BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 50, Number 1, January 2013, Pages 93–127
S 0273-0979(2012)01392-7
Article electronically published on October 2, 2012

THE STRUCTURE THEORY OF SET ADDITION REVISITED

TOM SANDERS

Abstract. In this article we survey some of the recent developments in the structure theory of set addition.

1. Introduction

The purpose of this survey is to review some recent advances in Freĭman's theorem, one of the central results in what is called the structure theory of set addition. This theory was first systematically developed by Freĭman in [Fre66, Fre73a], and a large part of it is concerned with the question, "What do approximate groups look like?"

In fact we shall be interested in what *Abelian* approximate *cosets of sub*groups look like. To craft a more concrete question, it is useful to have some notation: suppose, as it shall be throughout, that G is an Abelian group. Given $A, A' \subset G$, we write $A + A'$ for the **sumset** of A and A' which is defined by

$$A + A' := \{a + a' : a \in A, a' \in A'\}.$$

Given this, it is easy to check that a subset W of G is a coset (of a subgroup) if and only if

$$W \neq \emptyset \text{ and } |W + W| = |W|.$$

As indicated we are interested in approximate cosets and to this end we relax these requirements so that they are only approximately true. Relaxing the first requirement does not lead to an interesting generalisation; for the second we ask that the sumset be "not much larger" than the original set. To be clear, given $K \geq 1$ we say that A (non-empty) has **doubling**[1] K if $|A + A| \leq K|A|$ and are interested in which sets have this property.

We shall be interested in the case when the doubling is small, and to get a sense of what this means, it is worth noting that trivially *every* set has doubling $|A|$ since there cannot be more elements in $A + A$ than there are pairs in A^2. (In fact this can trivially be improved to $(|A| - 1)/2$ but our interest at this stage is really in orders of magnitude.)

It may be instructive on a first read to think of $K = O(1)$ as $|A| \to \infty$, although it will turn out later that we can allow K to grow (slowly) with $|A|$.

If A is a coset, then A has doubling 1 which is certainly small, but are there any other sets with small doubling? One way to create such sets is to take a large

Received by the editors July 16, 2012, and, in revised form, August 20, 2012.
2010 *Mathematics Subject Classification.* Primary 11B13.

[1]One might very reasonably suggest that one use the phrase "doubling ratio" instead of "doubling" here. While this would be sensible, this is not the terminology in use in the subject, and to maintain consistency with existing literature, we shall follow the standard terminology.

subset of a coset. In particular, suppose that W is a coset in G and $A \subset W$ is such that $|W| \leq K|A|$. Then since $A + A \subset W + W$, we conclude that A has doubling K.

It turns out that if K is small enough, then the above construction is characteristic in the sense that it is the *only* way to create sets with doubling K. This was proved by Freĭman in [Fre73b] and appears as [TV06, Exercise 2.6.5]. At this point it is worth remarking that the book [TV06] of Tao and Vu is the standard text for many of the better known aspects of the material we shall be discussing and, where possible, we have given references to that alongside the original source.

Proposition 1.1. *Suppose that $|A + A| \leq K|A|$ for some $K < 1.5$. Then there is a subgroup H of size at most $K|A|$ such that A is contained in a coset of H.*

(For the unfamiliar it may be worth saying that this result is *not* the Freĭman's theorem we shall ultimately be interested in.)

The result shows that the only way of creating sets with doubling less than 1.5 is the method described before Proposition 1.1 and, moreover, every set created in that way has doubling less than 1.5: the result characterises sets with doubling less than 1.5.

There is a good reason for the limitation of 1.5 above, and that is because there is a qualitatively new way of constructing sets with small doubling. Suppose that $H \leq G$, $x + H \in G/H$ has order 4 and put $A := H \cup (x + H)$. Then a short calculation shows that $|A + A| = 1.5|A|$, but any coset containing A has size at least $2|A|$.

Instead of taking large subsets of one coset, our new construction takes unions of cosets (of the same subgroup). In light of this we introduce a new piece of terminology:[2] we say that a set A is **k-covered** by B if there is a set X of size at most k such that $A \subset X + B$.

One can combine the two ways of creating sets with small doubling by considering (disjoint) unions of large subsets of cosets (of the same subgroup) to produce more sets with small doubling, and it turns out that while the doubling remains less than 2, this is the only way of creating sets with doubling less than 2.

Proposition 1.2. *Suppose that $|A + A| \leq K|A|$ for some $K < 2 - \epsilon$. Then there is a subgroup H of G such that $|H| \leq K|A|$ and A is $O_\epsilon(1)$-covered by H.*

Unlike Proposition 1.1 this result is *not* characteristic in that not every set satisfying the conclusion has doubling strictly less than 2. Indeed, the doubling of such sets may be much larger than 2. In fact a more precise characterisation of sets with doubling in this range is available in the form of Kneser's theorem ([Kne53] or [TV06, Theorem 5.5]) from which Proposition 1.2 follows via something called a covering argument of a sort we shall see later in §4.

The example to highlight the limitation of Proposition 1.1 was the first in a series of examples generated by longer and longer arithmetic progressions, and these examples go some way to explaining why 2 should be a critical point in Proposition 1.2. Indeed, if $G = \mathbb{Z}$ and A is a finite arithmetic progression, then A has doubling $(2-1/|A|)$. If some version of the conclusion of Proposition 1.2 were to hold without the dependence on ε, then we should need to cover A by $O(1)$ cosets of a subgroup $H \leq G$ of size $O(|A|)$. Of course the only finite subgroup of $\mathbb{Z}$ is $\{0\}$ and so this is not possible.

[2]This follows Green and Ruzsa [GR06], and has been much popularised by Tao [Tao08].

This last example shows us that if we are to have a hope of describing sets with doubling 2, then we shall need to admit another form of structure: long arithmetic progressions. An arithmetic progression can be thought of as a discrete representation of an interval, and in this light it can be seen as a special case of a more general structure which, it turns out, also has small doubling: lattices in convex bodies.

A **centred convex progression** is a set P in G, a symmetric convex body Q in $\mathbb{R}^d$ and a homomorphism $\phi : \mathbb{Z}^d \to G$ such that $\phi(\mathbb{Z}^d \cap Q) = P$. We say that P is **d-dimensional**, and we shall usually simply talk about the set P with Q and ϕ being implied (despite the fact that they are not necessarily well defined).

Given this definition, a (symmetric) arithmetic progression is a 1-dimensional centred convex progression, and all 1-dimensional centred convex progressions are (symmetric) arithmetic progressions.

A convex body in $\mathbb{R}^d$ has doubling 2^d, and it turns out that this doubling property is inherited by d-dimensional convex progressions in the sense that they have doubling $\exp(O(d))$. The proof of this is not very difficult and can be done using a covering argument. The details are in Lemma 4.2 to avoid breaking the flow.

Given a set of small doubling, we can always create a new set with small doubling by adding a subgroup. In light of this we define a **d-dimensional centred convex coset progression** to be a set of the form $P+H$ where P is a d-dimensional centred convex progression and H is a subgroup of G; this also has doubling $\exp(O(d))$. (Again, see Lemma 4.2 for a proof.)

With this new type of structure we can set about constructing a large class of sets with small doubling (small here meaning $O(1)$). In our earlier discussion we found two methods of producing sets with small doubling from subgroups: we could take large subsets, and we could take a union of a small number of cosets. We now replace "subgroup" in these constructions by "centred coset progression".

Suppose that A is $\exp(d)$-covered by a d-dimensional centred convex coset progression M of size at most $\exp(d)|A|$. Then by definition there is a set X of size at most $\exp(d)$ such that $A \subset X + M$, whence

$$|A + A| \leq |X + M + X + M| \leq |X|^2|M + M| = \exp(O(d))|A|, \tag{1.1}$$

so that A has doubling $\exp(O(d))$. Remarkably it turns out that the above is the *only* way of constructing sets of small doubling.

Theorem 1.3 (Green-Ruzsa theorem; Freĭman's theorem for Abelian groups)**.** *Suppose that $|A+A| \leq K|A|$. Then A is $\exp(d(K))$-covered by a $d(K)$-dimensional centred convex coset progression M of size at most $\exp(d(K))|A|$.*

The result above was first proved by Freĭman [Fre66] for the case of G torsion-free, and later a new proof with better bounds was given (for the same setting) by Ruzsa in [Ruz94]. In [Ruz99] Ruzsa proved the result for groups of bounded exponent which is, in some sense, at the other end of the spectrum from torsion-free. Then Green and Ruzsa in [GR07] established the result above for arbitrary (Abelian) groups with another proof appearing a little later in [TV06, Theorem 5.43].

While Theorem 1.3 resolves the qualitative question of the structure of sets with small doubling, the quantitative question remains, and this is where most of the recent advances have been. In their first proof of Theorem 1.3 Green and Ruzsa

showed that one may take

$$d(K) = O(K^{4+o(1)}).$$

Various strengthenings were available at that time for torsion-free and groups of bounded exponent. (See, for example, [Cha02] or the appendix to [Bou08] for the torsion-free case, and [GT09b] for the bounded exponent case.) Unfortunately, all bounds were of the form $d(K) = O(K^C)$ for some $C > 0$, and it was seen as a significant open problem to show $d(K) = O(K^{o(1)})$.

In [Sch11] Schoen made a striking breakthrough proving a bound of the form[3]

$$d(K) = O(\exp(O(\sqrt{\log K}))),$$

and then shortly after that Croot and Sisask came out with an important new argument in [CS10] which it turned out could be used to prove

$$d(K) = O(\log^{3+o(1)} K).$$

Establishing this is one of the main goals of this survey; to be clear we shall prove the following version of Theorem 1.3.

Theorem 1.4 (Green-Ruzsa theorem, good bounds)**.** *Suppose that* $|A+A| \leq K|A|$*. Then* A *is* $\exp(O(\log^{3+o(1)} K))$*-covered by an* $O(\log^{3+o(1)} K)$*-dimensional centred convex coset progression* M *of size at most* $\exp(O(\log^{3+o(1)} K))|A|$*.*

This result with a power of 6 instead of 3 was shown in [San10], and in the basic framework of this paper that 6 improves to a 4. An improvement of the 4 to a 3 is the result of a wonderful iterative application of our basic tool, which is due to Konyagin.

For comparison the calculation in the construction before Theorem 1.3 turns out to be tight, and it follows from this that $d(K) = \Omega(\log K)$, and this is conjecturally the correct order of magnitude. To see this, suppose that $|A + A| = K|A|$ and note by the calculation in (1.1) that

$$K|A| \leq \exp(2d(K)) \exp(O(d(K))) \exp(d(K))|A| = \exp(O(d(K)))|A|,$$

from which the lower bound on $d(K)$ follows.

Conjecture 1.5 (Polynomial Freĭman-Ruzsa conjecture)**.** *Suppose that* A *has* $|A + A| \leq K|A|$*. Then* A *is* $\exp(O(\log K))$*-covered by an* $O(\log K)$*-dimensional centred convex coset progression* M *of size at most* $\exp(O(\log K))|A|$*.*

We have skipped over a number of the details in this introduction, but before moving on to a more careful discussion, it is worth making a couple of remarks on why Freĭman's theorem is important.

First there is a practical reason: as a result of the celebrated work of Gowers [Gow98, Gow01] in the late 1990s, Freĭman's theorem has found a bevy of applications. For example, Gowers himself used it to spectacularly improve the bounds in Szemerédi's theorem; Szemerédi and Vu used it to investigate long arithmetic progressions in [SV06]; Tao and Vu used it to investigate random matrices in [TV07]; Schoen records many shorter consequences at the end of his paper [Sch11] on Freĭman's theorem; and Chang in [Cha09] collects together a number of other

[3]This is our first use of logs in this survey, and they will appear a lot more. We shall always think of the argument as being larger than some constant, but if the reader does not wish to concern themselves with this, then they may think of $\log x$ as denoting $\log(2 + x)$.

applications where good bounds would be particularly useful. There is some discussion of applications at the end of this article in §13.

Second, there are good theoretical reasons, three of which we shall record now. They may not all make precise sense at this point in the article, but part of our hope is that we shall be able to go some way towards explaining them.

(i) The hypothesis of the theorem is easily satisfied. In a sense we have seen that this is true empirically as a result of the many applications. From a theoretical perspective this is because convex coset progressions are ubiquitous in contrast to subgroups (in some groups). An example to bear in mind is $G = \mathbb{Z}/p\mathbb{Z}$ for p a prime. This has a very poor subgroup structure, but since arithmetic progressions are convex coset progressions, we see immediately that there is an abundance of convex progressions.

(ii) A convex coset progression supports a lot of structure. While it is not a coset, it behaves enough like a coset that it can support many commonly used analytic arguments, and in particular a sort of approximate harmonic analysis. This means that many results for groups can also be established for convex coset progressions. The pioneering work here is that of Bourgain [Bou99], which was framed in a level of generality that includes convex coset progressions by Green and the author in [GS08].

(iii) Finally, the result is a rough equivalence: any set satisfying the conclusion of the theorem satisfies the hypothesis with K replaced by $\exp(O(d(K)))$. Thus the better the bound on the function $d(K)$, the less loss there is in passing from the implicit algebraic data that a set has small doubling to the explicit algebraic data that it is generated from a convex coset progression.

The paper is organized as follows. In the next section, §2, we describe the main plan of attack on Freĭman's theorem which roughly splits it into two parts. The first part is covered in §§4–8; the second in §§9–11. There is a concluding section in §12, and also a section on Plünnecke's inequality in §3, which is a basic tool in the structure theory of set addition and which has recently received a fantastic new proof by Petridis.

2. Overview

The proof of Theorem 1.4 splits naturally into two parts: one covers the more combinatorial aspects, and one the more harmonic analytic aspects. This particular decoupling can be said to originate with the work of Green and Ruzsa [GR07], although their focus was much more on the second of the two, while the more recent improvements to the bounds have arisen (largely) from more careful combinatorial analysis in the first part of the argument.

The key definition is that of relative polynomial growth: to be clear, we say that a set X has **relative polynomial growth of order** d if

$$|nX| \leq n^d |X| \text{ for all } n \in \mathbb{N}.$$

One might reasonably wish to insert a constant in front of the term on the right-hand side, but we shall find that we are easily able to absorb this into the dimension at little cost to the quality of our eventual bounds.

It is worth noting that having relative polynomial growth is *a priori* stronger than a small doubling condition. It will turn out later (see Proposition 5.1) that the conditions are qualitatively equivalent in that doubling K implies relative polynomial

growth of order $O_K(1)$, but quantitatively this equivalence entails an exponential loss and is the reason for the exponential weakness of the original arguments of Green and Ruzsa.

With the definition above the argument splits into the following two parts.

(i) *From small doubling to relative polynomial growth.* Given a set A with $|A+A| \leq K|A|$, we find a **symmetric neighbourhood of the identity** X (meaning that $X = -X$ and $0_G \in X$) of size at most $O_K(|A|)$ with relative polynomial growth of order $O_K(1)$ such that A is $O_K(1)$-covered by X.

(ii) *From relative polynomial growth to convex coset progressions.* Given a symmetric neighbourhood of the identity X with relative polynomial growth of order d, we show that X is contained in an $O_d(1)$-dimensional centred convex coset progression of size at most $O_d(|X|)$.

