Sum Product Theorems and Applications (Spring 2022, Weikun He)

Ajorda Jiao

Contents

Theorem 0.1 (Erdös-Szemerédi Theorem)
There exists an absolute constant $c > 0$, such that for every finite set $A \subseteq \mathbb{R}$,

$$\max\{\sharp(A+A), \sharp AA\} \geqslant c(\sharp A)^{1+c}.$$

§1 Basic additive combinatorics

$(E, +)$ abelian group. $A, B \subseteq E$.

Notation 1.1. $A + B \coloneqq \{a + b : a \in A, b \in B\}$.

> **Question 1.2** (Freiman)**.** If $\sharp(A + A) \leqslant K\sharp A$, for some parameter K, what can we say about A?

Observation 1.3. If A is a **arithmetic progression**, then $\sharp(A + A) \leqslant 2\sharp A$. If A is a **generalized A.P.** of **rank** r, i.e.

$$A = \{a_0 + t_1 d_1 + \cdots + t_r d_r : \forall i, 1 \leqslant t_i \leqslant N_i\},$$

then $\sharp(A + A) \leqslant 2^r \sharp A$.

Freiman Type Theorem If $\sharp(A + A) \leqslant K\sharp A$, then exists

(i) $P \subseteq E$ is a generalized arithmetic progression of rank $O_K(1)$, $\sharp P = O_K(\sharp A)$.

(ii) $X \subseteq E$ finite, $\sharp X = O_K(1)$.

Such that $A \subseteq P + X$.

Theorem 1.4 (Szemerédi)
$A \subseteq \mathbb{N}$ with positive upper density, then A contains arbitrarily long A.P.

Lemma 1.5 (Ruzsa Triangle Inequality)
$A, B, C \subseteq (E, +)$ finite, then

$$\sharp(A - C)\sharp B \leqslant \sharp(A - B)\sharp(B - C).$$

Proof. Construct a map $(A - C) \times B \to (A - B) \times (B - C), (x, b) \mapsto (a_x - b, b - c_x)$, where $x = a_x - b_x$ is a typical decomposition, this map is an injective. $\square$

Definition 1.6. Define the **Ruzsa distance** between A, B by

$$d(A, B) = \log \frac{\sharp(A - B)}{(\sharp A)^{\frac{1}{2}}(\sharp B)^{\frac{1}{2}}}.$$

Lemma 1.7 (Ruzsa Covering Lemma)
$A, B \subseteq (E, +)$ finite, $K \geqslant 1$. If $\sharp(A+B) \leqslant K\sharp A$, then $\exists X \subseteq E, \sharp X \leqslant K$, such that $B \subset A - A + X$.

Proof. Let $X \subseteq B$ be the maximal set such that $(x+A)_{x\in X}$ is pointwise disjoint. □

Notation 1.8. $\mathbb{O}(K)$ denotes some subset of cardinality $\leqslant K$.

Remark 1.9 — Ruzsa Covering Lemma $\iff$ $B \subseteq A - A + \mathbb{O}\left(\frac{\sharp(A+B)}{\sharp A}\right)$.

Proposition 1.10 (Plünnecke-Ruzsa Inequality)
$A, B \subseteq E$ finite, $K \geqslant 1$. If $\sharp(A+B) \leqslant K\sharp A$, then $\forall k, l \geqslant 0$, we have

$$\sharp\left(\sum_k B - \sum_l B\right) \leqslant K^{k+l}\sharp A,$$

where $\sum_k B \coloneqq \underbrace{B + B + \cdots + B}_{k \text{ times}}$.

Lemma 1.11 (Petridis)
If $\sharp(A+B) \leqslant K\sharp A$, then $\exists A_0 \subseteq A$, such that for every $C \subset E$ finite,

$$\sharp(C + A_0 + B) \leqslant K\sharp(C + A_0).$$

Proof. Let $K_0 \coloneqq \inf_{A' \subseteq A} \frac{\sharp(A'+B)}{\sharp A'} \leqslant K$ and $A_0 \subseteq A$ such that $K_0 = \frac{\sharp(A_0+B)}{\sharp A_0}$. Applying induction to $\sharp C$, consider $C' = C \cup \{c\}$, where $c \notin C$. WLOG, assume $c = 0$. Then

$$\sharp(C' + A_0 + B) = \sharp(C + A_0 + B) + \sharp(A_0 + B) - \sharp((C + A_0 + B) \cap (A_0 + B)).$$

Observe that $((C + A_0) \cap A_0) + B \subseteq (C + A_0 + B) \cap (A_0 + B)$. By assumption,

$$(C + A_0) \cap A_0 \subseteq A \implies \sharp((C + A_0) \cap A_0) + B \geqslant K_0\sharp((C + A_0) \cap A_0).$$

Hence by inductive assumption,

$$\sharp(C' + A_0 + B) \leqslant K_0(\sharp(C + A_0) + \sharp A_0 - \sharp((C + A_0) \cap A_0)) = K_0\sharp(C' + A_0).$$

□

Proof of Plünnecke-Ruzsa Inequality 1.10. Applying lemma, we have

$$\sharp(B + A_0) \leqslant K\sharp A_0, \quad \sharp(B + B + A_0) \leqslant K\sharp(B + A_0) \leqslant K^2\sharp A_0, \quad \cdots.$$

Hence, $\sharp\left(\sum_k B + A_0\right) \leqslant K^k\sharp A_0$. Finally, applying Ruzsa triangle inequality, we have

$$\sharp\left(\sum_k B - \sum_l B\right) \leqslant \frac{\sharp\left(\sum_k B + A_0\right)\sharp\left(\sum_l B + A_0\right)}{\sharp A_0} \leqslant K^{k+l}\sharp A_0 \leqslant K^{k+l}\sharp A.$$

□

Question 1.12. If E is not an abelian group, does the arguments still hold?

Answer Ruzsa triangle inequality, Ruzsa covering lemma, Petridis lemma still hold, but Plünnecke-Ruzsa inequality **fails**. See the following examples.

Example 1.13

G non abelian group. Take $A = H \cup \{a\}$, where H is a subgroup of G and $a \notin H$. Then $AA = H \cup aH \cup Ha \cup \{a\}$. Assume $\sharp H = N$, then $\sharp(AA) \leqslant 3N + 1 \leqslant \sharp A$. Consider $AAA \supseteq HaH$, if $aHa^{-1} \cap H = \{1\}$, then $\sharp(HaH) = N^2$. Explicitly, we can choose $G = S_{N+1}$, $H = \langle(123 \cdots N)\rangle$ and $a = (N\ (N+1))$. Hence for any $N > 0$, there exists A such that $\sharp(AA) \leqslant 3\sharp A$ but $\sharp(AAA) \geqslant N\sharp A$.

§2 Sum-product theorems

Let $(E, 0, 1, +, \cdot)$ be a ring, $A \subseteq E$ finite set, $K \geqslant 1$ parameter.
Let $E^\times = \{\text{invertible elements in } E\}$.

Definition 2.1. Let $R(A, K) := \{x \in E : \sharp(A + xA) \leqslant K\sharp A\}$.

The following lemma shows that $R(A, K)$ has an "almost" ring structure.

Lemma 2.2

1. If $x \in R(A, K) \cap E^\times$, then $x^{-1} \in R(A, K)$.
2. If $1, x, y \in R(A, K)$, then $x + y, x - y, xy \in R(A, K^{O(1)})$, where $O(1) = 8$ is enough.

Proof. 1. Trivial.

2. If $x, y \in R(A, K)$, by Ruzsa covering lemma, we have

$$xA \subseteq A - A + \mathbb{O}(K), \quad yA \subseteq A - A + \mathbb{O}(K).$$

then $A + (x + y)A \subseteq \sum_3 A - \sum_2 A + \mathbb{O}(K^2)$. Because $1 \in R(A, K)$, hence by P-R, we have $\sharp\left(\sum_3 A - \sum_2 A\right) \leqslant K^5\sharp A$. Then $\sharp(A + (x + y)A) \leqslant K^7\sharp A$. Similarly, we can prove $\sharp(A + xyA) \leqslant K^8\sharp A$.

□

Notation 2.3. For $s \in \mathbb{N}$, let $\sum_{\leqslant s} A = \bigcup\limits_{1\leqslant k\leqslant s} \sum_k A$, let $\prod_{\leqslant s} A = \bigcup\limits_{1\leqslant k\leqslant s} \prod_k A$. Let

$$\langle A\rangle_s = \sum_{\leqslant s}\prod_{\leqslant s} A - \sum_{\leqslant s}\prod_{\leqslant s} A.$$

Notation 2.4. $O_s(1)$ denotes a constant which just depend on s.

Lemma 2.5 (Ring Version of P-R)

Assume $\sharp(A + AA) \leqslant K\sharp A$, then $\sharp\langle A\rangle_s \leqslant K^{O_s(1)}\sharp A$.

Remark 2.6 — $\sharp(A+A) \leqslant K\sharp A$ and $\sharp(AA) \leqslant K\sharp A$ do not imply $\sharp(A+AA) \leqslant K^{O(1)}\sharp A$. For a counter example, we consider $A = \sqrt{-1}\mathbb{F}_p \subseteq \mathbb{F}_p[\sqrt{-1}]$ for some $p = 4k+3$ and $K = 1$, then $\sharp(A+AA) = p^2 = p\sharp A$.

Proof. By R-covering, we have $AA \subseteq A - A + \mathbb{O}(K)$. Let $X = \mathbb{O}(K)$, note that X could be chose in AA. Because $A \subseteq R(A,K)$ and $1 \in R(A,K^2)$ for $\sharp A \geqslant 2$, then $AA \subseteq R(A,K^{O(1)})$. Then

$$AAA \subseteq AA - AA + \bigcup_{x\in X} xA \subseteq \sum_2 A - \sum_2 A + \mathbb{O}(K^2) + \bigcup_{x\in X}(A - A + \mathbb{O}(K^{O(1)})),$$

hence $AAA \subseteq \sum_3 A - \sum_3 A + \mathbb{O}(K^{O(1)})$. By induction, we can prove the theorem. □

As the consequence of this lemma, we have $\langle A\rangle_s \subseteq R(A,K^{O_s(1)})$ if $A \subseteq R(A,K)$.

From now on, let E be a field, $A \subset E$ finite, $K \geqslant 1$.

Notation 2.7. Denote $f \ll g$ if there is an absolute constant $C > 0$ such that $f \leqslant Cg$.

Theorem 2.8 (Sum-Product Theorem in Fields)

Assume $\sharp(A+AA) \leqslant K\sharp A$, then

(1) either $\sharp A \ll K^{10000}$.

(2) or $\exists$ finite subfield F, such that $A \subseteq F$ and $\sharp F \ll K^{10000}\sharp A$.

Remark 2.9 — If $E = \mathbb{R}$, then for every $A \subseteq \mathbb{R}$, $\sharp(A+AA) \geqslant (\sharp A)^{1+\frac{1}{10000}}$.

Lemma 2.10

For any $x \in E$, if $\sharp(A+xA) < (\sharp A)^2$, then $x \in \frac{A-A}{(A-A)\setminus\{0\}}$.

Proof of Theorem 2.8. Let $F = \frac{A-A}{(A-A)\setminus\{0\}}$. Consider $K = (\sharp A)^{\frac{1}{10000}}$, the lemma shows that $R(A,K^{9999}) \subseteq F$. By assumption, $A \subseteq R(A,K)$, hence $A \subseteq R(A,K^2)$ by P-R if $\sharp A \geqslant 2$. By "almost" ring structure, we have $A - A \subseteq R(A,K^{20})$ and $((A-A)\setminus\{0\})^{-1} \subseteq R(A,K^{20})$, hence $F \subseteq R(A,K^{200})$. Furthermore, $F+F, FF \subseteq R(A,K^{2000}) \subseteq F$. Hence F is a finite field.

Now, we estimate $\sharp F$. There are two methods. One way is to consider a map

$$F \times (A \setminus \{0\}) \to (AA - AA) \times (AA - AA), \quad (x,a) \mapsto (au_x, bv_x),$$

where $u_x, v_x \in A - A$ are typical of writing $x = \frac{u_x}{v_x}$. The map is injective, hence $(\sharp F)(\sharp A - 1) \leqslant (\sharp(AA - AA))^2 \leqslant K^4(\sharp A)^2$ by P-R.

