LeanAPAP

LeanAPAP project contributors

August 25, 2023

Chapter 1

Almost-Periodicity

Lemma 1.1 (Marcinkiewicz-Zygmund inequality). *Let $m \geq 1$. If $f : G \to \mathbb{R}$ is such that $\mathbb{E}_x f(x) = 0$ and $|f(x)| \leq 2$ for all x then*

$$\mathbb{E}_{x_1,\ldots,x_n} \left| \sum_{i=1}^{n} f(x_i) \right|^{2m} \leq (4mn)^m.$$

Proof. Let S be the left-hand side. Since $0 = \mathbb{E}_y f(y)$ we have, by the triangle inequality, and Hölder's inequality,

$$S = \mathbb{E}_{x_1,\ldots,x_n} \left| \sum_i f(x_i) - \mathbb{E}_{y_i} f(y_i) \right|^{2m} = \mathbb{E}_{x_1,\ldots,x_n} \left| \mathbb{E}_{y_i} \left(\sum_i f(x_i) - f(y_i) \right) \right|^{2m} \leq \mathbb{E}_{x_1,\ldots,y_n} \left| \sum_i f(x_i) - f(y_i) \right|^{2m}.$$

Changing the role of x_i and y_i makes no difference here, but multiplies the i summand by $\{-1, +1\}$, and therefore for any $\epsilon_i \in \{-1, +1\}$,

$$S \leq \mathbb{E}_{x_1,\ldots,y_n} \left| \sum_i \epsilon_i (f(x_i) - f(y_i)) \right|^{2m}.$$

In particular, if we sample $\epsilon_i \in \{-1, +1\}$ uniformly at random, then

$$S \leq \mathbb{E}_{\epsilon_i} \mathbb{E}_{x_1,\ldots,y_n} \left| \sum_i \epsilon_i (f(x_i) - f(y_i)) \right|^{2m}.$$

We now change the order of expectation and consider the expectation over just ϵ_i, viewing the $f(x_i) - f(y_i) = z_i$, say, as fixed quantities. For any z_i we can expand $\mathbb{E}_{\epsilon_i} |\sum_i \epsilon_i z_i|^{2m}$ and then bound it from above, using the triangle inequality and $|z_i| \leq 4$, by

$$4^{2m} \sum_{k_1 + \cdots + k_n = 2m} \binom{2m}{k_1, \ldots, k_n} \left| \mathbb{E} \epsilon_1^{k_1} \cdots \epsilon_n^{k_n} \right|.$$

The inner expectation vanishes unless each k_i is even, when it is trivially one. Therefore the above quantity is exactly

$$\sum_{l_1 + \cdots + l_n = m} \binom{2m}{2l_1, \ldots, 2l_n} \leq m^m n^m,$$

since for any $l_1 + \cdots + l_n = m$,

$$\binom{2m}{2l_1, \dots, 2l_n} \leq m^m \binom{m}{l_1, \dots, l_n}.$$

This can be seen, for example, by writing both sides out using factorials, yielding

$$\frac{(2m)!}{(2l_1)! \cdots (2l_n)!} \leq \frac{(2m)!}{2^m m!} \frac{m!}{l_1! \cdots l_n!} \leq m^m \frac{m!}{l_1! \cdots l_n!}.$$

$\square$

Lemma 1.2 (Complex-valued Marcinkiewicz-Zygmund inequality). *Let $m \geq 1$. If $f : G \to \mathbb{C}$ is such that $\mathbb{E}_x f(x) = 0$ and $|f(x)| \leq 2$ for all x then*

$$\mathbb{E}_{x_1, \dots, x_n} \left| \sum_{i=1}^{n} f(x_i) \right|^{2m} \leq (16mn)^m.$$

Proof. Test. $\square$

Lemma 1.3. *Let $\epsilon > 0$ and $m \geq 1$. Let $A \subseteq G$ and $f : G \to \mathbb{C}$. If $k \geq 64m\epsilon^{-2}$ then the set*

$$L = \{\vec{a} \in A^k : \|\tfrac{1}{k} \sum_{i=1}^{k} f(x - a_i) - \mu_A * f\|_{2m} \leq \epsilon \|f\|_{2m}\}.$$

has size at least $|A|^k / 2$.

