Discrete Fourier analysis

Structures in sumsets

CIMPA Research School, Shillong 2013

Olof Sisask
sisask@kth.se

These are some notes to accompany four hours of lectures given during the CIMPA Research School entitled *Fourier analysis of groups in combinatorics*, held at the North Eastern Hill University, Shillong, India in November 2013. Many thanks to the organisers and participants of this school for a most enjoyable occasion.

Contents

Part 1. Fundamental questions, concepts and techniques 5
1. Sumsets, and some questions 5
2. The Fourier transform and its properties 6
3. Convolution 10
4. Bohr sets 12
5. Notes and references 14

Part 2. Structures in sumsets: three or more summands 17
1. Strategy 17
2. Bogolyubov's lemma 17
3. Structures in $A + A + A$ 19
4. Notes and references 21

Part 3. Structures in sumsets: $A + A$ 23
1. Strategy 23
2. The law of large numbers: the Marcinkiewicz-Zygmund inequality 24
3. Approximation by short trigonometric polynomials and L^p-almost-periodicity 26
4. Long progressions in $A + A$ 27
5. Notes and references 28

Bibliography 31

PART 1

Fundamental questions, concepts and techniques

The aim of this course is to give an introduction to the use of discrete Fourier analysis in tackling additive combinatorial questions. Such use is widespread; we shall have to restrict ourselves to only a small part of the theory, focusing particularly on questions that are naturally phrased in terms of sumsets.

1. Sumsets, and some questions

For two subsets A, B of an abelian group G, we write $A+B := \{a+b : a \in A, b \in B\}$ for their *sumset*. Generally we shall work with G being the integers $\mathbb{Z}$, the group $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$ of integers modulo some number N, or a vector space $\mathbb{F}_p^n$ over a finite field $\mathbb{F}_p$ of size p.

Thus, for example, if $A = \{1, 2, \dots, n\} \subseteq \mathbb{Z}$, then $A + A = \{2, 3, \dots, 2n\}$, and if $A \subseteq \mathbb{F}_p^n$ is a subspace then $A + A = A$.

We shall be interested in the extent to which sumsets must be more additively structured than other sets. Of course, since $A = A + \{0\}$ can be considered a sumset itself, we need to place some kind of restrictions on the summands in order to for the question to be meaningful—we might restrict to A and B having similar sizes, for example. In the two examples just given, this was the case, and in each of these the sumset was highly structured—but then so too were the summands. Let us try taking a highly additively unstructured set.

Exercise 1.1. Let $A = \{1, 3, 9, 27, \dots, 3^{n-1}\}$ be a set of integers. What is $|A+A|$? Show that $A+A$ does not contain any arithmetic progressions x, $x+d$, $x+2d$, $\dots$, $x+(k-1)d$ of length $k \geqslant 4$ (with $d \neq 0$, of course). (*Hint:* *look at base 3 expansions.*)

Thus the sum of a set with itself need not be highly additively structured, at least in the sense of containing long arithmetic progressions, if we are allowed to consider arbitrary sets. On the other hand, in this example A was simply so 'spread out' that there was no real chance of the sumset being very structured; this was also detected by the sumset $A+A$ being much larger than A. Might it be that if A is not too spread out, then $A+A$ must be structured? It turns out that the answer is yes, in a certain sense: if one makes an assumption along the lines of $A + A$ being small compared to A, then one *can* find plenty of additive structure in A, and in particular long arithmetic progressions. Another (and related) way of ensuring that A is not 'too spread out' is to assume that it is a large subset of some 'unspread-out' set, such as the interval $\{1, \dots, N\}$ or one of the groups $\mathbb{Z}_N$ or $\mathbb{F}_p^n$; it is this type of setup that we shall work with in this course.

Here, then, are some natural questions along these lines.

1) How small can $|A+A|$ be in terms of $|A|$? How large can it be?
2) What structures can we find in $A+A$ if $A \subseteq \{1,\dots,N\}$ or $\mathbb{Z}_N$ and $|A| \geqslant \alpha N$?
3) What about $3A := A+A+A$, or $4A$?
4) What can we say about A if $|A+A| < K|A|$ for some number K? Say if $K=2$ or 100, and $|A|$ is large?

While certain answers to Question 1 are more-or-less an exercise in getting to know the definition, lots of interesting theory has been used and developed in order to provide good answers to the other questions. It is part of this theory that we shall study, focusing on Questions 2 and 3, with a quick look at (the highly related) Question 4 if there is time.

We begin by reviewing our most fundamental tool: the Fourier transform.

2. The Fourier transform and its properties

Most of the material in this section will have been covered in other parts of the school, but we include it here for ease-of-reference and since some of the details and notation differ[1].

2.1. Characters and dual groups.

Definition 2.2. Let G be a finite abelian group. By a *character* on G we mean a group homomorphism from G to the multiplicative group of non-zero complex numbers $\mathbb{C}^\times$. We denote the collection of all characters on G by $\widehat{G}$ and give this the group operation of pointwise multiplication of functions; we call this group the *dual* of G.

Thus $\gamma : G \to \mathbb{C}^\times$ is a character if $\gamma(a+b) = \gamma(a)\gamma(b)$ for all $a, b \in G$, and for any $\gamma_1, \gamma_2 \in \widehat{G}$ the character $\gamma_1\gamma_2 \in \widehat{G}$ is given by $(\gamma_1\gamma_2)(x) = \gamma_1(x)\gamma_2(x)$ for each $x \in G$. Note that we write the group operation multiplicatively on $\mathbb{C}^\times$ and hence on $\widehat{G}$ too.

Exercise 2.3. Suppose G is a finite abelian group with exponent n. Show that a character on G can only assume values in the n^{th} roots of unity $\{z \in \mathbb{C} : z^n = 1\}$. What is the identity element in $\widehat{G}$? Show that the inverse of a character $\gamma \in \widehat{G}$ is the conjugate character $\overline{\gamma}$, defined by $\overline{\gamma}(x) = \overline{\gamma(x)}$.

Example 2.4. Let N be a positive integer and let $G = \mathbb{Z}_N$. For any $r \in \mathbb{Z}_N$ the map $\gamma_r : x \mapsto e^{2\pi irx/N}$ is a character on G, and, conversely, any character is of this form (prove it). Furthermore, with this identification we have $\widehat{\mathbb{Z}_N} \cong \mathbb{Z}_N$.

This example is in fact a rather representative one thanks to the special role played by cyclic groups in the fundamental theorem of finite abelian groups. To be more precise, let us for two functions $f : X \to \mathbb{C}$ and $g : Y \to \mathbb{C}$ write $f \otimes g$ for the function from

[1]The material in this section and the remainder of Part 1 was largely extracted from some notes I had written for an undergraduate course on additive combinatorics, and so contains more detail than we dealt with during the school. I have left these details here in case they are useful to someone, but you may wish to skip or skim them.

$X \times Y$ to $\mathbb{C}$ given by

$$(f \otimes g)(x, y) = f(x)g(y).$$

Lemma 2.5. *Let G and H be finite abelian groups.*

(i) If G and H are isomorphic then so are $\widehat{G}$ and $\widehat{H}$.
(ii) For any characters $\gamma \in \widehat{G}$ and $\chi \in \widehat{H}$ the function $\gamma \otimes \chi$ is a character on $G \times H$. Moreover, this identification yields an isomorphism between $\widehat{G} \times \widehat{H}$ and the dual of $G \times H$.

Proof. Exercise. □

Proposition 2.6. *Let G be a finite abelian group. Then $G \cong \widehat{G}$. In particular, $|\widehat{G}| = |G|$.*

Proof. By the fundamental theorem of finite abelian groups and part (i) of the preceding lemma it suffices to prove the theorem for $G = \mathbb{Z}_{N_1} \times \cdots \times \mathbb{Z}_{N_k}$. The result then follows immediately from Example 2.4 and part (ii) of the lemma. □

Example 2.7. Endow the vector space $\mathbb{F}_p^n$ over the finite field $\mathbb{F}_p$ with the dot product $v \cdot w = v_1 w_1 + \cdots + v_n w_n$ for elements $v, w \in \mathbb{F}_p^n$. For any $r \in \mathbb{F}_p^n$, the map

$$\gamma_r : v \mapsto e^{2\pi i (r \cdot v)/p}$$

is then a character on $\mathbb{F}_p^n$. Moreover, every character on $\mathbb{F}_p^n$ has this form.

The fact that a finite abelian group and its dual are isomorphic can be useful to bear in mind, but is in a sense not essential to the theory of characters and Fourier transforms. The following result, on the other hand, is.

