} \subset \mathrm{Aff}(\mathrm{C})$ where $\mathrm{C} = g_*(\mathrm{A}) \cdot g_*(\mathrm{A})^{-1}$ for some fixed polynomial g. Proposition 3.2.5 therefore leads to

$$|\mathrm{Aff}_s(\mathrm{A})^{(3)}| \leqslant |\mathrm{C}|^2 \leqslant \kappa^2|\mathrm{A}|^2$$

where $\kappa \leqslant c_1\alpha^{d_1}$ for some constants c_1 and d_1.

Comparing the two bounds and using the fact that $|A| \leqslant p^{1-\gamma}$, we deduce that

$$\min(|A|^{\gamma/(1-\gamma)}, |A|^{1/3}) \leqslant \min(p^{\gamma}, |A|^{1/3}) \ll c_2\alpha^{d_2}$$

for some constants c_2 and d_2, which leads to the conclusion $\alpha \gg |A|^{\delta}$ for some $\delta > 0$ depending only on γ.

Finally, we reduce the case of a general subset A to this first case in exactly the same manner as was done at the end of the proof of Bourgain, Katz and Tao. □

We now prove Theorem 3.2.7. This uses two key ingredients from the elementary theory of approximate groups, which are due to Helfgott. The first is a version of the classical *orbit-stabilizer theorem.*

LEMMA 3.2.8 (Orbit-stabilizer lemma). *Let* G *be a group acting on a non-empty set* X. *Let* $x \in X$ *and denote by* H *the stabilizer of* x *in* G. *For any finite non-empty symmetric subset* A *of* G, *we have*

$$|A| \leqslant |H \cap A^{(2)}|\ |A \cdot x|,$$

where $A \cdot x = \{a \cdot x \mid a \in A\}$.

PROOF. Consider the surjective orbit map $f\colon A \to A \cdot x$ which maps a to $a \cdot x$. We have the equality

$$\sum_{y \in A \cdot x} |f^{-1}(y)| = |A|.$$

For a given $y = a \cdot x \in A$, the elements in $f^{-1}(y)$ are the $b \in A$ such that $a \cdot x = b \cdot x$. This condition holds if and only if $b^{-1}a \in H$, and since $b^{-1}a \in A^{(2)}$, this implies the bound $|f^{-1}(y)| \leqslant |H \cap A^{(2)}|$. Summing over y gives the lemma. □

REMARK 3.2.9. If $A = G$ (so G is finite) then the orbit-stabilizer formula is the equality $|G| = |H||G \cdot x|$, which explains the interpretation of this result.

The second fact we will use is the following, which states intuitively that if A is an approximate subgroup of G, then its intersection with any subgroup is also an approximate subgroup in that group.

LEMMA 3.2.10. *Let* G *be a group and* H *a subgroup of* G. *For any non-empty symmetric subset* A *of* G, *and any integer* $k \geqslant 1$, *we have*

$$\frac{|A^{(k+1)}|}{|A|} \geqslant \frac{|H \cap A^{(k)}|}{|H \cap A^{(2)}|}.$$

PROOF. We compute the size of $A^{(k+1)}$ by intersecting with cosets of H: we have

$$|A^{(k+1)}| = \sum_{xH \in G/H} |xH \cap A^{(k+1)}|.$$

We next observe that, as soon as $xH \cap A$ is not empty, say with $y \in xH \cap A$, the set $xH \cap A^{(k+1)}$ contains $y(H \cap A^{(k)})$, so that $|xH \cap A^{(k+1)}| \geqslant |H \cap A^{(k)}|$. Hence

$$|A^{(k+1)}| \geqslant |H \cap A^{(k)}|\,|X|$$

where X is the set of cosets xH which intersect A.

To obtain a lower bound for the size of X, we consider the analogue of the decomposition above for A itself:

$$|A| = \sum_{xH \in X} |xH \cap A|,$$

and observe that $|xH \cap A| \leqslant |H \cap A^{(2)}|$, so that

$$|A| \geqslant |X|\,|H \cap A^{(2)}|.$$

The combination of these bounds leads to the result. □

Before proving Theorem 3.2.7, we introduce some notation. The group $\mathrm{Aff}(\mathbf{F}_p)$ contains as subgroups the group of dilations or homotheties

$$D = \left\{ \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \mid a \in \mathbf{F}_p^\times \right\},$$

which is isomorphic to $\mathbf{F}_p^\times$, and the group of translations

$$U = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid b \in \mathbf{F}_p \right\},$$

which is isomorphic to $\mathbf{F}_p$. Since

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix},$$

and $D \cap U = \{1\}$, any element g of $\mathrm{Aff}(\mathbf{F}_p)$ has a unique representation $g = ud$ with $d \in D$ and $u \in U$; we call the corresponding elements of $\mathbf{F}_p^\times$ and $\mathbf{F}_p$ the *dilation factor* and *translation factor* of g, respectively. Note that the map $g \mapsto d$ is a group homomorphism.

For any subset A of $\mathrm{Aff}(\mathbf{F}_p)$, we will denote by A_{tr} the intersection $A \cap U(\mathbf{F}_p)$, identified with a subset of $\mathbf{F}_p$.

Proof of Theorem 3.2.7. We can assume that A contains at least one element different from 0 and 1, since otherwise the result is immediate.

The proof exploits two natural actions of $\mathrm{Aff}(\mathbf{F}_p)$. The first one, which exists for any group, is the action of the group on itself by conjugation; the second is the "obvious" one of $\mathrm{Aff}(\mathbf{F}_p)$ on $\mathbf{F}_p$ by affine-linear maps.

We first consider some properties of this second action. It is transitive, or in other words, the orbit of any $x \in \mathbf{F}_p$ is equal to $\mathbf{F}_p$ (indeed, this is already true for the action of the subgroup U by translation). By direct computation, the stabilizer of a given x is the subgroup T_x of matrices of the form

$$\begin{pmatrix} a & (1-a)x \\ 0 & 1 \end{pmatrix}$$

with $a \in \mathbf{F}_p^\times$. Each of these subgroups is isomorphic to $\mathbf{F}_p^\times$, by the homomorphism sending a matrix as above to a. The union of all T_x is the set of elements in $\mathrm{Aff}(\mathbf{F}_p)$ which have some fixed point, hence is the complement $\mathrm{Aff}(\mathbf{F}_p) - (U - \{1\})$ of the set of non-trivial translations. Moreover, we have $T_x \cap T_y = \{1\}$ if $x \neq y$ (because an affine linear transformation which is not the identity has at most one fixed point). Finally, we observe that for $g \in \mathrm{Aff}(\mathbf{F}_p)$ and $x \in \mathbf{F}_p$, we have $gT_xg^{-1} = T_{g \cdot x}$ (this is a general property of stabilizers in any group action).

Now the idea we will use is to study how $\mathrm{Aff}_s(A)^{(2)}$ is "shared" among the various subgroups T_x. Roughly speaking, there will be a dichotomy: either *all* T_x contain many elements of $\mathrm{Aff}_s(A)^{(2)}$, or there is such an accumulation in a *single* T_x that it suffices to produce growth using Lemma 3.2.10.

To implement this idea, we study $T_x \cap \mathrm{Aff}_s(A)^{(2)}$ for $x \in \mathbf{F}_p$. This set contains the identity, since $\mathrm{Aff}_s(A)$ is symmetric. We say that x is *involved* (with A) if $T_x \cap \mathrm{Aff}_s(A)^{(2)}$ is not reduced to the identity.

Suppose that $x \in \mathbf{F}_p$ is involved. There exists then $g \in \mathrm{Aff}_s(\mathrm{A})^{(2)} - \{1\}$ fixing x, and no other element of $\mathbf{F}_p$. We now observe that T_x is also the centralizer of g in $\mathrm{Aff}(\mathbf{F}_p)$, i.e., the stabilizer of g for the conjugation action of $\mathrm{Aff}(\mathbf{F}_p)$ on itself. Indeed, since T_x is abelian and contains g, it is contained in the centralizer of g; conversely, if $gh = hg$, we deduce that $g \cdot (h \cdot x) = h \cdot x$, so that $h \cdot x$ is a fixed point of g; by uniqueness of the fixed point of g, this means that $h \cdot x = x$, i.e., that $h \in \mathrm{T}_x$.

Applying Lemma 3.2.8 to the *conjugation action*, we deduce a lower bound

$$|\mathrm{T}_x \cap \mathrm{Aff}_s(\mathrm{A})^{(2)}| \geqslant \frac{|\mathrm{Aff}_s(\mathrm{A})|}{|g^{\mathrm{Aff}_s(\mathrm{A})}|},$$

where $g^{\mathrm{Aff}_s(\mathrm{A})} = \{hgh^{-1} \mid h \in \mathrm{Aff}_s(\mathrm{A})\}$ (a common notation that is used to avoid confusing the two actions).

At this point, we make a key computation: if $g = \begin{pmatrix} c & d \\ 0 & 1 \end{pmatrix}$, then

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} c & d \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} c & a^{-1}(b(c-1)+d) \\ 0 & 1 \end{pmatrix}, \tag{3.11}$$

which shows that all conjugates of g have the same dilation factor (which is simply because the dilation factor is multiplicative) and that, if we conjugate by an element of $\mathrm{Aff}_s(\mathrm{A})^2$, then the translation factor will belong to $f_*(\mathrm{A}) \cdot f_*(\mathrm{A})^{-1}$ for some fixed polynomial f. (Indeed, note first from

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} a^{-1} & -ba^{-1} \\ 0 & 1 \end{pmatrix}$$

we have (say) $\mathrm{Aff}_s(\mathrm{A}) \subset \mathrm{Aff}(\mathrm{A} - \mathrm{A} \cdot \mathrm{A}^{-1})$, and then that if, say

$$g = \begin{pmatrix} c_1 & d_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c_2 & d_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} c_1c_2 & c_1d_2 + d_1 \\ 0 & 1 \end{pmatrix},$$

with factors in $\mathrm{Aff}(\mathrm{A}_1)$, then in (3.11), we have

$$a^{-1}(b(c-1)+d) = a^{-1}(bc_1c_2 - b + c_1d_2 + d_1) \in \mathrm{A}_1^{-1} \cdot (\mathrm{A}_1^{(3)} - \mathrm{A}_1 + \mathrm{A}_1^{(2)} + \mathrm{A}_1),$$

which implies the result.)

We therefore have

$$|\mathrm{T}_x \cap \mathrm{Aff}_s(\mathrm{A})^{(2)}| \geqslant \frac{|\mathrm{Aff}_s(\mathrm{A})|}{|f_*(\mathrm{A}) \cdot f_*(\mathrm{A})^{-1}|}, \tag{3.12}$$

an inequality which is valid for any $x \in \mathbf{F}_p$ which is involved. (Note that if we had used the orbit-stabilizer lemma for the action on $\mathbf{F}_p$ to obtain a lower-bound for $|\mathrm{T}_x \cap \mathrm{Aff}_s(\mathrm{A})^{(2)}|$, the resulting lower bound would be $|\mathrm{Aff}_s(\mathrm{A})|/|\mathrm{Aff}_s(\mathrm{A}) \cdot x|$, but we have little control on an upper-bound for $|\mathrm{Aff}_s(\mathrm{A}) \cdot x|$; indeed, it follows from Proposition 1.3.9 that for *some* $x \in \mathbf{F}_p$ at least, the size of $\mathrm{Aff}_s(\mathrm{A}) \cdot x$ is close to that of $\mathrm{Aff}_s(\mathrm{A})$, in which case the lower-bound from the orbit-stabilizer lemma would be useless.)

We are now ready to start the final stages of the proof. First, we observe that some $x \in \mathbf{F}_p$ must be involved. Indeed, otherwise this means by definition that $\mathrm{Aff}_s(\mathrm{A})^{(2)}$ is contained in U, which is not the case since A contains an element distinct from 0 and 1. We then distinguish two cases.

Case 1. Assume that if x is involved and $g \in \mathrm{Aff}_s(\mathrm{A})$, then $g \cdot x \in \mathbf{F}_p$ is also involved. Since the subgroup generated by $\mathrm{Aff}_s(\mathrm{A})$ acts transitively on $\mathbf{F}_p$ (e.g., simply because it

contains a non-trivial translation), we deduce that every element of $\mathbf{F}_p$ is involved. Now we get

$$|\operatorname{Aff}_s(\mathrm{A})^{(2)}| \geqslant \sum_{x\in\mathbf{F}_p} |(\mathrm{T}_x - \{1\}) \cap \operatorname{Aff}_s(\mathrm{A})^{(2)}|$$

since the sets $\mathrm{T}_x - \{1\}$ are pairwise disjoint, and then

$$|\operatorname{Aff}_s(\mathrm{A})^{(2)}| \geqslant \frac{p|\operatorname{Aff}_s(\mathrm{A})|}{|f_*(\mathrm{A})\cdot f_*(\mathrm{A})^{-1}|} - p$$

by (3.12).

Case 2. There exists an x which is involved and some $g \in \operatorname{Aff}_s(\mathrm{A})$ such that $g\cdot x$ is *not* involved. We then apply Lemma 3.2.10 to the subgroup $\mathrm{T}_{g\cdot x}$, with $k \geqslant 1$ to be determined later; this leads to the inequality

$$\frac{|\operatorname{Aff}_s(\mathrm{A})^{(k+1)}|}{|\operatorname{Aff}_s(\mathrm{A})|} \geqslant \frac{|\mathrm{T}_{g\cdot x}\cap \operatorname{Aff}_s(\mathrm{A})^{(k)}|}{|\mathrm{T}_{g\cdot x}\cap \operatorname{Aff}_s(\mathrm{A})^{(2)}|} = |\mathrm{T}_{g\cdot x}\cap \operatorname{Aff}_s(\mathrm{A})^{(k)}|$$

since $g\cdot x$ is not involved. But

$$\mathrm{T}_{g\cdot x}\cap \operatorname{Aff}_s(\mathrm{A})^{(k)} = g\mathrm{T}_x g^{-1}\cap \operatorname{Aff}_s(\mathrm{A})^{(k)} \supset g(\mathrm{T}_x\cap \operatorname{Aff}_s(\mathrm{A})^{(k-2)})g^{-1}$$

since $g\in\operatorname{Aff}_s(\mathrm{A})$. We take $k=4$, and use the fact that x *is* involved to conclude that

$$\frac{|\operatorname{Aff}_s(\mathrm{A})^{(5)}|}{|\operatorname{Aff}_s(\mathrm{A})|} \geqslant |\mathrm{T}_x\cap \operatorname{Aff}_s(\mathrm{A})^{(2)}| \geqslant \frac{|\operatorname{Aff}_s(\mathrm{A})|}{|f_*(\mathrm{A})\cdot f_*(\mathrm{A})^{-1}|}$$

by (3.12), and so

$$|\operatorname{Aff}_s(\mathrm{A})^{(3)}| \geqslant |\operatorname{Aff}_s(\mathrm{A})|\Big(\frac{|\operatorname{Aff}_s(\mathrm{A})|}{|f_*(\mathrm{A})\cdot f_*(\mathrm{A})^{-1}|}\Big)^{1/3}$$

by Ruzsa's Lemma.

The combination of Cases 1 and 2 gives the claimed bound. □

Remark 3.2.11. A natural question is whether the formula of Theorem 3.2.7, assuming that the set A is "not too large", is really necessary, or if a statement similar to Helfgott's Theorem would be possible (without such an assumption, but with the alternative possibility that $\mathrm{A}^{(3)}$ is the whole group, or some other fixed multiple product set $\mathrm{A}^{(m)}$). This is not the case, because in contrast with $\mathrm{SL}_2(\mathbf{F}_p)$, the group $\operatorname{Aff}(\mathbf{F}_p)$ is not quasirandom enough. In fact, since we have the surjective group morphism $\operatorname{Aff}(\mathbf{F}_p)\to\mathbf{F}_p^\times$ mapping an affine-linear map to its dilation factor, there are many non-trivial characters of $\operatorname{Aff}(\mathbf{F}_p)$ if p is large, obtained by composition

$$\operatorname{Aff}(\mathbf{F}_p)\to\mathbf{F}_p^\times \xrightarrow{\chi} \mathbf{C}^\times,$$

hence $\operatorname{Aff}(\mathbf{F}_p)$ is only 1-quasirandom (if $p\geqslant 3$).

Concretely, pick some $\varepsilon>0$ arbitrarily small. Fix a large integer $m\geqslant 1$. For p large enough, if we identify $\mathbf{F}_p^\times$ with the cyclic group $\mathbf{Z}/(p-1)\mathbf{Z}$ (using a fixed generator of $\mathbf{F}_p^\times$), and consider an interval I of length about p/m in $\mathbf{F}_p^\times$, then the subset

$$\mathrm{A} = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \in \operatorname{Aff}(\mathbf{F}_p) \mid a\in \mathrm{I},\quad b\in\mathbf{F}_p \right\}$$

has size about $|\operatorname{Aff}(\mathbf{F}_p)|/m$, thus $|\mathrm{A}|\geqslant p^{1-\varepsilon}$ if p is large enough, but satisfies $\mathrm{A}^{(m-1)}\neq \operatorname{Aff}(\mathbf{F}_p)$, simply because we even have $\mathrm{I}^{(m-1)}\neq\mathbf{F}_p^\times$.

The same remark on the dependency of δ and γ applies to the sum-product theorem in $\mathbf{F}_p$ also, as was shown by Chang, and explained in the next exercise.

EXERCISE 3.2.12. Let p be a prime number. Fix a generator τ of the group $\mathbf{F}_p^\times$. Let $N \leqslant p$ be a positive integer.

(1) Let M be an integer such that $(pN)^{1/2} \leqslant M \leqslant 2(pN)^{1/2}$. Show that there exists an integer L such that

$$|\tau^{[M]} \cap (L + [M] \pmod p)| \gg N.$$

(2) Deduce that there exists a subset $A \subset \mathbf{F}_p$ such that $|A| \gg N$ and

$$\max(|A + A|, |A \cdot A|) \ll p^{1/2}|A|^{1/2}.$$

(3) Deduce an upper-bound for the possible exponent δ in Theorem 3.2.1 when $|A| \leqslant p^{1-\gamma}$ and γ is close to 1.

EXERCISE 3.2.13. Let p be a prime number. Let A_1, A_2 be subsets of $\mathbf{F}_p^\times$ and $A_3 \subset \mathbf{F}_p$. Let $G = \mathbf{F}_p^\times \times \mathbf{F}_p$ and consider the subsets

$$B = \{(x, x) \mid x \in A_1\} \subset G, \qquad C = A_2 \times A_3 \subset G.$$

(1) Show that $|B \star C| \leqslant |A_1 A_2||A_1 + A_3|$, where $\star$ refers to the group law in G.

(2) Find a suitable Sidon set $A \subset G$ such that $|A \cap B| = |B|$.

(3) Deduce that there exists a constant $c > 0$ such that

$$\max(|A_1 + A_3|, |A_1 A_2|) \geqslant c \min((|A_1|p)^{1/2}, |A_1|(|A_2||A_3|p^{-1})^{1/2})$$

(Use Exercise 2.3.21).

(4) What does this result, and that of the previous exercise, imply for the sum-product problem in $\mathbf{F}_p$?

The result of this exercise is due Cilleruelo [17, Th. 3.1]; the special case where $A_1 = A_2 = A_3$ was proved first by Garaev [39, Th. 1].

3.3. Approximate rings

We first prove Proposition 3.2.5. The argument follows the original one. We will use the notation $A \sqsubset_\alpha B$ for subsets of a group G, as in Definition 2.4.14, as well as the simple properties explained in Remark 2.4.16. Recall that $A \sqsubset_\alpha B$ means that there exists a subset $X \subset G$ with $|X| \leqslant \alpha$ such that $A \subset XB$.

LEMMA 3.3.1. *Let* R *be a commutative ring with unit. Let* A *be a finite subset of* R *containing* 1 *and* $\alpha \geqslant 1$ *a real number such that*

$$|A \cdot A - A \cdot A| \leqslant \alpha|A|.$$

(1) *We have* $A - A + A - A \sqsubset_{\alpha^5} A - A$.

(2) *We have* $A \cdot A \sqsubset_\alpha A - A$, *and in particular* $aA \sqsubset_\alpha A - A$ *For any* $a \in A$.

(3) *Let* x *and* $y \in R$, *and let* $\beta \geqslant 1$, $\gamma \geqslant 1$ *be such that* $xA \sqsubset_\beta A - A$ *and* $yA \sqsubset_\gamma A - A$. *Then we have*

$$(x + y)A \sqsubset_{\alpha^5\beta\gamma} A - A,$$
$$xyA \sqsubset_{\alpha^5\beta^2\gamma} A - A.$$

PROOF. (1) Since $1 \in A$, we have $A - A \subset A \cdot A - A \cdot A$, so that the assumption implies the bound $|A - A| \leqslant \alpha|A|$. The result is then a consquence of Plünnecke's Theorem in the form of Proposition 2.4.17.

(2) Since $1 \in A$, we have $|A \cdot A - A| \leqslant |A \cdot A - A \cdot A| \leqslant \alpha|A|$, so Ruzsa's Covering Lemma (Lemma 2.4.12) implies $A \cdot A \sqsubset_\alpha A - A$.

(3) By the elementary properties of Remark 2.4.16, combined with (1), the assumption implies that

$$(x+y)A \subset xA + yA \sqsubset_{\beta\gamma} A - A + A - A \sqsubset_{\alpha^5\beta\gamma} A - A.$$

There exist sets X and Y with $|X| \leqslant \beta$ and $|Y| \leqslant \gamma$ such that $xA \subset A - A + X$ and $yA \subset A - A + Y$. Then

$$xyA \subset A - A + A - A + xY + X - X,$$

so that (since $|xY + X - X| \leqslant \beta^2\gamma$), we get

$$xyA \sqsubset_{\beta^2\gamma} A - A + A - A \sqsubset_{\alpha^5\beta^2\gamma} A - A,$$

which concludes the proof. □

Proposition 3.2.5 follows easily from the following claim.

Proposition 3.3.2. *Let R be a commutative ring with unit. Let A be a finite subset of R containing 1 and $\alpha \geqslant 1$ a real number such that*

$$|A \cdot A - A \cdot A| \leqslant \alpha|A|.$$

Let $f \in \mathbf{Z}[X_1, \dots, X_d]$ be a polynomial. We have

$$f_*(A) \sqsubset_\beta A$$

where $\beta \leqslant c\alpha^d$ for some constants c, d depending only on f.

Proof. Using the first part of Lemma 3.3.1, it is enough to prove the result when $f = X^k$ for some integer $k \geqslant 0$. We proceed by induction on $k \geqslant 0$. The case $k = 0$ and $k = 1$ are elementary.

Suppose that $k \geqslant 2$ and that $A^k \sqsubset_\beta A - A$. Let X be a set with $|X| \leqslant \beta$ such that $A^k \subset A - A + X$ and $X \subset A^k - A + A$ (which we may assume). We have

$$A^k \subset A \cdot A - A \cdot A + A \cdot X.$$

By Lemma 3.3.1, (3), every $x \in X \subset A^k - A + A$ satisfies $xA \sqsubset_\gamma A - A$ for some $\gamma \leqslant c_1\alpha^{d_1}$, where c_1 and d_1 depend only on k. Writing

$$xA \subset A - A + Y_x, \qquad |Y_x| \leqslant \gamma,$$

for each $x \in X$, we get

$$A \cdot X \subset A - A + \bigcup_{x \in X} Y_x,$$

hence $A \cdot X \sqsubset_{|X|\gamma} A - A$. Consequently

$$A^k \sqsubset_{\beta\gamma} A \cdot A - A \cdot A + A - A,$$

and by Lemma 3.3.1, (2), this gives $A^k \sqsubset_{\alpha^2\beta\gamma} A - A + A - A + A - A$, so that we finish the induction using Lemma 3.3.1, (1) twice. □

Proof of Proposition 3.2.5. From the previous proposition, it only remains to estimate the size of $f_*(A) \cdot f_*(A)^{-1}$ for $f \in \mathbf{Z}[X_1, \dots, X_d]$, polynomially in terms of α such that $|A \cdot A - A \cdot A| \leqslant \alpha|A|$ (for a set A containing 1).

Denoting $B = f_*(A) - \{0\}$, we have

$$|f_*(A) \cdot f_*(A)^{-1}| \leqslant 1 + |B \cdot B^{-1}|,$$

where the 1 accounts for the possible appearance of 0 in the product set on the left-hand side. By Plünnecke's Theorem (Theorem 2.4.13 applied to the subset B of the group $\mathbf{F}_p^\times$) we have $|\mathrm{B}\cdot\mathrm{B}^{-1}|\leqslant\beta^2|\mathrm{B}|$, with β such that $|\mathrm{B}\cdot\mathrm{B}|\leqslant\beta|\mathrm{B}|$. By Proposition 3.3.2 (applied to the polynomial f_1f_2, where $f_1=f$ and f_2 results from f by replacing all variables by "new ones"), we have such a bound with $\beta\leqslant c\alpha^d$ for some $c\geqslant 0$ and $d\geqslant 1$ depending only on f. Hence

$$|f_*(\mathrm{A})\cdot f_*(\mathrm{A})^{-1}|\leqslant 1+c\alpha^d|\mathrm{A}|\leqslant(1+c)\alpha^d|\mathrm{A}|$$

since $|\mathrm{A}|\geqslant 1$, which concludes the proof. □

We now prove Proposition 3.2.3. We follow an argument of Bourgain, as presented by Tao and Vu [84, Lemma 2.53]. Remarkably, it will be (almost) possible to describe explicitly the set B.

Proof of Proposition 3.2.3. We assume that $0\notin\mathrm{A}$, which is certainly allowed. Writing here $x\mathrm{A}$ for the dilate of A by multiplication by some $x\in\mathbf{F}_p$, we claim that there exists $a_0\in\mathrm{A}$ such that

$$\sum_{a\in\mathrm{A}}|a\mathrm{A}\cap a_0\mathrm{A}|\geqslant\alpha^{-1}|\mathrm{A}|^2, \tag{3.13}$$

and that the set

$$\mathrm{B}=\{a\in\mathrm{A}\mid |a\mathrm{A}\cap a_0\mathrm{A}|\geqslant\tfrac{1}{2}\alpha^{-1}|\mathrm{A}|\}-\{a_0\}$$

then satisfies the required properties (3.8).

To prove (3.13), it suffices to prove that

$$\sum_{a,b\in\mathrm{A}}|a\mathrm{A}\cap b\mathrm{A}|\geqslant\frac{|\mathrm{A}|^3}{\alpha},$$

and this follows from the basic multiplicative energy inequality (2.15): indeed, the representation function r_{A} for $\mathrm{A}\cdot\mathrm{A}$ satisfies

$$\mathrm{E}(\mathrm{A},\mathrm{A})=\sum_x r_{\mathrm{A}}(x)^2=\sum_{x\in\mathrm{G}}\sum_{\substack{ac=x\\a,c\in\mathrm{A}}}\sum_{\substack{bd=x\\b,d\in\mathrm{A}}}1=\sum_{a,b\in\mathrm{A}}|a\mathrm{A}\cap b\mathrm{A}|,$$

so we obtain

$$\sum_{a,b\in\mathrm{A}}|a\mathrm{A}\cap b\mathrm{A}|=\mathrm{E}(\mathrm{A},\mathrm{A})\geqslant\frac{|\mathrm{A}|^4}{|\mathrm{A}\cdot\mathrm{A}|}\geqslant\frac{|\mathrm{A}|^3}{\alpha}$$

using the assumption $|\mathrm{A}\cdot\mathrm{A}|\leqslant\alpha\mathrm{A}$.

Once (3.13) is established, note that it implies that

$$|\mathrm{B}|\geqslant\frac{|\mathrm{A}|}{2\alpha}-1\geqslant\frac{|\mathrm{A}|}{4\alpha},$$

provided $|\mathrm{A}|\geqslant 4\alpha$, which we may also assume. This proves that B satisfies the first condition of (3.8), and it remains to check that

$$|\mathrm{B}\cdot\mathrm{B}-\mathrm{B}\cdot\mathrm{B}|\ll\alpha^d|\mathrm{B}|$$

for some constant d.

To do this, we introduce some notation. First, let ω be the product of all elements of A; we have $\omega\in\mathbf{F}_p^\times$ since we assumed that $0\notin\mathrm{A}$. Define then $\mathrm{C}=\mathrm{B}\cdot\mathrm{B}\cdot\omega\mathrm{B}^{-1}=\omega(\mathrm{B}\cdot\mathrm{B}\cdot\mathrm{B}^{-1})$. We first prove that if c_1 and c_2 are in C, then we have

$$|c_1\mathrm{A}-c_2\mathrm{A}|\ll\alpha^d|\mathrm{A}| \tag{3.14}$$

for some constant $d \geqslant 0$. To see this, we use the Ruzsa distance, in the additive group of $\mathbf{F}_p$, since by definition we have $|c_1\mathrm{A} - c_2\mathrm{A}| = |\mathrm{A}|\log(d(c_1\mathrm{A}, c_2\mathrm{A}))$, so that we need to bound the Ruzsa distance between $c_1\mathrm{A}$ and $c_2\mathrm{A}$.

This is done in multiple steps, and to simplifty notation we define a function

$$f\colon \mathbf{F}_p^2 \to \mathbf{R}$$

by $f(x,y) = d(x\mathrm{A}, y\mathrm{A})$. We first note explicitly some easy consequences of the definition of the Ruzsa distance. First, if $\mathrm{X} \subset \mathrm{Y}$ are non-empty subsets of a group G (with multiplicative notation) such that with $|\mathrm{Y}| \leqslant \beta|\mathrm{X}|$, then

$$d(\mathrm{X},\mathrm{Y}) = \log\Big(\frac{|\mathrm{X}\cdot\mathrm{Y}^{-1}|}{\sqrt{|\mathrm{X}||\mathrm{Y}|}}\Big) \leqslant \log\Big(\frac{\beta|\mathrm{Y}\cdot\mathrm{Y}|}{|\mathrm{X}|}\Big). \tag{3.15}$$

Second, if $b \in \mathbf{F}_p^\times$, then for any non-empty subsets X and Y of $\mathbf{F}_p$, we have $d(b\mathrm{X}, b\mathrm{Y}) = d(\mathrm{X},\mathrm{Y})$. In particular, this implies that $f(bx,by) = f(x,y)$ for any $(x,y) \in \mathbf{F}_p^2$.

We can then begin the proof of (3.14). First, the assumption $|\mathrm{A}+\mathrm{A}| \leqslant \alpha|\mathrm{A}|$ implies $d(\mathrm{A},-\mathrm{A}) \leqslant \log(\alpha)$, and by symmetry and the triangle inequality, it follows that $d(\mathrm{A},\mathrm{A}) \leqslant 2\log(\alpha)$. Next, for $a \in \mathrm{B}$, we apply (3.15) to $a\mathrm{A}\cap a_0\mathrm{A} \subset a\mathrm{A}$ and $a\mathrm{A}\cap a_0\mathrm{A} \subset a_0\mathrm{A}$, where we can take $\beta = 2\alpha$ by definition of B; we deduce the bound

$$f(a,a_0) = d(a\mathrm{A}, a_0\mathrm{A}) \leqslant d(a\mathrm{A}, a\mathrm{A}\cap a_0\mathrm{A}) + d(a\mathrm{A}\cap a_0\mathrm{A}, a_0\mathrm{A}) \leqslant 2\log(2\alpha^2),$$

using again the assumptions (in the form $|a\mathrm{A}+a\mathrm{A}| = |a_0\mathrm{A}+a_0\mathrm{A}| = |\mathrm{A}+\mathrm{A}| \leqslant \alpha|\mathrm{A}|$).

Now we proceed to bound $f(a_1a_2, a_0^2)$ for any a_1 and a_2 in B. Using the previous bound, invariance under dilation twice and the triangle inequality, we have

$$f(a_1a_2, a_0^2) \leqslant f(a_1a_2, a_1a_0) + f(a_1a_0, a_0^2) \leqslant 4\log(2\alpha^2).$$

Next, we bound $f(a_1a_2\omega^{-1}a_3, a_0^2\omega a_0^{-1})$ for a_1, a_2 and a_3 in B in a similar way:

$$\begin{aligned} f(a_1a_2\omega a_3^{-1}, a_0^2\omega a_0^{-1}) &\leqslant f(a_1a_2\omega a_3^{-1}, a_1a_2\omega a_0^{-1}) + f(a_1a_2\omega a_0^{-1}, a_0^2\omega a_0^{-1}) \\ &\leqslant f(ba_3, ba_0) + f(ca_1a_2, ca_0^2) \end{aligned}$$

where $b = a_1a_2\omega a_0^{-1}a_3^{-1}$ and $c = \omega a_0^{-1}$, so

$$f(a_1a_2\omega a_3^{-1}, a_0^2\omega a_0^{-1}) \leqslant 6\log(2\alpha^2).$$

Recalling that $\mathrm{C} = \mathrm{B}\cdot\mathrm{B}\cdot\omega\mathrm{B}^{-1}$, a last application of the triangle inequality gives

$$f(c_1,c_2) \leqslant 12\log(2\alpha^2)$$

for any c_1 and c_2 in C, which concludes the proof of (3.14). This, in turn, immediately gives

$$\sum_{c_1,c_2\in\mathrm{C}} |c_1\mathrm{A} - c_2\mathrm{A}| \ll \alpha^d|\mathrm{A}||\mathrm{C}|^2 \ll \alpha^{12+d}|\mathrm{A}|^3 \ll \alpha^{15+d}|\mathrm{B}|^3,$$

after noting that $|\mathrm{B}\cdot\mathrm{B}| \leqslant |\mathrm{A}\cdot\mathrm{A}| \leqslant \alpha|\mathrm{A}| \leqslant 4\alpha^2|\mathrm{B}|$ and using Plünnecke's Theorem in $\mathbf{F}_p^\times$ (see Theorem 2.4.13) to bound the size of C.

