Additive Combinatorics

Harry Armitage

March 2, 2021

Contents

0 Elementary Tools **2**

1 Sum Sets **2**
1.1 Sumset Calculus 4
1.2 Additive Energy 6
1.3 Covering 11
1.4 Span-Covering 13
1.5 Finding Arithmetic Progressions 14

2 Fourier Analysis **15**
2.1 Basic Concepts 15
2.2 Fourier Transform 16
2.3 Roth's Theorem in $\mathbb{F}_{p^n}$ 17
2.4 Bohr Sets under Addition 23
2.5 Roth's Theorem in the Integers. 26

3 Almost-Periodicity **30**
3.1 Applications of Almost-Periodicity 32
3.2 A Probabilistic Inequality 34
3.3 Almost-Periodicity via Random Sampling in Fourier Space 36

0 Elementary Tools

Asymptotic notation: for functions f, g we will write $f = O(g)$ to mean there is a constant $c > 0$ such that, for large enough x, $|f(x)| \leq c|g(x)|$. Sometimes, we will write $f = O_h(g)$ if c depends on h. We will also sometimes write $f \ll g$ to mean $f = O(g)$ (or indeed $g \gg f$, or $f \ll_h g$). We will also write such things as $(x+h)^2 = x^2 + O_h(x)$. We will often write $\log X$ - it will typically be irrelevant which base, but it will be assumed that X is large.

All sets will be finite and nonempty unless otherwise specified. Usually, they will be subsets of some abelian group, denoted by G. If finite, we will write $N = |G|$, for instance $G = \mathbb{Z}/N\mathbb{Z}$ or $\mathbb{F}_p^n$ (where $N = p^n$). Much of what we do is valid for any abelian group, some of it only for finite groups, and some only for specific groups. Often, it can be generalised and sometimes not - e.g. $\mathbb{F}_p^n$ vs $\mathbb{Z}/N\mathbb{Z}$.

We will write $\mathbb{1}_A(x)$ for the indicator function on a subset $A \subseteq G$, i.e. $\mathbb{1}_A(x) = 1$ if $x \in A$, and 0 if not. We also define the convolution of functions $f, g : G \to G$ by

$$(f * g)(x) = \sum_{y \in G} f(y)g(x-y) = \sum_{y+z=x} f(y)g(z)$$

We define the "difference convolution" by

$$(f \circ g)(x) = \sum_{z \in G} f(x+z)g(z) = \sum_{y-z=x} f(y)g(z)$$

We define an inner product on function $G \to \mathbb{C}$ by

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

We have the trivial but useful adjoint property that

$$\langle f * g, h \rangle = \langle f, h \circ \bar{g} \rangle$$

which follows from $x + y = z \iff x = z - y$.

1 Sum Sets

Given $A, B \subset G$ we define

$$A + B := \{a + b : a \in A, b \in B\}$$
$$A - B := \{a - b : a \in A, b \in B\}$$

Note that these are set operations, and don't necessarily follow any nice algebraic properties. For instance $(A + B) - B \neq A$, and in general is a much larger set.

We have trivial bounds $|A| \leq |A + B| \leq |A||B|$ - the first follows as, for fixed b, $\{a + b : a \in A\}$ is a set of size $|A|$ contained in $A + B$, and the second follows as $+$ is a surjection $A \times B \twoheadrightarrow A + B$.

When $A = B$, this surjection maps (a_1, a_2) and (a_2, a_1) to the same element, and hence $|A + A| \leq \frac{|A|(|A|+1)}{2} = \frac{1}{2}|A|^2 + O(|A|)$.

Are these trivial bounds sharp?

If A is a subgroup, then $|A + A| = |A|$ (or if A is a coset of a subgroup).

If $A = \{1, 2, 4, \ldots, 2^k\}$, then $|A + A| = \frac{|A|(|A|+1)}{2}$, as all pairwise sums are distinct. We can also have $A = \{1, 4, 16, \ldots, 2^{2k}\}$, $B = \{2, 8, \ldots, 2^{2k+1}\}$, and get $|A + B| = |A||B|$.

Lemma 1.1. $|A + A| \leq |A| \iff$ *A is a coset of a subgroup.*

Proof. Since both properties are invariant under translation, without loss of generality take $0 \in A$. So $A \subseteq A + A$, so since $|A + A| = |A|$, we must have $A = A + A$. Hence A is closed under addition. For any $a \in A$ there are $|A|$ many distinct translates $a + a'$ for $a' \in A$. By closure, these are all in A, and so one must be 0. Hence $a + a' = 0$, and we have inverses. □

In groups where there are many finite subgroups, we have many A such that $|A + A| = |A|$. What about in $\mathbb{Z}$? Here, there are no nontrivial finite subgroups, and so this result tells us that, if $|A| > 1$, then $|A + A| > |A|$.

Lemma 1.2. *If $A \subset \mathbb{Z}$ then $|A + A| \geq 2|A| - 1$. Equality holds if and only if A is an arithmetic progression (i.e. $|A + A| < 2|A| \implies$ A is an arithmetic progression).*

Proof. Suppose we order A like $\{a_1 < a_2 < \ldots a_n\}$. This induces an ordering on some elements of $A + A$:

$$2a_2 < a_1 + a_2 < a_1 + a_3 < \ldots < a_1 + a_n < a_2 + a_n < a_3 + a_n < \ldots < 2a_n$$

This gives $2n - 1$ distinct sums in $A + A$.

For the second part, it suffices to show that if equality holds then, for any $1 \leq i < n$ there is some $j > i$ with $a_j - a_i = a_2 - a_1$, as then $a_j = a_1 + (j - 1)(a_2 - a_1)$, and we have an arithmetic progression.

Consider for $2 \leq i < n$, the sum $a_2 + a_i \in A + A$. Since $|A + A| = 2n - 1$, the above ordered list is all of $A + A$, and so one of them is $a_2 + a_i$. Since $i < n$, $a_2 + a_i < a_j + a_n$ for all $2 \leq j \leq n$, and so $a_2 + a_i = a_1 + a_j$. □

What about if $A \subseteq \mathbb{F}_p$, for p a prime?

Lemma 1.3 (Cauchy-Davenport)**.** *If $A, B \subset \mathbb{F}_p$, then $|A + B| \geq \min(|A| + |B| - 1, p)$.*

Proof. We fix $B \subset \mathbb{F}_p$, and then prove the inequality for all $A \subset \mathbb{F}_p$ by induction on $|B|$.

When $|B| = 1$, this is trivial, since $|A + B| = |A| = \min(|A|, p)$.

Suppose that $|B| \geq 2$, so we have $b_1 \neq b_2$ distinct elements of B, and let $z = b_1 - b_2 \neq 0$. If $A + z \subset A$ then, for any fixed $a \in A$, we have $a + z \in A$, so $a + kz \in A$ for all $k \in \mathbb{N}$ by induction on k. Hence $A = \mathbb{F}_p$, as $\mathbb{F}_p$ is cyclic, and so we are done as $|A + B| = p = \min(|A| + |B| - 1, p)$.

Otherwise, $A + z \not\subseteq A$, and so there exists $a \in A$ such that $a + z \notin A$. Let $x = a - b_1$. Then $B + x$ contains some element not in A, namely $a + z = b_2 + x$, and some element in A, namely $a = b_1 + x$.

So $1 \leq |A \cap (B + x)| < |B|$. Now note that:

$$A + B \supseteq ((A - x) \cup B) + (A \cap (B + x))$$

To verify this, if $a' + b' \in RHS$, then either:

1. $a' \in A - x$. Then since $b' \in B + x$, $a' + b' \in A + B - x + x = A + B$.
2. $a' \in B$. Then since $b' \in A$, $a' + b' \in A + B$.

Hence if $A' = (A - x) \cup B$, $B' = A \cap (B + x)$, then $1 \leq |B'| < |B|$, $|A' + B'| \leq |A + B|$. By induction, $|A + B| \geq \min(|A'| + |B'| - 1, p)$.

But $|A'| = |(A-x) \cup B| = |A-x| + |B| - |(A-x) \cap B| = |A| + |B| - |B'|$. Hence $|A| + |B| = |A'| + |B'|$. □

This trick, transforming from $(A, B) \to ((A - x) \cup B, A \cap (B + x))$, is sometimes called the ***Dyson e-transform***.

We can characterise the case of equality as APs: if $|A + B| = |A| + |B| - 1$ then, aside from edge cases, A and B must be APs of the same step seize ("Vosper's Theorem").

We say that A has ***small doubling*** if $|A + A| \leq K|A|$ where K is "small" (e.g. $O(1)$). We've seen two examples of sets with small doubling:

- Cosets of subgroups ($K = 1$).
- Arithmetic progressions ($K = 2 - \frac{1}{|A|} = 2 + O(1)$).

There are two ways to generate new sets with small doubling from old.

1. <u>Pass to a large subset.</u> If $|A + A| \leq K|A|$ and $X \subset A$, then $|X + X| \leq |A + A| \leq \left(K\frac{|A|}{|X|}\right)|X|$. If $|X| \geq K^{-O(1)}|A|$, then A has doubling K implies X has doubling at most $K^{O(1)}$.
2. <u>Pass to a sumset.</u> If $|A + A| \leq K|A|$ and $X = A + Y$, then $|X + X| \leq |A + A + Y + Y| \leq |Y|^2|A + A| \leq (K|Y|^2)|X|$. So if A has doubling K and $X = A + Y$ where $|Y| \leq K^{O(1)}$, then X has doubling $\leq K^{O(1)}$. In fact it suffices if $|Y + Y| \leq K^{O(1)}|Y|$ and $|X| \geq K^{-O(1)}|A||Y|$.

This accounts for all sets with small doubling. That is, if $|A + A| \leq 100|A|$, then A is obtained from a coset or AP via these two operations. We'll state this more precisely and prove it later on.

1.1 Sumset Calculus

How are the sizes of sumsets related to each other?

Lemma 1.4 (Ruzsa's Triangle Inequality)**.** *For any sets A, B, C, $|A + B| \leq \frac{|A+C||B-C|}{|C|}$.*

Proof. For each $x = A + B$, fix an arbitrary representation $x = a_x + b_x$. Consider the map $C \times (A + B) \to (A + C) \times (B - C)$ given by $(c, x) \mapsto (c + a_x, b_x - c)$. This is clearly well defined - to finish the proof we will show that it is injective.

Note that we can recover x from the image of (c, x) by noting $(c + a_x) + (b_x - c) = a_x + b_x = x$. Then we can also recover c by subtracting a_x from $c + a_x$, and hence this map is injective. □

What about iterated sumsets? We write kA for the k-fold sumset of A, i.e. the set of all sums $a_1 + \ldots + a_k$ where $a_i \in A$.

Note that $|A| \leq |kA| \leq |A|^k$. This does not mean A dilated by k!

If $|A$ is small, we might want to know whether $|A + A|$ being small implies that $|kA|$ are also relatively small? The answer turns out to be yes! This is known as Plünnecke's inequality, which says that if $|A+A| \leq K|A|$, then $|kA| \leq K^k|A|$. The original proof by Plünnecke is graph-theoretic - it can be found in Tao & Vu. More recently, Petridis has found an alternative proof which is presented here. It uses the following lemma:

Lemma 1.5 (Plünnecke-Petridis Inequality)**.** *For any A, B there exists $X \subseteq A$ such that, for all C*

$$\frac{|C + X + B|}{|C + X|} \leq \frac{|A + B|}{|A|} \tag{$*$}$$

Proof. Note that $(*)$ holds whenever $|C| = 1$, and hence

$$\frac{|X + B|}{|X|} \leq \frac{|A + B|}{|A|} \tag{$**$}$$

We will chose $X \subseteq A$ such that $(**)$ holds and $\frac{|X+B|}{|X|}$ is minimal. Note that $X = A$ satisfies $(**)$, and so this choice makes sense.

We then show that $(*)$ holds for all C by induction on $|C|$. The case $|C| = 1$ is immediate.

Now suppose that $|C| > 1$, and choose some $c \in C$, and let $C' = C \setminus \{c\}$. Then:

$$C + X = (C' + X) \sqcup (c + X')$$

for some $X' \subseteq X$, maximal such that $C' + X$ and $c + X'$ are disjoint. By maximality, $c + (X \setminus X') \subseteq C' + X$, and so $c + (X \setminus X') + B \subseteq C' + X + B$. Hence:

$$C + X + B = (C' + X + B) \cup ((C + X + B) \setminus (c + X \setminus X' + B))$$

Taking cardinalities of both sides,

$$\begin{aligned}
|C + X + B| &\leq |C' + X + B| + |c + X + B| - |c + X \setminus X' + B| \\
&= |C' + X + B| + |X + B| - |X \setminus X' + B| \\
&\leq \frac{|A + B|}{|A|}|C' + X| + |X + B| - \frac{|X + B|}{|X|}|X \setminus X'|
\end{aligned}$$

using the inductive hypothesis and the fact that $\frac{|X\setminus X'+B|}{|X\setminus X'|} \geq \frac{|X+B|}{|X|}$ by minimality. Continuing, we have:

$$\begin{aligned}
|C + X + B| &\leq \frac{|A + B|}{|A|}|C' + X| + |X + B|\left(1 - \frac{|X \setminus X'}{|X|}\right) \\
&= \frac{|A + B|}{|A|}|C' + X| + \frac{|X + B|}{|X|}|X'| \\
&\leq \frac{|A + B|}{|A|}(|C' + X| + |X'|) \\
&= \frac{|A + B|}{|A|}|C + X|
\end{aligned}$$

□

Corollary 1.6 (Plünnecke's Inequality)**.** *If $|A+B| \leq K|A|$, then there exists $X \subseteq A$ such that, for all $k \geq 1$,*

$$|X + kB| \leq K^k|X|$$

In particular, $|kB| \leq K^k|A|$.

Proof. This follows immediately from the Plünnecke-Petridis Inequality by induction on k. For $k = 1$, we have $|X + B| \leq K|X|$, by taking any C with $|C| = 1$. In general, take $C = (k-1)B$, and then

$$|X + kB| = |C + X + B| \leq K|X + C| = K|X + (k-1)B| \leq K \cdot K^{k-1}|X| = K^k|X|$$

by induction. The second claim follows since trivially,

$$|kB| \leq |X + kB| \text{ and } K^k|X| \leq K^k|A|$$

□

By coupling this with Ruzsa's triangle Inequality (RTI), we get a form for mixed sums and differences.

Corollary 1.7. *Suppose $|A+B| \leq K|A|$. For any $k, \ell \in \mathbb{N}$ with $k + \ell \geq 2$,*

$$|kB - \ell B| \leq K^{k+\ell}|A|$$

Proof. By RTI with X as above,

$$|kB - \ell B| \leq \frac{|X + kB||X + \ell B|}{|X|} \leq \frac{K^k|X|K^\ell|X|}{|X|} = K^{k+\ell}|X| \leq K^{k+\ell}|A|$$

□

1.2 Additive Energy

The size of the sumset is one way to measure the 'structure' of A, but not the only way. One advantage of using $\frac{|A+A|}{|A|}$ is that it controls the same quantity for subsets automatically: i.e., if $X \subset A$, then $|X + X| \leq |A + A|$. However, it has the disadvantage that $\frac{|A+A|}{|A|}$ can grow out of control quickly if we add in a small number of elements. It's often more convenient to use "additive energy" to measuer structure.

Definition 1.8. *The* **additive energy** *of A, $E(A)$ is defined to be*

$$E(A) = \#\{(a, b, c, d) \in A^4 : a + b = c + d\}$$

Note that while small sumsets occur in 'structured' sets, here this corresponds to large additive energy. For example, if A is an arithmetic progression, then $E(A) \approx |A|^3$.

If A is a geometric progression, then $E(A) \approx |A|^2$, as $2^k + 2^\ell = 2^r + 2^s \iff \{k, \ell\} = \{r, s\}$.