Proof. Note that if $a \in A$ is chosen uniformly at random then, for any fixed $x \in G$,

$$\mathbb{E} f(x - a_i) = \frac{1}{|A|} \sum_{a \in A} f(x - a) = \frac{1}{|A|} 1_A * f(x) = \mu_A * f(x).$$

Therefore, if we choose $a_1, \dots, a_k \in A$ independently uniformly at random, for any fixed $x \in G$ and $1 \leq i \leq k$, the random variable $f(x - a_i) - f * \mu_A(x)$ has mean zero. By the Marcinkiewicz-Zygmund inequality Lemma 1.1, therefore,

$$\mathbb{E} \left| \frac{1}{k} \sum_i f(x - a_i) - f * \mu_A(x) \right|^{2m} \leq$$
$$(16m/k)^m k^{-1} \mathbb{E} \sum_i |f(x - a_i) - f * \mu_A(x)|^{2m}.$$

We now sum both sides over all $x \in G$. By the triangle inequality, for any fixed $1 \leq i \leq k$ and $a_i \in A$,

$$\sum_{x \in G} |f(x - a_i) - f * \mu_A(x)|^{2m} \leq 2^{2m-1} \sum_{x \in G} |f(x - a_i)|^{2m} + \sum_{x \in G} |f * \mu_A(x)|^{2m}$$
$$\leq 2^{2m-1} \left(\|f\|_{2m}^{2m} + \|f * \mu_A\|_{2m}^{2m} \right).$$

We note that $\|\mu_A\|_1 = \frac{1}{|A|} \sum_{x \in A} 1_A(x) = |A| / |A| = 1$, and hence by Young's inequality, $\|f * \mu_A\|_{2m} \leq \|f\|_{2m}$, and so

$$\sum_{x \in G} |f(x - a_i) - f * \mu_A(x)|^{2m} \leq 2^{2m} \|f\|_{2m}^{2m}.$$

2

It follows that

$$\mathbb{E}_{a_1,\dots,a_k \in A} \|\frac{1}{k}\sum_i \tau_{a_i} f - f * \mu_A\|_{2m}^{2m} \le (64m/k)^m \|f\|_{2m}^{2m}.$$

In particular, if $k \ge 64\epsilon^{-2}m$ then the right-hand side is at most $(\frac{\epsilon}{2}\|f\|_{2m})^{2m}$ as required. $\square$

Lemma 1.4. *Let $A \subseteq G$ and $f : G \to \mathbb{C}$. Let $\epsilon > 0$ and $m \ge 1$ and $k \ge 1$. Let*

$$L = \{\vec{a} \in A^k : \|\frac{1}{k}\sum_{i=1}^k f(x - a_i) - \mu_A * f\|_{2m} \le \epsilon\|f\|_{2m}\}.$$

If $t \in G$ is such that $\vec{a} \in L$ and $\vec{a} + (t,\dots,t) \in L$ then

$$\|\tau_t(\mu_A * f) - \mu_A * f\|_{2m} \le 2\epsilon\|f\|_{2m}.$$

Proof. Test. $\square$

Lemma 1.5. *Let $A \subseteq G$ and $k \ge 1$ and $L \subseteq A^k$. Then there exists some $\vec{a} \in L$ such that*

$$\#\{t \in G : \vec{a} + (t,\dots,t) \in L\} \ge \frac{|L|}{|A+S|^k}|S|.$$

Proof. Test. $\square$

Theorem 1.6 (L_p almost periodicity)**.** *Let $\epsilon \in (0,1]$ and $m \ge 1$. Let $K \ge 2$ and $A, S \subseteq G$ with $|A + S| \le K|A|$. Let $f : G \to \mathbb{C}$. There exists $T \subseteq G$ such that*