Note that if we had proved these two statements, then they combine to give Theorem 1.3. We now turn to look at these two parts in a little more detail.

2.1. **From small doubling to relative polynomial growth.** The starting point here is the covering arguments of Ruzsa, which will be developed in §4 and which will be related to relative polynomial growth in §5. As we shall see, it is possible to use these covering arguments to show that if $|A+A| \leq K|A|$, then A has relative polynomial growth of order $O(K^4)$ and from there it is a short step to the following corollary.

Corollary 2.2. *Suppose that $|A+A| \leq K|A|$. Then A is 1-covered by a symmetric neighbourhood of the identity of size at most $\exp(O(\log K))|A|$ and relative polynomial growth of order $O(K^4)$.*

This result is much weaker than we should like, but it turns out that it is essentially so because it provides a set which 1-covers. In §12 we discuss an example of a set A with doubling K such that any set 1-covering it must have either relative polynomial growth of order $\Omega(K)$ or size $\exp(\Omega(K))|A|$. Thus to improve the bound on the order of relative polynomial growth, we shall need to increase the covering number.

In §6 we discuss a general framework for improving the above Corollary 2.2 before §7 where we introduce a key new tool: the Croot-Sisask lemma. Section 7 includes the following result which can be seen as representing the state of the art prior to Schoen [Sch11] and Croot and Sisask [CS10] (although we shall use a special case of the Croot-Sisask lemma to prove it).

Proposition 2.3. *Suppose that $|A+A| \leq K|A|$. Then A is $\exp(O(K^{1+o(1)}))$-covered by a symmetric neighbourhood of the identity of size at most $\exp(O(\log K))|A|$ and relative polynomial growth of order $O(K^{1+o(1)})$.*

In §8 we shall then make much more effective use of the Croot-Sisask lemma to show the following.

Proposition 2.4. *Suppose that $|A+A| \leq K|A|$. Then A is $\exp(O(\log^4 K))$-covered by a symmetric neighbourhood of the identity of size at most $\exp(O(\log K))|A|$ and relative polynomial growth of order $O(\log^4 K)$.*

This result is where most of the more recent new material appears, but there is then also a combinatorial refinement following Konyagin which leads to our strongest result at the end of §8.

Proposition 2.5. *Suppose that* $|A+A| \leq K|A|$. *Then* A *is* $\exp(O(\log^{3+o(1)} K))$*-covered by a symmetric neighbourhood of the identity of size at most* $\exp(O(\log^{1+o(1)} K))|A|$ *and relative polynomial growth of order* $O(\log^{3+o(1)} K)$.

2.6. **From relative polynomial growth to convex coset progressions.** To pass from relative polynomial growth to convex coset progressions, it is useful to start by considering some examples of sets with relative polynomial growth. Of course, if Q is a convex body in $\mathbb{R}^d$, then $\mu(nQ) = n^d\mu(Q)$ for all $n \geq 1$, and so one expects that any d-dimensional centred convex coset progression has relative polynomial growth roughly d (in fact $d^{1+o(1)}$).

Now, if P is a d-dimensional centred convex coset progression and $X \subset P$ has size $\exp(-d^{1+o(1)})|P|$, then

$$|nX| \leq |nP| \leq n^{d^{1+o(1)}}|P| = O(n)^{d^{1+o(1)}}|X| = n^{d^{1+o(1)}}|X| \text{ for all } n \geq 1.$$

Crucially, though, a union of $\exp(d^{1+o(1)})$-translates of centred convex coset progressions will (generically) have relative polynomial growth of order $\exp(d^{1+o(1)})$ and *not* $d^{1+o(1)}$ so that relative polynomial growth distinguishes between covering and containment in a way that doubling does not.

It turns out that there is a matching result which tells us that essentially the only way of creating sets of relative polynomial growth is by the above method.

Theorem 2.7. *Suppose that* X *has relative polynomial growth of order* d. *Then there is a centred convex coset progression* M *such that*

$$X - X \subset M, \dim M = O(d\log^2 d) \text{ and } |M| \leq \exp(O(d\log^2 d))|X|.$$

The first thing to say is that the dimension here is tight up to factors of $\log d$. This can be seen by, for example, letting X be the cube of side length N in $\mathbb{Z}^d$. This has polynomial growth of order $\Omega(d)$, and any convex coset progression containing X has tripling at least 2^d by the discrete Brunn-Minkowski inequality (see, *e.g.*, [GT06, Lemma 2.4]) and so has dimension $\Omega(d)$.

This result is the part of the argument which uses harmonic analysis and itself splits into a number of parts. These are covered in the second part of the paper starting at §9.

We should remark that Theorem 1.4 follows immediately from Proposition 2.5 and Theorem 2.7.

3. Plünnecke's inequality

This section is the final section before we plunge into the proof of Freĭman's theorem, and it will cover the invaluable tool of Plünnecke's inequality following the exciting new work by Petridis [Pet11b, Pet11a]. The discussion in his papers is more comprehensive than ours, and we direct the reader interested in more details there, but we hope to cover the salient features in what follows.

Our starting point is the observation that *given* Freĭman's theorem if $|A+A| \leq K|A|$, then there is an $O_K(1)$-dimensional centred convex coset progression M of size $O_K(|A|)$ such that A is $O_K(1)$-covered by M. This means that there is some set X of size $O_K(1)$ such that $A \subset X + M$. On the other hand as remarked in §2.6 the set M has relative polynomial growth of order $O_K(1)$, hence

$$|nA| \leq |nX||nM| \leq |X|^n(n)^{O_K(1)}|M| = O_K(1)^n|A| \text{ for all } n \in \mathbb{N}.$$

It turns out that a much stronger inequality is true:

Theorem 3.1 (Plünnecke's inequality)**.** *Suppose that* $|A+A| \leq K|A|$*. Then*

$$|nA| \leq K^n|A| \text{ for all } n \in \mathbb{N}.$$

This result is due to Plünnecke [Plü69] and was rediscovered and greatly developed by Ruzsa [Ruz89]. Both Ruzsa and Plünnecke's arguments were graph theoretic and quite involved appealing to Menger's theorem (see [TV06, §6.5] for details). In [Pet11b] Petridis removed the need for Menger's theorem and then a little later in [Pet11a] he found a wonderful entirely new proof.

The core of Petridis's argument is the next lemma. The idea is that if we are given sets A and X such that $|A+X| \leq K|X|$, then it is a good idea to pass to the *best* possible subset of X; that is to say, to pass to the subset X' of X for which $|A+X'|/|X'|$ is minimal. Turning this around, if X is already the best subset, then every $X' \subset X$ has $|A+X'|/|X'|$ bigger than $|A+X|/|X|$. In this case Petridis proved the following beautiful lemma.

Lemma 3.2. *Suppose that* $|A+X| \leq K|X|$ *and* $|A+X'| \geq K|X'|$ *for all* $X' \subset X$*. Then for all (finite) sets* C *we have*

$$|A+X+C| \leq K|X+C|.$$

Proof. We iteratively decompose $X+C$ into disjoint sets contained in translates of X: $X+C = \bigsqcup_c X_c$ where $X_c \subset X+c$. Writing $Y_c := (X+c) \setminus X_c$, we have

$$\begin{aligned} |A+X+C| \leq \sum_c |A+X_c| &= \sum_c |A+((X+c)\setminus Y)| \\ &\leq \sum_c (|A+X+c| - |A+Y_c|) \\ &\leq \sum_c (K|X+c| - K|Y_c|) \\ &= \sum_c K|X_c| = K|X+C|. \end{aligned}$$

The result is proved. □

Given the idea of passing to this best possible X the proof is rather natural, but the reader should make no mistake: the idea to do this is very nice and has eluded many people!

Petridis then gives the following immediate corollary.

Corollary 3.3. *Suppose that* $|A+B| \leq K|B|$*. Then there is some non-empty* $X \subset B$ *such that*

$$|nA+X| \leq K^n|X| \text{ for all } n \in \mathbb{N}.$$

Proof. We can pick $X \subset B$ such that $|A+X|/|X|$ is minimal over (non-empty) subsets of B. In this case A and X satisfy the hypotheses of Petridis's lemma and hence the conclusion. Applying the conclusion with $C = (n-1)A$, we get that

$$|X+nA| \leq K|X+(n-1)A| \text{ for all } n \in \mathbb{N},$$

and this gives the result (by induction). □

Note that Plünnecke's inequality (Theorem 3.1) follows immediately from this applied to the set A and $B = A$ since $X \subset B = A$.

It is also possible to control jointly positive and negative sums of A using the following result called Ruzsa's triangle inequality [Ruz78] (see also [TV06, Lemma 2.6]).

Lemma 3.4 (Ruzsa's triangle inequality)**.** *Suppose that* $|A-B| \leq K|B|$ *and* $|B-C| \leq L|B|$. *Then*

$$|A+C| \leq KL|B|.$$

Proof. We consider the map $B \times (A+C) \to (A-B) \times (B-C)$ defined by $(b,s) \mapsto (a(s)-b, b-c(s))$ where $a(s)$ and $c(s)$ are functions on $A+C$ such that $a(s) \in A$, $c(s) \in C$ and $a(s)+c(s)=s$. It is easy to check that our map on $B \times (A+C)$ is an injection: suppose that

$$(a(s)-b, b-c(s)) = (a(s')-b', b'-c(s')),$$

then adding we get that $s = a(s)+c(s) = a(s')+c(s') = s'$ and so $s=s'$, and hence $b=b'$. It follows from this that $|B||A+C| \leq |A-B||B-C|$, and we have the result. □

It may be intuitively helpful to know that this can be seen as the triangle inequality for a certain pseudo-metric one can define on sets (in groups) called the Ruzsa distance. (See [TV06, §2.3] for more details.)

As an immediate corollary of our work so far we have the so-called Plünnecke-Ruzsa inequalities, which are slightly more general than Plünnecke's inequality.

Corollary 3.5 (Plünnecke-Ruzsa inequalities)**.** *Suppose that* $|A+A| \leq K|A|$. *Then*

$$|nA-mA| \leq K^{n+m}|A| \text{ for all } n,m \in \mathbb{N}.$$

Proof. Apply Corollary 3.3 to get a set $X \subset A$ such that $|rA+X| \leq K^r|X|$ for all $r \in \mathbb{N}$ so that in particular $|nA+X| \leq K^n|X|$ and $|-mA-X| \leq K^m|X|$. It follows from Ruzsa's triangle inequality that $|nA-mA| \leq K^{n+m}|X| \leq K^{n+m}|A|$ as required. □

4. Ruzsa's covering lemma

Plünnecke's inequality showed us how small doubling leads to small higher order sums. In [Ruz99] Ruzsa introduced another argument called a covering argument to the area which yields quantitatively similar order results to Plünnecke's inequality but has the advantage of also providing a little structure. This covering argument is the topic of this section and will already give us a version of Freĭman's theorem in groups of bounded exponent. We start with the basic lemma:

Lemma 4.1 (Ruzsa's covering lemma, [TV06, Lemma 2.14])**.** *Suppose that* $|A+S| \leq K|S|$. *Then there is a set* $T \subset A$ *with* $|T| \leq K$ *such that* $A \subset T+S-S$.

Proof. The technique here is very powerful, so it is worth developing in some detail: We let $T \subset A$ be maximal S-separated. (The set T is *S-separated* if every pair of distinct elements $t,t' \in T$ have $t+S$ and $t'+S$ disjoint.) It follows that $|T+S| = |T||S|$. On the other hand, since $T \subset A$, we have $T+S \subset A+S$, and so

$$|T||S| = |T+S| \leq |A+S| \leq K|S|.$$

We conclude that $|T| \leq K$.

Now we use the fact that T is maximal: if $a \in A$, then (either trivially if $a \in T$ or) by maximality there is some $t \in T$ such that $(t+S) \cap (a+S) \neq \emptyset$. It follows that $a \in t+S-S \subset T+S-S$, and the result is proved. □

It should be remarked that this has an extension developed by Tao in [Tao08] giving a non-Abelian version of a (slightly weak) Plünnecke inequality, although now Petridis's approach to Plünnecke's inequality also yields a non-Abelian version of the (almost) full strength Plünnecke inequality.

Lemma 4.1 (or rather the technique used to prove it) can also be used to show that d-dimensional centred convex progressions have doubling $\exp(O(d))$.

Lemma 4.2. *Suppose that M is a d-dimensional centred convex coset progression. Then $|M+M| \leq \exp(O(d))|M|$.*

Proof. To start with, we write $M = P + H$ for some centred convex progression P and $Q \subset \mathbb{R}^d$ for the convex body generating P. Given $\lambda \in \mathbb{R}_{>0}$, we write λQ for the set Q dilated by a factor λ so that $\mu(\lambda Q) = \lambda^d \mu(Q)$.

Now, let $X \subset 2Q$ be a maximal $\frac{1}{4}Q$-separated set so that by the same argument as in Ruzsa's covering lemma we have

$$2Q \subset X + \frac{1}{4}Q - \frac{1}{4}Q = X + \frac{1}{2}Q \text{ and } |X|4^{-d}\mu(Q) \leq (9/4)^d \mu(Q).$$

From the second of these it follows that $|X| \leq 9^d$. With the first we note that

$$P + P = \phi(Q \cap \mathbb{Z}^d) + \phi(Q \cap \mathbb{Z}^d) \subset \phi(2Q \cap \mathbb{Z}^d) \subset \bigcup_{x \in X} \phi((x + \frac{1}{2}Q) \cap \mathbb{Z}^d).$$

Let T be a set such that if $\phi((x + \frac{1}{2}Q) \cap \mathbb{Z}^d) \neq \emptyset$ for some $x \in X$, then T contains exactly one element of this set, so that $|T| \leq |X|$. Then if $t' \in \phi((x + \frac{1}{2}Q) \cap \mathbb{Z}^d)$, we have some $t \in T$ such that $t - t' \in \phi(Q \cap \mathbb{Z}^d) = P$, whence $P + P \subset T + P$. Adding H we get that $(P + H) + (P + H) \subset T + (P + H)$ since $H + H = H$, and the result follows given the bound on $|X|$ (and hence $|T|$). □

One informative illustration of why Ruzsa's covering lemma is so powerful is given in Ruzsa's original paper [Ruz99].

Proposition 4.3 (Freĭman-Ruzsa theorem for groups of bounded exponent)**.** *Suppose that G has exponent r and $|A+A| \leq K|A|$. Then $\langle A \rangle$, the group generated by A, has size at most $K^2 r^{K^4}|A|$.*

Proof. The idea is simply to apply Ruzsa's covering lemma to $2A - A$. By the Plünnecke-Ruzsa inequalities we have that $|(2A - A) + A| = |3A - A| \leq K^4|A|$ and so there is a set T of size at most K^4 such that

$$A + (A - A) = 2A - A \subset T + A - A.$$

By induction it follows that $nA + (A - A) \subset nT + (A - A)$ for all $n \in \mathbb{N}$. We write H for the group generated by T and note that $|H| \leq r^{|T|}$ and $nT + A - A \subset H + A - A$. We conclude that $nA \subset nA + (A - A) \subset H + A - A$ for all n and similarly for $-nA$ since H and $A - A$ are symmetric. It follows that $\langle A \rangle \subset H + A - A$, and we get the result since $|A - A| \leq K^2|A|$. □

Ruzsa has a further argument published in [DHP04] which improves the K^4 above to a K^3 using a slight refinement of the Plünnecke-Ruzsa inequalities. A refined covering argument of Green and Ruzsa [GR06] gives the best known result following from covering techniques, while the best known upper bound by any argument is due to Schoen [Sch11] who showed that the group generated by A has size at most $r^{K^{1+o(1)}}|A|$.