Lemma 2.8 (Sufficiently many characters)**.** *Let G be a finite abelian group. For any non-zero $x \in G$ there is a character $\gamma \in \widehat{G}$ such that $\gamma(x) \neq 1$.*

Proof. We may assume that $G = \mathbb{Z}_{N_1} \times \cdots \times \mathbb{Z}_{N_k}$ for some integers $N_i \geqslant 2$. Let $x = (x_1, \ldots, x_k) \in G$ be non-zero; we must then have $x_j \neq 0$ (in $\mathbb{Z}_{N_j}$) for some j. Fix such a j and define $\gamma \in \widehat{G}$ by $\gamma(y_1, \ldots, y_k) = e^{2\pi i y_j / N_j}$; this is a character on G such that $\gamma(x) \neq 1$. □

In order to discuss some of the useful properties of characters succinctly let us introduce a piece of notation. If X is a finite non-empty set and $f : X \to \mathbb{C}$ is a function then we write

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

for the average of f over X. If the set X is clear from the context, or is the whole domain of definition of f, then we may simply write $\mathbb{E}_x f(x)$ instead.

The following relations are known as the *character orthogonality relations*; they form the foundation of discrete Fourier analysis.

Proposition 2.9 (Orthogonality of characters)**.** *Let G be a finite abelian group and suppose $a, b \in G$ and $\gamma_1, \gamma_2 \in \widehat{G}$. Then*

$$\mathbb{E}_{x\in G}\gamma_1\overline{\gamma_2}(x) = \begin{cases} 1 & \text{if } \gamma_1 = \gamma_2 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\sum_{\gamma\in\widehat{G}} \gamma(a-b) = \begin{cases} |G| & \text{if } a = b \\ 0 & \text{otherwise.} \end{cases}$$

Proof. For the first claim it suffices to prove that for any non-identity character $\gamma \in \widehat{G}$ one has $\mathbb{E}_x\gamma(x) = 0$. Now, for any $\gamma \in \widehat{G}$ we have $\mathbb{E}_{x\in G}\gamma(x) = \mathbb{E}_{x\in G}\gamma(x+y) = \gamma(y)\mathbb{E}_{x\in G}\gamma(x)$ for each $y \in G$, the first equality being a result of a change of variables. Letting $y \in G$ be such that $\gamma(y) \neq 1$ if γ is not the identity character then proves the claim.
For the second claim of the proposition it similarly suffices to prove that $\sum_{\gamma\in\widehat{G}}\gamma(x) = 0$ if $x \neq 0$. This follows in a similar way to the previous part, now using Lemma 2.8 to find a character $\chi \in \widehat{G}$ such that $\chi(x) \neq 1$ and shifting the summation over $\widehat{G}$ by χ. □

Remark 2.10. We shall generally use averaged sums when summing over the group G and ordinary sums when summing over $\widehat{G}$; this means that one does not need to remember various normalisations in results to come.

2.11. The Fourier transform. The above relations are called orthogonality relations for the following reason. We can define an inner product $\langle\cdot,\cdot\rangle_G$ on the complex vector space $\mathbb{C}^G$ of complex-valued functions on G by setting

$$\langle f, g\rangle_G = \mathbb{E}_{x\in G} f(x)\overline{g(x)}$$

for $f, g : G \to \mathbb{C}$. The first relation then says that the characters on G form an orthonormal set in $\mathbb{C}^G$, and since there are $|G|$ characters (which is the dimension of $\mathbb{C}^G$) they in fact form an orthonormal basis. This observation leads us to the definition of the Fourier transform, which tells us how to express a function $f : G \to \mathbb{C}$ in terms of this basis.

Definition 2.12. Let G be a finite abelian group. Given a function $f : G \to \mathbb{C}$ we define the *Fourier transform* of f to be the function $\widehat{f} : \widehat{G} \to \mathbb{C}$ given by

$$\widehat{f}(\gamma) = \mathbb{E}_{x\in G} f(x)\overline{\gamma(x)}.$$

The complex numbers $\widehat{f}(\gamma)$ for $\gamma \in \widehat{G}$ are known as the *Fourier coefficients* of f.

Thus $\widehat{f}(\gamma) = \langle f, \gamma\rangle_G$. In particular, if $\gamma \in \widehat{G}$ then $\widehat{\gamma}(\chi) = 1$ if $\chi = \gamma$ and is 0 otherwise, so $\widehat{1}(\gamma) = 1$ if $\gamma = 1$ and is 0 otherwise. Note also that $\widehat{f}(1) = \mathbb{E}_{x\in G} f(x)$ is precisely the average of f, giving this Fourier coefficient a special status. Furthermore, the operation of taking the Fourier transform is linear: $\widehat{\alpha f + \beta g} = \alpha\widehat{f} + \beta\widehat{g}$ for any $\alpha, \beta \in \mathbb{C}$.

Importantly, the Fourier coefficients tell us how to express f as a linear combination of characters:

Theorem 2.13 (Fourier inversion formula)**.** *Let $f : G \to \mathbb{C}$ be a function on a finite abelian group G. Then, for any $x \in G$,*

$$f(x) = \sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\gamma(x).$$

In particular, f is completely determined by its sequence of Fourier coefficients.

Proof. This is the standard identity $f = \sum_\gamma \langle f, \gamma\rangle \gamma$ from linear algebra, valid since the characters form an orthonormal basis of $\mathbb{C}^G$. More directly,

$$\sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\gamma(x) = \mathbb{E}_{y \in G} f(y) \sum_{\gamma \in \widehat{G}} \gamma(x-y) = f(x),$$

by the character orthogonality relations. □

We also furnish $\mathbb{C}^{\widehat{G}}$, the vector space of complex-valued functions on $\widehat{G}$, with the inner product

$$\langle f, g\rangle_{\widehat{G}} = \sum_{\gamma \in \widehat{G}} f(\gamma)\overline{g(\gamma)}$$

for functions $f, g : \widehat{G} \to \mathbb{C}$. This differs from the inner product that we put on $\mathbb{C}^G$ only in the normalisation of the sum. There is of course a slight abuse of notation here since $\widehat{G}$ is a group in its own right, and so we could use our previous definition of the inner product on a group here too, but rather than introduce extra notation to circumvent this we shall simply employ the convention that if there is a $\widehat{\ }$ appearing in the subscript of the inner product then we use unnormalised sums and if there is not then we use averaged sums.

Theorem 2.14 (The Plancherel theorem and Parseval's identity)**.** *Let G be a finite abelian group and let $f, g : G \to \mathbb{C}$ be two functions. Then*

$$\langle f, g\rangle_G = \langle \widehat{f}, \widehat{g}\rangle_{\widehat{G}},$$

or, in other words,

$$\mathbb{E}_{x \in G} f(x)\overline{g(x)} = \sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\overline{\widehat{g}(\gamma)}.$$

In particular we have the identity

$$\mathbb{E}_{x \in G} |f(x)|^2 = \sum_{\gamma \in \widehat{G}} |\widehat{f}(\gamma)|^2.$$

Proof. Again these identities follow from the properties of orthonormal bases in inner product spaces. More directly,

$$\sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\overline{\widehat{g}(\gamma)} = \mathbb{E}_x \mathbb{E}_y f(x)\overline{g(y)} \sum_{\gamma \in \widehat{G}} \gamma(y-x),$$

whence the result follows from the character orthogonality relations. □

Note that we have followed our convention of using $\mathbb{E}$ for sums over the group and $\sum$ for sums over the dual group here; this is what ensures that we do not have to put in normalising factors of $|G|$ in the above identities.

Remark 2.15. We have defined the Fourier transform of a function $f : G \to \mathbb{C}$ to be a function $\widehat{f}$ on $\widehat{G}$, the dual of G. Now, we know that $\widehat{G} \cong G$ for finite abelian groups G, and so we could have defined $\widehat{f}$ to be a function on G. This approach may be preferable in some instances, but we shall avoid doing this in these notes as we find that distinguishing $\widehat{G}$ and G keeps things conceptually simpler in arguments. The added level of abstraction will hopefully not cause any problems given the explicit description one can give of characters in $\widehat{G}$ in terms of elements of G: we shall usually be interested in the groups $\mathbb{Z}_N$ and $\mathbb{F}_p^n$, and in these cases one can think of characters in terms of their explicit descriptions given in Examples 2.4 and 2.7.

3. Convolution

We can now start to describe some reasons why the Fourier transform is useful to us. Broadly speaking, our main objects of interest are subsets A of abelian groups, and such sets can be studied via their indicator functions[2] 1_A on the group. These indicator functions, in turn, can be studied via their Fourier transforms $\widehat{1_A}$. It might seem like this would make things more complicated, replacing a $\{0,1\}$-valued function by a complex-valued function, but it turns out this reparameterisation reveals a lot about A in a simple way. Indeed, characters are homomorphisms, so behave nicely under addition, and since $\widehat{1_A}$ tells us what 1_A looks like decomposed into characters, we should be kind of happy with this move[3]. A more tangible reason to be happy comes from the operation of *convolution*.