Another way is to use energy argument, see definition 3.1. Consider

$$(\sharp A)^4 = \sum_{x\in F} \sharp\left\{a,b,a',b' \in A : ax+b = a'x+b'\right\} \geqslant \sum_{x\in F}\frac{(\sharp A)^4}{\sharp(A+xA)} \geqslant \sharp F\frac{(\sharp A)^3}{K^{200}}.$$

Hence $\sharp F \leqslant K^{200}\sharp A$. □

Corollary 2.11

If $\sharp(AA) \leqslant K\sharp A, \sharp(A+A) \leqslant K\sharp A$, then

(1) either $\sharp A \ll K^{O(1)}$.

(2) or $\exists$ finite subfield F, $\exists a \in E$, such that $\sharp(A \cap aF) \gg \frac{\sharp A}{K^{O(1)}}$ and $\sharp F \ll K^{O(1)}\sharp A$.

Lemma 2.12 (Katz-Tao Lemma)

Assume $\sharp(A+A) \leqslant K\sharp A, \sharp(AA) \leqslant K\sharp A$. Then $\exists A' \subseteq A$ such that

$$\sharp A' \gg \frac{1}{K^{O(1)}}\sharp A \quad \text{and} \quad \sharp(A'A' - A'A') \ll K^{O(1)}\sharp A'.$$

Proof of Corollary 2.11 assuming Lemma 2.12. Take such A' in lemma, we choose $a \in A' \setminus \{0\}$, let $B = a^{-1}A'$. Then $1 \in B$ and $B - BB \subseteq BB - BB$, hence $\sharp(B - BB) \leqslant K^{O(1)}\sharp B$. Then $\sharp(B + BB) \leqslant K^{O(1)}\sharp B$ by P-R and R-covering. Applying Theorem 2.8 to B, the corollary follows. ☐

Notation 2.13. Denote $f \lesssim g$ if $f \ll K^{O(1)}g$, denote $f \sim g$ if $f \lesssim g$ and $g \lesssim f$.

Proof of Katz-Tao Lemma 2.12. Consider the function $\varphi = \sum_{a\in A} \mathbb{1}_{aA}$ defined on AA. Endowing AA with counting measure, then

$$(\sharp A)^4 = \|\varphi\|_1^2 \leqslant \|\varphi\|_2^2\,\|1\|_2^2 = \sharp(AA)\left\|\sum_{a,b\in A} \mathbb{1}_{aA\cap bA}\right\|_1 \leqslant K\sharp A \sum_{a,b\in A} \sharp(aA \cap bA).$$

Therefore, $\exists b \in A$ such that $\frac{1}{\sharp A}\sum_{a\in A}\sharp(aA\cap bA) \geqslant \frac{\sharp A}{K}$. Consider

$$A' := \left\{a \in A : \sharp(aA \cap bA) \geqslant \frac{\sharp A}{2K}\right\},$$

then $\sharp A' \geqslant \frac{\sharp A}{2K}$. Hence for every $a \in A'$, by R-triangle,

$$\sharp(aA + bA) \leqslant \frac{\sharp(aA + aA\cap bA)\sharp(bA - aA\cap bA)}{\sharp(aA\cap bA)} \lesssim \frac{\sharp(A+A)\sharp(A-A)}{\sharp A} \lesssim \sharp A.$$

By R-covering, $aA \subseteq bA - bA + \mathbb{O}(K^{O(1)})$. Then for every $a_1, a_2, a_3, a_4 \in A$,

$$(a_1a_2 - a_3a_4)A \subseteq b^2\left(\sum_4 A - \sum_4 A\right) + \mathbb{O}(K^{O(1)}).$$

Let $d = a_1a_2 - a_3a_4$, then $dA \subseteq \bigcup_{x\in X}\left(b^2\left(\sum_4 A - \sum_4 A\right) + x\right)$ where $\sharp X \lesssim 1$. Then $\exists x$ such that $\sharp\left(dA \cap \left(b^2\left(\sum_4 A - \sum_4 A\right) + x\right)\right) \gtrsim \sharp A$. Hence

$$\sharp\left\{u \in A - A : du \in b^2\left(\sum_8 A - \sum_8 A\right)\right\} \gtrsim \sharp A.$$

Consider $F = b^2\frac{\sum_8 A - \sum_8 B}{(A-A)\setminus\{0\}}$, then $\sharp F \leqslant \sharp(A-A)\sharp(\sum_8 A - \sum_8 A) \lesssim (\sharp A)^2$. On the other hand, $\sharp F \gtrsim \sharp A \sharp(A'A' - A'A')$ by the former deduction. Hence $\sharp(A'A' - A'A') \lesssim \sharp A$. ☐

§3 More additive combinatorics

$(E,+)$ abelian group.

Definition 3.1. For $A, B \subseteq (E,+)$, define the **additive energy** between A, B

$$\mathscr{E}_+(A,B) := \sharp\left\{(a,b,a',b') \in A \times B \times A \times B : a+b = a'+b'\right\}.$$

The trivial bound of energy is

$$\sharp A \sharp B \leqslant \mathscr{E}_+(A,B) \leqslant (\sharp A)^{\frac{3}{2}} (\sharp B)^{\frac{3}{2}}.$$

Let $r = \mathbb{1}_A * \mathbb{1}_B$, then $r(y) = \sharp\left\{(a,b) \in A \times B : a+b = y\right\}$. Endowing E with the counting measure, then

$$\mathscr{E}_+(A,B) = \sum_{y \in A+B} r(y)^2 = \|\mathbb{1}_A * \mathbb{1}_B\|_2^2.$$

Note that $\|\mathbb{1}_A * \mathbb{1}_B\|_1 = \|\mathbb{1}_A\|_1 \|\mathbb{1}_B\|_1 = \sharp A \sharp B$. By Cauchy-Schwarz,

$$\mathscr{E}_+(A,B) = \|\mathbb{1}_A * \mathbb{1}_B\|_2^2 \geqslant \frac{\|\mathbb{1}_A * \mathbb{1}_B\|_1^2}{\sharp \operatorname{supp} \mathbb{1}_A * \mathbb{1}_B} = \frac{(\sharp A)^2 (\sharp B)^2}{\sharp (A+B)}.$$

This inequality shows that if A and B have a small sum set, then the additive energy between A, B is big.

Remark 3.2 — The converse is **not** true. See the following example.

Example 3.3

Let $A = \{0,1,2,\cdots,N-1\} \cup \left\{N, 2N, \cdots, N^2\right\}$, then $\sharp A = 2N$. We have $\sharp(A+A) \asymp N^2$ and $\mathscr{E}_+(A,A) \geqslant \mathscr{E}_+(\{0,\cdots,N-1\},\{0,\cdots,N-1\}) \geqslant \frac{N^2}{2N} \gg N^3$. They both attain the trivial upper bound up to a constant.

Theorem 3.4 (Balog-Szemerédi-Gowers)

The following are equivalent, the parameter $K_i > 0$ differs from each other by at most a polynomial dependence:

(i) $\mathscr{E}_+(A,B) \geqslant \frac{1}{K_1} (\sharp A)^{\frac{3}{2}} (\sharp B)^{\frac{3}{2}}$.

(ii) $\exists A' \subseteq A, B' \subseteq B$ with $\sharp A' \geqslant \frac{\sharp A}{K_2}, \sharp B' \geqslant \frac{\sharp B}{K_2}$, such that $\sharp(A'+B') \leqslant K_2 (\sharp A)^{\frac{1}{2}} (\sharp B)^{\frac{1}{2}}$.

(iii) $\exists G \subseteq A \times B$ with $\sharp G \geqslant \frac{1}{K_3} \sharp A \sharp B$ such that $\sharp(A \overset{G}{+} B) \leqslant K_3 (\sharp A)^{\frac{1}{2}} (\sharp B)^{\frac{1}{2}}$, where $A \overset{G}{+} B := \{a+b : (a,b) \in G\}$.

Proof. (ii) $\Longrightarrow$ (i): Trivial.

(i) $\Longrightarrow$ (iii): Let $Y = \left\{y : r(y) \geqslant \frac{(\sharp A)^{\frac{1}{2}} (\sharp B)^{\frac{1}{2}}}{2K_1}\right\}$, $G = \{(a,b) \in A \times B : a+b \in Y\}$, then $A \overset{G}{+} B = Y$. The bound of energy $\mathscr{E}_+(A,B) \geqslant \frac{1}{K_1} (\sharp A)^{\frac{3}{2}} (\sharp B)^{\frac{3}{2}}$ immediately gives that $\sharp G \geqslant \frac{1}{2K_1} \sharp A \sharp B$. Besides,

$$\sharp Y \frac{\sharp A \sharp B}{4K_1^2} \leqslant \sum_{y \in Y} r(y)^2 \leqslant (\sharp A)^{\frac{3}{2}} (\sharp B)^{\frac{3}{2}},$$

hence $\sharp Y \ll K_1^2(\sharp A)^{\frac{1}{2}}(\sharp B)^{\frac{1}{2}}$.

For proving (iii) $\Longrightarrow$ (ii), we need some more preparations. □

Theorem 3.5 (Multiplicative Balog-Szemerédi-Gowers)
For every group $(H,\cdot)$, $A, B \subseteq H$ finite sets. The following are equivalent, the parameter $K_i > 0$ differs from each other by at most a polynomial dependence:

(i) $\mathscr{E}_+(A,B) \geqslant \frac{1}{K_1}(\sharp A)^{\frac{3}{2}}(\sharp B)^{\frac{3}{2}}$.

(ii) $\exists A' \subseteq A, B' \subseteq B$ with $\sharp A' \geqslant \frac{\sharp A}{K_2}, \sharp B' \geqslant \frac{\sharp B}{K_2}$, such that $\sharp(A'B') \leqslant K_2(\sharp A)^{\frac{1}{2}}(\sharp B)^{\frac{1}{2}}$.

(iii) $\exists G \subseteq A \times B$ with $\sharp G \geqslant \frac{1}{K_3}\sharp A \sharp B$ such that $\sharp(A \overset{G}{\cdot} B) \leqslant K_3(\sharp A)^{\frac{1}{2}}(\sharp B)^{\frac{1}{2}}$, where $A \overset{G}{\cdot} B \coloneqq \{ab : (a,b) \in G\}$.

Theorem 3.6 (Graph-Theoretic B-S-G)
Let A, B be finite sets, $G \subseteq A \times B$. Assume $\sharp G \geqslant \frac{1}{K}\sharp A \sharp B$. Then exists $A' \subseteq A, B' \subseteq B'$, $\sharp A' \gtrsim \sharp A, \sharp B' \gtrsim \sharp B$. And for every $a' \in A', b' \in B'$,

$$\sharp\left\{(a,b) \in A \times B : (a',b),(a,b),(a,b') \in G\right\} \gtrsim \sharp A \sharp B.$$

Proof of BSG assuming graph BSG. Let A', B' be given by graph B-S-G, for every $x \in A' \cdot B'$,

$$r_3(x) = \sharp\left\{(y_1,y_2,y_3) \in (A \overset{G}{\cdot} B)^3 : x = y_1 y_2^{-1} y_3\right\} \gtrsim \sharp A \sharp B.$$

Then

$$\sharp(A' \cdot B') \leqslant \frac{\sharp(A \overset{G}{\cdot} B)^3}{\sharp A \sharp B} \lesssim (\sharp A)^{\frac{1}{2}}(\sharp B)^{\frac{1}{2}}.$$

□

Notation 3.7. For $a \in A$, let $B(a) \coloneqq \{b \in B : (a,b) \in G\}$.

Proof of graph BSG. Let $A_1 \coloneqq \sharp\left\{a \in A : \sharp B(a) \geqslant \frac{\sharp B}{2K}\right\}$, then $\sharp A \geqslant \frac{\sharp A}{2K}$. Then

$$\sum_{a,a' \in A_1} \sharp B(a) \cap B(a') = \sum_{b \in B}\left(\sum_{a \in A_1} \mathbb{1}_{B(a)}(b)\right)^2 \geqslant \frac{\left(\sum_{a \in A_1} \sharp B(a)\right)^2}{\sharp B} \geqslant \frac{1}{4K^2}(\sharp A)^2 \sharp B.$$

Set $\varepsilon = \frac{1}{32K}$, let

$$U = \left\{(a,a') \in A_1 \times A_1 : \sharp B(a) \cap B(a') \leqslant \frac{\varepsilon}{4K^2}\sharp B\right\}.$$

Idea: we want $A' \subseteq A, B' \subseteq B$ such that:

(i) $\sharp A' \gtrsim \sharp A, \sharp B' \geqslant \sharp B$,

(ii) $\forall a \in A', \sharp A_1^U(a) \coloneqq \sharp\{a' \in A_1 : (a,a') \in U\} \leqslant \frac{\sharp A_1}{8K}$.

(iii) $\forall b \in B', \sharp A_1(b) \geqslant \frac{\sharp A_1}{4K}$.