$$|T| \ge K^{-512m\epsilon^{-2}}|S|$$

such that for any $t \in T$ we have

$$\|\tau_t(\mu_A * f) - \mu_A * f\|_{2m} \le \epsilon\|f\|_{2m}.$$

Proof. Test. $\square$

Theorem 1.7 (L_∞ almost periodicity)**.** *Let $\epsilon \in (0,1]$ and $m \ge 1$. Let $K \ge 2$ and $A, S \subseteq G$ with $|A + S| \le K|A|$. Let $B, C \subseteq G$. Let $\eta = \min(1, |C|/|B|)$. There exists $T \subseteq G$ such that*

$$|T| \ge K^{-4096\lceil \mathcal{L}\eta \rceil \epsilon^{-2}}|S|$$

such that for any $t \in T$ we have

$$\|\tau_t(\mu_A * 1_B * \mu_C) - \mu_A * 1_B * \mu_C\|_\infty \le \epsilon.$$

Proof. Let T be as given in 1.6 with $f = 1_B$ and $m = \lceil \mathcal{L}\eta \rceil$ and $\epsilon = \epsilon/e$. (The size bound on T follows since $e^2 \le 8$.) Fix $t \in T$ and let $F = \tau_t(\mu_A * 1_B) - \mu_A * 1_B$. We have, for any $x \in G$,

$$(\tau_t(\mu_A * 1_B * \mu_C) - \mu_A * 1_B * \mu_C)(x) = F * \mu_C(x) = \sum_y F(y)\mu_C(x - y) = \sum_y F(y)\mu_{x-C}(y).$$

By Hölder's inequality, this is (in absolute value), for any $m \ge 1$,

$$\|F\|_{2m}\|\mu_{x-C}\|_{1+\frac{1}{2m-1}}.$$

3

By the construction of T the first factor is at most $\frac{\epsilon}{e}\|1_B\|_{2m} = \frac{\epsilon}{e}|B|^{1/2m}$. We have by calculation

$$\|\mu_{x-C}\|_{1+\frac{1}{2m-1}} = |x-C|^{-1/2m} = |C|^{-1/2m}.$$

Therefore we have shown that

$$\|\tau_t(\mu_A * 1_B * \mu_C) - \mu_A * 1_B * \mu_C\|_\infty \leq \frac{\epsilon}{e}(|C|/|B|)^{-1/2m}.$$

The claim now follows since, by choice of m,

$$(|C|/|B|)^{-1/2m} \leq e$$

(dividing into cases as to whether $\eta = 1$ or not). $\qquad\square$

Theorem 1.8. *Let* $\epsilon \in (0,1]$ *and* $m \geq 1$ *and* $k \geq 1$. *Let* $K \geq 2$ *and* $A, S \subseteq G$ *with* $|A+S| \leq K|A|$. *Let* $B, C \subseteq G$. *Let* $\eta = \min(1, |C|/|B|)$. *There exists* $T \subseteq G$ *such that*

$$|T| \geq K^{-4096\lceil \mathcal{L}\eta\rceil k^2 \epsilon^{-2}}|S|$$

such that

$$\|\mu_T^{(k)} * \mu_A * 1_B * \mu_C - \mu_A * 1_B * \mu_C\|_\infty \leq \epsilon.$$

Proof. Let T be as in Theorem 1.7 with ϵ replaced by ϵ/k. Note that, for any $x \in G$,

$$\mu_T^{(k)} * \mu_A * 1_B * \mu_C(x) = \frac{1}{|T|^k} \sum_{t_1,\ldots,t_k \in T} \tau_{t_1+\cdots+t_k}\mu_A * 1_B * \mu_C(x).$$