The final flourish is the observation that

$$|\mathrm{B}|^2|\mathrm{B}\cdot\mathrm{B} - \mathrm{B}\cdot\mathrm{B}| = |\mathrm{B}|^2|\omega(\mathrm{B}\cdot\mathrm{B} - \mathrm{B}\cdot\mathrm{B})| \leqslant \sum_{c_1,c_2\in\mathrm{C}} |c_1\mathrm{A} - c_2\mathrm{A}|.$$

Indeed, this is due to the fact that for any $(a_1,\dots,a_4) \in \mathrm{B}^4$, and for any $(a,b) \in \mathrm{A}^2$, we have

$$\omega(a_1a_2 - a_3a_4) = a_1a_2\omega a^{-1}a - a_3a_4\omega b^{-1}b,$$

which means that $\omega(a_1a_2 - a_3a_4) \in c_1\mathrm{A} - c_2\mathrm{A}$, with $c_1 = a_1a_2\omega^{-1}a$ and $c_2 = a_3a_4\omega^{-1}b$, and thus this element is counted for $\geqslant |\mathrm{B}|^2$ values of c_1 and c_2 in the right-hand sum.

Combining this with the previous step finishes the proof. □

3.4. Applications of the sum-product theorem

The work of Bourgain, Katz and Tao was motivated at least in part by a number of applications which they deduced from Theorem 3.2.1. Moreover, many more applications, sometimes quite surprising, have appeared in the following years. We briefly discuss some examples here; in the next section, we will give full details of the proof of one of these results.

Incidence bound. Strikingly, Bourgain, Katz and Tao manage to *reverse* the argument of Elekes deriving a sum-product estimate from the incidence bound of Szemerédi and Trotter, obtaining a non-trivial incidence bound over finite fields from the sum-product theorem. They prove (see [11, Th. 6.2]):

THEOREM 3.4.1. *Let p be a prime number. Let P be a finite subset of $\mathbf{F}_p^2$ and L a finite set of affine lines $\ell \subset \mathbf{F}_p^2$. Assume that*

$$|\mathrm{L}| \leqslant |\mathrm{P}| \leqslant p^{\gamma}$$

where $0 < \gamma < 2$.

Let $k \geqslant 2$ be an integer such that $k \leqslant |\mathrm{P}|^{1/2}$. If the lower bound $|\ell \cap \mathrm{P}| \geqslant k$ holds for all $\ell \in \mathrm{L}$, then we have

$$|\mathrm{L}| \ll |\mathrm{P}|^2 k^{-2-\delta},$$

for some $\delta > 0$ depending only on γ, where the implied constant also only depends on γ.

REMARK 3.4.2. The precise statement in [11] is a bit different, but this version is a direct consequence. Precisely, from the result of Bourgain, Katz and Tao, one gets

$$k|\mathrm{L}| \ll \max(|\mathrm{P}|, |\mathrm{L}|)^{3/2-\delta}$$

for some $\delta > 0$ depending only on γ. Under our assumptions on the sizes of L, P and on k, this gives

$$|\mathrm{L}| \ll |\mathrm{P}|^{3/2-\varepsilon} k^{-1} \ll |\mathrm{P}|^{2-\delta} k^{-2} \ll |\mathrm{P}|^2 k^{-2-2\delta}.$$

Theorem 3.4.1 has been improved significantly, and in fact the most recent bounds are obtained independently of the sum-product theorem, which then permits the use of the method of Elekes also over finite fields to deduce improved versions of Theorem 3.2.1. Stevens and de Zeeuw prove, for instance, an incidence bound which implies that

$$\min(|\mathrm{A} + \mathrm{A}|, |\mathrm{A} \cdot \mathrm{A}|) \gg |\mathrm{A}|^{6/5}$$

for $|\mathrm{A}| \ll p^{5/8}$; this result had first been obtained, using a different method, by Roche–Newton, Rudnev and Shkredov [67], and this recovers the Elekes estimate for integers.

EXERCISE 3.4.3. Let p be a prime number. Let $\mathrm{P} \subset \mathbf{F}_p^2$ be a set of points and L a set of affine lines in $\mathbf{F}_p^2$. Assume that all lines are given by an equation $y = ax + b$ with $a \neq 0$ and that all $(u, v) \in \mathrm{P}$ satisfy $u \neq 0$.

(1) Find a large Sidon subset $\mathrm{A} \subset \mathbf{F}_p^\times \times \mathbf{F}_p$ and subsets B, $\mathrm{C} \subset \mathbf{F}_p^\times \times \mathbf{F}_p$ such that

$$|\{(b, c) \in \mathrm{B} \times \mathrm{C} \mid b + c \in \mathrm{A}\}| = |\{(p, \ell) \in \mathrm{P} \times \mathrm{L} \mid p \in \ell\}|.$$

(Hint: write the equations of the lines in the form $y = ax + b$ and the coordinates of the points as (u, v), and interpret the equation $au + b = v$.)

(2) Deduce from this and from the Exercise 2.3.21 that

$$|\{(p,\ell) \in \mathrm{P} \times \mathrm{L} \mid p \in \ell\}| = \frac{|\mathrm{P}||\mathrm{L}|}{p} + \mathrm{O}(p^{1/2}\sqrt{|\mathrm{P}||\mathrm{L}|}).$$

(3) When is this result (due to Cilleruelo [17, Th. 2.2]) interesting?

Bounds for Besicovitch sets over finite fields. A rather remarkable result of Besicovitch states that for $d \geqslant 2$, there exist subsets $\mathrm{B} \subset \mathbf{R}^d$ of Lebesgue measure 0 which contain a segment of length one in *any* direction. A question, raised by Kakeya, is whether such sets can also have small Hausdorff dimension, and it is expected that that answer should be No: the Hausdorff dimension of a *Besicovitch set* in $\mathbf{R}^d$ should be equal to 1. This problem turns out to have deep relevance to problems in harmonic analysis (see, for instance, the surveys of Tao [83] and Łaba [62] for accessible overviews, as well as the more recent discussion by Guth [47, Ch. 15]).

One of the applications of the sum-product theorem in the original paper of Bourgain, Katz and Tao was a lower bound for the analogue of the 3-dimensional version of this question over finite fields, namely the fact that for p prime, a subset $\mathrm{B} \subset \mathbf{F}_p^3$ which contains an affine line in every direction has cardinality $\gg p^{5/2+\delta}$ for some $\delta > 0$ independent of p.

This result was strikingly improved by Dvir [23] who, with completely different methods, actually solved the problem over finite fields. Precisely, for a prime p and an integer $d \geqslant 1$, we say that $\mathrm{B} \subset \mathbf{F}_p^d$ is a Besicovitch set if it contains a line in very direction, i.e., if for any $\xi \in \mathbf{F}_p^d - \{0\}$, there exists $x \in \mathbf{F}_p^d$ with $x + \mathbf{F}_p\,\xi \subset \mathrm{B}$.

THEOREM 3.4.4 (Dvir). *Let $d \geqslant 2$ be an integer. For any prime number p and any Besicovitch subset $\mathrm{B} \subset \mathbf{F}_p^d$, the lower bound*

$$|\mathrm{B}| \geqslant \frac{(p-1)^{d-1}}{(d-1)!}$$

holds. In fact, for any $\varepsilon > 0$, we have

$$|\mathrm{B}| \gg p^{d-\varepsilon},$$

where the implied constant depends only on d and ε.

Although the proof is very far from the sum-product phenomenon, we present it as an illustration of the variety of tools available to study problems related to additive combinatorics (more generally, this is an instance of the "polynomial method" in combinatorics; see for instance the book of Guth [47]).

PROOF. For any integers $k \geqslant 1$ and $d \geqslant 1$, we denote by $\mathscr{X}_{d,k}$ the vector space of homogeneous polynomials of degree k in d variables, with coefficients in $\mathbf{F}_p$. We will show that any Besicovitch set $\mathrm{B} \subset \mathbf{F}_p^d$ satisfies

$$|\mathrm{B}| \geqslant \dim \mathscr{X}_{d,p-2}. \tag{3.16}$$

Using the general formula

$$\dim \mathscr{X}_{d,k} = \binom{d+k-1}{d-1}$$

(see Proposition A.4.1), we conclude then that

$$|\mathrm{B}| \geqslant \frac{(p+d-3)\cdots(p-1)}{(d-1)!} \geqslant \frac{(p-1)^{d-1}}{(d-1)!}$$

which gives the first result.

The proof of (3.16) is by contradiction. In fact, let $k \geqslant 1$ be any integer such that $|\mathrm{B}| < \dim \mathscr{X}_{d,k}$. If we note that the conditions $f(b) = 0$, for all $b \in \mathrm{B}$, can be seen as *linear equations* for the coefficients of f, and that there are then $\dim \mathscr{X}_{d,k}$ unknowns and $|\mathrm{B}|$ equations, it follows from linear algebra that we can then find a non-zero polynomial $f \in \mathscr{X}_{d,k}$ such that $f(b) = 0$ for all $b \in \mathrm{B}$.

Let $\xi \in \mathbf{F}_p^d - \{0\}$. By assumption, we can find some $x \in \mathbf{F}_p^d$ such that the affine line

$$x + \mathbf{F}_p\xi = \{x + a\xi \mid a \in \mathbf{F}_p\}$$

is contained in B; we then define the one-variable polynomial $f_\xi = f(x\mathrm{X} + \xi)$, which has degree $\leqslant k$. Then, for any $a \in \mathbf{F}_p^\times$, we have

$$f_\xi(a) = f(ax + \xi) = f(a(x + a^{-1}\xi)) = a^k f(x + a^{-1}\xi) = 0,$$

since f is homogeneous of degree k and vanishes on $x + \mathbf{F}_p\xi$. Hence f_ξ has at least $p - 1$ zeros. If $k < p - 1$, so that $\deg(f_\xi) < p - 1$, this means that f_ξ is the zero polynomial. In particular, we then get $f_\xi(0) = f(\xi) = 0$. This is true for all non-zero $\xi \in \mathbf{F}_p^d$, but in addition $f(0) = 0$ because f is homogeneous of degree $k \geqslant 1$, so f in fact vanishes on all of $\mathbf{F}_p^d$. But the Schwarz–Zippel Lemma (see Proposition A.4.2) shows that the number of zeros of f in $\mathbf{F}_p^d$ is at most kp^{d-1}. If $k = p - 2$ (in particular, $k < p - 1$, so the above applies), we get a contradiction since $kp^{d-1} = p^d - 2p^{d-1} < p^d$.

We deduce the more precise statement using a trick: given $\varepsilon > 0$, let $r = \lceil \varepsilon^{-1} \rceil$. Note that if $\mathrm{B} \subset \mathbf{F}_p^d$ is a Besicovitch set, then the set B^r is a Besicovitch set in $\mathbf{F}_p^{dr}$ (indeed, let $\xi = (\xi_1, \dots, \xi_r) \in \mathbf{F}_p^{rd} - \{0\}$; for i such that $\xi_i \neq 0$, let $y_i \in \mathbf{F}_p^d$ be such that $y_i + \mathbf{F}_p\xi_i \subset \mathrm{B}$, and otherwise let y_i be an arbitrary element of B; then the line

$$(y_1, \dots, y_r) + \mathbf{F}_p\xi$$

with direction ξ is contained in B^r). The first result, applied to B^r in $\mathbf{F}_p^{rd}$, implies that

$$|\mathrm{B}| \geqslant \frac{(p-1)^{d-1/r}}{((dr-1)!)^{1/r}} \gg p^{d-\varepsilon},$$

where the implied constant depends only on d and ε. □

Exercise 3.4.5. Let p be a prime number, $d \geqslant 1$ an integer. Let $\mathrm{B} \subset \mathbf{F}_p^d$ be a Besicovitch set.

(1) If $|\mathrm{B}| \leqslant \binom{p+d-1}{d}$, show that there exists a non-zero polynomial $f \in \mathbf{F}_p[\mathrm{X}_1, \dots, \mathrm{X}_d]$, of degree $k \leqslant p - 1$, such which vanishes on B. (This polynomial need not be homogeneous.)

(2) Let f_d denote the homogeneous component of f of degree d. Show that for any $\xi \in \mathbf{F}_p^d$, we have $f_d(\xi) = 0$. (Hint: consider $f(x + \mathrm{X}\xi) \in \mathbf{F}_p[\mathrm{X}]$.)

(3) Conclude that the assumption on $|\mathrm{B}|$ cannot be satisfied and that $|\mathrm{B}| \gg p^d$, where the implied constant depends only on d.

(This improvement of Dvir's Theorem, which is now essentially best possible, was found by Alon and Tao, independently.)

Exercise 3.4.6. Let p be a prime number and $d \geqslant 1$ an integer. A subset $\mathrm{N} \subset \mathbf{F}_p^n$ is called a *Nikodym set* if, for every $x \in \mathbf{F}_p^d$, there exists an affine line $\ell_x \subset \mathbf{F}_p^d$ such that $x \in \ell_x$ and $\ell_x - \{x\} \subset \mathrm{N}$.

Prove that a Nikodym set $\mathrm{N} \subset \mathbf{F}_p^d$ satisfies $|\mathrm{N}| \gg p^d$, where the implied constant depends only on d (this result is also due to Dvir).

Bounds for exponential sums. An extremely important consequence of the sum-product theorem over finite fields was obtained by Bourgain, Glibichuk and Konyagin [10]. In its simplest form, it states that the Fourier coefficients of the characteristic function of a *multiplicative* subgroup of $\mathbf{F}_p^\times$ are small. Here, we recall that although $\mathbf{F}_p^\times$ is group-theoretically relatively "simple", since it is cyclic of order $p-1$, it may have many subgroups, of order given by the divisors of $p-1$. Since it is known that $p-1$ is, in many respects, a "typical" integer (e.g., as far as the number of prime factors, or the number of divisors, is concerned, among other things), such subgroups may have a wide variety of sizes.

THEOREM 3.4.7 (Bourgain, Glibichuk and Konyagin). *Let p be a prime number and let $\gamma > 0$ be a real number. There exists $\nu > 0$, depending only on γ, such that if $\mathrm{H} \subset \mathbf{F}_p^\times$ is a subgroup of $\mathbf{F}_p^\times$ with $|\mathrm{H}| \geqslant p^\gamma$, then we have*

$$\sum_{x\in \mathrm{H}} e\Big(\frac{ax}{p}\Big) \ll |\mathrm{H}|p^{-\nu}$$

for any $a \in \mathbf{F}_p^\times$, where the implied constant depends only on γ.

This result is another incarnation of the fact that a *multiplicatively structured* set (here, a multiplicative subgroup) should have little *additive structure*, here measured by the size of the discrete Fourier transform. The remarkable point of the theorem is the very weak assumption on the size of H: the well-established methods of number theory and harmonic analysis over finite fields are directly applicable to the problem, but they only succeed if H is quite large (for instance, when $|\mathrm{H}| \geqslant p^{1/2+\delta}$ for some $\delta > 0$, see Exercise 3.4.10 below).

COROLLARY 3.4.8. *Let $\gamma > 0$ be a real number Let p be a prime number and let $d \mid p-1$ be a divisor of $p-1$ such that $d \leqslant (p-1)p^{-\gamma}$. There exists $\nu > 0$, depending only on γ, such that*

$$\sum_{x\in \mathbf{F}_p} e\Big(\frac{ax^d}{p}\Big) = \mathrm{O}(p^{1-\nu})$$

for any $a \in \mathbf{F}_p^\times$.

PROOF. This is essentially a reformulation of Theorem 3.4.7, applied to the subgroup

$$\mathrm{H} = \{x^d \mid x \in \mathbf{F}_p^\times\}$$

which has order $(p-1)/d$. More precisely, since each $y \in \mathrm{H}$ is of the form $y = x^d$ for d different values of $x \in \mathbf{F}_p^\times$, we get

$$\sum_{x\in \mathbf{F}_p} e\Big(\frac{ax^d}{p}\Big) = 1 + \sum_{x\in \mathbf{F}_p^\times} e\Big(\frac{ax^d}{p}\Big) = 1 + d\sum_{y\in \mathrm{H}} e\Big(\frac{ay}{p}\Big) \ll 1 + (p-1)p^{-\nu}.$$

□

REMARK 3.4.9. One can wonder about even smaller subgroups, but some restriction is certainly needed since H could be of bounded order. For instance, if p is odd, there is always a subgroup of order 2, namely $\{-1, 1\}$, for which the behavior of the sums is quite clearly rather different. We note that papers of Garcia, Hyde and Lutz [40] and Duke, Garcia and Lutz [22], among others, have showed that some interesting statistical behavior can be observed in the context of sums over roots of unity of fixed order d, for primes $p \equiv 1 \pmod d$ (which ensures that $\mathbf{F}_p^\times$ contain all the d-th roots of unity).

EXERCISE 3.4.10. Let p be a prime number and let $\mathrm{H} \subset \mathbf{F}_p^\times$ be a multiplicative subgroup.

(1) For any character χ of $\mathbf{F}_p^\times$, show that

$$\sum_{x\in\mathrm{H}} \chi(x) = \begin{cases} |\mathrm{H}| & \text{if } \chi(x) = 1 \text{ for all } x \in \mathrm{H} \\ 0 & \text{otherwise,} \end{cases}$$

and that the number of χ such that the first case holds is equal to $(p-1)/|\mathrm{H}|$.

(2) Show that the characteristic function φ_{H} of H satisfies

$$\varphi_{\mathrm{H}}(x) = \frac{|\mathrm{H}|}{p-1} \sum_{\chi\in\mathrm{H}^\perp} \chi(x),$$

where $\mathrm{H}^\perp$ is the set of characters such that $\chi(x) = 1$ for all $x \in \mathrm{H}$.

(3) Let χ be a character of $\mathbf{F}_p^\times$ and let $a \in \mathbf{F}_p$. The sum

$$\tau(\chi, a) = \sum_{x\in\mathbf{F}_p^\times} \chi(x) e\Big(\frac{ax}{p}\Big)$$

is called a *Gauss sum*. Prove that $\tau(\chi, a) = -1$ if χ is trivial and $a \neq 0$, and that

$$|\tau(\chi, a)|^2 = p$$

if χ is non-trivial and $a \neq 0$.

(4) Deduce that if $a \in \mathbf{F}_p^\times$, we have

$$\Big|\sum_{x\in\mathrm{H}} e\Big(\frac{ax}{p}\Big)\Big| \leqslant \sqrt{p}.$$

(5) When is the last bound interesting?

3.5. Exponential sums and random walks

We will now prove Theorem 3.4.7 in detail, following essentially the account by Kurlberg [61]. We first outline the main steps and intermediate statements, before proving the latter.

Step 1. The key result is a statement according to which the existence of certain probability measures on $\mathbf{F}_p$ lead to the existence of subsets with small sum and product sets. We will state the result in a more probabilistic language than either the original paper of [61]. This requires some notation.

Given an $\mathbf{F}_p$-valued random variable X (defined on some probability space Ω which we do not need to specify), we denote by $(\mathrm{X}_1, \mathrm{X}_2)$ a couple of independent random variables, each distributed like X, and we define $\mathrm{Y} = \mathrm{X}_1 - \mathrm{X}_2$. We define then

$$\mathrm{N}_{\mathrm{X}} = \sum_{x\in\mathbf{F}_p} \mathbf{P}(\mathrm{X} = x)^2 = \mathbf{P}(\mathrm{Y} = 0).$$

The main statement of the first step is the following:

PROPOSITION 3.5.1. *Let p be a prime number. Let X be an $\mathbf{F}_p$-valued random variable, and let Y be as above. Define*

$$\phi(y) = \mathbf{P}(\mathrm{Y} = y)$$

for $y \in \mathbf{F}_p$. Let $\alpha \geqslant 1$ be a real number such that

$$\mathbf{E}(\phi(\mathrm{XY}))) \geqslant \frac{\phi(0)}{\alpha}. \tag{3.17}$$

Assuming that

$$\mathbf{P}(\mathrm{X}=0) \leqslant \frac{1}{4\alpha}, \qquad \mathbf{P}(\mathrm{Y}=0) \leqslant \frac{1}{4\alpha}, \tag{3.18}$$

there exists a subset $\mathrm{A} \subset \mathbf{F}_p^\times$ such that

$$\frac{1}{\alpha^d \phi(0)} \ll |\mathrm{A}| \ll \frac{\alpha}{\phi(0)}$$

with the property that

$$\max(|\mathrm{A}+\mathrm{A}|, |\mathrm{A}\cdot\mathrm{A}|) \ll \alpha^d |\mathrm{A}|,$$

where d is an absolute constant.

Remark 3.5.2. (1) Kurlberg [61, Prop. 3.1] shows that one can take here $d = 768$, and gives explicit values for all the implicit constants.

(2) The statement (and its proof) can also be presented (as in [61]) using measure-theoretic language. For instance, denoting

$$\mu(x) = \mathbf{P}(\mathrm{X}=x), \qquad \phi(y) = \mathbf{P}(\mathrm{Y}=y),$$

the assumption (3.17) is the requirement that

$$\sum_{x\in\mathbf{F}_p} \sum_{y\in\mathbf{F}_p} \mu(x)\phi(y)\phi(xy) \geqslant \alpha^{-1}.$$

Which language is more enlightening or transparent may depend on the reader – for those whose inclinations are less probabilistic, we can again recommend the presentation of Kurlberg, or the original papers.

Intuitively, if we rephrase the main assumption (3.17) in the form

$$\mathbf{E}\Big(\mathbf{P}(\mathrm{Y}=\mathrm{XY}')\Big) \geqslant \alpha^{-1}$$

for some random variable Y' distributed like Y, we see that it can interpreted as saying that, on average, there is a rather large probability that $\mathrm{X} = \mathrm{Y}_1\mathrm{Y}_2^{-1}$, with Y_1 and Y_2 distributed like Y, but conditioned to be non-zero. If Y_1 and Y_2 were uniformly distributed on some subset A of $\mathbf{F}_p^\times$, this would amount to saying that the average of the representation function $r_{\mathrm{A}\cdot\mathrm{A}^{-1}}(\mathrm{X})$ is large, and we could then use the Balog–Szemerédi–Gowers Theorem to extract from A a subset with small product set. The proof of Proposition 3.5.1 will start by showing that finding such an A is indeed possible even without such a normalizing assumption on Y, and then to show that the sumset $\mathrm{A}+\mathrm{A}$ is *also* under control.

Step 2. We now describe for which random variables we will apply Proposition 3.5.1. An elementary but crucial step is that the relevant quantity has a Fourier-analytic expression; using this, it will be possible to show that, assuming for contradiction that some of the exponential sums over H are "large", we obtain using Proposition 3.5.1 certain A which violate the sum-product theorem in $\mathbf{F}_p$.

We introduce first some necessary notation. For any $\mathbf{F}_p$-valued random variable X, we denote by φ_{X} the "characteristic function" of X (in the probabilistic sense, hence essentially its Fourier transform), namely the function on $\mathbf{F}_p$ defined by

$$\varphi_{\mathrm{X}}(a) = \mathbf{E}\Big(e\Big(\frac{ax}{p}\Big)\Big)$$

for $a \in \mathbf{F}_p$. We have $\varphi_{-\mathrm{X}} = \overline{\varphi_{\mathrm{X}}}$, and if X_1 and X_2 are independent, then

$$\varphi_{\mathrm{X}_1+\mathrm{X}_2} = \varphi_{\mathrm{X}_1}\varphi_{\mathrm{X}_2}.$$

We go back to the previous notation with a random variable X, the independent copies $(\mathrm{X}_1, \mathrm{X}_2)$ and $\mathrm{Y} = \mathrm{X}_1 - \mathrm{X}_2$. We have then $\varphi_{\mathrm{Y}} = |\varphi_{\mathrm{X}}|^2$. In particular, the characteristic function of Y is non-negative, and since $\varphi_{\mathrm{Y}}(0) = 1$, we can consider a random variable $\widehat{\mathrm{Y}}$ on $\mathbf{F}_p$, such that

$$\mathbf{P}(\widehat{\mathrm{Y}} = a) = \frac{\varphi_{\mathrm{Y}}(a)}{\mathrm{M}_{\mathrm{X}}} = \frac{|\varphi_{\mathrm{X}}(a)|^2}{\mathrm{M}_{\mathrm{X}}}$$

for $a \in \mathbf{F}_p$, where

$$\mathrm{M}_{\mathrm{X}} = \sum_{a\in\mathbf{F}_p} |\varphi_{\mathrm{X}}(a)|^2.$$

Moreover, we may (and do) insist that $\widehat{\mathrm{Y}}$ is independent from $(\mathrm{X}, \mathrm{X}_1, \mathrm{X}_2)$, hence also from Y. (Similarly, whenever we consider $\widehat{\mathrm{Z}}$ for some other random variable Z, it will be understood that $\widehat{\mathrm{Z}}$ is independent of any previously described random variables.)

The next lemma is the Fourier-theoretic description of $\mathbf{E}(\phi(\mathrm{XY}))$.

Lemma 3.5.3. *We have*

$$\phi(y) = \frac{\mathrm{M}_{\mathrm{X}}}{p}\varphi_{\widehat{\mathrm{Y}}}(y),$$

for any $y \in \mathbf{F}_p$ and

$$\mathbf{E}(\phi(\mathrm{XY})) = \phi(0)\,\mathbf{E}(|\varphi_{\mathrm{X}}(\mathrm{X}\widehat{\mathrm{Y}})|^2).$$

Proof. For any $y \in \mathbf{F}_p$, we have $\phi(y) = \mathbf{P}(\mathrm{Y} = y)$. We use the orthogonality of characters of $\mathbf{F}_p$ to represent the (set-theoretic!) characteristic function of y by

$$\frac{1}{p}\sum_{a\in\mathbf{F}_p} e\Big(\frac{a(x-y)}{p}\Big) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y, \end{cases}$$

and get

$$\phi(y) = \mathbf{E}\Big(\frac{1}{p}\sum_{a\in\mathbf{F}_p} e\Big(\frac{a(\mathrm{Y}-y)}{p}\Big)\Big) = \frac{1}{p}\sum_{a\in\mathbf{F}_p} e\Big(-\frac{ay}{p}\Big)\varphi_{\mathrm{Y}}(a) = \frac{\mathrm{M}_{\mathrm{X}}}{p}\varphi_{\widehat{\mathrm{Y}}}(-y).$$

This proves the first formula since $\phi(-y) = \phi(y)$. Since it implies in particular that $\phi(0) = \mathrm{M}_{\mathrm{X}}/p$, it further leads to

$$\mathbf{E}(\phi(\mathrm{XY})) = \phi(0)\,\mathbf{E}(\varphi_{\widehat{\mathrm{Y}}}(\mathrm{XY})),$$

and it only remains to appeal to the symmetry formula

$$\mathbf{E}(\varphi_{\widehat{\mathrm{Y}}}(\mathrm{XY})) = \mathbf{E}(|\varphi_{\mathrm{X}}(\mathrm{X}\widehat{\mathrm{Y}})|^2)$$

to conclude the proof. This last identity can be seen as a (very simple) instance of Fubini's formula:

$$\mathbf{E}(\varphi_{\widehat{\mathrm{Y}}}(\mathrm{XY})) = \mathbf{E}\Big(\mathbf{E}\Big(e\Big(\frac{\mathrm{XY}\widehat{\mathrm{Y}}}{p}\Big)\Big)\Big) = \mathbf{E}\Big(\mathbf{E}\Big(e\Big(\frac{\mathrm{X}(\mathrm{X}_1-\mathrm{X}_2)\widehat{\mathrm{Y}}}{p}\Big)\Big)\Big)$$
$$= \mathbf{E}\Big(\Big|\mathbf{E}\Big(e\Big(\frac{\mathrm{XX}_1\widehat{\mathrm{Y}}}{p}\Big)\Big)\Big|^2\Big) = \mathbf{E}(|\varphi_{\mathrm{X}}(\mathrm{X}_1\widehat{\mathrm{Y}}))|^2),$$

leading to the conclusion since X and X_1 are identically distributed. □

Remark 3.5.4. (1) Again, in concrete terms, with $\mu(x) = \mathbf{P}(\mathrm{X} = x)$ and $\phi(y) = \mathbf{P}(\mathrm{Y} = y)$, this statement means that

$$\sum_{x\in\mathbf{F}_p}\sum_{y\in\mathbf{F}_p}\mu(x)\phi(y)\phi(xy) = \frac{\mathrm{M}_{\mathrm{X}}}{p}\sum_{x\in\mathbf{F}_p}\sum_{a\in\mathbf{F}_p}\varphi_{\mathrm{X}}(ax)|\varphi_{\mathrm{X}}(a)|^2\mu(x).$$

This can be checked straightforwardly using elementary properties of the Fourier transform.

(2) The use of the random variable $\widehat{\mathrm{Y}}$ (which emphasizes values $a \in \mathbf{F}_p$ where $|\varphi_{\mathrm{X}}(a)|^2$ is "large") is reminiscent of the similar use of a non-uniform distribution in Schoen's proof of the Balog–Szemerédi–Gowers Theorem (see Theorem 2.7.1).

Given the subgroup $\mathrm{H} \subset \mathbf{F}_p^\times$, let S and $(\mathrm{S}_k)_{k\geqslant 1}$ be independent random variables all uniformly distributed on H (so that $\mathbf{P}(\mathrm{S}_j = x) = 0$ unless $x \in \mathrm{H}$, in which case the probability is $1/|\mathrm{H}|$). We then consider the random variables

$$\mathrm{X}_k = \mathrm{S}_1 - \mathrm{S}_2 + \cdots + \mathrm{S}_{2k-1} - \mathrm{S}_{2k}$$

for $k \geqslant 1$. Probabilistically, these correspond to a random walk on $\mathbf{F}_p$ where the steps are taken alternately from H and $-\mathrm{H}$ (so the picture could be simplified a bit in the case where $-1 \in \mathrm{H}$, since then each Y_i would distributed in the same way as $-\mathrm{Y}_i$, and we would have a "standard" random walk; in any case, the classical theory of such random walks – or of reversible finite Markov chains – implies that if k is allowed to go to infinity, the random variable X_k will converge in law to the uniform distribution on all of $\mathbf{F}_p$; thus, for k large, we should expect X_k to be quite well distributed, and this intuition helps to understand the quantitative statements which follow).

For $\nu > 0$, we define the set

$$\Lambda_\nu = \{a \in \mathbf{F}_p \mid |\varphi_{\mathrm{S}}(a)| > p^{-\nu}\}.$$

Note that $0 \in \Lambda_\nu$ in all cases, and that, since

$$\varphi_{\mathrm{S}}(a) = \frac{1}{|\mathrm{H}|}\sum_{x\in\mathrm{H}} e\Big(\frac{ax}{p}\Big),$$

we can restate our goal, Theorem 3.4.7, as the existence of some $\nu > 0$ such that Λ_ν *only* contains 0. This is therefore our objective.

The content of the second step is as follows:

Proposition 3.5.5. *Let $\eta > 0$ be a real number. Let p be a prime number. If p is large enough, depending only on η, then there exist an integer $k \geqslant 1$ and a positive real number $\nu < \frac{1}{2}\eta$, independent of p, such that*

$$p^{-\eta} \leqslant \frac{|\Lambda_\nu|}{\mathrm{M}_{\mathrm{X}_k}} \leqslant p^{\eta}, \tag{3.19}$$

and

$$\mathbf{E}(|\varphi_{\mathrm{X}_k}(\mathrm{X}_k\widehat{\mathrm{X}}_{2k})|^2) \geqslant p^{-10\eta}. \tag{3.20}$$

Since

$$\mathrm{X}_{2k} = \mathrm{S}_1 - \mathrm{S}_2 + \cdots + \mathrm{S}_{2k-1} - \mathrm{S}_{2k} + \mathrm{S}_{2k+1} - \mathrm{S}_{2k+2} + \cdots + \mathrm{S}_{4k-1} - \mathrm{S}_{4k}$$

is the difference of two independent random variables, each distributed like X_k, Lemma 3.5.3 shows that the estimate (3.20) is precisely the assumption (3.17) of Proposition 3.5.1 for the random variable X_k in place of X, and hence (up to checking the minor conditions (3.18)), we will be able to apply the latter to construct a set A with controlled sum and product growth.

Step 3. We can now conclude. Pick $\eta > 0$, and suppose that p is large enough so that the previous proposition applies.

By Lemma 3.5.3, the conclusion (3.20) of Proposition 3.5.5 implies that the random variables $\mathrm{X} = \mathrm{X}_k$ and $\mathrm{Y} = \mathrm{X}_{2k}$ satisfy $\mathbf{E}(\phi(\mathrm{X}_k\mathrm{Y}_k) \geqslant \alpha^{-1}$ with $\alpha = p^{10\eta}$. This verifies (3.17). Moreover, by induction on k, we have

$$\mathbf{P}(\mathrm{Y} = 0) \leqslant \mathbf{P}(\mathrm{X} = 0) \leqslant \max_{x\in\mathbf{F}_p} \mathbf{P}(\mathrm{Y} = x) = \frac{1}{|\mathrm{H}|},$$

and hence (if η is small enough compared to γ), the conditions (3.18) are also satisfied.