Definition 3.1. Let G be a finite abelian group. For any two functions $f, g : G \to \mathbb{C}$ we define their *convolution* $f * g$ to be the function from G to $\mathbb{C}$ given by

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

Lemma 3.2 (Basic properties of convolutions)**.** *The operation of convolution is commutative, associative and bilinear. In other words, for any $f, g, h : G \to \mathbb{C}$ we have*

(i) $f * g = g * f$,
(ii) $f * (g * h) = (f * g) * h$ *(and so we may write $f * g * h$), and*
(iii) $(f + g) * h = f * h + g * h$ *and* $f * (g + h) = f * g + f * h$.

Furthermore we have $\mathbb{E}_x f * g(x) = (\mathbb{E}_x f(x))(\mathbb{E}_x g(x))$.

Proof. We prove commutativity and leave the proof of the other claims as an exercise. For any $x \in G$ we have

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

the second equality following from changing the order of summation by setting $z = x - y$. □

[2] If X is a set then we write 1_X for the *indicator function* of X, defined by $1_X(x) = 1$ if $x \in X$ and $1_X(x) = 0$ if $x \notin X$.

[3] Incidentally, expressing an indicator function as a sum of other functions in a non-trivial way is a generally useful trick: we saw it used in Anirban Mukhopadhyay's number theory course, for example, where we at several points used a decomposition of the form $1_{\{1\}}(n) = \sum_{d|n} \mu(d)$.

Convolution is directly relevant to additive combinatorics through the following lemma. For a function $f : X \to \mathbb{C}$ we write $\mathrm{supp}(f) = \{x \in X : f(x) \neq 0\}$ for the *support* of f.

Lemma 3.3. *Let G be a finite abelian group and let $A, B, A_1, \dots, A_k \subseteq G$. Then*

$$A_1 + \cdots + A_k = \mathrm{supp}(1_{A_1} * \cdots * 1_{A_k}).$$

In particular we have that $A + B = \mathrm{supp}(1_A * 1_B)$. Moreover we have the more specific relationship

$$1_{A_1} * \cdots * 1_{A_k}(x) = \frac{|\{(a_1, \dots, a_k) \in A_1 \times \cdots \times A_k : a_1 + \cdots + a_k = x\}|}{|G|^{k-1}}; \qquad (3.1)$$

this convolution thus gives a normalised count of how many ways there are of writing x as $a_1 + \cdots + a_k$ with $a_i \in A_i$. Finally we have

$$1_A * 1_B(x) = \frac{|A \cap (x - B)|}{|G|}.$$

Proof. Exercise. □

Thus, not only can we study sumsets $A + A = \mathrm{supp}(1_A * 1_A)$ and $A + A + A = \mathrm{supp}(1_A * 1_A * 1_A)$ using convolutions, but we can study how 'robustly' an element lies in the sumset, this being measured by $1_A * 1_A(x)$ or $1_A * 1_A * 1_A(x)$.

Exercise 3.4. What is $f * 1$? What is $1_A * 1_{\{x\}}$?

It may at this stage seem more natural not to normalise as we have done in the definition of convolution, but it turns out that this normalisation is useful in many situations—in particular it fits with our convention of taking averages over G (and sums over $\widehat{G}$) from our discussion of the Fourier transform. Note that a particular consequence of it is that we always have $0 \leqslant 1_{A_1} * \cdots * 1_{A_k}(x) \leqslant 1$ for any $x \in G$. More generally, the convolution of two $[0, 1]$-valued functions is itself a $[0, 1]$-valued function.

A key property of the Fourier transform is that it interacts nicely with convolutions. Specifically, Fourier transforms convert the somewhat complex operation of convolution to the simple operation of taking pointwise products.

Lemma 3.5. *Let $f, g : G \to \mathbb{C}$ be two functions on a finite abelian group G. Then*

$$\widehat{f * g} = \widehat{f} \cdot \widehat{g}$$

We again leave the proof to the reader.

Note, then, that the Fourier inversion formula tells us that

$$f * g(x) = \sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\widehat{g}(\gamma)\gamma(x)$$

and more generally that

$$f_1 * \cdots * f_k(x) = \sum_{\gamma \in \widehat{G}} \widehat{f_1}(\gamma) \cdots \widehat{f_k}(\gamma)\gamma(x);$$

expressions we shall find very useful shortly.

Exercise 3.6. Use this to prove that the two-variable sum

$$\sum_{x,d\in G} 1_A(x)1_A(x+d)1_A(x+2d) \tag{3.2}$$

counting the number of three-term progressions in A is equal to a multiple of the one-variable sum

$$\sum_{\gamma\in\widehat{G}} \widehat{1_A}(\gamma)^2\widehat{1_A}(\gamma^{-2}) \tag{3.3}$$

when expressed in terms of $\widehat{1_A}$, at least if $|G|$ is odd. (You should be able to guess the multiple from our conventions!) This will be very useful in another part of this school.

4. Bohr sets

We mentioned some different groups of interest at the start: $\mathbb{Z}$, $\mathbb{Z}_N$ and $\mathbb{F}_p^n$. We saw different types of sets with small sumset in these groups: in the first two we had the *arithmetic progressions*, whereas in the third we had *subspaces*. In this section we introduce the notion of a Bohr set, which to some extent extends to arbitrary groups both the concepts of subspaces and arithmetic progressions simultaneously. This notion will crop up fairly naturally in the course of our arguments.

Definition 4.1. Let G be a finite abelian group, let $\Gamma \subseteq \widehat{G}$ be a set of characters and let $\delta \geqslant 0$. Then we define

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

and call this a *Bohr set* with *rank* $|\Gamma|$ and *radius* δ. If the group G is clear from the context we may drop the subscript from the Bohr_G.

A Bohr set is thus an *approximate annihilator* of a certain set of characters; it should be compared with the definition of the annihilator of a subspace V in linear algebra.

Before we make these relationships between Bohr sets, subspaces and arithmetic progressions more precise, let us note some elementary properties of Bohr sets in the form of some exercises. In each of these exercises we suppose that G is a finite abelian group.

Exercise 4.2. Let $\Gamma \subseteq \widehat{G}$ be a set of characters. Show that $\mathrm{Bohr}_G(\Gamma,\delta)$ is symmetric, that is that $\mathrm{Bohr}_G(\Gamma,\delta) = -\mathrm{Bohr}_G(\Gamma,\delta)$. Show also that $\mathrm{Bohr}_G(\Gamma,\delta) = G$ if $\delta \geqslant 2$.

Exercise 4.3. Let $\Gamma_1,\Gamma_2 \subseteq \widehat{G}$ be sets of characters and suppose $\delta \geqslant 0$. Show that $\mathrm{Bohr}(\Gamma_1 \cup \Gamma_2,\delta) = \mathrm{Bohr}(\Gamma_1,\delta) \cap \mathrm{Bohr}(\Gamma_2,\delta)$.

Exercise 4.4. Let $\Gamma \subseteq \widehat{G}$ be a set of characters and suppose $\delta_1,\delta_2 \geqslant 0$. Show that $\mathrm{Bohr}(\Gamma,\delta_1) + \mathrm{Bohr}(\Gamma,\delta_2) \subseteq \mathrm{Bohr}(\Gamma,\delta_1+\delta_2)$.

We next show that, in $\mathbb{F}_p^n$, Bohr sets of not too large rank contain large subspaces.

Proposition 4.5 (Bohr sets contain subspaces)**.** *Let p be a prime and let $G = \mathbb{F}_p^n$ for some $n \geqslant 1$. Let $\Gamma \subseteq \widehat{\mathbb{F}_p^n}$ be a set of characters and suppose $\delta \geqslant 0$. Then the Bohr set $\mathrm{Bohr}(\Gamma,\delta)$ contains a subspace of $\mathbb{F}_p^n$ of dimension at least $n - |\Gamma|$.*

Proof. The point here is that $\mathrm{Bohr}(\Gamma,\delta)$ contains the subspace $\langle\Gamma\rangle^\perp$, the actual annihilator of the subspace generated by Γ. The result then follows as this subspace has dimension $n-\dim\langle\Gamma\rangle$. But let us be a bit more explicit: let us identify $\widehat{\mathbb{F}_p^n}$ with $\mathbb{F}_p^n$ itself as in Example 2.7. Then, writing $\gamma_1,\dots,\gamma_k$ for the elements of Γ, we have[4] $\gamma_i(x)=e((r_i\cdot x)/p)$ for some vectors $r_i\in\mathbb{F}_p^n$, where $v\cdot w$ is the usual dot product of vectors. Let

$$V=\{x\in\mathbb{F}_p^n : r_i\cdot x=0 \text{ for all } i=1,\dots,k\}=\ker f,$$

where $f:\mathbb{F}_p^n\to\mathbb{F}_p^k$ is defined by

$$f(x)=(r_1\cdot x,\dots,r_k\cdot x).$$

This is a linear map, and since the dimension of its image is clearly at most k we have, by the rank-nullity theorem, that

$$\dim V=n-\operatorname{rank} f\geqslant n-k.$$

We also have that $V\subseteq\mathrm{Bohr}(\Gamma,\delta)$, since for any $x\in V$ and any $i\in\{1,\dots,k\}$

$$|\gamma_i(x)-1|=|e((r_i\cdot x)/p)-1|=0\leqslant\delta,$$

and so we are done. □

So Bohr sets in $\mathbb{F}_p^n$ are highly structured: they contain (and are often equal to) huge subspaces. What about in $\mathbb{Z}_N$? We shall use the following lemma from Diophantine approximation to get a handle on the situation; $\|x\|_{\mathbb{R}/\mathbb{Z}}$ represents the distance from x to the nearest integer.