It therefore suffices (by the triangle inequality) to show, for any fixed $x \in G$ and $t_1, \ldots, t_k \in T$, that with $F = \mu_A * 1_B * \mu_C$, we have

$$|\tau_{t_1+\cdots+t_k}F(x) - F(x)| \leq \epsilon.$$

This follows by the triangle inequality applied k times if we knew that, for $1 \leq l \leq k$,

$$|\tau_{t_1+\cdots+t_l}F(x) - \tau_{t_1+\cdots+t_{l-1}}F(x)| \leq \epsilon/k.$$

We can write the left-hand side as

$$|\tau_{t_1+\cdots+t_l}F(x) - \tau_{t_1+\cdots+t_{l-1}}F(x)| = |\tau_{t_l}F(x-t_1-\cdots-t-l-1) - F(x-t_1-\cdots-t-l-1)|.$$

The right-hand side is at most

$$\|\tau_{t_l}F - F\|_\infty$$

and we are done by choice of T. $\qquad\square$

Chapter 2

Chang's lemma

Definition 2.1 (Large spectrum)**.** *Let G be a finite abelian group and $f : G \to \mathbb{C}$. Let $\eta \in \mathbb{R}$. The η-large spectrum is defined to be*

$$\Delta_\eta(f) = \{\gamma \in \widehat{G} : |\widehat{f}(\gamma)| \geq \eta\|f\|_1\}.$$

Definition 2.2 (Weighted energy)**.** *Let $\Delta \subseteq \widehat{G}$ and $m \geq 1$. Let $\nu : G \to \mathbb{C}$. Then*

$$E_{2m}(\Delta; \nu) = \sum_{\gamma_1, \dots, \gamma_{2m} \in \Delta} |\widehat{\nu}(\gamma_1 + \cdots - \gamma_{2m})|.$$

Definition 2.3 (Energy)**.** *Let G be a finite abelian group and $A \subseteq G$. Let $m \geq 1$. We define*

$$E_{2m}(A) = \sum_{a_1, \dots, a_{2m} \in A} 1_{a_1 + \cdots - a_{2m} = 0}.$$

Definition 2.4 (Dissociation)**.** *We say that $A \subseteq G$ is dissociated if, for any $m \geq 1$, and any $x \in G$, there is at most one $A' \subset A$ of size $|A'| = m$ such that*

$$\sum_{a \in A'} a = x.$$

Lemma 2.5. *Let G be a finite abelian group and $f : G \to \mathbb{C}$. Let $\nu : G \to \mathbb{R}_{\geq 0}$ be such that whenever $|f| \neq 0$ we have $\nu \geq 1$. Let $\Delta \subseteq \Delta_\eta(f)$. Then, for any $m \geq 1$.*

$$\eta^{2m} \frac{\|f\|_1^2}{\|f\|_2^2} |\Delta|^{2m} \leq E_{2m}(\Delta; \nu).$$

Proof. By definition of $\Delta_\eta(f)$ we know that

$$\eta\|f\|_1 |\Delta| \leq \sum_{\gamma \in \Delta} |\widehat{f}(\gamma)|.$$

There exists some $c_\gamma \in \mathbb{C}$ with $|c_\gamma| = 1$ for all γ such that

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

Interchanging the sums, therefore,

$$\eta \|f\|_1 |\Delta| \le \sum_{x \in G} f(x) \sum_{\gamma \in \Delta} c_\gamma \overline{\gamma(x)}.$$

By Hölder's inequality the right-hand side is at most

$$\left(\sum_x |f(x)| \right)^{1-1/m} \left(\sum_x |f(x)| \left| \sum_{\gamma \in \Delta} c_\gamma \overline{\gamma(x)} \right|^m \right)^{1/m}.$$

Taking mth powers, therefore, we have

$$\eta^m |\Delta|^m \|f\|_1 \le \sum_x |f(x)| \left| \sum_{\gamma \in \Delta} c_\gamma \overline{\gamma(x)} \right|^m.$$

By assumption we can bound $|f(x)| \le |f(x)| \nu(x)^{1/2}$, and therefore by the Cauchy-Schwarz inequality the right-hand side is bounded above by

$$\|f\|_2 \left(\sum_x \nu(x) \left| \sum_{\gamma \in \Delta} c_\gamma \overline{\gamma(x)} \right|^{2m} \right)^{1/2}.$$