We have an alternate definition of $E(A)$ - note that

$$\begin{aligned} E(A) &= \sum_{a,b\in A}\sum_{c,d\in A}\sum_{x} \mathbb{1}_{a+b=x}\mathbb{1}_{c+d=x} \\ &= \sum_{x}\left(\sum_{a,b\in A}\mathbb{1}_{a+b=x}\right)^2 \\ &= \sum_{x}\mathbb{1}_A * \mathbb{1}_A(x)^2 \\ &= \|\mathbb{1}_A * \mathbb{1}_A\|_2^2 \end{aligned}$$

(recall that the L^p norm of $f : G \to \mathbb{C}$ is defined by $\|f\|_p = (\sum_x |f(x)|^p)^{1/p}$, and $\|f\|_\infty = \sum_{x\in G}|f(x)|$).

Its status as an L^2-norm explains why $E(A)$ is so useful in analytic arguments. Also note that the size of the sumset $|A+A|$ is the size of the support of $\mathbb{1}_A * \mathbb{1}_A$, sometimes called the "$L^0$-norm". Note that the L^1-norm of A is $\sum_{a,b\in A}\sum_x \mathbb{1}_{a+b=x} = \sum_{a,b\in A} 1 = |A|^2$.

It should also be noted that

$$\begin{aligned} E(A) &= \#\{(a,b,c,d)\in A^4 : a-c = b-d\} \\ &= \sum_x \mathbb{1}_A \circ \mathbb{1}_A(x)^2 = \|\mathbb{1}_A \circ \mathbb{1}_A\|_2^2 \end{aligned}$$

Lemma 1.9. *$|A|^2 \le E(A) \le |A|^3$ and $E(A) \ge \frac{|A|^4}{|A+A|}$.*

i.e., small sumset $\implies$ large energy.

Proof. $|A|^2 \le E(A)$ comes from counting solutions to $a+b=c+d$ where $a=c$ and $b=d$.

$E(A) \le |A|^3$ is because once we have fixed a,b,c, then either there are 1 or 0 choices for d.

To connect $E(A)$ with the size of the sumset, we use the Cauchy-Schwarz inequality, i.e.

$$\sum a_i b_i \le \left(\sum |a_i|^2\right)^{1/2}\left(\sum |b_i|^2\right)^{1/2}$$

We have

$$\begin{aligned} |A|^2 &= \sum_x \mathbb{1}_A * \mathbb{1}_A(x) \times 1 \\ &= \sum_x \mathbb{1}_A * \mathbb{1}_A(x) \times \mathbb{1}_{A+A}(x) \\ &\le \left(\sum_x \mathbb{1}_A * \mathbb{1}_A(x)^2\right)^{1/2}\left(\sum_x \mathbb{1}_{A+A}(x)^2\right)^{1/2} \\ &= E(A)^{1/2}|A+A|^{1/2} \end{aligned}$$

Similarly, we have $E(A) \ge \frac{|A|^4}{|A-A|}$. □

The converse to small sumset $\implies$ large energy does not hold. If we take A to be the union of an arithmetic progression and geometric progression of equal sizes, then $|A+A| \gg |A|^2$ (from the geometric progression), but $E(A) \gg |A|^3$ (from the arithmetic progression).

A itself is not structured much, but a large subset of A is. One might hope that this construction is the only thing that can go wrong. More precisely, that if A has large energy, then there is a large subset of A with small doubling. The answer to this hope is yes! - the Balog-Szemerédi-Gowers Lemma.

Lemma 1.10 (Balog-Szemerédi-Gowers)**.** *If $E(A) \geq K^{-1}|A|^3$ for some $K \geq 1$, then there exists $A' \subseteq A$ such that:*

1. $|A'| \gg K^{-1}|A|$
2. $|A' - A'| \ll K^7|A|$.

Note that the dependence on K is polynomial, and also that we can recover a bound for $|A'+A'|$ via the Ruzsa triangle inequality:

$$|A'+A'| \leq \frac{|A'+A'| \cdot |A'+A'|}{|A'|} \leq K^{14}|A|$$

We will use the "first moment method" - a simple application of probability which is very useful in combinatorics. It essentially says that if X is a real-valued random variable then $X \geq \mathbb{E}\, X$ with positive probability.

Proof (following Schoen). Firstly, we will find a large subset $X \subset A$ such that there are 'many' differences in $X - X$ with 'many' different representations as elements of $A - A$. Note that, if all of $A - A$ had $\geq K^{-O(1)}|A|$ representations, then $E(A) = \sum_x \mathbb{1}_A \circ \mathbb{1}_A(x)^2 \geq |A - A|K^{-O(1)}|A|^2$, and on the other hand $\leq |A|^3$, then $|A - A| \leq K^{O(1)}|A|$.

We will then find some $X' \subseteq X$ such that $|X' - X'|$ is small as required. We'll break these steps up into the following lemmas. □

Lemma 1.11. *If $E(A) \geq K^{-1}|A|^3$ then for any $0 < c < 1$ there is some $X \subseteq A$ such that $|X| \gg K^{-1}|A|$ and for all but at most $c|X|^2$ many pairs $(a, b) \in X^2$,*

$$\mathbb{1}_A \circ \mathbb{1}_A(a - b) \gg c^2K^{-3}|A|$$

Proof. The set X will be of the form $A \cap (A + s)$ for some $s \in A - A$ randomly chosen. We choose $s \in A - A$ with probability $\frac{\mathbb{1}_A \circ \mathbb{1}_A(s)}{|A|^2}$. Then:

$$\begin{aligned}
\mathbb{E}\,|X| &= \sum_s \mathbb{P}(s \text{ chosen})|A \cap (A+s)| \\
&= \sum_s \frac{\mathbb{1}_A \circ \mathbb{1}_A(s)}{|A|^2} \sum_{a \in A} \mathbb{1}_A(a - s) \\
&= \frac{1}{|A|^2} \sum_s \mathbb{1}_A \circ \mathbb{1}_A(s)^2 \\
&= \frac{E(A)}{|A|^2}
\end{aligned}$$

For any $G \subset A^2$, we calculate $\mathbb{E}\,|X^2 \cap G|$.
Using linearity of expectation:

$$\begin{aligned}
\mathbb{E}\,|X^2 \cap G| &= \sum_{(a,b)\in G} \mathbb{P}(a,b \in X) \\
&= \sum_{(a,b)\in G} \sum_{s} \frac{\mathbb{1}_A \circ \mathbb{1}_A(s)}{|A|^2} \mathbb{1}_A(a-s)\mathbb{1}_A(b-s) \\
&= \frac{1}{|A|^2} \sum_{(a,b)\in G} \sum_{s} \mathbb{1}_A \circ \mathbb{1}_A(s)\mathbb{1}_A(a-s)\mathbb{1}_A(b-s)
\end{aligned}$$

We bound the inner sum above by Cauchy-Schwarz:

$$\begin{aligned}
\sum_{s} \mathbb{1}_A \circ \mathbb{1}_A(s)\mathbb{1}_A(a-s)\mathbb{1}_A(b-s) &\leq \left(\sum_{s} \mathbb{1}_A \circ \mathbb{1}_A(s)^2\right)^{1/2} \left(\sum_{s\in G} \mathbb{1}_A(a-s)\mathbb{1}_A(b-s)\right)^{1/2} \\
&= E(A)^{1/2}\mathbb{1}_A \circ \mathbb{1}_A(a-b)^{1/2}
\end{aligned}$$

It then follows that:

$$\mathbb{E}\,|X^2 \cap G| \leq \frac{E(A)^{1/2}}{|A|^2} \sum_{(a,b)\in G} \mathbb{1}_A \circ \mathbb{1}_A(a-b)^{1/2}$$

In particular, if $G \subset A^2$ is the set of pairs (a,b) such that:

$$\mathbb{1}_A \circ \mathbb{1}_A(a-b) \leq \frac{c^2}{4}\frac{E(A)^3}{|A|^8}$$

then

$$\mathbb{E}\,|X^2 \cap G| \leq \frac{c}{2}\frac{E(A)^2}{|A|^4}$$

using the trivial bound that $|G| \leq |A|^2$.
Furthermore, by Cauchy-Schwarz again, $\mathbb{E}\,|X|^2 \geq (\mathbb{E}\,|X|)^2$, so $\mathbb{E}\,|X|^2 \geq \frac{E(A)^2}{|A|^4}$.
Hence, by the first moment method, there is some $X \subseteq A$ such that

$$|X|^2 - \frac{1}{c}|X^2 \cap G| \geq \frac{1}{2}\frac{E(A)^2}{|A|^4}$$

Hence we have:

1. $|X|^2 \geq \frac{1}{2}\frac{E(A)^2}{|A|^4} \geq \frac{1}{2}\frac{|A|^2}{K^2}$, so $|X| \gg \frac{|A|}{K}$.
2. $|X^2 \cap G| \leq c|X|^2$

□

It remains to find some large $A' \subset X$ such that $|A' - A'|$ is small.

Proof of BSG.. We apply the previous lemma to A with $c = \frac{1}{8}$, to obtain some X as stated.

Since $\mathbb{1}_A \circ \mathbb{1}_A(a-b) = \mathbb{1}_A \circ \mathbb{1}_A(b-a)$, we can define a graph H with vertex set X, such that a, b are connected by an edge in H if and only if $a \neq b$ and $\mathbb{1}_A \circ \mathbb{1}_A(a-b) \gg K^{-3}|A|$.

By **Lemma 4.4**, there are at most $\frac{1}{8}|X|^2$ many pairs $(a,b) \in X^2$ where this condition fails, and hence H has at least $\binom{|X|}{2} - \frac{1}{16}|X|^2$ edges.

If $d(x)$ denotes the degree of a vertex $x \in H$, then:

$$\sum_{x \in X} d(x) \geq \frac{7}{8}|X|^2 - |X|$$

Let A' be the subset of X consisting of those elements of degree at least $\frac{3}{4}|X|$ in H.

The contribution to the degree count from $x \notin A'$ is at most $\frac{3}{4}|X|^2$, and hence:

$$|X||A'| \geq \sum_{x \in A'} d(x) \geq \frac{1}{8}|X|^2 - |X|$$

and hence $|A'| \gg |X|$.

We now claim that, for any $x \in A' - A'$, there at $\gg K^{-6}|A|^2|X|$ many quadruples $(a_1, a_2, a_3, a_4) \in A^4$ such that $x = a_1 - a_2 + a_3 - a_4$. Assuming this, since the number of such quadruples is at most $|A|^4$, we have:

$$|A|^4 \gg |A' - A'|K^{-6}|A|^2|X|$$

Then, rearranging we will have:

$$|A' - A'| \ll K^6 \frac{|A|^2}{|X|} \ll K^7|A|$$

Fix some $a, b \in A'$ such that $x = a - b$. By choice of A', there are $\geq \frac{1}{2}|X|$ many $c \in X$ such that $(a,c), (b,c)$ are edges in H.

Hence there are $\gg K^{-3}|A|$ many $(a_1, a_2) \in A^2$ such that $a_1 - a_2 = a - c$ and $\gg K^{-3}|A|$ many $(a_3, a_4) \in A^2$ such that $a_3 - a_4 = b - c$. Since different c give different quadruples for fixed x, in total we have:

$$\gg \left(\frac{1}{2}|X|\right)\left(K^{-3}|A|\right)\left(K^{-3}|A|\right) \gg |X|K^{-6}|A|$$

many such quadruples. □

Addendum/Big Picture Remarks. The goal of BSG is to go from $E(A)$ is large (i.e. $\geq K^{-1}|A|^3$) to getting $A' \subset A$ with $|A' - A'| \ll |A'|$, where A' is a 'large' subset of A.

Recall that $E(A)$ is $\sum_x \mathbb{1}_A \circ \mathbb{1}_A(x)^2$, where $\mathbb{1}_A \circ \mathbb{1}_A(x)$ is the number of ways of writing x as a difference in $A - A$. dd Let S dbe the set $\{x : \mathbb{1}_A \circ \mathbb{1}_A(x) \geq \frac{1}{2K}|A|\}$ (note that trivially $\mathbb{1}_A \circ \mathbb{1}_A(x) \leq |A|$), so $x \in S$ implies that x really can be written as a difference in a large number of ways. We want to know how big S is.

One the one hand,

$$\sum_{x \notin S} \mathbb{1}_A \circ \mathbb{1}_A(x)^2 < \frac{1}{2K}|A| \sum_x \mathbb{1}_A \circ \mathbb{1}_A(x) = \frac{1}{2K}|A|^3$$

Now since $\sum_x \mathbb{1}_A \circ \mathbb{1}_A(x)^2 = \frac{1}{K}|A|^3$, we must have:

$$|S||A|^2 \geq \sum_{x \in S} \mathbb{1}_A \circ \mathbb{1}_A(x)^2 \geq \frac{1}{2K}|A|^3$$

and hence

$$|S| \gg \frac{1}{K}|A|$$

For an upper bound on S, we note that $|S|\frac{1}{2K}|A| \leq \sum_{x \in S} \mathbb{1}_A \circ \mathbb{1}_A(x) \leq \sum_x \mathbb{1}_A \circ \mathbb{1}_A(x) = |A|^2$, and so $|S| \ll K|A|$. So up to some constant terms, $|S|$ is essentially $|A|$.

Since $x \in S$ implies that x is represented many times in $A - A$, we must have $S \subset A - A$. So we would be done if we could find $A' \subset A$ such that $A' - A' \subset S$.

This can't be done in general, but we can find a large $A' \subset A$ such that "99%" of all pairs $(a', b') \in A'^2$ have $a' - b' \in S$ - this is **4.4**. Such an S is often called a "symmetry set". We then want to get to a "100%" statement: we know that 99% of the (a', b') have $a' - b' \in S$, so we expect for most $a', b' \in S$, there are $\gg |A'|$ many $c \in A'$ such that $a' - c, b' - c \in S$.

We then write $a' - b' = (a' - c) - (b' - c) \in S - S$. So we can write $a' - c \in S$ as $a_1 - a_2$ in many ways, and likewise for $b' - c$. So $a' - b'$ can be written in $\gg K|A'||A|^2$ many ways, and so in total have $\gg |A' - A'||A|^3$ many such quadruples. But trivially there are at most $|A|^4$ quadruples, so $|A' - A'| \ll K|A|$.

1.3 Covering

We say that A is K-covered by B if there is X with $|X| \leq K$ such that $A \subseteq X + B$. I.e., we can contain A in at most K-many translates of B.

For example, if A is K-covered by B, where $|B| \ll |A|$ and $|B|$ has small doubling, then A must also have small doubling:

$$|A + A| \leq |X + X + B + B| \ll K^2|B + B| \ll_K |B| \ll |A|$$

We can often get more from covering than this suggests. For instance, we can control iterated sums: if $A + B$ is efficiently covered by A itself, then so is $A + nB$ for small n. If $A + B \subset A + X$, then $A + nB \subset A + nX$.

Before we use this concept, we need to show how to produce efficient coverings. The first way is due to Ruzsa, and is simple and useful.

Lemma 1.12 (Ruzsa Covering Lemma)**.** *If $|A + B| \leq K|B|$, then A is K-covered by $B - B$.*

Proof. Let X be maximal such that $(x + B)_{x \in X}$ are disjoint.

We can control the size of X:

$$|X||B| = |X + B| \leq |A + B| \leq K|A|$$

and so $|X| \leq K$.

If $a \in X$, then obviously $a \in X + B - B$. If $a \in A \setminus X$, then $a + B$ must intersect $x + B$ for some $x \in X$ (by maximality). I.e. there are $b_1, b_2 \in B$ such that $a + b_1 = x + b_2$. But then $a = x + b_1 - b_2 \in X + B - B$, and so $A \subseteq X + B - B$. □

Note that we are covering A by $B - B$, not B itself. This is because $B - B$ is much 'smoother' than B itself - e.g. what if B is a large random subset of G. On average, G needs $\gg \log |G|$ translates of B to cover everything, so even though $|G + B| \leq |G| \ll |B|$, but if $G \subset X + B$, then $X \gg \log |G|$ (with high probability). So G is not efficiently covered by B, even though $|G + B| \ll |B|$. But $B - B = G$ with high probability. Indeed, any $B \subset G$ with $|B| > \frac{1}{2}|G|$ has $B - B = G$, since for any $x \in G$, consider B and $x + B$. These are subsets of G of size $> \frac{1}{2}|G|$, so they must intersect. I.e. there are $b_1, b_2 \in B$ with $b_1 = x + b_2$.