Thus we can apply Proposition 3.5.1, and deduce that there exists a set $\mathrm{A} \subset \mathbf{F}_p^\times$ with

$$\max(|\mathrm{A} + \mathrm{A}|, |\mathrm{A} \cdot \mathrm{A}|) \ll \alpha^d|\mathrm{A}| \tag{3.21}$$

and

$$\frac{1}{\alpha^d\phi(0)} \ll |\mathrm{A}| \ll \frac{\alpha}{\phi(0)},$$

where $\phi(0) = \mathbf{P}(\mathrm{Y} = 0)$. Moreover, since $\phi(0) = \mathrm{M}_\mathrm{X}/p$, the inequalities (3.19) imply that

$$p^{-1-\eta}|\Lambda_\nu| \leqslant \phi(0) \leqslant p^{-1+\eta}|\Lambda_\nu|.$$

Now comes a key observation which exploits the specific structure of the random variable Y as a uniform random variable over H: the set Λ_ν is stable under multiplication by H. (Indeed, for S uniformly distributed on H, the random variable $x\mathrm{S}$ has the same law as S for any $x \in \mathrm{H}$, and therefore $\varphi_{x\mathrm{S}} = \varphi_\mathrm{S}$, which implies the claim.) Hence, as soon as there exists some $a \in \Lambda_\nu - \{0\}$, we have $|\Lambda_\nu| \geqslant |\mathrm{H}|$ (this key step is reminiscent of the use of quasi-randomness, in the sense of Gowers, for instance in the last stages of the proof of Helfgott's Growth Theorem).

Thus, assuming that Λ_ν is not reduced to 0, we deduce from (3.19), and the assumption on $|\mathrm{H}|$, that the estimate

$$|\mathrm{A}| \ll \frac{\alpha}{\phi(0)} \ll \frac{p^{1+11\eta}}{|\Lambda_\nu|} \ll \frac{p^{1+11\eta}}{|\mathrm{H}|} \ll p^{1+11\eta-\gamma}$$

holds. Assuming that (say) $11\eta < \frac{1}{2}\gamma$, this means that $|\mathrm{A}| \ll p^{1-\gamma/2} \leqslant p^{1-\gamma/4}$, for p large enough. Hence the bound (3.21) will contradict Theorem 3.2.1 if η is also small enough so that $\alpha^d = p^{10d\eta} \ll |\mathrm{A}|^\delta$, where $\delta > 0$ is the exponent from the sum-product theorem, applied for the parameter $\gamma/4$.

We claim that, for η sufficiently small, this will indeed be the case. Indeed, since

$$\sum_{a\in\mathbf{F}_p} |\varphi_\mathrm{S}(a)|^2 = \frac{p}{|\mathrm{H}|},$$

we have $|\Lambda_\nu| \leqslant p^{1+2\nu}|\mathrm{H}|^{-1}$ by Chebychev's inequality. Thus

$$|\mathrm{A}| \gg \frac{1}{\alpha^d\phi(0)} \gg \frac{p^{1-(10d+1)\eta}}{|\Lambda_\nu|} \gg p^{-2\nu-(10d+1)\eta+\gamma},$$

and we reach the desired contradiction if η is small enough, depending only on γ (recall that $2\nu < \eta$ by assumption).

We conclude finally that $\Lambda_\nu = \{0\}$, and (by definition) this means that

$$\Big|\frac{1}{|\mathrm{H}|}\sum_{x\in\mathrm{H}} e\Big(\frac{ax}{p}\Big)\Big| \leqslant p^{-\nu}$$

for all $a \in \mathbf{F}_p^\times$, provided p is large enough.

We have thus reduced the proof of Theorem 3.4.7 to that of Propositions 3.5.1 and 3.5.5. This will occupy the remainder of this section.

We begin with Proposition 3.5.1. In the proof, the set A will be constructed by two applications of the Balog–Szemerédi–Gowers Theorem. The necessary energy assumptions will be deduced from probabilistic properties of certain random variables, which we state separately and in greater generality than strictly needed here.

Lemma 3.5.6. *Let* G *be a finite group and let* A *be a subset of* G. *Let* X *be a* G*-valued random variable. We assume that* $\beta \geqslant 1$ *is such that*

$$\mathbf{E}(r_{\mathrm{A}\cdot\mathrm{A}^{-1}}(\mathrm{X})) \geqslant \beta^{-1}|\mathrm{A}|.$$

Let

$$\mathrm{N_X} = \sum_{x\in\mathrm{G}} \mathbf{P}(\mathrm{X}=x)^2.$$

We then have

$$e(\mathrm{A}) \geqslant \frac{1}{4\beta^4\mathrm{N_X}|\mathrm{A}|}.$$

Proof. Let

$$\mathrm{L} = \{x \in \mathrm{G} \mid r_{\mathrm{A}\cdot\mathrm{A}^{-1}}(x) \geqslant \tfrac{1}{2}\beta^{-1}|\mathrm{A}|\},$$

so that we have the lower-bound

$$\mathrm{E}(\mathrm{A}) = \sum_{x\in\mathrm{G}} r_{\mathrm{A}\cdot\mathrm{A}^{-1}}(x)^2 \geqslant \sum_{x\in\mathrm{L}} r_{\mathrm{A}\cdot\mathrm{A}^{-1}}(x)^2 \geqslant \beta^{-2}|\mathrm{A}|^2|\mathrm{L}|.$$

Noting that $r_{\mathrm{A}\cdot\mathrm{A}^{-1}}(x) \leqslant |\mathrm{A}|$ for all x, the assumption implies that

$$\mathbf{P}(\mathrm{L}) = \mathbf{P}\Big(r_{\mathrm{A}\cdot\mathrm{A}^{-1}}(\mathrm{X}) \geqslant \tfrac{1}{2}\beta^{-1}|\mathrm{A}|\Big) \geqslant \frac{1}{2\beta}$$

(see (A.3)), but the Cauchy–Schwarz inequality and positivity imply that

$$\mathbf{P}(\mathrm{L}) = \sum_{x\in\mathrm{L}} \mathbf{P}(\mathrm{X}=x) \leqslant |\mathrm{L}|^{1/2}\Big(\sum_{x\in\mathrm{G}} \mathbf{P}(\mathrm{X}=x)^2\Big)^{1/2} = |\mathrm{L}|^{1/2}\mathrm{N_X}^{1/2},$$

and hence $|\mathrm{L}| \geqslant (2\beta)^{-2}\mathrm{N_X}^{-1}$. The previous lower-bound gives

$$\mathrm{E}(\mathrm{A}) \geqslant 2^{-2}\beta^{-4}\mathrm{N_X}^{-1}|\mathrm{A}|^2,$$

which is equivalent to the desired result. □

Proof of Proposition 3.5.1. We will use frequently the fact that $\phi(y) \leqslant \phi(0)$ for all $y \in \mathbf{F}_p$, which follows for instance from the first formula in Lemma 3.5.3.

We define

$$\mathrm{A}_1 = \{y \in \mathbf{F}_p \mid \phi(y) \geqslant \frac{\phi(0)}{8\alpha}\}$$

and $\mathrm{A}_2 = \mathrm{A}_1 - \{0\} \subset \mathbf{F}_p^\times$ (note that $0 \in \mathrm{A}_1$). By Chebychev's inequality simply, we have

$$|\mathrm{A}_2| \leqslant |\mathrm{A}_1| \leqslant \frac{8\alpha}{\phi(0)}.$$

We now claim that (3.17) implies that

$$\mathbf{E}(\phi(\mathrm{XY})\mathbf{1}_{\mathrm{X}\neq 0,\, \mathrm{Y}\in \mathrm{A}_1\cap \mathrm{X}^{-1}\mathrm{A}_1}) \geqslant \frac{\phi(0)}{2\alpha}. \tag{3.22}$$

This is a matter of showing that the contributions to $\mathbf{E}(\phi(\mathrm{XY}))$ from the complementary event, where $\mathrm{X} = 0$ or $\mathrm{Y} \notin \mathrm{A}_1$, or $\mathrm{XY} \notin \mathrm{A}_1$, are small enough. And indeed, first of all the first part of (3.18) gives the upper bound

$$\mathbf{E}\Big(\phi(\mathrm{XY})\mathbf{1}_{\mathrm{X}=0}\Big) = \phi(0)\,\mathbf{P}(\mathrm{X} = 0) \leqslant \frac{\phi(0)}{4\alpha},$$

while

$$\mathbf{E}(\phi(\mathrm{XY})\mathbf{1}_{\mathrm{X}\neq 0,\ \mathrm{XY}\notin \mathrm{A}_1}) \leqslant \frac{1}{8\alpha}\,\mathbf{E}(\phi(\mathrm{XY})) \leqslant \frac{\phi(0)}{8\alpha}.$$

To bound the last contribution with $\mathrm{X} \neq 0$ and $\mathrm{Y} \notin \mathrm{A}_1$, we write

$$\mathbf{E}(\phi(\mathrm{XY})\mathbf{1}_{\mathrm{X}\neq 0,\ \mathrm{Y}\notin \mathrm{A}_1}) = \sum_{y\notin \mathrm{A}_1} \mathbf{E}(\phi(\mathrm{XY})\mathbf{1}_{\mathrm{X}\neq 0,\, \mathrm{Y}=y}) = \sum_{y\notin \mathrm{A}_1} \mathbf{E}(\phi(y\mathrm{X})\mathbf{1}_{\mathrm{X}\neq 0,\, \mathrm{Y}=y}).$$

Using the independance of X and Y, we deduce that

$$\begin{aligned}\mathbf{E}(\phi(\mathrm{XY})\mathbf{1}_{\mathrm{X}\neq 0,\ \mathrm{Y}\notin \mathrm{A}_1}) &= \sum_{y\in \mathbf{F}_p - \mathrm{A}_1} \mathbf{P}(\mathrm{Y} = y)\,\mathbf{E}(\phi(y\mathrm{X})\mathbf{1}_{\mathrm{X}\neq 0}) \\ &\leqslant \frac{\phi(0)}{8\alpha}\,\mathbf{E}\Big(\sum_{y\notin \mathrm{A}_1} \phi(y\mathrm{X})\mathbf{1}_{\mathrm{X}\neq 0}\Big) \leqslant \frac{\phi(0)}{8\alpha}\,\mathbf{E}\Big(\sum_{y\in \mathbf{F}_p} \phi(y\mathrm{X})\mathbf{1}_{\mathrm{X}\neq 0}\Big) \leqslant \frac{\phi(0)}{8\alpha},\end{aligned}$$

using in the last step the fact that, for any given $x \neq 0$, we have

$$\sum_{y\in \mathbf{F}_p} \phi(yx) = \mathbf{P}(\mathrm{Y} \neq 0) \leqslant 1.$$

We next deduce from (3.22) a lower-bound for $|\mathrm{A}_1|$ complementing the previous upper-bound, namely

$$\frac{1}{2\alpha\phi(0)} \leqslant |\mathrm{A}_1| \leqslant \frac{8\alpha}{\phi(0)}, \tag{3.23}$$

which in turn implies $|\mathrm{A}_1| \geqslant 2$ (by (3.18) since $\phi(0) = \mathbf{P}(\mathrm{Y} = 0)$), and therefore $|\mathrm{A}_2| = |\mathrm{A}_1| - 1 \geqslant \frac{1}{2}|\mathrm{A}_1|$, i.e.

$$\frac{1}{4\alpha\phi(0)} \leqslant |\mathrm{A}_2| \leqslant \frac{8\alpha}{\phi(0)}, \tag{3.24}$$

Indeed, we obtain (3.23) by noting that, by (3.22), we have

$$\frac{\phi(0)}{2\alpha} \leqslant \mathbf{E}(\phi(\mathrm{XY})\mathbf{1}_{\mathrm{X}\neq 0,\ \mathrm{Y}\in \mathrm{A}_1}) \leqslant \phi(0)\,\mathbf{P}(\mathrm{Y} \in \mathrm{A}_1) \leqslant \phi(0)^2|\mathrm{A}_1|.$$

The next step is to relate the bound (3.22) to the representation function r_2 for $\mathrm{A}_2 \cdot \mathrm{A}_2^{-1}$. For this, we start with the formula

$$\mathbf{E}(r_2(\mathrm{X})) = \sum_{y,z\in \mathrm{A}_2} \mathbf{P}(\mathrm{X} = y^{-1}z) = \sum_{y\in \mathrm{A}_2} \mathbf{E}\Big(\sum_{z\in \mathrm{A}_2} \mathbf{P}(y\mathrm{X} = z)\Big) = \sum_{y\in \mathrm{A}_2} \mathbf{P}(y\mathrm{X} \in \mathrm{A}_2).$$

On the other hand, by independance of X and Y, we have

$$\begin{aligned}\mathbf{E}(\phi(\mathrm{XY})\mathbf{1}_{\mathrm{X}\neq 0,\, \mathrm{Y}\in \mathrm{A}_1\cap \mathrm{X}^{-1}\mathrm{A}_1}) &= \sum_{y\in \mathrm{A}_1} \phi(y)\, \mathbf{E}(\phi(y\mathrm{X})\mathbf{1}_{\mathrm{X}\neq 0,\, y\mathrm{X}\in \mathrm{A}_1}) \\ &\leqslant \phi(0)^2\, \mathbf{E}\Big(\sum_{y\in \mathrm{A}_1} \mathbf{1}_{\mathrm{X}\neq 0,\ y\mathrm{X}\in \mathrm{A}_1}\Big) = \phi(0)^2 \sum_{y\in \mathrm{A}_1} \mathbf{P}(\mathrm{X}\neq 0 \text{ and } y\mathrm{X}\in \mathrm{A}_1).\end{aligned}$$

Isolating the contribution of $y = 0 \in \mathrm{A}_1$, we then have

$$\sum_{y\in \mathrm{A}_1} \mathbf{P}(\mathrm{X}\neq 0 \text{ and } y\mathrm{X}\in \mathrm{A}_1) = \mathbf{P}(\mathrm{X}\neq 0) + \mathbf{E}(r_2(\mathrm{X})) \leqslant 1 + \mathbf{E}(r_2(\mathrm{X})),$$

and thus (3.22) implies that

$$\frac{\phi(0)}{2\alpha} \leqslant \phi(0)^2\, \mathbf{E}(r_2(\mathrm{X})) + \phi(0)^2.$$

The assumption $\mathbf{P}(\mathrm{Y} = 0) = \phi(0) \leqslant (4\alpha)^{-1}$ (see (3.18)) now leads to the lower-bound

$$\mathbf{E}(r_2(\mathrm{X})) \geqslant \frac{1}{4\alpha\phi(0)} \geqslant \frac{|\mathrm{A}_2|}{32\alpha^2}.$$

Applying Lemma 3.5.6 to the random variable X on $\mathbf{F}_p^\times$, with $\beta = 32\alpha^2$, we obtain

$$e(\mathrm{A}_2) \geqslant \frac{1}{2^{22}\alpha^8 \mathrm{N}_{\mathrm{X}}|\mathrm{A}_2|} = \frac{1}{2^{22}\alpha^8\phi(0)|\mathrm{A}_2|} \geqslant \frac{1}{2^{25}\alpha^9},$$

and therefore, by the Balog–Szemerédi–Gowers Theorem (Theorem 2.7.1, applied to $\mathrm{A}_2 \subset \mathbf{F}_p^\times$), there exists a subset $\mathrm{A}_3 \subset \mathrm{A}_2$ with $|\mathrm{A}_2| \ll \alpha^d|\mathrm{A}_3|$ and $|\mathrm{A}_3 \cdot \mathrm{A}_3| \ll \alpha^d|\mathrm{A}_3|$, where d and the implied constants are absolute.

We now consider the additive growth. Let r_3 be the representation function for $\mathrm{A}_3 - \mathrm{A}_3$. We first show that the random variable $r_3(\mathrm{Y})$ is quite large on average. Recall that $\mathrm{Y} = \mathrm{X}_1 - \mathrm{X}_2$, where X_1 and X_2 are independent and distributed like X. We then have

$$\mathbf{E}(r_3(\mathrm{Y})) = \sum_{a,b\in \mathrm{A}_3} \mathbf{P}(\mathrm{Y} = a - b) = \sum_{a,b\in \mathrm{A}_3} \mathbf{P}(\mathrm{X}_1 - a = \mathrm{X}_2 - b),$$

and this implies that

$$\begin{aligned}\mathbf{E}(r_3(\mathrm{Y})) &= \sum_{y\in \mathbf{F}_p} \sum_{a,b\in \mathrm{A}_3} \mathbf{P}(\mathrm{X}_1 - a = y \text{ and } \mathrm{X}_2 - b = y) \\ &= \sum_{y\in \mathbf{F}_p} \sum_{a,b\in \mathrm{A}_3} \mathbf{P}(\mathrm{X}_1 - a = y)\, \mathbf{P}(\mathrm{X}_2 - b = y) = \sum_{y\in \mathbf{F}_p} \mathbf{P}(\mathrm{X}_1 \in y + \mathrm{A}_3)^2.\end{aligned}$$

The "reversed" Cauchy–Schwarz inequality now shows that for any choice of $f(y) \geqslant 0$ for $y \in \mathbf{F}_p$, not all zero, we have

$$\mathbf{E}(r_3(\mathrm{Y})) \geqslant \frac{\mathrm{V}^2}{\mathrm{W}}$$

with

$$\mathrm{V} = \sum_{y \in \mathbf{F}_p} f(y)\, \mathbf{P}(\mathrm{X}_1 \in y + \mathrm{A}_3), \qquad \mathrm{W} = \sum_{y \in \mathbf{F}_p} f(y)^2.$$

We pick $f(y) = \mathbf{P}(\mathrm{X}_2 = y)$; in this case, we have

$$\mathrm{V} = \mathbf{P}(\mathrm{Y} \in \mathrm{A}_3), \qquad \mathrm{W} = \mathbf{P}(\mathrm{Y} = 0),$$

and therefore

$$\mathbf{E}(r_3(\mathrm{Y})) \geqslant \frac{\mathbf{P}(\mathrm{Y} \in \mathrm{A}_3)^2}{\phi(0)} \geqslant \frac{\phi(0)}{2^6 \alpha^2} |\mathrm{A}_3|^2,$$

where the last step follows from the fact that $\mathrm{A}_3 \subset \mathrm{A}_1$, so that $\mathbf{P}(\mathrm{Y} = y) \geqslant 2^{-3}\alpha^{-1}\phi(0)$ for $y \in \mathrm{A}_3$. Now recall that $|\mathrm{A}_3| \gg \alpha^{-d}|\mathrm{A}_2| \gg \alpha^{-d-1}\phi(0)^{-1}$ (see (3.24)), and thus $\phi(0)|\mathrm{A}_3| \gg \alpha^{-d-1}$. We deduce then that

$$\mathbf{E}(r_3(\mathrm{Y})) \gg \alpha^{-d_1}|\mathrm{A}_3|,$$

for some absolute constant d_1. Applying Lemma 3.5.6 to the random variable Y on $\mathbf{F}_p$ and the set A_3, with β a multiple of α^{d_1}, we get

$$e(\mathrm{A}_3) \geqslant \frac{1}{4\beta^4 \mathrm{N}_\mathrm{Y} |\mathrm{A}_3|}.$$

But we have

$$\mathrm{N}_\mathrm{Y} = \sum_{y \in \mathbf{F}_p} \mathbf{P}(\mathrm{Y} = y)^2 \leqslant \mathbf{P}(\mathrm{Y} = 0) \sum_{y \in \mathbf{F}_p} \mathbf{P}(\mathrm{Y} = y) = \mathbf{P}(\mathrm{Y} = 0) = \phi(0),$$

(a simple instance of the "flattening" effect of random walks) so we get finally the lower bound

$$e(\mathrm{A}_3) \geqslant \frac{1}{4\beta^4 \phi(0) |\mathrm{A}_3|} \gg \alpha^{-d_2}$$

for some absolute constant d_2. Applying Theorem 2.7.1 to $\mathrm{A}_3 \subset \mathbf{F}_p$, we find a subset $\mathrm{A}_4 \subset \mathrm{A}_3$ with $|\mathrm{A}_3| \ll \alpha^{d_3}|\mathrm{A}_4|$ and $|\mathrm{A}_4 + \mathrm{A}_4| \ll \alpha^{d_3}|\mathrm{A}_4|$ for some absolute constant d_3. Since, in addition

$$|\mathrm{A}_4 \cdot \mathrm{A}_4| \leqslant |\mathrm{A}_3 \cdot \mathrm{A}_3| \ll \alpha^d |\mathrm{A}_3| \ll \alpha^{d+d_3}|\mathrm{A}_4|,$$

we finally have proved Proposition 3.5.1 with the set A equal to A_4. □

We now come to the proof of Proposition 3.5.5. This splits into a more general statement where the distribution of the steps of the random walk are quite arbitrary, and which provides the estimate (3.19) for a suitable value of k, and a last step where (3.20) is obtained when this distribution is uniform on a multiplicative subgroup.

Recall the definition

$$\mathrm{X}_k = \sum_{i=1}^{k} (\mathrm{S}_{2i-1} - \mathrm{S}_{2k}), \qquad k \geqslant 1,$$

of the random walk, and note that it implies the formula

$$\varphi_{\mathrm{X}_k}(a) = |\varphi_\mathrm{S}(a)|^{2k}$$

for any $k \geqslant 1$ and $a \in \mathbf{F}_p$, since the summands are independent.

Proof of Proposition 3.5.5. We observe first that for any integer $k \geqslant 1$ and $\nu > 0$, provided the condition $4k\nu \leqslant \eta$ is satisfied, we have

$$\frac{|\Lambda_\nu|}{\mathrm{M}_{\mathrm{X}_k}} \leqslant p^\eta,$$

since then

$$\mathrm{M}_{\mathrm{X}_k} \geqslant \frac{|\Lambda_\nu|}{p^{4k\nu}} \geqslant \frac{|\Lambda_\nu|}{p^\eta}$$

by definition.

We now claim that if p is large enough, depending only on η, then we can find the integer $k \geqslant 1$ and $\nu < \frac{1}{2}\eta$, independent of p, such that $4k\nu \leqslant \eta$ and

$$p^{-\eta} \leqslant \frac{|\Lambda_\nu|}{\mathrm{M}_{\mathrm{X}_k}}. \tag{3.25}$$

To prove the claim, we first note that there is a general upper bound

$$\mathrm{M}_{\mathrm{X}_k} \leqslant |\Lambda_{1/k}| + p \cdot (p^{-4k})^k = |\Lambda_{1/k}| + p^{-3} \leqslant |\Lambda_{1/k}|(1+p^{-3}),$$

valid for any integer $k \geqslant 1$. Now, given $k \geqslant 1$, we denote $k_+ = \lceil \frac{\eta}{k^2} \rceil$. If the inequality $\mathrm{M}_{\mathrm{X}_k} > p^\eta |\Lambda_{1/k_+}|$ holds, then it follows that

$$|\Lambda_{1/k_+}| \leqslant |\Lambda_{1/k}| p^{-\eta}(1+p^{-3}).$$

Iterating this observation m times, starting from $k = 4$, we see that *either* we find $k \geqslant 1$ such that (3.25) holds for $\nu = 1/k_+$, or we have

$$|\Lambda_{1/k}| \leqslant p^{1-m\eta}(1+p^{-3})^m$$

for $m \geqslant 1$ and some k depending on m. But for suitable m, we obtain $|\Lambda_{1/k}| < 1$, which is a contradiction since $0 \in \Lambda_\nu$ for all ν.

Our next goal is the crucial inequality

$$\mathbf{E}(|\varphi_{\mathrm{X}_k}(a\mathrm{X}_k)|^2) \geqslant \varphi_{\mathrm{S}}(a)^{4k} \tag{3.26}$$

for all $k \geqslant 1$ and $a \in \mathbf{F}_p$, which depends on the specific choice of random walk. Indeed, we have

$$\mathbf{E}(\varphi_{\mathrm{X}_k}(a\mathrm{X}_k)^2) = \mathbf{E}(\varphi_{\mathrm{X}_k}(a\mathrm{X}_{2k})) = \mathbf{E}(|\varphi_{\mathrm{S}}(a\mathrm{X}_{2k})|^{2k}) \geqslant \mathbf{E}(\varphi_{\mathrm{S}}(a\mathrm{X}_{2k}))^{2k}$$

by Jensen's inequality. But, by symmetry, we have

$$\mathbf{E}(\varphi_{\mathrm{S}}(a\mathrm{X}_{2k})) = \mathbf{E}(|\varphi_{\mathrm{X}_k}(a\mathrm{S})|^2)$$

and finally $\mathbf{E}(|\varphi_{\mathrm{X}_k}(a\mathrm{S})|^2) = \varphi_{\mathrm{X}_k}(a)^2$ since $\varphi_{\mathrm{X}_k}(a\mathrm{S}) = \varphi_{\mathrm{X}_k}(a)$, which concludes the proof.

We can then deduce (3.20) straightforwardly. First, From (3.25), we deduce the lower bound

$$\mathbf{P}(\widehat{\mathrm{X}}_{2k} \in \Lambda_\nu) \geqslant p^{-\eta}\frac{|\Lambda_\nu|}{\mathrm{M}_{\mathrm{X}_k}} \geqslant p^{-2\eta},$$

and then from (3.26), we get

$$\mathbf{E}(|\varphi_{\mathrm{X}_k}(\mathrm{X}_k\widehat{\mathrm{X}}_{2k})|^2) \geqslant \mathbf{E}(\varphi_{\mathrm{X}_k}(\widehat{\mathrm{X}}_{2k})^{4k}) \geqslant p^{-4k^2\nu}\,\mathbf{P}(\widehat{\mathrm{X}}_{2k} \in \Lambda_\nu) \geqslant p^{-4k^2\nu-2\eta} \geqslant p^{-10\eta}.$$

□

3.6. Final remarks

As indicated by Bourgain, Katz and Tao, their proof was inspired by a paper of Edgar and Miller [25], who established that any *subring* A of $\mathbf{R}$ which is a so-called "analytic set" (i.e., the image of a Borel set by a continuous map $\mathbf{R}^k \to \mathbf{R}$ for some k) must be either equal to $\mathbf{R}$ or rather small, in the sense that its Hausdorff dimension is zero. (This means that, for any real numbers $s > 0$, $\delta > 0$ and $\varepsilon > 0$, we can find a sequence $(\mathrm{I}_j)_{j\geqslant 1}$

of intervals in $\mathbf{R}$ of length $< \delta$ such that A is contained in the union of the I_j's and the inequality

$$\sum_j \lambda(\mathrm{I}_j)^s < \varepsilon$$

holds, where λ denotes the length of the intervals.) Interestingly, some set-theoretic regularity condition, such that the assumption that A is a Borel set, is necessary: it is known by work of Davies (as explained by Edgar and Miller [25, p. 1122]) that *assuming the continuum hypothesis* there exist subrings of $\mathbf{R}$ of any Hausdorff dimension between 0 and 1.

CHAPTER 4

Arithmetic progressions

4.1. Introduction: structure and randomness

The topic of this chapter is that of *arithmetic progressions* in subsets of abelian groups. This was already mentioned in the introduction, with the theorem of van der Waerden and that of Szemerédi. Our discussion will be very incomplete and will focus on (elementary) quantitative aspects, which have been at the forefront of many fundamental developments of additive combinatorics and its applications.

We begin by introducing some notation.

DEFINITION 4.1.1. Let $k \geqslant 1$ be an integer and let G be an abelian group.

A k-term arithmetic progression in G is an arithmetic progression of length k.

For any finite subset A of G, we denote by $\mathscr{F}_k(\mathrm{A})$ the maximal size of a subset of A which *does not* contain a *proper* k-term arithmetic progression, i.e., for which there does not exist $a_0 \in \mathrm{A}$ and $a \in \mathrm{G}$ such that A contains the elements a_0, a_0+a, ..., $a_0+(k-1)a$, and moreover these elements are pairwise distinct.

REMARK 4.1.2. (1) A more customary notation is $r_k(\mathrm{A})$, but we want to avoid a clash of notation with representation functions. The initial $\mathscr{F}$ is meant to indicate sets Free of k-term arithmetic progressions.

(2) With this notation, Szemerédi's Theorem can be summarized as the fact that, for k fixed, we have

$$\lim_{\mathrm{N}\to+\infty} \frac{\mathscr{F}_k([\mathrm{N}])}{\mathrm{N}} = 0,$$

(with [N] viewed as a subset of $\mathbf{Z}$).

The quantitative aspect we are considering is the question of finding "explicit" functions f defined for positive integers, with $f(\mathrm{N})$ tending to $+\infty$ as $\mathrm{N} \to +\infty$, so that

$$\mathscr{F}_k([\mathrm{N}]) \leqslant \frac{\mathrm{N}}{f(\mathrm{N})}.$$

(3) If G has no torsion (for instance, if $\mathrm{G} = \mathbf{Z}$), a k-term arithmetic progression $\{a_0 + ia\}$ is a proper progression if and only if $a \neq 0$.

The following fact is, as usual, elementary, but very useful.

PROPOSITION 4.1.3. *Let $k \geqslant 1$ be an integer. Let* G *and* H *be abelian groups. If* A *is a subset of* G *and $f\colon \mathrm{A} \to \mathrm{H}$ is a Freiman 2-morphism, then the image by f of an arithmetic progression in* A *is an arithmetic progression in* H.

In particular, if f is injective, then $\mathscr{F}_k(f(\mathrm{A})) \leqslant \mathscr{F}_k(\mathrm{A})$, and if f defines by restriction a Freiman 2-isomorphism from A *to $f(\mathrm{A})$, then $\mathscr{F}_k(f(\mathrm{A})) = \mathscr{F}_k(\mathrm{A})$.*

PROOF. The fact that the image by f of an arithmetic progression in A is one in H follows directly from the characterization of arithmetic progressions as images of Freiman 2-morphisms from intervals of $\mathbf{Z}$ to H (Example 2.2.5, (4)).

If f is injective, then the image of a proper arithmetic progression is a proper arithmetic progression. In this case, if $X \subset f(A)$ contains no k-terms arithmetic progression, the inverse image $f^{-1}(X)$ is contained in A (by injectivity) and contains no k-terms arithmetic progression by the above. It follows that $\mathscr{F}_k(A) \geqslant |f^{-1}(X)| = |X|$, hence $\mathscr{F}_k(A) \geqslant \mathscr{F}_k(f(A))$. □

EXERCISE 4.1.4. (1) Let G be a finite abelian group and let $H = \mathbf{Z}/p\mathbf{Z}$ for some prime number p. Let $A \subset G$ and $B \subset H$ be subsets of G and H, respectively. Show that $\mathscr{F}_k(A)\mathscr{F}_k(B) \leqslant \mathscr{F}_k(A \times B)$, with $A \times B \subset G \times H$.

(2) Show by an example that the previous fact is *false* if H is an arbitary finite abelian group.

(3) For $n \geqslant 1$, show that a proper 3-term progression in $\mathbf{F}_3^n$ is an affine line in this $\mathbf{F}_3$-vector space. Moreover, show that such a line ℓ is of the form $\ell = \{x_1, x_2, x_3\}$ where $x_i = (x_{i,1}, \ldots, x_{i,n})$ and for $j = 1, \ldots, n$, *either*

$$x_{1,j} = x_{2,j} = x_{3,j}$$

or

$$\{x_{1,j}, x_{2,j}, x_{3,j}\} = \mathbf{F}_3.$$

(4) For $n \geqslant 1$, show that $\mathscr{F}_3(\mathbf{F}_3^n) \geqslant 2^n$.

4.2. Sets without arithmetic progressions

It is rather natural to attempt to get a feeling for the problem by trying to construct specific examples of "large" sets which do *not* contain k-terms in arithmetic progressions. The simplest example is probably the following:

EXAMPLE 4.2.1. Let $N \geqslant 1$ be an integer and let $S \subset [N]$ be the set of elements whose ternary expansion does not involve the digit 2, i.e., the integers $n \leqslant N$ of the form

$$a_0 + 3a_1 + \cdots + 3^k a_k$$

with $a_i \in \{0, 1\}$. The set S does not contain any 3-term arithmetic progression: if a, b, c are elements of S, with ternary digits (a_i), (b_i) and (c_i), then the computation of $a + b$ as well as that of 2 can be done "without carry", so

$$a + b = \sum_i (a_i + b_i)3^i, \qquad 2c = \sum_i (2c_i)3^i.$$

Since $0 \leqslant a_i, b_i, c_i \leqslant 1$, the equality $a + b = 2c$ is only possible if, whenever $c_i = 1$, we have $a_i = b_i = 1$, and whenever $c_i = 0$, we have $a_i = b_i = 0$. But this combines to say that $a = b$, proving the assertion.

It was apparently conjectured by Szekeres (as mentioned by Erdős and Turán [29, p. 263]) that the previous example was best possible, and in particular that a set $A \subset [N]$ with no 3-term progressions has size $|A| \ll N^\alpha$, where $\alpha = (\log 2)/(\log 3)$ (the "dimension" of the set of integers whose ternary expansion omits the digit 2). This was disproved by Salem and Spencer [73], who constructed examples showing that *no* inequality of the form $|A| \ll N^{1-\delta}$, with $\delta > 0$ fixed, could hold. Behrend [5] showed how to improve the construction using a beautiful trick inspired by geometry, providing examples which are still close to the best known.

PROPOSITION 4.2.2 (Behrend). *For* N *large enough, we have*

$$\mathscr{F}_3([\mathrm{N}]) \gg \frac{\mathrm{N}}{\exp(c\sqrt{\log \mathrm{N}})}$$

where $c > 0$ *is an absolute constant, or in other words, for* $\mathrm{N} \geqslant 1$ *large enough, there exists a set* $\mathrm{B} \subset [\mathrm{N}]$ *which contains no 3-term arithmetic progression and satisfies*

$$|\mathrm{B}| \gg \frac{\mathrm{N}}{\exp(c\sqrt{\log \mathrm{N}})}.$$

PROOF. The beautiful idea is to combine an "obvious" construction in a high-rank group free abelian $\mathbf{Z}^d$ with a Freiman 2-morphism to $\mathbf{Z}$, and to apply Proposition 4.1.3.