Squaring and simplifying, we deduce that

$$\eta^{2m} |\Delta|^{2m} \frac{\|f\|_1^2}{\|f\|_2^2} \le \sum_x \nu(x) \left| \sum_{\gamma \in \Delta} c_\gamma \overline{\gamma(x)} \right|^{2m}.$$

Expanding out the power, the right-hand side is equal to

$$\sum_x \nu(x) \sum_{\gamma_1, \dots, \gamma_{2m}} c_{\gamma_1} \cdots \overline{c_{\gamma_{2m}}} (\overline{\gamma_1} \cdots \gamma_{2m})(x).$$

Changing the order of summation this is equal to

$$\sum_{\gamma_1, \dots, \gamma_{2m}} c_{\gamma_1} \cdots \overline{c_{\gamma_{2m}}} \widehat{\nu}(\gamma_1 \cdots \overline{\gamma_{2m}}).$$

The result follows by the triangle inequality. $\qquad \square$

Lemma 2.6. *Let G be a finite abelian group and $f : G \to \mathbb{C}$. Let $\Delta \subseteq \Delta_\eta(f)$. Then, for any $m \ge 1$.*

$$N^{-1} \eta^{2m} \frac{\|f\|_1^2}{\|f\|_2^2} |\Delta|^{2m} \le E_{2m}(\Delta).$$

Proof. Apply Lemma 2.5 with $\nu \equiv 1$, and use the fact that $\sum_x \lambda(x)$ is N if $\lambda \equiv 1$ and 0 otherwise. $\qquad \square$

Lemma 2.7. *If $A \subset G$ and $m \ge 1$ then*

$$E_{2m}(A) = \sum_x 1_A^{(m)}(x)^2.$$

Proof. Expand out definitions. $\qquad \square$

Lemma 2.8. *If $A \subseteq G$ is dissociated then $E_{2m}(A) \le m! \, |A|^m$.*

Proof. By Lemma 2.7

$$E_{2m}(A) = \sum_x 1_A^{(m)}(x)^2.$$

By the definition of dissociativity, $1_A^{(m)}(x) \le m!$ for all $x \in G$. We are done since

$$\sum_x 1_A^{(m)}(x) = |A|^m.$$

TODO(Thomas): This proof is wrong. $\qquad\square$

Lemma 2.9. *If $A \subseteq G$ contains no dissociated set with $\ge K + 1$ elements then there is $A' \subseteq A$ of size $|A'| \le K$ such that*

$$A \subseteq \left\{ \sum_{a \in A'} c_a a : c_a \in \{-1, 0, 1\} \right\}.$$

Proof. Let $A' \subseteq A$ be a maximal dissociated subset (this exists and is non-empty, since trivially any singleton is dissociated). We have $|A'| \le K$ by assumption.

Let S be the span on the right-hand side. It is obvious that $A' \subseteq S$. Suppose that $x \in A \backslash A'$. Then $A' \cup \{x\}$ is not dissociated by maximality. Therefore there exists some $y \in G$ and two distinct sets $B, C \subseteq A' \cup \{x\}$ such that

$$\sum_{b \in B} b = y = \sum_{c \in C} c.$$

If $x \notin B$ and $x \notin C$ then this contradicts the dissociativity of A'. If $x \in B$ and $x \in C$ then we have

$$\sum_{b \in B \backslash x} b = y - x = \sum_{c \in C \backslash x} c,$$

again contradicting the dissociativity of A'. Without loss of generality, therefore, $x \in B$ and $x \notin C$. Therefore

$$x = \sum_{c \in C} c - \sum_{b \in B \backslash x} b$$

which is in the span as required. $\qquad\square$

Theorem 2.10 (Chang's lemma). *Let G be a finite abelian group and $f : G \to \mathbb{C}$. Let $\eta > 0$ and $\alpha = N^{-1} \|f\|_1^2 / \|f\|_2^2$. There exists some $\Delta \subseteq \Delta_\eta(f)$ such that*

$$|\Delta| \le \lceil e \mathcal{L}(\alpha) \eta^{-2} \rceil$$

and

$$\Delta_\eta(f) \subseteq \left\{ \sum_{\gamma \in \Delta} c_\gamma \gamma : c_\gamma \in \{-1, 0, 1\} \right\}.$$