This is enough, but condition (ii) is too much. Instead, we want $A' \subseteq A_2 \subseteq A_1, B' \subseteq B$ such that

(i) $\sharp A' \gtrsim \sharp A, \sharp B' \geqslant \sharp B$,

(ii) $\forall a \in A', \sharp A_2^U(a) \leqslant \frac{\sharp A_2}{8K}$.

(iii) $\forall b \in B', \sharp A_2(b) \geqslant \frac{\sharp A_2}{4K}$.

Candidate $A_2 = A_1(b)$ for some $b \in B$. Notice that

$$\sum_{b\in B} \sharp(A_1(b) \times A_1(b)) = \sum_{a,a'\in A_1} \sharp(B(a) \cap B(a')) \geqslant \frac{(\sharp A_1)^2 \sharp B}{4K^2},$$

$$\sum_{b\in B} \sharp(A_1(b) \times A_1(b) \cap U) = \sum_{(a,a')\in U} \sharp(B(a) \cap B(a')) \leqslant \frac{\varepsilon(\sharp A_1)^2 \sharp B}{4K^2}.$$

Hence $\exists b \in B$, write $A_2 = A_1(b)$ such that

$$\sharp(A_2 \times A_2) - \frac{1}{2\varepsilon}\sharp(A_2 \times A_2 \cap U) \geqslant \frac{(\sharp A_1)^2}{8K^2}.$$

Then $\sharp A_2 \geqslant \frac{\sharp A_1}{2\sqrt{2K}}$ and $\sharp(U \cap (A_2 \times A_2)) \leqslant 2\varepsilon(\sharp A_2)^2$. Let $A' = \left\{a \in A' : \sharp A_2^U(a) \leqslant \frac{\sharp A_2}{8K}\right\}$, by

$$\sum_{a\in A_2} \sharp A_2^U(a) = \sharp(U \cap (A_2 \times A_2)) \leqslant \frac{(\sharp A_2)^2}{16K},$$

it shows $\sharp A' \gtrsim \sharp A$. Let $B' = \left\{b \in B', \sharp A_2(b) \geqslant \frac{\sharp A_2}{4K}\right\}$, then

$$\sum_{b\in B} \sharp A_2(b) = \sum_{a\in A_2 \subseteq A_1} \sharp B(a) \geqslant \frac{\sharp A_2 \sharp A}{2K},$$

hence $\sharp B' \geqslant \frac{\sharp B}{4K}$. $\square$

§4 A product theorem

Let $(G, \cdot)$ be a group, $A \subseteq G$ finite subset.

Notation 4.1. Let $\prod_k A = \underbrace{AA\cdots A}_{k \text{ times}}$, $A^{-1} = \left\{a^{-1} : a \in A\right\}$.

Lemma 4.2 1. If $\sharp(AAA) \leqslant K\sharp A$, then $\sharp \prod_3 (A \cup \{1\} \cup A^{-1}) \ll K^3 \sharp A$.

2. If $\sharp \prod_3 (A \cup \{1\} \cup A^{-1}) \leqslant K\sharp A$, then for every $k \geqslant 3$,

$$\sharp \prod_k (A \cup \{1\} \cup A^{-1}) \leqslant K^{k-2} \sharp A.$$

Proof.

1. By Ruzsa-triangle,

$$\sharp(AAA^{-1}) \leqslant \frac{\sharp(AAA)\sharp(A^{-1}A^{-1})}{\sharp A^{-1}} \leqslant K^2\sharp A,$$

$$\sharp(AA^{-1}A) \leqslant \frac{\sharp(AA^{-1}A^{-1})\sharp(AA)}{\sharp A} \leqslant K^3\sharp A,$$

The result follow.

2. Assume $1 \in A = A^{-1}$, the statement follows by Ruzsa-triangle.

□

Definition 4.3. Finite set $A \subseteq G$ is called a K-approximate subgroup, if

(i) $1 \in A, A^{-1} = A$,

(ii) $\exists X \subseteq G, \sharp X \leqslant K$, such that $AA \subseteq XA$.

Lemma 4.4 (Reformulation of lemma 4.2)
If $\sharp(AAA) \leqslant \sharp A$, then $B = \prod_2(A \cup \{1\} \cup A^{-1})$ is a $O(K^{O(1)})$-approximate subgroup.

Problem 4A. Does $\sharp(AAA) \leqslant K\sharp(AA)$ implies $\sharp\prod_k A \leqslant K^{O_k(1)}\sharp A$.

Theorem 4.5 (Helfgott)
$\forall\delta > 0, \exists\varepsilon > 0$, let $G = \mathrm{SL}(2, \mathbb{F}_p), p$ is a prime number. Let $A \subseteq G, \langle A\rangle = G$, then either

(1) $\sharp(AAA) \geqslant c(\sharp A)^{1+\varepsilon}$,

(2) or $\sharp A \geqslant p^{3-\delta}$.

Theorem 4.6 (Equivalent formulation of Helfgott's Theorem)
If $A \subseteq G = \mathrm{SL}(2, \mathbb{F}_p)$ is a K-approximate subgroup, then either

(i) $\langle A\rangle \neq G$.

(ii) or $\sharp A \lesssim 1$.

(iii) or $\sharp A \gtrsim \sharp G$.

Exercise 4.7. Prove two statements above are equivalent.

Remark 4.8 — $\mathrm{PSL}(2, \mathbb{F}_p)$ is a simple group for $p > 5$.

Remark 4.9 — Such result does not hold for abelian group.

Lemma 4.10 (Orbit-Stabalizer Formula)
$A \curvearrowright X$, then for every $x \in X$, we have

$$\sharp A \leqslant \sharp(A.x)\sharp(\mathrm{Stab}(x) \cap A^{-1}A).$$

Remark 4.11 — If A is a subgroup, then identity holds.

Definition 4.12. $T \subseteq \mathrm{SL}(2, \overline{\mathbb{F}}_p)$ is called a torus if $T = g\begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}g^{-1}$ for some $g \in \mathrm{SL}(2, \overline{\mathbb{F}}_p)$.

Lemma 4.13
Assume A is K-approximate subgroup, $\exists T \subseteq \mathrm{SL}(2, \overline{\mathbb{F}}_p)$ a torus such that

$$\sharp(T \cap AA) \gtrsim \sharp\,\mathrm{tr}(A) - 2,$$

where $\mathrm{tr}(A) = \{\mathrm{tr}(a) : a \in A\}$.

Proof. Consider $B \subseteq A$ with $\sharp B = \sharp\,\mathrm{tr}(A) - 2, \pm 2 \notin \mathrm{tr}(B)$ and $\mathrm{tr}(b), b \in B$ are pairwise distinct. Consider the conjugation, we have

$$\sharp B \sharp A = \sum_{b \in B} \sharp\left\{aba^{-1} : a \in A\right\} \sharp(C_G(b) \cap AA) \leqslant \sharp(AAA) \max_{b \in B} \sharp(C_G(b) \cap AA),$$

hence there are some $b \in B$ such that $\sharp(C_G(b) \cap AA) \geqslant \frac{\sharp B}{K}$. □

Definition 4.14. An affine variety over $\overline{\mathbb{F}}_p$ of complexity $\leqslant M$ is $V \subseteq \overline{\mathbb{F}}_p^n$,

$$V = \left\{\underline{x} \in \overline{\mathbb{F}}_p^n : f_1(\underline{x}) = \cdots = f_s(\underline{x}) = 0\right\},$$

where $f_1, \cdots, f_s \in \overline{\mathbb{F}}_p[x_1, x_2, \cdots, x_n]$ and $s, n, \deg f_1, \cdots, \deg f_s \leqslant M$.

Proposition 4.15 (Escape from Subvarieties)
$\forall M > 0, \exists p_0 = p_0(M)$, such that for every $p > p_0$ prime, $G = \mathrm{SL}(2, \overline{\mathbb{F}}_p)$, $V \subsetneq G$ a proper subvariety of complexity $\leqslant M$. $A \subseteq \mathrm{SL}(2, \mathbb{F}_p)$, assume $\langle A \rangle = \mathrm{SL}(2, \mathbb{F}_p)$, then $\exists g \in \prod_m(\{1\} \cup A)$, such that $g \notin V$, where m depends only on M.

Remark 4.16 — $\mathrm{SL}(2, \mathbb{F}_p)$ is not Zariski dense in G, i.e., $\exists$ proper subvariety V such that $\mathrm{SL}(2, \mathbb{F}_p) \subseteq V$, hence we need an additional condition on complexity.

Definition 4.17. An affine subvariety V is **irreducible** if V can not be written as $V = V_1 \cup V_2$ where V_1, V_2 are both subvarieties and $V_1, V_2 \neq V$.

Definition 4.18. **Krull dimension** of a subvariety V is defined as

$$\dim V := \max \{k : \exists V_1 \subsetneqq V_2 \subsetneqq \cdots \subsetneqq V_k \subseteq V, V_1, \cdots, V_k \text{ irreducible}\}.$$

Proof. $G = \left\{(x_{11}, x_{12}, x_{21}, x_{22}) \in \overline{\mathbb{F}}_p^4 : x_{11}x_{22} - x_{12}x_{21} = 1\right\}$ is of complexity 4. Let

$$\overline{\mathbb{F}}_p[G] := \overline{\mathbb{F}}_p[x_{11}, \cdots, x_{22}]/(\det \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} - 1).$$

For every $V \subseteq G$ subvariety, with complexity $\leqslant M$, let

$$I_V := \left\{f \in \overline{\mathbb{F}}_p[G] : \forall x \in V, f(x) = 0\right\},$$

which is an ideal. There exists $d = d(M)$ such that $I = I_V \cap \overline{\mathbb{F}}_p[G]_{\deg \leqslant d} = I_V$. Consider $G \curvearrowright \overline{\mathbb{F}}_p[G]$ given by $(g.f)(\cdot) = f(g^{-1}\cdot)$. Hence $G \curvearrowright \overline{\mathbb{F}}_p[G]_{\deg\leqslant d}$, let $m = \dim \overline{\mathbb{F}}_p[G]_{\deg \leqslant d}$. Assume for a contradiction, $\prod_m (A \cup \{1\}) \subseteq V$. Then there exists $g_1, \cdots, g_s \in \prod_m (A \cup \{1\})$ such that

$$J = I + g_1^{-1} I + \cdots + g_s^{-1} I$$

is $\langle A \rangle$-invariant. Let $H = \{g \in G : g.I = I\}$, then

1. H is a subgroup, $A \subseteq H$.
2. $H \subseteq V$. Indeed, $\forall h \in H, f \in I$, $h^{-1}.f \in J$. Hence $\exists f_0, f_1, \cdots, f_s \in I$, such that

   $$h^{-1} f = f_0 + g_1^{-1} f_1 + \cdots + g_s^{-1} f_s.$$

   Take $x = 1_G$, we have $h \in V$.
3. Complexity of H is $O_M(1)$.

By a Schwarz-Zippel (Lang-Weil) theorem, we have

$$\sharp(H \cap \mathrm{SL}_2(\mathbb{F}_p)) \ll_M p^{\dim H} \ll_M p^{\dim V}.$$

But $\sharp\langle A \rangle \asymp p^3$, if V is proper, then $\dim V < \dim G = 3$. A contradiction. $\square$

Proof of Theorem 4.6. We separate the proof into following four steps.

I. $\exists T \subseteq G$ torus such that $\sharp(T \cap AA) \gtrsim \sharp\,\mathrm{tr}(A) - 2$.

II. There exists some integers of $O(1)$ such that $\sharp\,\mathrm{tr}(\prod_{O(1)} A) \gg (\sharp A)^{\frac{1}{3}}$.

III. T torus, finite $V \subseteq T$, then $\exists g \in \prod_{O(1)} A$ such that one of the following holds:
   (1) $\sharp VVV \geqslant K' \sharp V$.
   (2) $\sharp\,\mathrm{tr}(\prod_{20} V g \prod_{20} V g^{-1}) \geqslant K' \sharp V$.
   (3) $\sharp V \lesssim 1$.
   (4) $\sharp V \gtrsim p$.

IV. T torus, finite $V \subseteq T$, then $\exists g \in \prod_{O(1)} A$ such that $\sharp(VgVg^{-1}V) \gg (\sharp V)^3$.