We now present an application of Ruzsa's covering lemma and Plünnecke's inequality - a non-trivial inverse sumset result.

As we saw earlier, if H is a subgroup of G and $A \subset x + H$ with $|A| \gg |H|$, then $|A + A| \ll |A|$. We will show that this is the only way that A can have small doubling in groups of small torsion (e.g. $G = \mathbb{F}_p^n$, not in subsets of $\mathbb{Z}$ or $\mathbb{Z}/N\mathbb{Z}$).

Theorem 1.13. *If $A \subset \mathbb{F}_p^n$ is such that $|A + A| \leq K|A|$, then there is a subgroup $H \leq \mathbb{F}_p^n$ and x such that $A \subseteq x + H$ and $|A| \gg_{p,K} |H|$.*

Proof. Without loss of generality, $0 \in A$. The idea is to take $H = \langle A \rangle$. But trivially we only have $|H| \leq p^{|A|}$ - we will use the small doubling of A to control the size of H.

We want to use Ruzsa's covering lemma - the obvious way is $|A + A| \leq K|A|$, so $|A|$ is K-covered by $A - A$. But A is 1-covered by $A - A$ anyway, so this is useless. Instead, we will cover $A + A - A$ by $A - A$. Plünnecke's inequality says that:

$$|A + A - A + A| \leq K^4|A|$$

and hence by Ruzsa's covering lemma, there is $X \subset 2A - A$ with $|X| \leq K^4$ and:

$$\begin{aligned} A + A - A &\subseteq X + A - A \\ A + A + A - A &\subseteq X + (A + A - A) \subseteq X + X + A - A \end{aligned}$$

and so on. In general, we have for all $n \geq 1$, $nA + (A - A) \subseteq nX + (A - A)$. If $H_0 = \langle X \rangle$, then $|H_0| \ll_{p,K} 1$, because $|H_0| \leq p^{|X|} \leq p^{K^4} = C(p, K) \cdot 1$.

Let $H = \langle A \rangle$. If $h \in H$, then $h \in nA$ for some $n \geq 1$, so:

$$h \in nA \subseteq nA + (A - A) \subseteq nX + (A - A) \subseteq H_0 + (A - A)$$

Hence $H \subseteq H_0 + (A - A)$. So $|A| \leq |H_0||A - A| \ll_{p,K} 1 \cdot K^2|A| \ll_{p,K} |A|$. □

Note the essential use of bounded torsion, when we bounded $|H_0| \ll_{p,K} 1$. If we tried this in $\mathbb{Z}$, then $|H_0| = \infty$. Even in $\mathbb{Z}/N\mathbb{Z}$, if $|X| \geq 2$, then $H_0 = \langle X \rangle = \mathbb{Z}/N\mathbb{Z}$ (if N is prime). Something similar does hold however if we replace 'subgroup' by 'arithmetic progression'.

The dependency on constants is $|H| \leq (K^2 p^{K^4})|A|$. This has been improved - the best known bound is due to Green-Tao, Lovett, and Zohar, and is of the form $|H| \leq p^{o(K)}|A|$.

This is the best possible: if $|A| = K$ and A is linearly independent over $\mathbb{F}_p$, then $|A + A| \leq |A|^2 \leq K|A|$, and the smallest cost containing A has size p^K.

But, part of the need for this exponential dependence is that we insisted A is 1-covered by H.

Note in the example above, A, is K-covered by $\{0\}$. Now we expect that we can do polynomial bounds in K, but this is still open.

Conjecture 1.14 (Marton)**.** *If $A \subset F_p^n$ has $|A + A| \leq K|A|$, then A is C-covered by $H \leq \mathbb{F}_{p^n}$ with $|H| \leq |A|$, where $C = O_p(K^{O_p(1)})$.*

Conjecture 1.15 (Polynomial Freiman-Ruzsa conjecture for $\mathbb{F}_p^n$)**.** *If $|A| \subset \mathbb{F}_p^n$ has $|A + A| \leq K|A|$, then there is $H \leq \mathbb{F}_p^n$ such that $|H| \ll_p K^{O_p(1)}|A|$ and $|A \cap (x + H)| \gg_p K^{-O_p(1)}|A|$.*

These conjectures are equivalent: assume Marton's conjecture. Then there is X with $|X| \ll K^{O(1)}$ such that $A \subset X + H$ and $|H| \leq |A|$. Then by pigeonholing, there is some $x \in X$ such that $|A \cap (x + H)| \geq K^{-O(1)}|A|$. We'll see the other direction later on.

A similar form of PFR is conjectured for any abelian group, again replacing subgroups by arithmetic progressions.

1.4 Span-Covering

In covering, we wanted a "small" X such that $A \subset X + B$. Sometimes, it's enough for X to have "low complexity" (i.e. X is spanned by a small number of elements).

Definition 1.16. *We say A is K-span covered by B if there is X of size $|X| \leq K$ such that:*

$$A \subset Span(X) + B$$

where $Span(X) = \{\sum_{x \in X} c_x x : c_x \in \{-1, 0, 1\}\}$.

Lemma 1.17. *If $|A + A| \leq K|A|$ and $|A + B| \leq K'|B|$. Then A is $O(K \log(KK'))$-span covered by $B - B$.*

Note that if K is fixed, we can view this as $O(\log K')$ dependence. So $|\mathrm{Span}(X)| \approx 3^{\log(K')} \approx (K')^c$.

Proof. We construct a sequence of sets $B_0, B_1, \ldots$ as follows: first, let $B_0 = B$. We then construct B_{n+1} in terms of B_n.

Suppose there are $2K$ many disjoint translates of B_n of the form $a + B_n$ for $a \in A'_n$ disjoint, where $|A'_n| = 2K, A'_n \subseteq A$. Then we just let $B_{n+1} A'_n + B_n$, and note that $|B_{n+1}| = 2K|B_n|$.

Otherwise, let $A' \subseteq A$ be maximal such that $(a + B_n)$ for $a \in A'$ are disjoint, so $|A'| < 2K$. Now let $B_{n+1} = A' + B_n$ and stop the construction.

We first show this construction must stop eventually. By induction, $B_n \subset B + nA$, so $(2K)^n|B| = |B_n| \leq |B_n A|$. Now,

$$\begin{aligned} |B + nA| &\leq \frac{|B + A||nA - A|}{|A|} && (RTI) \\ &\leq K'|B|K^{n+1} && (\text{Plünnecke}) \end{aligned}$$

So $(2K)^n \leq K'K^{n+1}$, so $2^n \leq K'K$, so $n \leq \log_2(K'K) \ll \log(K'K)$. Hence we must stop at some $n = O(\log(KK'))$.

We now claim that the span of $X = A'_0 \cup \ldots \cup A'_n$ covers A by $B - B$. Then we are done, since $|X| \leq \sum_{i=0}^n |A'_n| \leq 2K(n + 1) \ll K \log(KK')$.

Suppose that $a \in A$. We know that $a + B_n$ is not disjoint from B_{n+1}, since A' was chosen to be maximally disjoint. So $a \in B_{n+1} - B_n$. Hence $A \subseteq B_{n+1} - B_n$. But $B_{n+1} = A'_n + A'_{n-1} + \ldots + A'_0 + B$, $B_n = A'_{n-1} + \ldots + A'_0 + B$.

So $A \subseteq B_{n+1} - B_n = B - B + (A'_0 - A'_0) + \ldots + (A'_{n-1} - A'_{n-1}) + A'_n \subseteq B - B + \mathrm{Span}(X)$. □

1.5 Finding Arithmetic Progressions

What kind of conditions on a set of integers guarantee the existence of an arithmetic progression?

Conjecture 1.18 (Baudet, 1920s)**.** *If $\mathbb{Z}$ is coloured red/blue, then there exist arbitrarily long APs of the same colour.*

For instance, any colouring of $\{1, \ldots, 9\}$ contains a monochromatic 3AP.

Theorem 1.19 (van der Waerden)**.** *For any $r, k \geq 1$, if $\mathbb{Z}$ is r-coloured then there exists a monochromatic k-term arithmetic progression.*

This is proven in the Part III course Ramsey Theory.

We might also ask which colour class contains the k-AP. Perhaps it is always the largest?

Conjecture 1.20. *For any $\delta > 0$, if $A \subset \mathbb{Z}$ has 'density' δ, then A has arbitrarily long APs.*

This would imply VDW, since in an r-colouring there must be a colour class with density $\geq \frac{1}{r}$. This also turns out to be true:

Theorem 1.21 (Szemerédi, 1975)**.** *For any $\delta > 0$ and $k \geq 1$ there exists $N \ll_{k,\delta} 1$ such that, if $N' \geq N$ and $A \subset \{1, \ldots, N'\}$ with $|A| \geq \delta N'$, then A contains a k-term AP.*

There are many different proofs - the combinatorial one by Szémeredi, as well as ones via ergodic theory, hypergraph regularity, and an analytic proof by Gowers.

Let $r_k(N)$ be the size of the largest $A \subseteq \{1, \ldots, N\}$ such that A has no k-term APs, so that Szémeredi's theorem is that $\frac{r_k(N)}{N} \to 0$ as $N \to \infty$.

Theorem 1.22 (Gowers)**.** *For all $k \geq 3$,*

$$r_k(N) \ll \frac{N}{(\log\log N)^{c_k}}$$

In this course we will focus on the case $k = 3$. Let $r(N) = r_3(N)$.

Theorem 1.23 (Roth, 1953)**.** $\frac{r(N)}{N} \to 0$ *as* $N \to \infty$.

In fact, Roth proved $r(N) \ll \frac{N}{\log\log N}$. Roth used Fourier analysis. We will develop enough Fourier analysis to prove:

Theorem 1.24 (Bourgain, 1999)**.** $r(N) \ll \left(\frac{\log\log N}{\log N}\right)^{1/2} N$

2 Fourier Analysis

2.1 Basic Concepts

We will only be considering finite abelian groups. Let G be a finite abelian group. A ***character*** on G is a group homomorphism $\gamma : G \to \mathbb{C}^\times$. Lower-case Greek letters will be reserved for characters, and lower-case Roman letters are for elements of G.

The set of characters can be made into a group with the group operation pointwise multiplication, and the identity the trivial character $\mathbb{1}(x) = 1\ \forall x \in G$. This group is $\widehat{G}$, the ***dual group*** of G. Since multiplication in $\mathbb{C}$ is abelian, $\widehat{G}$ is abelian. Since any character is a homomorphism, $\gamma(0) = 1$. Also, $\gamma(x)$ is a $|G|^{\text{th}}$ root of unity for any $x \in G$. Hence $\widehat{G}$ is a finite abelian group.

Lemma 2.1. $\widehat{G} \cong G$

Proof. Omitted. Straightforward with classification of finite abelian groups. Note it is obvious is G is cyclic. □

For example, if $G = \mathbb{Z}/N\mathbb{Z}$, then any $\gamma \in \{1, \dots, N\}$ induces a character by $\gamma(x) = e\left(\frac{\gamma x}{N}\right)$, where $e(z) = \exp(2\pi i z)$. Or, if $G = \mathbb{F}_p^n$, then $\gamma \in \mathbb{F}_p^n$ gives a character $\gamma \in \widehat{G}$ where $\gamma(x) = e\left(\frac{\gamma \cdot x}{p}\right)$.

When talking about G, we use the counting measure, the sum. With $\widehat{G}$, we use the 'normalised probability measure' $\frac{1}{N}\Sigma$, which we denote by $\mathbb{E}$.

We have a natural inner product for functions $f, g : G \to \mathbb{C}$:

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

And, if $f, g : \widehat{G} \to \mathbb{C}$, we use:

$$\langle f, g \rangle = \mathop{\mathbb{E}}_{\gamma \in \widehat{G}} f(\gamma)\overline{g(\gamma)}$$

Lemma 2.2 (Orthogonality)**.** *For any $\gamma \in \widehat{G}$:*

$$\sum_{x \in G} \gamma(x) = \begin{cases} |G| & \gamma = \mathbb{1} \\ 0 & \textit{otherwise} \end{cases}$$

For any $x \in G$:

$$\mathop{\mathbb{E}}_{\gamma \in \widehat{G}} \gamma(x) = \begin{cases} 1 & x = 0 \\ 0 & \textit{otherwise} \end{cases}$$

Proof. If $\gamma = \mathbb{1}$ or $x = 0$, this is trivial.

Now suppose $\gamma \neq \mathbb{1}$, so there is $y \in G$ such that $\gamma(y) \neq 1$. If $S = \sum_{x \in G} \gamma(x)$, then:

$$\begin{aligned}
\gamma(y)S &= \gamma(y) \sum_{x \in G} \gamma(x) \\
&= \sum_{x \in G} \gamma(y)\gamma(x) \\
&= \sum_{x \in G} \gamma(y + x) \\
&= \sum_{x \in G} \gamma(z) = S
\end{aligned}$$

Hence $S = 0$. The second claim is similar, using $\lambda \in \widehat{G}$ with $\lambda(x) \neq 1$. □

2.2 Fourier Transform

If $f : G \to \mathbb{C}$, then we define $\widehat{f} : \widehat{G} \to \mathbb{C}$ by:

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

Lemma 2.3 (Parseval's Identity)**.** *For any $f, g : G \to \mathbb{C}$, we have:*

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

In particular, $\|f\|_2 = \|\widehat{f}\|_2$.

Proof.

$$\begin{aligned}
\langle f, g \rangle &= \sum_{x \in G} f(x)\overline{g(x)} \\
&= \sum_{x,y \in G} f(x)\overline{g(y)}\mathbb{1}_{x=y} \\
&= \sum_{x,y \in G} f(x)\overline{g(y)} \mathop{\mathbb{E}}_{\gamma \in \widehat{G}} \gamma(y - x) \\
&= \mathop{\mathbb{E}}_{\gamma \in \widehat{G}} \left(\sum_{x \in G} f(x)\gamma(-x) \right) \left(\sum_{y \in G} \overline{g(y)}\gamma(y) \right) \\
&= \mathop{\mathbb{E}}_{\gamma \in \widehat{G}} \left(\sum_{x \in G} f(x)\overline{\gamma(x)} \right) \overline{\left(\sum_{y \in G} g(y)\overline{\gamma(y)} \right)} \\
&= \langle \widehat{f}, \widehat{g} \rangle
\end{aligned}$$

□

Lemma 2.4 (Diagonalises convolution)**.** *Given $f, g : G \to \mathbb{C}$, we have:*

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

Proof. For any $\gamma \in \widehat{G}$,

$$\begin{aligned}
\widehat{f * g}(\gamma) &= \sum_{z \in G} f * g(z)\overline{\gamma}(z) \\
&= \sum_{z \in G} \sum_{x,y \in G} f(x)g(y)\mathbb{1}_{x+y=z}\overline{\gamma(z)} \\
&= \sum_{x,y \in G} f(x)g(y)\overline{\gamma(x+y)} \\
&= \left(\sum_{x \in G} f(x)\overline{\gamma(x)}\right)\left(\sum_{y \in G} g(y)\overline{\gamma(y)}\right)
\end{aligned}$$

□

2.3 Roth's Theorem in $\mathbb{F}_{p^n}$

Our goal is to prove Roth's thoerm in $\mathbb{Z}$ - if $A \subset \{1, \ldots, N\}$ and $|A|/N$ is large, then A contains a 3AP. We will first prove a version of this in $\mathbb{F}_p^n$, and the proof for $\mathbb{Z}$ will be a "translation" of this proof.

Fix some odd prime $p \geq 3$ (in $p = 2$, then 3APs are trivial as $x, x + d, x$ is a 3AP). Then a 3AP is a set of the form $\{x, x + d, x + 2d\}$ where $x, d \in \mathbb{F}_p^n$. We allow $d = 0$, but call a 3AP with $d \neq 0$ 'non-trivial'.

We're given $A \subseteq \mathbb{F}_p^n$ and we know its size or equivalently its ***density*** $\alpha = |A|/p^n$. We want to deduce that α large implies A has a non-trivial 3AP.