The "obvious" construction is as follows: given any integer $d \geqslant 1$ and any radius $r > 0$, the (euclidean) sphere of radius r in $\mathbf{R}^d$, i.e., the set $\mathrm{S}_{\mathbf{R}^d}(r)$ of all $x = (x_i)_{1 \leqslant i \leqslant d}$ in $\mathbf{R}^d$ such that

$$\|x\|^2 = x_1^2 + \cdots + x_d^2 = r^2,$$

contains no 3-term arithmetic progression. This is because the equation $x + y = 2z$, with x, y and z in $\mathbf{R}^d$, implies if $x \neq y$ that $z = \frac{1}{2}(x+y)$ belongs to the line passing through x and y, which is not possible if x and y are on the same sphere (so $\|x\| = \|y\|$) since a line intersects the sphere in at most two points, and $z \notin \{x, y\}$. (Of course, this can also be checked easily using the equation of the sphere.)

If we now assume that r^2 is a positive integer, then the set $\mathrm{S}_{\mathbf{Z}^d}(r) = \mathbf{Z}^d \cap \mathrm{S}_{\mathbf{R}^d}(r)$ contains also no 3-term arithmetic progressions, and may have rather large size if r^2 is suitably chosen. Noting that $\mathrm{S}_{\mathbf{Z}^d}(r) \subset [-r, r]^d$, the image of this set by any Freiman 2-isomorphism $[-r, r]^d \to \mathrm{X}$, where $\mathrm{X} \subset [\mathrm{N}]$, is a subset of $[\mathrm{N}]$ which does not contain 3-term arithmetic progressions. Behrend's bound follows by finding suitables choices of the parameters which are involved.

We do this without looking for any optimality. First, consider $d \geqslant 1$ arbitrary. Consider all $(x_i) \in \mathbf{Z}^d$ with $1 \leqslant x_i \leqslant \mathrm{M}$ for some integer M; there are M^d such elements, and they all have euclidean norm squared $\leqslant d\mathrm{M}^2$. By the pigeon-hole principle, we can find a radius $r^2 \leqslant d\mathrm{M}^2$ such that $|\widetilde{\mathrm{S}}_{\mathbf{Z}^d}(r)| \geqslant d^{-1}\mathrm{M}^{d-2}$, where $\widetilde{\mathrm{S}}_{\mathbf{Z}^d}(r)$ restricts the integral solutions to $\|x\|^2 = r^2$ to have $1 \leqslant x_i \leqslant \mathrm{M}$. As in Example 2.2.5, (3), we find a Freiman 2-isomorphism f from $[\mathrm{M}]^d$ to a subset of $[\mathrm{N}]$, where $\mathrm{N} = (2\mathrm{M}+1)^d \leqslant (3\mathrm{M})^d$. Then

$$\mathscr{F}_3([\mathrm{N}]) \geqslant \frac{\mathrm{M}^{d-2}}{d}.$$

We need to transcribe this inequality in terms of N: we first have

$$\frac{\mathrm{M}^{d-2}}{d} \geqslant \frac{\mathrm{N}}{d3^d\mathrm{M}^2} \geqslant \frac{\mathrm{N}}{4d3^d\mathrm{N}^{2/d}} \gg \frac{\mathrm{N}}{4^d\mathrm{N}^{2/d}}.$$

Since the parameter d is free, and we want the right-hand side to be as large as possible, we choose d to equalize the negative effect of dividing by 4^d with the positive effect of having $\mathrm{N}^{2/d}$ in the denominator. In other words, we select d so that 4^d is (approximately) equal to $\mathrm{N}^{2/d}$, say

$$d = \left\lceil \left(\frac{2\log \mathrm{N}}{\log 4}\right)^{1/2} \right\rceil.$$

Then

$$\mathscr{F}_3([\mathrm{N}]) \geqslant \frac{\mathrm{M}^{d-2}}{d} \gg \frac{\mathrm{N}}{4^{2d}} \gg \mathrm{N}\exp(-c\sqrt{\log \mathrm{N}})$$

for some suitable $c > 0$ and N large. □

REMARK 4.2.3. (1) Functions of the type $N \exp(-c\sqrt{\log N})$ also occur naturally in number theory in the error term for the prime number theorem and related results (see, e.g., [53, Th. 5.13, Th. 5.27]). The basic qualitative property of this function is that it is growing *faster* than $N^{1-\delta}$ for any fixed $\delta > 0$, but *slower* than $N(\log N)^{-A}$ for any $A > 0$.

(2) One can be much more precise concerning the number of points with integral coordinates on a sphere of radius r with r^2 a positive integer, especially with d large. This more precise information is however not really needed here.

EXERCISE 4.2.4. Construct an example of a coloring of the set of positive integers in two colors, in such a way that there is no *infinite* arithmetic progression of either color.

4.3. Three-term progressions

The first non-trivial case of Szemerédi's Theorem is that of 3-term progressions, or in other words, that of finding an element a of A and an integer $d \geqslant 1$ such that $a + d$ and $a + 2d$ both belong to A.

An equivalent formulation, which turns out to be convenient, is that if a, b and c are elements of A with $a \neq b$, then $a + b \neq 2c$. (Indeed, there is a bijection

$$\{(a, d) \in A \times \mathbf{N} \mid \{a, a + d, a + 2d\} \subset A\} \to \{(a, b, c) \in A^3 \mid a \leqslant b \text{ and } a + b = 2c\}$$

mapping (a, d) to $(a, a+2d, a+d)$, with reciprocal bijection mapping (a, b, c) to $(a, c-a)$, and the degenerated progressions with $d = 0$ map to (a, b, c) with $a = b$.)

REMARK 4.3.1. This may suggest that the question is close to that of *sum-free sets* discussed in Section 2.8, but we will see that in fact the behavior of those two problems is completely different.

The first quantitative form of existence of arithmetic progressions in sets of positive density was found by Roth [68], who in fact then proved in a rather strong form the first non-trivial case of Szemerédi's Theorem, when $k = 3$.

THEOREM 4.3.2 (Roth). *Let $N \geqslant 1$ be an integer and let $A \subset [N]$ be a set of positive integers such that A does not contain a 3-term arithmetic progression. We have the bound*

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

where the implied constant is absolute. In particular, we have

$$\lim_{N \to +\infty} \frac{\mathscr{F}_3([N])}{N} = 0.$$

REMARK 4.3.3. Numerous successive breakthroughs have led rather recently to spectacular improvements to the understanding of sets without 3-term arithmetic progressions. First, Bloom and Sisask [6] succeeded in proving that

$$\mathscr{F}_3([N]) \ll \frac{N}{(\log N)^{1+\delta}}$$

for all $N \geqslant 2$ and some fixed real number $\delta > 0$, which in particular implies immediately that any set of prime numbers with positive density (among the primes) contains some 3-term arithmetic progression (a fact which had been first proved, using very different methods, by Green [44]). Then Kelley and Meka [54, Th. 1.2] obtained the estimate

$$\mathscr{F}_3([N]) \ll N \exp(-c(\log N)^{1/12})$$

for some constant $c > 0$, thus reaching the "same type" of condition as Behrend's example.

The basic ideas of the proof, in the way it is presented in modern texts, are the following.

Step 1. One can express the "number" of 3-term arithmetic progressions contained in a subset $A \subset [N]$ in terms of Fourier analysis. Here, although Roth used ideas from the classical circle method of analytic number theory, it can be replaced by discrete Fourier analysis after applying a Freiman 2-isomorphism to reduce to $\mathbf{Z}/q\mathbf{Z}$ for a suitable integer q.

Step 2. Studying the formula obtained in Step 1, one sees that if A does *not* contain a 3-term arithmetic progression, then some Fourier coefficient of the characteristic function of A must be relatively large.

Step 3. Going further, one proves that, under the same assumption that $A \subset [N]$ contains no 3-term arithmetic progression, there exist $M < N$, with $\log M \gg \log N$, a subset $B \subset [M]$ such that B contains no 3-term progressions but the densities $\beta = |B|/M$ and $\alpha = |A|/N$ of A and B satisfy

$$\beta \geqslant \alpha + c\alpha^2$$

for some (absolute) constant $c > 0$. This crucial step is known as a *density increment* property.

Step 4. The conclusion is now obtained by repeatedly applying the density increment in Step 3, which cannot be done *ad infinitum* since the density β must always remain $\leqslant 1$: after a number of steps, the assumption in Step 2 that the sets we have obtained do not contain 3-term arithmetic progressions must fail, and it turns out that this leads to the precise form of Roth's Theorem that we have given.

We will now proceed to discuss all these steps in turn.

Definition 4.3.4. Let G be a finite abelian group, with additive notation. The 3-*term detector* is the trilinear map

$$\mathrm{AP}_3 \colon \mathrm{C}(G)^3 \to \mathbf{C}$$

defined by

$$\begin{aligned}\mathrm{AP}_3(f_1, f_2, f_3) &= \mathop{\mathbf{E}}_{a_0,a} (f_1(a_0) f_2(a_0 + a) f_3(a_0 + 2a)) \\ &= \frac{1}{|G|^2} \sum_{a \in G} \sum_{a_0 \in G} f_1(a_0) f_2(a_0 + a) f_3(a_0 + 2a).\end{aligned}$$

Note that we perform the summation over all a_0 and a, including $a = 0$, which means including also the "degenerate" 3-term arithmetic progressions. The contribution of $a = 0$ is however quite explicit, namely

$$\frac{1}{|G|^2} \sum_{a_0 \in G} f_1(a_0) f_2(a_0) f_3(a_0).$$

In particular, a subset A of a finite abelian group G contains a 3-term arithmetic progression if and only if $\mathrm{AP}_3(\varphi_A, \varphi_A, \varphi_A)$ is different from this contribution, where φ_A is the characteristic function of the set A, or in other words, if

$$\mathrm{AP}_3(\varphi_A, \varphi_A, \varphi_A) \neq \frac{|A|}{|G|^2}.$$

The crucial starting point of the proof of Roth's Theorem is that this detector function has a simple Fourier-theoretic expression. We recall here the definition

$$\widehat{f}(\xi) = \frac{1}{\sqrt{|\mathrm{G}|}} \sum_{x\in \mathrm{G}} f(x)\overline{\xi(x)}$$

of the unitarily normalized Fourier transform of a function $f\colon \mathrm{G} \to \mathbf{C}$, where $\xi \in \widehat{\mathrm{G}}$ is a character of G.

Proposition 4.3.5. *Let* G *be a finite abelian group. For* f_1, f_2, $f_3 \in \mathrm{C}(\mathrm{G})$, *we have*

$$\mathrm{AP}_3(f_1, f_2, f_3) = \frac{1}{|\mathrm{G}|^{3/2}} \sum_{\xi\in\widehat{\mathrm{G}}} \widehat{f_1}(\xi)\widehat{f_2}(\xi^{-2})\widehat{f_3}(\xi).$$

Proof. We insert the expansion of the functions in terms of their Fourier transforms in the definition of the detector: since

$$f(x) = \frac{1}{\sqrt{|\mathrm{G}|}} \sum_{\xi\in\widehat{\mathrm{G}}} \widehat{f}(\xi)\xi(x)$$

for any $x \in \mathrm{G}$, we have

$$\mathrm{AP}_3(f_1, f_2, f_3) = \frac{1}{|\mathrm{G}|^{7/2}} \sum_{a_0\in\mathrm{G}} \sum_{a\in\mathrm{G}} \sum_{\xi_1,\xi_2,\xi_3\in\widehat{\mathrm{G}}} \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)\xi_1(a_0)\xi_2(a_0+a)\xi_3(a_0+2a).$$

Summing over a and a_0 first, and using the orthogonality relations

$$\frac{1}{|\mathrm{G}|}\sum_{a_0\in\mathrm{G}} (\xi_1\xi_2\xi_3)(a_0) = \begin{cases} 1 & \text{if } \xi_1\xi_2\xi_3 = 1,\\ 0 & \text{otherwise,}\end{cases}$$

$$\frac{1}{|\mathrm{G}|}\sum_{a_0\in\mathrm{G}} (\xi_2\xi_3^2)(a_0) = \begin{cases} 1 & \text{if } \xi_2\xi_3^2 = 1,\\ 0 & \text{otherwise,}\end{cases}$$

for characters (see (A.7)), we deduce

$$\mathrm{AP}_3(f_1, f_2, f_3) = \frac{1}{|\mathrm{G}|^{3/2}} \sum_{\substack{\xi_1,\xi_2,\xi_3\\ \xi_1\xi_2\xi_3=1\\ \xi_2\xi_3^2=1}} \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3).$$

The set of triples of characters (ξ_1, ξ_2, ξ_3) satisfying the summation conditions can be parameterized by the single character $\xi = \xi_3$, with $\xi_2 = \xi^{-2}$ (by the last condition) and $\xi_1 = \xi$ (inserting the value of ξ_2 in the first equation). Therefore

$$\mathrm{AP}_3(f_1, f_2, f_3) = \frac{1}{|\mathrm{G}|^{3/2}} \sum_{\xi\in\widehat{\mathrm{G}}} \widehat{f_1}(\xi)\widehat{f_2}(\xi^{-2})\widehat{f_3}(\xi),$$

as claimed. □

The next corollary is immediate, but crucial.

Corollary 4.3.6. *Let* G *be a finite abelian group* of odd order. *For* f_1, f_2, $f_3 \in \mathrm{C}(\mathrm{G})$, *we have*

$$|\mathrm{AP}_3(f_1, f_2, f_3)| \leqslant \frac{1}{|\mathrm{G}|^{1/2}} \|f_1\|\ \|f_2\|\ \|\widehat{f_3}\|_\infty$$

where $\|f_1\|$ *and* $\|f_2\|$ *are the Hilbert space norms in* $\mathrm{C}(\mathrm{G})$ *and*

$$\|\widehat{f_3}\|_\infty = \max_{\xi\in\widehat{\mathrm{G}}} |\widehat{f_3}(\xi)|.$$

Proof. By the triangle inequality followed by the Cauchy-Schwarz inequality, we get

$$\begin{aligned}|\mathrm{AP}_3(f_1,f_2,f_3)| &\leqslant \frac{1}{|\mathrm{G}|^{3/2}}\|\widehat{f_3}\|_\infty \sum_{\xi\in\widehat{\mathrm{G}}} |\widehat{f_1}(\xi)||\widehat{f_2}(\xi^{-2})| \\ &\leqslant \frac{1}{|\mathrm{G}|^{1/2}}\|\widehat{f_3}\|_\infty \Big(\frac{1}{|\mathrm{G}|}\sum_{\xi\in\widehat{\mathrm{G}}} |\widehat{f_1}(\xi)|^2\Big)^{1/2}\Big(\frac{1}{|\mathrm{G}|}\sum_{\xi\in\widehat{\mathrm{G}}} |\widehat{f_2}(\xi^{-2})|^2\Big)^{1/2}.\end{aligned}$$

Since the Fourier transform is unitary, we have

$$\Big(\frac{1}{|\mathrm{G}|}\sum_{\xi\in\widehat{\mathrm{G}}} |\widehat{f_1}(\xi)|^2\Big)^{1/2} = \|f_1\|.$$

Moreover, since $|\mathrm{G}|$ is odd, so is $|\widehat{\mathrm{G}}|$, and this implies that the map $\xi \mapsto \xi^{-2}$ is a bijection on $\widehat{\mathrm{G}}$. Thus

$$\Big(\frac{1}{|\mathrm{G}|}\sum_{\xi\in\widehat{\mathrm{G}}} |\widehat{f_2}(\xi^{-2})|^2\Big)^{1/2} = \|f_2\|,$$

and the inequality above gives the corollary. □

This concludes the first step of the proof. For the second step, we consider an arbitrary subset $\mathrm{A} \subset \mathrm{G}$ (where G is now only assumed to have odd order, so that the corollary applies). We denote its characteristic function φ_A and put $\alpha = |\mathrm{A}|/|\mathrm{G}|$. We then apply the Fourier decomposition to the function $\psi_\mathrm{A} = \varphi_\mathrm{A} - \alpha$, which is often called the "balanced" (characteristic) function of A. By definition, it satisfies

$$\widehat{\psi}_\mathrm{A}(1) = \frac{1}{|\mathrm{G}|^{1/2}}\sum_{x\in\mathrm{G}}(\varphi_\mathrm{A}(x)-\alpha) = \frac{1}{|\mathrm{G}|^{1/2}}\Big(\frac{|\mathrm{A}|}{|\mathrm{G}|}-\alpha\Big) = 0.$$

and also

$$\widehat{\psi}_\mathrm{A}(\xi) = \widehat{\varphi}_\mathrm{A}(\xi)$$

if ξ is not the trivial character.

Proposition 4.3.7. *Let* G *be a finite abelian group of odd order. Let* $\mathrm{A} \subset \mathrm{G}$ *be a subset which does not contain a proper* 3*-term arithmetic progression. Let* $\alpha = |\mathrm{A}|/|\mathrm{G}|$.

Either we have

$$|\mathrm{G}| < \frac{2}{\alpha^2}, \tag{4.1}$$

or there exists a character $\xi \neq 1$ *of* G *such that*

$$|\widehat{\psi}_\mathrm{A}(\xi)| \geqslant |\mathrm{G}|^{1/2}\frac{\alpha^2}{2^4}. \tag{4.2}$$

Proof. As a first step, we show that the assumption implies either (4.1) or that there exist functions f_1, f_2, both bounded by 1, such that

$$\mathrm{AP}_3(f_1,f_2,\psi_\mathrm{A}) \geqslant \frac{\alpha^3}{2^4}.$$

To see this, we apply Proposition 4.3.5 with all functions equal to φ_A. The assumption means that

$$\mathrm{AP}_3(\varphi_A, \varphi_A, \varphi_A) = \frac{|A|}{|G|^2} = \frac{\alpha}{|G|},$$

because the only improper 3-term arithmetic progressions in a group of odd order are the those progressions $\{a_0, a_0+a, a_0+2a\}$ with $a = 0$ (indeed, if $a \neq 0$, then $a_0 \neq a_0 + a$ and $a_0 + a \neq a_0 + 2a$, but also $a_0 \neq a_0 + 2a$ because $|G|$ is odd).

On the other hand, if we write $f_i = \varphi_A + \alpha$ and use the trilinearity of AP_3, then $\mathrm{AP}_3(\varphi_A, \varphi_A, \varphi_A)$ is the sum of eight terms. One of these is

$$\mathrm{AP}_3(\alpha, \alpha, \alpha) = \alpha^3$$

while (by symmetry) the others are equal to $\mathrm{AP}_3(f_1, f_2, \psi_A)$ for f_1 and f_2 either equal to ψ_A or to the constant function α.

If the inequality (4.1) is not true, then $\alpha^3 - \alpha/|G| \geqslant \frac{1}{2}\alpha^3$, so for some choice of f_1, f_2 either α or ψ_A, we must have

$$\mathrm{AP}_3(f_1, f_2, \psi_A) \geqslant \frac{1}{7} \times \frac{\alpha^3}{2} \geqslant \frac{\alpha^3}{2^4}.$$

We note that $\|f_1\| \leqslant \sqrt{\alpha}$ and $\|f_2\| \leqslant \sqrt{\alpha}$ in all cases: indeed, the norm of the constant function α is α and

$$\|\psi_A\|^2 = \frac{1}{|G|}\sum_{x\in G} |\varphi_A(x) - \alpha|^2 = \alpha - \alpha^2 \leqslant \alpha$$

(by a simple computation), so $\|f_1\|\,\|f_2\| \leqslant \alpha$. Corollary 4.3.6 implies that

$$\frac{\alpha^3}{2^4} \leqslant \frac{1}{|G|^{1/2}}\|f_1\|\,\|f_2\|\,\|\widehat{\psi_A}\|_\infty \leqslant \frac{\alpha}{|G|^{1/2}}\|\widehat{\psi_A}\|_\infty,$$

hence

$$\|\widehat{\psi_A}\|_\infty \geqslant \frac{|G|^{1/2}\alpha^2}{2^4},$$

which concludes the proof. □

Remark 4.3.8. (1) The situation of groups G of even order is (in general) genuinely different. Indeed, if $G = (\mathbf{Z}/2\mathbf{Z})^d$ for some integer $d \geqslant 1$, then a 3-term progression is of the form $\{a_0, a_0 + a, a_0 + 2a\}$, so it is never a proper progression.

On the Fourier side, this is reflected in the fact that every character ξ satisfies $\xi^2 = 1$, so that for functions f_1, f_2 and f_3 on G, we have

$$\mathrm{AP}_3(f_1, f_2, f_3) = \frac{\widehat{f_2}(1)}{|G|^{3/2}}\sum_{\xi\in\widehat{G}} \widehat{f_1}(\xi)\widehat{f_3}(\xi).$$

If $f_1 = f_2 = f_3$ is the characteristic function of a set A, then this gives

$$\mathrm{AP}_3(\varphi_A, \varphi_A, \varphi_A) = \Big(\frac{|A|}{|G|}\Big)^2.$$

(2) We can express the conclusion of Proposition 4.3.7 in the form

$$\Big|\sum_{x\in A} \xi(x)\Big| \geqslant \frac{\alpha^2}{2^4}|G| = \frac{|A|^2}{16|G|}.$$

In particular, this shows that although Roth's Theorem concentrates on relatively dense sets, as is necessary to avoid examples like Behrend's, there are nevertheless many

rather interesting smaller sets which contain 3-term progressions, namely any set A such that

$$\max_{\xi\in\widehat{\mathrm{G}}-\{1\}}\Big|\sum_{x\in\mathrm{A}}\xi(x)\Big| < \frac{|\mathrm{A}|^2}{16|\mathrm{G}|}.$$

We then come to the key "density increment" step. We switch to the case of the group $\mathrm{G} = \mathbf{Z}/\mathrm{N}\mathbf{Z}$ for some odd integer N. In this case, recall that the characters ξ are parameterized by the elements a of $\mathbf{Z}/\mathrm{N}\mathbf{Z}$, where a corresponds to the character $x \mapsto e(ax/\mathrm{N})$. We identify this way the Fourier transform of a function $f\colon \mathbf{Z}/\mathrm{N}\mathbf{Z} \to \mathbf{C}$ with a function $\mathbf{Z}/\mathrm{N}\mathbf{Z} \to \mathbf{C}$.

PROPOSITION 4.3.9. *Let $\mathrm{N} \geqslant 1$ be an integer and $\beta > 0$ a real number. Let f be a function in $\mathrm{C}(\mathbf{Z}/\mathrm{N}\mathbf{Z})$ which is bounded by 1 in modulus and which satisfies*

$$\widehat{f}(0) = 0, \qquad \max_{a\in\mathbf{Z}/\mathrm{N}\mathbf{Z}} |\widehat{f}(a)| \geqslant \beta \mathrm{N}^{1/2}.$$

There exists an arithmetic progression $\mathrm{P} \subset \mathbf{Z}/\mathrm{N}\mathbf{Z}$ with $|\mathrm{P}| \geqslant \frac{\beta}{2^6}\mathrm{N}^{1/2}$ such that

$$\Big|\sum_{x\in\mathrm{P}} f(x)\Big| \geqslant \frac{\beta|\mathrm{P}|}{2}.$$

If f is real-valued, then we can find P such that

$$\sum_{x\in\mathrm{P}} f(x) \geqslant \frac{\beta|\mathrm{P}|}{4}.$$

PROOF. The assumption means that there exists $a \in \mathbf{Z}/\mathrm{N}\mathbf{Z}$ such that

$$\Big|\sum_{x\in\mathbf{Z}/\mathrm{N}\mathbf{Z}} f(x)e\Big(\frac{ax}{\mathrm{N}}\Big)\Big| \geqslant \beta\mathrm{N},$$

and a must be non-zero (since the left-hand side vanishes when $a = 0$).

Let $\mathrm{M} < \mathrm{N}$ be a positive integer to be determined later. According to Dirichlet's approximation theorem (see Theorem A.1.1), we can find integers q and r, with $1 \leqslant m \leqslant \mathrm{M}$ such that

$$\Big|\frac{a}{\mathrm{N}} - \frac{r}{m}\Big| \leqslant \frac{1}{m\mathrm{M}}.$$

Consider then the arithmetic progressions of the form $\mathrm{P} = n+m[\mathrm{I}]$ in $\mathbf{Z}/\mathrm{N}\mathbf{Z}$, where $\mathrm{I} \geqslant 1$ is another parameter and n varies.[1] The point is that for suitable choice of I, the character $x \mapsto e(ax/\mathrm{N})$ is close to being constant on these progressions, and consequently the sum of $f(x)e(ax/\mathrm{N})$ over P is very close to the sum of $f(x)$ over P.

Precisely, for $x = n + mi \in \mathrm{P}$, we have by definition

$$\frac{ax}{\mathrm{N}} = \frac{an}{\mathrm{N}} + \frac{ami}{\mathrm{N}} = \frac{an}{\mathrm{N}} + ri + mi\Big(\frac{a}{\mathrm{N}} - \frac{r}{m}\Big)$$

and since $e(ri) = 1$, we get

$$e\Big(\frac{ax}{\mathrm{N}}\Big) = e\Big(\frac{an}{\mathrm{N}}\Big)e\Big(mi\Big(\frac{a}{\mathrm{N}} - \frac{r}{m}\Big)\Big).$$

The choice of m and the fact that $1 \leqslant i \leqslant \mathrm{I}$ ensure that

$$\Big|mi\Big(\frac{a}{\mathrm{N}} - \frac{r}{m}\Big)\Big| \leqslant \frac{\mathrm{I}}{\mathrm{M}},$$

[1] Here, the multiplication by m indicates dilation.

from which it follows that

$$\Big|e\Big(\frac{ax}{\mathrm{N}}\Big) - e\Big(\frac{an}{\mathrm{N}}\Big)\Big| \leqslant \frac{2\pi\mathrm{I}}{\mathrm{M}}$$

for all $x \in n + m[\mathrm{I}]$. It follows next that

$$\Big|\sum_{x\in n+m[\mathrm{I}]} f(x)e\Big(\frac{ax}{\mathrm{N}}\Big) - e\Big(\frac{an}{\mathrm{N}}\Big)\sum_{x\in n+m[\mathrm{I}]} f(x)\Big| \leqslant \frac{2\pi\mathrm{I}^2}{\mathrm{M}}, \tag{4.3}$$

since $|f(x)| \leqslant 1$ for all x.

If we assume that the arithmetic progression $m[\mathrm{I}] \subset \mathbf{Z}/\mathrm{N}\mathbf{Z}$ is a proper progression (of length I), then summing over all $n \in \mathbf{Z}/\mathrm{N}\mathbf{Z}$, the sets $n + m[\mathrm{I}]$ cover each element x exactly I times:

$$\sum_{n\in\mathbf{Z}/\mathrm{N}\mathbf{Z}}\sum_{x\in n+m[\mathrm{I}]} f(x)e\Big(\frac{ax}{\mathrm{N}}\Big) = \sum_{x\in\mathbf{Z}/\mathrm{N}\mathbf{Z}} f(x)e\Big(\frac{ax}{\mathrm{N}}\Big)\sum_{x\in n+m[\mathrm{I}]} 1,$$

so that

$$\frac{1}{\mathrm{I}}\sum_{n\in\mathbf{Z}/\mathrm{N}\mathbf{Z}}\sum_{x\in n+m[\mathrm{I}]} f(x)e\Big(\frac{ax}{\mathrm{N}}\Big) = \sum_{x\in\mathbf{Z}/\mathrm{N}\mathbf{Z}} f(x)e\Big(\frac{ax}{\mathrm{N}}\Big).$$

The assumption combined with (4.3) therefore implies that

$$\Big|\frac{1}{\mathrm{I}}\sum_{n\in\mathbf{Z}/\mathrm{N}\mathbf{Z}} e\Big(\frac{an}{\mathrm{N}}\Big)\sum_{x\in n+m[i]} f(x)\Big| \geqslant \beta\mathrm{N} - \frac{2\pi\mathrm{IN}}{\mathrm{M}}.$$

If we also pick I so that $2\pi\mathrm{I} \leqslant \frac{1}{2}\beta\mathrm{M}$, this implies that

$$\Big|\frac{1}{\mathrm{I}}\sum_{n\in\mathbf{Z}/\mathrm{N}\mathbf{Z}} e\Big(\frac{an}{\mathrm{N}}\Big)\sum_{x\in n+m[\mathrm{I}]} f(x)\Big| \geqslant \frac{\beta\mathrm{N}}{2}, \tag{4.4}$$

and it follows that there exists some $n \in \mathbf{Z}/\mathrm{N}\mathbf{Z}$ such that

$$\Big|\frac{1}{\mathrm{I}}\sum_{x\in n+m[\mathrm{I}]} f(x)\Big| \geqslant \frac{\beta}{2}. \tag{4.5}$$

In order to ensure that $m[\mathrm{I}]$ is a proper arithmetic progression modulo N, it suffices to ensure that $\mathrm{MI} < \mathrm{N}$. Since we also want I to be as large as possible (to get a progression $n + m[\mathrm{I}]$ which is also as large as possible), we take M and I of size comparable to $\mathrm{N}^{1/2}$, for instance

$$\mathrm{M} = \lfloor \mathrm{N}^{1/2} \rfloor, \qquad \mathrm{I} = \Big\lfloor \frac{\beta\mathrm{M}}{4\pi} \Big\rfloor.$$

The conclusion (4.5) then gives the first assertion of the proposition with $\mathrm{P} = n+m[\mathrm{I}]$, since $|\mathrm{P}| = |\mathrm{I}| \geqslant \beta\mathrm{N}^{1/2}/2^6$ (say, if N large enough). If f is real-valued, then define

$$\mathrm{F}(n) = \frac{1}{\mathrm{I}}\sum_{x\in n+m[\mathrm{I}]} f(x),$$

for $n \in \mathbf{Z}/\mathrm{N}\mathbf{Z}$, and observe that from $\widehat{f}(0) = 0$, it follows that

$$\sum_{n\in\mathbf{Z}/\mathrm{N}\mathbf{Z}} \mathrm{F}(n) = 0,$$

so that (4.4) leads to

$$\sum_{n\in\mathbf{Z}/\mathrm{N}\mathbf{Z}} (\mathrm{F}(n) + |\mathrm{F}(n)|) \geqslant \frac{\beta\mathrm{N}}{2},$$

which implies the existence of some n for which $\mathrm{F}(n) \geqslant \beta/4$. □

Remark 4.3.10. One can argue that the need to restrict to "short intervals" in $\mathbf{Z}/\mathrm{N}\mathbf{Z}$, and essentially reducing these to intervals in $\mathbf{Z}$, means that the argument might better have been framed from the start in the group of integers. Indeed, this is how it is presented in [84, Lemma 10.25]. (The reader should note that if we sum over *all* $n+m[\mathrm{I}]$, then the multiplicity of covering is equal to N instead of I, which means that the final steps do not achieve the desired result.)

We summarize the outcome of the previous steps in the following corollary.

Corollary 4.3.11. *Let* N *be an odd integer. Let* $\mathrm{A} \subset \mathbf{Z}/\mathrm{N}\mathbf{Z}$ *be a subset and let* $\alpha = |\mathrm{A}|/\mathrm{N}$. *One of the following statements holds:*

(1) *We have* $\mathrm{N} \leqslant 2\alpha^{-2}$.

(2) *The set* A *contains a proper* 3*-term arithmetic progression.*

(3) *There exists an odd integer* $\mathrm{M} \geqslant \frac{\alpha^2}{2^{10}}\mathrm{N}^{1/2}$ *and a subset* $\mathrm{B} \subset \mathbf{Z}/\mathrm{M}\mathbf{Z}$ *which contains no* 3*-term arithmetic progressions such that* $\beta = |\mathrm{B}|/\mathrm{M}$ *satisfies*

$$\beta \geqslant \alpha + \frac{\alpha^2}{2^6}.$$

Proof. Let $\psi_{\mathrm{A}} = \varphi_{\mathrm{A}} - \alpha$. Proposition 4.3.7 states that, if both (1) and (2) are false, then we can find $a \in \mathbf{Z}/\mathrm{N}\mathbf{Z}$ non-zero such that

$$\Big|\sum_{x\in\mathbf{Z}/\mathrm{N}\mathbf{Z}} \psi_{\mathrm{A}}(x)e\Big(\frac{ax}{\mathrm{N}}\Big)\Big| \geqslant \frac{\alpha^2}{2^4}.$$

We then apply Proposition 4.3.9 to ψ_{A} and $\beta = \alpha^2/2^4$; we obtain an arithmetic progression $\mathrm{P} \subset \mathbf{Z}/\mathrm{N}\mathbf{Z}$ of size $\geqslant \alpha^2\mathrm{N}^{1/2}/2^{10}$ such that

$$\sum_{x\in\mathrm{P}} \psi_{\mathrm{A}}(x) \geqslant \frac{\alpha^2|\mathrm{P}|}{2^6},$$

which translates to

$$\frac{|\mathrm{P}\cap\mathrm{A}|}{|\mathrm{P}|} \geqslant \alpha + \frac{\alpha^2}{2^6}.$$

The arithmetic progression P is Freiman 2-isomorphic to $\mathbf{Z}/\mathrm{M}\mathbf{Z}$ for some integer $\mathrm{M} \geqslant |\mathrm{P}|$, which we may assume to be odd (by adding an extra element if needed), and the subset $\mathrm{P}\cap\mathrm{A}$ is then identified with a subset B of $\mathbf{Z}/\mathrm{M}\mathbf{Z}$ without proper 3-term arithmetic progression, which shows that the third statement holds. □

Remark 4.3.12. If α is sufficiently close to 1 (so that $\alpha^2/2^6 + \alpha > 1$, which happens for $\alpha > 99/100$, for instance), we already see that the last alternative is impossible.

We finally complete the proof of Roth's Theorem. Using a Freiman 2-isomorphism, it is enough to prove that

$$\mathscr{F}_3(\mathbf{Z}/\mathrm{N}\mathbf{Z}) \ll \frac{\mathrm{N}}{\log\log\mathrm{N}}$$

for N odd.