After those four steps, we can prove the theorem. Applying II, we have $\sharp\,\mathrm{tr} \prod_{O(1)} A \gg (\sharp A)^{\frac{1}{3}}$. By I, there is T torus, let $V = T \cap \prod_{O(1)} A$, such that $\sharp V \gtrsim (\sharp A)^{\frac{1}{3}}$. For every $g \in \prod_{O(1)} A$, we have $\sharp\,\mathrm{tr}(\prod_{O(1)} A) \geqslant \sharp\,\mathrm{tr}(\prod_{20} V g \prod_{20} V g^{-1})$. By I, there is some $V' = T' \cap \prod_{O(1)} A$ such that

$$\sharp V' \gtrsim \max\left\{\sharp\,\mathrm{tr}(\prod\nolimits_{20} V g \prod\nolimits_{20} V g^{-1}), \sharp VVV\right\}.$$

By IV, there exists $h \in \prod_{O(1)} A$, such that

$$\sharp A \gtrsim \sharp \prod_{O(1)} A \gg \sharp(V'hV'h^{-1}V') \gg (\sharp V')^3.$$

Hence, $\max\left\{\sharp\operatorname{tr}(\prod_{20} Vg \prod_{20} Vg^{-1}), \sharp VVV\right\} \lesssim (\sharp A)^{\frac{1}{3}}$. By III, take a suitable $K' = O(K^{O(1)})$, then there exists $g \in \prod_{O(1)} A$ such that $\sharp V \lesssim 1$ or $\sharp V \gtrsim p$. Which shows that $\sharp A \lesssim 1$ or $\sharp A \gtrsim p^3$. $\square$

Proof of II. For every $g, h \in G$, consider

$$\Phi_{g,h} : G \to (\overline{F}_p)^3, \quad x \mapsto (\operatorname{tr}(x), \operatorname{tr}(gx), \operatorname{tr}(hx)).$$

Then

$$\begin{aligned} &\{(g,h) \in G \times G : \Phi_{g,h} \text{ has fiber of positive dimension}\} \\ = &\{(g,h) \in G \times G : \Phi_{g,h} \text{ has fiber of } \sharp > 2\} \end{aligned}$$

is a proper subvariety of $G \times G$ of complexity $O(1)$. By "escape"(4.15), there exists $g, h \in \prod_{O(1)}(A \cup \{1\})$ such that each fiber of $\Phi_{g,h}$ has $\sharp \leqslant 2$, hence $\sharp A \ll (\sharp\operatorname{tr}(\prod_{O(1)} A))^3$. $\square$

Proof of IV. For every $g \in G$, consider

$$\phi_g : T^3 \to G, \quad (x, y, z) \mapsto xgyg^{-1}z.$$

Then

$$\{g \in G : \phi_g \text{ has fiber of positive dimension}\}$$

is a proper subvariety of G of complexity $O(1)$. By "escape"(4.15), there exists $g \in \prod_{O(1)}(A \cup \{1\})$ such that each fiber of ϕ_g is of 0-dimensional. Because the complexity is of $O(1)$, hence each fiber of ϕ_g is of $\sharp \leqslant O(1)$. Therefore, $\sharp\phi_g(V^3) \gg (\sharp V)^3$. $\square$

Proof of III. Assume $V \subseteq T = \left\{\begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}\right\}, g = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then

$$\operatorname{tr}\left(\begin{bmatrix} x & 0 \\ 0 & x^{-1} \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} y & 0 \\ 0 & y^{-1} \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}\right) = ad \cdot w(xy) - bc \cdot w(xy^{-1}),$$

where $w(x) = x + x^{-1}$. Then the statement is equivalent to the following proposition. $\square$

Proposition 4.19
$\widehat{V} \subseteq \overline{\mathbb{F}}_p^{\times}, a_1, a_2 \in \overline{\mathbb{F}}_p^{\times}$, assume $\widehat{V}$ is K-approximate subgroup of $\overline{\mathbb{F}}_p$ and

$$\left\{a_1 w(xy) + a_2 w(xy^{-1}) : x, y \in \prod_{20} \widehat{V}\right\} \leqslant K\sharp\widehat{V},$$

then either $\sharp\widehat{V} \lesssim 1$ or $\sharp\widehat{V} \gtrsim p$.

Proof. We just prove a special case of $a_1 = a_2 = 1$. Let $E = \left\{(w(xy), w(xy^{-1})) : x, y \in \widehat{V}\right\}$, by assumption, $\sharp(w(\prod_2 \widehat{V}) \overset{E}{+} w(\prod_2 \widehat{V})) \lesssim \sharp\widehat{V}$. At the same time, $\sharp E \gg (\sharp\widehat{V})^2$, hence by B-S-G(3.4) and P-R, there exists $V' \subseteq \prod_2 \widehat{V}, \sharp V' \gtrsim \sharp\widehat{V}$ such that

$$\sharp(w(V') + w(V')) \lesssim \sharp\widehat{V}.$$

Notice that $w(x)w(y) = w(xy) + w(xy^{-1})$, then $w(V')w(V') \leqslant K\sharp\widehat{V}$. By sum-product, either $\sharp w(V') \lesssim 1$ or $\sharp w(V') \gtrsim p$. $\square$

Exercise 4.20. Prove the general cases.

Remark 4.21 — Another view of this proposition is given by Eleke-Ronyai problem. Which shows that there exists $\varepsilon > 0$, such that for every $f \in \mathbb{R}[x,y]$ or $f \in \mathbb{R}(x,y)$, then

(1) either $\forall A \subseteq \mathbb{R}$ finite, $\sharp A = N$, we have $\sharp f(A \times A) \gg N^{1+\varepsilon}$,

(2) or $\exists g, h, \phi : \mathbb{R} \to \mathbb{R}$ analytic such that $f(x,y) = \phi(g(x) + h(y))$.

§5 Expansion in $\mathrm{SL}(2,\mathbb{F}_p)$

Let $S \subseteq \mathrm{SL}(2,\mathbb{Z})$ be a finite subset, $S = S^{-1}$. Let $G_p = \mathrm{SL}(2,\mathbb{F}_p) = \mathrm{SL}(2,\mathbb{Z})/\ker \pi_p$, where

$$\pi_p : \mathrm{SL}(2,\mathbb{Z}) \to \mathrm{SL}(2,\mathbb{F}_p)$$

is the projection by $\mod p$. Let $\Gamma = \mathrm{SL}(2,\mathbb{Z})$, then there is a natural action $\Gamma \curvearrowright G_p$. Consider **Koopman representation** $\Gamma \curvearrowright L^2(G_p)$ given by

$$\gamma \mapsto T_p(\gamma) \in U(L^2(G_p)), \quad T_p(\gamma) f(\,\cdot\,) = f(\gamma^{-1}\,\cdot\,).$$

Let $\chi_S = \frac{1}{\sharp S}\mathbb{1}_S$, define

$$T_p(\chi_S) f(\,\cdot\,) = \frac{1}{\sharp S}\sum_{\gamma\in S} f(\gamma^{-1}\,\cdot\,) = \chi_S * f,$$

then $T_p(\chi_S) \in \mathrm{End}(L^2(G_p))$.

Remark 5.1 — If $S = S^{-1}$, then $T_p(\chi_S)$ is self-adjoint.

Consider the spectrum of $T_p(\chi_S)$. Note that $\|T_p(\chi_S)\| \leqslant 1$ and $1 \in \mathrm{Spec}(T_p(\chi_S))$. Let

$$L_0^2(G_p) := \mathbb{1}_G^{\perp} = \left\{ f \in L^2(G_p) : \int f = 0 \right\},$$

then $T_{p,0}(\chi_S) : L_0^2(G_p) \to L_0^2(G_p)$.

Theorem 5.2 (Uniform Expansion in $\mathrm{SL}(2,\mathbb{F}_p)$, Bourgain-Gamburd)
Assume $\langle S \rangle \subseteq \mathrm{SL}(2,\mathbb{Z})$ is not virtually solvable, then $\{T_{p,0}(\chi_S)\}_p$ has a **uniform spectral gap**, i.e., there exists $c > 0$, such that for every p prime,

$$\mathrm{Spec}(T_{p,0}(\chi_S)) \cap [1-c, 1] = \varnothing.$$

Exercise 5.3. Prove that the conclusion is equivalent to $\exists \varepsilon > 0$, such that $\forall p$ prime, for every $f \in L_0^2(G_p)$, there exists $s \in S$,

$$\|f - T_p(s) f\| \geqslant \varepsilon \|f\|.$$

(We say $\bigoplus_p L_0^2(G_p)$ has no almost invariant vector).

Remark 5.4 — As a consequence of the exercise, let $S' \subseteq \langle S \rangle$ be a finite symmetric set, if $\{T_p(\chi_{S'})\}_p$ has a uniform spectral gap, then $\{T_p(\chi_S)\}_p$ has a uniform spectral gap.

Proposition 5.5 (Tits Alternative for $\mathrm{SL}(2, \mathbb{Z})$)
$\Gamma' \subseteq \mathrm{SL}(2, \mathbb{Z})$ subgroup, then

(1) either Γ' contains non-abelian free subgroup,

(2) or Γ' is virtually solvable.

Proof. Consider $\Gamma(3) = \ker \pi_3 = \{g \in \mathrm{SL}(2, \mathbb{Z}) : g \equiv 1 \bmod 3\}$, then $[\Gamma : \Gamma(3)] < \infty$. Note that $\Gamma(3) = \pi_1(\mathbb{H}/\Gamma(3))$ which is a free group. By Nielson-Schreien's argument, $\Gamma' \cap \Gamma(3) \subseteq \Gamma(3)$ is of finite index and hence is also a free group. Then, $\Gamma' \cap \Gamma(3) = 1, \mathbb{Z}$, or a non-abelian free group. $\square$

Remark 5.6 — Finite index subgroup of finite generated group is also finite generated.

Remark 5.7 — This proposition allows us to reduce the statement of Theorem 5.2 to the case that S freely generates a non-abelian free group.

Theorem 5.8 (B-S-G weighted version)
Let μ, ν be two probability measures on G, $K \geqslant 2$, if

$$\|\mu * \nu\| \geqslant K^{-1} \|\mu\|^{\frac{1}{2}} \|\nu\|^{\frac{1}{2}},$$

then there exists an $O(K^{O(1)})$-approximate subgroup H, $a, b \in G$, such that

$$\sharp H \sim \|\mu\|^{-2} \sim \|\nu\|^{-2}, \quad \mu(aH) \gtrsim 1, \nu(aH) \gtrsim 1.$$

Remark 5.9 — If $\mu = \frac{1}{\sharp A} \mathbb{1}_A$, then $\|\mu\|^2 = \frac{1}{\sharp A}$. This shows that the exponent -2 is reasonable.

Remark 5.10 — $\|\mu\|^2 \leqslant \|\mu\|_\infty \|\mu\|_1 \leqslant 1$, and $\|\mu\| = 1$ iff μ is Dirac. $\|\mu\|^2 \geqslant \frac{1}{\sharp G}$, the equality holds iff $\mu = \chi_G$.

Remark 5.11 — $\|\mu * \nu\| \leqslant \|\mu\|_1 \|\nu\| = \|\nu\|$, hence if $\|\mu\|^{\frac{1}{2}} \|\nu\|^{\frac{1}{2}} \lesssim \|\mu * \nu\|$, then $\|\mu\| \lesssim \|\nu\|$. Therefore, $\|\mu\| \sim \|\nu\|$.