Two observations:

1. This is easy if A is very structured, e.g. if A is a subspace.
2. This is easy if A is random. The expected number of 3APs in A is approximately the sum over all 3APs in $\mathbb{F}_p^n$ of the probability that 3AP is contained in A, which is about $p^{2n} \times \alpha^3$. The number of trivial 3APs in A is just $|A| = \alpha p^n$, so provided $\alpha^3 p^{2n}$ is much larger than αp^n, the non-trivial count dominates, and there must be some non-trivial 3APs.

Tao called this the "structred vs randomness" dichotomy - all sets should be structured to some degree and random to some degree, and so the theorem should be true.

Our proof will use the "density increment strategy".

Suppose that $A \subseteq \mathbb{F}_p^n$ has no non-trivial 3APs. We want to show that $\alpha = |A|/p^n$ is small.

Either:

1. A has $\gg \alpha^3 p^{2n}$ many 3APs, and hence either:
   (a) α is small - done.
   (b) A has a non-trivial 3AP - contradiction.

2. A is not well-distributed, and there is a large subspace $W \leq \mathbb{F}_p^n$ and a coset of W, say $x+W$, on which A has increased density (i.e. $\frac{|A\cap(W+x)|}{|W|} \geq \alpha(1+c)$ for $c > 0$).

   Then we zoom in on $A \cap (W+x)$ (or rather $(A-x)\cap W$, as being a 3AP is invariant under translation). $W \cong \mathbb{F}_p^{n-m}$ for some small m, and repeat.

We can't repeat forever since α is bounded above by 1, so we eventually stop and are done.

The goal is:

Theorem 2.5 (Meshulam, 1957). *If $A \subseteq \mathbb{F}_p^n$ has no non-trivial 3APs, then $|A| \ll_p \frac{p^n}{n}$, and in particular, $\frac{|A|}{p^n} \to 0$ as $n \to \infty$.*

Our main tool is:

Lemma 2.6. *Let V be an n-dim vector space over $\mathbb{F}_p^n$ and let $A \subseteq V$ with density α. Suppose that A has no non-trivial 3APs. Then either:*

1. *$\alpha \leq \frac{\sqrt{2}}{p^{n/2}}$ or*
2. *There is a subspace $V' \leq V$ of codimension 1, and $x \in V$ with:*

$$\frac{|(A-x)\cap V'|}{|V'|} \geq (1+\frac{\alpha}{4})\alpha$$

*Proof of Meshulam using **Lemma 2.6**.* Let $A \subseteq \mathbb{F}_p^n$ with density $\alpha > 0$ and no nontrivial 3APs. We want $\alpha \ll \frac{1}{n}$.

If $\alpha \leq p^{-n/4}$, then we're done.

It's also enough to prove this for large n - if our proof showed that $\alpha \leq \frac{100}{n}$ for all $n \geq 1000$, then using $\alpha \leq 1$, we get that $\alpha \leq \frac{1000}{n}$ for all $n \geq 1$, and so $\alpha \ll \frac{1}{n}$ for all $n \geq 1$.

Let $k \geq 0$ be maximal such that the following holds: there is a sequence of sets $A_0, \ldots, A_k$ and associated $V_0, \ldots, V_k$ vector spaces over $\mathbb{F}_p$ such that:

1. $A_0 = A, V_0 = \mathbb{F}_p^n$.
2. $A_i \subseteq V_i$.
3. A_i has no nontrivial 3APs.
4. If $\alpha_i = \frac{|A_i|}{|V_i|}$ then $\alpha_{i+1} \geq (1+\frac{1}{4}\alpha_i)\alpha_i$ for all $0 \leq i < k$.
5. $|V_{i+1}| = \frac{1}{p}|V_i|$.

Note that $k = 0$ and the sequences A_0, V_0 work, so some such maximal k exists, and k cannot be infinite, since $1 \leq |V_i| = p^{-i}|V_0| = p^{n-i}$.

How large can k be? Induction shows that:

$$\alpha_i \geq (1+\alpha/4)^i\alpha \geq (1+i\alpha/4)\alpha$$

and in particular, after $i \geq \lceil 4/\alpha \rceil$ steps, $\alpha_i \geq 2\alpha$. After another $\lceil 4/2\alpha \rceil$ steps, $\alpha_i \geq 4\alpha$, and so on. Hence:

$$k \leq \sum_{j=0}^{O(\log(1/\alpha))} \lceil 4/2^j\alpha \rceil \leq \sum_{j=0}^{\infty} \lceil 4/2^j\alpha \rceil \ll \alpha^{-1}$$

We can assume in particular that

$$k \leq \frac{n}{10}$$

or else $n \ll \alpha^{-1}$, so $\alpha \ll \frac{1}{n}$ and we're done.

Now apply **Lemma 2.6** to $A_k \subseteq V_k$. If the second conclusion of **Lemma 2.6** holds, then there exists $V' \leq V_k$ and $x \in V_k$ such that

$$\frac{|(A_k - x) \cap V'|}{|V'|} \geq (1 + \frac{\alpha_k}{4})\alpha_k$$

But then we could let $A_{k+1} = (A_k - x) \cap V'$ and $V_{k+1} = V'$, contradicting the maximality of k. Hence the first conclusion must be held and

$$\alpha_k \leq \frac{\sqrt{2}}{p^{n/2}} \implies p^{-n/4} < \alpha \leq \alpha_k \ll \frac{1}{|V_k|^{\frac{1}{2}}}$$

By induction $|V_k| = p^{n-k} \geq p^{\frac{9}{10}n}$, and so $p^{-n/4} \ll p^{-9n/20}$, which is a contradiction. □

We now try to prove Meshulam's theorem. The strategy is to:

1. Write the difference between the number of 3APs in A and the expected number in a generic set of density α as an inner product involving $\mathbb{1}_A$ and the balanced function $\mathbb{1}_A - \alpha\mathbb{1}_G$.
2. If A has no nontrivial 3APs and is not too small, then this difference is large in absolute value.
3. Apply Parseval's identity to this inner product.
4. Deduce that there is some $\gamma \neq \mathbb{1}$ at which the Fourier transform of $\mathbb{1}_A - \alpha\mathbb{1}_G$ is large in absolute value.
5. Use this large Fourier coefficient to get a density increment, by choosing V' a subspace orthogonal to γ.

Proof. Think of V as $\mathbb{F}_p^n$. We first note that 3APs are exactly the sets $\{x, y, z\}$ such that $x + y = 2z$ (as $z - x = y - z$, if the AP is, in order, x, z, y).

In these terms, a 'trivial' 3AP (where $d = 0$) corresponds to the trivial solutions $x = y = z$.

Hence the number of 3APs in A is

$$\sum_{x,y,z \in A} \mathbb{1}_{x+y=2z} = \sum_{x,y \in A} \sum_{w \in 2\cdot A} \mathbb{1}_{x+y=w} = \sum_{w \in 2\cdot A} \mathbb{1}_A * \mathbb{1}_A(w) = \langle \mathbb{1}_A * \mathbb{1}_A, \mathbb{1}_{2\cdot A} \rangle$$

where $2 \cdot A$ is A dilated by 2, not $A + A$. Since p is odd, $g \mapsto 2g$ is a bijection, so $|2 \cdot A| = |A|$.

We want to compare this to the expected count. Introduce the 'first-order approximation' to $\mathbb{1}_A$ as $\alpha \cdot \mathbb{1}_G$. Note that this agrees with $\mathbb{1}_A$ on average, i.e. $\sum_x \mathbb{1}_A(x) = |A| = \alpha p^n = \sum_x \alpha\mathbb{1}_A(x)$.

Then:

$$\begin{aligned}\langle \alpha \mathbb{1}_G * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle &= \alpha\langle \mathbb{1}_G * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle \\ &= \alpha\langle \mathbb{1}_G, \mathbb{1}_{2\cdot A} \circ \mathbb{1}_A\rangle \\ &= \alpha \sum_{x\in G}\sum_{a,b\in A} \mathbb{1}_{2a-b=x} \\ &= \alpha|A|^2 \\ &= \alpha^3 p^{2n}\end{aligned}$$

(Note that this is the expected number of 3APs in a random set of density α)

We will compare this to the actual count of 3APs in A.

$$\begin{aligned}\langle \mathbb{1}_A * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle - \langle \alpha\mathbb{1}_G * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle &= \langle (\mathbb{1}_A - \alpha\mathbb{1}_G) * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle \\ &= \langle F_A * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle\end{aligned}$$

where $F_A = \mathbb{1}_A - \alpha\mathbb{1}_G$ is the "balanced" function.

So $\langle F_A * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle = |A| - \alpha^3 p^{2n} = \alpha p^n(1-\alpha^2 p^n)$. In particular, if $|A| \geq (2p^n)^{1/2}$ (so that the first condition fails), then $1 - \alpha^2 p^n \leq -\frac{1}{2}\alpha^2 p^n$, and

$$|\langle F_A * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle| \geq \frac{1}{2}\alpha^3 p^{2n}$$

We then use the theory about Fourier transforms:

$$\begin{aligned}\langle F_A * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle &= \langle \widehat{F_A * \mathbb{1}_A}, \widehat{\mathbb{1}_{2\cdot A}}\rangle \\ &= \langle \widehat{F_A}\cdot\widehat{\mathbb{1}_A}, \widehat{\mathbb{1}_{2\cdot A}}\rangle \\ &= \mathop{\mathbb{E}}_{\gamma} \widehat{F_A}(\gamma)\widehat{\mathbb{1}_A}(\gamma)\overline{\widehat{\mathbb{1}_{2\cdot A}}(\gamma)}\end{aligned}$$

By the triangle inequality, $\mathop{\mathbb{E}}_{\gamma} |\widehat{F_A}(\gamma)||\widehat{\mathbb{1}_A}(\gamma)||\widehat{\mathbb{1}_{2\cdot A}}(\gamma)| \geq \frac{1}{2}\alpha^3 p^{2n}$.

Note that $\widehat{F_A}(\mathbb{1}) = \sum_x F_A(x) = \sum_x(\mathbb{1}_A(x) - \alpha\mathbb{1}_G(x)) = |A| - \alpha|G| = 0$. Hence:

$$\begin{aligned}\mathop{\mathbb{E}}_{\gamma\neq\mathbb{1}} |\widehat{F_A}(\gamma)||\widehat{\mathbb{1}_A}(\gamma)||\widehat{\mathbb{1}_{2\cdot A}}(\gamma)| &\geq \frac{1}{2}\alpha^3 p^{2n} \\ &\leq (\sup_{\gamma\neq\mathbb{1}} |\widehat{F_A}(\gamma)|)(\mathop{\mathbb{E}}_{\gamma} |\widehat{\mathbb{1}_A}(\gamma)||\widehat{\mathbb{1}_{2\cdot A}}(\gamma)|)\end{aligned}$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathop{\mathbb{E}}_{\gamma} |\widehat{\mathbb{1}_A}(\gamma)||\widehat{\mathbb{1}_{2A}}(\gamma)| &\leq (\mathop{\mathbb{E}}_{\gamma} |\widehat{\mathbb{1}_A}(\gamma)|^2)^{1/2}(\mathop{\mathbb{E}}_{\gamma} |\widehat{\mathbb{1}_{2\cdot A}}(\gamma)|^2)^{1/2} \\
&= \|\widehat{\mathbb{1}_A}\|_2 \cdot \|\widehat{\mathbb{1}}_{2\cdot A}\|_2 \\
&= \|\mathbb{1}_A\|_2 \cdot \|\mathbb{1}_{2\cdot A}\|_2 \\
&= |A|^{1/2}|2 \cdot A|^{1/2} = |A| = \alpha p^n
\end{aligned}$$

Hence

$$\sup_{\gamma \neq \mathbb{1}} |\widehat{F_A}(\gamma)| \geq \frac{1}{2}\alpha^2 p^n$$

(Compare this to the trivial upper bound $|\widehat{F_A}(\gamma)| \leq \sum_x |F_A(x)| \leq |A| + \alpha p^n = 2\alpha p^n$).

Let $V' \leq \mathbb{F}_p^n$ be the subspace which 'annihilates' γ. I.e., thinking of $\gamma : \mathbb{F}_p^n \to \mathbb{C}$, this means

$$V' = \{x \in \mathbb{F}_p^n : \gamma(x) = 1\}$$

or, thinking of $\gamma \in \mathbb{F}_p^n$ so the corresponding character is $x \mapsto e\left(\frac{\gamma \cdot x}{p}\right)$ where $e(z) = \exp(2\pi i z)$, then

$$V' = \{x \in \mathbb{F}_p^n : \gamma \cdot x = 0\}$$

Then V' is a subspace of codimension 1. The key observation is that γ is constant on cosets of V'.

We know that $|\widehat{F_A}(\gamma)|$ is large. To relate this to $|A \cap (x + V')|$, let $V_1', \ldots, V_p'$ be the cosets of V'. Then:

$$\begin{aligned}
\widehat{F_A}(\gamma) &= \sum_x (\mathbb{1}_A(x) - \alpha \mathbb{1}_G(x))\overline{\gamma(x)} \\
&= \sum_{i=1}^{p} \left(\sum_{x \in V_i'} (\mathbb{1}_A(x) - \alpha \mathbb{1}_G(x))\overline{\gamma(x)} \right) \\
&= \sum_{i=1}^{p} \overline{\gamma(v_i)}(|A \cap V_i'| - \alpha p^{n-1})
\end{aligned}$$

We want to say there exists some i such that $|A \cap V_i'| - \alpha p^{n-1} \geq \frac{1}{4}\alpha^2 p^{n-1}$.

Let $c \in \mathbb{C}$ such that $|\widehat{F_A}(\gamma)| = c\widehat{F_A}(\gamma)$, so $|c| = 1$, and consider $\langle F_A, c\gamma + 1\rangle$.

$$\begin{aligned}
\langle F_A, c\gamma + 1\rangle &= c\langle F_A, \gamma\rangle + \langle F_A, 1\rangle \\
&= c\widehat{F_A}(\gamma) \\
&= |\widehat{F_A}(\gamma)|
\end{aligned}$$

Now note that $x \mapsto c\gamma(x) + 1$ is constant on cosets of V' - say it takes the value x_i on V'_i. Split this inner product over $V'_1, \ldots, V'_p$ as above:

$$\langle F_A, c\gamma + 1\rangle = \sum_i x_i \left(|A \cap V'_i| - \alpha p^{n-1}\right)$$

Take real parts:

$$\mathfrak{Re}\langle F_A, c\gamma + 1\rangle = \mathfrak{Re}\,|\widehat{F_A}(\gamma)| \geq \frac{1}{2}\alpha^2 p^n$$

So, by averaging, there is some $1 \leq i \leq p$ such that

$$\mathfrak{Re}(x_i)\left(|A \cap V'_i| - \alpha p^{n-1}\right) \geq \frac{1}{2}\alpha^2 p^{n-1}$$

Finally, note that, since $x_i = c\gamma(v_i) + 1$, and $c\gamma(v_i)$ is on S^1, we have $\mathfrak{Re}(x_i) \in [0, 2]$, and hence

$$|A \cap V'_i| - \alpha p^{n-1} \geq \frac{1}{4}\alpha^2 p^{n-1}$$

□

What goes wrong with this in $\mathbb{Z}/N\mathbb{Z}$? There is no large subgroup on which $\gamma(x) = 1$ (in particular, if N is prime, there are no large subgroups at all!)

Instead, we will pass to the subset where $|\gamma(x) - 1| \leq \varepsilon$ for some small $\varepsilon > 0$. These sets form 'Bohr sets' - they act like subspaces even when there aren't any.

Definition 2.7. *Let $\Gamma \subset \widehat{G}$ and $\rho \in [0, 2]$. Then the* **Bohr set** *with frequency set Γ and width ρ is*

$$Bohr(\Gamma; \rho) = \{x \in G : |1 - \gamma(x)| \leq \rho \ \forall \gamma \in \Gamma\}$$

If $\lambda > 0$ and $B = Bohr(\Gamma; \rho)$, then we write B_λ for $Bohr(\Gamma, \lambda\rho)$, the dilate of B by λ. The size of Γ is called the **rank** *of B.*

Important note: The Bohr set does not uniquely determine Γ and ρ, e.g. $\mathrm{Bohr}(\Gamma; 2) = G$ for all Γ, and hence the rank is not necessarily unique. Formally, we should speak of triples $(\mathrm{Bohr}(\Gamma; \rho), \Gamma; \rho)$, but this is cumbersome, so by convention whenever we speak about a Bohr set we are implicitly fixing Γ and ρ.