Let $\mathrm{N} \geqslant 1$ be an odd integer and $\mathrm{A} \subset \mathbf{Z}/\mathrm{N}\mathbf{Z}$ a subset which does not contain a proper 3-term arithmetic progression. In Corollary 4.3.11, this means that the possible outcomes are the first and third statements. Whenever the third applies, we can apply the corollary to the subset $\mathrm{B} \subset \mathbf{Z}/\mathrm{M}\mathbf{Z}$ thus provided, and continue this process. The

next lemma shows that this can only be done finitely many times, because the density β given by the corollary must be $\leqslant 1$.

LEMMA 4.3.13. *Let $c > 0$ be a real number. Let $(\alpha_i)_{1\leqslant i\leqslant k}$ be a finite family of real numbers such that $0 < \alpha_1 \leqslant 1$ and $\alpha_{i+1} \geqslant \alpha_i + c\alpha_i^2$ for $i \geqslant 1$. If $\alpha_k \leqslant 1$, then $k \leqslant \lceil 2(c\alpha_1)^{-1}\rceil$.*

PROOF. For $0 \leqslant j \leqslant k-1$, the assumption implies by induction that

$$\alpha_{1+j} \geqslant \alpha_1 + jc\alpha_1^2,$$

and in particular

$$\alpha_{1+j} \geqslant 2\alpha_1$$

if $j \geqslant (c\alpha_i)^{-1}$.

By induction again, this first step gives $\alpha_{1+j} \geqslant 2^m\alpha_1$ for $0 \leqslant j \leqslant k-1$ such that

$$j \geqslant \frac{1}{c\alpha_i}\Big(1 + \frac{1}{2} + \cdots + \frac{1}{2^{m-1}}\Big).$$

Since $2^m\alpha_1 > 1$ for $m = \lceil \log(\alpha_1^{-1})/\log 2\rceil$, the assumption implies that the corresponding inequality cannot hold, which means that

$$k \leqslant \frac{1}{c\alpha_1}\Big(1 + \frac{1}{2} + \cdots + \frac{1}{2^{m-1}}\Big) \leqslant \lceil 2(c\alpha_1)^{-1}\rceil,$$

as claimed. □

We denote by k the number of times the corollary can be applied with the third alternative occuring; by the lemma, with $c = 2^8$, we have

$$k \leqslant \lceil (2^{-7}\alpha)^{-1}\rceil \leqslant \frac{c}{\alpha}$$

We then have a sequence of odd integers $(\mathrm{N}_i)_{1\leqslant i\leqslant k}$ and subsets $\mathrm{A}_i \subset \mathbf{Z}/\mathrm{N}_i\mathbf{Z}$ of density $\alpha_i = |\mathrm{A}_i|/\mathrm{N}_i$, with $\mathrm{N}_1 = \mathrm{N}$, $\mathrm{A}_1 = \mathrm{A}$, such that

$$\mathrm{N}_i \geqslant \alpha_i^2 2^{-10}\mathrm{N}_{i-1}^{1/2}, \qquad \alpha_i \geqslant \alpha_{i-1} + 2^{-6}\alpha_{i-1}^2,$$

for $2 \leqslant i \leqslant k$ and

$$\mathrm{N}_k \leqslant \frac{2}{\alpha_k^2}.$$

Noting simply that $\alpha_i \geqslant \alpha_1 = \alpha$ for all $i \leqslant k$, these properties imply by induction that

$$\beta^k \mathrm{N}^{2^{-k}} \leqslant \mathrm{N}_k \leqslant \frac{2}{\alpha^2} \leqslant \beta^{-1},$$

where $\beta = \alpha^2 2^{-10}$, and hence

$$-k\log 2 + \log\log \mathrm{N} \leqslant \log\log(\beta^{-k-1}) = \mathrm{O}(\log k).$$

Since $k \leqslant c\alpha^{-1}$, this translates to

$$\alpha = \mathrm{O}\Big(\frac{1}{\log\log \mathrm{N}}\Big),$$

which concludes the proof of the theorem.

EXERCISE 4.3.14. For positive integers n_0, n and k, we write $\mathrm{P}_{n_0,n}(k)$ for the k-term arithmetic progression $\{n_0, n_0+n, \ldots, n_0+(k-1)n\}$ in positive integers.

(1) Let $\gamma > 0$ be a real number. Show that there exists an integer $\mathrm{N}_1 \geqslant 1$ with the following property: if $\mathrm{N} \geqslant \mathrm{N}_1$ and $\mathrm{A} \subset [\mathrm{N}]$ satisfies $|\mathrm{A}| \geqslant \gamma\mathrm{N}$, then A contains elements a, b and c with $a + c = 2b$ and $a \neq c$.

(2) Let A be a set of positive integers. Let $k \geqslant 1$ be an integer and $\gamma > 0$ a real number. Show that there exists an integer $\mathrm{K} \geqslant 1$ such that any proper k-term arithmetic progression P of positive integers with $k \geqslant \mathrm{K}$ and $|\mathrm{P} \cap \mathrm{A}| \geqslant \gamma k$ contains a proper 3-term progression which is also contained in A.

In the remainder of the exercise, we fix a real number $\delta > 0$, an integer $\mathrm{N} \geqslant 1$ and a subset $\mathrm{A} \subset [\mathrm{N}]$ such that $|\mathrm{A}| \geqslant \delta \mathrm{N}$.

(3) Let $k \geqslant 1$ be an integer. Show that if a is such that $kn < \delta \mathrm{N}/k$, then we have

$$\sum_{\substack{n_0 \geqslant 1 \\ n_0+(k-1)n \leqslant \mathrm{N}}} |\mathrm{P}_{n_0,n}(k) \cap \mathrm{A}| \geqslant \delta k\Big(1 - \frac{2}{k}\Big)\mathrm{N}.$$

(Hint: for given $a \in \mathrm{A}$, show that if $kn \leqslant a \leqslant \mathrm{N} - kn$, then a belongs to k among those arithmetic progressions, then estimate how many a satisfy this property.)

(4) For given $n \geqslant 1$, let $\mathscr{G}_n$ be the set of integers $n_0 \geqslant 1$ such that

$$|\mathrm{P}_{n_0,n}(k) \cap \mathrm{A}| \geqslant \frac{\delta k}{2}.$$

Show that

$$\sum_{\substack{n_0 \geqslant 1 \\ n_0+(k-1)n \leqslant \mathrm{N}}} |\mathrm{P}_{n_0,n}(k) \cap \mathrm{A}| \leqslant \frac{\delta k \mathrm{N}}{2} + k|\mathscr{G}_n|.$$

(5) Deduce that if $kn < \delta \mathrm{N}/k$, then we have

$$|\mathscr{G}_n| \geqslant \frac{\delta \mathrm{N}}{4}.$$

(6) Show that the number of values of (n_0, n) such that $|\mathrm{P}_{n_0,n}(k) \cap \mathrm{A}| \geqslant \delta k/2$ is at least $\delta^2 \mathrm{N}^2/(4k^2)$.

(7) Let (a, b, c) be elements of A such that $a + c = 2b$ and $a < c$. Show that if (n_0, n) are such that $\{a, b, c\} \subset \mathrm{P}_{n_0,n}(k)$, then $b - a$ divides n.

(8) Deduce that the number of (n_0, n) such that $\{a, b, c\} \subset \mathrm{P}_{n_0,n}(k)$ is bounded by a constant depending only on k.

(9) Conclude that there exists $\mathrm{N}_2 \geqslant 1$ and $c > 0$, depending only on δ, such that if $\mathrm{N} \geqslant \mathrm{N}_2$, then A contains at least $c\mathrm{N}^2$ different arithmetic progressions of length 3. (Hint: apply the preceeding results for a value $k = \mathrm{K}$ given by an application of (b).)

The result of this exercise is known of *Varnavides's Theorem* (see [85]); a similar argument applies to Szemerédi's Theorem, and shows that a "weak" statement of existence of at least one k-term progression in any suitably dense set in fact implies the existence of *many* progressions.

4.4. Gowers norms

The success of Roths's approach to 3-terms arithmetic progressions, in contrast with the intricacy of Szemerédi's work, and the quantitative weaknesses of Furstenberg's approach, lead naturally to wonder if some variant of his methods based on Fourier analysis could possibly work with 4-term and longer arithmetic progressions. A basic example shows that this can certainly not be straightforward, in the sense that classical Fourier analysis is unable to capture the corresponding analytic quantities.

EXAMPLE 4.4.1. Let $p \geqslant 5$ be a prime number. Consider the quadrilinear form

$$\mathrm{AP}_4 \colon \mathrm{C}(\mathbf{F}_p)^4 \to \mathbf{C}$$

given by

$$\mathrm{AP}_4(f_1, \ldots, f_4) = \frac{1}{p^2} \sum_{a_0 \in \mathbf{F}_p} \sum_{a \in \mathbf{F}_p} f_1(a_0) f_2(a_0 + a) f_3(a_0 + 2a) f_4(a_0 + 3a)$$

which one wants to use to count arithmetic progressions of length 4. We claim that there is *no* inequality of the type

$$|\mathrm{AP}_4(f_1, \ldots, f_4)| \leqslant \mathrm{C} \|f_1\| \, \|f_2\| \, \|f_3\| \, \|\widehat{f_4}\|_\infty$$

for some C *independent of* p, so that Corollary 4.3.6 has no direct analogue. This follows from the simple example of

$$f_1(x) = e(x^2/p), \quad f_2(x) = e(-3x^2/p), \quad f_3(x) = e(3x^2/p), \quad f_4(x) = e(-x^2/p).$$

Indeed, since

$$(a_0^2 - 3(a_0 + a)^2 + 3(a_0 + 2a)^2 - (a_0 + 3a)^2 = 0$$

for all choices of a_0 and a, we have

$$f_1(a_0) f_2(a_0 + a) f_3(a_0 + 2a) f_4(a_0 + 3a) = e\Big(\frac{1}{p}(a_0^2 - 3(a_0 + a)^2 + 3(a_0 + 2a)^2 - (a_0 + 3a)^2)\Big) = 1$$

and therefore

$$\mathrm{AP}_4(f_1, f_2, f_3, f_4) = 1.$$

On the other hand, we have $\|f_i\| = 1$ for $1 \leqslant i \leqslant 3$, but

$$\|\widehat{f_4}\|_\infty = \mathrm{O}(p^{-1/2}).$$

Gowers [43] has also described examples of sets whose characteristic functions have small norm, and for which the number of 4-terms arithmetic progressions is significantly *smaller* than for random sets of the same density (a possibility which he had in fact previously conjectured not to be possible, see [41, Conj. 4.1]).

DEFINITION 4.4.2 (Gowers norms). Let G be a finite abelian group.

(1) Let $h \in \mathrm{G}$. The h-translation operator γ_h is the linear map $\mathrm{C}(\mathrm{G}) \to \mathrm{C}(\mathrm{G})$ defined by

$$\gamma_h(f)(x) = f(x + h),$$

for all $f \in \mathrm{C}(\mathrm{G})$ and $x \in \mathrm{G}$. The discrete multiplicative h-derivative on G is the map $\tau_h \colon \mathrm{C}(\mathrm{G}) \to \mathrm{C}(\mathrm{G})$ such that

$$\tau_h(f)(x) = f(x + h)\overline{f(x)} = \gamma_h(f)(x)\overline{f(x)}$$

for all $f \in \mathrm{C}(\mathrm{G})$ and $x \in \mathrm{G}$.

(2) The family of Gowers functionals is the family of maps $f \mapsto \|f\|_{(k)}$ on $\mathrm{C}(\mathrm{G})$ defined inductively for integers $k \geqslant 1$ by

$$\|f\|_{(1)} = |\,\mathbf{E}(f)| = \Big|\frac{1}{|\mathrm{G}|} \sum_{x \in \mathrm{G}} f(x)\Big|$$

for $f \in \mathrm{C}(\mathrm{G})$ and

$$\|f\|_{(k+1)}^{2^{k+1}} = \mathop{\mathbf{E}}_h\Big(\|\tau_h(f)\|_{(k)}^{2^k}\Big) = \frac{1}{|\mathrm{G}|} \sum_{h \in \mathrm{G}} \|\tau_h(f)\|_{(k)}^{2^k},$$

for $k \geqslant 1$ and $f \in \mathrm{C}(\mathrm{G})$

As we will see, the Gowers functionals are norms for $k \geqslant 2$ (and the definition shows that the Gowers functional for $k = 1$ is a seminorm), but this is not completely obvious, so we will soon change the name to reflect this fact.

Example 4.4.3. (1) For $k = 2$, spelling out the definition, we obtain the formula

$$\begin{aligned}\|f\|_{(2)}^4 &= \frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\|\tau_h(f)\|_{(1)}^2 \\ &= \frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\Big|\frac{1}{|\mathrm{G}|}\sum_{x\in\mathrm{G}} f(x+h)\overline{f(x)}\Big|^2.\end{aligned}$$

Expanding the square, this expression gives

$$\|f\|_{(2)}^4 = \frac{1}{|\mathrm{G}|^3}\sum_{x,y\in\mathrm{G}}\sum_{h\in\mathrm{G}} f(y+h)\overline{f(y)f(x+h)}f(x).$$

It is customary to rewrite the formula by changing the variables from (x, y, h) to (x, h_1, h_2) where $h_2 = h$ and $y = x + h_1$; this gives

$$\|f\|_{(2)}^4 = \frac{1}{|\mathrm{G}|^3}\sum_{x\in\mathrm{G}}\sum_{h_1,h_2\in\mathrm{G}} f(x+h_1+h_2)\overline{f(x+h_1)f(x+h_2)}f(x), \tag{4.6}$$

which is interpreted as a sum (or weighted count) over all "rectangles" in G, i.e., all sets of the form $\{x, x + h_1, x + h_2, x + h_1 + h_2\}$.

We can also interpret the definition as an average of inner products of f with its translates:

$$\|f\|_{(2)}^4 = \frac{1}{\mathrm{G}}\sum_{h\in\mathrm{G}}|\langle\gamma_h(f), f\rangle|^2.$$

This immediately suggests an expression in terms of the Fourier transform: by unitarity, we have $\langle\tau_h f, f\rangle = \langle\widehat{\tau_h f}, \widehat{f}\rangle$, and since

$$\widehat{\gamma_h f}(\chi) = \frac{1}{|\mathrm{G}|^{1/2}}\sum_{x\in\mathrm{G}} f(x+h)\overline{\chi(x)} = \sum_{y\in\mathrm{G}} f(y)\overline{\chi(y-h)} = \chi(h)\widehat{f}(\chi)$$

for any $\chi \in \widehat{\mathrm{G}}$ and $h \in \mathrm{G}$, it follows that

$$\langle\gamma_h f, f\rangle = \frac{1}{|\mathrm{G}|}\sum_{\chi\in\widehat{\mathrm{G}}}\overline{\chi(h)}|\widehat{f}(\chi)|^2$$

hence

$$\begin{aligned}\|f\|_{(2)}^4 &= \frac{1}{\mathrm{G}}\sum_{h\in\mathrm{G}}\Big|\frac{1}{|\mathrm{G}|}\sum_{\chi}\overline{\chi(h)}|\widehat{f}(\chi)|^2\Big|^2 \\ &= \frac{1}{|\mathrm{G}|^2}\sum_{\chi_1,\chi_2\in\widehat{\mathrm{G}}}|\widehat{f}(\chi_1)|^2|\widehat{f}(\chi_2)|^2\frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\overline{\chi_1(h)}\chi_2(h) = \frac{1}{|\mathrm{G}|^2}\sum_{\chi\in\widehat{\mathrm{G}}}|\widehat{f}(\chi)|^4. \qquad (4.7)\end{aligned}$$

This shows that the second Gowers functional of a function f is the same as the L^4-norm of its Fourier transform (up to normalization); in particular, it already shows that the second Gowers functional is in fact a norm on $\mathrm{C}(\mathrm{G})$.

From the point of view of arithmetic progressions, this computation will also reveal that the approach through Gowers norms, in the case of progressions of length 3, is

equivalent to the approach through Fourier analysis. This link will however also be shown to disappear for the higher Gowers seminorms.

(2) The formula (4.6) can be generalized by induction to give a formula for the Gowers functional as a weighted count for (product of values of) the function f over "hypercube patterns" in G. To define these, we will denote by $\mathbf{B}_k$ the set $\{0,1\}^k \subset \mathbf{Z}^k$, which we interpret as the set of binary strings of length k. A k-dimensional hypercube in G is a subset of the form

$$x + \mathbf{B}_k \cdot h$$

for some $x \in \mathrm{G}$ and some $h \in \mathrm{G}^k$, where we denote

$$c \cdot h = c_1 h_1 + \cdots + c_k h_k.$$

Thus, for instance, if $k = 2$, a 2-dimensional cube is a set of the form

$$\{x, x + h_1, x + h_2, x + h_1 + h_2\},$$

corresponding to the expressions appearing in formula (4.6).

If f is real-valued, the general formula for $\|f\|_{(k)}$ takes the form

$$\|f\|_{(k)}^{2^k} = \frac{1}{|\mathrm{G}|^{k+1}} \sum_{x \in \mathrm{G}} \sum_{h \in \mathrm{G}^k} \prod_{c \in \mathbf{B}_k} f(x + c \cdot h),$$

where $|c| = c_1 + \cdots + c_k$ for $c \in \mathbf{B}_k$.

If f is not necessarily real-valued, we must handle the fact that sometimes conjugates of values of f appear. Let $\sigma \colon \mathbf{C} \to \mathbf{C}$ be the complex conjugation function $z \mapsto \bar{z}$. Then one gets

$$\|f\|_{(k)}^{2^k} = \frac{1}{|\mathrm{G}|^{k+1}} \sum_{x \in \mathrm{G}} \sum_{h \in \mathrm{G}^k} \prod_{c \in \mathbf{B}_k} \sigma^{|c|}(f(x + c \cdot h)). \tag{4.8}$$

(3) For some functions, one can compute exactly the Gowers functionals. For instance, note that $\|1\|_{(k)} = 1$ for all k in the case of the constant function 1. Moreover, it is straightforward (by induction) that when f is bounded by 1 in modulus, then we have also $\|f\|_{(k)} \leqslant 1$ for all $k \geqslant 1$.

Exercise 4.4.4. Let $\chi \in \widehat{\mathrm{G}}$ be a non-trivial character of G. Show that $\|\chi\|_{(1)} = 0$ and $\|\chi\|_{(k)} = 1$ for $k \geqslant 2$.

The basic properties of Gowers functionals and their link with arithmetic progressions are given by the next propositions.

Proposition 4.4.5 (Gowers). *Let* G *be a finite abelian group.*

(1) *For any integer $k \geqslant 1$ and $f \in \mathrm{C}(\mathrm{G})$, we have*

$$\|f\|_{(k)} \leqslant \|f\|_{(k+1)}.$$

(2) *For any $k \geqslant 2$, the Gowers functional $f \mapsto \|f\|_{(k)}$ is a norm on* $\mathrm{C}(\mathrm{G})$.

In order to prove this, it is convenient to use the map

$$\mathscr{G}_k \colon \mathrm{C}(\mathrm{G})^{\mathbf{B}_k} \to \mathrm{C}(\mathrm{G})$$

(which we call the k-th Gowers operator) defined by

$$\mathscr{G}_k((f_c)_{c\in\mathbf{B}_k}) = \frac{1}{|\mathrm{G}|^{k+1}}\sum_{x\in\mathrm{G}}\sum_{h\in\mathrm{G}^k}\prod_{c\in\mathbf{B}_k}\sigma^{|c|}(f_c(x+c\cdot h))$$
$$= \mathop{\mathbf{E}}_{x,h}\Big(\prod_{c\in\mathbf{B}_k}\sigma^{|c|}(f_c(x+c\cdot h))\Big)$$

for any $(f_c)_{c\in\mathbf{B}_k}$ in $\mathrm{C}(\mathrm{G})^{\mathbf{B}_k}$. This map is linear with respect to each variable f_c such that $|c|$ is even, and conjugate-linear with respect to the other variables. In particular, it is linear in each variable when only linear combinations with real coefficients are considered.

By the formula (4.8) above, we have $\|f\|_{(k)}^{2^k} = \mathscr{G}_k(f)$, where on the right-hand side, we consider the family (f_c) with $f_c = f$ for all c.

Lemma 4.4.6 (Gowers–Cauchy–Schwarz inequality). *Let $k \geqslant 1$ be an integer. For any family $(f_c)_{c\in\mathbf{B}_k}$, we have*

$$|\mathscr{G}_k((f_c))| \leqslant \prod_{c\in\mathbf{B}_k}\|f_c\|_{(k)}.$$

Proof. For a set X and $x \in \mathrm{X}^k$, we will denote here by $x' = (x_1,\dots,x_{k-1})$ the projection of x to the first $k-1$ coordinates.

Let $1 \leqslant j \leqslant k$ be fixed. For any $c \in \mathbf{B}_k$ and $i \in \{0,1\}$, write $\widetilde{c}_i$ for the binary string identical to c except that the j-th digit, counted from the left, is replaced by i. (So, for instance, if $k = 5$, $c = 00101$ and $j = 3$, then $\widetilde{c}_0 = 00001$ and $\widetilde{c}_1 = 00101$.)

Let $\boldsymbol{f} = (f_c) \in \mathrm{C}(\mathrm{G})^{\mathbf{B}_k}$. The key step is the inequality

$$|\mathscr{G}_k(\boldsymbol{f})| \leqslant \sqrt{\mathscr{G}_k(\boldsymbol{f}_0)\mathscr{G}_k(\boldsymbol{f}_1)}, \tag{4.9}$$

where $\boldsymbol{f}_i = (f_{\widetilde{c}_i})_{c\in\mathbf{B}_k}$. Note that these new families depend on j, but we omit this from the notation for simplicity. The point is that, even if all f_c are distinct functions, both $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ have at most 2^{k-1} distinct coordinates, and that each f_c appears in a single one of these families.

To prove (4.9), we assume $j = k$ to simplify the notation, leaving to the reader to check the other cases (it amounts to using below the j-th variable h_j instead of h_k). By rearranging the definition, we see that $\mathscr{G}_k(\boldsymbol{f})$ is equal to

$$\frac{1}{|\mathrm{G}|^{k+1}}\sum_{h'\in\mathrm{G}^{k-1}}\sum_{x\in\mathrm{G}}\prod_{c'\in\mathbf{B}_{k-1}}\sigma^{|c'|}(f_{c',0}(x+c'\cdot h'))\sum_{h_k\in\mathrm{G}}\prod_{c'\in\mathbf{B}_{k-1}}\sigma^{|c'|+1}(f_{c',1}(x+c'\cdot h'+h_k)).$$

For each value of h' and x, we make the change of variable $y = x + h_k$ in the last sum, and this gives

$$\frac{1}{|\mathrm{G}|^{k+1}}\sum_{h'\in\mathrm{G}^{k-1}}\sum_{x\in\mathrm{G}}\prod_{c'\in\mathbf{B}_{k-1}}\sigma^{|c'|}(f_{c',0}(x+c'\cdot h'))\sum_{y\in\mathrm{G}}\prod_{c'\in\mathbf{B}_{k-1}}\sigma^{|c'|+1}(f_{c',1}(y+c'\cdot h')).$$

We can then apply the Cauchy–Schwarz inequality to the sum over h', and this leads to the inequality $|\mathscr{G}_k(\boldsymbol{f}| \leqslant \sqrt{\mathrm{A}_0\mathrm{A}_1}$, where

$$\mathrm{A}_i = \frac{1}{|\mathrm{G}|^{k+1}}\sum_{h'\in\mathrm{G}^{k-1}}\Big|\sum_{x\in\mathrm{G}}\prod_{c'\in\mathbf{B}_{k-1}}\sigma^{|c'|}(f_{c',0}(x+c'\cdot h'))\Big|^2,$$

for $i \in \{0,1\}$.

Using $|z|^2 = z\bar{z}$, we have

$$\Big|\sum_{x\in \mathrm{G}} \prod_{c'\in \mathbf{B}_{k-1}} \sigma^{|c'|}(f_{c',0}(x+c'\cdot h'))\Big|^2 = \prod_{c'\in \mathbf{B}_{k-1}} \sigma^{|c'|}(f_{c',0}(x+c'\cdot h')) \prod_{c'\in \mathbf{B}_{k-1}} \sigma^{|c'|+1}(f_{c',0}(y+c'\cdot h'))$$

for each $h' \in \mathrm{G}^{k-1}$. Now note that

$$x + c'\cdot h' = x + (c', 0)\cdot(h', y-x), \qquad y + c'\cdot h' = x + (c', 1)\cdot(h', y-x)$$

for any given $(x, y) \in \mathrm{G}^2$, $h' \in \mathrm{G}^{k-1}$ and $c' \in \mathbf{B}_{k-1}$. If we define $h_k = y - x$ and write $h = (h', h_k)$, this gives

$$\prod_{c'\in \mathbf{B}_{k-1}} \sigma^{|c'|}(f_{c',0}(x + c'\cdot h')) \prod_{c'\in \mathbf{B}_{k-1}} \sigma^{|c'|+1}(f_{c',0}(y + c'\cdot h')) =$$
$$\prod_{c'\in \mathbf{B}_{k-1}} \sigma^{|c'|}(f_{c',0}(x + (c', 0)\cdot h)) \prod_{c'\in \mathbf{B}_{k-1}} \sigma^{|c'|+1}(f_{c',0}(y + (c', 1)\cdot h))$$
$$= \prod_{c\in \mathbf{B}_k} \sigma^{|c|}(f_{c',0}(x + c\cdot h)).$$

Thus

$$\mathrm{A}_i = \frac{1}{|\mathrm{G}|^{k+1}} \sum_{h'\in \mathrm{G}^{k-1}} \sum_{x\in \mathrm{G}} \sum_{y\in \mathrm{G}} \prod_{c'\in \mathbf{B}_{k-1}} \sigma^{|c'|}(f_{c',0}(x + c'\cdot h')) \prod_{c'\in \mathbf{B}_{k-1}} \sigma^{|c'|+1}(f_{c',0}(y + c'\cdot h'))$$
$$= \frac{1}{|\mathrm{G}|^{k+1}} \sum_{h\in \mathrm{G}^{k}} \sum_{x\in \mathrm{G}} \prod_{c\in \mathbf{B}_k} \sigma^{|c|}(f_{c',0}(x + c\cdot h)) = \mathscr{G}_k(\boldsymbol{f}_i),$$

which concludes the proof of (4.9), in the case $j = k$.

We now iterate (4.9), but apply it in order with $j = k$, then $j = k - 1$, and so on. This leads to an inequality

$$|\mathscr{G}_k(\boldsymbol{f})| \leqslant \prod_{c\in \mathbf{B}_k} \mathscr{G}_k(\boldsymbol{f}_c)^{2^{-k}},$$

where each tuple $\boldsymbol{f}_c \in \mathrm{C}(\mathrm{G})^{\mathbf{B}_k}$ has all coefficients the same, and one checks that they are equal to the function f_c. Using $\|f_c\|_{(k)} = \mathscr{G}_k(\boldsymbol{f}_c)$ then concludes the proof of the lemma. □

Proof of Proposition 4.4.5. (1) We consider $(f_c) \in \mathrm{C}(\mathrm{G})^{\mathbf{B}_k}$ defined by $f_{c',0} = f$ and $f_{c',1} = 1$ for all $c \in \mathbf{B}_k$, and apply Lemma 4.4.6. The right-hand side of the inequality is then $\|f\|_{(k)}^{2^{k-1}}$ (since $\|1\|_{(k)} = 1$). On the left-hand side, note that

$$\sigma^{|c|}(f_c(x + c\cdot h)) = \begin{cases} \sigma^{|c'|}(f(x + c'\cdot h) & \text{if } c = (c', 0) \\ 1 & \text{if } c = (c', 1), \end{cases}$$

so that

$$\mathscr{G}_k((f_c)) = \|f\|_{(k-1)}^{2^{k-1}},$$

by the formula (4.8) again. The bound follows.

(2) Suppose that $k \geqslant 2$. Let $f \in \mathrm{C}(\mathrm{G})$. It is straightforward to check by induction that $\|\lambda f\|_{(k)} = |\lambda|\|f\|_{(k)}$ for any $\lambda \in \mathbf{C}$. Moreover, $\|f\|_{(k)} \geqslant 0$, and if equality holds, we deduce from (1) by induction that $\|f\|_{(2)} = 0$. But then the L^4-norm of the Fourier transform of f is zero by (4.7), so $\widehat{f} = 0$, and hence $f = 0$.

It remains only to check the triangle inequality to conclude that the Gowers seminorms are norms for $k \geqslant 2$. Let f and g be elements of $\mathrm{C}(\mathrm{G})$. Then, denoting also by $f + g$ the

family in $\mathrm{C}(\mathrm{G})^{\mathbf{B}_k}$ with all coefficients equal to $f+g$, we know that $\|f+g\|_{(k)}^{2^k}$ is equal to $\mathscr{G}_k(f+g)$. Using the multilinearity of $\mathscr{G}_k$ (note that we have a linear combination with integral coefficients, so the fact that the Gowers operator is conjugate-linear in some variables has no consequence), we get

$$\|f+g\|_{(k)}^{2^k} = \mathscr{G}_k(f+g) = \sum_{\mathrm{I}\subset\mathbf{B}_k} \mathscr{G}_k(f_{\mathrm{I}}),$$

where $f_{\mathrm{I}} = (f_{\mathrm{I},c})$ is such that $f_{\mathrm{I},c} = f$ if $c \in \mathrm{I}$ and $f_{\mathrm{I},c} = g$ otherwise. From Lemma 4.4.6, we have

$$|\mathscr{G}_k(f_{\mathrm{I}})| \leqslant \prod_{c\in\mathbf{B}_k} \|f_{\mathrm{I},c}\|_{(k)}$$

for each $\mathrm{I} \subset \mathbf{B}_k$, and therefore

$$\sum_{\mathrm{I}\subset\mathbf{B}_k} \mathscr{G}_k(f_{\mathrm{I}}) \leqslant \sum_{\mathrm{I}\subset\mathbf{B}_k} \|f_{\mathrm{I},c}\|_{(k)} = (\|f\|_{(k)} + \|g\|_{(k)})^{|\mathbf{B}_k|},$$

from which the desired inequality follows. □

We can now prove the relation between Gowers norms and counting arithmetic progressions. We first define the k-linear function $\mathrm{AP}_k\colon \mathrm{C}(\mathrm{G})^k \to \mathbf{C}$ by

$$\begin{aligned}\mathrm{AP}_k(f_1,\dots,f_k) &= \mathop{\mathbf{E}}_{a_0,a}(f_1(a_0)\cdots f_k(a_0+(k-1)a)\\ &= \frac{1}{|\mathrm{G}|^2}\sum_{a\in\mathrm{G}}\sum_{a_0\in\mathrm{G}} f_1(a_0)f_2(a_0+a)\cdots f_k(a_0+(k-1)a).\end{aligned}$$

Proposition 4.4.7. *Let* G *be a finite abelian group. Let* $k \geqslant 1$ *be an integer. Assume that* G *contains no element of order* $\leqslant k-1$. *For any* $(f_1,\dots,f_k) \in \mathrm{C}(\mathrm{G})^k$, *and any* $j \in [k]$, *we have*

$$\mathrm{AP}_k(f_1,\dots,f_k) \leqslant \Big(\prod_{i\neq j}\|f_i\|_\infty\Big)\,\|f_j\|_{(k-1)}.$$

Remark 4.4.8. (1) The numerology is thus that the Gowers k-norm controls progressions of length $k-1$.

(2) The assumption on G and k is equivalent to asking that $|\mathrm{G}|$ has no prime factor $\leqslant k-1$, or to asking that $|\mathrm{G}|$ is coprime to $(k-1)!$.

This assumption implies that an arithmetic progression $\{a_0, a_0+a, \dots, a_0+(k-1)a\}$ is proper if and only if $a \neq 0$. (Indeed, if the progression is not proper, we have an equality $a_0+ia = a_0+ja$ with $0 \leqslant i,j \leqslant k-1$; it follows that $(i-j)a = 0$, and since $|i-j| \leqslant k-1$, we then deduce that $a = 0$.)

It also implies that for any integer j with $1 \leqslant j \leqslant k-1$, the map $x \mapsto jx$ on G is bijective (since it is then injective).

Proof. The proof is by induction on k.

Suppose first that $k = 2$. For functions f_1, f_2 in $\mathrm{C}(\mathrm{G})$, we then have

$$\mathrm{AP}_2(f_1,f_2) = \frac{1}{|\mathrm{G}|^2}\sum_{a_0,a\in\mathrm{G}} f_1(a_0)f_2(a_0+a) = \frac{1}{|\mathrm{G}|^2}\sum_{a_0,b_0\in\mathrm{G}} f_1(a_0)f_2(b_0),$$

by the change of variable from (a_0,a) to (a_0,a_0+a). Thus

$$|\mathrm{AP}_2(f_1,f_2)| = \|f_1\|_{(1)}\|f_2\|_{(1)} \leqslant \|f_1\|_\infty\|f_2\|_{(1)}.$$

We now assume that $k \geqslant 3$ and that the bound holds for $k-1$. We will then prove the result for AP_k, assuming that $j = k$ and leaving to the reader the care of checking that the argument extends to all j.