Proof. Let $m = \frac{1}{16K^4}$, $M = 4K^4$, let

$$A_0 = \left\{x \in G : m \|\mu\|^2 \leqslant \mu(x) \leqslant M \|\mu\|^2\right\},$$

$$A_- = \left\{x \in G : \mu(x) < m \|\mu\|^2\right\}, \quad A_+ = \left\{x \in G : \mu(x) > M \|\mu\|^2\right\}.$$

Consider $\mu_0 = \mu \mathbb{1}_{A_0}, \mu_- = \mu \mathbb{1}_{A_-}, \mu_+ = \mu \mathbb{1}_{A_+}$, then $\mu = \mu_0 + \mu_- + \mu_+$. Similarly, write $\nu = \nu_0 + \nu_- + \nu_+$. We have

$$\|\mu_- * \nu\| \leqslant \|\mu_-\| \leqslant m \|\mu\| \leqslant mK \|\mu\|^{\frac{1}{2}} \|\nu\|^{\frac{1}{2}},$$

$$\|\mu_+ * \nu\| \leqslant \|\mu_+\|_1 \|\nu\| \leqslant \frac{1}{M} \|\nu\| = \frac{K}{M} \|\mu\|^{\frac{1}{2}} \|\nu\|^{\frac{1}{2}}.$$

Hence

$$\|\mu_0 * \nu_0\| \geqslant \frac{1}{2K} \|\mu\|^{\frac{1}{2}} \|\nu\|^{\frac{1}{2}}.$$

On the other hand,

$$\mu_0 * \nu_0 \sim \|\mu\|^2 \|\nu\|^2 \mathbb{1}_{A_0} * \mathbb{1}_{B_0}, \quad \text{pointwise}.$$

Notice that $\sharp A_0 \sim \|\mu\|^{-2}$, recall the additive energy, it shows that

$$\mathscr{E}_+(A_0, B_0) = \|\mathbb{1}_{A_0} * \mathbb{1}_{B_0}\|^2 \gtrsim \|\mu\|^{-3} \|\nu\|^{-3} \gtrsim (\sharp A_0)^{\frac{3}{2}} (\sharp B_0)^{\frac{3}{2}}.$$

By B-S-G, $\exists A \subseteq A_0, B \subseteq B_0$, $\sharp A \gtrsim \sharp A_0, \sharp B \gtrsim \sharp B_0$ such that $\sharp(AB) \lesssim (\sharp A_0)^{\frac{1}{2}} (\sharp B_0)^{\frac{1}{2}}$. We have $\mu(A) = \mu_0(A) \gtrsim 1, \nu(B) \gtrsim 1$, it suffices to show the following lemma. $\square$

Lemma 5.12

Assume $\sharp AB \leqslant K(\sharp A)^{\frac{1}{2}} (\sharp B)^{\frac{1}{2}}$, then there exists $K^{O(1)}$-approximate subgroup H, $\exists a, b \in G$ such that

$$\sharp(A \cap aH) \gtrsim \sharp A, \quad \sharp(B \cap Hb) \gtrsim \sharp B.$$

Exercise 5.13. Assume $\sharp A \cdot A^{-1} \leqslant K \sharp A$. Then $\exists S \subseteq G$ symmetric such that

$$\sharp S \geqslant \frac{\sharp A}{2K} \quad \text{and} \quad \sharp \left(A \left(\prod_n S \right) A^{-1} \right) \leqslant 2^n K^{2n+1} \sharp A, \quad \forall n \geqslant 0.$$

Show this statement by the following steps.

I. $\mathscr{E}(A, A^{-1}) = \mathscr{E}(A^{-1}, A)$.

II. Let $S = \left\{x \in G : r_{A^{-1} \cdot A}(x) \geqslant \frac{1}{2K} \sharp A\right\}$, show that $\sharp S \geqslant \frac{1}{2K} \sharp A$.

III. $\forall a, b \in A$, $\forall x_1, \cdots, x_n \in S$, bounded from below the number of ways to write $a x_1 x_2 \cdots x_n b^{-1}$ as $y_1 y_2 \cdots y_{n+1}$, where $y_j \in AA^{-1}$.

IV. Conclude

Proof of Lemma assuming Exercise. By R-triangle, we have $\sharp AA^{-1} \lesssim \sharp A$. Take S as in the exercise, let $H = SS$. Then $\sharp(SSS) \lesssim \sharp A \lesssim \sharp S$, hence H is a $O(K^{O(1)})$-approximate subgroup. Besides $\sharp(AH) \lesssim \sharp H$, by R-covering, there holds $A \subseteq XHH \subseteq X'H$, where $\sharp X \lesssim 1, \sharp X' \lesssim 1$. Then theres is some $x \in X'$ such that $\sharp(A \cap xH) \gtrsim \sharp A$. $\square$

Proposition 5.14 (Bourgain-Gamburd expansion machine)

Γ group, $S \subseteq \Gamma$ finite, $S = S^{-1}$. Assume G is a finite quotient of Γ and $\pi : \Gamma \to G$ is the natural projection. Let $\chi_S = \frac{1}{\sharp S}\mathbb{1}_S$ and $\mu = \pi_*\chi_S$. Assume that

- (quasi-randomness) minimal degree of non-trivial irreducible linear representation of G over $\mathbb{C}$ is at least $(\sharp G)^\kappa$.
- (non-concentration in approximate subgroup) $\exists n_0 \leqslant C\log \sharp G$, such that $\forall K$-approximate subgroup $H \subseteq G$,

$$\text{either}\quad \sharp H \geqslant \frac{1}{CK^C}\sharp G, \quad \text{or}\quad \mu^{*2n_0}(H) \leqslant CK^C(\sharp G)^{-\kappa}.$$

Then $\mathrm{Spec}(T_0(\chi_S)) \cap [1-c,1] = \varnothing$ for some $c = c(\kappa, C) > 0$.

Lemma 5.15 (L^2-flattening)

Same assumption as above, $\forall\delta > 0, \exists\varepsilon = \varepsilon(\delta,\kappa) > 0$, let $\nu = \mu^{*n}$ where $n \geqslant n_0$. Assume $\|\nu\|^2 \geqslant (\sharp G)^{-1+\delta}$, then $\|\nu * \nu\| \leqslant (\sharp G)^{-\varepsilon}\|\nu\|$.

Proof. Assume for a contradiction. Let $K = (\sharp G)^\varepsilon$, by B-S-G, there exists $H \subseteq G$ an $O(K^{O(1)})$-approximate subgroup such that $\sharp H \sim \|\nu\|^{-2} \leqslant (\sharp G)^{1-\delta}$ and $\nu(aH) \gtrsim 1$ for some $a \in G$. For every $x \in G$, we have

$$\mu^{*n_0}(xH)^2 = \mu^{*n_0}(Hx^{-1})\mu^{*n_0}(xH) \leqslant \mu^{*2n_0}(HH).$$

Because HH is also an $O(K^{O(1)})$-approximate subgroup, by the assumption, at least one of the followings holds:

(1) $(\sharp G)^{1-\delta} \gtrsim \sharp(HH) \gtrsim \sharp G$.

(2) $\mu^{*2n_0}(HH) \lesssim (\sharp G)^{-\kappa}$, then $1 \lesssim \nu(aH) \lesssim (\sharp G)^{-\frac{\kappa}{2}}$.

Take $\varepsilon = \varepsilon(\delta,\kappa)$ sufficiently small, both cases lead to a contradiction. $\square$

Proof of Proposition 5.14. Consequently, $\exists C_0 = C_0(\delta,\kappa)$ such that $\left\|\mu^{*C_0n_0}\right\| \leqslant (\sharp G)^{-1+\delta}$. Let $n_1 = C_0n_0$, let λ be an eigenvalue of $T_0(\chi_S)$, let m_λ be the multiplicity of λ. Consider $L^2(G)$ as the regular representation of G, then

$$L^2(G) = \bigoplus\nolimits_{\rho\in\widehat{G}}(\deg\rho)\rho.$$

Because $T(\chi_S) \in \mathbb{C}[\widehat{G}]$, hence it preserves each ρ, then $m_\lambda \geqslant \deg\rho \geqslant (\sharp G)^\kappa$.

On the other hand,

$$\mathrm{tr}(T(\chi_S)^{2n_1}) = \sum_{g\in G}\left\langle T(\chi_S)^{2n_1}\delta_g, \delta_g\right\rangle = \sum_{g\in G}\|T(\chi_S)^{n_1}\delta_g\|^2 = \sharp G\,\|\mu^{*n_1}\|^2 \leqslant (\sharp G)^\delta.$$

Hence $m_\lambda\lambda^{2n_1} \leqslant (\sharp G)^\delta$, take $\delta = \frac{\kappa}{2}$, then $\lambda^{2n_1} \leqslant (\sharp G)^{-\frac{\kappa}{2}}$. Therefore,

$$\log\lambda \leqslant -\frac{\kappa}{4}\frac{\log(\sharp G)}{C_0n_0} \leqslant -\frac{\kappa}{4CC_0} \implies \lambda \leqslant 1-c.$$

$\square$

Quasi-randomness

Remark 5.16 — Gowers shows that if finite group G is κ-quasi-randomness, then Cayley graph of G for some generator sets is quasi-randomness graph.

Theorem 5.17 (Frobenius)
Let $G = \mathrm{SL}(2,\mathbb{F}_p)$, let ρ be a non-trivial irreducible linear representation of G, then $\deg\rho \geqslant \frac{p-1}{2}$.

Proof. Let $(\rho,\mathcal{H})$ be a non-trivial linear representation of G. Consider $U = \left\{\begin{bmatrix}1 & * \\ & 1\end{bmatrix}\right\} \subseteq G$, then $U \cong \mathbb{F}_p$ is abelian. For $a \in \mathbb{F}_p$, let $\chi_a : \mathbb{F}_p \to \mathbb{C}, x \mapsto e(\frac{xa}{p})$. Then we have a decomposition

$$\mathcal{H} = \sum_{a\in\mathbb{F}_p} \mathcal{H}_a, \quad \mathcal{H}_a = \{\xi \in \mathcal{H} : \forall u \in U : \rho(u)\xi = \chi_a(u)\xi\}.$$

For $a_t = \begin{bmatrix}t & \\ & t^{-1}\end{bmatrix}, u \in U$, we have $a_t^{-1}ua_t = u^{-t^2}$. Then $\forall \xi \in \mathcal{H}_a, u \in U$,

$$\rho(u)\rho(a_t)\xi = \rho(a_t)\rho(a_t^{-1}ua_t)\xi = \rho(a_t)\chi_a(u)^{t^{-2}}\xi = \chi_{t^{-2}a}\rho(a_t)\xi.$$

Given $a \in \mathbb{F}_p$, the orbit $\{t^{-2}a : t \in \mathbb{F}_p^\times\}$ is either $\{0\}$ or have $\frac{p-1}{2}$ elements. Then $\dim\mathcal{H} \geqslant \frac{p-1}{2}$, otherwise $\mathcal{H} = \mathcal{H}_0$. In the second case, $U \in \ker\rho$, but $\ker\rho$ is a normal subgroup of G, hence ρ is trivial. □

Non-concentration in approximate subgroup

Proposition 5.18
Let $S \subseteq \mathrm{SL}(2,\mathbb{Z})$ be a finite set, $S = S^{-1}$, freely generates a non-abelian free group. Then $\exists\kappa > 0, \exists C > 0$, such that for every prime p, there is some $n_0 \leqslant C\log p$, such that for every K-approximate subgroup $H \subseteq G_p$,

$$\text{either} \quad \sharp H \gtrsim \sharp G_p \asymp p^3, \quad \text{or} \quad \mu^{*2n_0}(H) \leqslant p^{-\kappa}.$$

Lemma 5.19 (Kesten)
Assume $\sharp S = 2k$, then $\exists c > 0$,

$$\max_{g\in\mathrm{SL}(2,\mathbb{Z})} \chi_S^{*2n}(g) = \chi_S^{*2n}(1) \leqslant \left(\frac{\sqrt{2k-1}}{k}\right)^n \leqslant e^{-cn}.$$

Exercise 5.20. Find a recursive relation and use generating function to prove the lemma.

Remark 5.21 — Let $B_n := \prod_n(\{1\}\cup S)$ be the ball of word metric. Then there is some $c > 0$, such that for every prime p and every $n \leqslant c\log p$, $\pi_p : B_n \mapsto G_p$ is injective. This is because the norms of elements in B_n are with at most exponential growth.

Proof of Proposition 5.18. Let H be a K-approximate subgroup of G_p, by Helfgott's Theorem (4.6), there are three cases:

(1) $\sharp H \lesssim 1$, then $\mu^{*n}(H) \leqslant e^{-cn}\sharp H \lesssim e^{-cn}$.

(2) $\sharp H \gtrsim \sharp G_p$.

(3) $\langle H \rangle \neq G_p$, we need a more technical theorem to deal with this case.

Theorem 5.22 (Dickson)
Let prime $p \geqslant 5$, assume $H \subseteq G_p$ and $\langle H \rangle \neq G_p$, then $\langle H \rangle$ is one of the followings:

(1) dihedral group $D_{2\frac{p\pm 1}{2}}$ or its subgroup.

(2) Borel subgroup $\left\{ \begin{bmatrix} * & * \\ & * \end{bmatrix} \right\} \subseteq G_p$.

(3) A_4, A_5, S_4.

Remark 5.23 — The third case in this theorem is similar with the case $\sharp H \lesssim 1$. For other two cases, we should notice that $\langle H \rangle$ is always a meta-abelian group, i.e.,

$$[[\langle H \rangle, \langle H \rangle], [\langle H \rangle, \langle H \rangle]] = \{1\}.$$

Continued Proof of Proposition 5.18. Take $n = \frac{c}{16} \log p$, we have

$$\mu^{*n}(H) \leqslant e^{-cn}\sharp(B_n \cap \pi_p^{-1}(H)).$$

Let $X = B_n \cap \pi_p^{-1}(H)$, we claim that $\sharp X \ll n^2$. Note that $[[X, X], [X, X]] \subseteq B_{16n}$, hence π_p is injective on it, which shows $[[X, X], [X, X]] = \{1\}$.