Lemma 4.6. *Let N be a positive integer and let $x_1,\dots,x_k\in\mathbb{Z}_N$. Then there is a non-zero element $d\in\mathbb{Z}_N$ such that*

$$\left\|\frac{dx_i}{N}\right\|_{\mathbb{R}/\mathbb{Z}}\leqslant N^{-1/k} \text{ for all } i=1,\dots,k.$$

Proof. This is Dirichlet's boxing argument. Define $\boldsymbol{x}=(x_1,\dots,x_k)\in\mathbb{Z}_N^k$ and consider the elements $\mathbf{0},\boldsymbol{x},2\boldsymbol{x},\dots,(N-1)\boldsymbol{x}$ in $\mathbb{Z}_N^k$. Base a cube of side $\lceil N^{1-1/k}\rceil$ at each of these points; since there are only N^k points in total in $\mathbb{Z}_N^k$, two of these cubes must intersect. Hence two of the points, say $m\boldsymbol{x}$ and $n\boldsymbol{x}$ with $m\neq n$, must have each pair of components within a distance of at most $\lfloor N^{1-1/k}\rfloor$ of each other. In other words, for each i the element $(n-m)x_i$ must lie between $-N^{1-1/k}$ and $N^{1-1/k}$ in $\mathbb{Z}_N$. We may thus take $d=n-m$. □

Proposition 4.7 (Bohr sets contain arithmetic progressions)**.** *Let N be a prime, let $\Gamma\subseteq\widehat{\mathbb{Z}_N}$ be a set of characters and suppose $\delta>0$. Then $\mathrm{Bohr}_{\mathbb{Z}_N}(\Gamma,\delta)$ contains an arithmetic progression P of length*

$$|P|\geqslant\tfrac{1}{2\pi}\delta N^{1/|\Gamma|}.$$

Proof. We identify $\widehat{\mathbb{Z}_N}$ with $\mathbb{Z}_N$, via Example 2.4 so that, writing $\gamma_1,\dots,\gamma_k$ for the elements of Γ, we have $\gamma_i(x)=e((r_ix)/N)$ for some $r_i\in\mathbb{Z}_N$. By the preceding lemma we can find a non-zero element $d\in\mathbb{Z}_N$ such that $\|dr_i/N\|_{\mathbb{R}/\mathbb{Z}}\leqslant N^{-1/k}$ for each i. Let M be a positive integer to be determined later and define an arithmetic progression P by

$$P=\{-Md,\dots,-d,0,d,2d,\dots,Md\}\subseteq\mathbb{Z}_N;$$

[4]We employ here and elsewhere the standard notation $e(\theta)=e^{2\pi i\theta}$.

we claim that this is contained in $\mathrm{Bohr}(\Gamma, \delta)$ provided we pick M not too large. Indeed, if $n \in [-M, M]$ then, for any $i \in \{1, \dots, k\}$, we have[5] that

$$|\gamma_i(nd) - 1| \leqslant 2\pi \|ndr_i/N\|_{\mathbb{R}/\mathbb{Z}} \leqslant 2\pi |n| \|dr_i/N\|_{\mathbb{R}/\mathbb{Z}} \leqslant 2\pi M/N^{1/k},$$

and this will be at most δ provided we pick $M = \lfloor \delta N^{1/k}/2\pi \rfloor$. Since $|P| = 2M + 1$ we are done. □

5. Notes and references

At the end of each part of these notes are some remarks and references. These are in no way meant to be encyclopaedic; instead the intention is that they might offer starting points if one wants to deepen one's knowledge, and to provide at least some sort of historical context. I apologise in advance to anyone who ought to have been mentioned but was not.

Most of the content of this part of the notes is standard, and the standard reference is the book *Additive Combinatorics* by Tao and Vu [**14**], which contains lots of useful and interesting material.

Before we move on to the course material proper, let us just mention one further type of structure that is of high interest in additive combinatorics:

Generalised progressions. It turns out that much more is true than is suggested by Proposition 4.7: in $\mathbb{Z}_N$ Bohr sets contain large *generalised* arithmetic progressions. We shall not really use this fact in this course, but it is worth bearing in mind.

Definition 5.1 (Generalised arithmetic progressions)**.** A *generalised arithmetic progression* in an abelian group G is any set[6] of the form

$$P = \{a + \lambda_1 x_1 + \cdots + \lambda_d x_d : M_i \leqslant \lambda_i \leqslant N_i\},$$

where $a, x_1, \dots, x_d \in G$ and the M_i and N_i are integers with $M_i \leqslant N_i$. We call $x = (x_1, \dots, x_d)$ the *base* or *basis* of P and d the *dimension* or *rank* of P. P is called *proper* if all the elements $\lambda_1 x_1 + \cdots + \lambda_d x_d$ with $M_i \leqslant \lambda_i \leqslant N_i$ are distinct.

There are several useful ways of thinking about generalised arithmetic progressions. For example, one can think of a generalised arithmetic progression in $\mathbb{Z}$ as a linear projection of a box in $\mathbb{Z}^d$ to $\mathbb{Z}$. One can also think of a generalised arithmetic progression with base set $\{x_1, \dots, x_d\}$ as (a translate of) a truncated version of the subgroup generated by the x_i. Yet another way to think about a generalised arithmetic progression is as a sum of ordinary arithmetic progressions.

By using more sophisticated arguments than those appearing in the proofs of Lemma 4.6 and Proposition 4.7, relying on the Geometry of Numbers, one can prove the following.

[5] Note that $4\|\theta\|_{\mathbb{R}/\mathbb{Z}} \leqslant |e(\theta) - 1| \leqslant 2\pi\|\theta\|_{\mathbb{R}/\mathbb{Z}}$ for all $\theta \in \mathbb{R}$, as follows from drawing a picture.

[6] There is an abuse of terminology here: really the progression should be defined as a tuple consisting of the parameters involved in its definition rather than as a set, but it would get very tedious to talk about such tuples and so we refer simply to the set instead.

Proposition 5.2 (Bohr sets contain GAPs)**.** *Let N be a prime, let $\Gamma \subseteq \widehat{\mathbb{Z}_N}$ be a set of d characters and suppose $\delta \in (0, 2]$. Then* $\mathrm{Bohr}_{\mathbb{Z}_N}(\Gamma, \delta)$ *contains a proper generalised arithmetic progression P of rank at most d and size*

$$|P| \geqslant \left(\frac{\delta}{2\pi d}\right)^d N.$$

For this and more, see §4.4 in Tao–Vu.

PART 2

Structures in sumsets: three or more summands

We now come to answering some of the questions we set out at the start of these notes about structures in sumsets. The more sets we add, the more structure we expect to see, and indeed it turns out that sets like $A + A$ are somewhat trickier to analyse than sets like $A + A + A$. We therefore begin with sums of at least three sets.

1. Strategy

We shall focus on establishing results of the following form.

Theorem 1.1. *Let $\alpha > 0$ and suppose $A \subseteq [1, N]$ is a set of size at least αN. Then $3A = A + A + A$ contains an arithmetic progression P of length*

$$|P| \geqslant c\alpha N^{c\alpha^3},$$

where $c > 0$ is an absolute constant.

In order to prove this we shall use Fourier analysis to study iterated convolutions, and then apply these results about convolutions to obtain information about sumsets using the fact that $A + A + A$ is the support of $1_A * 1_A * 1_A$ (see Lemma 1.3.3). Specifically, we shall show that convolutions are almost-periodic in a certain sense: we shall show that there is a long arithmetic progression P such that, for any element x,

$$|1_A * 1_A * 1_A(x + t) - 1_A * 1_A * 1_A(x)| \leqslant \epsilon \text{ for all } t \in P.$$

We shall then use the fact that there is an x for which $1_A * 1_A * 1_A(x)$ is large to conclude that, for this x, $1_A * 1_A * 1_A(x+t)$ must be non-zero for all $t \in P$, and so $x+P \subseteq A+A+A$.

In the next section we shall employ a related but somewhat simpler argument to deal with sets of the form $2A - 2A$: in addition to being a useful result, this is also a good warm-up in getting to know the tools.

2. Bogolyubov's lemma

We begin with the following result, known as Bogolyubov's lemma. It says that sets of the form $2A - 2A$ must be highly structured if A is large[1].

[1]To avoid repeating the word 'non-empty' everywhere, we shall assume wherever appropriate in these notes that the sets given to us are non-empty and that related constants α are non-zero.