Proof. By Lemma 2.9 it suffices to show that $\Delta_\eta(f)$ contains no dissociated set with at least

$$K = \lceil e\mathcal{L}(\alpha)\eta^{-2}\rceil + 1$$

many elements. Suppose not, and let $\Delta \subseteq \Delta_\eta(f)$ be a dissociated set of size K. Then by Lemma 2.8 we have, for any $m \geq 1$,

$$E_{2m}(\Delta) \leq m! K^m.$$

On the other hand, by Lemma 2.6,

$$\eta^{2m}\alpha K^{2m} \leq E_{2m}(\Delta).$$

Rearranging these bounds, we have

$$K^m \leq m!\alpha^{-1}\eta^{-2m} \leq m^m\alpha^{-1}\eta^{-2m}.$$

Therefore $K \leq \alpha^{-1/m}m\eta^{-2}$. This is a contradiction to the choice of K if we choose $m = \mathcal{L}(\alpha)$, since $\alpha^{-1/m} \leq e$. $\qquad\square$

Chapter 3

Unbalancing

Lemma 3.1. *For any function $f : G \to \mathbb{R}$ and integer $k \geq 0$*

$$\mathbb{E}_x f \circ f(x)^k \geq 0.$$

Proof. Test. $\square$

Lemma 3.2. *Let $\epsilon \in (0,1)$ and $\nu : G \to \mathbb{R}_{\geq 0}$ be some probability measure such that $\hat{\nu} \geq 0$. Let $f : G \to \mathbb{R}$. If $\|f \circ f\|_{p(\nu)} \geq \epsilon$ for some $p \geq 1$ then $\|f \circ f + 1\|_{p'(\nu)} \geq 1 + \frac{1}{2}\epsilon$ for $p' = 120\epsilon^{-1}\log(3/\epsilon)$.*

Proof. Up to gaining a factor of 5 in p', we can assume that $p \geq 5$ is an odd integer. Since the Fourier transforms of f and ν are non-negative,

$$\mathbb{E}\nu f^p = \hat{\nu} * \hat{f}^{(p)}(0_{\widehat{G}}) \geq 0.$$

It follows that, since $2\max(x,0) = x + |x|$ for $x \in \mathbb{R}$,

$$2\langle \max(f,0), f^{p-1}\rangle_\nu = \mathbb{E}\nu f^p + \langle |f|, f^{p-1}\rangle_\nu \geq \|f\|_{p(\nu)}^p \geq \epsilon^p.$$

Therefore, if $P = \{x : f(x) \geq 0\}$, then $\langle 1_P, f^p\rangle_\nu \geq \frac{1}{2}\epsilon^p$. Furthermore, if $T = \{x \in P : f(x) \geq \frac{3}{4}\epsilon\}$ then $\langle 1_{P \setminus T}, f^p\rangle_\nu \leq \frac{1}{4}\epsilon^p$, and hence by the Cauchy-Schwarz inequality,

$$\nu(T)^{1/2}\|f\|_{2p(\nu)}^p \geq \langle 1_T, f^p\rangle_\nu \geq \frac{1}{4}\epsilon^p.$$

On the other hand, by the triangle inequality

$$\|f\|_{2p(\nu)} \leq 1 + \|f + 1\|_{2p(\nu)} \leq 3,$$

or else we are done, with $p' = 2p$. Hence $\nu(T) \geq (\epsilon/3)^{3p}$. It follows that, for any $p' \geq 1$,

$$\|f + 1\|_{p'(\nu)} \geq \langle 1_T, |f + 1|^{p'}\rangle_\nu^{1/p'} \geq (1 + \tfrac{3}{4}\epsilon)(\epsilon/3)^{3p/p'}.$$