It may also be worth noting that by letting A be $2K+1$ independent elements the upper bound is at least $r^{2K+1}|A|$ so that Schoen's result is tight up to the $o(1)$-term. (In the case when $r=2$, this $o(1)$-term has been eliminated via some arguments from extremal set theory introduced by Green and Tao. We shall not pursue this here, but see [GT09a, Kon08] and [Zoh11] for details.)

5. Relative polynomial growth

In §2 we made it clear that relative polynomial growth was going to be a key concept for us, and it arises naturally when we compare the results of §3 with those of §4, as we shall now see.

Suppose that $|A+A| \leq K|A|$. By Plünnecke's inequality we have that

$$|nA| \leq K^n|A| \text{ for all } n \in \mathbb{N}.$$

On the other hand, by an inductive application of Ruzsa's covering lemma (as in the proof of Proposition 4.3) we have that $(n-1)A+(A-A) \subset (n-1)T+A-A$ for all $n \in \mathbb{N}$ and some T of size at most K^4. Now since G is Abelian, we have

$$|(n-1)T| \leq \binom{|T|+n-2}{|T|-1} \leq n^{|T|},$$

and so

$$|nA| \leq |(n-1)T|.|A-A| \leq K^2|(n-1)T||A| \leq n^{O(K^4)}|A|$$

for all $n \in \mathbb{N}$. For small values of n this is much weaker than Plünnecke's inequality, but for large values of n, the estimate from Plünnecke is exponential while this is polynomial.

Proposition 4.3 does not adapt directly to the case of general Abelian groups because when G does not have bounded exponent, we cannot expect the group generated by A to be finite (consider, for example, $G=\mathbb{Z}$), but as we saw above it is sufficient to give relative polynomial growth.

Proposition 5.1. *Suppose that $A \subset G$ has $|A+A| \leq K|A|$. Then A has relative polynomial growth of order $O(K^4)$.*

We have an immediate corollary of this in the following.

Corollary (Corollary 2.2)**.** *Suppose that $|A+A| \leq K|A|$. Then A is 1-covered by a symmetric neighbourhood of the identity of size at most $\exp(O(\log K))|A|$ and relative polynomial growth of order $O(K^4)$.*

Proof. Since A has relative polynomial growth of order $O(K^4)$ by Proposition 5.1, we see by Ruzsa's triangle inequality that $|n(A-A)| \leq |nA-A||-A-nA|/|A| = n^{O(K^4)}|A|$ and so $A-A$ also has relative polynomial growth of order $O(K^4)$. On the other hand $|A-A| \leq K^2|A|$ and $A-A$ is a symmetric neighbourhood of the identity which 1-covers A, and so we are done. □

The weakness of this result is that it is exponentially expensive to apply: if A has relative polynomial growth of order d, then A trivially has doubling at most 2^d. This means that if A has doubling K and we apply the proposition, we get that A has polynomial growth of order $O(K^4)$, but then we conclude that A has doubling at most $\exp(O(K^4))$ — an exponential loss. Incidentally, this exponential loss is exactly the reason for the exponential loss in Green and Ruzsa's first version of Freĭman's theorem.

To deal with this situation, we have a slight refinement of Ruzsa's covering lemma due to Chang [Cha02]. Chang observed that if a set has a sort of relative sub-exponential growth on one scale, then the covering set T in Ruzsa'c overing lemma can be made to be highly structured and hence get relative polynomial growth of much lower order on all scales. To be clear we need some notation: write

$$\mathrm{Span}(X) := \{\sigma.X := \sum_{x\in X}\sigma_x x : \sigma \in \{-1,0,1\}^X\}.$$

Then we have the following result.

Lemma 5.2 (A variant of Chang's covering lemma)**.** *Suppose that $|kA+S| < 2^k|S|$ (and $0_G \in A$). Then there is a set $T \subset A$ with $|T| < k$ such that $A \subset \mathrm{Span}(T) + S - S$.*

Proof. Let T be a maximal S-*dissociated* subset of A, that is a maximal subset of A such that

$$(\sigma.T + S) \cap (\sigma'.T + S) = \emptyset \text{ for all } \sigma \neq \sigma' \in \{0,1\}^T.$$

Now suppose that $x' \in A\setminus T$ and write $T' := T\cup\{x'\}$. By maximality of T there are elements $\sigma, \sigma' \in \{0,1\}^{T'}$ such that $(\sigma.T' + S) \cap (\sigma'.T' + S) \neq \emptyset$. Now if $\sigma_{x'} = \sigma'_{x'}$, then $(\sigma|_T.T + S) \cap (\sigma'|_T.T + S) \neq \emptyset$ contradicting the fact that T is S-dissociated. Hence, without loss of generality, $\sigma_{x'} = 1$ and $\sigma'_{x'} = 0$, whence

$$x' \in \sigma'|_T.T - \sigma|_T.T + S - S \subset \mathrm{Span}(T) + S - S.$$

We are done unless $|T| \geq k$; assume it is, and let $T' \subset T$ be a set of size k. Denote $\{\sigma.T' : \sigma \in \{0,1\}^{T'}\}$ by P and note that $P \subset kA$ (since $0_G \in A$), whence

$$2^k|S| = |P + S| \leq |kA + S| < 2^k|S|.$$

This contradiction completes the proof. □

Dissociativity is a very important concept in harmonic analysis and the relative version introduced in the above proof also has many uses. The reader interested in learning more is directed to [TV06, §4.5] or the book [Rud90].

This result yields the following useful corollary.

Corollary 5.3. *Suppose that $X \subset G$ is a symmetric neighbourhood and $|(3k + 1)X| < 2^k|X|$ for some $k \in \mathbb{N}$. Then X has relative polynomial growth of order $O(k)$.*

Proof. Apply Lemma 5.2 to the sets $3X$ and X to get a set T of size less than k such that $3X \subset \mathrm{Span}(T) + 2X$. It follows that $nX \subset (n-2)\,\mathrm{Span}(T) + 2X$, and so

$$|nX| \leq |(n-2)\,\mathrm{Span}(T)||2X| \leq (2n-3)^k|2X| = O(n)^k|X|,$$

provided $n \geq 2$. We conclude that X has relative polynomial growth $O(k)$ as required. □

This will often be combined with the following useful application of Ruzsa's covering lemma.

Lemma 5.4. *Suppose that X is a set of relative polynomial growth of order d and $|A + X| \leq K|X|$. Then A is K-covered by $X - X$, a symmetric neighbourhood of the identity having relative polynomial growth of order $O(d)$.*

Proof. We just apply Ruzsa's covering lemma to get that A is K-covered by $X-X$. This is a symmetric neighbourhood of the identity and $|n(X-X)| \leq n^{O(d)}|X|$ by Ruzsa's triangle inequality and the fact that X has relative polynomial growth of order d. The result follows. □

6. Bogolyubov-Ruzsa-type lemmas

In the last section we proved Proposition 5.1 which converted our small doubling condition into a relative polynomial growth of low order condition. As mentioned there, this was not a particularly efficient process, and so we set about proving Corollary 5.3 to do better. In this section we shall discuss a general framework for using this corollary.

To start with, suppose that X (is symmetric and) has $|X+X| \leq K|X|$. By Plünnecke's inequality we have that

$$|(3k+1)X| \leq K^{3k+1}|X| \text{ for all } k \geq 1,$$

which is *not* smaller than 2^k (unless K is very small, which is a case we have already discussed in the introduction). To get a subexponential estimate, it will be useful to have a result of the following shape.

Proposition 6.1 (Weak Bogolyubov-Ruzsa-type lemma)**.** *Suppose that X is symmetric with $|X+X| \leq K|X|$ and $m \in \mathbb{N}$. Then there is a symmetric neighbourhood of the identity, T, such that*

$$|T| = \Omega_{m,K}(|X|) \text{ and } mT \subset 4X.$$

Before remarking on the proof or history, we should see how such a result can be used to give a set with relative subexponential growth. Given X (symmetric) with $|X+X| \leq K|X|$, we apply the lemma with some parameter m to get a set T as described. On the other hand by Plünnecke's inequality with parameter l, we have that $|4lX| \leq K^{4l}|X|$, and it follows that

$$|mlT| \leq K^{4l}|X| = O_{m,K}(K^{4l}|T|) = \exp(O_K(l + O_m(1)))|T|.$$

At this point put $3k+1 = ml$ and letting $m \to \infty$ very slowly with l, we get

$$|(3k+1)T| = \exp(o_K(k))|T|.$$

It follows that for k sufficiently large in terms of K the right-hand side can be made to be at most 2^k and so Corollary 5.3 can be applied to the set T. Whether this turns out to be useful or not depends entirely on the quality of the lower bound in Proposition 6.1 and establishing results of that type with good bounds will be a major part of the remainder of the paper.

Returning to the history, in the case when X is a thick set (meaning $|X| = \Omega(|G|)$) Proposition 6.1 follows from work of Bogolyubov [Bog39]. Ruzsa in [Ruz94] introduced results of this type to Freĭman's theorem, and the above Proposition does follow from his work. The difference here is that both Bogolyubov and Ruzsa prove stronger statements, in particular showing that the set $4A$ contains a low-dimensional Bohr set (see §9 for a definition); the set T can then be identified as a $1/m$-dilate of this Bohr set.

The structurally weaker version of the Bogolyubov-Ruzsa lemma which we need here is fortunately rather easier to prove and results in stronger bounds. Since our objective is one of bounds, this works out well.

7. The Croot-Sisask lemma

One of the key recent tools that has made advances in Freĭman's theorem possible is called the Croot-Sisask lemma. This was first proved by Croot and Sisask in [CS10] and then refined by Croot, Łaba and Sisask in [CŁS11]. The aim of this section is to give a proof of the Croot-Sisask lemma and then immediately give an application to Freĭman's theorem.

Before starting, we shall need a little notation. As we are interested in sumsets, it will not come as too much of a surprise that we should be using the convolution of functions. First, recall that for $p \in [1,\infty)$ the space $\ell^p(G)$ is the space of functions $f : G \to \mathbb{C}$ endowed with the norm

$$\|f\|_{\ell^p(G)} := \left(\sum_{x\in G} |f(x)|^p\right)^{1/p}.$$

For infinity we take the usual convention that

$$\|f\|_{\ell^\infty(G)} := \max\{|f(x)| : x \in G\},$$

and apart from ℓ^∞ there is one other ℓ^p space of particular importance, and that is ℓ^2. This is also a Hilbert space with inner product defined by

$$\langle f, g\rangle = \sum_{x\in G} f(x)\overline{g(x)} \text{ for all } f, g \in \ell^2(G).$$

Now, given $f, g \in \ell^1(G)$ we define their **convolution** to be the function $f * g$ determined pointwise by

$$f * g(x) := \sum_{y+z=x} f(y)g(z) \text{ for all } x \in G.$$

Given a finite set $A \subset G$, we write μ_A for the uniform probability mass function supported on A. (If G were locally compact rather than discrete, then we should define μ_A as a measure, but we do not need to involve the additional analysis here.)

There are two ways in which convolution is useful. The first is because it is an average: in particular if $f \in \ell^1(G)$ and $A \subset G$ is finite, then $f * \mu_A(x)$ is the average value of f on $x - A$. In general this means that the convolution of two functions is smoother than the constituent functions, and hence the convolution is easier to analyse.

Secondly, convolution is useful to us because

$$A + B := \operatorname{supp} 1_A * 1_B \text{ for all } A, B \subset G,$$

so that we can analyse $A + B$ through the (much easier to understand) function $1_A * 1_B$.

In a certain sense convolution comes from integrating the regular representation, and it will be useful to have some notation for this. We write

$$\rho : G \to \operatorname{Aut}(\ell^2(G)); x \mapsto (f \mapsto \rho_x(f) : G \to \mathbb{C}; y \mapsto f(x + y)).$$

To be concrete, with the regular representation in hand we have that

$$f * g(x) = \langle f, \rho_{-x}(\tilde{g})\rangle \text{ for all } x \in G,$$

where $\tilde{g}(x) = \overline{g(-x)}$ for all $x \in G$.

With this notation we can describe the idea behind the Croot-Sisask lemma. Suppose that S is an arithmetic progression and T is a much shorter arithmetic progression with the same common difference so that $|S+T| \approx |S|$.

Now the Croot-Sisask lemma will tell us that for $f \in \ell^p(G)$ the function $f * \mu_S$ does not change much when we translate by elements of T. To see this, we recall from earlier that $f * \mu_S(x)$ is the average value of f on $x - S$. Then if $t \in T$, we have $x - S + t \approx x - S$ so that the average of f over $x - S + t$ is approximately the same as the average of f over $x - S$.

The full Croot-Sisask lemma is the following much stronger version of this argument replacing arithmetic progressions by any set with small doubling.

Lemma 7.1 (Croot-Sisask)**.** *Suppose that $f \in \ell^p(G)$ for some $p \geq 2$, $S, T \subset G$ are such that $|S+T| \leq L|S|$, and $\eta \in (0,1]$ and $p \in [2,\infty)$ are parameters. Then the set of x is such that*

$$\|\rho_x(f * \mu_S) - f * \mu_S\|_{\ell^p(G)} \leq \eta \|f\|_{\ell^p(G)}$$

is a symmetric neighbourhood of the identity and has size at least $(2L)^{-O(\eta^{-2}p)}|T|$.

The proof proceeds by random sampling. The idea is that since $f * \mu_S$ is pointwise the average value of f on translates of S, this can be well approximated by the average value of f on a small set of "typical" elements of S. We are then done if we let X be the set of elements of G such that translating these typical elements does not vary them very much. To make the notion of being well approximated precise, we shall need an inequality called the Marcinkiewicz-Zygmund inequality, and for this we require a little more notation.

Given $p \in [1,\infty)$ and (X,μ) a measure space, we write $L^p(\mu)$ for the space (of equivalence classes of) measurable functions on X endowed with the norm

$$\|f\|_{L^p(\mu)} := \left(\int |f(x)|^p d\mu(x)\right)^{1/p}.$$

Theorem 7.2 (Marcinkiewicz-Zygmund inequality)**.** *Suppose that $p \in [2,\infty)$ and we are given independent random variables $X_1, \dots, X_n \in L^p(\mathbb{P})$ with $\mathbb{E}\sum_i X_i = 0$. Then*

$$\|\sum_i X_i\|_{L^p(\mathbb{P})} = O\left(\sqrt{p}\|\sum_i |X_i|^2\|^{1/2}_{L^{p/2}(\mathbb{P})}\right).$$

Intuitively, one might like to think of the X_i's are independent variance one, mean zero random variables. Then the central limit theorem suggests that $\sqrt{n}^{-1}\sum_i X_i \sim N(0,1)$ and the pth moments of the normal distribution are well known (and in any case easily computed). We have

$$\|\sum_i X_i\|^p_{L^p(\mathbb{P})} = n^{p/2} \cdot \frac{2^{p/2}\Gamma((p+1)/2)}{\sqrt{\pi}} = O\left(\sqrt{p}\|\sum_i |X_i|^2\|^{1/2}_{L^{p/2}(\mathbb{P})}\right).$$

Thus the Marcinkiewicz-Zygmund inequality can be thought of as saying that nothing much worse than this can happen.