Some trivial properties:

1. B is symmetric, so $B = -B$, and contains 0, as $\gamma(-x) = \overline{\gamma(x)}$ and $\gamma(0) = 1$.
2. If $\Gamma \supset \Gamma'$, then $\mathrm{Bohr}(\Gamma; \rho) \subseteq \mathrm{Bohr}(\Gamma', \rho)$. They are decreasing in frequency sets.
3. If $\rho \leq \rho'$, then $\mathrm{Bohr}(\Gamma; \rho) \subseteq \mathrm{Bohr}(\Gamma, \rho')$. They are increasing in width.
4. $\mathrm{Bohr}(\Gamma; \rho_1) + \mathrm{Bohr}(\Gamma; \rho_2) \subseteq \mathrm{Bohr}(\Gamma, \rho_1 + \rho_2)$, by the triangle inequality.

Note also that a Bohr set of rank d is the inverse image of a cube of dimension d: given $G \to \mathbb{C}^d$ by $x \mapsto (\gamma(x))_{\gamma \in \Gamma}$, then $\mathrm{Bohr}(\Gamma, \rho)$ is the inverse image of a cube centered at 1 and side length 2ρ.

Examples. For any θ, $|1 - e(\theta)| = |e(-\theta/2) - e(\theta/2)| = 2|\sin(\pi\theta)|$. Jordan's inequality gives:

$$\frac{2}{\pi}|x| \leq |\sin(x)| \leq |x|$$

for $x \in (-\pi/2, \pi/2]$, and in particular, if $\|\theta\|$ is the distance of θ from the nearest integer, then

$$4\|\theta\| \leq |1 - e(\theta)| \leq 2\pi\|\theta\|$$

In $\mathbb{F}_p^n$, we identify $\widehat{G}$ with $\mathbb{F}_p^n$ where $\gamma \in \mathbb{F}_p^n$ induces a character via the dot product, $x \mapsto e(\frac{\gamma \cdot x}{p})$. In particular, if $\rho < 4/p$, then $|1 - \gamma(x)| \leq \rho$, so $\|\frac{\gamma \cdot x}{p}\| < \frac{1}{p}$. So this forces $\gamma \cdot x = 0$, and so, provided $\rho < 4/p$, $\mathrm{Bohr}(\Gamma, \rho)$ is the subspace orthogonal to Γ, giving an equivalence between subspaces and Bohr sets of sufficiently small width.

In $\mathbb{Z}/N\mathbb{Z}$ for N prime, identify $\widehat{G}$ with $\mathbb{Z}/N\mathbb{Z}$ by multiplication, so $\gamma(x) = e(\frac{x\gamma}{N})$. Then, using $|1 - e(\theta)| \approx \|\theta\|$, and so $\mathrm{Bohr}(\Gamma, \rho) \approx \{x \in \mathbb{Z}/N\mathbb{Z} : \|\frac{\gamma x}{N}\| \ll \rho\}$. In the rank 1 case, say $\gamma = 1$, then $x \in \mathrm{Bohr}(\Gamma; \rho) \iff |x| \ll \rho N$, i.e. the interval centered at 0 of width $O(\rho N)$. If we change γ, it just dilates this interval. So Bohr sets of rank 1 correspond here to symmetric APs containing 0.

If the rank > 1, things are more complicated, but they correspond to generalised APs of rank d, i.e. terms of the form $a + n_1 v_1 + \ldots + n_d v_d$ for $0 \leq n_i < K_i$.

Heuristically, if $\gamma(x)$ were distributed randomly along the unit circle, then $|1 - \gamma(x)| \leq \rho$ is true with probability $\approx \rho$, and so assuming independence of $\gamma \in \Gamma$, and so $|B| \approx \rho^d N$. If there is some dependence among the elements of Γ, this will only help, so this should be an approximate lower bund. We will also show that dilating a Bohr set changes it's size by a factor exponential in d.

Lemma 2.8. *If B is a Bohr set of rank d and width $\rho \in [0, 1]$, then:*

$$|B| \geq \left(\frac{\rho}{8}\right)^d N$$

Also, $|B_{1/2}| \geq 8^{-d}|B|$.

Proof. Let $B = \mathrm{Bohr}(\Gamma; \rho)$. We can cover the unit circle by $\leq \lceil 2\pi/\rho \rceil$ circles of radius $\rho/2$. It follows that G is covered by sets of the form $\{x \in G : \gamma(x) \in D_\gamma\}$ where D_γ are these discs of radius $\rho/2$. There are at most $\lceil 2\pi/\rho \rceil^d$ such sets. If X is any such set, then $X - X \subseteq B$, as if $\gamma(x_1), \gamma(x_2)$ both lie in a disc centered on a of radius $\rho/2$, then $|1 - \gamma(x_1 - x_2)| \leq \rho$ by the triangle inequality.

In particular, $|X| \leq |B|$, and hence $N \leq \lceil 2\pi/\rho \rceil^d |B|$. Then, as $2\pi + 1 < 8$, we get the bound.

The second part is similar - cover the part of the unit circle which is distance $\leq \rho$ from 1 by circles of radius $\rho/4$. If X' is any such set, then $X' - X' \subseteq B_{1/2}$, so $|X'| \leq |B_{1/2}|$. As before, we get $|B| \leq 8^d |B_{1/2}|$. □

2.4 Bohr Sets under Addition

Bohr sets may grow exponentially under addition - $\mathrm{Bohr}(\Gamma; \rho) + \mathrm{Bohr}(\Gamma; \rho) \subseteq \mathrm{Bohr}(\Gamma; 2\rho)$, and we expect $|B| \approx \rho^d N$. So if this containment is sharp, $|B + B| \approx (2\rho)^d N \approx 2^d |B|$.

Instead, we consider $B + B_\delta$ where δ is small. The above heuristic now suggests $|B + B_\delta| \approx (1 + \delta)^d |B|$, so if $\delta = o(1/d)$, then $|B + B_\delta| \approx (1 + o(1))|B|$.

Although this heuristic can fail, as we only showed a lower bound and no upper bound, in practice we can deal with this.

Definition 2.9 (Regularity). *A Bohr set B of rank d is* **regular** *if for all $0 \leq \delta \leq \frac{1}{200d}$, we have*

$$|B_{1+\delta}| \leq (1 + 200d\delta)|B|$$

and

$$|B_{1-\delta}| \geq (1 - 200d\delta)|B|$$

The constant 200 can be improved, but it's a reasonably nice round number that's big enough. In particular we then have, if $0 \leq \delta \leq \frac{\varepsilon}{200d}$, then

$$|B + B_\delta| \leq |B_{1+\delta}| \leq (1 + \varepsilon)|B|$$

Not all Bohr sets are regular! Take $G = \mathbb{F}_2^n$ and let $\Gamma \subset \mathbb{F}_2^n$ be linearly independent of size d, and let $B = \mathrm{Bohr}(\Gamma; 2 - \frac{1}{1000d})$.

Since $\gamma(x) \in \{\pm 1\}$, then $|1 - \gamma(x)| < 2 \iff \gamma(x) = 1$, so $B = \{x \in \mathbb{F}_2^n : \gamma \cdot x = 0 \ \forall \gamma \in \Gamma\}$, and so $|B| = 2^{n-d}$.

If $\delta = \frac{1}{200d}$, then since $(1 + \delta)(2 - \frac{2}{1000d}) > 2$, so $B_{1+\delta} = G = \mathbb{F}_2^n$.

The only thing that can go wrong is some malicious choice of initial width. Bourgain showed that, for any B, B can be made regular by changing the width.

Lemma 2.10 (Bourgain's Regularity Lemma). *For any Bohr set B, there is $\lambda \in [\frac{1}{2}, 1]$ such that B_λ is regular.*

We will use the following lemma:

Lemma 2.11. *Let $\mathcal{I}$ be a collection of open intervals in $\mathbb{R}$ whose union contains a closed interval of length λ.*

Then there is an finite collection $I_1, \ldots, I_n \in \mathcal{I}$ of disjoint intervals of total length $\geq \frac{\lambda}{2}$.

Proof. By compactness, without loss of generality we can assume $\mathcal{I}$ is finite. Let $\mathcal{I}' \subseteq \mathcal{I}$ be a minimal collection covering this closed interval.

Fix $x \in \mathbb{R}$, and suppose that there are two intervals $I = (a_I, b_I), J = (a_J, b_J)$ in $\mathcal{I}'$ containing x. Choose these such that $a_I < x$ is minimal and $b_J > x$ is maximal. If $(a, b) \in \mathcal{I}'$ also contains x, then $a_I \leq a$ and $b_J \geq b$. Then $(A, b) \subseteq I \cup J$, so by minimality of $\mathcal{I}'$, $(a, b) = I$ or J. So any x is contained in at most 2 intervals from $\mathcal{I}'$.

Now order the intervals in $\mathcal{I}'$ as $I_1, \ldots, I_n$ by their lower bound, $a_1 \leq a_2 \leq \ldots \leq a_n$.

Then we must have $a_1 \leq a_2 \leq b_1 \leq a_3 \leq b_2 \leq a_4 \leq \ldots$.

All the odd intervals are disjoint, as are the evens, so by pigeonhole one of them must union to total length $\geq \frac{\lambda}{2}$. □

Regularity says (roughly) that changing the width by $O(1/d)$ doesn't change the size by more than $O(1)$. If we keep failing being regular for $\lambda \in [1/2, 1]$, then there are $\gg d$ many 'steps' of width, going from $\frac{1}{2}\rho$ to ρ, each time multiplying the size of the Bohr set by some constant C. Then in the end $|B| \gg C^d |B_{1/2}|$, but $|B| \leq 8^d |B_{1/2}|$, giving a contradiction.

Proof of Bourgain's Regularity Lemma. Let $B = \mathrm{Bohr}(\Gamma; \rho)$, and $B(\delta) = B_\delta$. Suppose the lemma is false. Then for every $\lambda \in [1/2, 1]$, there is $0 \leq \delta_\lambda \leq \frac{1}{200d}$ such that either

$$|B((1+\delta_\lambda)\lambda)| > (1 + 200\delta_\lambda d)|B(\lambda)|$$

or

$$|B((1-\delta_\lambda)\lambda)| < (1 - 200\delta_\lambda d)|B(\lambda)|$$

In either case,

$$|B((1+\delta_\lambda)\lambda)| > (1 + 100\delta_\lambda d)|B((1-\delta_\lambda)\lambda)|$$

Consider the intervals $I_\lambda = ((1-2\delta_\lambda)\lambda, (1+2\delta_\lambda)\lambda)$ for all $\lambda \in [\frac{1}{2} + \frac{1}{100d}, 1 - \frac{1}{100d}]$.

By the previous lemma, there is some finite $\lambda_1 < \lambda_2 < \ldots < \lambda_k$ such that the I_{λ_i} are disjoint and have total measure:

$$\sum_{i=1}^{k} 4\delta_{\lambda_i}\lambda_i \geq \frac{1}{2}(\frac{1}{2} - \frac{1}{50}d) \geq \frac{1}{5}$$

and so $\sum_{i=1}^{k} \delta_{\lambda_i} \geq \frac{1}{20}$.

Since $(1-\delta_{\lambda_1})\lambda_1 \geq \frac{1}{2}$ and $(1+\delta_{\lambda_k})\lambda_k \leq 1$, we know:

$$\frac{|B(1/2)|}{|B|} \leq \frac{|B((1-\delta\lambda_1)\lambda_1)|}{|B((1+\delta_{\lambda_k})\lambda_k)|}$$

By disjointedness of the I_{λ_i}, we know that $(1+\delta_{\lambda_i})\lambda_i < (1-\delta_{\lambda_{i+1}})\lambda_{i+1}$.

So $\frac{|B((1-\delta_{\lambda_{i+1}})\lambda_{i+1})|}{|B((1+\delta_{\lambda_i})\lambda_i)|} \geq 1$.

Therefore:

$$\frac{|B(1/2)|}{|B|} \leq \frac{|B((1-\delta_{\lambda_1})\lambda_1)|}{|B((1-\delta_{\lambda_k})\lambda_k)|} \leq \prod_{i=1}^{k} \frac{|B((1-\delta_{\lambda_i})\lambda_i)|}{|B((1+\delta_{\lambda_i})\lambda_i)|} \leq \prod_{i=1}^{k}(1 + 100\delta_{\lambda_i} d)^{-1}$$

Using $1 + x \geq e^{x/2}$ for $0 \leq x \leq 1$, we have:

$$\frac{|B(1/2)|}{|B|} \leq e^{-\frac{1}{2}\cdot 100\cdot\frac{1}{20}\cdot d} \leq e^{-\frac{5}{2}d}$$

But we also have $|B(1/2)| \geq 8^{-d}|B|$, and so $8^{-d} \leq e^{-\frac{5}{2}d}$, giving a contradiction, as $5/2 > \log(8)$. □

We now use regularity to prove a convenient lemma for removing convolutions by B_δ.

Lemma 2.12. *If B is a regular Bohr set of rank d and $B' \subseteq B_\delta$, where $0 < \delta < \frac{1}{200d}$, then for any function f with $|f(x)| \leq m$, we have:*

$$\langle f, \mathbb{1}_B * \mathbb{1}_{B'} \rangle = \langle f, \mathbb{1}_B \rangle |B'| + O(\delta d m |B||B'|)$$

In particular, if $A \subseteq B$, then

$$\langle \mathbb{1}_A, \mathbb{1}_B * \mathbb{1}_{B'} \rangle = |A||B'| + O(\delta d_{|B||B'|})$$

In future, we will say "by regularity" to mean "by **Lemma 2.12**" - this is the critical thing we need regularity for.

Proof.

$$\langle f, \mathbb{1}_B * \mathbb{1}_{B'} \rangle - \langle f, \mathbb{1}_B \rangle |B'| = \sum_x f(x)(\mathbb{1}_B * \mathbb{1}_{B'}(x) - \mathbb{1}_B(x)|B'|)$$

By the triangle inequality, this is at most:

$$M \sum_x |\mathbb{1}_B * \mathbb{1}_{B'}(x) - \mathbb{1}_B(x)|B'|| = M \sum_x \left| \sum_{y \in B'} \mathbb{1}_B(x-y) - \mathbb{1}_B(x)\mathbb{1}_{B'}(y) \right| \leq M \sum_{y \in B'} \sum_x |\mathbb{1}_B(x-y) - \mathbb{1}_B(x)\mathbb{1}_{B'}(y)|$$

This is equal to:

$$M \sum_{y \in B'} |B \triangle (B+y)|$$

Now $B_{1-\delta} \subseteq B + y \subseteq B_{1+\delta}$ for all $y \in B_\delta$. So:

$$|B \triangle (B+y)| \ll |B_{1+\delta} \setminus B| + |B \setminus B_{1-\delta}| \ll \delta d |B|$$

□

2.5 Roth's Theorem in the Integers.

Theorem 2.13 (Bourgain, 1999). *If $A \subset \{1, \ldots, N\}$ has no non-trivial 3APs, then $|A| \ll \left(\frac{\log\log N}{\log N}\right)^{1/2} N$. In particular, $\frac{|A|}{N} \to 0$ as $N \to \infty$.*

This will be proven using:

Theorem 2.14. *Let G be a finite abelian group of odd order N. If $A \subseteq G$ has no nontrivial 3APs then*

$$|A| \ll \left(\frac{\log\log N}{\log N}\right)^{1/2} N$$

*Proof of Bourgain's theorem, assuming **2.14**.* Let $A \subset \{1, \ldots N\}$ have no nontrivial 3APs.

Let $M = 2N - 1$ (odd). Suppose A had a non-trivial 3AP mod M. This means there are $x, y, z \in A$ distinct such that $x + y \equiv 2z \mod M$, with $1 \leq x, y, z \leq N$.

Then $-M < 2 - 2N \leq x + y - 2z \leq 2N - 2 < M$, so $x + y - 2z = 0$. This is a 3AP in A, giving a contradiction.