We write

$$\frac{1}{|\mathrm{G}|^2}\sum_{a\in\mathrm{G}}\sum_{a_0\in\mathrm{G}} f_1(a_0)f_2(a_0+a)\cdots f_k(a_0+(k-1)a) = \frac{1}{|\mathrm{G}|}\sum_{a_0\in\mathrm{G}}\sum_{a\in\mathrm{G}} g(a_0,a)f_1(a_0)$$

where

$$g(a_0,a) = \frac{1}{|\mathrm{G}|}f_2(a_0+a)\cdots f_k(a_0+(k-1)a).$$

We can recognize here a special case of the type of sums appearing in the bilinear forms of Lemma 1.6.1 (with $\mathrm{N} = |\mathrm{G}|$ and all β_n equal to 1 there), and the reader can check that the next steps is just a repetition of the general argument. By the Cauchy–Schwarz inequality, we deduce that

$$|\mathrm{AP}_k(f_1,\dots,f_k)|^2 \leqslant \Big(\frac{1}{|\mathrm{G}|}\sum_{a_0\in\mathrm{G}}|f_1(a_0)|^2\Big)\Big(\frac{1}{|\mathrm{G}|}\sum_{a_0\in\mathrm{G}}\Big|\sum_{a\in\mathrm{G}} g(a_0,a)\Big|^2\Big).$$

The first term on the right-hand side $\|f_1\| \leqslant \|f_1\|_\infty$. Opening the square, the second is equal to

$$\frac{1}{|\mathrm{G}|}\sum_{a,b\in\mathrm{G}}\sum_{a_0\in\mathrm{G}}\overline{g(a_0,a)}g(a_0,b) = \frac{1}{|\mathrm{G}|}\sum_{a,h\in\mathrm{G}}\sum_{a_0\in\mathrm{G}}\overline{g(a_0,a)}g(a_0,a+h).$$

We now observe that, for any (a_0,a,h), we have

$$\begin{aligned}\overline{g(a_0,a)}g(a_0,a+h) &= \frac{1}{|\mathrm{G}|^2}\overline{f_2(a_0+a)}\cdots\overline{f_k(a_0+(k-1)a)}\times\\ &\qquad f_2(a_0+a+h)\cdots f_k(a_0+(k-1)(a+h))\\ &= \frac{1}{|\mathrm{G}|^2}\overline{f_2(a_0+a)}f_2(a_0+a+h)\cdots\overline{f_k(a_0+(k-1)a)}f_k(a_0+(k-1)(a+h)),\end{aligned}$$

which we recognize as equal to

$$\frac{1}{|\mathrm{G}|^2}\prod_{j=2}^{k}(\tau_{(j-1)h}f_j)(a_0+ja).$$

This implies that

$$\begin{aligned}\frac{1}{|\mathrm{G}|}\sum_{a,b\in\mathrm{G}}\sum_{a_0\in\mathrm{G}}\overline{g(a_0,a)}g(a_0,b) &= \frac{1}{|\mathrm{G}|^3}\sum_{a,b\in\mathrm{G}}\sum_{a_0\in\mathrm{G}}\prod_{j=2}^{k}(\tau_{(j-1)h}f_j)(a_0+ja)\\ &= \frac{1}{|\mathrm{G}|^3}\sum_{a,b\in\mathrm{G}}\sum_{b_0\in\mathrm{G}}\prod_{j=1}^{k-1}(\tau_{(j-1)h}f_j)(b_0+ja)\end{aligned}$$

by putting $b_0 = a_0 + a$, and we now recognize here the average

$$\frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\mathrm{AP}_{k-1}(\tau_h f_2,\dots,\tau_{(k-1)h}f_k).$$

Applying the induction hypothesis, and noting that $\|\tau_{(j-1)h}f_j\|_\infty \leqslant \|f_j\|_\infty^2$ for all j, we deduce that

$$\Big|\frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\mathrm{AP}_{k-1}(\tau_h f_2,\dots,\tau_{(k-1)h}f_k)\Big| \leqslant \frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\Big(\prod_{j=2}^{k-1}\|f_j\|_\infty\Big)^2\,\|\widetilde{f}_{k,h}\|_{(k-2)}.$$

We then observe that

$$\frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\|\widetilde{f}_{k,h}\|_{(k-2)} = \frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\|\tau_{(k-1)h}f_k\|_{(k-2)} = \frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\|\tau_h f_k\|_{(k-2)},$$

because $k-1$ is coprime to the size of G. Applying Hölder's inequality and the inductive definition of Gowers norms gives

$$\frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\|f\|_{(k-2)} \leqslant \Big(\frac{1}{|\mathrm{G}|}\sum_{h\in\mathrm{G}}\|f\|_{(k-2)}^{2^{k-2}}\Big)^{2^{-(k-2)}} = \|f\|_{(k-1)}^2,$$

and gathering the previous steps leads to

$$|\mathrm{AP}_k(f_1,\dots,f_k)| \leqslant \prod_{j=1}^{k-1}\|f_j\|_\infty\ \|f_k\|_{(k-1)},$$

as claimed. □

We have then an analogue of the trichotomy concerning arithmetic progressions which appeared previously in the case $k=3$.

Corollary 4.4.9. *Let* G *be a finite abelian group. Let* $k\geqslant 2$ *be an integer. Assume that* G *has no element of order* $\leqslant k-1$. *Let* $\mathrm{A}\subset\mathrm{G}$ *be a subset of* G, *and denote* $\alpha = |\mathrm{A}|/|\mathrm{G}|$.

Then at least one of the following properties holds:

(1) *We have*

$$|\mathrm{G}| < \frac{2}{\alpha^{k-1}}.$$

(2) *We have*

$$\|\psi_\mathrm{A}\|_{(k-2)} \geqslant \frac{\alpha^k}{2^{k+1}},$$

where $\psi_\mathrm{A} = \varphi_\mathrm{A} - \alpha$ *is the balanced characteristic function of* G.

(3) *There exists a proper* k*-term arithmetic progression in* A.

Proof. This is similar to the proof of Proposition 4.3.7. If A does not contain a k-term proper arithmetic progression, then we have

$$\mathrm{AP}_k(\varphi_\mathrm{A},\dots,\varphi_\mathrm{A}) = \frac{\alpha}{|\mathrm{G}|},$$

because the assumption implies that the only improper k-term arithmetic progressions in G are those with common difference 0 (if $ia = ja$ with $1\leqslant i\neq j\leqslant k$, then $(i-j)a=0$ implies that $a=0$).

Writing $\varphi_\mathrm{A} = \alpha+\psi_\mathrm{A}$, and using the multilinearity of AP_k, we therefore have

$$\begin{aligned}\frac{\alpha}{|\mathrm{G}|} &= \mathrm{AP}_k(\alpha+\psi_\mathrm{A},\dots,\alpha+\psi_\mathrm{A})\\ &= \alpha^k + \sum_{\substack{\mathrm{I}\subset[k]\\ \mathrm{I}\neq\emptyset}}\mathrm{AP}_k(\psi_{\mathrm{I},1},\dots,\psi_{\mathrm{I},k})\end{aligned}$$

with $\psi_{\mathrm{I},j} = \psi_{\mathrm{A}}$ if $j \in \mathrm{I}$ and $\psi_{\mathrm{I},j} = \alpha$ otherwise. If $\alpha^k \geqslant 2\alpha/|\mathrm{G}|$, then this implies that

$$\Big|\sum_{\substack{\mathrm{I}\subset[k]\\ \mathrm{I}\neq\emptyset}} \mathrm{AP}_k(\psi_{\mathrm{I},1},\dots,\psi_{\mathrm{I},k})\Big| \geqslant \frac{\alpha^k}{2},$$

hence there exists a subset $\mathrm{I} \neq \emptyset$ such that

$$|\mathrm{AP}_k(\psi_{\mathrm{I},1},\dots,\psi_{\mathrm{I},k})| \geqslant \frac{\alpha^k}{2^{k+1}}.$$

and therefore some $j \in [k]$ and functions f_i for $i \neq j$ with $|f_i| \leqslant 1$ such that

$$|\mathrm{AP}_k(f_1,\dots,f_{j-1},\psi_{\mathrm{A}},f_{j+1},\dots,f_k)| \geqslant \frac{\alpha^k}{2^{k+1}}.$$

By Proposition 4.4.7, we conclude that

$$\|\psi_{\mathrm{A}}\|_{(k)} \geqslant \frac{\alpha^k}{2^{k+1}}.$$

In summary, if neither property (2) nor property (3) is valid, we have must have $\alpha/|\mathrm{G}| \leqslant \frac{1}{2}\alpha^k$, which translates to the validity of (1), thus finishing the proof. □

This corollary shows that the problem of proving Szemerédi's Theorem for cyclic groups $\mathrm{G} = \mathbf{Z}/\mathrm{N}\mathbf{Z}$ is related to the *inverse problem for Gowers norms*: for $k \geqslant 2$ and $\beta > 0$ fixed, given a function $f \in \mathrm{C}(\mathbf{Z}/\mathrm{N}\mathbf{Z})$ which is bounded by one and satisfies

$$\|f\|_{(k)} \geqslant \beta,$$

what kind of structural information concerning f can one deduce? In the case $k = 2$ (corresponding to 3-term progressions), a suitable consequence was found in the course of proving (4.2), namely the existence of a large Fourier coefficient. This we interpret now as saying that there exists a non-zero $a \in \mathbf{Z}/\mathrm{N}\mathbf{Z}$ such that

$$\Big|\sum_{x\in\mathrm{G}} f(x)e\Big(-\frac{ax}{\mathrm{N}}\Big)\Big|$$

is very large, so that f correlates very strongly with a character of $\mathbf{Z}/\mathrm{N}\mathbf{Z}$.

The following example shows that, at the very least, there are quite a few bounded functions on $\mathbf{Z}/\mathrm{N}\mathbf{Z}$ for which $\|f\|_{(k)}$ is large.

Example 4.4.10. Let $\mathrm{N} \geqslant 1$ be an integer and let $\mathrm{P} \in \mathbf{Z}/\mathrm{N}\mathbf{Z}[\mathrm{X}]$ be a polynomial of degree $d \geqslant 0$. Define a function $f \in \mathrm{C}(\mathbf{Z}/\mathrm{N}\mathbf{Z})$ by

$$f(x) = e\Big(\frac{\mathrm{P}(x)}{\mathrm{N}}\Big),$$

which has modulus 1. We claim that $\|f\|_{(k)} = 1$ for all $k \geqslant d+1$.

The point is that for a function of this type, the function $\tau_h(f)$ is of the same kind, but for a polynomial P_h of degree $\leqslant d-1$, reflecting the fact that τ_h is similar to a (discrete) derivative; after d iterations of the differentiation process, we get a constant polynomial, for which the corresponding function has modulus 1.

To be precise, we proceed by induction on d. For $d = 0$, the function f is constant so $\|f\|_{(1)} = |f| = 1$. Assume now that $d \geqslant 1$ and that the result holds when f is defined using a polynomial of degree $\leqslant d-1$. For any $h \in \mathbf{Z}/\mathrm{N}\mathbf{Z}$ and $x \in \mathbf{Z}/\mathrm{N}\mathbf{Z}$, we have

$$\tau_h(f)(x) = e\Big(\frac{\mathrm{P}(x+h)-\mathrm{P}(x)}{\mathrm{N}}\Big) = e\Big(\frac{\mathrm{P}_h(x)}{\mathrm{N}}\Big)$$

where $\mathrm{P}_h = \mathrm{P}(\mathrm{X}+h)-\mathrm{P}(\mathrm{X})$ is a polynomial of degree at most $d-1$ (the terms of degree d cancel out in the difference). If $k \geqslant d+1$, then $k-1 \geqslant d$, so that by induction we have $\|\tau_h(f)\|_{(k-1)} = 1$, hence

$$\|f\|_{(k)}^{2^k} = \frac{1}{\mathrm{N}} \sum_{h \in \mathbf{Z}/\mathrm{N}\mathbf{Z}} \|\tau_h(f)\|_{(k-1)}^{2^{k-1}} = 1.$$

The threshold $k \geqslant d+1$ is sharp, in the sense that for $k \leqslant d$, the k-th Gowers norm of a function like f is quite small. More precisely, one can show that

$$\|f\|_{(k)} \leqslant \|f\|_{(d)} \leqslant c_d \mathrm{N}^{-\gamma_d}$$

if $k \leqslant d$, for some constants c_d and $\gamma_d > 0$. If $\mathrm{N} = p$ is a prime number and $d < p$, one can be even more precise, and prove that

$$\|f\|_{(k)} \leqslant (5(d+1))^{k+1} p^{-2^{-k}}$$

(in other words, $\gamma_d = 2^{-d}$ is possible). This was proved by Fouvry, Kowalski and Michel (see [34, Ex. 1.8]); the bound

$$\|f\|_{(k)} \leqslant \|f\|_{(d)} \leqslant (d-1)^{2^{-d}} p^{-2^{-d}},$$

which is only slightly weaker, can be proved elementarily (see for instance [84, Exercise 11.1.12]).

The various estimates also show clearly that the various Gowers norms are really different: for each k, there are functions f with $\|f\|_{(k)}$ small, but $\|f\|_{(k+1)}$ large.

Example 4.4.11. Results like those mentioned in the previous example lead to non-trivial cases where the trichotomy can be used to prove the existence of proper k-term arithmetic progressions in some interesting subsets of $\mathbf{Z}/p\mathbf{Z}$. In particular, the results of Fouvry, Kowalski and Michel can be used to prove that if $\mathrm{A} \subset \mathbf{Z}/p\mathbf{Z}$ is "uniformly algebraically defined" for an infinite sequence of prime numbers p, and satisfies $|\mathrm{A}| \gg p^{1-\gamma_k}$ for some (very small) constant $\gamma_k > 0$, then A contains a k-term progression if p is large enough.

The precise definition of "uniformly algebraically defined" is somewhat technical, but the following are examples of suitable sets A:

- A is the set of squares modulo p;
- More generally, $\mathrm{A} = \mathrm{P}(\mathbf{F}_p)$ for some fixed integral polynomial $\mathrm{P} \in \mathbf{Z}[\mathrm{X}]$ of degree $\geqslant 1$ (for instance, the set of $x \in \mathbf{F}_p$ such that $x = y^3 + y + 1$ for some y also in $\mathbf{F}_p$).

The point is that, in each of these examples (and many others), the balanced characteristic function of A has a description as a linear combination with "small" coefficients of functions $f \in \mathrm{C}(\mathbf{F}_p)$ of very special algebraic nature, the so-called "tame trace functions" over finite fields. For such a function f, it is proved in [34, Cor. 1.6] that $\|f\|_{(k)} \ll p^{-2^{-k}}$ for *all* values of k, where the implied constant depends only on a complexity invariant of f. The latter is uniformly bounded as p varies when A is one of the sets above, and the consequence is that we then obtain an estimate of the form

$$\|\psi_{\mathrm{A}}\|_{(k)} \ll p^{-2^{-k}}$$

for all primes p, where the implied constant depends only on the way the sets $\mathrm{A} \subset \mathbf{F}_p$ are "uniformly" defined.

From the point of view above, these examples are highly suggestive, because the proof goes through showing that *either* a function f of "algebraic type" satisfies the

bound $\|f\|_{(k)} \ll p^{-2^{-k}}$, *or* there exists a polynomial $\mathrm{P} \in \mathbf{F}_p[\mathrm{X}]$, of degree $d \leqslant k-1$, such that

$$\Big|\sum_{x\in\mathbf{F}_p} f(x)e\Big(-\frac{\mathrm{P}(x)}{p}\Big)\Big| \gg p.$$

In other words, for such functions, we have essentially a "perfect" inverse theorem, where having large Gowers norm is exactly equivalent to having high correlation with a function of the simplest type which itself has large Gowers norms, namely with $g(x) = e(\mathrm{P}(x)/\mathrm{N})$ for some polynomial P. The estimate for $\|\psi_{\mathrm{A}}\|_{(k)}$ mentioned above follows then from the fact that, in this case, one can prove that ψ_{A} *does not* correlate with such functions.

(As a last remark, the results of [34] rely ultimately on extremely deep results, namely Deligne's most general form of the Riemann Hypothesis over finite fields [20], one of the most sophisticated and important result in number theory of the 20th Century.)

APPENDIX A

Reminders

We summarize here, with precise references when needed, a number of elementary facts that are used in the rest of the text.

A.1. Dirichlet's Theorem

The following result is one of the first and most useful applications of the pigeon-hole principle (see [21, p. 636], where already the higher-dimensional version is proved this way).

THEOREM A.1.1 (Dirichlet). *Let $x \in [0,1]$ be a real number. For any integer $\mathrm{Q} \geqslant 1$, there exists a positive integer $q \geqslant \mathrm{Q}$ and an integer a such that*

$$\left|x - \frac{a}{q}\right| \leqslant \frac{1}{q\mathrm{Q}} \leqslant \frac{1}{\mathrm{Q}^2}.$$

PROOF. Among the fractional parts of the $\mathrm{Q}+1$ numbers 0, x, …, $\mathrm{Q}x$, at least two must fall in the same of the Q intervals

$$[0, \mathrm{Q}^{-1}[, \quad [\mathrm{Q}^{-1}, 2\mathrm{Q}^{-1}[, \quad \dots, \quad [(\mathrm{Q}-1)\mathrm{Q}^{-1}, 1[,$$

say $\{ix\}$ and $\{jx\}$ with $0 < leqi < j \leqslant \mathrm{Q}$ satisfy

$$\frac{k}{\mathrm{Q}} \leqslant \{ix\} \leqslant \frac{k+1}{\mathrm{Q}}, \qquad \frac{k}{\mathrm{Q}} \leqslant \{jx\} \leqslant \frac{k+1}{\mathrm{Q}}.$$

Let $q = j - i$, so that $1 \leqslant j \leqslant \mathrm{Q}$. We have $\{ix\} = ix - u$ and $\{jx\} = jx - v$ for some integers u and v, and so the inequality $|\{jx\} - \{ix\}| \leqslant 1/\mathrm{Q}$ leads to

$$|(j-i)x - (v-u)| \leqslant \frac{1}{\mathrm{Q}}$$

from which the result follows with $a = v - u$ by dividing by q. □

REMARK A.1.2. By dividing through common factors, if need be, one can also assume in Dirichlet's Theorem that a is coprime to q.

A.2. Summation by parts

The following lemmas are variants of the standard formula of integration by parts, and are often known as "summation by parts".

LEMMA A.2.1. (1) *Let $(a_n)_{n\geqslant 1}$ be a sequence of complex number and define*

$$\mathrm{A}(x) = \sum_{1 \leqslant n \leqslant x} a_n.$$

for $x \geqslant 0$. Let $f\colon [0, +\infty[\to \mathbf{C}$ be a function of class C^1. For $x \geqslant 0$, we then have

$$\sum_{1 \leqslant n \leqslant x} a_n f(n) = \mathrm{A}(x) f(x) - \int_1^x \mathrm{A}(t) f'(t) dt. \tag{A.1}$$

(2) *Let* $\mathrm{N} \geqslant 1$ *be an integer and let* $(a_n)_{1\leqslant n\leqslant \mathrm{N}}$ *and* $(b_n)_{1\leqslant n\leqslant \mathrm{N}}$ *be sequences of complex numbers. Define*

$$\mathrm{A}_k = \sum_{n=1}^{k} a_n$$

for $0 \leqslant k \leqslant \mathrm{N}$*, in particular* $\mathrm{A}_0 = 0$*. We then have*

$$\sum_{n=1}^{\mathrm{N}} a_n b_n = a_{\mathrm{N}} b_{\mathrm{N}} - \sum_{n=1}^{\mathrm{N}-1} \mathrm{A}_n (b_{n+1} - b_n). \tag{A.2}$$

Proof. We leave the first statement as an exercise to the reader. For the second, just observe that $a_n = \mathrm{A}_n - \mathrm{A}_{n-1}$, hence

$$\begin{aligned}\sum_{n=1}^{\mathrm{N}} a_n b_n &= \sum_{n=1}^{\mathrm{N}} (\mathrm{A}_n - \mathrm{A}_{n-1}) b_n \\ &= \sum_{n=1}^{\mathrm{N}} \mathrm{A}_n b_n - \sum_{n=0}^{\mathrm{N}-1} \mathrm{A}_n b_{n+1} = \mathrm{A}_{\mathrm{N}} b_{\mathrm{N}} - \sum_{n=1}^{\mathrm{N}-1} \mathrm{A}_n (b_{n+1} - b_n),\end{aligned}$$

as claimed. □

A.3. Probability theory

We use frequently the basic Markov and Chebychev inequalities.

Another useful, somewhat less standard, inequality is the following:

Lemma A.3.1. *Let* X *be a bounded non-negative random variable. Let* $\mathrm{M} \geqslant 0$ *be such that* $\mathrm{X} \leqslant \mathrm{M}$ *and* $\alpha \geqslant 1$ *such that*

$$\mathbf{E}(\mathrm{X}) \geqslant \frac{\mathrm{M}}{\alpha}.$$

We have

$$\mathbf{P}\Big(\mathrm{X} \geqslant \frac{\mathrm{M}}{2\alpha}\Big) \geqslant \frac{1}{2\alpha}. \tag{A.3}$$

Proof. Let φ be the characteristic function of the event $\{\mathrm{X} \geqslant \frac{1}{2}\alpha^{-1}\mathrm{M}\}$. Using the bounds

$$0 \leqslant (1-\varphi)\mathrm{X} \leqslant \frac{\mathrm{M}}{2\alpha}, \qquad 0 \leqslant \varphi \mathrm{X} \leqslant \mathrm{M}\varphi,$$

and the fact that $\mathbf{E}(\varphi) = \mathbf{P}(\mathrm{M} \geqslant \frac{1}{2}\alpha^{-1}\mathrm{M})$, we obtain the inequalities

$$\frac{\mathrm{M}}{\alpha} \leqslant \mathbf{E}(\mathrm{X}) = \mathbf{E}((1-\varphi)\mathrm{X}) + \mathbf{E}(\varphi \mathrm{X}) \leqslant \frac{\mathrm{M}}{2\alpha} + \mathrm{M}\,\mathbf{P}(\mathrm{M} \geqslant \tfrac{1}{2}\alpha^{-1}\mathrm{M}),$$

from which the conclusion follows. □

The inclusion-exclusion principle (in probability theory, this probably goes back to the study of the "problème des rencontres" in the early 18th Century, by de Montfort, N. Bernoulli II and de Moivre, see [82]) is often useful, as well as its truncated versions which give upper or lower bounds only.

Proposition A.3.2. *Let* μ *be a positive measure on a set* X*, and let* $(\mathrm{X}_i)_{i\in \mathrm{I}}$ *be a finite family of measurable subsets of* X*. We have*

$$\mu\Big(\bigcup_{i\in \mathrm{I}} \mathrm{X}_i\Big) = \sum_{\substack{\mathrm{J}\subset \mathrm{I} \\ \mathrm{J}\neq \emptyset}} (-1)^{|\mathrm{J}|-1} \mu\Big(\bigcup_{j\in \mathrm{J}} \mathrm{X}_j\Big).$$

Moreover, if $j \geqslant 1$ is an odd integer, then

$$\mu\Big(\bigcup_{i\in \mathrm{I}} \mathrm{X}_i\Big) \leqslant \sum_{\substack{\mathrm{J}\subset \mathrm{I}\\ 1\leqslant|\mathrm{J}|\leqslant j}} (-1)^{|\mathrm{J}|-1}\mu\Big(\bigcup_{j\in \mathrm{J}} \mathrm{X}_j\Big), \tag{A.4}$$

whereas if $j \geqslant 1$ is even, then

$$\mu\Big(\bigcup_{i\in \mathrm{I}} \mathrm{X}_i\Big) \leqslant \sum_{\substack{\mathrm{J}\subset \mathrm{I}\\ 1\leqslant|\mathrm{J}|\leqslant j}} (-1)^{|\mathrm{J}|-1}\mu\Big(\bigcup_{j\in \mathrm{J}} \mathrm{X}_j\Big). \tag{A.5}$$

Proof. Let φ_i be the characteristic function of X_i, and for any subset J of I, let φ_{J} be the characteristic function of the intersection of the sets X_j for $j \in \mathrm{J}$. Consider the function

$$\widetilde{\varphi} = \sum_{\substack{\mathrm{J}\subset \mathrm{I}\\ \mathrm{J}\neq\emptyset}} (-1)^{|\mathrm{J}|-1}\varphi_{\mathrm{J}}.$$

We will show that $\widetilde{\varphi}$ coincides with the characteristic function φ of the union of the X_i for $i \in \mathrm{I}$; computing then the integral with respect to μ, the formula follows.

To check the claim, take $x \in \mathrm{X}$. Define $\mathrm{K} \subset \mathrm{I}$ to be the set of those $k \in \mathrm{I}$ such that $x \in \mathrm{X}_k$. If K is empty, then we have $\varphi(x) = 0$, and also $\widetilde{\varphi}(x) = 0$ since $\varphi_{\mathrm{J}}(x) = 0$ for any subset J. On the other hand, if K is not empty, then $\varphi(x) = 1$, and we have $\varphi_{\mathrm{J}}(x) = 1$ if and only if $\mathrm{J} \subset \mathrm{K}$, so that

$$\widetilde{\varphi}(x) = \sum_{\substack{\mathrm{J}\subset \mathrm{K}\\ \mathrm{J}\neq\emptyset}} (-1)^{|\mathrm{J}|-1}1 = -\Big(\sum_{\mathrm{J}\subset \mathrm{K}} (-1)^{|\mathrm{J}|} - 1\Big) = -\Big(\sum_{j=0}^{|\mathrm{K}|} \binom{|\mathrm{K}|}{j}(-1)^j - 1\Big) = 1 = \varphi(x)$$

by the binomial theorem.

(Another version of the proof goes as follows: with the same notation, the characteristic function $1 - \varphi$ of the *complement* of the union of the X_i's satisfies

$$1 - \varphi = \prod_{i\in \mathrm{I}} (1 - \varphi_i) = \sum_{\mathrm{J}\subset \mathrm{I}} (-1)^{|\mathrm{J}|}\varphi_{\mathrm{J}},$$

and since $\varphi_\emptyset = 1$, we get again

$$\varphi = -\sum_{\substack{\mathrm{J}\subset \mathrm{I}\\ \mathrm{J}\neq\emptyset}} (-1)^{|\mathrm{J}|}\varphi_{\mathrm{J}},$$

and conclude as before.)

We also see from this proof that (A.4) and (A.5) result from the inequalities

$$\sum_{0\leqslant i\leqslant j} (-1)^i \binom{k}{i} \leqslant 0, \qquad \text{if } j \leqslant k \text{ is odd},$$

$$\sum_{0\leqslant i\leqslant j} (-1)^i \binom{k}{i} \geqslant 0, \qquad \text{if } j \leqslant k \text{ is even},$$

which hold for alternating sums of binomial coefficients with $k \geqslant 0$. These follow immediately from the exact formula

$$\sum_{0\leqslant i\leqslant j} (-1)^i \binom{k}{i} = (-1)^j \binom{k-1}{j},$$

for $0 \leqslant j \leqslant k$. This last statement may be proved, for instance, by induction on k. □

EXAMPLE A.3.3. In the simplest example where X_i are finite sets and μ is the counting measure on X, we get

$$\Big|\bigcup_{i\in \mathrm{I}} \mathrm{X}_i\Big| = \sum_{i\in \mathrm{I}} |\mathrm{X}_i| - \sum_{\substack{\{i,j\}\subset \mathrm{I}\\ i\neq j}} |\mathrm{X}_i \cap \mathrm{X}_j| + \sum_{\substack{\{i,j,k\}\subset \mathrm{I}\\ i,j,k \text{ distinct}}} |\mathrm{X}_i \cap \mathrm{X}_j \cap \mathrm{X}_k| + \cdots,$$

and the first two truncations lead to the inequalities

$$\sum_{i\in \mathrm{I}} |\mathrm{X}_i| - \frac{1}{2}\sum_{i,j\in \mathrm{I}} |\mathrm{X}_i \cap \mathrm{X}_j| \leqslant \Big|\bigcup_{i\in \mathrm{I}} \mathrm{X}_i\Big| \leqslant \sum_{i\in \mathrm{I}} |\mathrm{X}_i|.$$

REMARK A.3.4. The truncated form of the inclusion and exclusion principle is only the simplest possible refinement of the exact formula. Variants of this idea are at the source of the fundamental *sieve methods* in analytic number theory, which provide one of the most powerful techniques in multiplicative number theory, and especially in the study of prime numbers. For further information on this topic, the best source is the book [36] of Friedlander and Iwaniec.

A.4. Polynomials

The following standard fact will be used in the proof of Dvir's Theorem.

PROPOSITION A.4.1. *Let* E *be a field. Let* $d \geqslant 1$ *and* $k \geqslant 0$ *be integers. The vector space* $\mathscr{X}_{d,k}(\mathrm{E})$ *of homogeneous polynomials of degree* k *in* d *variables over* E *has dimension*

$$\binom{d+k-1}{d-1} = \binom{d+k-1}{k}.$$

For instance, for $d = 1$, the space $\mathscr{X}_{1,k}(\mathrm{E})$ is generated by X^k, hence has dimension $1 = \binom{k}{0}$, and for $d = 2$, the space $\mathscr{X}_{2,k}(\mathrm{E})$ is spanned by

$$\mathrm{X}^k, \quad \mathrm{X}^{k-1}\mathrm{Y}, \quad \cdots, \quad \mathrm{X}\mathrm{Y}^{k-1}, \quad \mathrm{Y}^k,$$

and has dimension $k+1 = \binom{k+1}{1}$.

PROOF. The question is to count the number of tuples $(k_1, \ldots, k_d)$ of non-negative integers which satisfy

$$k_1 + \cdots + k_d = k,$$

since these correspond bijectively to the monomials

$$\mathrm{X}_1^{k_1} \cdots \mathrm{X}_d^{k_d}$$

of degree k which form a basis of $\mathscr{X}_{d,k}(\mathrm{E})$.

There are different ways to do this. We can for instance consider the generating series

$$f(z) = \sum_{k\geqslant 0} \dim(\mathscr{X}_{d,k}(\mathrm{E}))z^k,$$

as a formal power series with integer coefficients. Using the above observation, we can express the dimension as the sum

$$\sum_{\substack{k_1+\cdots+k_d=k\\ k_i\geqslant 0}} 1,$$

and by inserting it into the generating function, we obtain the expression

$$f(z) = \Big(\sum_{k\geqslant 0} z^k\Big)^d = \frac{1}{(1-z)^d}.$$

If we denote $g(z) = (1-z)^{-1}$, we have then the equality

$$f(z) = \frac{1}{(d-1)!} g^{(d-1)}(z),$$

which, by computing the derivative, is equal to

$$\frac{1}{(d-1)!} \sum_{k \geqslant 0} (k+d-1) \cdots (k+1) z^k,$$

so that by equating coefficients we obtain the formula

$$\dim(\mathscr{X}_{d,k}(\mathrm{E})) = \frac{(k+d-1)\cdots(k+1)}{(d-1)!} = \binom{k+d-1}{d-1},$$

as claimed. □

Another extremely useful result concerning polynomials is the Schwarz–Zippel Lemma, which controls the number of possible zeros of multivariable polynomials in certain sets, generalizing the fact that a non-zero one-variable polynomial has at most as many zeros in a field as its degree.

Proposition A.4.2. *Let* E *be a field. Let* $\mathrm{S} \subset \mathrm{E}$ *be a finite subset of* E*. For any integer* $d \geqslant 1$ *and any non-zero polynomial* $f \in \mathrm{E}[\mathrm{X}_1, \ldots, \mathrm{X}_d]$*, we have*

$$|\{x \in \mathrm{S}^d \mid f(x) = 0\}| \leqslant \deg(f) |\mathrm{S}|^{d-1}.$$

Proof. The proof proceeds by induction on the number d of variables. For $d = 1$, the result is just the statement that a non-zero polynomial of degree $k \geqslant 0$ has at most k zeros in E.

We now assume that $d \geqslant 2$ and that the statement is valid for polynomials in $d-1$ variables. We write

$$f = \sum_{j=0}^{k} a_j \mathrm{X}_d^j$$

where $a_j \in \mathrm{E}[\mathrm{X}_1, \ldots, \mathrm{X}_{d-1}]$ and $a_k \neq 0$. For a given $x' \in \mathrm{S}^{d-1}$, the equation $f(x', x_d) = 0$ with unknown $x_d \in \mathrm{S}$ has at most k roots if $a_k(x') \neq 0$, and at most $|\mathrm{S}|$ roots if $a_k(x') = 0$. Thus

$$|\{x \in \mathrm{S}^d \mid f(x) = 0\}| \leqslant k|\mathrm{S}|^{d-1} + |\mathrm{S}|\,|\{x' \in \mathrm{S}^{d-1} \mid a_k(x') = 0\}|.$$

By induction, we get

$$|\{x' \in \mathrm{S}^{d-1} \mid a_k(x') = 0\}| \leqslant \deg(a_k)|\mathrm{S}|^{d-2},$$

and we conclude that

$$|\{x \in \mathrm{S}^d \mid f(x) = 0\}| \leqslant k|\mathrm{S}|^{d-1} + \deg(a_k)|\mathrm{S}|^{d-1} \leqslant \deg(f)|\mathrm{S}|^{d-1},$$

since $k + \deg(a_k) \leqslant \deg(f)$, which completes the induction. □

A.5. Graphs

We use only simple undirected graphs in this book.

Definition A.5.1 (Graph). A *graph* is a pair (V, E) of sets, called *vertices* and *edges*, respectively, such that E is a set of subsets of V with $|e| = 2$ for any $e \in \mathrm{E}$.

A graph is said to be finite if V is finite, in which case E is also finite.

Given a vertex $x \in \mathrm{V}$, the *neighbours* of x are the $y \in \mathrm{V}$ such that $\{x, y\}$ is an edge of the graph. The degree of the graph at x is the number of neighbours of x, and is

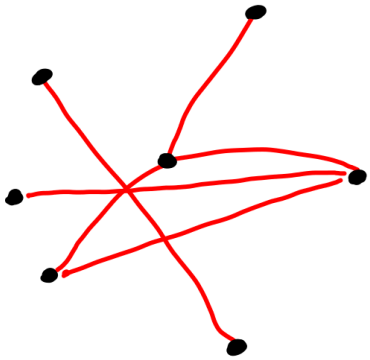


FIGURE A.1. A graph with crossing number 4.

denoted $\deg(x)$. If all vertices have the same degree, say $k \geqslant 0$, then the graph is said to be k-regular.