There is a special case of the Marcinkiewicz-Zygmund inequality called Khintchine's inequality, which can be used in the proof of the former.

Theorem 7.3 (Khintchine's inequality)**.** *Suppose that $p \in [2,\infty)$, and we are given independent random variables $X_1,\dots,X_n \in L^p(\mathbb{P})$ with $\mathbb{P}(X_i = a_i) = \mathbb{P}(X_i = -a_i) = 1/2$. Then*

$$\|\sum_i X_i\|_{L^p(\mathbb{P})} = O\left(\sqrt{p}\|\sum_i |X_i|^2\|^{1/2}_{L^{p/2}(\mathbb{P})}\right) = O\left(\sqrt{p}\left(\sum_i |a_i|^2\right)^{1/2}\right).$$

Khintchine's inequality is proved by restricting to the case when p is an even integer (the other cases follow by nesting of norms) and then raising the left-hand side to the power p, multiplying it out and collecting together terms. There are more elegant proofs, but this gives the main idea.

Given this, to prove the Marcinkiewicz-Zygmund inequality, one can proceed by a process of symmetrisation. First, if the variables are complex, then the result follows from taking real and imaginary parts, and so one may as well assume they are real. We then take copies $Y_1,\dots,Y_n$ of $X_1,\dots,X_n$ such that $X_i \sim Y_i$ and $X_1,\dots,X_n,Y_1,\dots,Y_n$ are mutually independent. Following this, we apply Khintchine's inequality to the variables $X_i - Y_i$ restricted to atoms of the sample space on which they are symmetric and only take two values. Collecting all this together gives the result.

Proof of Lemma 7.1. Let $z_1,\dots,z_k$ be independent uniformly distributed S-valued random variables, and for each $y \in G$ define $Z_i(y) := \rho_{-z_i}(f)(y) - f*\mu_S(y)$. For fixed y, the variables $Z_i(y)$ are independent and have mean zero, so it follows by the Marcinkiewicz-Zygmund inequality and Hölder's inequality that

$$\begin{aligned}\|\sum_{i=1}^k Z_i(y)\|^p_{L^p(\mu_S^k)} &\leq O(p)^{p/2}\int\left(\sum_{i=1}^k |Z_i(y)|^2\right)^{p/2} d\mu_S^k\\ &\leq O(p)^{p/2}k^{p/2-1}\sum_{i=1}^k\int |Z_i(y)|^p d\mu_S^k.\end{aligned}$$

Summing over y and interchanging the order of summation, we get

$$\sum_{y\in G}\|\sum_{i=1}^k Z_i(y)\|^p_{L^p(\mu_S^k)} \leq O(p)^{p/2}k^{p/2-1}\int\sum_{i=1}^k\sum_{y\in G}|Z_i(y)|^p d\mu_S^k. \tag{7.1}$$

On the other hand,

$$\left(\sum_{y\in G}|Z_i(y)|^p\right)^{1/p} = \|Z_i\|_{\ell^p(G)} \leq \|\rho_{-z_i}(f)\|_{\ell^p(G)} + \|f*\mu_S\|_{\ell^p(G)} \leq 2\|f\|_{\ell^p(G)}$$

by the triangle inequality. Dividing (7.1) by k^p and inserting the above and the expression for the Z_i's, we get that

$$\int\sum_{y\in G}\left|\frac{1}{k}\sum_{i=1}^k \rho_{-z_i}(f)(y) - f*\mu_S(y)\right|^p d\mu_S^k(z) = O(pk^{-1}\|f\|^2_{\ell^p(G)})^{p/2}.$$

Pick $k = O(\eta^{-2}p)$ such that the right-hand side is at most $(\eta\|f\|_{\ell^p(G)}/4)^p$, and write $\mathcal{L}$ for the set of $x \in S^k$ for which the integrand above is at most $(\eta\|f\|_{\ell^p(G)}/2)^p$; by averaging $\mu_S^k(\mathcal{L}^c) \leq 2^{-p}$ and so $\mu_S^k(\mathcal{L}) \geq 1 - 2^{-p} \geq 1/2$.

Now, $\Delta := \{(t,\dots,t) : t \in T\}$ has $\mathcal{L} + \Delta \subset (S+T)^k$, whence $|\mathcal{L}+\Delta| \leq 2L^k|\mathcal{L}|$ and so

$$\langle 1_\Delta * 1_{-\Delta}, 1_{-\mathcal{L}} * 1_{\mathcal{L}} \rangle_{\ell^2(G^k)} = \|1_{\mathcal{L}} * 1_\Delta\|^2_{\ell^2(G^k)} \geq |\Delta|^2|\mathcal{L}|/2L^k,$$

by the Cauchy-Schwarz inequality.

By averaging it follows that at least $|\Delta|^2/2L^k$ pairs $(z,y) \in \Delta \times \Delta$ have $1_{-\mathcal{L}} * 1_{\mathcal{L}}(z-y) > 0$, and hence there are at least $|\Delta|/2L^k = |T|/2L^k$ distinct elements $x \in T - T \subset G$ with $1_{-\mathcal{L}} * 1_{\mathcal{L}}(x,\dots,x) > 0$; write X for this set.

By design for each $x \in X$ there is some $z(x) \in \mathcal{L}$ and $y(x) \in \mathcal{L}$ such that $y(x)_i = z(x)_i + x$. But then by the triangle inequality, we get that

$$\begin{aligned}\|\rho_{-x}(f*\mu_S) - f*\mu_S\|_{\ell^p(G)} &\leq \|\frac{1}{k}\sum_{i=1}^k \rho_{-y(x)_i}(f) - f*\mu_S\|_{\ell^p(G)} \\ &\quad + \|\rho_{-x}\left(\frac{1}{k}\sum_{i=1}^k -\rho_{z(x)_i}(f) - f*\mu_S\right)\|_{\ell^p(G)}.\end{aligned}$$

However, since ρ_x is isometric on $\ell^p(G)$, we see that

$$\begin{aligned}\|\rho_{-x}(f*\mu_S) - f*\mu_S\|_{\ell^p(G)} &\leq \|\frac{1}{k}\sum_{i=1}^k \rho_{-y(x)_i}(f) - f*\mu_S\|_{\ell^p(G)} \\ &\quad + \|\frac{1}{k}\sum_{i=1}^k \rho_{-z(x)_i}(f) - f*\mu_S\|_{\ell^p(G)} \\ &\leq 2(\eta\|f\|_{\ell^p(G)}/2),\end{aligned}$$

since $z(x), y(x) \in \mathcal{L}$. □

The real strength here is the quality of the bounds for large p. For $p = 2$ a stronger result follows from Chang's theorem (at least in the case of good modeling in the sense of Green and Ruzsa [GR07]) which can actually be used to show that the set on which $f * \mu_S$ is approximately invariant is not just large, but it actually contains a large Bohr set. The techniques for proving this are Fourier analytic in nature and yield doubly exponential dependence on p if they are used to prove a version of the above result.

In the next section we shall make more careful use of the above result for large p, but here we just use the $p = 2$ case to give a set of polynomial growth following the outline in the previous section.

Proposition (Proposition 2.3)**.** *Suppose that* $|A + A| \leq K|A|$. *Then* A *is* $\exp(O(K \log K))$*-covered by a symmetric neighbourhood of the identity of size at most* $\exp(O(\log K))|A|$ *and relative polynomial growth of order* $O(K \log^3 K)$.

Proof. We put $f = 1_A$ and apply the Croot-Sisask lemma with $p = 2$, $S = T = A$, and a parameter η/m (where η and m are to be optimised later) to get a symmetric neighbourhood of the identity, X, with $|X| \geq (2K)^{-O(m^2\eta^{-2})}|A|$ such that

$$\|\rho_x(1_A * \mu_A) - 1_A * \mu_A\|^2_{\ell^2(G)} \leq \eta^2 m^{-2}|A| \text{ for all } x \in X.$$

It follows by the triangle inequality that

$$\|\rho_x(1_A * \mu_A) - 1_A * \mu_A\|^2_{\ell^2(G)} \leq \eta^2|A| \text{ for all } x \in mX,$$

and then multiplying out the ℓ^2-norm, we see that

$$2\|1_A * \mu_A\|_{\ell^2(G)}^2 - 2\langle \rho_x(1_A * \mu_A), 1_A * \mu_A\rangle_{\ell^2(G)} \leq \eta^2|A|.$$

Of course by the Cauchy-Schwarz inequality we have that

$$\|1_A * \mu_A\|_{\ell^2(G)}^2 \geq \frac{1}{|A+A|}\|1_A * \mu_A\|_{\ell^1(G)}^2 \geq |A|/K,$$

thus if we set $\eta^2 = 1/K$, we get

$$\langle \rho_x(1_A * \mu_A), 1_A * \mu_A\rangle_{\ell^2(G)} \geq |A|/2K.$$

It follows that $x \in 2A-2A$, and so $mX \subset 2A-2A$. Now by Plünnecke's inequality we have that $|(3l+1)(2A-2A)| \leq K^{4(3l+1)}|A|$ and so

$$|(3ml+1)X| \leq |(3l+1)mX| \leq (2K)^{O(l+m^2K)}|X|.$$

We put $l = m^2K + O(1)$ and write $k := ml = m^3K + O(m)$ so that

$$|(3k+1)X| \leq (2K)^{O(m^2K)}|X| = \exp(O(km^{-1}\log K))|X|.$$

We can then pick $m = O(\log K)$ such that the right-hand side is strictly less than $2^k|X|$ and hence $|(3k+1)X| < 2^k|X|$. Thus by Corollary 5.3 we have that X has relative polynomial growth of order $O(k) = O(K\log^3 K)$.

On the other hand, since X is symmetric we have $X - X \subset 2A - 2A$ and so $|X-X| \leq K^4|A|$ by the Plünnecke-Ruzsa inequalities, but also $X + A \subset 3A - 2A$. Of course with these choices $|X| \geq \exp(-O(K\log K))|A|$ and hence $|X+A| \leq \exp(O(K\log K))|X|$ by the Plünnecke-Ruzsa inequalities. With this information Lemma 5.4 completes the proof. □

This result gives bounds of roughly the same order as those of Green and Ruzsa [GR07], and more or less represents the state of the art prior to Schoen's work [Sch11].

8. A weak Bogolyubov-Ruzsa-type lemma with strong bounds

This section contains most of the newest material, and we shall start with a proof of an asymmetric weak Bogolyubov-Ruzsa-type lemma with good bounds in line with the aims of §6.

Before diving in, it is worth making a few motivating remarks. Our starting point is the argument at the end of the last section (the proof of Proposition 2.3). The weakness there was that we had to apply the Croot-Sisask lemma with a very small choice of η. This was because we have the lower bound

$$\|1_A * \mu_A\|_{\ell^2(G)}^2 \geq |A|/K,$$

which is small when compared with the trivial upper bound of $|A|$. We should like something somewhat larger, but as it is the lower bound may well be nearly this small. In [Sch11] Schoen addressed this problem by proving the following important combinatorial lemma.

Lemma 8.1 ([Sch11, Lemma 3])**.** *Suppose that $|A+A| \leq K|A|$ and $\epsilon \in (0,1]$. Then there are sets $X \subset A-A$ and $Y \subset A+A$ such that $|X| \geq \exp(-O(2^{\epsilon^{-1}}\log K))|A|$ and $|Y| \geq |A|$ such that*

$$\|1_Y * \mu_X\|_{\ell^2(G)}^2 \geq K^{-2\epsilon}|Y|.$$

The proof of this is a beautiful induction using an observation of Katz and Koester [KK10], which we shall not, unfortunately, have time to pursue here.

Given this lemma we proceed along the lines of the proof of Proposition 2.3, but using the sets Y and X given by the lemma instead of A which yields the following proposition.

Proposition 8.2. *Suppose that $|A+A| \leq K|A|$. Then A is $\exp(\exp(O(\sqrt{\log K})))$-covered by a symmetric neighbourhood of the identity of size at most $\exp(O(\log K))|A|$ and relative polynomial growth of order $\exp(O(\sqrt{\log K}))$.*

Our approach here is somewhat different, and instead of taking the inner product of $1_A * 1_A$ (or $1_Y * 1_X$) with itself, we take a different function following López and Ross [LR75]:

$$\langle 1_A * \mu_A, 1_{A+A}\rangle = |A|.$$

Given the above identity, we should like to analyse $1_{A+A} * \mu_A$ using the Croot-Sisask lemma; we do this now in the more convenient case of symmetric sets although the argument is not essentially different.

Proposition 8.3. *Suppose that $S \subset G$ is symmetric and $|S+S| \leq K|S|$, T has $|S+T| \leq L|S|$, and $m \in \mathbb{N}$ is a parameter. Then there is a symmetric neighbourhood of the identity X with*

$$|X| \geq \exp(-O(m^2 \log K \log L))|T| \text{ and } mX \subset 4S.$$

Proof. We put $f = 1_{S+S}$ and apply the Croot-Sisask lemma with a parameter η (to be optimised later) to get a symmetric neighbourhood of the identity X with $|X| \geq (2L)^{-O(\eta^{-2}p)}|T|$ such that

$$\|\rho_x(1_{S+S} * \mu_S) - 1_{S+S} * \mu_S\|_{\ell^p(G)} \leq \eta \|1_{S+S}\|_{\ell^p(G)} \text{ for all } x \in X.$$

It follows by the triangle inequality that

$$\|\rho_x(1_{S+S} * \mu_S) - 1_{S+S} * \mu_S\|_{\ell^p(G)} \leq \eta m \|1_{S+S}\|_{\ell^p(G)} \text{ for all } x \in mX.$$

Taking an inner product with μ_S, we see that

$$|\langle \rho_x(1_{S+S} * \mu_S), \mu_S\rangle - \langle 1_{S+S} * \mu_S, \mu_S\rangle| \leq \eta m \|1_{S+S}\|_{\ell^p(G)} \|\mu_S\|_{\ell^{p'}(G)},$$

where p' is the conjugate exponent to p. Now

$$\langle 1_{S+S} * \mu_S, \mu_S\rangle = \langle 1_{S+S}, \mu_S * \mu_S\rangle = 1,$$

since S is symmetric and $\operatorname{supp} \mu_S * \mu_S \subset S+S$. Thus

$$|\mu_S * 1_{S+S} * \mu_S(x) - 1| \leq \eta m \|1_{S+S}\|_{\ell^p(G)} \|\mu_S\|_{\ell^{p'}(G)} \leq \eta m K^{1/p}.$$

We take $p = 2 + \log K$, and then $\eta = \Omega(m^{-1})$ such that the term on the right is at most $1/2$ to get the desired conclusion. □

As a consequence of this we already get the following poly-logarithmic bounds.