Hence A in $\mathbb{Z}/M\mathbb{Z}$ has no nontrivial 3APs, so by **2.14**,

$$|A| \ll \left(\frac{\log\log M}{\log M}\right)^{1/2} M \ll \left(\frac{\log\log N}{\log N}\right)^{1/2} N$$

□

To prove this, we will proceed as in $\mathbb{F}_p^n$ - prove the density increment, and then iteratively apply it.

Lemma 2.15. *Let B be a regular Bohr set of rank d and width ρ. Let $A \subseteq B$ with density $\alpha = \frac{|A|}{|B|}$, and no non-trivial 3APs.*

There is a constant $c > 0$ such that either:

1. $|A| \ll (d/\alpha)^{O(d)}|B|^{1/2}$
2. *There is a density increment: a regular Bohr set $B' \subseteq B$ of rank $\leq d+1$ and width $\gg (\alpha/d)^{O(1)}\rho$ and x such that*
$$\frac{|(A-x)\cap B'|}{|B'|} \geq (1+c\alpha)\alpha$$

*Proof of **2.14**, assuming **2.15**.* Let $A \subseteq G$ with density α, no non-trivial 3APs. Assume WLOG that $\alpha \geq \frac{1}{\log N}$. Let $k \geq 0$ be maximal such that there is a sequence of sets $A_0, \ldots, A_k$, regular Bohr sets $B_0, \ldots, B_k$ with ranks $d_0, \ldots, d_k$, widths $\rho_0, \ldots, \rho_k$ such that:

1. $A_0 = A, B_0 = G, d_0 = 1, \rho_0 = 2$
2. $A_i \subseteq B_i$
3. A_i has no non-trivial 3APs
4. $\alpha_i = |A_i|/|B_i|$, then $\alpha_{i+1} \geq (1+c\alpha_i)\alpha_i$, where c is the constant from **2.15**
5. $d_i \leq i+1$
6. $\rho_{i+1} \gg (\alpha/d_i)^{O(1)}\rho_i$

Just as in the $\mathbb{F}_p^n$ case, (4) implies $k \ll \alpha^{-1}$.

We now apply **2.15** to $A_k \subseteq B_k$. By maximality of k, the second condition of **2.15** cannot hold, so $\alpha_k \ll (d_k/\alpha)^{O(d_k)}|B_k|^{-1/2}$. $d_k \leq k+1 \ll \alpha^{-1}$, and so $1/\log N \leq \alpha \leq \alpha_k \ll \alpha^{-O(\alpha^{-1})}|B_k|^{-1/2}$

Compart this to our lower bound for $|B_k|$. For $1 \leq i \leq k$, $\rho_i \gg \alpha^{O(1)}\rho_i$.

So $\rho_k \gg \alpha^{O(d_k)} \times 2 \gg \alpha^{O(\alpha^{-1})}$.

Hence, $|B_k| \geq (\rho_K/8)^{d_k}N \geq \alpha^{O(\alpha^{-2})}N$, and thus we have:

$$\frac{1}{\log N} \ll \alpha^{-O(\alpha^{-1})} \times \alpha^{-O(\alpha^{-2})} \times N^{-1/2} \ll \alpha^{-O(\alpha^{-2})}N^{-1/2}$$

Rearranging, $\alpha^{-2}\log(1/\alpha) \gg \log N$, and since $\alpha \geq \frac{1}{\log N}$, $\log(1/\alpha) \ll \log\log N$, and so $\alpha^{-2} \gg \frac{\log N}{\log\log N}$, or $\alpha \ll \left(\frac{\log\log N}{\log N}\right)^{-1/2}$. □

It remains to prove **2.15**, but first, 2 technical lemmas. The first compensates for not being able to count $\langle \mathbb{1}_B * \mathbb{1}_A, \mathbb{1}_{2\cdot A}\rangle$ - recall in the $\mathbb{F}_p^n$ proof we used that this is equal to $\alpha^2|V|^2$. We will show that we can recover a suitable lower bound if instead we replace $2\cdot A$ by $2\cdot B_\delta$ for δ small enough, or we get a density increment.

Lemma 2.16. *Let B be a regular Bohr set of rank d and width ρ. Suppose that $\delta \leq c_0\frac{\alpha}{d}$ for some absolute constant $c_0 > 0$, and that also B_δ is regular.*

Let $A \subseteq B$ with density $\alpha = |A|/|B|$. Then either:

1. $\langle \mathbb{1}_A * \mathbb{1}_A, \mathbb{1}_{2\cdot B_\delta}\rangle \geq \frac{1}{2}\alpha^2|B||B_\delta|$, *or:*

2. *There is a regular B' of rank d, width $\gg \delta^2\rho$ and x such that $|(A-x)\cap B'|/|B'| \geq (1+\frac{1}{256})\alpha$.*

Proof. If condition 1 fails, then

$$\frac{1}{2}\alpha^2|B||B_\delta| > \langle \mathbb{1}_A * \mathbb{1}_A, \mathbb{1}_{2\cdot B_\delta}\rangle = \langle \mathbb{1}_A, \mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A\rangle$$

Roughly, this suggests there can't be too many elements of A where $\mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A$ is large. Let $A = A_{\text{large}} \sqcup A_{\text{small}}$ where $A_{\text{large}} = \{x \in A : \mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A(x) > \frac{3}{4}\alpha|B_\delta|\}$.

We then have

$$\frac{1}{2}\alpha^2|B||B_\delta| > \langle \mathbb{1}_A, \mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A\rangle \geq |A_{\text{large}}| \times \frac{3}{4}\alpha|B_\delta|$$

and so $|A_{\text{large}}| < \frac{2}{3}|A|$. In particular, $|A_{\text{small}}| \geq \frac{1}{3}|A|$.

So we know there are 'many' elements of A where $\mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A$ is small. We know upgrade this to find many elements of B where this is so.

Let $c \in [\frac{1}{2}, 1]$ be such that $B_{c\delta^2}$ is regular. We note that, by regularity of B_δ, for any $z \in B_{c\delta^2}$, we have:

$$|(2\cdot B_\delta - 2z) \setminus 2\cdot B_\delta| = |(B_\delta - z) \setminus B_\delta| \leq |B_{(1+c\delta)\delta} \setminus B_\delta| \ll \delta d|B_\delta|$$

(assuming $c\delta \leq \frac{1}{200d}$, which can be assumed by taking c_0 small.

Hence for any $y \in A_{\text{small}}$, $\mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A(y+2z) = |(2\cdot B_\delta - 2z) \cap (A+y)| \leq |2\cdot B_\delta \cap (A+y)| + O(\delta d|B_\delta|)$, which is equal to $\mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A(y) + O(\delta d|B_\delta|)$. In particular, for any $x \in A_{\text{small}} + 2\cdot B_{c\delta^2}$, since $\delta \leq c_0\frac{\alpha}{d}$, provided c_0 is sufficiently small, we have

$$\mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A(x) \leq \mathbb{1}_{2\cdot B_\delta} * \mathbb{1}_A(y) + O(\delta d|B_\delta|) \leq \frac{3}{4}\alpha|B_\delta| + O(c_0\alpha|B_\delta|) \leq \frac{7}{8}\alpha|B_\delta|$$

Now let $B_{\text{small}} = (A_{\text{small}} + 2\cdot B_{c\delta^2}) \cap B$, so that $x \in B_{\text{small}} \implies \mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A(x) \leq \frac{7}{8}\alpha|B_\delta|$.

We now split into cases:

1. $|B_{\text{small}} \leq \frac{1}{16}|B|$. Here consider $\langle \mathbb{1}_{A_{\text{small}}} * \mathbb{1}_{2\cdot B_{c\delta^2}}, \mathbb{1}_B\rangle = (*)$. By regularity and the fact that $2\cdot B_{c\delta^2} \subseteq B_{2c\delta^2}$, we have:

$$(*) = |A_{\text{small}}||B_{c\delta^2}| + O(\delta^2 d|B||B_{c\delta^2}|) \geq \frac{1}{3}|A||Bc\delta^2| + \frac{1}{10}\alpha|B||B_{c\delta^2}|$$

provided that $\delta^2 \leq c'\frac{\alpha}{d}$ for some small enough c', which we can make true, and so in particular:

$$(*) \geq \frac{1}{8}|A||B_{c\delta^2}|$$

On the other hand, we have an upper bound for this: $\mathbb{1}_{A_{\text{small}}} * \mathbb{1}_{2\cdot B_{c\delta^2}}$ is supported on B_{small}, so

$$(*) \leq |B_{\text{small}}| \max_x(\mathbb{1}_{A_{\text{small}}} * \mathbb{1}_{2\cdot B_{c\delta^2}}(x)) \leq \frac{1}{16}|B| \max_x(\mathbb{1}_A * \mathbb{1}_{2\cdot Bc\delta^2})(x))$$

so there is an x such that

$$|(A-x) \cap 2\cdot B_{c\delta^2}| = \mathbb{1}_A * \mathbb{1}_{2\cdot Bc\delta^2}(x) \geq 2\alpha|B_{c\delta^2}|$$

and we have a density increment $B' = 2\cdot B_{c\delta^2}$.

2. Otherwise, we have $|B_{\text{small}}| \geq \frac{1}{16}|B|$. Here, consider $(*) = \langle \mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A, \mathbb{1}_B \rangle$. By regularity, provided δ is small, then $(*) = |A||B_\delta| - O(\delta d|B_\delta||B|) \geq (1 - \frac{1}{256})|A||B_\delta|$.

   For an upper bound, $(*) \leq |B_{\text{small}}|\frac{7}{8}\alpha|B_\delta| + (|B| - |B_{\text{small}}|)\max_x \mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A(x)$.

   If $\max_x \mathbb{1}_{2\cdot B_\delta} \circ \mathbb{1}_A(x) \geq (1+\frac{1}{256})\alpha|B_\delta|$, then we have a density increment $B' = 2\cdot B_\delta$, otherwise:

$$(*) \leq |B_{\text{small}}|\frac{7}{8}\alpha|B_\delta| + (|B| - |B_{\text{small}}|)(1 + \frac{1}{256})\alpha|B_\delta| \geq (1 - \frac{1}{256})|A||B_\delta|$$

   Rearranging gives:

$$(1/8 + 1/256)|B_{\text{small}}| \leq \frac{1}{128}|B|$$

   but also $\geq \frac{1}{16} \times (\frac{1}{8} + \frac{1}{256})|B|$, giving a contradiction.

□

This lemma shows the need to replace $\langle \mathbb{1}_A * \mathbb{1}_A, \mathbb{1}_{2\cdot A} \rangle$ by $\langle \mathbb{1}_A * \mathbb{1}_A, \mathbb{1}_{2\cdot B_\delta} \rangle$, but $A \subseteq B$ not B_δ, and we don't know that $A \cap B_\delta$ is large, or even non-empty, unlike in the $\mathbb{F}_p^n$ case. We will show that there is some translate of A which is large enough in both a narrowed copy of B and a doubly narrowed copy of B.

Lemma 2.17. *Let $\varepsilon > 0$. Let B be regular, of rank d, and $A \subseteq B$ with density α. Suppose $B', B'' \subseteq B_\delta$ where $\delta \leq c_0\alpha\varepsilon/d$ for some small absolute constant $c_0 > 0$. Then either:*

1. *There is x such that $\frac{|(A-x)\cap B'|}{|B'|} \geq (1-2\varepsilon)\alpha$ and $\frac{|(A-x)\cap B''|}{|B''|} \geq (1-2\varepsilon)\alpha$, or*
2. *There is y such that either $\frac{|(A-y)\cap B'|}{|B'|} \geq (1+\varepsilon)\alpha$ or $\frac{|(A-y)\cap B''|}{|B''|} \geq (1+\varepsilon)\alpha$.*

Proof. By regularity,

$$\langle \mathbb{1}_A * \mathbb{1}_{B;}, \mathbb{1}_B \rangle = \langle \mathbb{1}_A, \mathbb{1}_B \circ \mathbb{1}_{B'} \rangle = |A||B'| + O(\delta d|B||B'|) \geq (1 - \varepsilon/2)\alpha|B||B'|$$

if δ is small enough, and similarly $\langle \mathbb{1}_A * \mathbb{1}_{B''}, \mathbb{1}_B \rangle \geq (1 - \varepsilon/2)\alpha|B||B'|$.

Let $\mu_{B'} = \frac{1}{|B'|}\mathbb{1}_{B'}$, and $\mu_{B''} = \frac{1}{|B''|}\mathbb{1}_{B''}$, so adding bounds together, we have:

$$\langle \mathbb{1}_A * \mu_{B'} + \mathbb{1}_A * \mu_{B''}, \mathbb{1}_B \rangle \geq (2\varepsilon)\alpha|B|$$

and so by averaging there is some $x \in B$ such that $\mathbb{1}_A * \mu_{B'}(x) + \mathbb{1}_A * \mu_{B''}(x) \geq (2-\varepsilon)\alpha$.

If $\mathbb{1}_A * \mu_{B'}(x) \geq (1+\varepsilon)\alpha$, in the second case, similarly if $\mathbb{1}_A * \mu_{B''}(x) \geq (1+\varepsilon)\alpha$, so if not, then $\mathbb{1}_A * \mu_{B'}(x) \geq (1-2\varepsilon)\alpha$ and $\mathbb{1}_A * \mu_{B''}(x) \geq (1-2\varepsilon)\alpha$. □

*Proof of **2.15**.* Let $A \subseteq B$ of density α, and B regular of rank d and width ρ. We will need some layers of Bohr sets:

- $B^{(1)} = B_{\delta_1}$
- $B^{(2)} = B^{(1)}_{\delta_2} = B_{\delta_1\delta_2}$

where $\delta_i = c_i\alpha^2/d$ where c_i are small absolute constants chosen so that $B^{(1)}$ and $B^{(2)}$ are regular by Bourgain's regularity lemma.

Then $A \subseteq B \subset B^{(1)} \subset B^{(2)}$.

We start by applying **2.17** with $B' = B^{(1)}$, $B'' = B^{(2)}$, and $\varepsilon = c\alpha$, where c is a small constant we'll choose later. If the second case holds, then we have our density increment and are done.

Otherwise, there is x such that, if $A_1 = (A - x) \cap B^{(1)}$ and $A_2 = (A - x) \cap B^{(2)}$. Let $\alpha_1 = \frac{|A_1|}{|B^{(1)}|}$ and $\alpha_2 = \frac{|A_2|}{|B^{(2)}|}$. Then $\min(\alpha_1, \alpha_2) \geq (1 - 2\varepsilon)\alpha$. In particular, since $\varepsilon \leq \frac{1}{4}$, we have $\alpha_1, \alpha_2 \geq \frac{\alpha}{2}$. Also, we can assume $\max(\alpha_1, \alpha_2) \leq 2\alpha$, or else we have a density increment.

Because A has no 3APs and 3APs are translation invariant, there are no non-trivial solution to $x + y = 2z$ with $x, y \in A_1, z \in A_2$, and so $\langle \mathbb{1}_{A_1} * \mathbb{1}_{A_1}, \mathbb{1}_{2\cdot A_2}\rangle = |A_2|$. By **2.16**, we have a density increment or $\langle \mathbb{1}_{A_1} * \mathbb{1}_{A_1}, \mathbb{1}_{2\cdot B^{(2)}}\rangle \geq \frac{1}{2}\alpha_1^2|B^{(1)}||B^{(2)}|$.

So if $f = \mathbb{1}_{2\cdot A_2} - \alpha_2\mathbb{1}_{2\cdot B^{(2)}}$, then $\langle \mathbb{1}_{A_1} * \mathbb{1}_{A_1}, f\rangle \leq |A_2| - \frac{1}{2}\alpha_1^2|B^{(1)}||A_2| \leq -\frac{1}{4}\alpha^2|B_1||A_2$ or $\alpha_1 < 2|B_1|^{-1/2}$, but by repeated applications of **2.8**, $|B^{(1)}| \geq (\delta_1)^{O(d)}|B|$, and so $\alpha \leq 2\alpha_1 \ll (\delta_1)^{-O(d)}|B|^{-1/2}$, and so $\alpha \ll (d/\alpha)^{-O(d)}|B|^{-1/2}$, and we are in the first case of **2.15**.

So we can assume $|\langle \mathbb{1}_{A_1} * \mathbb{1}_{A_1}, f\rangle| \geq \frac{1}{4}\alpha_1^2|B^{(1)}||A_2|$, and we want a density increment.