Let $z \in [X, X] \setminus \{1\}$, we have $[X, X] \in C(z)$. But S freely generates a non-abelian free group, we can show that

$$\sharp[X, X] \leqslant \sharp(C(z) \cap B_{4n}) \ll n.$$

Then there is $y \in X, b \in [X, X]$ such that

$$\sharp \{x \in X : [x, y] = b\} \gg \frac{\sharp X}{n}.$$

Take some x, then

$$\frac{\sharp X}{n} \ll \sharp(B_n \cap xC(y)) \ll n \implies \sharp X \ll n^2.$$

□

Combining above discussions, given $S \in \mathrm{SL}(2, \mathbb{Z})$, we can show that $(G_p, (\pi_p)_*\chi_S)$ satisfies the quasi-randomness condition and the non-concentration condition with parameters C, κ independent with p. By B-G expansion machine (5.14), $T_{p,0}(\chi_S)$ has a uniform spectral gap. This concludes the uniform expansion in $\mathrm{SL}(2, \mathbb{F}_p)$ (5.2). □

§6 Discretized sum-product theorems

The discretized settings: $A \subseteq \mathbb{R}$ bounded, $\delta > 0$.

Definition 6.1. The **δ-covering number (metric entropy)** of A is defined as

$$\mathcal{N}_\delta(A) := \min\left\{k \in \mathbb{N} : \exists x_1, x_2, \cdots, x_k, A \subseteq \bigcup_{i=1}^{n} B(x_i, \delta)\right\}.$$

Notation 6.2. $|A|$ denotes the Lebesgue measure of A. $A^{(\delta)} = A + B(0,\delta)$ be the δ-neighborhood of A.

Definition 6.3. A is called **δ-separate** if $\forall a \neq a' \in A, d(a,a') > \delta$.

We can also consider

$$\frac{|A^{(\delta)}|}{|B(0,\delta)|}, \quad \sharp\widetilde{A} \text{ with } \widetilde{A} \text{ maximal } \delta\text{-separated subset},$$

$$\sharp\left\{k \in \mathbb{Z} : k\delta \in A^{(\delta)}\right\}, \quad \sharp\left\{k \in \mathbb{Z} : [k\delta, (k+1)\delta[\cap A = \varnothing\right\}.$$

Exercise 6.4. Show that all the quantities are big O of each other.

Remark 6.5 — How to understand $\mathcal{N}_\delta(A)$? We will always view δ as the size of a pixel or the resolution. Then think of $\mathcal{N}_\delta(A)$ as the number of pixels A needed at this resolution.

Some similar results hold:

1. (Ruzsa triangle) $\mathcal{N}_\delta(A-C)\mathcal{N}_\delta(B) \ll \mathcal{N}_\delta(A-B)\mathcal{N}_\delta(B-C)$.
2. (Ruzsa covering) If $\mathcal{N}_\delta(A+B) \leqslant K\mathcal{N}_\delta(A)$, then $B \subseteq A - A + \mathbb{O}(K) + B(0,\delta)$.
3. (Plünnecke-Ruzsa) If $\mathcal{N}_\delta(A+B) \leqslant K\mathcal{N}_\delta(A)$, then

$$\mathcal{N}_\delta\left(\sum_k B - \sum_l B\right) \ll_{k,l} K^{k+l}\mathcal{N}_\delta(A), \quad \forall k,l \in \mathbb{N}.$$

Definition 6.6. Let $\varphi : A \to \mathbb{R}$, the **$\varphi$-energy** of A at scale δ is

$$\mathscr{E}_\delta(\varphi, A) = \mathcal{N}_\delta\left((a,a') \in A \times A : |\varphi(a) - \varphi(a')| \leqslant \delta\right).$$

Remark 6.7 — We fix a norm on $\mathbb{R}^2$ to talk about $\mathcal{N}_\delta(B)$ with $B \subseteq \mathbb{R}^2$.

In particular, the additive energy between $A, B \subseteq \mathbb{R}$ at scale δ is

$$\mathscr{E}_\delta(+, A \times B), \quad \text{where } + : \mathbb{R} \times \mathbb{R} \to \mathbb{R}.$$

Theorem 6.8 (B-S-G)
The following are equivalent, the parameter $K_i > 0$ differs from each other by at most a polynomial dependence:

(i) $\mathscr{E}_\delta(+, A \times B) \geqslant \frac{1}{K_1}\mathcal{N}_\delta(A)^{\frac{3}{2}}\mathcal{N}_\delta(B)^{\frac{3}{2}}$.

(ii) $\exists G \subseteq A \times B$ such that

$$\mathcal{N}_\delta(G) \geqslant \frac{1}{K_2}\mathcal{N}_\delta(A)\mathcal{N}_\delta(B) \quad \text{and} \quad \mathcal{N}_\delta(A \overset{G}{+} B) \leqslant K_2\mathcal{N}_\delta(A)^{\frac{1}{2}}\mathcal{N}_\delta(B)^{\frac{1}{2}}.$$

(iii) $\exists A' \subseteq A, B' \subseteq B$ such that $\mathcal{N}_\delta(A') \geqslant \frac{1}{K_3}\mathcal{N}_\delta(A), \mathcal{N}_\delta(B') \geqslant \frac{1}{K_3}\mathcal{N}_\delta(B)$ and

$$\mathcal{N}_\delta(A' + B') \leqslant K_3\mathcal{N}_\delta(A)^{\frac{1}{2}}\mathcal{N}_\delta(B)^{\frac{1}{2}}.$$

Lemma 6.9
$\varphi : A \to \mathbb{R}$, then

$$\mathscr{E}_\delta(\varphi, A)\mathcal{N}_\delta(\varphi(A)) \gg \mathcal{N}_\delta(A)^2.$$

Sum-product estimate

Notation 6.10. $R_\delta(A, K) = \{x \in \mathbb{R} : \mathcal{N}_\delta(A + xA) \leqslant K\mathcal{N}_\delta(A)\}$.

Assume $A \subseteq B(0, 1) \subseteq \mathbb{R}$, let $K, L \geqslant 1$, there are some properties:

1. $R_\delta(A, K)^{(K\delta)} \subseteq R_\delta(A, O(K^2))$.
2. $\forall s \geqslant 1$, $\langle R_\delta(A, K)\rangle_s \subseteq R_\delta(A, O_s(K^{O_s(1)}))$.
3. If $x \in R_\delta(A, K) \setminus B(0, L^{-1})$, then $x^{-1} \in R_\delta(A, KL)$.
4. If $\mathcal{N}_\delta(A + A) \leqslant K\mathcal{N}_\delta(A)$ and $\mathcal{N}_\delta(A + AA) \leqslant K\mathcal{N}_\delta(A)$, then

$$\mathcal{N}_\delta(\langle A\rangle_s) \ll_s K^{O_s(1)}\mathcal{N}_\delta(A), \quad \forall s \geqslant 1.$$

Remark 6.11 — $\mathcal{N}_\delta(AA)$ can be smaller than $\mathcal{N}_\delta(A)$. For example, let $A = B(0, \delta^{\frac{1}{2}})$, than $\mathcal{N}_\delta(A) \approx \delta^{-\frac{1}{2}}$ and $\mathcal{N}_\delta(AA) = 1$. That is, at scale δ, some points are somehow nilpotent.

Definition 6.12. The **Minkowski lower/upper dimension** are defined as

$$\underline{d}_M(A) = \liminf_{\delta \to 0^+} -\frac{\log \mathcal{N}_\delta(A)}{\log \delta}, \quad \overline{d}_M(A) = \limsup_{\delta \to 0^+} -\frac{\log \mathcal{N}_\delta(A)}{\log \delta}.$$

Theorem 6.13 (Bourgain Sum-Product Theorem)
$\forall \sigma \in (0,1), \exists \varepsilon = \varepsilon(\sigma) > 0$ such that for every $A \subseteq B(0,1) \subseteq \mathbb{R}, \delta > 0$ sufficiently small, assume that

- $\mathcal{N}_\delta(A) \leqslant \delta^{-\sigma-\varepsilon}$.
- (Frostman type non-concentration)

$$\forall \rho \geqslant \delta, \quad \max_{x\in\mathbb{R}} \mathcal{N}_\delta(A \cap B(x,\rho)) \leqslant \delta^{-\varepsilon}\rho^\sigma \mathcal{N}_\delta(A).$$

Then $\mathcal{N}_\delta(A + AA) \geqslant \delta^{-\varepsilon}\mathcal{N}_\delta(A)$.

Remark 6.14 — The conclusion does not hold without the non-concentration condition, for example, $A = B(0, \delta^{\frac{1}{2}})$.

Remark 6.15 — By a variant of Katz-Tao lemma (2.12), the conclusion can be replaced by $\max\{\mathcal{N}_\delta(A + A), \mathcal{N}_\delta(AA)\} \geqslant \delta^{-\varepsilon}\mathcal{N}_\delta(A)$.

Observation 6.16. For $A \subseteq \mathbb{R}, \delta < \delta'$, we have $\mathcal{N}_{\delta'}(A) \leqslant \mathcal{N}_\delta(A) \ll \frac{\delta'}{\delta}\mathcal{N}_{\delta'}(A)$.

Observation 6.17. For $A, B \subseteq \mathbb{R}, B \subseteq B(0,\rho)$, we have $\mathcal{N}_\delta(A + B) \geqslant \mathcal{N}_\rho(A)\mathcal{N}_\delta(B)$.

Proof. Let $\gamma = \gamma(\delta) > 0$ to be determined, let

$$F = \frac{A - A}{(A - A) \setminus B(0, \delta^\gamma)}.$$

Assume for a contradiction that

$$\mathcal{N}_\delta(A + AA) \leqslant \delta^{-\varepsilon}\mathcal{N}_\delta(A).$$

Let $\rho = \delta^{\frac{\varepsilon}{\sigma}}$, then $A \setminus B(0, \delta^{\frac{\varepsilon}{\sigma}}) \neq \varnothing$ by the non-concentration condition. Then

$$\mathcal{N}_\delta(AA) \geqslant \delta^{O(\frac{\varepsilon}{\sigma})}\mathcal{N}_\delta(A),$$

By the assumption and P-R, we have

$$\mathcal{N}_\delta(A + A) \leqslant \delta^{-O(\varepsilon + \frac{\varepsilon}{\sigma})}\mathcal{N}_\delta(A).$$

This shows that $\langle A\rangle_s \subseteq R_\delta(A, O_s(\delta^{O_s(\varepsilon)}))$ for every $s \geqslant 0$.

Claim Let $\delta_1 = \delta^{1-2\gamma}$, then either $F^{(2\delta_1)} \supseteq [0,1]$ or $\exists x \in F$, $\frac{x+1}{2} \notin F^{(\delta_1)}$ or $\frac{x}{2} \notin F^{(\delta_1)}$.
Proof of Claim. Assume $\forall x \in F, \frac{x+1}{2}, \frac{x}{2} \in F^{(\delta_1)}$. Then for every $x \in F^{(2\delta_1)}$, we have $\frac{x+1}{2}, \frac{x}{2} \in F^{(2\delta_1)}$. Because $0, 1 \in F \subseteq F^{(2\delta_1)}$, then $[0,1] \subseteq F^{(2\delta_1)}$.

Dense case: $F^{(2\delta_1)} \supseteq [0,1]$.
Then $\mathcal{N}_{\delta_1}(F) \gg \delta_1^{-1}$. Let $\widetilde{F} \subseteq F, \widetilde{A} \subseteq A \setminus B(0, \delta^\gamma)$ be maximal δ_1-separated sets. Consider

$$\widetilde{A} \times \widetilde{F} \to (AA - AA) \times (AA - AA), \quad (a, x) \mapsto (au_x, av_x), x = \frac{u_x}{v_x}.$$

We show that this map is injective and the image is $\frac{\delta}{C}$-separated. Assume $a'u_{x'} = au_x + O(\frac{\delta}{C}), a'v_{x'} = av_x + O(\frac{\delta}{C})$, then

$$|a|, |v_x| \geqslant \delta^\gamma \implies x' = \frac{au_{x'}}{av_{x'}} = \frac{au_x + O(\frac{\delta}{C})}{av_x + O(\frac{\delta}{C})} = \frac{u_x}{v_x} + O\left(\frac{\delta_1}{C}\right).$$

Choose C large enough, it implies that $|x-x'| \leqslant \delta_1$ and hence $x' = x$. By $\widetilde{A}$ is δ_1-separated, we have $a' = a$. Hence, by P-R,

$$\sharp\widetilde{A}\sharp\widetilde{F} \ll \mathcal{N}_\delta(AA - AA)^2 \leqslant \delta^{-O(\varepsilon)}\mathcal{N}_\delta(A)^2.$$

Because $\sharp\widetilde{F} \asymp \mathcal{N}_{\delta_1}(F) \asymp \delta_1^{-1} = \delta^{-1+2\gamma}$, and

$$\sharp\widetilde{A} \asymp \mathcal{N}_{\delta_1}(A \setminus B(0,\delta^\gamma)) \gg \delta^{-2\gamma}\mathcal{N}_\delta(A \setminus B(0,\delta^\gamma)) \gg \delta^{-2\gamma}(\mathcal{N}_\delta(A) - \delta^{-\varepsilon}\delta^{\gamma\sigma}\mathcal{N}_\delta(A)).$$

Choose γ small such that $\delta^{\gamma\sigma-\varepsilon} \leqslant \frac{1}{2}$, then

$$\mathcal{N}_\delta(A) \gg \delta^{-1+O(\gamma)+O(\varepsilon)}$$

contradict with $\mathcal{N}_\delta(A) \leqslant \delta^{-\varepsilon-\sigma}$ when γ, ε small enough.