The desired bound now follows if we choose $p' = 24\epsilon^{-1}\log(3/\epsilon)p$, using $1 - x \leq e^{-x}$. $\square$

Chapter 4

Dependent random choice

Lemma 4.1. *Let $p \geq 2$ be an even integer. Let $B_1, B_2 \subseteq G$ and $\mu = \mu_{B_1} \circ \mu_{B_2}$. For any finite set $A \subseteq G$ and function $f : G \to \mathbb{R}_{\geq 0}$ there exist $A_1 \subseteq B_1$ and $A_2 \subseteq B_2$ such that*

$$\langle \mu_{A_1} \circ \mu_{A_2}, f \rangle \| 1_A \circ 1_A \|_{p(\mu)}^p \leq 2 \langle (1_A \circ 1_A)^p, f \rangle_\mu$$

and

$$\min \left(\frac{|A_1|}{|B_1|}, \frac{|A_2|}{|B_2|} \right) \geq \frac{1}{4} |A|^{-2p} \| 1_A \circ 1_A \|_{p(\mu)}^{2p}.$$

Proof. First note that the statement is trivially true (with $A_1 = B_1$ and $A_2 = B_2$, say) if $\| 1_A \circ 1_A \|_{p(\mu)}^p = 0$. We can therefore assume this is $\neq 0$.

For $s \in G^p$ let $A_1(s) = B_1 \cap (A + s_1) \cap \cdots \cap (A + s_p)$, and similarly for $A_2(s)$. Note that

$$
\begin{aligned}
\langle (1_A \circ 1_A)^p, f \rangle_\mu &= \sum_x \mu_{B_1} \circ \mu_{B_2}(x)(1_A \circ 1_A(x))^p f(x) \\
&= \sum_{b_1, b_2} \mu_{B_1}(b_1) \mu_{B_2}(b_2) 1_A \circ 1_A (b_1 - b_2)^p f(b_1 - b_2) \\
&= \sum_{b_1, b_2} \mu_{B_1}(b_1) \mu_{B_2}(b_2) \left(\sum_{t \in G} 1_{A+t}(b_1) 1_{A+t}(b_2) \right)^p f(b_1 - b_2) \\
&= \sum_{b_1, b_2} \mu_{B_1}(b_1) \mu_{B_2}(b_2) \sum_{s \in G^p} 1_{A_1(s)}(b_1) 1_{A_2(s)}(b_2) f(b_1 - b_2) \\
&= |B_1|^{-1} |B_2|^{-1} \sum_{s \in G^p} \langle 1_{A_1(s)} \circ 1_{A_2(s)}, f \rangle.
\end{aligned}
$$

In particular, applying this with $f \equiv 1$ we see that

$$\| 1_A \circ 1_A \|_{p(\mu)}^p = |B_1|^{-1} |B_2|^{-1} \sum_s |A_1(s)| |A_2(s)|$$

and

$$\frac{\langle (1_A \circ 1_A)^p, f \rangle_\mu}{\| 1_A \circ 1_A \|_{p(\mu)}^p} = \frac{\sum_s \langle 1_{A_1(s)} \circ 1_{A_2(s)}, f \rangle}{\sum_s |A_1(s)| |A_2(s)|} = \eta,$$

say. Let $M > 0$ be some parameter, and let

$$g(s) = \begin{cases} 1 & \text{if } 0 < |A_1(s)| |A_2(s)| < M^2 \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

10

Then we have
$$\sum_s g(s) \left| A_1(s) \right| \left| A_2(s) \right| < \sum_s M \left| A_1(s) \right|^{1/2} \left| A_2(s) \right|^{1/2}.$$

To see why, note first that each summand on the left-hand side is $\leq$ the corresponding summand on the right-hand side, arguing by cases on whether $g(s) = 1$ or not. It therefore suffices to show that there exists some s such that the summand on the left-hand side is $<$ the corresponding summand on the right-hand side.