Theorem 2.1 (Bogolyubov's lemma)**.** *Let G be a finite abelian group and suppose $A \subseteq G$ has size $\alpha|G|$. Then there is a set $\Gamma \subseteq \widehat{G}$ of size at most $1/\alpha^2$ such that*

$$\mathrm{Bohr}(\Gamma, \sqrt{2}) \subseteq 2A - 2A.$$

Proof. We use the fact that $2A-2A$ is the support of $1_A * 1_A * 1_{-A} * 1_{-A}$. Applied to this function, the Fourier inversion formula together with the convolution identity $\widehat{f * g} = \widehat{f} \cdot \widehat{g}$ yields

$$1_A * 1_A * 1_{-A} * 1_{-A}(x) = \sum_{\gamma \in \widehat{G}} |\widehat{1_A}(\gamma)|^4 \gamma(x) = \sum_{\gamma \in \widehat{G}} |\widehat{1_A}(\gamma)|^4 \operatorname{Re} \gamma(x); \tag{2.1}$$

our aim is to show that this is strictly greater than 0 for all x in some Bohr set. Define

$$\Gamma = \{\gamma \in \widehat{G} : |\widehat{1_A}(\gamma)| \geqslant \delta\alpha\},$$

the 'large spectrum' of 1_A, where $\delta = \alpha^{1/2}$. By Parseval's identity we have that Γ cannot be very large:

$$|\Gamma|\alpha^3 \leqslant \sum_{\gamma \in \widehat{G}} |\widehat{1_A}(\gamma)|^2 = \mathbb{E}_{x \in G} |1_A(x)|^2 = \alpha,$$

so $|\Gamma| \leqslant 1/\alpha^2$. Now, for any $x \in \mathrm{Bohr}(\Gamma, \sqrt{2})$ and any $\gamma \in \Gamma$ the real part of $\gamma(x)$ is non-negative (draw a picture). Together with the fact that $\widehat{1_A}(1)^4 = \alpha^4$, this implies from (2.1) that

$$\begin{aligned} 1_A * 1_A * 1_{-A} * 1_{-A}(x) &\geqslant \alpha^4 + \sum_{\gamma \in \widehat{G} \setminus \Gamma} |\widehat{1_A}(\gamma)|^4 \operatorname{Re} \gamma(x) \\ &\geqslant \alpha^4 - \sum_{\gamma \in \widehat{G} \setminus \Gamma} |\widehat{1_A}(\gamma)|^4 \\ &> \alpha^4 - \delta^2\alpha^2 \sum_{\gamma \in \widehat{G} \setminus \Gamma} |\widehat{1_A}(\gamma)|^2 \\ &\geqslant 0, \end{aligned}$$

the last inequality following from Parseval's identity. Thus $x \in 2A - 2A$, and since $x \in \mathrm{Bohr}(\Gamma, \sqrt{2})$ was arbitrary, we are done. □

Combining this theorem with Proposition 1.4.7 immediately yields the following nice combinatorial result.

Corollary 2.2 (APs in $2A - 2A$)**.** *Let N be a prime and suppose $A \subseteq \mathbb{Z}_N$ has size αN. Then $2A - 2A$ contains an arithmetic progression P of length*

$$|P| \geqslant \tfrac{1}{5} N^{\alpha^2}.$$

Similarly we have the following result in $\mathbb{F}_p^n$.

Corollary 2.3 (Subspaces in $2A - 2A$)**.** *Let p be a prime and suppose $A \subseteq \mathbb{F}_p^n$ has size αp^n. Then $2A - 2A$ contains a subspace of dimension at least $n - 1/\alpha^2$.*

Note the discrepancy between the sizes of the sets found in $2A - 2A$ in the two setups: in the former we obtain a small power of $|G|$ whereas in the latter we obtain something proportional to $|G|$. This is due to us restricting to an arithmetic progression in the

former result: the enveloping Bohr set has size proportional to $|G|$ in each case, and so too would the size of a generalised arithmetic progression in $2A - 2A$, as follows from Proposition 1.5.2.

Now, in the above argument we used the particular form of the convolution $1_A * 1_A * 1_{-A} * 1_{-A}$, but it turns out that one may argue similarly for more general convolutions of three or more functions; we turn to this next.

3. Structures in $A + A + A$

We would like to also say something similar to Bogolyubov's lemma about sets of the form $A + A + A$. In the proof of this result we made use of the fact that the function $1_A * 1_A * 1_{-A} * 1_{-A}$ had non-negative Fourier coefficients, which need no longer hold true for convolutions of the form $1_A * 1_A * 1_A$. We, may however, use very similar techniques to deal with such convolutions. We start with a preliminary lemma that encodes the heart of the matter.

Lemma 3.1. *Let G be a finite abelian group, let $f : G \to [0,1]$ and write $\alpha = \mathbb{E}_x f(x)$. Suppose $\epsilon > 0$. Then there is a set of characters $\Gamma \subseteq \widehat{G}$ of size at most $4/\alpha\epsilon^2$ such that*

$$|f * f * f(x+t) - f * f * f(x)| < \epsilon\alpha^2 \text{ for all } t \in \mathrm{Bohr}(\Gamma, \epsilon)$$

holds for any $x \in G$.

Remark 3.2. This lemma is saying that three-fold convolutions of the form $f * f * f$ are *almost-periodic*: we have a large, structured supply of almost-periods t. (Recall at this point what a period of a function is.) One can think of this as being a continuity property: one can shift the function by certain translates without affecting its value by much.

Proof. Similarly to the previous proof, let $\Gamma = \{\gamma \in \widehat{G} : |\widehat{f}(\gamma)| \geqslant \delta\alpha\}$ be the large spectrum of f, where $\delta = \epsilon/2$. Parseval's identity again tells us that this set of characters cannot be very large: $|\Gamma| \leqslant 1/\alpha\delta^2 = 4/\alpha\epsilon^2$. Now let $t \in \mathrm{Bohr}(\Gamma, \epsilon)$ and let $x \in G$ be arbitrary. By the Fourier inversion formula and the convolution identity $\widehat{f * g} = \widehat{f} \cdot \widehat{g}$ we have

$$\begin{aligned}
|f * f * f(x+t) - f * f * f(x)| &= |\sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)^3 \gamma(x)(\gamma(t) - 1)| \\
&\leqslant \sum_{\gamma \in \widehat{G}} |\widehat{f}(\gamma)|^3 |\gamma(t) - 1| \\
&\leqslant \sum_{\gamma \in \Gamma} |\widehat{f}(\gamma)|^3 |\gamma(t) - 1| + 2 \sum_{\gamma \in \widehat{G} \setminus \Gamma} |\widehat{f}(\gamma)|^3.
\end{aligned}$$

We can bound the first sum using the fact that, for $\gamma \in \Gamma$, $|\gamma(t) - 1| \leqslant \epsilon$, and we can bound the second sum using the fact that $|\widehat{f}(\gamma)| < \delta\alpha$ for any $\gamma \notin \Gamma$. Using these bounds,

and the bound $|\widehat{f}(\gamma)| \leqslant \alpha$, valid for any $\gamma \in \widehat{G}$, we obtain

$$|f*f*f(x+t) - f*f*f(x)| < \epsilon\alpha \sum_{\gamma\in\Gamma} |\widehat{f}(\gamma)|^2 + \epsilon\alpha \sum_{\gamma\in\widehat{G}\setminus\Gamma} |\widehat{f}(\gamma)|^2$$
$$\leqslant \epsilon\alpha^2,$$

the final inequality following from the consequence $\sum_\gamma |\widehat{f}(\gamma)|^2 = \mathbb{E}_x |f(x)|^2 \leqslant \alpha$ of Parseval's identity. □

This kind of almost-periodicity result is extremely useful. Applying it to $f = 1_A$ we almost immediately obtain the following result.

Theorem 3.3. *Let G be a finite abelian group and suppose $A \subseteq G$ has size $\alpha|G|$. Then there is some element $x \in G$ such that*

$$x + \mathrm{Bohr}(\Gamma, \alpha) \subseteq A + A + A,$$

where $|\Gamma| \leqslant 4/\alpha^3$.

Proof. Apply the previous lemma to $f = 1_A$ with $\epsilon = \alpha$ to obtain a set $\Gamma \subseteq \widehat{G}$ with $|\Gamma| \leqslant 4/\alpha^3$ such that

$$|1_A * 1_A * 1_A(x+t) - 1_A * 1_A * 1_A(x)| < \alpha^3$$

for each $t \in \mathrm{Bohr}(\Gamma, \alpha)$. By the last part of Lemma 1.3.2 we have $\mathbb{E}_x 1_A * 1_A * 1_A(x) = \alpha^3$; hence there is an element $x \in G$ such that $1_A * 1_A * 1_A(x) \geqslant \alpha^3$. For this element we thus have, for any $t \in \mathrm{Bohr}(\Gamma, \alpha)$,

$$1_A * 1_A * 1_A(x+t) > 1_A * 1_A * 1_A(x) - \alpha^3 \geqslant 0,$$

whence $x + t \in A + A + A$. Hence $x + \mathrm{Bohr}(\Gamma, \alpha) \subseteq A + A + A$. □

We again obtain the following immediate corollaries from the results of Section 1.4.