Proposition (Proposition 2.4)**.** *Suppose that $|A+A| \leq K|A|$. Then A is $\exp(O(\log^4 K))$-covered by a symmetric neighbourhood of the identity of size at most $\exp(O(\log K))|A|$ and relative polynomial growth of order $O(\log^4 K)$.*

Proof. We apply the previous result with $T = S = A - A$ and a parameter $m \in \mathbb{N}$ to be optimised later to get a symmetric neighbourhood of the identity X with

$$|X| \geq \exp(-O(m^2 \log^2 K))|A - A| \text{ and } mX \subset 4(A - A).$$

Given $l \in \mathbb{N}$ also to be optimised later, by the Plünnecke's inequality we have that

$$|(3ml + 1)X| \leq |(3l + 1)4(A - A)| \leq K^{O(l)}|A - A| \leq K^{O(l)} \exp(O(m^2 \log^2 K))|X|.$$

We now put $l = m^2 \log K + O(1)$ and write $k := ml = m^3 \log K + O(m)$ so that we have

$$|(3k + 1)X| \leq \exp(O(m^2 \log^2 K))|X| = \exp(O(km^{-1} \log K))|X|.$$

We can then pick $m = O(\log K)$ such that the right-hand side is strictly less than $2^k|X|$ and hence $|(3k + 1)X| < 2^k|X|$. Thus by Corollary 5.3 we have that X has relative polynomial growth of order $O(k) = O(\log^4 K)$.

On the other hand we have $X - X \subset 4A - 4A$ and so $|X - X| \leq K^8|A|$ by the Plünnecke-Ruzsa inequalities, but also $X + A \subset 5A - 4A$. Hence $|X + A| \leq \exp(O(\log^4 K))|X|$ by the Plünnecke-Ruzsa inequalities. With this information Lemma 5.4 completes the proof. □

We saw Proposition 8.3 with $S = T$ is already rather powerful, but Konyagin introduced a rather nice bootstrapping technique whereby the result is first applied iteratively to reduce L to $O(1)$. To do this we first note the following corollary of Proposition 8.3.

Corollary 8.4. *Suppose that $S \subset G$ is a symmetric neighbourhood of the identity and $|S + S| \leq K|S|$, T is a symmetric neighbourhood of the identity with $|S + T| \leq L|S|$, and $D \geq 1$ is a parameter. Then there is some symmetric neighbourhood of the identity, T', such that*

$$|T'| \geq \exp(-O(D^2 \log L \log K))|T|,$$

and a symmetric neighbourhood of the identity S' with $S \subset S' \subset 5S$ and $|S' + T'| \leq K^{1/D}|S'|$.

Proof. Let k be a natural number to be optimised later, and apply Proposition 8.3 to get a symmetric neighbourhood of the identity, T', such that

$$|T'| \geq \exp(-O(k^2 \log L \log K))|T| \text{ and } 4S \supset kT'.$$

It follows that $|S + kT'| \leq |5S| \leq K^5|S|$ by Plünnecke's inequality. Thus by the pigeon-hole principle there is some $l \in \{0, \dots, k - 1\}$ such that

$$|(S + lT') + T'| \leq K^{5/k}|S + lT'|.$$

Of course we can pick $k = O(D)$ such that $K^{5/k} \leq K^{1/D}$ and so putting $S' := S + lT'$ the corollary is proved. □

The pigeon-holing trick was developed by Tao in [Tao10] to establish a Freĭman-type result in the non-Abelian setting, but it has since found use in the Abelian setting.

We are now in a position to apply the above corollary iteratively.

Proposition 8.5. *Suppose that $|A + A| \leq K|A|$. Then there is some symmetric neighbourhood of the identity, T, a natural number $r = O(\log \log K)^{O(1)}$, and a symmetric neighbourhood of the identity S with $A - A \subset S \subset r(A - A)$ and*

$$|S + T| = O(|S|) \text{ and } |T| \geq \exp(-O(\log \log K)^{O(1)} \log^3 K)|S|.$$

Proof. We define two sequences of sets $(S_i)_i$ and $(T_i)_i$, and a sequence of reals $(L_i)_i$ such that S_i and T_i are symmetric neighbourhoods of the identity, and

$$A-A \subset S_i \subset 5^i(A-A) \text{ and } |S_i+T_i| \leq L_i|S_i|,$$

where $L_i = \exp(4(\log 2K)^{2^{-i}})$. To start with, we put $S_0 := A-A$ and $T_0 := A-A$, which satisfies the requirements by the Plünnecke-Ruzsa inequalities. At stage i we note that

$$|S_i+S_i| \leq |2\cdot 5^i(A-A)| \leq K^{4\cdot 5^i}|A-A| \leq K^{4\cdot 5^i}|S_i|$$

by the Plünnecke-Ruzsa inequalities. We apply Corollary 8.4 to the sets S_i and T_i with parameter $D_i := 1+(\log(2K^{4\cdot 5^i}))^{1-2^{-(i+1)}}$ to get symmetric neighbourhoods of the identity S_{i+1} and T_{i+1}, with

$$|T_{i+1}| \quad \geq \quad \exp(-O(D_i^2(\log L_i)(\log K^{4\cdot 5^i})))|T_i| \geq \exp(-O(\exp(O(i))\log^3 K))|T_i|,$$

$$A-A \subset S_i \subset S_{i+1} \subset 5S_i \subset 5^{i+1}(A-A),$$

and

$$|S_{i+1}+T_{i+1}| \leq \exp(4(\log 2K)^{2^{-(i+1)}})|S_{i+1}|.$$

We terminate the iteration when $2^{i+O(1)} = \log 2\log 2K$ and find that the result is proved with $S=S_i$ and $T=T_i$. □

Finally, we have the strongest result of the section and the driving ingredient in this survey:

Proposition (Proposition 2.5)**.** *Suppose that* $|A+A| \leq K|A|$. *Then* A *is* $\exp(O(\log^{3+o(1)} K))$*-covered by a symmetric neighbourhood of the identity of size at most* $\exp(O(\log^{1+o(1)} K))|A|$ *and relative polynomial growth of order* $O(\log^{3+o(1)} K)$.

Proof. We apply Proposition 8.5 to the set A to get symmetric neighbourhoods of the identity S and T, and a natural number $r = O(\log^{o(1)} K)$ such that

$$A-A \subset S \subset r(A-A), |S+T| = O(|S|) \text{ and } |T| \geq \exp(-O(\log^{3+o(1)} K))|S|.$$

Now, by Proposition 8.3 applied to the sets S and T with a parameter m to be optimised later, we get a symmetric neighbourhood of the identity X with

$$|X| \geq \exp(-O(m^2\log^{1+o(1)} K))|T| \text{ and } mX \subset 4S.$$

Given $l \in \mathbb{N}$ also to be optimised later, by Plünnecke's inequality we have that

$$\begin{aligned} |(3ml+1)X| &\leq |(3l+1)4S| \leq K^{O(l)}|S| \\ &\leq K^{O(l)}\exp(O(m^2\log^{1+o(1)} K)+O(\log^{3+o(1)} K))|X|. \end{aligned}$$

We now put $l = m^2\log^{o(1)} K$ and write $k := ml = m^3\log^{o(1)} K$, so that we have

$$|(3k+1)X| \leq \exp(O(k(m^{-1}\log^{1+o(1)} K + m^{-3}\log^{3+o(1)} K))|X|.$$

We can then pick $m = \log^{1+o(1)} K$ such that the right-hand side is strictly less than $2^k|X|$ and hence $|(3k+1)X| < 2^k|X|$. Thus by Corollary 5.3 we have that X has relative polynomial growth of order $O(k) = O(\log^{3+o(1)} K)$.

On the other hand we have $X-X \subset 4S \subset 4r(A-A)$, and so, by the Plünnecke-Ruzsa inequalities, we have

$$|X-X| \leq |4rA-4rA| \leq K^{8r}|A| \leq \exp(O(\log^{1+o(1)} K))|A|.$$

This set inclusion (and the fact that $0_G \in X$) also tells us that $X + A \subset 4r(A-A)+A$. Hence, by the Plünnecke-Ruzsa inequalities again, and the fact that $|A| \leq |S| \leq \exp(O(\log^{3+o(1)} K))|X|$, we have

$$|X+A| \leq K^{8r+1}|A| = \exp(O(\log^{1+o(1)} K))|A| \leq \exp(O(\log^{3+o(1)} K))|X|.$$

With this information Lemma 5.4 completes the proof. □

It may be worth saying that all the $\log^{o(1)} K$ terms in the above proposition can be replaced by $(\log\log K)^{O(1)}$ terms if desired.

9. From relative polynomial growth to convex coset progressions

Our aim in the next few sections it to prove Theorem 2.7 which we restate now for convenience.

Theorem (Theorem 2.7)**.** *Suppose that X has relative polynomial growth of order d. Then there is a centred convex coset progression M such that*

$$X - X \subset M, \dim M = O(d\log^2 d) \text{ and } |M| \leq \exp(O(d\log^2 d))|X|.$$

We shall make considerable use of harmonic analysis on discrete groups to do this and so it will be useful to record some definitions. The classic reference is Rudin [Rud90], although the reader will be equally well served by Tao and Vu [TV06].

We have already introduced convolution, and the Fourier transform is defined to diagonalise the operators induced by convolution, so we are already have quite a bit of what we need.

Given G (discrete) we write $\widehat{G}$ for the set of homomorphisms $\gamma : G \to S^1$ where $S^1 := \{z \in \mathbb{C} : |z| = 1\}$. These homomorphisms are called **characters** and the set $\widehat{G}$ naturally supports the structure of a topological group, in particular a compact Abelian group under pointwise multiplication of characters, called the **dual group** of G.

The dual group is naturally endowed with a translation invariant probability measure called the Haar probability measure, and we are now in a position to define the Fourier transform. Given $f \in \ell^1(G)$, we define the **Fourier transform** of f to be the function $\widehat{f} \in L^\infty(\widehat{G})$ determined by

$$\widehat{f}(\gamma) := \sum_{x\in G} f(x)\overline{\gamma(x)} \text{ for all } \gamma \in \widehat{G}.$$

This has the property that $\widehat{f*g} = \widehat{f}\cdot\widehat{g}$. More than this we have Plancherel's formula which tells us that

$$\langle f, g\rangle_{\ell^2(G)} = \langle \widehat{f}, \widehat{g}\rangle_{L^2(\widehat{G})} \text{ for all } f, g \in \ell^2(G).$$

We have already indicated that $\widehat{G}$ has a natural topology, and in fact if G is small enough this topology is induced by a metric. There are then a range of metrics which define different topologies of $\widehat{G}$ reflecting the subgroup structure of $\widehat{G}$. These can be defined by bases of what are called Bohr sets.

Given a neighbourhood Γ of characters on G and a parameter $\delta \in (0,2]$, we define the **Bohr set** with **frequency set** Γ and **width** δ to be the set

$$\mathrm{Bohr}(\Gamma,\delta) := \{x \in G : |\gamma(x) - 1| \leq \delta \text{ for all } \gamma \in \Gamma\}.$$

One rather useful property of Bohr sets which we use repeatedly is the fact that they are balls in a pseudo-metric. What we mean by this is that for a character $\gamma \in \widehat{G}$, we have the very useful triangle inequality

$$|1-\gamma(x+y)| = |1-\gamma(x)+(1-\gamma(y))\gamma(x)| \leq |1-\gamma(x)|+|1-\gamma(y)|$$

for all $x, y \in G$.

The first ingredient in proving Theorem 2.7 is to show that in some sense the topology determined by a set X is roughly the same as that determined by certain Bohr sets.

Proposition 9.1. *Suppose that X has relative polynomial growth of order d. Then there is a neighbourhood of characters Γ and a natural number $k = O(d \log^2 d)$ such that*

$$X - X \subset \mathrm{Bohr}(\Gamma, 1/(4(3k+1))) \text{ and } |\mathrm{Bohr}(\Gamma, 1/2)| < 2^k |X|.$$

Now we shall see later that Bohr sets are already convex progressions, and if they satisfy a certain growth condition of the form used in Chang's covering lemma, then they turn out to be low dimensional. In particular we shall show the following which combines with the previous result to yield Theorem 2.7.

Proposition 9.2. *Suppose that $\mathrm{Bohr}(\Gamma, \delta)$ is a finite Bohr set and $k \in \mathbb{N}$ is such that*

$$|\mathrm{Bohr}(\Gamma, (3k+1)\delta)| < 2^k |\mathrm{Bohr}(\Gamma, \delta)| \text{ for some } \delta < 1/(4(3k+1)).$$

Then $\mathrm{Bohr}(\Gamma, \delta)$ is an (at most) k-dimensional centred convex coset progression.

10. Relative polynomial growth and Bohr sets

In this section we show how to pass from sets with relative polynomial growth to a Bohr set which (effectively) has polynomial growth of relatively low order. Shortly we shall see that Bohr sets are convex coset progressions (provided the width parameter is sufficiently small), but for now we think of them as a sort of "approximate annihilator".

To find an appropriate Bohr set, we shall need to examine the (very) large spectrum of a finite set A, which is defined to be the set

$$\mathrm{LSpec}(A, \epsilon) := \{\gamma \in \widehat{G} : \|1-\gamma\|_{L^2(\mu_A * \mu_{-A})} \leq \epsilon\}.$$

(Note immediately that $\mathrm{LSpec}(A, \epsilon)$ is a neighbourhood since A is finite.) The definition of LSpec we have given takes the form it does for ease of use of the triangle inequality: if $\gamma \in \mathrm{LSpec}(A, \epsilon)$ and $\gamma' \in \mathrm{LSpec}(A, \epsilon')$, then $\gamma + \gamma' \in \mathrm{LSpec}(A, \epsilon + \epsilon')$ by the triangle inequality:

$$\begin{aligned} \|1-\gamma\gamma'\|_{L^2(\mu_A * \mu_{-A})} &= \|(1-\gamma')+(1-\gamma)\gamma'\|_{L^2(\mu_A * \mu_{-A})} \\ &\leq \|1-\gamma'\|_{L^2(\mu_A * \mu_{-A})} + \|(1-\gamma)\gamma'\|_{L^2(\mu_A * \mu_{-A})} \\ &= \|1-\gamma\|_{L^2(\mu_A * \mu_{-A})} + \|1-\gamma'\|_{L^2(\mu_A * \mu_{-A})}. \end{aligned}$$

On the other hand to connect the definition to the idea that LSpec should represent the large spectrum, we have the following useful identity,

$$\|1-\gamma\|^2_{L^2(\mu_A * \mu_{-A})} = 2(1-|\widehat{\mu_A}(\gamma)|^2),$$

so that

$$\|1-\gamma\|_{L^2(\mu_A * \mu_{-A})} \leq \epsilon \text{ if and only if } |\widehat{\mu_A}(\gamma)| \geq \sqrt{1-\epsilon^2/2}.$$

This fact will be used extensively in the remainder of this section.

We have two key tools for establishing our main proposition (Proposition 9.1). The first of these uses an approximation developed by Schoen in [Sch03] which was imported into this context by Green and Ruzsa in [GR07].