As in the $\mathbb{F}_p^n$ case we apply Parseval + triangle inequality to get:

$$\mathop{\mathbb{E}}_{\gamma\in\widehat{G}} |\widehat{f}(\gamma)||\widehat{\mathbb{1}_{A_1}}(\gamma)|^2 \gg \alpha_1^2|B^{(1)}||A_2|$$

and since $\mathop{\mathbb{E}}_{\gamma\in\widehat{G}} |\widehat{\mathbb{1}_{A_1}}(\gamma)|^2 = |A_1| = \alpha_1|B^{(1)}|$ by Parseval, and some $\lambda \in \widehat{G}$ such that $|\widehat{f}(\lambda)| \gg \alpha_1|A_2|$.

Note that, if $f_A = \mathbb{1}_{A_2} - \alpha_2\mathbb{1}_{B^{(2)}}$ then $f(2x) = f_A(x)$, so $\widehat{f}(2\lambda) = \sum_x f(2x)\overline{\lambda(2x)} = \sum_y f(y)\overline{\lambda(y)} = \widehat{f}(\lambda)$, and so there is γ with $|\widehat{f_A}(\gamma)| \gg \alpha_1|A_2|$.

Now let B' be the Bohr set obtained by adding γ to the frequency set of $B^{(2)}$, so rank of B' is $\leq d + 1$, and multiply the width of $B^{(2)}$ by $c_3\alpha^2/d$ where $c_3 > 0$ is a small constant chosen such that B' is regular. First replace $f_A = \mathbb{1}_{A_2} - \alpha\mathbb{1}_{B^{(2)}}$ by $f'_A = \mathbb{1}_{A_2} - \alpha\mathbb{1}_{B^{(2)}+B'}$ by regularity. Namely, $|\widehat{f_A}(\gamma) - \widehat{f_{A'}}(\gamma)| \leq \alpha_2|(B^{(2)} + B') \setminus B^{(2)}| \ll c_3\alpha_2\alpha^2|B^{(2)}|$.

In particular, we can assume $|\widehat{f_{A'}}(\gamma)| \gg \alpha_1|A_2|$. □

3 Almost-Periodicity

Let τ_t be the "translation by t operator", so that if $f : G \to \mathbb{C}$, then $\tau_t f : G \to \mathbb{C}$ is $\tau_t f(x) = f(x-t)$. A ***period*** of f is some t such that $\tau_t f = f$. For example, 0 is a period of every f. Or 2π is a period of $\sin(x)$. More relevant here, if $H \leq G$, then any $t \in H$ is a period for $\mathbb{1}_H$.

Most functions however have no non-zero periods. We instead consider $\tau_t f \approx f$, in which case t is an almost-period, for some definition of $\approx$. We are particularly interested in almost-periods of convolutions. This is because a sum set $A + B$ is just the support of $\mathbb{1}_A * \mathbb{1}_B$, so if there is some x such that $\mathbb{1}_A * \mathbb{1}_B(x) > 0$, and if t is an almost-period for $\mathbb{1}_A * \mathbb{1}_B$, then $\mathbb{1}_A * \mathbb{1}_B(x-t) \approx \mathbb{1}_A * \mathbb{1}_A(x) > 0$,

and so $x - t \in A + B$, so if T is the set of almost-periods, then $x - T \subseteq A + B$. This makes almost-periodicity very useful for finding structure in sum sets, for example "If $A \subset \{1, \ldots, N\}$ where $|A|$ is large, then how long an AP must $A + A$ contain?"

The modern treatment of almost-periodicity dates back to Croot and Sisask, around 2010, and is a sort of 'natural complement' to Fourier analysis.

We can think of almost-periodicity as a sort of continuity. If $f : \mathbb{R} \to \mathbb{R}$ is uniformly continuous, then for all $\varepsilon > 0$ there is δ such that, for $|t| < \delta$, we have $\|\tau_t f - f\|_\infty < \varepsilon$, so the ball $|t| < \delta$ is a set of L^∞-almost periods, with error ε. We will allow any kind of L^p-norm here (where $\|f\|_p = (\sum_{x \in G} |f(x)|^p)^{1/p}$).

These are norms, so they satisfy the triangle inequality, and are invariant under τ_t.

Definition 3.1. *Let $1 \leq p \leq \infty$ and $X > 0$. For any $f : G \to \mathbb{C}$, the set of L^p-almost-periods for f with error X is defined to be:*

$$AP_p(f; X) = \{t \in G : \|\tau_t f - f\|_p \leq x\}$$

For any p, f, X, this AP set is symmetric and contains 0, as $\|\tau_{-t} f - f\| = \|\tau_t(\tau_{-t} f - f)\| = \|f - \tau_t f\| = \|\tau_t f - f\|$. These sets are increasing in X, and for any p, f, for large enough X, $AP_p(f; X) = G$.

We also have $AP_p(f; X_1) + AP_p(f; X_2) \subseteq AP_p(f; X_1 + X_2)$ from the triangle inequality.

We will prove two main results concerning almost periods.

- For any f with $\|\widehat{f}\|_1 = \mathbb{E}_\gamma |\widehat{f}(\gamma)|$ small, there is a large Bohr set of almost periods.

Theorem 3.2. *Let $\varepsilon > 0$ and m some integer ≥ 1. For any $f : G \to \mathbb{C}$, the set of L^{2m}-almost periods of f with error $\varepsilon \|\widehat{f}\|_1 N^{\frac{1}{2m}}$ contains a Bohr set of rank $O(m\varepsilon^{-2})$ and width $\gg \varepsilon$.*

- Instead of finding a Bohr set of almost periods, we just show there are many almost periods, in the case $f = \mathbb{1}_A * \mathbb{1}_B$ where A is structured.

Theorem 3.3 ("Dense Global Version")**.** *Let $\varepsilon > 0, m \geq 1$. Suppose that $A \subseteq G$ with density $\alpha = |A|/N$. For any set B:*

$$|AP_{2m}(\mathbb{1}_A * \mathbb{1}_B; \varepsilon |A||B|^{\frac{1}{2m}})| \geq \alpha^{O(m\varepsilon^{-2})} N$$

Note that $\|\mathbb{1}_A * \mathbb{1}_B\|_{2m}^{2m} = \sum_x (\sum_{a \in A} \mathbb{1}_B(x - a))^{2m} \leq |A|^{2m-1} \left(\sum_x \sum_{a \in A} \mathbb{1}_{B(x-a)}\right) = |A|^{2m}|B|$, by Hölder's inequality.

So trivially we always have

$$\|\tau_t \mathbb{1}_A * \mathbb{1}_B - \mathbb{1}_A * \mathbb{1}_B\|_{2m} \leq 2|A||B|^{\frac{1}{2m}}$$

and so this is non-trivial for $\varepsilon < 2$.

Theorem 3.4 ("Local Version")**.** *Let $\varepsilon > 0, m \geq 1$, and $|A + S| \leq K|A|$. Then for any B,*

$$|AP_{2m}(\mathbb{1}_A * \mathbb{1}_B; \varepsilon |A||B|^{\frac{1}{2m}})| \geq K^{-O(m\varepsilon^{-2})}|S|$$

Note that **3.4** implies **3.3** by letting $S = G$.

3.1 Applications of Almost-Periodicity

An essential tool will be Hölder's inequality: for all $a_i, b_i \in \mathbb{R}$, and any $p \geq 1$,

$$\left|\sum a_i b_i\right| \leq \|a_i\|_p \|b_i\|_p$$

At $p = 2$ this is just Cauchy-Schwarz. $p = 1$ is trivial.

In particular, for any $f : G \to \mathbb{C}$ and any $A \subseteq G$, for any $m \geq 1$,

$$|\langle \mathbb{1}_A, f\rangle| \leq |A|^{1-1/2m} \|f\|_{2m}$$

We will show that, if A is a set with small doubling, then $A + A - A - A$ is very structured, and in particular contains k-fold sumsets for large k. E.g. if $|A + A| \ll |A|$ then there is A' such that $|A'| \gg |A|$ and $100A' \subseteq A + A - A - A$.

Theorem 3.5. *If $|A + A| \leq K|A|$ then for any $k \geq 1$ there is a symmetric set T such that $0 \in T$ and $|T| \geq \exp(-O(k^2(\log K)^2))|A|$ and $kT = T + \ldots + T \subseteq A + A - A - A$.*

Proof. We apply **3.4** with $S = A$ and $B = A - A$, and $\varepsilon > 0, m \geq 1$ to be chosen later.

Then if T is the set of L^{2m}-almost periods for $\mathbb{1}_A * \mathbb{1}_{A-A}$ with error $\varepsilon|A||A - A|^{1/2m}$ then $|T| \geq K^{-O(m\varepsilon^{-2})}|A|$ since $|A + S| = |A + A| \leq K|A|$ by assumption.

Addition of almost-periods adds the errors, so if $t \in kT$ then

$$\|\tau_t(\mathbb{1}_A * \mathbb{1}_{A-A}) - \mathbb{1}_A * \mathbb{1}_{A-A}\|_{2m} \leq k\varepsilon|A||A - A|^{1/2m}$$

Suppose for contradiction that there is $t \in kT$ such that $t \notin A+A-A-A$. Then $\mathbb{1}_A * \mathbb{1}_{A-A} \circ \mathbb{1}_A(t) = 0$, and hence $\langle \mathbb{1}_A, \tau_t(\mathbb{1}_A * \mathbb{1}_{A-A})\rangle = 0$.

But $\langle \mathbb{1}_A, \mathbb{1}_A * \mathbb{1}_{A-A}\rangle = \sum_{a,b\in A} \sum_{c\in A-A} \mathbb{1}_{a-b=c} = |A|^2$.

Hence:

$$\begin{aligned}
|\langle \mathbb{1}_A, (\tau_t(\mathbb{1}_A * \mathbb{1}_{A-A}) - \mathbb{1}_A * \mathbb{1}_{A-A})\rangle &= |A|^2 \\
&\leq |A|^{1-1/2m} \|\tau_t(\mathbb{1}_A * \mathbb{1}_{A-A}) - \mathbb{1}_A * \mathbb{1}_{a-A}\|_{2m} \\
&\leq k\varepsilon|A|^2 \left(\frac{|A - A|}{|A|}\right)^{1/2m}
\end{aligned}$$

By Plünnecke, $|A - A| \leq K^2|A|$, and so this is $\leq k\varepsilon|A|^2 K^{1/m}$, so if we choose $m = \lceil \log K \rceil$, this is $\leq k\varepsilon|A|^2 e$.

So $|A|^2 \leq k\varepsilon e|A|^2$, giving a contradiction if we choose $\varepsilon = \frac{1}{2ek}$. Hence $kT \subseteq A + A - A - A$ with this choice of ε, m, and

$$|T| \geq K^{-O(m\varepsilon^{-2})}|A| \geq \exp(-O(k^2(\log K)^2))|A|$$

□

The second application will be using **3.2** (finding a Bohr set of APs), to answer the question "Given a large subset $A \subseteq \{1, \ldots, N\}$, how large an arithmetic progression must $A + A$ contain?"

Last chapter, we showed A contained an AP of length 3. Here we can do much better, replacing A by $A + A$.

Theorem 3.6. *There exist constants $c > 0, C \geq 1$ such that, if $A \subset \{1, \dots N\}$ has size $|A| = \alpha N$ where $\alpha \geq C \frac{\log\log(N)}{\sqrt{\log N}}$, then $A + A$ contains an AP of length $\gg \exp(c\alpha\sqrt{\log N})$.*

This is already non-trivial and interesting in the case $\alpha = \frac{1}{4}$. This proof is due to Croot, Laba, and Sisusk (2011). This was first proved by Ben Green using Fourier methods. The dependence on N is still the best-known. A construction of Ruzsa gives a subset $A \subseteq \{1, \dots, N\}$ of size $\geq \frac{1}{4}N$ such that the longest AP in $A + A$ has length $\ll \exp(O((\log N)^{2/3} \log\log N))$.

Before we prove this, we will first prove:

Lemma 3.7. *If N is prime and $B \subseteq \mathbb{Z}/N\mathbb{Z}$ is a Bohr set of rank d and width ρ, then B contains an AP of length $\gg \rho N^{1/d}$.*

Proof. Let $\ell \geq 1$ be maximal such that $(\rho/8\ell)^d \geq 2/N$. Note that $\ell \gg \rho N^{1/d}$.

We know that $|B_{1/2}| \geq 8^{-d}|B|$, so for any $k \geq 1$, $|B_{1/2^k}| \geq 8^{-kd}|B|$.

So, choosing k such that $1/2^{k+1} \leq 1/\ell < 1/2^k$, then $|B_{1/\ell}| \geq |B_{1/2^{k+1}}| \geq (1/2)^{3kd}|B| \geq (1/\ell)^{3d}|B|$.

We have $|B_{1/\ell}| \geq (\rho/8\ell)^d N \geq 2$, so in particular there is $x \in B_{1/\ell}, x \neq 0$. By the triangle inequality, $\ell B_{1/\ell} \subseteq B$ and in particular $P = \ell\{0, x\} = \{0, \ell, \dots, \ell x\} \subseteq B$. □

*Proof of **3.6**.* It suffices to prove the result with $\{1, \dots, N\}$ replaced with $\mathbb{Z}/N\mathbb{Z}$ with N prime. Let $4N < M \leq 8N$ be prime, which exists by Bertrand's postulate.

Then $A \subseteq \{1, \dots, N\} \implies A + A \subseteq \{1, \dots, 2N\}$.

We claim that any AP modulo M inside $A + A$ is a genuine AP in the integers.

Let P be an AP modulo M inside $A + A$, say $\{x_0, \dots, x_k\} \subseteq A + A$ where $x_i \equiv x + id \mod m$, for some $d \in \mathbb{Z}$. Then $d \equiv x_1 - x_0 \mod M$, and wlog $d = x_2 - x_0$.

By induction, for any $k \geq i \geq 0$, $x_i = x + id$: suppose not, and let $i \geq 2$ be minimal such that this fails. Then $x_i - (x + id) = x_i - (x_{i-1} + d) = x_i + x_0 - x_{i-1} - x_1 \in \{-4N, \dots, 4N\}$. So $|x_i - (x + id)| < M$, and so $x_i \equiv x + id \mod M$ implies $x_i = x_i d$. So P is indeed a genuine AP of the same length. Hence it suffices to prove the following:

There exist constants $c > 0, C \geq 1$ such that, for all N prime, if $A \subseteq \mathbb{Z}/N\mathbb{Z}$ with $|A| = \alpha N$ where $\alpha \geq C \frac{\log\log(N)}{\sqrt{\log N}}$, then $A + A$ contains an AP of length $\gg \exp(c\alpha\sqrt{\log N})$.

We use **3.1** to find a large Bohr set of almost-periods for $\mathbb{1}_A * \mathbb{1}_A$, find a progression P inside this Bohr set, then show some translate of P must be in $A + A$.

Let T be the set of L^{2m}-almost-periods for $\mathbb{1}_A * \mathbb{1}_A$ with error $\frac{\alpha}{4}|A|N^{1/2m}$ - this is the subject of **3.1** with $\varepsilon = \alpha/4$, since $\mathbb{E}_\gamma |\widehat{\mathbb{1}_A * \mathbb{1}_A}(\gamma)| = \mathbb{E}_\gamma |\widehat{\mathbb{1}_A}(\gamma)|^2 = |A|$ by Parseval.

So T contains a Bohr set of rank $O(m\alpha^{-2})$ and width $\gg \alpha$.