Corollary 3.4. *Let N be a prime and suppose $A \subseteq \mathbb{Z}_N$ has size αN. Then $A + A + A$ contains an arithmetic progression P of length*

$$|P| \geqslant \tfrac{1}{2\pi}\alpha N^{\alpha^3/4}.$$

Corollary 3.5. *Let p be a prime and suppose $A \subseteq \mathbb{F}_p^n$ has size αp^n. Then $A + A + A$ contains a translate of a subspace of dimension at least $n - 4/\alpha^3$.*

Exercise 3.6. Deduce Theorem 1.1 from Corollary 3.4.

A number of improvements have been made to these results over the years, but it is still not known what the optimal results are in terms of bounds. If we have time, we shall discuss this at the end of the course.

Exercise 3.7. What kinds of variants of Lemma 3.1 can you prove for $f_1 * f_2 * f_3$ with three functions f_1, f_2, f_3 instead of just a single f?

Exercise 3.8. Formulate a version of Lemma 3.1 for functions $f : G \to \mathbb{C}$. What conditions on f are natural?

4. Notes and references

The ideas behind the above results have their ideas in fairly classical work on almost-periodicity, first studied by Harald Bohr. For some background on Bogolyubov's lemma see §4.6 of Tao–Vu. Ruzsa is credited with bringing out Bogolyubov's work [**1**], which was related to an elementary proof of Bohr's approximation theorem for almost-periodic functions, into the context of the results presented above and in particular to the result described above as Bogolyubov's lemma. The results on $3A$ are essentially due to Freiman, Halberstam and Ruzsa [**5**], who again acknowledged Bogolyubov's work. Incidentally, this latter paper seems to have inspired Bourgain's initial work on arithmetic progressions in $A + A$, the topic to which we turn in the next part.

The bounds in the results on structures in $3A$ can be improved, at least as far as the dependence on the density in concerned. For $3A$, it is (with modern eyes) a fairly straightforward matter to use Chang's theorem [**14**, Lemma 4.36] to improve the exponent α^3 to $\alpha^{2+\epsilon}$; one can also show this directly from the results of the next part of these notes. The recent paper [**7**] of Henriot contains a useful summary of results on $3A$ and $A + B + C$, setting into context the earlier papers [**6, 10, 3**]. There is a huge gap between the best known lower bounds and the best known upper bound, which is something like $N^{C/\log(1/\alpha)}$ [**5**].

For $2A - 2A$ a vastly better density dependence is known. One of the most recent breakthroughs is due to Sanders, who in [**11**] proved a result of the form of Theorem 2.1 with something like $C(\log 1/\alpha)^4$ in place of the $1/\alpha^2$, using ideas on almost-periodicity from [**4**]. See the survey [**12**] of Sanders for a detailed history, particularly related to Freiman's theorem on the classification of sets with small sumset—Question 4 of the introduction. This is currently a hot topic of research, with one of the ultimate goals relating to the so-called Polynomial Freiman-Ruzsa conjecture, which would be implied by something along the lines of one being able to replace the $1/\alpha^2$ in Theorem 2.1 with $C\log(1/\alpha)$.

PART 3

Structures in sumsets: $A + A$

1. Strategy

We turn now to single sumsets $A + A$, which are, from our perspective, slightly more subtle. Our results for $3A$ in the previous section depended in a simple way on a certain almost-periodicity result, which said that functions of the form $f * f * f$ have lots of almost-periods: shifts t for which $f * f * f(x + t) \approx f * f * f(x)$ for all x; we then applied this with $f = 1_A$. Such a result is no longer true with good bounds for single convolutions $f * f$, but we shall nevertheless be able to prove the following.

Theorem 1.1. *Let A be a subset of $[1, N]$ of size at least αN. Then $A + A$ contains an arithmetic progression of length at least*

$$\exp\left(c(\alpha^2 \log N)^{1/2} - \log\log N\right),$$

where $c > 0$ is some absolute constant[1].

Note that this is smaller than $N^{s(\alpha)}$ for any fixed α, but it is also much, much larger than any fixed power of $\log N$. (Take logs!)

But if the almost-periodicity result corresponding to Lemma 2.3.1 fails for $f * f$, then what can we do? It turns out we can prove a (in some sense) weaker but still useful type of result.

Theorem 1.2 (Fourier-based L^p-almost-periodicity)**.** *Let $p \geqslant 2$ and $\epsilon \in (0, 1)$ be parameters. Let G be a finite abelian group and suppose $f : G \to \mathbb{C}$ is a function. Then there is a Bohr set $T \subseteq G$ of rank at most Cp/ϵ^2 and radius $c\epsilon$ such that*

$$\|f(x + t) - f(x)\|_{L^p(x)} \leqslant \epsilon \|\widehat{f}\|_{\ell^1} \text{ for each } t \in T.$$

Note that we have stated this for an arbitrary function f and not a convolution, but one needs $\|\widehat{f}\|_{\ell^1} = \sum_{\gamma \in \widehat{G}} |\widehat{f}(\gamma)|$ to be small for the result to be useful, which is the case for convolutions by Parseval's identity.

The expression on the left-hand side is simply

$$(\mathbb{E}_{x \in G} |f(x + t) - f(x)|^p)^{1/p},$$

which, as $p \to \infty$, tends to $\|f(\cdot + t) - f\|_\infty$, the quantity we bounded for convolutions of three functions in Lemma 2.3.1—a hint as to why the result might be useful to us.

[1] We shall have quite a lot of constants floating about whose values we do not care much about; from here on we therefore employ the convention that c and C stand for some absolute constants, but which may vary from occurrence to occurrence.

But before we say more about that, let us prove the result itself. For this we shall take a little detour into probability theory.

Exercise 1.3. Prove a result of the form of Theorem 1.2 in the case $p = 2$, by using Parseval's identity.

Exercise 1.4. By expanding the square, or otherwise, deduce a version of Bogolyubov's lemma from this $p = 2$ case, if perhaps with worse constants. Can you deduce the $f * f * f$ almost-periodicity lemma?

2. The law of large numbers: the Marcinkiewicz-Zygmund inequality

If one takes a bunch of independent samples from a numerical population, then one expects their mean (the sample mean) to be a fair approximation to the population mean, as long as the number of samples is not very small. Various versions of this principle are associated with the phrase *the law of large numbers*; we shall use the following version.

Lemma 2.1 (Marcinkiewicz-Zygmund inequality)**.** *Let $p \geqslant 2$, and suppose $X_1, \dots, X_n$ are independent, mean-zero complex-valued random variables with $\mathbb{E}|X_j|^p < \infty$. Then*

$$\mathbb{E}|\sum_{j=1}^{k} X_j|^p \leqslant (Cp)^{p/2}\, \mathbb{E}\left[\left(\sum_{j=1}^{k}|X_j|^2\right)^{p/2}\right]. \tag{2.1}$$

Exercise 2.2. What does the $p = 2$ case of this inequality say? Prove that one actually has equality in this case, with no factor $(Cp)^{p/2}$: *the variance of a sum of independent random variables is the sum of the variances.* (Our proof of the general case will follow a natural argument for this quite closely.) Lemma 2.1 says that this is still true up to some constant even for higher moments/L^p-norms.

How does this link with the law of large numbers? Let us assume that $|X_j| \leqslant 1$. If we divide throughout by k^p, and assume that k is large enough, we see that the 'sample mean' $\mathbb{E}_{j\in[k]}X_j$ is pretty close (in L^p expectation) to its actual mean (0): the error is at most $(Cp/k)^{p/2}$, which is small provided our number of samples k is large enough.

To prove this inequality, which is relatively central to our argument, we begin with the following special case in which the random variables are assumed to be symmetric. For discrete random variables, as in our applications, this just means that $\mathbb{P}(X = x) = \mathbb{P}(X = -x)$ for all x. Note that this automatically implies that $\mathbb{E}X = 0$.

Lemma 2.3 (Khintchine's inequality)**.** *Let $p \geqslant 2$, and suppose $X_1, \dots, X_n$ are independent, symmetric (real-valued) random variables with $\mathbb{E}|X_j|^p < \infty$. Then*

$$\mathbb{E}|\sum_{j=1}^{k} X_j|^p \leqslant (Cp)^{p/2}\mathbb{E}\left[\left(\sum_{j=1}^{k} X_j^2\right)^{p/2}\right].$$

Remark 2.4. This is usually stated for the case when $X_j = \pm c_j$ with equal probability; the version above is more convenient for us and follows directly from Khintchine's proof.