Proposition 10.1. *Suppose that X has relative polynomial growth of order d. Then*

$$|\operatorname{Bohr}(\operatorname{LSpec}(X,\epsilon),1/2)| \leq \exp(O(d\log\epsilon^{-1}d))|X|.$$

Proof. By Plancherel's theorem and the Cauchy-Schwarz inequality, we have

$$\int |\widehat{1_X}(\gamma)|^{2k}d\gamma = \|1_X^{(k)}\|_{\ell^2(G)}^2 \geq \frac{\|1_X^{(k)}\|_{\ell^1(G)}^2}{|\operatorname{supp} 1_X^{(k)}|} = \frac{|X|^{2k}}{|kX|}. \tag{10.1}$$

We shall show that most of this mass is supported on the set of characters where the Fourier transform of 1_X is very large. In particular note that

$$\begin{aligned}\int_{\operatorname{LSpec}(X,\epsilon)^c} |\widehat{1_X}(\gamma)|^{2k}d\gamma &\leq (\sqrt{1-\epsilon^2/2}|X|)^{2k-2}\int |\widehat{1_X}|^2 d\gamma\\ &= (1-\epsilon^2/2)^{k-1}|X|^{2k-1},\end{aligned}$$

by Parseval's theorem.

Since X has polynomial growth of order d, we have that $|kX|\leq k^d|X|$ for $k\geq 1$, so there is a positive integer k with $k=O(\epsilon^{-2}d\log\epsilon^{-1}d)$ and

$$(1-\epsilon^2/2)^{k-1}\leq 1/2k^d \leq |X|/2|kX|,$$

whence

$$\int_{\operatorname{LSpec}(X,\epsilon)^c} |\widehat{1_X}(\gamma)|^{2k}d\gamma \leq \frac{|X|^{2k}}{2|kX|}.$$

Thus, by (10.1) we have

$$\int_{\operatorname{LSpec}(X,\epsilon)} |\widehat{1_X}(\gamma)|^{2k}d\gamma \geq \frac{|X|^{2k}}{2|kX|}.$$

Now, let B be a finite subset of $\operatorname{Bohr}(\operatorname{LSpec}(X,\epsilon),1/2)$. Integrating, we get that $|1-\widehat{\mu_B}(\gamma)|\leq 1/2$ for any $\gamma\in\operatorname{LSpec}(X,\epsilon)$ and it follows by the triangle inequality that $|\widehat{\mu_B}(\gamma)|\geq 1/2$. Consequently,

$$\int |\widehat{1_X}(\gamma)|^{2k}|\widehat{\mu_B}(\gamma)|^2 d\gamma \geq 2^{-2}\int_{\operatorname{LSpec}(X,\epsilon)} |\widehat{1_X}(\gamma)|^{2k}d\gamma \geq \frac{|X|^{2k}}{2^3|kX|}.$$

On the other hand

$$\begin{aligned}\int |\widehat{1_X}(\gamma)|^{2k}|\widehat{\mu_B}(\gamma)|^2 d\gamma &\leq |X|^{2k-2}\|1_X*\mu_B\|_{\ell^2(G)}^2\\ &\leq |X|^{2k-2}\|1_X*\mu_B\|_{\ell^1(G)}\|1_X*\mu_B\|_{\ell^\infty(G)}\end{aligned}$$

by the Hausdorff-Young inequality, Parseval's theorem and then Hölder's inequality. Since $\|1_X*\mu_B\|_{\ell^1(G)}=|X|$, we conclude that

$$\frac{|X|}{2^3|kX|}\leq \|1_X*\mu_B\|_{\ell^\infty(G)}\leq \frac{|X|}{|B|}.$$

This gives the desired upper bound, but on B rather than $\operatorname{Bohr}(\operatorname{LSpec}(X,\epsilon),1/2)$. The result follows since B was an arbitrary finite subset of $\operatorname{Bohr}(\operatorname{LSpec}(X,\epsilon),1/2)$. □

Our second key tool is yet another of the developments of Green and Ruzsa from [GR07]. It is only slightly more general than [TV06, Proposition 4.39].

Proposition 10.2. *Suppose that* $|X+S| \leq K|S|$ *and* $\epsilon \in (0,1]$ *is a parameter. Then*

$$X - X \subset \mathrm{Bohr}(\mathrm{LSpec}(X+S,\epsilon), O(\epsilon\sqrt{K})).$$

Proof. Write $\delta = 1-\sqrt{1-\epsilon^2/2}$, and suppose that $\gamma \in \mathrm{LSpec}(X+S,\epsilon)$. Then there is a phase $\omega \in S^1$ such that

$$\sum_{x\in G} 1_{X+S}(x)\omega\gamma(x) = \omega\widehat{1_{X+S}}(\gamma) = |\widehat{1_{X+S}}(\gamma)|.$$

Since the right-hand side is real, we conclude that

$$\sum_{x\in G} 1_{X+S}(x)\,\mathrm{Re}\,\omega\gamma(x) = \mathrm{Re}\sum_{x\in G} 1_{X+S}(x)\omega\gamma(x) = |\widehat{1_{X+S}}(\gamma)| \geq (1-\delta)|X+S|.$$

It follows that

$$\sum_{x\in G} 1_{X+S}(x)|1-\omega\gamma(x)|^2 = 2\sum_{x\in G} 1_{X+S}(x)(1-\mathrm{Re}\,\omega\gamma(x)) \leq 2\delta|X+S|.$$

If $y_0, y_1 \in X$, then

$$\sum_{x\in G} 1_S(x)|1-\omega\gamma(y_i)\gamma(x)|^2 \leq \sum_{x\in G} 1_{X+S}(x)|1-\omega\gamma(x)|^2 \leq 2\delta|X+S|.$$

The 2-variable Cauchy-Schwarz inequality applied to $1-\omega\gamma(y_0)\gamma(x)$ and $1-\omega\gamma(y_1)\gamma(x)$ tells us that

$$\begin{aligned} |1-\gamma(y_0-y_1)|^2 &= |(1-\omega\gamma(y_0)\gamma(x)) - (1-\omega\gamma(y_1)\gamma(x))|^2 \\ &\leq 2(|1-\omega\gamma(y_0)\gamma(x)|^2 + |1-\omega\gamma(y_1)\gamma(x)|^2) \end{aligned}$$

for all $x \in G$ since $|\omega| = 1$ and $|\gamma(x)| = 1$, whence

$$|S||1-\gamma(y_0-y_1)|^2 = \sum_{x\in G} 1_S(x)|1-\gamma(y_0-y_1)|^2 \leq 2^3\delta|X+S|.$$

The result follows since $\delta = O(\epsilon^2)$. □

With these two results we are in a position to prove the main result of this section.

Proposition (Proposition 9.1)**.** *Suppose that* X *has relative polynomial growth of order* d*. Then there is a neighbourhood of characters* Γ *and a natural number* $k = O(d\log^2 d)$ *such that*

$$X - X \subset \mathrm{Bohr}(\Gamma, 1/(4(3k+1))) \text{ and } |\mathrm{Bohr}(\Gamma, 1/2)| < 2^k|X|.$$

Proof. Since X has relative polynomial growth of order d, we may apply the pigeonhole principle to pick $l = O(d\log d)$ such that $|X+lX| = O(|lX|)$. Let ϵ be a parameter to be optimised later. By Proposition 10.1 applied to the set $(l+1)X$ which has relative polynomial growth of order $O(d\log d)$, we see that for $\Gamma := \mathrm{LSpec}(X+lX,\epsilon)$ (which is closed) we have

$$|\mathrm{Bohr}(\Gamma, 1/2)| \leq \exp(O(d\log^2\epsilon^{-1}d))|X|.$$

On the other hand, by Proposition 10.2 applied to the sets X and lX we see that

$$\mathrm{Bohr}(\Gamma, O(\epsilon)) \supset X - X.$$

We now pick $k = \Omega(\epsilon^{-1})$ such that the width parameter above is at most $1/(4(3k+1))$ and the size bound is less than 2^k. This is possible with $\epsilon = \Omega(1/(d\log^2 d))$. The result is proved. □

11. Ruzsa's embedding and convex coset progressions

In his paper [Ruz94], Ruzsa developed an important embedding for relating Bohr sets and convex coset progressions. Given a set Γ of characters on G, write $B(\Gamma,\mathbb{R})$ for the vector space of bounded real-valued functions on Γ. Now, we define the map

$$\begin{aligned} R_\Gamma : G &\to B(\Gamma,\mathbb{R}) \\ x &\mapsto R_\Gamma(x) : \Gamma \to \mathbb{R}; \gamma \mapsto \frac{1}{2\pi i}\log\gamma(x), \end{aligned}$$

where the logarithm takes its principal value. (Since $|\gamma(x)| = 1$ this means that the logarithm lies in $(-\pi i, \pi i]$ and so the functions are bounded.)

The map R_Γ preserves inverses provided $\|R_\Gamma(x)\|_\infty < 1/2$, meaning that $R_\Gamma(-x) = -R_\Gamma(x)$; and furthermore we see that if

$$\|R_\Gamma(x_1)\|_\infty + \cdots + \|R_\Gamma(x_d)\|_\infty < 1/2,$$

then

$$R_\Gamma(x_1+\cdots+x_d) = R_\Gamma(x_1)+\cdots+R_\Gamma(x_d).$$

This essentially encodes the idea that R_Γ behaves like a Freĭman morphism,[4] although we shall not formalise this notion here. We use this embedding to establish the following proposition.

Proposition (Proposition 9.2)**.** *Suppose that* $\mathrm{Bohr}(\Gamma,\delta)$ *is a finite Bohr set and* $d \in \mathbb{N}$ *is such that*

$$|\mathrm{Bohr}(\Gamma,(3d+1)\delta)| < 2^d|\mathrm{Bohr}(\Gamma,\delta)| \text{ for some } \delta < 1/(4(3d+1)).$$

Then $\mathrm{Bohr}(\Gamma,\delta)$ *is an (at most) d-dimensional centred convex coset progression.*

Proof. We shall prove that if $L := \bigcap\{\ker\gamma : \gamma\in\Gamma\}$ is trivial, then $\mathrm{Bohr}(\Gamma,\delta)$ is a d-dimensional centred convex progression. The result then follows from this by quotienting out by L (which does not impact the hypotheses of the proposition) to get a homomorphism $\phi : \mathbb{Z}^d \to G/L$ and a symmetric convex body $Q \subset \mathbb{R}^d$ such that $\mathrm{Bohr}(\Gamma,\delta)/L = \phi(Q\cap\mathbb{Z}^d)$.

Let $e_1,\dots,e_d$ be the standard set of generators for $\mathbb{Z}^d$, and for each $i\in\{1,\dots,d\}$ let $h_i \in G$ be a representative of $\phi(e_i)$. Since $\mathbb{Z}^d$ is free define $\tilde\phi : \mathbb{Z}^d \to G$ by extension from its value at the generators $\tilde\phi(e_i) := h_i$, and note that

$$\mathrm{Bohr}(\Gamma,\delta) = \bigcup \mathrm{Bohr}(\Gamma,\delta)/L = \bigcup \phi(Q\cap\mathbb{Z}^d) = \tilde\phi(Q\cap\mathbb{Z}^d) + L.$$

The result follows.

For notational convenience we write $B_\eta := \mathrm{Bohr}(\Gamma,\eta)$ for any $\eta\in(0,2]$. To start with note that if $x\in B_\eta$, then

$$\|R_\Gamma(x)\|_\infty \le \frac{1}{2\pi}\arccos(1-\eta^2/2) \le 2\eta.$$

Since $2(3d+1)\delta < 1/2$, we have that if $x_1,\dots,x_{3d+1}\in B_\delta$, then

$$R_\Gamma(x_1+\cdots+x_{3d+1}) = R_\Gamma(x_1)+\cdots+R_\Gamma(x_{3d+1}). \tag{11.1}$$

[4] We direct the unfamiliar reader to [TV06, Chapter 5.3].

By hypothesis we then have that

$$|(3d+1)R_\Gamma(B_\delta)| = |R_\Gamma((3d+1)B_\delta)| \leq |(3d+1)B_\delta| \leq |B_{(3d+1)\delta}| < 2^d|B_\delta|.$$

Now $|B_\delta| = |R_\Gamma(B_\delta)|$ since R_Γ is injective on B_δ. To see this, note that if $x, y \in B_\delta$ have $R_\Gamma(x) = R_\Gamma(y)$, then $R_\Gamma(x-y) = 0$ by (11.1) and the fact that R_Γ preserves inverses on B_δ. It then follows that $\gamma(x-y) = 1$ for all $\gamma \in \Gamma$, and since L is trivial, we conclude that $x = y$.

In light of all this we have that $|3dR_\Gamma(B_\delta)+R_\Gamma(B_\delta)| < 2^d|R_\Gamma(B_\delta)|$, and so by the variant of Chang's covering lemma in Lemma 5.2 applied to the sets $3R_\Gamma(B_\delta)$ and $R_\Gamma(B_\delta)$ (both of which are symmetric neighbourhoods since R_Γ preserves inverses and the identity, and B_δ is symmetric), we get a set $X \subset 3R_\Gamma(B_\delta)$ with $|X| < d$ such that

$$3R_\Gamma(B_\delta) \subset \mathrm{Span}(X) + 2R_\Gamma(B_\delta) \subset \langle X\rangle + 2R_\Gamma(B_\delta).$$

Here, of course, $\langle X\rangle$ denotes the group generated by X. It follows that for all $n \in \mathbb{N}$ we have

$$(n+2)R_\Gamma(B_\delta) \subset \langle X\rangle + 2R_\Gamma(B_\delta).$$

Now, for each $v \in R_\Gamma(B_\delta)$ and $n \in \mathbb{N}$ there is some $v_n \in 2R_\Gamma(B_\delta)$ such that $nv \in \langle X\rangle + v_n$. However, since $2R_\Gamma(B_\delta)$ is finite, it follows that there are distinct natural numbers $n \neq m$ such that $v_n = v_m$, whence

$$(n-m)v = nv - mv \in (\langle X\rangle + v_n) - (\langle X\rangle + v_m) = \langle X\rangle.$$

Thus to every $v \in R_\Gamma(B_\delta)$ there is some natural number l_v such that $l_v v \in \langle X\rangle$. Let L be the lowest common multiple of all the natural numbers $(l_v)_{v\in R_\Gamma(B_\delta)}$ so that $Lv \in \langle X\rangle$ for all $v \in R_\Gamma(B_\delta)$. It follows that $v \in \langle x/L : x \in X\rangle$ and so $R_\Gamma(B_\delta)$ generates a lattice Λ in $B(\Gamma, \mathbb{R})$ of dimension $k \leq |X| < d$.

Let $v_1, \dots, v_k$ be a basis for Λ and for each $j \in \{1, \dots, k\}$ write $v_j = \sum_{x\in B_\delta} z_{j,x}R_\Gamma(x)$ for some integers $(z_{j,x})_{x\in B_\delta}$. We now put $h_j := \sum_{x\in B_\delta} z_{j,x}x$ and define a homomorphism

$$\phi : \mathbb{Z}^k \to G; (n_1, \dots, n_k) \mapsto n_1h_1 + \cdots + n_kh_k.$$

Finally, write V for the subspace of $B(\Gamma, \mathbb{R})$ generated by X and $\psi : V \to \mathbb{R}^k$ for the change of basis taking v_i to the canonical basis vector e_i of $\mathbb{R}^k$, and let Q be the cube in $B(\Gamma, \mathbb{R})$ centred at the origin and with side length 2δ. The set $\psi(Q\cap V)$ is a symmetric convex body in $\mathbb{R}^k$, and it remains to check that $\phi(\psi(Q\cap V)\cap\mathbb{Z}^k) = B_\delta$.