Gap case: $\exists x \in F$, such that $\frac{x+1}{2} \notin F^{(\delta_1)}$ or $\frac{x}{2} \notin F^{(\delta_1)}$.

Write $\frac{x+1}{2}$ or $\frac{x}{2}$ as $\frac{u}{v}$, then $u, v \in A - A + A - A$ and $|v| \geqslant \delta^\gamma$. We know $u, v \in R_\delta(A, O(\delta^{-O(\varepsilon)}))$, by R-covering and P-R, we have $\mathcal{N}_\delta(A + uA + vA) \ll \delta^{-O(\varepsilon)}\mathcal{N}_\delta(A)$. We want to prove a lower bound on $\mathcal{N}_\delta(uA + vA)$. Consider

$$\varphi : A \times A \to \mathbb{R}, \quad (a,b) \mapsto ua + vb,$$

it suffices to give an upper bound for $\mathscr{E}_\delta(\varphi, A\times A)$. For $a, b, c, d \in A$, if $|u(a-c)+v(b-d)| \leqslant \delta$, then

$$\left|\frac{u}{v} - \frac{d-b}{a-c}\right| \leqslant \frac{\delta}{|v||a-c|}.$$

Because $\frac{u}{v} \notin F^{(\delta_1)}, |v| \geqslant \delta^\gamma$, then $|a-c| \leqslant \delta^\gamma$. Now we estimate the choices of (a,b,c,d) :

- Choice for $a : \mathcal{N}_\delta(A)$ choices, choice for $b : \mathcal{N}_\delta(A)$ choices.
- Fix a, choice for $c : \mathcal{N}_\delta(A \cap B(a, \delta^\gamma)) \leqslant \delta^{-\varepsilon+\gamma\sigma}\mathcal{N}_\delta(A)$.
- Fix a, b, c, choice for $d : \mathcal{N}_\delta(A \cap B(-, \frac{\delta}{|v|})) \leqslant \delta^{-\varepsilon}(\frac{\delta}{|v|})^\sigma\mathcal{N}_\delta(A)$.

Then

$$\mathscr{E}_\delta(\varphi, A\times A) \leqslant \delta^{-O(\varepsilon)+\gamma\sigma+\sigma}|v|^{-\sigma}\mathcal{N}_\delta(A)^4 \implies \mathcal{N}_\delta(uA+vA) \geqslant |v|^\sigma\delta^{O(\varepsilon)-\gamma\sigma-\sigma}.$$

Because

$$\mathcal{N}_\delta(A) \leqslant \mathcal{N}_{2|v|}(A)\max_x \mathcal{N}_\delta(A \cap B(x, 2|v|)) \ll \delta^{-\varepsilon}|v|^\sigma\mathcal{N}_\delta(A),$$

and notice that $uA + vA \subseteq B(0, 2|v|)$, then

$$\mathcal{N}_\delta(A + uA + vA) \gg \mathcal{N}_{2|v|}(A)\mathcal{N}_\delta(uA+vA) \gg |v|^{-\sigma}|v|^\sigma\delta^{O(\varepsilon)-\gamma\sigma-\sigma}.$$

Then $\delta^{-\sigma-\varepsilon} \geqslant \mathcal{N}_\delta(A) \geqslant \delta^{-\sigma-\gamma\sigma-O(\varepsilon)}$, choose γ, ε small enough, a contradiction. $\square$

Remark 6.18 — The idea of this proof is like the original sum-product theorem.

I. We first show that F is not much bigger than A, in the dense case. Where if we choose γ, ε small enough, we can get $\sharp\widetilde{F}$ is not much bigger than $\sharp\widetilde{A}$.

II. In the gap case, if there are some $x \notin F^{(\delta)}$, we can conclude that $\mathcal{N}_\delta(A + xA)$ is big. This is similar to the fact in the original sum-product theorem: if $\sharp(A + xA) \leqslant (\sharp A)^2$, then $x \in \frac{A-A}{A-A}$. If we can show $F \subseteq R_\delta(A, \delta^{-O(\varepsilon)})$ and some "ring structure" of F, the conclusion will follow.

Theorem 6.19 (Bourgain Sum-Product Theorem, another version)
$\forall \sigma \in (0,1), \kappa > 0, \exists \varepsilon = \varepsilon(\sigma, \kappa) > 0$ such that for every $A \subseteq B(0,1) \subseteq \mathbb{R}$ and $\delta > 0$ sufficiently small, assume that

- $\mathcal{N}_\delta(A) \leqslant \delta^{-\sigma-\varepsilon}$.
- $\forall \rho \geqslant \delta, \mathcal{N}_\rho(A) \geqslant \delta^{\varepsilon}\rho^{-\kappa}$.

Then $\mathcal{N}_\delta(A + AA) \geqslant \delta^{-\varepsilon}\mathcal{N}_\delta(A)$.

Proof. We prove a special case of $\kappa = \sigma$. Assume $\mathcal{N}_\delta(A + AA) \leqslant \delta^{-\varepsilon}\mathcal{N}_\delta(A)$, consider $\rho = \delta^{\frac{\varepsilon}{\sigma}}$, we can also have $A \setminus B(0,\rho) \neq \varnothing$. A same argument, we have $\mathcal{N}_\delta(A + A + AA) \leqslant \delta^{-O(\varepsilon)}\mathcal{N}_\delta(A)$. Hence

$$\delta^{-O(\varepsilon)}\mathcal{N}_\delta(A) \geqslant \mathcal{N}_\delta(A + A + AA) \geqslant \mathcal{N}_\delta(A + A) \geqslant \mathcal{N}_\rho(A)\max_{x\in\mathbb{R}}\mathcal{N}_\delta(A \cap B(x,\rho)),$$

then $\max_{x\in\mathbb{R}}\mathcal{N}_\delta(A \cap B(x,\rho)) \leqslant \delta^{-O(\varepsilon)}\rho^{\sigma}\mathcal{N}_\delta(A)$. Gives the condition in last version. □

§7 Projection theorem

Let $\mathbb{S}^1 = \left\{\theta \in \mathbb{R}^2 : \|\theta\| = 1\right\}$ be the unit circle in $\mathbb{R}^2$, for every $\theta \in \mathbb{S}^1$, let

$$\mathrm{proj}_\theta : \mathbb{R}^2 \to \mathbb{R} \cdot \theta$$

be the orthogonal projection.

Theorem 7.1 (Bourgain's Projection Theorem)
For every $\alpha \in (0,1), \kappa > 0$, there exists $\varepsilon = \varepsilon(\alpha, \kappa) > 0$, the following holds for $\delta > 0$ sufficiently small. Let $A \subseteq B_{\mathbb{R}^2}(0,1)$, assume

- $\mathcal{N}_\delta(A) \leqslant \delta^{-2\alpha}$.
- $\forall \rho \geqslant \delta, \forall x \in \mathbb{R}^2$, $\mathcal{N}_\delta(A \cap B(x,\rho)) \leqslant \delta^{-\varepsilon}\rho^{2\alpha}\mathcal{N}_\delta(A)$.

Write

$$\mathscr{E} = \left\{\theta \in \mathbb{S}^1 : \mathcal{N}_\delta(\mathrm{proj}_\theta A) \leqslant \delta^{-\alpha-\varepsilon}\right\},$$

then $\mathscr{E}$ does not support a probability measure μ satisfying

$$\mu(B_{\mathbb{S}^1}(\theta,\rho)) \leqslant \delta^{-\varepsilon}\rho^{\kappa}, \quad \forall \rho \geqslant \delta, \theta \in \mathbb{S}^1.$$

Remark 7.2 — $\mathscr{E}$ refers to an exception.

Example 7.3

Let $A = B(0, \delta^{\frac{1}{2}})$, then $\mathcal{N}_\delta(A) \asymp \delta^{-1}$. Notice that $\mathcal{N}_\delta(\mathrm{proj}_\theta A) \asymp \delta^{-\frac{1}{2}}$, hence $\mathscr{E} = \mathbb{S}^1$. This is a contradiction. The reason is that A does not satisfy the second condition (non concentration condition).

Example 7.4

$A = C \times C$, where C is a Cantor set. We can choose C let $\mathcal{N}_\delta(C - C) \asymp \mathcal{N}_\delta(C)$, then when θ near $0, \frac{\pi}{2}, \frac{\pi}{4}$, $\mathcal{N}_\delta(\mathrm{proj}_\theta A)$ is small. For more other θ's, $\mathcal{N}_\delta(\mathrm{proj}_\theta A)$ is large.

Idea Write $\theta = \theta_t = \frac{(1,t)}{\sqrt{1+t^2}}$, where $t \in [\frac{1}{2}, 2]$. Then

$$\mathrm{proj}_\theta(x, y) = \frac{\langle \theta, (x, y)\rangle}{\langle \theta, \theta\rangle}\theta = (x + ty)\frac{\theta}{\sqrt{1+t^2}}.$$

Consider a special case for $A = A_0 \times A_0$, then

$$\mathcal{N}_\delta(\mathrm{proj}_\theta A) \asymp \mathcal{N}_\delta(A_0 + tA_0).$$

Then $\mathscr{E}$ is almost the set $R_\delta(A_0, \delta^{-\varepsilon})$.

Theorem 7.5 (Bourgain Sum-Product Theorem, another version)

$\forall \alpha \in (0, 1), \forall \kappa > 0, \exists \varepsilon = \varepsilon(\alpha, \kappa) > 0$, such that for every $A_0 \subseteq B(0, 1) \subseteq \mathbb{R}$ and $\delta > 0$ sufficiently small, assume that

- $\mathcal{N}_\delta(A_0) \leqslant \delta^{-\alpha-\varepsilon}$.
- $\forall \rho \geqslant \delta, x \in \mathbb{R}$, $\mathcal{N}_\delta(A_0 \cap B(x, \rho)) \leqslant \delta^{-\varepsilon}\rho^\kappa \mathcal{N}_\delta(A_0)$.

Then for every $B_0 \subseteq \mathbb{R}$ such that $\forall \rho \geqslant \delta$, $\mathcal{N}_\rho(B_0) \geqslant \delta^\varepsilon \rho^{-\kappa}$, there exists $t \in B_0$, such that $\mathcal{N}_\delta(A_0 + tA_0) \geqslant \delta^{-\varepsilon}\mathcal{N}_\delta(A_0)$.

Remark 7.6 — The condition $\forall \rho \geqslant \delta$, $\mathcal{N}_\rho(B_0) \geqslant \delta^\varepsilon \rho^{-\kappa}$ is strictly weaker than the non concentration condition for B_0.