Proof. At the cost of increasing the constant C in the conclusion slightly, we may assume that $p = 2m$ is an even integer, since if $p \leqslant q$ we have $(\mathbb{E}|X|^p)^{1/p} \leqslant (\mathbb{E}|X|^q)^{1/q}$ (this is the nesting of L^p-norms; cf Jensen's inequality)[2].
We may then expand the left-hand side combinatorially: by the multinomial theorem

$$\mathbb{E}\left(\sum_{j\in[k]} X_j\right)^{2m} = \mathbb{E}\sum_{\substack{r_1,\dots,r_k\geqslant 0\\ r_1+\cdots+r_k=2m}} \binom{2m}{r_1,\dots,r_k} X_1^{r_1}\cdots X_k^{r_k}, \tag{2.2}$$

the multinomial coefficient being

$$\binom{2m}{r_1,\dots,r_k} = \frac{(2m)!}{r_1!\cdots r_k!}.$$

By the linearity of expectation we may take the $\mathbb{E}$ inside the sum, and since the X_j are independent we have $\mathbb{E}X_1^{r_1}\cdots X_k^{r_k} = (\mathbb{E}X_1^{r_1})\cdots(\mathbb{E}X_k^{r_k})$. But if some r_j is odd then, since the variables are symmetric, $\mathbb{E}X_j^{r_j} = 0$. Thus we may restrict the summation in (2.2) to where all the r_j are even, say $r_j = 2s_j$, and so

$$\begin{aligned}\mathbb{E}\left(\sum_{j\in[k]} X_j\right)^{2m} &= \sum_{s_1+\cdots+s_k=m} \binom{2m}{2s_1,\dots,2s_k} (\mathbb{E}X_1^{2s_1})\cdots(\mathbb{E}X_k^{2s_k})\\ &\leqslant \frac{(2m)!}{2^m m!}\sum_{s_1+\cdots+s_k=m} \binom{m}{s_1,\dots,s_k} (\mathbb{E}X_1^{2s_1})\cdots(\mathbb{E}X_k^{2s_k})\\ &= \frac{(2m)!}{2^m m!}\,\mathbb{E}\left(\sum_{j\in[k]} X_j^2\right)^m,\end{aligned}$$

the inequality being a consequence of the trivial bound $(2s_j)! \geqslant 2^{s_j} s_j!$. Since $(2m)!/2^m m! \leqslant m^m$, we are done. □

Exercise 2.5. What does this say when $X_j = c_j$ and $X_j = -c_j$ each with probability $1/2$, for constants $c_j \in \mathbb{R}$? (This is the traditional setting of Khintchine's inequality.)

Of course, not all random variables are symmetric—in our application it certainly need not be the case. However, it turns out that we can deduce the full Marcinkiewicz-Zygmund inequality from the above special case by the very useful trick of *symmetrisation*. The basic idea behind this is very simple: if one has a random variable X then, letting X' be an independent copy of X, the random variable $X - X'$ is a symmetric random variable whose statistics are highly related to those of X.

Since this trick can cause some confusion, let us be completely explicit about it here, in the following lemma-with-definition.

Lemma 2.6. *For a random variable X on a sample space Ω, let the* symmetrisation $X^{(s)}$ *be the random variable on $\Omega\times\Omega$ defined by $X^{(s)}(\omega,\nu) = X(\omega) - X(\nu)$. For $p \geqslant 1$, if $\mathbb{E}X = 0$ and $\mathbb{E}|X|^p < \infty$, then,*

$$\mathbb{E}|X^{(s)}|^p \geqslant \mathbb{E}|X|^p.$$

[2]Alternatively, just assume p is an even integer in the statement and only apply it in this case.

Note that the first expectation in this inequality is with respect to the product measure on $\Omega \times \Omega$, whereas the second one is with respect to the measure on Ω itself.

Proof. By the mean-zero hypothesis, we have

$$\mathbb{E}_\omega |X(\omega)|^p = \mathbb{E}_\omega |X(\omega) - \mathbb{E}_\nu X(\nu)|^p = \mathbb{E}_\omega |\mathbb{E}_\nu X(\omega) - X(\nu)|^p \leqslant \mathbb{E}_\omega \mathbb{E}_\nu |X(\omega) - X(\nu)|^p,$$

the inequality being a case of Jensen's inequality. This gives the result, by Fubini. □

Proof of Lemma 2.1. The complex-valued case follows from the real-valued one by looking at real and imaginary parts, so we assume that the X_j are real-valued[3]. Using Lemma 2.6 and applying Lemma 2.3 to the symmetric random variables $X_j^{(s)}$ yields

$$\mathbb{E}|\sum_{j\in[k]} X_j|^p \leqslant \mathbb{E}|\sum_{j\in[k]} X_j^{(s)}|^p \leqslant (Cp)^{p/2}\mathbb{E}\left[\left(\sum_{j\in[k]}(X_j^{(s)})^2\right)^{p/2}\right].$$

The right-hand side cannot be much bigger than the right-hand side of (2.1): applying the inequality $(a-b)^2 \leqslant 2(a^2+b^2)$ and the triangle inequality for the $L^{p/2}$ norm gives

$$\mathbb{E}\left(\sum_{j\in[k]}(X_j(\omega)-X_j(\nu))^2\right)^{p/2} \leqslant 2^{p/2}\mathbb{E}\left(\sum_{j\in[k]} X_j(\omega)^2 + \sum_{j\in[k]} X_j(\nu)^2\right)^{p/2} \leqslant 2^p \mathbb{E}\left(\sum_{j\in[k]} X_j^2\right)^{p/2},$$

whence we are done. □

This (essentially) completes our detour into probability; next we come to applying these results.

3. Approximation by short trigonometric polynomials and L^p-almost-periodicity

As we shall see, a natural way to prove Theorem 1.2 is by approximating the function f by a short trigonometric polynomial, in the following sense.

Theorem 3.1 (Short trigonometric approximation)**.**
Let $p \geqslant 2$ and $\epsilon > 0$. Let $f : G \to \mathbb{C}$ be normalised so that $\|\widehat{f}\|_{\ell^1} = 1$. Then there are characters $\gamma_1, \dots, \gamma_k \in \widehat{G}$ and coefficients $c_1, \dots, c_k \in \mathbb{C}$ with $|c_j| = 1$, where $k \leqslant Cp/\epsilon^2$, such that

$$\|f - \frac{1}{k}\sum_{j\in[k]} c_j\gamma_j\|_p \leqslant \epsilon.$$

Proof. Write, as we may by the Fourier inversion formula,

$$f = \sum_{\gamma\in\widehat{G}} \widehat{f}(\gamma)\gamma = \sum_{\gamma\in\widehat{G}} |\widehat{f}(\gamma)| c_\gamma \gamma$$

where $|c_\gamma| = 1$. Now let $\chi : G \to \mathbb{C}^\times$ be a random 'twisted' character picked according to the distribution $\mathbb{P}(\chi = c_\gamma\gamma) = |\widehat{f}(\gamma)|$, so that

$$\mathbb{E}\chi = \sum_\gamma c_\gamma \gamma\, \mathbb{P}(\chi = c_\gamma\gamma) = f.$$

[3]One can also modify the proof of Lemma 2.3 to work directly with complex-valued random variables.

Let $\chi_1, \dots, \chi_k$ be independent samples according to this distribution for some k to be specified. Then, by the Marcinkiewicz-Zygmund inequality, Lemma 2.1,

$$\begin{aligned}\mathbb{E}\,\|\mathbb{E}_{j\in[k]}\chi_j - f\|_p^p &= \mathbb{E}_{x\in G}\,\mathbb{E}\,|\mathbb{E}_{j\in[k]}\chi_j(x) - f(x)|^p \\ &\leqslant \left(\frac{Cp}{k}\right)^{p/2} \mathbb{E}_{x\in G}\,\mathbb{E}\left[\left(\mathbb{E}_{j\in[k]}|\chi_j(x) - f(x)|^2\right)^{p/2}\right] \\ &\leqslant \left(\frac{Cp}{k}\right)^{p/2},\end{aligned}$$

the last inequality following from the fact that $|\chi_j(x)|, |f(x)| \leqslant 1$. Pick $k = \lceil Cp/\epsilon^2 \rceil$ so that the final bound is at most ϵ^p. Since the expectation over all twisted characters is this small, certainly there is some choice of $\chi_1 = c_{\gamma_1}\gamma_1$, …, $\chi_k = c_{\gamma_k}\gamma_k$ such that $\|\mathbb{E}_{j\in[k]}\chi_j - f\|_p^p \leqslant \epsilon^p$, and we are done. □

Thus, up to a small error in L^p, we may approximate any f with $\|\widehat{f}\|_{\ell^1}$ small by a short trigonometric polynomial[4]. Indeed, we can do so with a number of characters that is in a sense *independent of* $|G|$, whereas a priori one needs $|G|$ characters to express f exactly. Since such short sums are actually L^∞-almost-periodic, this leads very quickly to Theorem 1.2, which we now recall.