If $x_0 \in B_\delta$, then $R_\Gamma(x_0) \in \Lambda$ and $R_\Gamma(x_0) \in Q$, and so

$$R_\Gamma(x_0) = n_1v_1 + \cdots + n_kv_k \text{ for some } n \in \psi(Q \cap V) \cap \mathbb{Z}^k.$$

Given the definition of the v_i's, we have that

$$R_\Gamma(x_0) = \sum_{j=1}^{k} n_j \sum_{x\in B_\delta} z_{j,x}R_\Gamma(x).$$

Exponentiating this pointwise (via $x \mapsto \exp(2\pi ix)$ which is a homomorphism from $B(\Gamma, \mathbb{R}) \to B(\Gamma, S^1)$) tells us that

$$\gamma(x_0) = \prod_{j=1}^{k}\left(\prod_{x\in B_\delta}\gamma(x)^{z_{j,x}}\right)^{n_j} = \gamma(\sum_{j=1}^{k} n_j \sum_{x\in B_\delta} z_{j,x}x) \text{ for all } \gamma \in \Gamma.$$

Since L is trivial, we conclude that

$$x_0 = \sum_{j=1}^{k} n_j \sum_{x\in B_\delta} z_{j,x}x = n_1h_1 + \cdots + n_kh_k.$$

It follows that $\phi(n) = x_0$, and so $x_0 \in \phi(\psi(Q \cap V) \cap \mathbb{Z}^k)$.

In the other direction suppose that $x_0 \in \phi(\psi(Q \cap V) \cap \mathbb{Z}^k)$ and $v_0 \in Q \cap \Lambda$ is such that $x_0 = \phi(\psi(v_0))$. Then $v_0 \in \Lambda$ and so

$$v_0 = n_1v_1 + \cdots + n_kv_k \text{ for some } n \in \mathbb{Z}^k,$$

and so

$$v_0 = \sum_{j=1}^{k} n_j \sum_{x\in B_\delta} z_{j,x}R_\Gamma(x).$$

We exponentiate pointwise as before to get that

$$\exp(2\pi i v_0) = \prod_{j=1}^{k}\left(\prod_{x\in B_\delta}\gamma(x)^{z_{j,x}}\right)^{n_j} = \gamma(\sum_{j=1}^{k} n_j \sum_{x\in B_\delta} z_{j,x}x) \text{ for all } \gamma \in \Gamma.$$

But $v_0 \in Q$ and so $|1 - \exp(2\pi i v_0)| \leq \delta$ for all $\gamma \in \Gamma$ and hence

$$x_0 = \phi(n) = \sum_{j=1}^{k} n_jh_j = \sum_{j=1}^{k} n_j \sum_{x\in B_\delta} z_{j,x}x \in B_\delta$$

as required. The result is proved. □

In light of the start of the proof here it might be more natural to define a centred convex coset progression to be a set of the form $\bigcup \phi(Q \cap \mathbb{Z}^d)$ where $\phi : \mathbb{Z}^d \to G/H$ is a homomorphism, $H \leq G$ and Q is a symmetry convex body in $\mathbb{R}^d$. This sort of consideration becomes more relevant as one moves to the non-Abelian setting, but this is not our concern here.

12. Concluding remarks

First we should note that Theorem 1.4 follows immediately from combining Proposition 2.5 and Theorem 2.7, and all the $\log^{o(1)} K$ terms can be replaced by $(\log\log K)^{O(1)}$ terms for those interested.

It may also be worth noting that there are really three different functions in Theorem 1.3; we really show the following.

Theorem 12.1. *Suppose that $A \subset G$ has $|A + A| \leq K|A|$. Then A is $\exp(h(K))$-covered by a $d(K)$-dimensional centred convex coset progression M of size at most $\exp(f(K))|A|$.*

The quantities $h(K)$, $d(K)$, and $f(K)$ can be traded off between each other to some extent, but there is an associated cost. The precise relationships are a little *ad hoc* because they reflect different combinations of our three main examples. Let us recall these now:

(i) *Cosets of subgroups.* Suppose that H is a finite subgroup of G and X is an H-separated set of $2K + O(1)$ points. Then letting $A := X + H$ we have $|A + A| \sim K|A|$.

(ii) *Convex progressions.* Suppose that M is a d-dimensional convex coset progression. Then we have seen that $|M+M| \leq \exp(O(d))|M|$. On the other hand if A is a cube in $\mathbb{Z}^d$ (so that all the side lengths are the same), then in fact $|A+A| \sim 2^d|A|$ so that the doubling of A really is this large.

(iii) *Subsets of subgroups.* Suppose that H is a finite subgroup of G and A is a randomly chosen subset of H, taking $x \in H$ with probability $1/K$. Then with high probability $|A| \sim |H|/K$ and $|A+A| \sim |H|$ so that $|A+A| \sim K|A|$.

Each of these suggests a lower bound on (respectively) $h(K)$, $d(K)$, and $f(K)$, but they do not all give such bounds, and there is no one example which forces lower bounds on all of them simultaneously. This is because of the previously mentioned ability to trade, which we shall now explain in a little more depth. We assume that we are given Theorem 12.1 with some functions $h(K), d(K)$, and $f(K)$.

12.2. **Reducing $h(K)$ in exchange for $d(K)$.** One can eliminate $h(K)$ entirely and replace "$\exp(h(K))$-covered by" in Theorem 12.1 by "contained in" at the expense of replacing $d(K)$ by $d(K) + \exp(h(K))$, and $f(K)$ by $2f(K)$. This is a little fiddly, but not difficult to do.

Removing the dependence on the covering number is the additional requirement which is made in traditional statements of Freĭman-type theorems; indeed, Green and Ruzsa in [GR07] actually proved the following.

Theorem 12.3 (Green-Ruzsa theorem, original version)**.** *Suppose that $|A+A| \leq K|A|$. Then A is contained in a $K^{4+o(1)}$-dimensional centred convex coset progression M of size at most $\exp(K^{4+o(1)})|A|$.*

This has been slightly improved, with the power of $4+o(1)$ being replaced by $1+o(1)$ but the reason we do not use this formulation is that the dimension bound must be at least $\Omega(K)$ — exponentially worse than in the Polynomial Freĭman-Ruzsa conjecture. This is, of course, suggested by the fact that reducing the covering number has a cost of $\exp(h(K))$ rather than $h(K)$ associated with it.

To see the difficulty directly, suppose that A is a set of $2K + O(1)$ generators of a torsion-free group. Then $|A+A| \sim K|A|$, but any convex coset progression containing A has dimension at least $2K - O(1)$.

12.4. **Reducing $d(K)$ in exchange for $f(K)$.** In general one cannot trade in all of the dimension for size, but one can if the group has bounded exponent (meaning every element has order bounded by an absolute constant). Then one may reduce $d(K)$ to 0 at the expense of replacing $f(K)$ by $\exp(f(K) + O(d(K)))$. In Theorem 1.4 this gives the following result.

Theorem 12.5. *Suppose that G is a group of bounded exponent and $A \subset G$ has $|A+A| \leq K|A|$. Then A is $\exp(O(\log^{3+o(1)} K))$-covered by a subgroup M of size at most $\exp(O(\log^{3+o(1)} K))|A|$.*

Conjecturally, one can do much better, and here the Polynomial Freĭman-Ruzsa conjecture becomes the following, which was one of its original motivations.

Conjecture 12.6 (Marton's conjecture)**.** *Suppose that G is a group of bounded exponent and $A \subset G$ has $|A+A| \leq K|A|$. Then A is $\exp(O(\log K))$-covered by a subgroup M of size at most $\exp(O(\log K))|A|$.*

12.7. **Reducing $d(K)$ in exchange for $h(K)$.** We just saw how to trade in dimension for size in the case where the group has bounded exponent. In general one cannot trade in all of the dimension for size but Green and Tao in [GT06] show (in torsion-free groups) how to reduce the dimension of the progression to $O(\log K)$ while incurring an exponential cost in the covering number so that $h(K) = \Theta(K)$. (They get a larger polynomial in K in their work but this can be removed given the recent stronger bounds in Freĭman's theorem.)

The paper [GT06] is, in general, rather useful as a source of tools for giving the lower bounds on the order of relative polynomial growth of sets, and we direct the reader interested in the more precise relationships between $h(K), d(K)$, and $f(K)$ there.

As a final remark it is worth saying that convex progressions may not be quite the right notion to deal with, and one might like to ask for a convex progression of a particular type. There is some discussion of this in [GT06], but we shall not pursue this here except to remark that Freĭman's theorem is usually stated using generalised arithmetic progressions which are a special type of (translate of a centred) convex progression defined by a cube. Specifically a set M is a **generalised arithmetic progression** if

$$M = \{x_0 + z_1x_1 + \cdots + z_dx_d : |l_i| \leq L_i\}$$

for some natural numbers $L_1, \ldots, L_d$ and elements $x_0, \ldots, x_d \in G$. If we define a homomorphism

$$\phi : \mathbb{Z}^d \to G; (z_1, \ldots, z_d) \mapsto z_1x_1 + \cdots + z_dx_d$$

and a convex set $Q := \prod_{i=1}^d [-L_i, L_i]$, then $M = x_0 + \phi(\mathbb{Z}^d \cap Q)$. A **coset progression** (as defined by Green and Ruzsa in [GR07]) is then a set of the form $M + H$ where $H \leq G$ and M is a generalised arithmetic progression in G. Proving the results of this paper for coset progressions instead of convex coset progressions is not conceptually harder, but does seem to involve some additional technical difficulties.

Generalised arithmetic progressions have been studied in their own right and there are various questions concerning whether they are proper or not, meaning whether ϕ is injective on $Q \cap \mathbb{Z}^d$. Bilu in [Bil99] has a nice discussion of this (see also [TV06, §3.1]).

13. Applications

As indicated in the introduction there are numerous applications of Freĭman's theorem, and for completeness we shall discuss a few of these here. These are chosen mainly because they do not require too much additional material to develop rather than because they are necessarily the most exciting. This section is of a much more sketchy nature than the rest of the paper: it is intended to indicate directions in which one can take the results discussed in this paper; it is not intended to cover them in detail, and the interested reader is referred to the papers indicated in each subsection below for more comprehensive discussions.

One thing it is worth remembering is that while Freĭman's theorem is very attractive at a qualitative level, in applications one can often squeeze a little more juice out of the situation by using the methods of this paper rather than the results. In particular the combinatorial arguments on their own are often enough for what one hopes to do. In this regard it should be mentioned that there are many direct applications of the techniques of Croot and Sisask in [CS10] and [CŁS11],

which can also be proved using Freĭman's theorem but which only really require the Croot-Sisask lemma.

A second remark is due with regard to Roth's theorem. The reader may be hoping for a discussion of bounds in Roth's theorem in this survey, but this is not really the place for that. In particular, while the results of Proposition 2.4 are relevant to that work, nothing else from the paper is, and a discussion of the combinatorial techniques of Katz and Koester [KK10] and the regular Bohr set technology of Bourgain [Bou99] would be required.

The U^3-inverse theorem. Gowers' work [Gow98] marks the start of an explosion of applications of Freĭman's theorem after he made the crucial observation that it can be combined with the Balog-Szemerédi lemma [BS94]. Gowers used Freĭman's theorem to improve the bounds in Szemerédi's theorem for arithmetic progressions of length four and a little after that Green and Tao expressed Gowers' ideas in a framework often described as "quadratic Fourier analysis". Indeed, Gowers' original aim seems to have included finding a proof of Szemerédi's theorem which was closer to Roth's proof of Roth's theorem for arithmetic progressions of length three and Green and Tao's framework helps highlight these parallels. This subsection is more thoroughly explained in the paper [GT08].

Roth's proof of Roth's theorem has, at its core, something now called a U^2-inverse theorem. The U^2-norm of a function f on a finite (compact) Abelian group G is defined by

$$\|f\|_{U^2(G)}^4 = \mathbb{E}_{x,y,z\in G} f(x)\overline{f(x+y)f(x+z)}f(x+y+z).$$

It turns out that this is a norm, and if A and B are two sets in G with $\|1_A - 1_B\|_{U^2(G)}$ small, then the number of three-term arithmetic progressions in A is close to that in B. This is why the U^2-norm is useful for understanding problems about three-term arithmetic progressions. It turns out that if a function does not have small U^2-norm, then it has a linear bias in the following sense.

Theorem 13.1 ($U^2(\mathbb{F}_2^n)$-inverse theorem)**.** *Suppose that $f \in L^\infty(\mathbb{F}_2^n)$ has $\|f\|_{U^2(\mathbb{F}_2^n)} \geq \delta\|f\|_{L^\infty(\mathbb{F}_2^n)}$. Then there is a linear polynomial $l : \mathbb{F}_2^n \to \mathbb{F}_2$, meaning a map $x \mapsto r \cdot x$ for some $r \in \mathbb{F}_2^n$, such that*

$$|\langle f, (-1)^l\rangle_{L^2(\mathbb{F}_2^n)}| \geq \delta^{O(1)}\|f\|_{L^\infty(\mathbb{F}_2^n)}.$$

This is essentially trivial to prove and, a version for the group $G = \mathbb{Z}/N\mathbb{Z}$ rather than $\mathbb{F}_2^n$, can be used as the basis for an iteration to prove Roth's theorem on three-term arithmetic progressions.

Now suppose that one is interested in four-term arithmetic progressions. In this case if we have two sets A and B with $\|1_A - 1_B\|_{U^2(G)}$ small it is *not* necessarily the case that A and B have similar numbers of four-term arithmetic progressions. There is, however, a stronger norm called the U^3-norm for which this is true. The U^3-norm of a function f on a finite (compact) Abelian group G is defined by

$$\begin{aligned}\|f\|_{U^3(G)}^8 &= \mathbb{E}_{x,y,z,w\in G}\Big(f(x)\overline{f(x+y)f(x+z)f(x+w)}\cdot\\ &\qquad \times f(x+y+z)f(x+y+w)f(x+z+w)\overline{f(x+y+z+w)}\Big).\end{aligned}$$

It turns out that this is also a norm and there is a U^3-inverse theorem. This is where Theorem 1.4 can be inserted into the various proofs of the inverse theorem.

For $\mathbb{F}_2^n$ this is due to Samorodnitsky [Sam07] (see also [Wol09]) for $\mathbb{F}_2^n$, and one gets the following.

Theorem 13.2 ($U^3(\mathbb{F}_2^n)$-inverse theorem)**.** *Suppose that $f \in L^\infty(\mathbb{F}_2^n)$ has $\|f\|_{U^3(\mathbb{F}_2^n)} \geq \delta\|f\|_{L^\infty(\mathbb{F}_2^n)}$. Then there is a quadratic polynomial $q : \mathbb{F}_2^n \to \mathbb{F}_2$, meaning a map $x \mapsto x \cdot Ax$ where A is an upper triangular matrix $\mathbb{F}_2^n \to \mathbb{F}_2^n$, such that*

$$|\langle f, (-1)^q\rangle_{L^2(\mathbb{F}_2^n)}| \geq \exp(-O(\log^{3+o(1)}\delta^{-1}))\|f\|_{L^\infty(\mathbb{F}_2^n)}.$$

This is much harder to prove than the $U^2(\mathbb{F}_2^n)$-inverse theorem, and there is actually a close relationship between this and Marton's conjecture. Indeed, Green and Tao in [GT10] and Lovett in [Lov10] showed that Marton's conjecture for $\mathbb{F}_2^n$ is equivalent to the following.