Lemma 7.7

$\kappa, \varepsilon > 0$, let μ be a probability measure on $\mathbb{R}$, $\operatorname{supp}\mu \subseteq B(0, 1)$, satisfying

$$\mu(B(x, \rho)) \leqslant \delta^{-\varepsilon}\rho^\kappa, \quad \forall \rho \geqslant \delta, x \in \mathbb{R}.$$

Then $\exists B_0 \subseteq \operatorname{supp}\mu$ satisfying

$$\mathcal{N}_\delta(B_0 \cap B(x, \rho)) \leqslant \delta^{-O(\varepsilon)}\rho^\kappa \mathcal{N}_\delta(B_0), \quad \forall \rho \geqslant \delta, x \in \mathbb{R}.$$

Proof. Let
$$\mathcal{Q} := \{[k\delta, (k+1)\delta[: k \in \mathbb{Z}\},$$
for every $i \in \mathbb{N}$, define
$$\mathcal{Q}_i = \left\{Q \in \mathcal{Q} : 2^{-i-1} < \mu(Q) \leqslant 2^{-i}\right\}.$$
Observe that
$$\mu\left(\bigcup_{\substack{Q\in\mathcal{Q}_i,\\ i\geqslant 2|\log\delta|}} Q\right) \ll \delta^{-1}2^{-2|\log\delta|} \leqslant \delta < \frac{1}{2}.$$
Then $\exists i \in [0, 2|\log\delta|] \cap \mathbb{N}$, such that (for δ sufficiently small)
$$\mu\left(\bigcup_{Q\in\mathcal{Q}_i} Q\right) \geqslant \frac{1}{4|\log\delta|} \geqslant \delta^{\varepsilon}.$$
Fix this i, let $B_1 = \bigcup_{Q\in\mathcal{Q}_i} Q$ and $B_0 = B_1 \cap \operatorname{supp}\mu$. Then for every $\rho \geqslant \delta, x \in \mathbb{R}$,
$$\begin{aligned}\mathcal{N}_\delta(B_0 \cap B(x,\rho)) &\asymp \sharp\left\{Q \in \mathcal{Q}_i : Q \cap B(x,\rho) \neq \varnothing\right\}\\ \ll \frac{\mu(B_0 \cap B(x, 2\rho))}{\min_{Q\in\mathcal{Q}_i}\mu(Q)} &\ll \frac{\delta^{-\varepsilon}\rho^{\kappa}}{\min_{Q\in\mathcal{Q}_i}\mu(Q)} \leqslant \delta^{-2\varepsilon}\rho^{\kappa}\sharp\mathcal{Q}_i \asymp \delta^{-2\varepsilon}\rho^{\kappa}\mathcal{N}_\delta(B_0).\end{aligned}$$
□

Lemma 7.7 + Theorem 7.5 $\implies$ the special case $A = A_0 \times A_0$ of Theorem 7.10.

Proof of General Case of Theorem 7.10. Assume for a contradiction that $\mathscr{E}$ supports such a probability measure μ. In particular, there exists $\theta_1, \theta_2 \in \mathscr{E}$ with $d(\theta_1, \theta_2) \geqslant \delta^{\frac{\varepsilon}{\kappa}} = \delta^{O(\varepsilon)}$. After a rotation and affine transformation of norm at most $\delta^{O(\varepsilon)}$, we can assume that x-axis and y-axis are both in $\mathscr{E}$. Let B, C be the projection of A to the x-axis and y-axis, respectively. For a $\theta_t = \frac{(1,t)}{\sqrt{1+t^2}} \in \mathscr{E}$, we have
$$\mathcal{N}_\delta(B \overset{A}{+} tC) \leqslant \delta^{-\alpha-\varepsilon},$$
here we abuse a notation $\overset{A}{+}$ to refer to $a = (b,c) \in A, b \in B, c \in C$. We have
$$\delta^{-2\alpha+O(\varepsilon)} \leqslant \mathcal{N}_\delta(A) \ll \mathcal{N}_\delta(B)\mathcal{N}_\delta(C),$$
hence $\mathcal{N}_\delta(B), \mathcal{N}_\delta(C) \geqslant \delta^{-\alpha+O(\varepsilon)}$. By B-S-G (6.8), there exists $B_t \subseteq B, C_t \subseteq C$ such that
$$\mathcal{N}_\delta(B_t) \geqslant \delta^{-\alpha+O(\varepsilon)}, \quad \mathcal{N}_\delta(C_t) \geqslant \delta^{-\alpha+O(\varepsilon)}, \quad \mathcal{N}_\delta(B_t + tC_t) \leqslant \delta^{-\alpha-O(\varepsilon)}.$$
If B_t, C_t are independent of t, then done. We need a following lemma.

Lemma 7.8 (popularity argument)
(X, λ) is a finite measure space, (T, ν) is a probability space, $K \geqslant 2$. If $\forall t \in T, X_t \subseteq X$ with $\lambda(X_t) \geqslant \frac{1}{K}\lambda(X)$. Then $\exists t_\star \in T$, such that
$$\nu\left\{t \in T : \lambda(X_{t_\star} \cap X_t) \geqslant \frac{1}{2K^2}\lambda(X)\right\} \geqslant \frac{1}{2K^2}.$$

Exercise 7.9. Prove the lemma. Hint: applying C-S to $x \mapsto \int \mathbb{1}_{X_t}(x) d\nu(t)$.

Continued Proof of Theorem 7.10. Use lemma for $X = B^{(\delta)} \times C^{(\delta)}, \lambda = \text{Leb}$. Let $X_t = B_t^{(\delta)} \times C_t^{(\delta)}$, let ν be the push forward of μ under $\theta_t \mapsto t$. Then $\exists t_\star \in \mathbb{R}, D \subseteq \mathbb{R}$ with $\nu(D) \geqslant \delta^{O(\varepsilon)}$ such that for every $t \in D$,

$$\mathcal{N}_\delta(B_t \cap B_\star) \geqslant \delta^{-\alpha+O(\varepsilon)}, \quad \mathcal{N}_\delta(C_t \cap C_\star) \geqslant \delta^{-\alpha+O(\varepsilon)},$$

where $B_\star = B_{t_\star}$ and $C_\star = C_{t_\star}$.

We use notation $E \approx F$ to refer to $\mathcal{N}_\delta(E-F) \ll \delta^{-O(\varepsilon)} \mathcal{N}_\delta(E)^{\frac{1}{2}} \mathcal{N}_\delta(F)^{\frac{1}{2}}$. Now, we do some Ruzsa calculus. We know $B_t \approx -tC_t$, hence $B_t \approx B_t$, and then $B_t \approx B_t \cap B_\star$. Moreover, for every $t \in D$, we have $B_\star \approx B_t \cap B_\star \approx B_t$ and $C_\star \approx C_t \cap C_\star \approx C_t$. Because $B_\star \approx -t_\star C_\star$, we have

$$B_\star \approx B_t \approx -tC_t \approx -tC_\star \approx \frac{t}{t_\star} B_\star, \quad \forall t \in D \subseteq [\frac{1}{2}, 2].$$

This will contradict with the Sum-Product Theorem (7.5) when ε is small. $\square$

Theorem 7.10 (Bourgain's Projection Theorem, adapted version)

For every $\alpha \in (0,1), \kappa > 0$, there exists $\varepsilon = \varepsilon(\alpha, \kappa) > 0$, the following holds for $\delta > 0$ sufficiently small. Let $A \subseteq B_{\mathbb{R}^2}(0,1)$, assume

- $\mathcal{N}_\delta(A) \leqslant \delta^{-2\alpha}$.
- $\forall \rho \geqslant \delta, \forall x \in \mathbb{R}^2, \mathcal{N}_\delta(A \cap B(x,\rho)) \leqslant \delta^{-\varepsilon} \rho^{2\alpha} \mathcal{N}_\delta(A)$.

Write

$$\mathscr{E} = \left\{ \theta \in \mathbb{S}^1 : \exists A' \subseteq A, \mathcal{N}_\delta(A') \geqslant \delta^\varepsilon \mathcal{N}_\delta(A) \text{ and } \mathcal{N}_\delta(\text{proj}_\theta A') \leqslant \delta^{-\alpha-\varepsilon} \right\},$$

then $\mathscr{E}$ does not support a probability measure μ satisfying

$$\mu(B_{\mathbb{S}^1}(\theta, \rho)) \leqslant \delta^{-\varepsilon} \rho^\kappa, \quad \forall \rho \geqslant \delta, \theta \in \mathbb{S}^1.$$

Corollary 7.11

$\forall \alpha \in (0,1), \exists \varepsilon = \varepsilon(\alpha) > 0$, let $A \subseteq \mathbb{R}^2$ be a Borel subset. If $\dim_H A = 2\alpha$, then

$$\dim_H \left(\left\{ \theta \in \mathbb{S}^1 : \dim_H \text{proj}_\theta A \leqslant \alpha + \varepsilon \right\} \right) = 0.$$

Remark 7.12 — To compare with Marstrand's Theorem: if $\alpha < \frac{1}{2}$, then

$$\text{Leb} \left\{ \theta \in \mathbb{S}^1 : \dim_H \text{proj}_\theta A < 2\alpha \right\} = 0.$$

Recall Hausdorff dimension of A, $\dim_H A \leqslant \alpha$ if and only if $\forall \varepsilon > 0$, $\exists x_i \in \mathbb{R}^2, 0 < r_i < \varepsilon, i \in \mathbb{N}$, such that

$$A \subseteq \bigcup_{i \in \mathbb{N}} B(x_i, r_i), \quad \sum_{i=0}^{\infty} r_i^{\alpha+\varepsilon} < \varepsilon.$$

In other word, $\dim_H A = \inf \{\text{such } \alpha\}$.

Lemma 7.13 (Frostman Lemma)
If $A \subseteq \mathbb{R}^2, \dim_H A > \alpha$, then $\exists$ a finite nonzero Borel measure μ on $\mathbb{R}^2, \operatorname{supp}\mu \subseteq A$, and for every $\rho > 0, x \in \mathbb{R}^2$, $\mu(B(x,\rho)) < \rho^\alpha$.

Remark 7.14 — Such a measure is said to be α-Frostman (or α-Hölder?).

Proof of Corollary 7.11. Assume for a contradiction, let

$$\mathscr{E} = \left\{\theta \in \mathbb{S}^1 : \dim_H \operatorname{proj}_\theta A \leqslant \alpha + \varepsilon\right\},$$

assume $\dim_H \mathscr{E} > \kappa > 0$. By Frostman lemma, there exists μ on $\mathscr{E}$ which is κ-Frostman. There exists ν on A is $(2\alpha - \varepsilon)$-Frostman. For every $\theta \in \mathscr{E}$, we can cover $\operatorname{proj}_\theta A$ by $\bigcup_{i\in\mathbb{N}} B(x_i, r_i)$ with $r_i \leqslant \delta = 2^{k_0}$ and $\sum r_i^{\alpha+2\varepsilon} \leqslant \varepsilon$. WLOG, we can assume that $r_i = 2^{-k_i}$ where $k_i \in \mathbb{N}$, then

$$\operatorname{proj}_\theta A \subseteq \bigcup_{k\geqslant k_0} B_{\theta,k}, \quad \text{where } B_{\theta,k} := \bigcup_{x\in X_{\theta,k}} B(x, 2^{-k}).$$

We also have an estimate for every $k \geqslant k_0$, $\sharp X_{\theta,k} \leqslant 2^{k(\alpha+2\varepsilon)}$. Then

$$\nu(A)\mu(\mathscr{E}) = \int \nu\left(\bigcup\nolimits_{k\geqslant k_0} \operatorname{proj}_\theta^{-1} B_{\theta,k}\right) \mathrm{d}\mu(\theta) \leqslant \int \sum_{k\geqslant k_0} \nu(\operatorname{proj}_\theta^{-1} B_{\theta,k}) \mathrm{d}\mu(\theta).$$

Let $A_{\theta,k} = \operatorname{proj}_\theta^{-1} B_{\theta,k}$, then $\exists k \geqslant k_0$, such that

$$\frac{1}{\nu(A)\mu(\mathscr{E})} \int \nu(A_{\theta,k}) \mathrm{d}\mu(\theta) \geqslant \frac{6}{\pi^2}\frac{1}{k} \gg |\log\delta|^{-2} \geqslant \delta^{-\varepsilon}.$$

Choose $\delta_0 = 2^{-k_0}$ sufficiently small, fix a such k, let $\delta = 2^{-k} \leqslant \delta_0$. We have $A_{\theta,k} \subseteq A$ and

$$\mathcal{N}_\delta(\operatorname{proj}_\theta A_{\theta,k}) \leqslant \sharp X_{\theta,k} \leqslant \delta^{-\alpha-2\varepsilon}.$$

Then $\exists D \subseteq \mathscr{E} \subseteq \mathbb{S}^1$ such that $\mu(D) \geqslant \delta^\varepsilon \mu(\mathscr{E})$ and $\forall \theta \in D$, $\nu(A_{\theta,k}) \geqslant \delta^\varepsilon \nu(A)$. We want to find $B \subseteq A$ such that

$$\mathcal{N}_\delta(A_{\theta,k} \cap B) \geqslant \delta^\varepsilon \mathcal{N}_\delta(B), \quad \mathcal{N}_\delta(B \cap B(x,\rho)) \leqslant \delta^{-O(\varepsilon)} \rho^{2\alpha} \mathcal{N}_\delta(B).$$

A similar argument in the proof of Lemma 7.7, we can find such a B and some $D' \subseteq D$, which contradicts with the Sum-Product theorem. $\square$