In the proof of the Szemerédi–Trotter Theorem, we use a result concerning *planar graphs*. How to define this in general would require some care if one allows curved realizations of the edges, but we will only need the simplest examples, and we can restrict to the following definition.

DEFINITION A.5.2. Let $\gamma = (\mathrm{V}, \mathrm{E})$ be a graph. A *planar realization* of γ is a subset $\mathrm{P} \subset \mathbf{R}^2$ given with a bijection $f\colon \mathrm{V} \to \mathrm{P}$ such that, for any distinct edges $\{u, v\}$ and $\{x, y\}$ of γ, the open segments $]f(u), f(v)[$ and $]f(x), f(y)[$ in $\mathbf{R}^2$ are either disjoint or intersect in a single point.

The *crossings* of the planar realization are the pairs of edges for which the corresponding open segments intersect.

REMARK A.5.3. We allow the possibility that the segments corresponding to three (or more) edges intersect at a single point; in that case, the intersection point contributes as many crossings as there are pairs of edges involved (for instance, 3 for a triple point, 6 for a quadruple point), see Figure A.1 for an illustration.

THEOREM A.5.4 (Crossing number inequality). *Let $\gamma = (\mathrm{V}, \mathrm{E})$ be a finite graph such that $|\mathrm{E}| \geqslant 4|\mathrm{V}|$. Any planar realization of γ has at least $2^{-6}|\mathrm{E}|^3|\mathrm{V}|^{-2}$ crossings.*

This is due to Ajtai, Chvátal, Newborn and Szemerédi and to Leighton, independently; see for instance [84, Th. 8.1] for a proof. The key fact linking the combinatorics of graphs with the geometry is Euler's formula for the Euler–Poincaré characteristic of a finite planar graph: in any planar realization of a finite graph γ, we have

$$|\mathrm{F}| - |\mathrm{E}| + |\mathrm{V}| = 2$$

where F is the number of "faces" of the planar realization.

A.6. Finite fields

We review here the basic theory of finite fields. The main facts are summarized in the following theorem, where we recall that for any field E, finite or not, there is a unique integer p, which is either 0 or a prime number, such that $p \cdot 1 = 0$. This integer is called the characteristic of the field.

THEOREM A.6.1 (Finite fields). (1) *For any prime number p, the ring $\mathbf{Z}/p\mathbf{Z}$ is a finite field, which is denoted $\mathbf{F}_p$; it has characteristic p.*

(2) *For any prime number p and any integer $\nu \geqslant 1$, there exists a finite field* E *with* $|\mathrm{E}| = p^\nu$. *Conversely, if* E *is a finite field, its order is of the form p^ν for some prime p and some integer $\nu \geqslant 1$; the prime p is the characteristic of the field* E.

(3) *For any finite field of size p^ν for some prime p and some integer $\nu \geqslant 1$, the map $x \mapsto x^p$ from* E *to itself is a field automorphism. In fact, this holds for* any *field of characteristic p.*

(4) *If* E *and* F *are finite fields of the same size, then they are isomorphic.*

(5) *If* E *is a finite field of size p^ν for some prime p and some integer $\nu \geqslant 1$, then* E *contains a unique subfield* F *of size p^d for any positive integer d dividing ν, and contains no other subfields. Moreover we have*

$$\mathrm{F} = \{x \in \mathrm{E} \mid x^{p^d} = x\}.$$

We sketch the classical argument very quickly, using basic abstract algebra to do so. Readers who are not fully familiar with such ideas can find a detailed elementary discussion in the book [78, Ch. 4, 5] of Soundararajan.

Sketch of proof. We first note that (1) is valid; since it is clear from the definition that $p \cdot 1 = 0$ in $\mathbf{Z}/p\mathbf{Z}$, we only need to check that this is a field. Two ways to see this are:

(1) (abstractly) because of the general fact that the quotient of a commutative ring with unit by a maximal ideal is a field, and $p\mathbf{Z}$ is a maximal ideal in $\mathbf{Z}$;

(2) (concretely) because one can deduce from the euclidean algorithm for the computation of the gcd of two integers that for any integer a not divisible by p, there exist integers u and v such that

$$au + pv = 1$$

(Bezout's equation), hence the class of u modulo p is a multiplicative inverse of the class of a modulo p. Since proving that every non-zero element of $\mathbf{Z}/p\mathbf{Z}$ is the only axiom of fields really requiring a proof for this ring, this proves that it is a field.

Next, we recall that (3) holds because, defining $f(x) = x^p$ for x in a field E, we obtain a map $\mathrm{E} \to \mathrm{E}$ such that $f(0) = 0$, $f(1) = 1$ and $f(xy) = f(x)f(y)$ for all x and y in E, without restrictions on E. Moreover, we know that the binomial coefficient $\binom{p}{a}$ is divisible by p for any integer a not divisible by p, so the binomial theorem gives

$$(x+y)^p = x^p + \sum_{j=1}^{p-1} \binom{p}{j} x^j y^{p-j} + y^p = x^p + y^p$$

for all x, y in E, whenever $p \cdot 1$ is zero in a field E, i.e., whenever E has characteristic p.

We will now show that (1) and (3) imply all other parts of the statement, using one abstract fact from algebra: any field E is contained in an algebraically closed field F, which is unique up to isomorphism.

We fix a prime number p and apply this fact to $\mathrm{E} = \mathbf{F}_p$, obtaining an algebraic closure F of E. The field F also has characteristic p, and we denote by $f\colon \mathrm{F} \to \mathrm{F}$ the automorphism $x \mapsto x^p$. Then, for any integer $\nu \geqslant 1$, composition of f with itself ν times (i.e., the map $f^\nu\colon x \mapsto x^{p^\nu}$ on F) is also an automorphism of fields, and it follows formally that the set of fixed points

$$\mathrm{E}_\nu = \{x \in \mathrm{F} \mid f^\nu(x) = x\}$$

is a subfield of F containing E. Crucially, the equation $f^\nu(x) = x$ is equivalent to $x^{p^\nu} - x = 0$, hence is a polynomial equation of degree p^ν. It follows that $|\mathrm{E}_\nu| \leqslant p^\nu$, and in fact that there is equality, because F is algebraically closed, and the polynomial $\mathrm{X}^{p^\nu} - \mathrm{X}$ is *separable* (i.e., it does not have multiple roots, which can be checked by noting that its derivative is -1, which is never zero). This establishes the existence part in (2).

For the converse, we just observe that any finite field containing E is an E-vector space; it has finite dimension if it is finite, and if this dimension is $\nu \geqslant 1$, then looking at the expansion in a fixed basis, it is elementary that it has size p^ν. This gives the second assertion in (2).

(TODO) □

This abstract theorem does not by itself provide a way to describe concretely a finite field with a given number $q = p^\nu$ of elements, and to perform computations in it (e.g., to solve polynomials equations in the field). In order to this, the usual method is to find a polynomial $f \in \mathbf{F}_p[\mathrm{X}]$ which is irreducible of degree ν, and to observe that basic algebra implies then that the quotient ring $\mathrm{E} = \mathbf{F}_p[\mathrm{X}]/f\mathbf{F}_p[\mathrm{X}]$ (of the polynomial ring with coefficients in $\mathbf{F}_p$ by the principal ideal generated by f) is then a finite field with p^ν elements, hence gives a model of such a field. If we write

$$f = \mathrm{X}^\nu + a_{\nu-1}\mathrm{X}^{\nu-1} + \cdots + a_1\mathrm{X} + a_0,$$

with coefficients $a_i \in \mathbf{F}_p$, and denote by α a root of this polynomial (this can be seen formally as the class of the indeterminate X in the quotient ring $\mathbf{F}_p[\mathrm{X}]/f\mathbf{F}_p[\mathrm{X}]$, or as a root in some algebraic closure of $\mathbf{F}_p$), then any element x of E have a unique expression of the form

$$x = \lambda_0 + \lambda_1\alpha + \cdots + \lambda_{\nu-1}\alpha^{d-1},$$

with $\lambda_i \in \mathbf{F}_p$. These expressions can be added and multiplied using the usual rules of algebra and the unique extra property that

$$\alpha^\nu = -(a_{\nu-1}\alpha^{\nu-1} + \cdots + a_1\alpha + a_0),$$

which reflects the fact that $f(\alpha) = 0$.

Remark A.6.2. It is often customary to denote by $\mathbf{F}_q$ a finite field with q elements. One must be careful when using this notation that although such a field is well-defined up to isomorphism, this isomorphism is not unique. This means that "concrete" computations performed by two different mathematicians using two different representations of a field with q elements might not be identical. Since one can show quite easily that the number of irreducible polynomials of degree $\nu \geqslant 2$ with coefficients in $\mathbf{F}_p$ is about p^ν/d, there are *many* ways to represent finite fields.

Example A.6.3. We consider some computations in a field E with 8 elements. To do this as described above, we need a polynomial f of degree 3 with coefficients in $\mathbf{F}_2 = \mathbf{Z}/p\mathbf{Z}$ which is irreducible. It is not too hard to find such an f, since a reducible polynomial of degree 3 has a root, and in $\mathbf{F}_2$, only two elements can be such a root. So for instance, the polynomial

$$f = \mathrm{X}^3 + \mathrm{X} + 1$$

will do, since $f(0) = 1$ and $f(1) = 1 + 1 + 1 = 1$ in $\mathbf{F}_2$. Denoting by α a root of f, we can represent the elements of E as combinations

$$x = a_0 + a_1\alpha + a_2\alpha^2$$

with $a_i \in \mathbf{F}_2$. For instance, we can then compute

$$(1 + \alpha^2)(1 + \alpha) = 1 + \alpha + \alpha^2 + \alpha^3 = 1 + \alpha + \alpha^2 + (1 + \alpha) = \alpha^2,$$

since $\alpha^3 = -(1+\alpha) = 1+\alpha$ (remember that the coefficients are in $\mathbf{F}_2$ to $-1 = 1$).

How do we compute the inverse of some non-zero element x of E? A systematic way (applicable in general finite fields) is to find an equation satisfied by x with coefficients in $\mathbf{F}_2$. There is always such an equation of degree at most 3 (in general, at most ν if the field has p^ν elements), and if it has the form

$$b_0 + b_1 x + b_2 x^2 + b_3 x^3 = 0,$$

with $b_i \in \mathbf{F}_2$ and b_0 non-zero (which is not a restriction since we would otherwise divide as often as needed by the non-zero element x to obtain another equation, of smaller degree even, with this property), then we get

$$x(b_3 x^2 + b_2 x + b_1) = -b_0, \quad \text{so} \quad \frac{1}{x} = -\frac{b_3}{b_0}x^2 - \frac{b_2}{b_0}x - \frac{b_1}{b_0},$$

which we can further express in terms of α.

For instance, let $x = 1+\alpha$. We have

$$x^2 = 1+\alpha^2, \qquad x^3 = (1+\alpha)(1+\alpha^2) = \alpha^2$$

(by the computation done before), so we get the equation

$$x^3 + x^2 + 1 = 0,$$

which gives

$$\frac{1}{x} = -x^2 - x = (1+\alpha^2) + 1 + \alpha = \alpha + \alpha^2.$$

Exercise A.6.4. (1) Find a concrete "model" of a finite field E with $25 = 5^2$ elements, using a root α of a quadratic equation with coefficients in $\mathbf{F}_5$.

(2) Find the expression of $(\alpha-3)^{-1}$ and α^3 in terms of α; determine if α^4+1 is invertible or not, and if Yes, compute $(3\alpha+2)/(\alpha^4+1)$. (Note that whether α^4+1 is invertible or not will depend on the choice of the quadratic polynomial chosen to "compute" in E.)

(3) Determine if the equation $\mathrm{X}^3 + 3\mathrm{X} + 2 = 0$ has a root in E, or not. (The answer to this question will *not* depend on the choice of defining equation.)

It is also useful and important to know the group structure of finite fields and their groups of invertible elements.

Proposition A.6.5. *Let* E *be a finite field.*

(1) *The group* $\mathrm{E}^\times = \mathrm{E} - \{0\}$ *of invertible elements of* E *is cyclic of order* $|\mathrm{E}| - 1$.

(2) *Writing* $|\mathrm{E}| = p^\nu$ *for some prime number* p *and some integer* $\nu \geqslant 1$, *the additive group of* E *is isomorphic to* $(\mathbf{Z}/p\mathbf{Z})^\nu$.

Proof. (1) The key point is that, for any integer $n \geqslant 1$, the number of elements of order dividing n in $\mathrm{E}^\times$ is the number of solutions of the equation $x^n - 1 = 0$; since this is a polynomial equation of degree n with unknown in a field, this number must be $\leqslant n$. One can then check that a finite abelian group with this property is necessarily cyclic. (Abstractly, this is because a non-cyclic group always contains a subgroup H isomorphic to $(\mathbf{Z}/d\mathbf{Z})^2$ for some integer $d \geqslant 2$, and all d^2 elements of H satisfy $x^d = 1$.)

(2) is simply because E is a vector space of dimension ν over the field $\mathbf{F}_p$. □

A.7. Harmonic analysis on finite abelian groups

We summarize in this section the basic properties of characters of finite abelian groups, and the resulting discrete form of the Fourier transform and Fourier analysis.

Definition A.7.1. Let G be a group. A *character* of G is a group morphism from G to the group $\mathbf{C}^\times$ of non-zero complex numbers.

The set of characters of G is denoted $\widehat{\mathrm{G}}$.

The set $\widehat{\mathrm{G}}$ has itself a group structure, and in fact it is abelian: for any characters χ_1 and χ_2, the product $x \mapsto \chi_1(x)\chi_2(x)$ is also a character, since

$$\chi_1(xy)\chi_2(xy) = \chi_1(x)\chi_1(y)\chi_2(x)\chi_2(y) = (\chi_1\chi_2)(x)(\chi_1\chi_2)(y),$$

and we have $\chi_1\chi_2 = \chi_2\chi_1$ because the product in $\mathbf{C}$ is commutative.

Since moreover the function 1 on G is a character such that $1 \cdot \chi = \chi$ for any character χ, and since $\chi^{-1}\colon x \mapsto \chi(x)^{-1}$ is a character such that $\chi \cdot \chi^{-1} = 1$, we have checked all required properties to see that $\widehat{\mathrm{G}}$ is an abelian group. When G itself is commutative, it is called the *dual group*.

We will consider mostly finite groups. In this case, since any element x has finite order (we have $x^r = 1$ for some $r \geqslant 1$), the values $\chi(x)$ of any character of G are roots of unity, and in particular have modulus 1. This also implies that $\chi^{-1}(x) = \overline{\chi(x)}$ for any $x \in \mathrm{G}$ and $\chi \in \widehat{\mathrm{G}}$.

Example A.7.2. Let $q \geqslant 1$ be a positive integer, and let $\mathrm{G} = \mathbf{Z}/q\mathbf{Z}$ be the cyclic group of order q. The characters of G are then easy to determine: since G is generated by a single element (by the class of 1, for instance), a morphism $\chi\colon \mathrm{G} \to \mathbf{C}^\times$ is uniquely determined by the image $\chi(1) \in \mathbf{C}^\times$ (we then have $\chi(n) = \chi(1)^n$ for all n). This image cannot be arbitrary: since the generator has order q, we have also $\chi(1)^q = 1$, i.e., $\chi(1)$ must be a q-th root of unity. So there exists an integer h such that $\chi(1) = \exp(2i\pi h/q)$, and the character is then given by

$$\chi(n \,(\mathrm{mod}\, q)) = \chi(1)^n = e\Big(\frac{nh}{q}\Big)$$

for all $n \in \mathbf{Z}$.

Conversely, it is elementary that the right-hand side, as a function of n, is actually well-defined for n taken modulo q (i.e., it doesn't change when n is replaced by $n + kq$ for any $k \in \mathbf{Z}$), and it defines a character of G. Moreover, the characters defined by two integers h_1 and h_2 are equal if and only if h_1 and h_2 are equal modulo q (since the q-th roots of unity are equal only under this condition).

We conclude that the characters of $\mathrm{G} = \mathbf{Z}/q\mathbf{Z}$ are parameterized uniquely by $\mathbf{Z}/q\mathbf{Z}$. In other words, the dual group of G is isomorphic to G.

The key fact of harmonic analysis on finite abelian groups is the following statement:

Theorem A.7.3. *Let* G *be a finite abelian group. The set of characters of* G *is an orthonormal basis of the space* $\mathrm{C}(\mathrm{G})$ *of complex-valued functions on* G*, with the inner product*

$$\langle f, g\rangle = \mathop{\mathbf{E}}_{x\in\mathrm{G}} f\bar{g} = \frac{1}{|\mathrm{G}|}\sum_{x\in\mathrm{G}} f(x)\overline{g(x)}.$$

In particular, we have $|\widehat{\mathrm{G}}| = \dim \mathrm{C}(\mathrm{G}) = |\mathrm{G}|$.

In order to prove this, we need one fundamental lemma, which informally means that there are "enough" characters. Note that in this proof, although G is abelian, we usually write the group operation multiplicatively.

LEMMA A.7.4. *Let* G *be a finite abelian group and let* $y \neq 1$ *be a non-trivial element of* G. *There exists a character* χ *of* G *such that* $\chi(y) \neq 1$.

PROOF. We first give a proof assuming that G has the form

$$\mathrm{G} = \mathbf{Z}/q_1\mathbf{Z} \times \cdots \times \mathbf{Z}/q_k\mathbf{Z},$$

for some integer $k \geqslant 1$ and integers $q_i \geqslant 2$. (This is in fact no restriction, since every finite abelian group is isomorphic to a group of this type, but the proof of this is rather more complicated; the point is however that many abelian groups will be given *a priori* as such a product.)

Writing $y = (y_1, \ldots, y_k)$ with $y_i \in \mathbf{Z}/q_i\mathbf{Z}$, we select i such that $y_i \neq 0$ (in $\mathbf{Z}/q_i\mathbf{Z}$). Then we can define

$$\chi(x) = e\Big(\frac{x_i}{q_i}\Big)$$

for any $x = (x_i) \in \mathrm{G}$. Indeed, it is straightforward that this is a character of G (see Example A.7.2), and $\chi(y) = e(y_i/q_i) \neq 1$.

Now we give a general proof, which is a bit more abstract, but also relies essentially on Example A.7.2. We define H to be the largest subgroup (in size) of G containing y such that there exists a character $\eta \in \widehat{\mathrm{H}}$ with $\eta(y) \neq 1$. This is well-defined, because there is at least one such subgroup: y is in the cyclic subgroup that it generates, and if we denote by q the order of y in G, then we can define η on the subgroup generated by y by $\eta(y^n) = e(n/q)$ for $n \in \mathbf{Z}$; this is again more or less the same as Example A.7.2.

The goal is now to show that in fact $\mathrm{H} = \mathrm{G}$. Indeed, if this is not the case, then we could find $z \in \mathrm{G}$ that is not in H. Let H' be the subgroup generated by H and z; we have $|\mathrm{H}'| > |\mathrm{H}|$, and we will get a contradiction by extending the character η of H to H'.

Let Z be the subgroup generated by z, and let q be its order. We first consider the product group $\mathrm{H} \times \mathrm{Z}$, and observe that for any q-th root of unity θ, there is a character $\widetilde{\eta}$ of $\mathrm{H} \times \mathrm{Z}$ defined by

$$\widetilde{\eta}(x, z^n) = \eta(x)\theta^n$$

for any $n \in \mathbf{Z}$. This satisfies $\widetilde{\eta}(x, 1) = \eta(x)$ for any $x \in \mathrm{H}$. There is furthermore a surjective "product" morphism $f\colon \mathrm{H} \times \mathrm{Z} \to \mathrm{H}'$ sending (x, z^n) to xz^n. So to construct a character of H' extending η, it suffices to find a value of θ for which $\ker(f) \subset \ker(\widetilde{\eta})$, as this will ensure that $\widetilde{\eta}$ "descends" to a character η' of H' such that $\eta'(f(x, z^n)) = \widetilde{\eta}(x, z^n)$, and in particular $\eta'(y) = \eta(y) \neq 1$.

But we can find such a θ: indeed, $\ker(f)$ is the set of (x, z^n) such that $xz^n = 1$, so it is isomorphic to $\mathrm{H} \cap \mathrm{Z}$ (by sending $x \in \mathrm{H} \cap \mathrm{Z}$ to (h, h^{-1})), in particular to a subgroup of Z. Such a subgroup is cyclic of some order d dividing q, and taking θ to be a d-th root of unity will give the desired property. - □

PROOF OF THEOREM A.7.3. We first prove that the characters form an orthonormal system in the space C(G).

Let $\chi \in \widehat{\mathrm{G}}$. Since χ has modulus 1, as we observed above, we get

$$\langle \chi, \chi \rangle = \frac{1}{|\mathrm{G}|} \sum_{x \in \mathrm{G}} |\chi(x)|^2 = 1.$$

Let now $\chi_1 \neq \chi_2$ be distinct characters of G. We have

$$\langle \chi_1, \chi_2 \rangle = \frac{1}{|\mathrm{G}|} \sum_{x \in \mathrm{G}} \chi_1(x)\overline{\chi_2(x)},$$

and we need to show that this is zero. Since $\chi_1 \neq \chi_2$, the character $\chi = \chi_1\overline{\chi_2} = \chi_1\chi_2^{-1}$ is not the trivial character. Picking an element x_0 such that $\chi_1(x_0) \neq \chi_2(x_0)$, and making the (bijective) change of variable $x = x_0 y$ (with inverse $y = x_0^{-1}x$, we get

$$\frac{1}{|\mathrm{G}|} \sum_{x \in \mathrm{G}} \chi_1(x)\overline{\chi_2(x)} = \frac{\chi_1(x_0)\overline{\chi_2(x_0)}}{|\mathrm{G}|} \sum_{y \in \mathrm{G}} \chi_1(y)\overline{\chi_2(y)}.$$

This is only possible if the sum $\langle \chi_1, \chi_2 \rangle$ is zero, hence this inner product vanishes.

In order to conclude the proof of the theorem, it only remains to prove that the set of characters is a generating set for C(G). Among various possibilities, we do this by showing that the $\langle f, \chi \rangle$ "compute" the norm square of a function $f \in \mathrm{C}(\mathrm{G})$. Namely, we compute

$$\begin{aligned} \sum_{\chi \in \widehat{\mathrm{G}}} |\langle f, \chi \rangle|^2 &= \sum_{\chi \in \widehat{\mathrm{G}}} \Big| \frac{1}{|\mathrm{G}|} f(x)\overline{\chi(x)} \Big|^2 \\ &= \frac{1}{|\mathrm{G}|^2} \sum_{x,y \in \mathrm{G}} f(x)\overline{f(y)} \sum_{\chi \in \widehat{\mathrm{G}}} \chi(x)\overline{\chi(y)}. \end{aligned}$$

We consider the inner sum over characters

$$\sum_{\chi \in \widehat{\mathrm{G}}} \chi(x)\overline{\chi(y)} = \sum_{\chi \in \widehat{\mathrm{G}}} \chi(xy^{-1}). \tag{A.6}$$

If $x \neq y$, then Lemma A.7.4 provides some character χ_0 such that $\chi_0(xy^{-1}) \neq 1$. We make the change of variable $\chi = \eta\chi_0$, and get

$$\sum_{\chi \in \widehat{\mathrm{G}}} \chi(xy^{-1}) = \chi_0(xy^{-1}) \sum_{\eta \in \widehat{\mathrm{G}}} \eta(xy^{-1}),$$

which implies that the sum vanishes in that case. Hence only the terms with $x = y$ remain, in which case the sum over χ is equal to $|\widehat{\mathrm{G}}|$, so that

$$\sum_{\chi \in \widehat{\mathrm{G}}} |\langle f, \chi \rangle|^2 = \frac{|\widehat{\mathrm{G}}|}{|\mathrm{G}|^2} \sum_{x \in \mathrm{G}} |f(x)|^2.$$

This equality, valid for any $f \in \mathrm{C}(\mathrm{G})$, proves that the space of functions orthogonal to all characters is reduced to the zero function. Hence the characters span C(G), and form an orthonormal basis (and since the dimension of C(G) is |G|, we get $|\widehat{\mathrm{G}}| = \dim \mathrm{C}(\mathrm{G}) = |\mathrm{G}|$). □

The following formulas occured in the proof, and are worth emphasizing separately.

Corollary A.7.5 (Orthogonality relations). *Let* G *be a finite abelian group.*

(1) *For any characters* χ_1 *and* χ_2*, we have*

$$\sum_{x \in \mathrm{G}} \chi_1(x)\overline{\chi_2(x)} = \begin{cases} |\mathrm{G}| & \text{if } \chi_1 = \chi_2 \\ 0 & \text{otherwise.} \end{cases} \tag{A.7}$$

(2) *For any elements x and y of* G*, we have*

$$\sum_{\chi\in\widehat{\mathrm{G}}}\chi(x)\overline{\chi(y)}=\begin{cases}|\mathrm{G}| & \text{if } x=y\\ 0 & \text{otherwise.}\end{cases} \tag{A.8}$$

Proof. The first relation re-states the fact that the characters form an orthonormal basis of $\mathrm{C}(\mathrm{G})$. The second re-states the formula for the quantity (A.6), combined with the fact that $|\mathrm{G}|=|\widehat{\mathrm{G}}|$. □

The theorem means that for any $f\colon \mathrm{G}\to\mathbf{C}$, we have an equality

$$f=\sum_{\chi\in\widehat{\mathrm{G}}}\langle f,\chi\rangle\chi$$

between functions. The collection of inner products $\langle f,\chi\rangle$, which therefore characterizes f, can be seen as a function on the dual group. It is often convenient to normalize these differently, and to define the *(unitary) Fourier transform* of f, which is the function $\widehat{f}\colon\widehat{\mathrm{G}}\to\mathbf{C}$ such that

$$\widehat{f}(\chi)=\frac{1}{\sqrt{|\mathrm{G}|}}\sum_{x\in\mathrm{G}}f(x)\overline{\chi(x)}=\sqrt{|\mathrm{G}|}\ \langle f,\chi\rangle.$$

The point of the normalization is the following result.

Theorem A.7.6. *Let* G *be a finite abelian group. The Fourier transform* $f\mapsto\widehat{f}$ *is an isometry from the space* $\mathrm{C}(\mathrm{G})$ *to the space* $\mathrm{C}(\widehat{\mathrm{G}})$.

Proof. Let $f\in\mathrm{C}(\mathrm{G})$. According to Theorem A.7.3, we have

$$\|f\|^2=\sum_{\chi\in\widehat{\mathrm{G}}}|\langle f,\chi\rangle|^2.$$

The left hand side is

$$\frac{1}{|\mathrm{G}|}\sum_{\chi\in\widehat{\mathrm{G}}}|\mathrm{G}|\ |\langle f,\chi\rangle|^2=\frac{1}{|\mathrm{G}|}\sum_{\chi\in\widehat{\mathrm{G}}}|\widehat{f}(\chi)|^2=\|\widehat{f}\|^2,$$

hence the result. □

Concretely, as displayed in the proof, this means that for any function $f\in\mathrm{C}(\mathrm{G})$, we have the *discrete Plancherel formula*

$$\sum_{x\in\mathrm{G}}|f(x)|^2=\sum_{\chi\in\widehat{\mathrm{G}}}|\widehat{f}(\chi)|^2. \tag{A.9}$$

Similarly, by construction, the Fourier transform satisfies the *Fourier inversion formula*:

$$f(x)=\sum_{\chi\in\widehat{\mathrm{G}}}\langle f,\chi\rangle\chi(x)=\frac{1}{\sqrt{|\mathrm{G}|}}\sum_{\chi\in\widehat{\mathrm{G}}}\widehat{f}(\chi)\chi(x)$$

for $f\in\mathrm{C}(\mathrm{G})$ and for all $x\in\mathrm{G}$.

This result gives a natural motivation for another important operation on $\mathrm{C}(\mathrm{G})$. Indeed, since the Fourier transform is bijective, given two functions f, g on G, there exists a unique function with Fourier transform equal to the product $\widehat{f}\cdot\widehat{g}$ of the Fourier transforms. What is it?

PROPOSITION A.7.7. *Let* G *be a finite abelian group. For any functions* f, $g\colon \mathrm{G} \to \mathbf{C}$, *we have* $\widehat{f * g} = \widehat{h}$ *where the function* h *is defined by*

$$h(x) = \frac{1}{|\mathrm{G}|^{1/2}} \sum_{y \in \mathrm{G}} f(y) g(y^{-1}x)$$

for all $x \in \mathrm{G}$.

PROOF. By Fourier inversion, the unique function with Fourier transform $\widehat{f}\,\widehat{g}$ sends $x \in \mathrm{G}$ to

$$\frac{1}{\sqrt{|\mathrm{G}|}} \sum_{\chi \in \widehat{\mathrm{G}}} \widehat{f}(\chi)\widehat{g}(\chi)\chi(x).$$

We replace in this expression the Fourier transforms by their values, and obtain

$$\frac{1}{|\mathrm{G}|^{3/2}} \sum_{\chi \in \widehat{\mathrm{G}}} \Big(\sum_{y \in \mathrm{G}} f(y)\overline{\chi(y)} \Big) \Big(\sum_{z \in \mathrm{G}} g(z)\overline{\chi(z)} \Big) \chi(x).$$

Exchanging the order of the sums, this is equal to

$$\frac{1}{|\mathrm{G}|^{1/2}} \sum_{y \in \mathrm{G}} \sum_{z \in \mathrm{G}} f(y)g(z) \; \frac{1}{|\mathrm{G}|} \sum_{\chi \in \widehat{\mathrm{G}}} \overline{\chi(yz)}\chi(x) = \frac{1}{|\mathrm{G}|^{1/2}} \sum_{\substack{y \in \mathrm{G},\, z \in \mathrm{G} \\ yz = x}} f(y)g(z),$$

which is the formula which we claimed. □

The function h is one possible definition of the *convolution* of g and g. There are however other normalizations which may be used to define precisely the convolution product. We will use the following:

DEFINITION A.7.8. Let f and g be functions in $\mathrm{C}(\mathrm{G})$. The *convolution* $f * g$ of f and g is defined by

$$(f * g)(x) = \frac{1}{|\mathrm{G}|} \sum_{y \in \mathrm{G}} f(y)g(y^{-1}x),$$

and the *normalized convolution* is defined by

$$(f \star g)(x) = \frac{1}{|\mathrm{G}|^{1/2}} \sum_{y \in \mathrm{G}} f(y)g(y^{-1}x).$$

REMARK A.7.9. (1) According to the proposition, the normalized convolution satisfies

$$\widehat{f} \cdot \widehat{g} = \widehat{f \star g}$$

for any two functions f and g, and therefore

$$\widehat{f * g} = \frac{1}{|\mathrm{G}|^{1/2}} \widehat{f} \cdot \widehat{g}.$$

On the other hand, by comparing with the relation between the Fourier transform and the expansion in the basis of characters of $\mathrm{C}(\mathrm{G})$, we otain the relation

$$\langle f * g, \chi \rangle = \langle f, \chi \rangle \; \langle g, \chi \rangle$$

for any character χ of G.

(2) A third common normalization of the convolution is given by the function

$$x \mapsto \sum_{y \in \mathrm{G}} f(y)g(y^{-1}x). \tag{A.10}$$

Its behavior with respect to Fourier expansions is not as nice as the previous ones, but it interacts well with representation functions for product sets: given sets A and B with characteristic functions φ_{A} and φ_{B}, we have

$$r_{\mathrm{A,B}}(x) = \sum_{y\in\mathrm{G}} \varphi_{\mathrm{A}}(y)\varphi_{\mathrm{B}}(y^{-1}x),$$

which is the formula (A.10) applied to φ_{A} and φ_{B}.

All three choices can be justified, and are natural from a certain point of view. For instance:

(1) The value $(f * g)(x)$ of the convolution at x can be interpreted as the inner product in $\mathrm{C}(\mathrm{G})$ of the functions f and $y \mapsto \overline{g(y^{-1}x)}$, so we can for instance deduce immediately that its value at x is at most $\|f\|\,\|g\|$ by the Cauchy-Schwarz inequality.

(2) The normalized convolution $f \star g$ has the feature that the convolution of two "random" functions with roughly constant values on average will have the same property (this is because the sum

$$\sum_{y\in\mathrm{G}} f(y)g(y^{-1}x)$$

for such functions will be a sum of $|\mathrm{G}|$ complex numbers with oscillating phases and roughly constant modulus, which on probabilistic grounds is of size roughly $|\mathrm{G}|^{1/2}$ in general).

(3) From a certain point of view (motivated by aspects of arithmetic geometry, as in work of Katz and Forey, Fresán and Kowalski, explained in [32]), it would pay to introduce somewhat systematically a family of operations $(f, g) \mapsto f *_{[k]} g$, for $k \geqslant 0$ in $\mathbf{R}$, defined by

$$(f *_{[k]} g)(x) = \frac{1}{|\mathrm{G}|^k} \sum_{y\in\mathrm{G}} f(y)g(y^{-1}x),$$

so that for instance $f \star g = f *_{[1/2]} g$.

From Proposition A.7.7, or by direct computations, it is straightforward that, for instance, the convolution product is commutative and associative, in the sense that

$$f * (g * h) = (f * g) * h, \qquad f * g = g * f$$

for all f, g and h in $\mathrm{C}(\mathrm{G})$. The same applies to the normalized convolution.