Theorem 1.2. *Let $p \geqslant 2$ and $\epsilon \in (0,1)$. For any $f : G \to \mathbb{C}$ there is a Bohr set $T \subseteq G$ of rank at most Cp/ϵ^2 and radius $c\epsilon$ such that*

$$\|f(x+t) - f(x)\|_{L^p(x)} \leqslant \epsilon\|\widehat{f}\|_{\ell^1} \text{ for each } t \in T.$$

Proof. Apply Theorem 3.1 to $f' := f/\|\widehat{f}\|_{\ell^1}$ with parameters p and $\epsilon/3$ to get a set of characters $\Gamma \subseteq \widehat{G}$ of size at most Cp/ϵ^2 and coefficients c_j of absolute value 1 such that the function $g = \mathbb{E}_{\gamma\in\Gamma}\, c_\gamma\gamma$ approximates f' well: $\|f'-g\|_p \leqslant \frac{1}{3}\epsilon$. For any $t \in T := \mathrm{Bohr}(\Gamma, \epsilon/3)$ we then have, for any $x \in G$,

$$|g(x+t) - g(x)| = |\mathbb{E}_{\gamma\in\Gamma}\, c_\gamma\gamma(x)(\gamma(t)-1)| \leqslant \mathbb{E}_{\gamma\in\Gamma}\,|\gamma(t)-1| \leqslant \epsilon/3,$$

and so

$$\|f'(\cdot+t) - f'\|_p \leqslant \|f'(\cdot+t) - g(\cdot+t)\|_p + \|g(\cdot+t) - g\|_p + \|g - f'\|_p \leqslant \epsilon. \qquad \square$$

4. Long progressions in $A + A$

It is now a short matter to deduce Theorem 1.1. We prove the following variant for cyclic groups, and leave it as an exercise to deduce the version for integers from this.

Theorem 4.1. *Let N be a prime and let $A \subseteq \mathbb{Z}_N$ be a set of size at least αN. Then $A + A$ contains an arithmetic progression of length at least*

$$\exp\left(c(\alpha^2 \log N)^{1/2} - \log\log N\right).$$

[4]Why polynomial? In the case of $G = \mathbb{Z}_N$ all the characters are of the form γ^n for some fixed γ.

Proof. Apply Theorem 1.2 to $f := 1_A * 1_A$ with parameters p and ϵ to be determined. We get a Bohr set T of rank at most Cp/ϵ^2 and radius $c\epsilon$ such that, for each $t \in T$,

$$\|f(\cdot + t) - f\|_p \leqslant \epsilon \|\widehat{f}\|_{\ell^1} = \epsilon \sum_{\gamma \in \widehat{G}} |\widehat{1_A}(\gamma)|^2 = \epsilon\alpha,$$

where we have used the convolution identity and Parseval's identity. Let $P \subseteq T$ be an arithmetic progression—we shall attempt to find a translate of this in $A + A$. We have

$$\begin{aligned}\mathbb{E}_{x\in G} \sup_{t\in P} |f(x+t) - f(x)| &\leqslant \mathbb{E}_{x\in G} \left(\sum_{t\in P} |f(x+t) - f(x)|^p \right)^{1/p} \\ &\leqslant \left(\sum_{t\in P} \mathbb{E}_x |f(x+t) - f(x)|^p \right)^{1/p} \\ &\leqslant \epsilon\alpha |P|^{1/p}.\end{aligned}$$

If we can show that this is less than $\alpha^2 = \mathbb{E}_{x\in G} f(x)$, then we can be happy, as there then exists some $x \in G$ such that $|f(x + t) - f(x)| < f(x)$ for all $t \in P$, whence $f(x + t) > 0$ and so $x + P \subseteq \operatorname{supp} f = A + A$. So now we just have to optimise the parameters to make this the case. We pick $\epsilon = \alpha/2$, and so all we need is for $|P|$ to be at most 2^p. On the other hand, we can find an arithmetic progression in T of any length up to $c\alpha N^{c\alpha^2/p}$, by Proposition 1.4.7. If α is very small, the result is trivial; otherwise we pick $p = C\alpha\sqrt{\log N}$ which then gives the desired bound. □

Remark 4.2. So how did we use the fact that we had L^p-almost-periods for a large p? We did indeed do an $L^p \longleftrightarrow L^\infty$ approximation as hinted at earlier, though not over the whole group, but rather on a much smaller scale, namely over (the relatively small set) P.

Exercise 4.3. Show that in $\mathbb{F}_p^n$, $A + A$ contains a translate of a subspace of dimension at least $c\alpha^2 n$.

Exercise 4.4. Can you extend the results to $A + B$?

5. Notes and references

The first result of the form of Theorem 1.1 was proved by Bourgain [**2**], who essentially proved the result with an exponent $1/3$ instead of $1/2$. Green [**6**] subsequently proved Theorem 1.1 itself. The proof given above follows [**3**], in which a somewhat better density dependence is also proved. An exposition of Bourgain's argument with a focus on almost-periodicity can be found in [**13**].

For a fixed density α, say $1/1000$, it is not currently known whether the kind of bound in Theorem 1.1 is best possible or not. By way of upper bounds, there is a remarkable construction due to Ruzsa [**9**] which establishes that one cannot replace the exponent $1/2$ in the theorem by anything larger than $2/3$, even if one makes A a symmetric subset of $\mathbb{Z}_N$ of size almost $N/2$. This is particularly striking, since as soon as $|A| > N/2$ one has $A + A = \mathbb{Z}_N$.

The result on short approximations can be vastly generalised; the proof given above essentially constitutes a proof of Maurey's lemma in Banach space theory [**8**]. This perspective unites parts of the arguments from [**3**] and [**4**]; see [**3**] for something of a discussion on this.

There is a lot more that could be said here, but we did not have time for much more during the course, and it is probably a good idea to keep these notes somewhat focused. Many of the topics we have touched on in the course lead on to active areas of research; the course content will hopefully allow anyone who wants to get up to speed with the latest developments in these areas to do so.

Bibliography

[1] N. N. Bogolioùboff, *Sur quelques propriétés arithmétiques des presque-périodes*, Ann. Chaire Math. Phys. Kiev **4** (1939), 185–194.

[2] J. Bourgain, *On arithmetic progressions in sums of sets of integers*, A tribute to Paul Erdős, 105–109 (CUP, 1990).

[3] E. Croot, I. Łaba and O. Sisask, *Arithmetic progressions in sumsets and L^p-almost-periodicity*, Combin. Probab. Comput. **22** (2013), no. (3), 351–365.
http://arxiv.org/abs/1103.6000

[4] E. Croot and O. Sisask, *A probabilistic technique for finding almost-periods of convolutions*, Geom. Funct. Anal. **20** (2010), no. 6, 1367–1396.
http://arxiv.org/abs/1003.2978

[5] G.A. Freiman, H. Halberstam, I. Ruzsa, *Integer sum sets containing long arithmetic progressions*, Jour. London Math. Soc. **46** (1992), no. 2, 193–201.
At the time of writing, can be downloaded for free at
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.101.3922&rep=rep1&type=pdf.

[6] B. Green, *Arithmetic progressions in sumsets*, Geom. Funct. Anal. **12** (2002), no. 3, 584–597.
http://people.maths.ox.ac.uk/greenbj/papers/arithmetic-progressions-in-sumsets.pdf

[7] K. Henriot, *On Arithmetic Progressions in $A + B + C$*, Int. Math. Res. Not., first published online June 2013, doi:10.1093/imrn/rnt121.
http://arxiv.org/abs/1211.4917

[8] G. Pisier, *Remarques sur un résultat non publié de B. Maurey*, Seminar on Functional Analysis, 1980–1981, Exp. No. V, 13, École Polytech., Palaiseau, 1981.
http://www.numdam.org/item?id=SAF_1980-1981____A5_0

[9] I. Z. Ruzsa, *Arithmetic progressions in sumsets*, Acta Arith. **60** (1991), no. 2, 191–202.
At the time of writing, can be downloaded for free at
http://matwbn.icm.edu.pl/ksiazki/aa/aa60/aa6027.pdf.

[10] T. Sanders, *Additive structures in sumsets*, Math. Proc. Cambridge Philos. Soc. **144** (2008), no. 2, 289–316.
http://arxiv.org/abs/math/0605520

[11] ———, *On the Bogolyubov–Ruzsa lemma*, Anal. PDE **5** (2012), no. 3, 627–655.
http://arxiv.org/abs/1011.0107

[12] ———, *The structure theory of set addition revisited*, Bull. Amer. Math. Soc. **50** (2013), 93–127.
http://arXiv.org/abs/1212.0458

[13] O. Sisask, *Bourgain's proof of the existence of long arithmetic progressions in $A + B$*, available at
http://www.math.kth.se/~sisask/

[14] T. Tao and V. H. Vu, *Additive Combinatorics* (Cambridge University Press, 2006).