By Hölder, for any $f : G \times G \to \mathbb{C}$ and any $P \subseteq G$,

$$\begin{aligned}\sum_{x\in G} \sup_{t\in P} |f(x,y)| &\leq \sum_{x\in G} (\sum_{t\in P} |f(x,t)|^{2m})^{1/2m} \\ &\leq N^{1-1/2m} (\sum_{t\in P} \sum_{x\in G} |f(x,t)|^{2m})^{1/2m} \\ &\leq N^{1-1/2m} |P|^{1/2m} \sup_{t\in P} \|f(\cdot,t)\|_{2m}\end{aligned}$$

In particular, for any $P \subseteq T$,

$$\begin{aligned}\sum_{x\in G} \sup_{t\in P} |\mathbb{1}_A * \mathbb{1}_A(x+t) - \mathbb{1}_A * \mathbb{1}_A(x)| &\leq |P|^{1/2m} N^{1-1/2m} \sup_{t\in P} \|\tau_t(\mathbb{1}_A * \mathbb{1}_A) - \mathbb{1}_A * \mathbb{1}_A\|_{2m} \\ &\leq |P|^{1/2m} N^{1-1/2m} \frac{\alpha}{4} |A| N^{1/2m} \\ &= \frac{1}{4} |A|^2 |P|^{1/2m}\end{aligned}$$

In particular, if $|P| < 4^{2m}$, this is $< |A|^2$, and $|A|^2 = \sum_{x\in G} \mathbb{1}_A * \mathbb{1}_A(x)$, so by pigeonhole, there is some $x \in G$ such that $\sup_{t\in P} |\mathbb{1}_A * \mathbb{1}_A(x+t) - \mathbb{1}_A * \mathbb{1}_A(x)| < \mathbb{1}_A * \mathbb{1}_A(x)$.

So $\mathbb{1}_A * \mathbb{1}_A(x+t) \neq 0$ for all $t \in P$, and so $x + P \subseteq A + A$. We just need to choose m such that T contains a progression P of length $|P| < 4^{2m}$. By the above lemma, T contains a Bohr set of rank $O(m\alpha^{-2})$ and width $\gg \alpha$, which contains a progression P of length $\geq c_1 \alpha N^{c_2\alpha^2/m}$, for some constants $c_1, c_2 > 0$. So we need m such that $N^{\alpha^2/m} \ll C^m$, i.e. $m \gg \alpha\sqrt{\log N}$.

So, choosing $m = \lfloor \alpha\sqrt{\log N} \rfloor$, then $m \geq 1$ by assumption, and there is a progression of length $\ell = \lfloor c_1 \alpha \exp(c_2 \alpha \sqrt{\log N}) \rfloor$, and in particular, $\ell \leq \exp(c_2\alpha\sqrt{\log N}) \leq 4^{2m}$, and so we have an AP of length ℓ inside $A + A$. Noting that $\lfloor x \rfloor \geq x/2$ if $x \geq 1$, and $\alpha \geq C \frac{\log\log N}{\sqrt{\log N}}$ for large $C \geq 1$, and so $c_1 \alpha \exp(c_2\alpha\sqrt{\log N}) \geq \exp(c_2/2\alpha\sqrt{\log N}) \geq 1$, we have the result. □

3.2 A Probabilistic Inequality

In this section we will use expectation in the probabilistic sense. All random variables are assumed to be taken over a finite probability space, so that $\mathbb{E}\, X = \sum_{s\in S} \mu(s) X(s)$.

Suppose we have n independent random variables $X_1, \ldots, X_n$. Can we control the L^p norm of $\sum X_i$ by the L^p norms of the X_i? Crudely, the answer is yes, via Hölder's inequality + the triangle inequality. We can say:

$$\mathbb{E}\, |\sum_i X_i|^p \leq n^{p-1} \mathbb{E} \sum_i |X_i|^p = p^{p-1} \sum_i \mathbb{E}\, |X_i|^p$$

For our applications, we want to improve this n^{p-1} factor. We can do this if the random variables have expectation 0 - then we can improve by a factor of $(2p/n)^{p/2}$, which is a big saving. More precisely, we have:

Lemma 3.8 (Marcinkiewicz-Zygmund Inequality)**.** *Let $m \geq 1$. If $X_1, \ldots, X_n$ are independent complex valued R.V.s, over some finite probability space, with $\mathbb{E}\, X_i = 0$ for all i. Then:*

$$\mathbb{E}\, |\sum_i X_i|^{2m} \leq (4m)^m \mathbb{E} \left(\sum_i |X_i|^2 \right)^m$$

In particular,

$$\mathbb{E}\,|\sum_i X_i|^{2m} \le (4m)^m n^{m-1}\,\mathbb{E}\sum_i |X_i|^{2m}$$

Proof. The second inequality follows from the first by Hölder.

Our goal is to bound $S = \mathbb{E}\,|\sum_i X_i|^m$.

Introduce Y_i, distributed identically to X_i by independent. Then $\mathbb{E}\,Y_i = 0$ also, and so:

$$S = \mathop{\mathbb{E}}_{X_i}|\sum_i X_i|^{2m} = \mathop{\mathbb{E}}_{X_i}|\sum_i (X_i - \mathop{\mathbb{E}}_{Y_i} Y_i)|^{2m} = \mathop{\mathbb{E}}_{X_i}|\mathop{\mathbb{E}}_{Y_i}(\sum_i X_i - Y_i)|^{2m}$$

By the triangle inequality, this is bounded above by:

$$\mathop{\mathbb{E}}_{X_i,Y_i}|\sum_i X_i - Y_i|^{2m}$$

Now $X_i - Y_i$ has the same distribution as $Y_i - X_i$, and so for any choice of signs $\varepsilon_i \in \{-1,+1\}$, we have

$$\mathop{\mathbb{E}}_{X_i,Y_i}|\sum_i (X_i - Y_i)|^{2m} = \mathop{\mathbb{E}}_{X_i,Y_i}|\sum_i \varepsilon_i(X_i - Y_i)|^{2m}$$

Now choose the ε_i uniformly at random. Then:

$$S \le \mathop{\mathbb{E}}_{\varepsilon_i}\mathop{\mathbb{E}}_{X_i,Y_i}|\sum_i \varepsilon_i(X_i - Y_i)|^{2m}$$

Now change the order of expectation. So for fixed X_i, Y_i, consider

$$\mathop{\mathbb{E}}_{\varepsilon_i}|\sum_i \varepsilon_i(X_i - Y_i)|^{2m}$$

Let $x_i = X_i - Y_i$. Expanding out, this is:

$$\mathop{\mathbb{E}}_{\varepsilon_i}\sum_{k_1+\ldots+k_n=2m}\binom{2m}{k_1,\ldots,k_n}x_1^{k_1}\ldots x_n^{k_n}\varepsilon_1^{k_1}\ldots\varepsilon_n^{k_n} = \sum_{k_1+\ldots+k_n=2m}\binom{2m}{k_1,\ldots,k_n}x_1^{k_1}\ldots x_n^{k_n}\varepsilon_1^{k_1}\left(\mathop{\mathbb{E}}_{\varepsilon_i}\varepsilon_1^{k_1}\ldots\varepsilon_n^{k_n}\right) = (*)$$

and this expectation factors as

$$\prod_{i=1}^n\left(\mathop{\mathbb{E}}_{\varepsilon_i}\varepsilon_i^{k_i}\right)$$

and clearly $\mathbb{E}_\varepsilon\,\varepsilon^k = \begin{cases}1 & k \text{ even}\\ 0 & k \text{ odd}\end{cases}$. Hence this expectation vanishes unless all the k are even. So the quantity of interest is equal to

$$(*) = \sum_{\ell_1+\ldots+\ell_n=m}\binom{2m}{2\ell_1,\ldots,2\ell_n}x_1^{2\ell_1}\ldots x_n^{2\ell_n}$$

We have a bound on the multinomial coefficients given by

$$\binom{2m}{2\ell_1,\ldots,2\ell_n} \le m^m\binom{m}{\ell_1,\ldots,\ell_n}$$

and so it follows that

$$(*) \leq m^m \sum_{\ell_1+\ldots+\ell_n=m} \binom{m}{\ell_1,\ldots,\ell_n} (x_1^2)^{\ell_1} \ldots (x_n^2)^{\ell_n} = m^m \left(\sum_i x_i^2\right)^m$$

In particular, we have:

$$S \leq m^m \mathop{\mathbb{E}}_{X_i,Y_i} \left(\sum_i |X_i - Y_i|^2\right)^m$$

Now C-S inequality gives

$$S \leq 2^m m^m \mathop{\mathbb{E}}_{X_i} \left(\sum_i |X_i|^2 + \sum_i |Y_i|^2\right)^m$$

Since $(a+b)^m \leq 2^{m-1}(a^m + b^m)$, we have

$$S \leq \frac{1}{2} 4^m m^m \mathop{\mathbb{E}}_{X_i,Y_i} \left(\left(\sum_i |X_i|^2\right)^m + \left(\sum_i |Y_i|^2\right)^m\right) = 4^m m^m \mathop{\mathbb{E}}_{X_i} \left(\sum_i |X_i|^2\right)^m$$

□

3.3 Almost-Periodicity via Random Sampling in Fourier Space

Here we aim to prove **Theorem 3.2**. The idea of the proof is to find some g such that g is the sum of not too many characters, and $g \approx f$ in a L^{2m} sense. Then take the Bohr set on the characters in G. If $t \in \mathrm{Bohr}(\gamma; \varepsilon)$, then t is an L^∞-almost period for γ with error ε, so for any $x \in G$:

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

By the triangle inequality, almost periods for g will also be almost periods for f.

How do we find g? We randomly choose $\gamma \in \widehat{G}$ and let $g(x) = \frac{1}{k}\sum_{i=1}^k \gamma_i(x)$.

Lemma 3.9. *Let $\varepsilon > 0$ and $m \geq 1$. Then for any $f : G \to \mathbb{C}$, there is $k = O(m\varepsilon^{-2})$ and $\gamma_1, \ldots, \gamma_k \in \widehat{G}$ not necessarily distinct with $c_i, \ldots, c_k \in \mathbb{C}, |c_i| = 1$, such that, if $g(x) = \frac{\|\widehat{f}\|_1}{k}\sum_{i=1}^k c_i\gamma_i(x)$, then $\|g - f\|_{2m} \leq \varepsilon\|\widehat{f}\|_1 N^{1/2m}$.*

Proof. Dilating f if necessary, we may take $\|\widehat{f}\|_1 = \frac{1}{N}$, and so $\sum_{\gamma\in\widehat{G}} |\widehat{f}(\gamma)| = 1$, so $|\widehat{f}(\gamma)|$ is a probability measure on $\widehat{G}$. Consider a random function $h(x) = \frac{1}{N}c_\gamma\gamma(x)$ where $\gamma \in \widehat{G}$ is chosen with probability $|\widehat{f}(\gamma)|$, and $c_\gamma \in \mathbb{C}$ such that $\widehat{f}(\gamma) = c_\gamma|\widehat{f}(\gamma)|$.

Then observe that, for any fixed $x \in G$,

$$\mathbb{E}\, h(x) = \sum_{\gamma\in\widehat{G}} |\widehat{f}(\gamma)| h(x) = \sum_{\gamma\in\widehat{G}} |\widehat{f}(\gamma)| \frac{1}{N} c_\gamma \gamma(x) = \frac{1}{N}\sum_{\gamma\in\widehat{G}} \widehat{f}(\gamma)\gamma(x) = f(x)$$

So $h(x) - f(x)$ is a random variable with mean zero, hence we may apply Marcinkiewicz-Zygmund to k copies of h, say $h_1, \ldots, h_k$. Then:

$$\mathbb{E}\left|\sum_{i=1}^k (h_i(x) - f(x))\right|^{2m} \leq (4m)^m k^{m-1} \mathbb{E}\sum_{i=1}^k |h_i(x) - f(x)|^{2m}$$

Note that, for any x,

$$|f(x)| = \frac{1}{N}\left|\sum_{\gamma\in\widehat{G}} \widehat{f}(\gamma)\gamma(x)\right| \leq \frac{1}{N}\sum_{\gamma} |\widehat{f}(\gamma)| = \frac{1}{N}$$

and also $|h_i(x)| = \frac{1}{N}$. So $|h_i(x) - f(x)| \leq 2/N$. So:

$$\mathbb{E}\left|\frac{1}{k}\sum_{i=1}^{k}(h_i(x) - f(x))\right|^{2m} \leq \frac{(4m)^m k^{m-1}}{k^{2m}}\,\mathbb{E}\sum_{i=1}^{k}\left(\frac{2}{N}\right)^{2m} = \left(\frac{4m}{k}\right)^m\left(\frac{2}{N}\right)^{2m} = \left(\frac{16m}{k}\right)^m N^{-2m}$$

Summing both sides over all $x \in G$, using linearity of expectation, we obtain

$$\mathbb{E}\sum_{x\in G}\left|\frac{1}{k}\sum_{i=1}^{k}(h_i(x) - f(x))\right|^{2m} \leq \left(\frac{16m}{k}\right)^m N^{-2m+1}$$

That is,

$$\mathbb{E}\|g - f\|_{2m}^{2m} \leq \left(\frac{16m}{k}\right)^m N^{-2m+1}$$

where $g = \frac{1}{k}\sum_{i=1}^{k} h_i(x)$. In particular, there exists some such g such that

$$\|g - f\|_{2m} \leq \left(\frac{16m}{k}\right)^{1/2} N^{-1+1/2m} = \left(\frac{16m}{k}\right)^{1/2} \|\widehat{f}\|_1 N^{1/2m}$$

Choose $k = \lceil \frac{16m}{\varepsilon^2} \rceil$, so this is $\leq \varepsilon\|\widehat{f}\|_1 N^{1/2m}$. □

Proof of **3.2**. Let $f : G \to \mathbb{C}$, $m \geq 1$, $\varepsilon > 0$, $k \ll m\varepsilon^{-2}$. By the previous lemma there are characters $\gamma_1, \ldots, \gamma_k \in \widehat{G}$ and twists $c_1, \ldots, c_k \in \mathbb{C}$, $|c_i| = 1$, such that if $g(x) = \frac{1}{k}\sum c_i\gamma_i(x)$ then

$$\|g - f\|_{2m} \leq \frac{1}{3}\varepsilon\|\widehat{f}\|_1 N^{1/2m}$$

Let $B = \mathrm{Bohr}(\{\gamma_1, \ldots, \gamma_k\}; \varepsilon/3)$, which has rank $\ll m\varepsilon^{-2}$ and width $\gg \varepsilon$ as required.

For any $t \in B$ and $x \in G$, by the triangle inequality,

$$|g(x+t) - g(x)| = \frac{\|\widehat{f}\|_1}{k}\left|\sum c_i(\gamma_i(x+t) - \gamma_i(x))\right| \leq \frac{\|\widehat{f}\|_1}{k}\sum_{i=1}^{k}|\gamma_i(x+t) - \gamma_i(x)| = \frac{\|\widehat{f}\|_1}{k}\sum_{i=1}^{k}|\gamma_i(t) - 1|$$

Now $|\gamma_i(t) - 1)| \leq \frac{\varepsilon}{3}$, so:

$$|g(x+t) - g(x)| \leq \frac{\varepsilon}{3}\|\widehat{f}\|_1$$

In particular, since, for any f,

$$\|f\|_{2m} = \left|\sum_x |f(x)|^{2m}\right|^{1/2m} \leq (\max|f(x)|)N^{1/2m}$$

and so

$$\|\tau_t g - g\|_{2m} \leq \|\tau_t g - g\|_\infty N^{1/2m} \leq \varepsilon/3\|\widehat{f}\|_1 N^{1/2m}$$

By the triangle inequality, we have

$$\|\tau_t f - f\|_{2m} \leq \varepsilon \|\widehat{f}\|_1 N^{1/2m}$$

□

The proof of the other almost-periodicity result goes via random sampling in physical space. Instead of approximating f by a sum of characters, we will approximate $\mathbb{1}_A * \mathbb{1}_B(x)$ by $\frac{|A|}{k} \sum_{i=1}^{k} \mathbb{1}_B(x - a_i)$, where a_i are randomly chosen from A uniformly. I.e., we replace the A in $\mathbb{1}_A * \mathbb{1}_B$ by a random subset of A of size k. Unlike above, there are no obvious almost periods for $\sum_{i=1}^{k} \mathbb{1}_B(x - a_i)$. We will use the small doubling hypothesis - we actually show that, for almost all $(a_1, \ldots, a_k)$, this convolution is $\approx \mathbb{1}_A * \mathbb{1}_B(x)$. Then show, using small doubling, that there must be many t and $(a_1, \ldots, a_k)$ such that $(a_1 + t, \ldots, a_k + t)$ is good for all such t.