Moreover, the function δ_1 equal to $|\mathrm{G}|$ on the neutral element and to 0 elsewhere is a unit, in the sense that

$$\delta_1 * f = f * \delta_1 = f$$

for all $f \in \mathrm{C}(\mathrm{G})$ (this reflects the fact that the function always equal to 1 is the unit for multiplication of functions).

Example A.7.10. We translate the previous results and notation in the case of a cyclic group $\mathrm{G} = \mathbf{Z}/q\mathbf{Z}$ (with additive notation). In Example A.7.2, we saw that the character group can also be identified with $\mathbf{Z}/q\mathbf{Z}$, with the element $h \in \mathbf{Z}/q\mathbf{Z}$ corresponding to the character

$$\chi_h(x) = e\Big(\frac{hx}{q}\Big).$$

Thus, the Fourier transform of $f\colon \mathbf{Z}/q\mathbf{Z} \to \mathbf{C}$ can be identified with the function $\widehat{f}\colon \mathbf{Z}/q\mathbf{Z} \to \mathbf{C}$ given by

$$\widehat{f}(h) = \sqrt{q}\langle f, \chi_n\rangle = \frac{1}{\sqrt{q}} \sum_{x\in\mathbf{Z}/q\mathbf{Z}} f(x)e\Big(-\frac{hx}{q}\Big).$$

The function f is then recovered from $\widehat{f}$ by

$$f(x) = \frac{1}{\sqrt{q}} \sum_{h\in\mathbf{Z}/q\mathbf{Z}} \widehat{f}(h)e\Big(\frac{hx}{q}\Big),$$

and the Plancherel formula is

$$\sum_{x\in\mathbf{Z}/q\mathbf{Z}} |f(x)|^2 = \sum_{h\in\mathbf{Z}/q\mathbf{Z}} |\widehat{f}(h)|^2.$$

The convolution is defined by

$$f * g(x) = \frac{1}{q} \sum_{y\in\mathbf{Z}/q\mathbf{Z}} f(y)g(x-y).$$

Remark A.7.11. The theory of characters of finite abelian group and the corresponding applications to Fourier transform and convolution has a generalization to all *locally compact* abelian groups (i.e., abelian groups which are also topological spaces in a compatible manner, and where the topology is locally compact), provided only continuous characters are considered. This is the topic of *Pontryagin duality*; see, for instance, [8, Ch. 2].

Two important special cases appear in classical Fourier analysis: the case of $\mathrm{G} = \mathbf{R}/\mathbf{Z}$ (or equivalently of the group of complex numbers of modulus 1, which is isomorphic by $x \mapsto e(x)$), and that of $\mathrm{G} = \mathbf{R}$.

In both cases, there are now restrictions (such as integrability, etc) to impose on functions so that the inner product

$$\langle f, g\rangle = \int_{\mathrm{G}} f(x)\overline{g(x)}dx$$

of functions $f\colon \mathrm{G} \to \mathbf{C}$ is defined, where the integral is taken in the Lebesgue sense usually (for $\mathrm{G} = \mathbf{R}/\mathbf{Z}$, it can be viewed as an integral from 0 to 1).

In the first case, one can check that the characters are the complex exponentials $e_n\colon x \mapsto e(nx)$, parameterized by $n \in \mathbf{Z}$. The expansion

$$f(x) = \sum_{n\in\mathbf{Z}} \langle f, e_n\rangle e_n(x),$$

when it holds, is the expansion of f in *Fourier series*, with coefficients

$$\langle f, e_n\rangle = \int_0^1 f(x)e(-nx)dx.$$

The usual Parseval formula is the identity

$$\int_0^1 |f(x)|^2 dx = \sum_{n\in\mathbf{Z}} |\langle f, e_n\rangle|^2,$$

valid when the integral makes sense. The convolution of two functions is defined by

$$f * g(x) = \int_0^1 f(t)g(x-t)dt,$$

again when the integral exists.

In the second case, the group of characters can be identified again with $\mathbf{R}$, with $y \in \mathbf{R}$ corresponding to the character $x \mapsto e(xy)$, and the Fourier transform of $f : \mathrm{G} \to \mathbf{C}$ is the function on $\mathbf{R}$ such that

$$\widehat{f}(y) = \int_{\mathbf{R}} f(x)e(-xy)dx,$$

at least when these integrals make sense. Other normalizations are sometimes used, but the one above satisfies the isometry property

$$\int_{\mathbf{R}} |f(x)|^2 dx = \int_{\mathbf{R}} |\widehat{f}(y)|^2 dy,$$

without extra constant factor.

The Fourier inversion formula takes the form

$$f(x) = \int_{\mathbf{R}} \widehat{f}(y)e(xy)dy,$$

but there are various conditions required to ensure its validity.

A.8. Harmonic analysis on finite groups

It is possible to extend much of the results of the previous section to all finite groups, not necessarily abelian, but this requires new ideas. Indeed, although characters can be defined for arbitrary groups, there exist many interesting finite groups which have no non-trivial character (for instance, any finite simple group which is not abelian); in any case, if G is not commutative, there are "not enough" characters, in the sense that Lemma A.7.4 fails to hold in that case.

Certain notions are however easily adapted. The most straightforward is in fact the convolution operation, whose definition makes sense for any finite group.

Definition A.8.1. Let G be a finite group and C(G) the vector space of complex-valued functions on G. The *convolution* of f and g is defined by

$$(f * g)(x) = \frac{1}{|\mathrm{G}|} \sum_{y \in \mathrm{G}} f(y)g(y^{-1}x) = \sum_{yz=x} f(y)g(z),$$

for all $x \in \mathrm{G}$.

Even without the interpretation of Proposition A.7.7, one can check directly that the convolution product is associative, so that $f * (g * h) = (f * g) * h$ for all f, g and h in C(G), and that the function δ_1 equal to $|\mathrm{G}|$ on the neutral element and to 0 elsewhere is a unit. However, the convolution is commutative only if the group itself is commutative.

We will present briefly (some of) the basic facts of Fourier analysis on a finite group; our approach is a bit unusual, but is chosen to be accessible and suggestive for readers with a less algebraic background.

The point of view we take is that of trying to decompose the finite-dimensional Hilbert space C(G) in suitable orthogonal subspaces, one of which should be the one-dimensional space of the constant functions.

Definition A.8.2. Let E be a non-zero finite-dimensional Hilbert space. The *model space* associated to E is the finite-dimensional Hilbert spaces H such that H equal to the space of linear transformations on E, with the inner product given by

$$\langle u, v \rangle = \dim(\mathrm{E})\ \mathrm{Tr}(uv^*),$$

for u and v linear maps from E to itself, where v^* is the Hilbert space adjoint of v.

EXAMPLE A.8.3. Concretely, one should think of $\mathrm{E} = \mathbf{C}^d$ for some integer $d \geqslant 1$, and then H can be identified with the space of complex matrices of type $d \times d$, equipped with the hermitian inner-product such that

$$\|\mathrm{A}\|_{\mathrm{H}}^2 = d \sum_{i,j} |a_{i,j}|^2$$

for any matrix $\mathrm{A} = (a_{i,j})_{1 \leqslant i,j \leqslant d}$ in H. Indeed, the (i,j)-coefficient of AA^* is

$$\sum_{k=1}^{d} a_{i,k} \overline{a_{j,k}},$$

hence the trace of AA^* is equal to

$$\sum_{i=1}^{d} \sum_{k=1}^{d} a_{i,k} \overline{a_{i,k}} = \sum_{i,k} |a_{i,k}|^2.$$

This example, after choosing an orthonormal basis of E, shows that every model space can be thought of having this form for some d, which shows in particular that this is indeed a Hilbert space (i.e., the hermitian form is positive definite).

The simplest case is $d = 1$, in which case the model space is also 1-dimensional, and can be identified with $\mathbf{C}$, with the scalar product $(w, z) \mapsto w\bar{z}$, the norm being the modulus of complex numbers.

REMARK A.8.4. Given a finite-dimensional Hilbert space E, the norm $u \mapsto \mathrm{Tr}(uu^*)$ on the space of linear maps on E is also called the *Hilbert–Schmidt norm*.

The existence of an orthonormal basis for a finite-dimensional Hilbert space E can be interpreted as stating that E is isometric to an orthogonal direct sum of model spaces of dimension 1. However, if the space E has additional structure, it may be more convenient to find a "rougher" decomposition which respects this structure. In the case of interest in this section, the space is $\mathrm{C}(\mathrm{G})$ for some finite group G, with the Hilbert space structure defined by

$$\langle f, g \rangle = \frac{1}{|\mathrm{G}|} \sum_{x \in \mathrm{G}} f(x) \overline{g(x)},$$

and the additional structure is the convolution product $(f, g) \mapsto f * g$. Having a compatible decomposition of $\mathrm{C}(\mathrm{G})$ means finding an orthogonal decomposition

$$\mathrm{C}(\mathrm{G}) = \bigoplus_{i \in \mathrm{I}} \mathrm{H}_i,$$

where each space H_i, viewed as a subspace of $\mathrm{C}(\mathrm{G})$, is *stable* by the convolution product, in the sense that if we decompose

$$f = \sum_{i \in \mathrm{I}} f_i, \qquad g = \sum_{i \in \mathrm{I}} g_i,$$

with f_i, $g_i \in \mathrm{H}_i$, then we should have

$$f * g = \sum_{i \in \mathrm{I}} f_i * g_i,$$

with $f_i * g_i \in \mathrm{H}_i$ also. In fact, if H_i is a model space associated to E_i, then f_i and g_i can *also* be thought of as linear maps on E_i, and we may want to also have $f_i * g_i$ correspond to the composition $f_i \circ g_i$ of linear maps (concretely, if we view f_i and g_i as matrices, then $f_i * g_i$ should correspond to their product).

It turns out that such a decomposition does exist.

THEOREM A.8.5 (Fourier analysis on G). *Let* G *be a finite group. There exists a finite set* $\widehat{\mathrm{G}}$ *and a family* $(\mathrm{E}_\varrho)_{\varrho\in\widehat{\mathrm{G}}}$ *of finite-dimensional Hilbert spaces, with associated model spaces* H_ϱ*, such that*

(1) *The cardinality of* $\widehat{\mathrm{G}}$ *is the number of conjugacy classes in* H*.*

(2) *We have an isometric isomorphism*

$$\mathrm{C}(\mathrm{G}) = \bigoplus_{\varrho\in\widehat{\mathrm{G}}} \mathrm{H}_\varrho,$$

where the direct sum is orthogonal, allowing us to view each model space H_ϱ *as a subspace of* $\mathrm{C}(\mathrm{G})$*.*

(3) *For any* f *and* g *in* $\mathrm{C}(\mathrm{G})$*, represented as the sums*

$$f = \sum_{\varrho\in\widehat{\mathrm{G}}} f_\varrho, \quad g = \sum_{\varrho\in\widehat{\mathrm{G}}} g_\varrho,$$

with f_ϱ *and* g_ϱ *in* H_ϱ*, we have*

$$f * g = \sum_{\varrho\in\widehat{\mathrm{G}}} f_\varrho * g_\varrho,$$

with the function $f_\varrho * g_\varrho$ *also identified to the linear map* $f_\varrho \circ g_\varrho$ *in the model space* H_ϱ*.*

(4) *There exists* $\varrho_0 \in \widehat{\mathrm{G}}$ *such that* E_{ϱ_0} *is one-dimensional, and such that, for any* $f \in \mathrm{C}(\mathrm{G})$*, we have*

$$f_{\varrho_0} = \frac{1}{|\mathrm{G}|}\sum_{x\in\mathrm{G}} f(x) \in \mathrm{H}_{\varrho_0} = \mathbf{C}.$$

(5) *For any* $\varrho \in \widehat{\mathrm{G}}$*,* $f \in \mathrm{C}(\mathrm{G})$ *and* $x \in \mathrm{G}$*, we have*

$$|f_\varrho(x)| \leqslant \|f_\varrho\|_\varrho,$$

where on the left we view f_ϱ *as an element of* $\mathrm{C}(\mathrm{G})$ *and on the right as a linear transformation, and the norm is the model norm.*

PROOF. Since this is an unusual formulation of the theory, we explain how to deduce this from basic representation theory, taking [58, Ch. 4] as reference (among many suitable texts). We assume therefore some familiarity with this language.

We denote by $\widehat{\mathrm{G}}$ a set of representative of isomorphism classes of irreducible unitary representations of G, and by ϱ_0 the trivial one-dimensional representation. Each ϱ is a group morphism from G to the unitary group of some finite-dimensional Hilbert space E_ϱ, which we use to define the model space H_ϱ. The subspace of $\mathrm{C}(\mathrm{G})$ corresponding to H_ϱ is the space of matrix coefficients of ϱ, i.e., the space of functions of the form

$$f(x) = \langle v, \varrho(g)w\rangle, \qquad x \in \mathrm{G}$$

for some v and w in E_ϱ, where the inner product is the one on that space. The isomorphism with H_ϱ is obtained by mapping a linear map $u\colon \mathrm{E}_\pi \to \mathrm{E}_\pi$ to the function $f_u \in \mathrm{H}_\varrho$ defined by

$$f_u(x) = \langle \varrho(x), u\rangle = \dim(\mathrm{E}_\pi)\,\mathrm{Tr}(\varrho(x)u^*)$$

for $x \in \mathrm{G}$. □

REMARK A.8.6. For $\varrho \in \widehat{\mathrm{G}}$, we denote by $\deg(\varrho)$ the dimension of the Hilbert space E_π, and call it the *degree* of ϱ.

EXAMPLE A.8.7. Let $f \in \mathrm{C}(\mathrm{G})$. Concretely, after isolating the term corresponding to ϱ_0, the above means that we have a decomposition

$$f = \frac{1}{|\mathrm{G}|}\sum_{x\in\mathrm{G}} f(x) + \sum_{\substack{\varrho\in\widehat{\mathrm{G}}\\ \varrho\neq\varrho_0}} f_\varrho \tag{A.11}$$

where f_ϱ is in the space H_ϱ, and with the following properties:

(1) The *Parseval formula*:

$$\|f\|^2 = \frac{1}{|\mathrm{G}|}\sum_{x\in\mathrm{G}} |f(x)|^2 = \sum_{\varrho\in\widehat{\mathrm{G}}} \|f_\varrho\|^2 = \sum_{\varrho\in\widehat{\mathrm{G}}} \deg(\varrho)\,\mathrm{Tr}(f_\varrho f_\varrho^*). \tag{A.12}$$

(2) The pointwise estimates:

$$|f_\varrho(x)| \leqslant \|f_\varrho\|_\varrho.$$

for any ϱ and any $x \in \mathrm{G}$.

(3) The convolution property: if $f = f_1 * f_2$ for some functions $f_i \in \mathrm{C}(\mathrm{G})$, then

$$f_\varrho = f_{1,\varrho} \circ f_{2,\varrho}$$

in the sense that the ϱ-component of f, as a linear map, is the composition of the ϱ-components of f_1 and f_2.

EXAMPLE A.8.8. Suppose that G is commutative. The Fourier decomposition of Theorem A.8.5 corresponds in the following way to the discussion of Section A.7: the set $\widehat{\mathrm{G}}$ can be identified with the set of characters, with ϱ_0 corresponding to the trivial character, and for each χ, the corresponding model space has dimension 1; in fact, the subspace of $\mathrm{C}(\mathrm{G})$ which is associated to H_χ can be identified with the one-dimensional space spanned by χ itself (since $\chi \in \mathrm{C}(\mathrm{G})$), and for $f\colon \mathrm{G} \to \mathbf{C}$, we have

$$f_\chi(x) = \langle f, \chi\rangle\ \chi(x),$$

so that the decomposition

$$f = \sum_\chi f_\chi$$

is identical with the expansion of f in the orthonormal basis of characters.

This means, for instance, that the convolution property reflects the identity

$$\langle f * g, \chi\rangle = \langle f, \chi\rangle\ \langle g, \chi\rangle,$$

which can be checked straightforwardly:

$$\begin{aligned}\langle f * g, \chi\rangle &= \frac{1}{|\mathrm{G}|^2}\sum_{x\in\mathrm{G}}\Big(\sum_{ab=x} f(a)g(b)\Big)\overline{\chi(x)}\\ &= \frac{1}{|\mathrm{G}|^2}\sum_{a,b\in\mathrm{G}} f(a)g(b)\overline{\chi(ab)}\\ &= \langle f, \chi\rangle\ \langle g, \chi\rangle\end{aligned}$$

(where we see that the crucial fact is that $\chi(ab) = \chi(a)\chi(b)$).

The pointwise estimate is also elementary here: for any $x \in \mathrm{G}$, we have in fact $|f_\chi(x)| = |\langle f, \chi\rangle| = \|f_\chi\|_\chi$.

It may be surprising to see how far one can go in using Fourier analysis in concrete problems with nothing more than the previous properties. We add to these some useful estimates which only concern the model spaces. We recall that given Hilbert (or Banach spaces) E and F and a linear map $u\colon \mathrm{E} \to \mathrm{F}$, the operator norm of u is defined by

$$\|u\| = \sup_{\substack{x\in \mathrm{E}\\ x\neq 0}} \frac{\|u(x)\|_{\mathrm{F}}}{\|x\|_{\mathrm{E}}}.$$

By definition, this means that we have

$$\|u(x)\| \leqslant \|u\|\,\|x\|$$

for any $x \in \mathrm{E}$ and $u\colon \mathrm{E} \to \mathrm{F}$.

Proposition A.8.9. *Let* H *be a model space associated to a finite-dimensional Hilbert space* E*. Denote by* $u \mapsto \|u\|_{\mathrm{H}}$ *its model norm, and by* $u \mapsto \|u\|$ *the operator norm on linear maps on* E*.*

(1) *For any* $u \in \mathrm{H}$*, we have*

$$\|u\| \leqslant \dim(\mathrm{E})^{-1/2}\|u\|_{\mathrm{H}}.$$

(2) *For any* $u \in \mathrm{H}$*, we have* $\|u\|_{\mathrm{H}} = \|u^*\|_{\mathrm{H}}$.

(3) *For any* u_1 *and* u_2 *in* H*, we have*

$$\|u_1 \circ u_2\|_{\mathrm{H}} \leqslant \|u_1\|\,\|u_2\|_{\mathrm{H}},$$

and

$$\|u_1 \circ u_2\|_{\mathrm{H}} \leqslant \|u_1\|_{\mathrm{H}}\,\|u_2\|.$$

Proof. (1) Since $\|u\|_{\mathrm{H}}^2 = \dim(\mathrm{E})^{1/2}\,\mathrm{Tr}(uu^*)$, it is enough to prove that the inequality $\|u\| \leqslant \mathrm{Tr}(uu^*)^{1/2}$ holds. But if we pick an orthonormal basis $(e_i)_{i\in \mathrm{I}}$ of E, we obtain

$$\mathrm{Tr}(uu^*) = \mathrm{Tr}(u^*u) = \sum_{i\in \mathrm{I}} \langle u^*u(e_i), e_i\rangle = \sum_{i\in \mathrm{I}} \|u(e_i)\|^2. \tag{A.13}$$

Now let $x \in \mathrm{E} - \{0\}$. We can find an orthonormal basis which contains the vector $x/\|x\|$, and therefore

$$\frac{\|u(x)\|^2}{\|x\|^2} \leqslant \sum_{i\in \mathrm{I}} \|u(e_i)\|^2 \leqslant \mathrm{Tr}(uu^*).$$

Since x is arbitrary, we obtain the bound $\|u\|_{\mathrm{H}} \leqslant \mathrm{Tr}(uu^*)^{1/2}$ by taking the square-root and the supremum over $x \neq 0$.

(2) We get $\|u^*\|_{\mathrm{H}} = \|u\|_{\mathrm{H}}$ from the equalities $\mathrm{Tr}(uu^*) = \mathrm{Tr}(u^*u)$ and $(u^*)^* = u$.

(3) From the definition again, it is enough to prove the bounds

$$\mathrm{Tr}((u_1 \circ u_2)(u_1 \circ u_2)^*) \leqslant \|u_1\|^2\,\mathrm{Tr}(u_2u_2^*),$$
$$\mathrm{Tr}((u_1 \circ u_2)(u_1 \circ u_2)^*) \leqslant \|u_2\|^2\,\mathrm{Tr}(u_1u_1^*).$$

For the first one, we apply (A.13) to $u = u_1 \circ u_2$ and any orthonormal basis $(e_i)_{i\in \mathrm{I}}$, and we get

$$\mathrm{Tr}((u_1 \circ u_2)(u_1 \circ u_2)^*) = \sum_{i\in \mathrm{I}} \|u_1(u_2(e_1))\|^2 \leqslant \|u_1\|^2 \sum_{i\in \mathrm{I}} \|u_2(e_1)\|^2 = \|u_1\|^2\,\mathrm{Tr}(u_2u_2^*),$$

by the defining property of the operator norm and (A.13) again.

For the second, we combine (2) and the first inequality applied to u_2^* and u_1^*, together with the fact that $\|u_2^*\| = \|u_2\|$, to obtain

$$\|u_1 \circ u_2\|_{\mathrm{H}} = \|(u_2^* \circ u_1^*)^*\|_{\mathrm{H}} = \|u_2^* \circ u_1^*\| \leqslant \|u_2^*\| \, \|u_1^*\|_{\mathrm{H}} = \|u_1\|_{\mathrm{H}} \, \|u_2\|.$$

□

Bibliography

[1] L. Babai and V.T. Sós: *Sidon sets in groups and induced subgraphs of Cayley graphs*, European J. Combin. 6 (1985), 101–114.

[2] A. Balog and E. Szemerédi: *A statistical theorem of set addition*, Combinatorica 14 (1994), 263–268.

[3] J. Balogh, Z. Füredi and S. Roy: *An upper bound on the size of Sidon sets*, Amer. Math. Monthly 130 (2023), 437–445.

[4] J. Beck: *On the lattice property of the plane and some problems of Dirac, Motzkin and Erdős*, Combinatorica 3 (1983), 281–297.

[5] F.A. Behrend: *On sets of integers which contain no three terms in arithmetic progression*, Proc. Nat. Acad. Sci. 32 (1946), 331–332.

[6] T. Bloom and O. Sisask: *Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions*, preprint (2020), arXiv:2007.03528.

[7] R.C. Bose: *An affine analogue of Singer's theorem*, J. Indian Math. Soc. (N.S.) 6 (1942), 1–15.

[8] N. Bourbaki: *Éléments de mathématique: Théories spectrales*, chapitres I et II, Springer, 2019.

[9] J. Bourgain and A. Gamburd: *Uniform expansion bounds for Cayley graphs of* $\mathrm{SL}_2(\mathbf{F}_p)$, Annals of Math. 167 (2008), 625–642.

[10] J. Bourgain, A.A. Glibichuk and S. Konyagin: *Estimates for the number of sums and products and for exponential sums in fields of prime order*, J. London Math. Soc. 73 (2006), 380–398.

[11] J. Bourgain, N.H. Katz and T. Tao: *A sum-product estimate in finite fields, and applications*, GAFA 14 (2004), 27–57.

[12] E. Breuillard and H. Oh (eds): *Thin groups and super-strong approximation*, MSRI Publications Vol. 61, Cambridge Univ. Press, 2014.

[13] E. Breuillard: *A brief introduction to approximate groups*, in [12], 23–50.

[14] T.D. Browning, L. Matthiesen and A.N. Skorobogatov: *Rational points on pencils of conics and quadrics with many degenerate fibers*, Annals of Math. 180 (2014), 381–402.

[15] W. Burnside: *Theory of groups of finite order*, 2nd edition, Cambridge Univ. Press, 1911.

[16] A.L. Cauchy: *Recherches sur les nombres*, J. École Polytechnique 9 (1813), 99–116; http://gallica.bnf.fr/ark:/12148/bpt6k90193x/f45

[17] J. Cilleruelo: *Combinatorial problems in finite fields and Sidon sets*, Combinatorica 32 (2012), 497–511.

[18] J. Cilleruelo: *Infinite Sidon sequences*, Adv. Math. 255 (2014), 474–486.

[19] H. Davenport: *On the addition of residue classes*, J. London Math. Soc. 10 (1935), 30–32.

[20] P. Deligne: *La conjecture de Weil, II*, Publ. Math. IHÉS 52 (1980), 137–252.

[21] P.G.L. Dirichlet: *Verallgemeinerung eines Satzes aus der Lehre von den Kettenbrüchen nebst einigen Anwendungen auf die Theorie der Zahlen*, 1842; in G. Lejeune Dirichlet's Werke, Vol. 1, G. Reimer, Berlin, 1889; https://archive.org/details/glejeunedirichl01dirigoog/page/635/mode/1up?view=theater.

[22] W. Duke, S.R. Garcia and B. Lutz: *The graphic nature of Gaussian periods*, Proc. Amer. Math. Soc. 143 (2015), 1849–1863.

[23] Z. Dvir: *On the size of Kakeya sets in finite fields*, J. Amer. Math. Soc. 22 (2009), 1093–1097.

[24] S. Eberhard and F. Manners: *The apparent structure of dense Sidon sets*, Electron. J. Combin. 30 (2023), 1–33.

[25] G.A. Edgar and C. Miller: *Borel subrings of the reals*, Proceedings AMS 131 (2002), 1121–1129.

[26] M. Einsiedler and T. Ward: *Ergodic theory: with a view towards number theory*, Grad. Texts in Math. 259, Springer, 2011.

[27] G. Elekes: *On the number of sums and products*, Acta Arithmetica 81 (1997), 365–367.

[28] P. Erdős and E. Szemerédi: *On sums and products of integers*, in "Studies in pure mathematics to the memory of Paul Turán", edited by P. Erdős, L. Alpár, G. Halász and A. Sárközy, Studies in Pure Mathematics, Birkhäuser (1983).
[29] P. Erdős and P. Turán: *On some sequences of integers*, J. London Math. Soc. 11 (1936), 261–264.
[30] P. Erdős and P. Turán: *On a problem of Sidon in additive number theory, and on some related problems*, J. London Math. Soc. 16 (1941), 212–215.
[31] K. Ford: *The distribution of integers with a divisor in a given interval*, Annals of Math. 168 (2008), 367–433.
[32] A. Forey, J. Fresán and E. Kowalski: *Arithmetic Fourier transforms over finite fields*, preprint, 2021; available at https://www.math.ethz.ch/~kowalski/convolution-equidistribution.pdf.
[33] A. Forey, J. Fresán and E. Kowalski: *Sidon sets in algebraic geometry*, IMRN (2023), doi.org/10.1093/imrn/rnad1692023.
[34] É. Fouvry, E. Kowalski and Ph. Michel: *An inverse theorem for Gowers norms of trace functions over* $\mathbf{F}_p$, Math. Proc. Cambridge Phil. Soc. 155 (2013), 277–295.
[35] G.A. Freiman: *Foundations of a structural theory of set addition*, translated from the Russian, Translations of Mathematical Monographs 37, A.M.S, 1973.
[36] J. Friedlander and H. Iwaniec: *Opera de cribro*, Colloquium Publ. 57, A.M.S, 2010.
[37] H. Furstenberg: *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. 31 (1977), 204–256.
[38] M.J. Ganley: *Direct product difference sets*, Journal Combinat. Theory A 23 (1977), 321–332.
[39] M.Z. Garaev: *The sum-product estimate for large subsets of prime fields*, Proc. A.M.S 136 (2008), 2735–2739.
[40] S.R. Garcia, T. Hyde and B. Lutz: *Gauss's hidden menagerie: from cyclotomy to supercharacters*, Notices AMS 62 (2015), 878–888.
[41] T. Gowers: *A new proof of Szemerédi's Theorem*, Geom. Funct. Analysis 11 (2001), 465–588.
[42] T. Gowers: *Quasirandom groups*, Comb. Probab. Comp. 17 (2008), 363–387.
[43] T. Gowers: *A uniform set with fewer than expected arithmetic progressions of length* 4, Acta Math. Hungar. 161 (2020) 756–767.
[44] B.J. Green: *Roth's Theorem in the primes*, Annals of Math. 161 (2005), 1609–1636.
[45] B.J. Green and T. Tao: *The primes contain arbitrarily long arithmetic progressions*, Annals of Math. 167 (2008), 481–547.
[46] B.J. Green, T. Tao and T. Ziegler: *An inverse theorem for the Gowers* $\mathrm{U}^{s+1}[\mathrm{N}]$ *norms*, Annals of Math. 176 (2012), 1231–1372.
[47] L. Guth: *Polynomial methods in combinatorics*, University Lectures Series 64, A.M.S, 2016.
[48] G.H. Hardy and E.M. Wright: *An introduction to the theory of numbers*, 5th edition, Oxford, 1979.
[49] H. Helfgott: *Growth and generation in* $\mathrm{SL}_2(\mathbf{Z}/p\mathbf{Z})$, Annals of Math. 167 (2008), 601–623.
[50] H. Helfgott: *Growth in groups: ideas and perspectives*, Bull. Amer. Math. Soc. 52 (2015), 357–413.
[51] S. Hoory, N. Linial and A. Wigderson: *Expander graphs and their applications*, Bull. Amer. Math. Soc. 43 (2006), 439–561.
[52] D. R. Hughes: *Planar division neo-rings*, Transactions of the AMS 80 (1955), 502–527.
[53] H. Iwaniec and E. Kowalski: *Analytic Number Theory*, Colloquium Publ. 53, A.M.S, 2004.
[54] Z. Kelley and R. Meka: *Strong bounds for* 3*-progressions*, preprint (2023), arXiv:2302.05537.
[55] A. Khintchine: *Three pearls of number theory*, Dover Publications (1998); available at https://archive.org/details/khinchin-three-pearls-of-number-theory.
[56] S. Konyagin: *A sum-product estimate in fields of prime order*, preprint (2003), arXiv:0304217.
[57] E. Kowalski: *Explicit growth and expansion for* $\mathrm{SL}_2(\mathbf{F}_p)$, IMRN 2013, 5645–5708.
[58] E. Kowalski: *An introduction to the representation theory of groups*, Graduate Studies in Math. 155, A.M.S, 2014.
[59] E. Kowalski: *An introduction to expander graphs*, Cours Spécialisés 26, Soc. Math. France, 2019.
[60] E. Kowalski: *An introduction to probabilistic number theory*, Cambridge Studies in Advanced Math. 192, Cambridge University Press, 2021.
[61] P. Kurlberg: *Bounds on exponential sums over small multiplicative subgroups*, in "Additive combinatorics", CRM Proc. Lecture Notes, 43, A.M.S, 2007.
[62] I. Łaba: *From harmonic analysis to arithmetic combinatorics*, Bulletin A.M.S 45 (2008), 77–115.

[63] B. Murphy: *Upper and lower bounds for rich lines in grids*, Amer. J. Math. 143 (2021), 577–611.
[64] G. Petridis: *New proofs of Plünnecke-type estimates for product sets in groups*, Combinatorica 32 (2012), 721–733.
[65] H. Plünnecke: *Eine zahlentheoretische Anwendung der Graphtheorie*, J. reine angew. Math. 243 (1970), 171–183.
[66] L. Pyber and E. Szabó: *Growth in finite simple groups of Lie type*, J. American Math. Soc. 29 (2016), 95–146.
[67] O. Roche–Newton, M. Rudnev and I. Shkredov: *New sum-product type estimates over finite fields*, Adv. Math. 293 (2016) 589–605.
[68] K.F. Roth: *On certain sets of integers*, J. London Math. Soc. 28 (1953), 245–252.
[69] M. Rudnev and I. Shkredov: *On growth rate in* $\mathrm{SL}_2(\mathbf{F}_p)$*, the affine group and sum-product type implications*, Mathematika (2022), 738–783.
[70] I. Ruzsa: *Solving a linear equation in a set of integers, I*, Acta Arith. 65 (1993), 259–282.
[71] I. Ruzsa: *Generalized arithmetical progressions and sumsets*, Acta Math. Hung. 65 (1994), 379–388.
[72] I. Ruzsa: *An infinite Sidon sequence*, J. Number Theory 68 (1998), 63–71.
[73] R. Salem and D.C. Spencer: *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. 28 (1942), 561–563.
[74] T. Schoen: *New bounds in Balog–Szemerédi–Gowers*, Combinatorica 35 (2015), 695–701.
[75] S. Sidon: *Ein Satz über trigonometrische Polynome und seine Anwendung in der Theorie der Fourier-Reihen*, Math. Annalen 106 (1931), 536–539.
[76] J. Singer: *A theorem in finite projective geometry and some applications to number theory*, Transactions of the AMS 43 (1938), 377–385.
[77] J. Solymosi: *Bounding multiplicative energy by the sumset*, Advances in Math. 222 (2009), 402–408.
[78] K. Soundararajan: *Finite fields, with applications to combinatorics*, Student Math. Library 99, American Math. Soc., 2022.
[79] S. Stevens and F. de Zeeuw: *An improved point-line incidence bound over arbitrary fields*, Bull. Lond. Math. Soc. 49 (2017), 842–858.
[80] E. Szemerédi: *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.
[81] E. Szemerédi and W.T. Trotter: *Extremal problems in discrete geometry*, Combinatorica 3 (1983), 381–392.
[82] L. Takács: *The problem of coincidences*, Arch. Hist. Exact Sci. 21 (1979/80), 229–244.
[83] T. Tao: *From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE*, Notices of the AMS 48 (2001), 294–303.
[84] T. Tao and V. Vu: *Additive combinatorics*, Cambridge Studies in Advanced Math. 105, Cambridge Univ. Press, 2006.
[85] P. Varnavides: *On certain sets of positive density*, Journal London Math. Soc. 34 (1959), 358–360.
[86] B. van der Waerden: *Beweis einer Baudet'schen Vermutung*, Nieuw Archief voor Wiskunde 15 (1928), 212–216; available at https://www.delpher.nl/nl/tijdschriften/view?coll=dts&identifier=MMKWG01:022157001:00228
[87] Y. Zhao: *Graph theory and additive combinatorics*, Cambridge Univ. Press, 2023.