Conjecture 13.3 (Polynomial $U^3(\mathbb{F}_2^n)$-inverse conjecture)**.** *Suppose that $f \in L^\infty(\mathbb{F}_2^n)$ has $\|f\|_{U^3(\mathbb{F}_2^n)} \geq \delta\|f\|_{L^\infty(\mathbb{F}_2^n)}$. Then there is a quadratic polynomial $q : \mathbb{F}_2^n \to \mathbb{F}_2$ such that*

$$|\langle f, (-1)^q\rangle_{L^2(\mathbb{F}_2^n)}| \geq \exp(-O(\log\delta^{-1}))\|f\|_{L^\infty(\mathbb{F}_2^n)}.$$

If true, this would bring the $U^3(\mathbb{F}_2^n)$-inverse state of affairs in line with the U^2 situation.

Again, the analogue of the $U^3(\mathbb{F}_2^n)$-inverse theorem for the group $G = \mathbb{Z}/N\mathbb{Z}$ can be used to give a proof of Szemerédi's theorem for progressions of length four, and, of course, there are higher analogues called U^k-norms for longer progressions, but again we do not discuss this here.

Long arithmetic progressions in sumsets. The question of finding long arithmetic progressions in sets of integers is one of central interest in additive combinatorics. The basic question has the following form: suppose that $A_1, \dots, A_k \subset \{1, \dots, N\}$ all have density at least α. How long an arithmetic progression can we guarantee that $A_1 + \cdots + A_k$ contains?

For one set this is addressed by the notoriously difficult Szemerédi's theorem [Sze69, Sze75], where the best quantitative work is that of Gowers [Gow98, Gow01] (as mentioned in the previous subsection). For two sets the longest progression is much longer with the state of the art due to Green [Gre02] (see also Croot and Sisask [CS10]). For three sets or more the results get even stronger with the work of Freĭman, Halberstam and Ruzsa [FHR92]; and finally, for eight sets or more, longer again by the recent work of Schoen [Sch11].

The ideas around Theorem 1.4 (see [San10]) can be used to give an improvement for four sets or more, and in particular we have the following theorem.

Theorem 13.4. *Suppose that $A_1, \dots, A_4 \subset \{1, \dots, N\}$ all have density at least α. Then $A_1 + \cdots + A_4$ contains an arithmetic progression of length $N^{O(\log^{-O(1)} 2\alpha^{-1})}$.*

$\Lambda(4)$-estimate for the squares. A wonderful conjecture of Rudin [Rud60] asserts that the squares are a $\Lambda(4)$-set. In symbols this is the following conjecture.

Conjecture 13.5. *Suppose that $n_1, \dots, n_k$ are natural numbers. Then*

$$\int \left|\sum_{i=1}^k \exp(2\pi i n_i^2\theta)\right|^4 d\theta = O(k^{2+o(1)}).$$

Inserting ideas around Theorem 1.4 (see [San10]) into the work of [Cha04] (itself developed from an argument of Bourgain in [JL01]) yields the following result.

Theorem 13.6. *Suppose that $n_1, \dots, n_k$ are natural numbers. Then*

$$\int \left| \sum_{i=1}^{k} \exp(2\pi i n_i^2 \theta) \right|^4 d\theta = O(k^3 \exp(-\Omega(\log^{\Omega(1)} 2k))).$$

This is essentially equivalent to inserting Theorem 1.4 into the proof of [Sch11, Theorem 8] and Gowers' [Gow98] version of the Balog-Szemerédi Lemma [BS94]. Of course, this is far form Rudin's conjecture, but it is still the best known result at this time.

The Konyagin-Łaba theorem. Ideas around Theorem 1.4 (see [San10]) inserted into the argument at the end of [Sch11] yield the following quantitative improvement to a result from [KŁ06].

Theorem 13.7 (Konyagin-Łaba theorem)**.** *Suppose that A is a set of reals and $\alpha \in \mathbb{R}$ is transcendental. Then*

$$|A + \alpha.A| = \exp(\Omega(\log^{\Omega(1)} 2|A|))|A|.$$

What is particularly interesting here is that there is a simple construction which shows that there are arbitrarily large sets A with $|A+\alpha.A| = \exp(O(\sqrt{\log |A|}))|A|$.

Acknowledgments

The author should very much like to thank Andrew Granville for a very thorough reading of this paper and supplying a much clearer proof of Proposition 9.2, Sergei Konyagin for a talk on his improvements at the Paul Turán memorial conference 2011, Olof Sisask for directing the author's attention to a better proof of the Marcinkiewicz-Zygmund inequality, and an anonymous referee for a very thorough reading of this paper which has made it immeasurably clearer.

It should also be apparent that the author is heavily influenced by the work of Ben Green, Imre Ruzsa and Terry Tao, and this survey would not exist without their numerous insights. Ben, in particular, has been exceptionally generous with his ideas and conversations.

About the author

Tom Sanders is a research fellow at Oxford University, and his research on the structure theory of set addition was part of the work for which he was awarded the Adams Prize in 2011.

References

[Bil99] Y. Bilu. Structure of sets with small sumset. *Astérisque*, (258):xi, 77–108, 1999. Structure theory of set addition. MR1701189 (2000h:11109)

[Bog39] N. Bogolioùboff. Sur quelques propriétés arithmétiques des presque-périodes. *Ann. Chaire Phys. Math. Kiev*, 4:185–205, 1939. MR0020164 (8:512b)

[Bou99] J. Bourgain. On triples in arithmetic progression. *Geom. Funct. Anal.*, 9(5):968–984, 1999. MR1726234 (2001h:11132)

[Bou08] J. Bourgain. Roth's theorem on progressions revisited. *J. Anal. Math.*, 104:155–192, 2008. MR2403433 (2009g:11011)

[BS94] A. Balog and E. Szemerédi. A statistical theorem of set addition. *Combinatorica*, 14(3):263–268, 1994. MR1305895 (95m:11019)

[Cha02] M.-C. Chang. A polynomial bound in Freĭman's theorem. *Duke Math. J.*, 113(3):399–419, 2002. MR1909605 (2003d:11151)

[Cha04] M.-C. Chang. On problems of Erdős and Rudin. *J. Funct. Anal.*, 207(2):444–460, 2004. MR2032997 (2004j:11022)

[Cha09] M.-C. Chang. Some consequences of the polynomial Freĭman-Ruzsa conjecture. *C. R. Math. Acad. Sci. Paris*, 347(11-12):583–588, 2009. MR2532910 (2010e:11005)

[CŁS11] E. S. Croot, I. Łaba, and O. Sisask. Arithmetic progressions in sumsets and L^p-almost-periodicity. 2011, arXiv:1103.6000.

[CS10] E. S. Croot and O. Sisask. A probabilistic technique for finding almost-periods of convolutions. *Geom. Funct. Anal.*, 20(6):1367–1396, 2010. MR2738997 (2012d:11019)

[DHP04] J.-M. Deshouillers, F. Hennecart, and A. Plagne. On small sumsets in $(\mathbb{Z}/2\mathbb{Z})^n$. *Combinatorica*, 24(1):53–68, 2004. MR2057683 (2005f:11231)

[FHR92] G. A. Freĭman, H. Halberstam, and I. Z. Ruzsa. Integer sum sets containing long arithmetic progressions. *J. London Math. Soc. (2)*, 46(2):193–201, 1992. MR1182477 (93j:11008)

[Fre66] G. A. Freĭman. *Nachala strukturnoi teorii slozheniya mnozhestv*. Kazan. Gosudarstv. Ped. Inst, 1966. MR0360495 (50:12943)

[Fre73a] G. A. Freĭman. *Foundations of a structural theory of set addition*. American Mathematical Society, Providence, R. I., 1973. Translated from the Russian, Translations of Mathematical Monographs, Vol 37.

[Fre73b] G. A. Freĭman. Groups and the inverse problems of additive number theory. In *Number-theoretic studies in the Markov spectrum and in the structural theory of set addition (Russian)*, pages 175–183. Kalinin. Gos. Univ., Moscow, 1973.

[Gow98] W. T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998. MR1631259 (2000d:11019)

[Gow01] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001. MR1844079 (2002k:11014)

[GR06] B. J. Green and I. Z. Ruzsa. Sets with small sumset and rectification. *Bull. London Math. Soc.*, 38(1):43–52, 2006. MR2201602 (2006i:11027)

[GR07] B. J. Green and I. Z. Ruzsa. Freĭman's theorem in an arbitrary abelian group. *J. Lond. Math. Soc. (2)*, 75(1):163–175, 2007. MR2302736 (2007m:20087)

[Gre02] B. J. Green. Arithmetic progressions in sumsets. *Geom. Funct. Anal.*, 12(3):584–597, 2002. MR1924373 (2003i:11148)

[GS08] B. J. Green and T. Sanders. A quantitative version of the idempotent theorem in harmonic analysis. *Ann. of Math. (2)*, 168(3):1025–1054, 2008, arXiv:math/0611286. MR2456890 (2010c:11013)

[GT06] B. J. Green and T. C. Tao. Compressions, convex geometry and the Freĭman-Bilu theorem. *Q. J. Math.*, 57(4):495–504, 2006. MR2277597 (2007g:11013)

[GT08] B. J. Green and T. C. Tao. An inverse theorem for the Gowers $U^3(G)$ norm. *Proc. Edinb. Math. Soc. (2)*, 51(1):73–153, 2008. MR2391635 (2009g:11012)

[GT09a] B. J. Green and T. C. Tao. Freĭman's theorem in finite fields via extremal set theory. *Combin. Probab. Comput.*, 18(3):335–355, 2009. MR2501431 (2010f:11176)

[GT09b] B. J. Green and T. C. Tao. A note on the Freĭman and Balog-Szemerédi-Gowers theorems in finite fields. *J. Aust. Math. Soc.*, 86(1):61–74, 2009. MR2495998 (2010d:11015)

[GT10] B. J. Green and T. C. Tao. An equivalence between inverse sumset theorems and inverse conjectures for the U^3 norm. *Math. Proc. Cambridge Philos. Soc.*, 149(1):1–19, 2010. MR2651575 (2011g:11019)

[JL01] W. B. Johnson and J. Lindenstrauss, editors. *Handbook of the geometry of Banach spaces. Vol. I.* North-Holland Publishing Co., Amsterdam, 2001. MR1863689 (2003f:46013)

[KK10] N. H. Katz and P. Koester. On additive doubling and energy. *SIAM J. Discrete Math.*, 24(4):1684–1693, 2010. MR2746716 (2012d:11020)

[KŁ06] S. V. Konyagin and I. Łaba. Distance sets of well-distributed planar sets for polygonal norms. *Israel J. Math.*, 152:157–179, 2006. MR2214458 (2006m:11032)

[Kne53] M. Kneser. Abscätzungen der symptoticschen dichte von summenmengen. *Math. Z.*, 58:459–484, 1953. MR0056632 (15:104c)

[Kon08] S. V. Konyagin. On Freĭman's theorem in finite fields. *Mat. Zametki*, 84(3):472–474, 2008. MR2473762 (2009i:11013)

[Lov10] S. Lovett. Equivalence of polynomial conjectures in additive combinatorics. 2010, arXiv:1001.3356.

[LR75] J. M. López and K. A. Ross. *Sidon sets*. Marcel Dekker Inc., New York, 1975. Lecture Notes in Pure and Applied Mathematics, Vol. 13. MR0440298 (55:13173)

[Pet11a] G. Petridis. New proofs of Plünnecke-type estimates for product sets in groups. 2011, arXiv:1101.3507.

[Pet11b] G. Petridis. Plünnecke's inequality. 2011, arXiv:1101.2532. MR2847275

[Plü69] H. Plünnecke. *Eigenschaften und Abschätzungen von Wirkungsfunktionen*. BMwF-GMD-22. Gesellschaft für Mathematik und Datenverarbeitung, Bonn, 1969. MR0252348 (40:5569)

[Rud60] W. Rudin. Trigonometric series with gaps. *J. Math. Mech.*, 9:203–227, 1960. MR0116177 (22:6972)

[Rud90] W. Rudin. *Fourier analysis on groups*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1990. Reprint of the 1962 original, A Wiley-Interscience Publication. MR1038803 (91b:43002)

[Ruz78] I. Z. Ruzsa. On the cardinality of $A + A$ and $A - A$. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 933–938. North-Holland, Amsterdam, 1978. MR519317 (80c:05016)

[Ruz89] I. Z. Ruzsa. An application of graph theory to additive number theory. *Scientia, Ser. A.*, 3:97–109, 1989. MR2314377

[Ruz94] I. Z. Ruzsa. Generalized arithmetical progressions and sumsets. *Acta Math. Hungar.*, 65(4):379–388, 1994. MR1281447 (95k:11011)

[Ruz99] I. Z. Ruzsa. An analog of Freĭman's theorem in groups. *Astérisque*, (258):xv, 323–326, 1999. Structure theory of set addition. MR1701207 (2000h:11111)

[Sam07] A. Samorodnitsky. Low-degree tests at large distances. In *STOC '07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 506–515. ACM, New York, 2007. MR2402476 (2009f:68077)

[San10] T. Sanders. On the Bogolyubov-Ruzsa lemma. Anal. PDE, to appear, 2010, arXiv:1011.0107.

[Sch03] T. Schoen. Multiple set addition in $\mathbb{Z}_p$. *Integers*, 3:A17, 6 pp. (electronic), 2003. MR2036483 (2004j:11012)

[Sch11] T. Schoen. Near optimal bounds in Freĭman's theorem. *Duke Math. J.*, 158:1–12, 2011. MR2794366 (2012f:11018)

[SV06] E. Szemerédi and V. Vu. Long arithmetic progressions in sumsets: thresholds and bounds. *J. Amer. Math. Soc.*, 19(1):119–169, 2006. MR2169044 (2006j:11015)

[Sze69] E. Szemerédi. On sets of integers containing no four elements in arithmetic progression. *Acta Math. Acad. Sci. Hungar.*, 20:89–104, 1969. MR0245555 (39:6861)

[Sze75] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975. Collection of articles in memory of Juriĭ Vladimirovič Linnik. MR0369312 (51:5547)

[Tao08] T. C. Tao. Product set estimates for non-commutative groups. *Combinatorica*, 28(5):547–594, 2008. MR2501249 (2010b:11017)

[Tao10] T. C. Tao. Freĭman's theorem for solvable groups. *Contrib. Disc. Math.*, 5(2):137–184, 2010. MR2791295 (2012e:05063)

[TV06] T. C. Tao and H. V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006. MR2289012 (2008a:11002)

[TV07] T. C. Tao and V. H. Vu. On the singularity probability of random Bernoulli matrices. *J. Amer. Math. Soc.*, 20(3):603–628 (electronic), 2007. MR2291914 (2008h:60027)

[Wol09] J. Wolf. A local inverse theorem in $\mathbb{F}_2^n$. Preprint, 2009.

[Zoh11] C. E. Zohar. On sums of generating sets in $\mathbb{Z}_2^n$. 2011, arXiv:1108.4902.

Mathematical Institute, University of Oxford, 24-29 St. Giles', Oxford OX1 3LB, England

E-mail address: tom.sanders@maths.ox.ac.uk