If $g(s) = 0$ for all s then choose some s such that $\left| A_1(s) \right| \left| A_2(s) \right| \geq M^2$ (this must exist or else $\left| A_1(s) \right| \left| A_2(s) \right| = 0$ for all s, but then $\left\| 1_A \circ 1_A \right\|_{p(\mu)}^p = 0$ by the above calculation). Otherwise, choose some s such that $g(s) = 1$, and note that for this s we have, by definition of s,
$$\left| A_1(s) \right| \left| A_2(s) \right| < M \left| A_1(s) \right|^{1/2} \left| A_2(s) \right|^{1/2}.$$

We now choose
$$M = \tfrac{1}{2} \left| A \right|^{-p} \left(\left| B_1 \right| \left| B_2 \right| \right)^{1/2} \left\| 1_A \circ 1_A \right\|_{p(\mu)}^p,$$

so that, by the Cauchy-Schwarz inequality,

$$\sum_s g(s) \left| A_1(s) \right| \left| A_2(s) \right| < M \sum_s \left| A_1(s) \right|^{1/2} \left| A_2(s) \right|^{1/2}$$

$$\leq M \left(\sum_s \sum_{x \in G} 1_{A_1(s)}(x) \right)^{1/2} \left(\sum_s \sum_{x \in G} 1_{A_2(s)}(x) \right)^{1/2}$$

$$= M \left| A \right|^p \left(\left| B_1 \right| \left| B_2 \right| \right)^{1/2}$$

$$= \frac{1}{2} \sum_s \left| A_1(s) \right| \left| A_2(s) \right|$$

and so
$$\sum_s (1 - g(s)) \left| A_1(s) \right| \left| A_2(s) \right| > \frac{1}{2} \sum_s \left| A_1(s) \right| \left| A_2(s) \right|$$

whence

$$\sum_s \langle 1_{A_1(s)} \circ 1_{A_2(s)}, f \rangle = \eta \sum \left| A_1(s) \right| \left| A_2(s) \right| < 2\eta \sum_s \left| A_1(s) \right| \left| A_2(s) \right| (1 - g(s)).$$

In particular there must exist some s such that

$$\langle 1_{A_1(s)} \circ 1_{A_2(s)}, f \rangle < 2\eta \left| A_1(s) \right| \left| A_2(s) \right| (1 - g(s)).$$

We claim this s meets the requirements. The first is satisfied since the right-hand side is $\leq 2\eta \left| A_1(s) \right| \left| A_2(s) \right|$. The second is satisfied since the left-hand side is trivially ≥ 0 and hence such an s must satisfy $g(s) = 0$, whence either $\left| A_1(s) \right| \left| A_2(s) \right| \geq M^2$, that is,

$$\left| A_1(s) \right| \left| A_2(s) \right| \geq \frac{1}{4} \left| A \right|^{-2p} \left| B_1 \right| \left| B_2 \right| \left\| 1_A \circ 1_A \right\|_{p(\mu)}^{2p},$$

or $\left| A_1(s) \right| \left| A_2(s) \right| = 0$, which can't happen because then the right-hand side is $= 0$.

The final bound now follows since $xy \leq \min(x, y)$ when $x, y \leq 1$. $\qquad \square$

Lemma 4.2. *Let $\epsilon, \delta > 0$ and $p \geq \max(2, \epsilon^{-1}\log(2/\delta))$ be an even integer. Let $B_1, B_2 \subseteq G$, and let $\mu = \mu_{B_1} \circ \mu_{B_2}$. For any finite set $A \subseteq G$, if*

$$S = \{x \in G : 1_A \circ 1_A(x) > (1-\epsilon)\|1_A \circ 1_A\|_{p(\mu)}\},$$

then there are $A_1 \subseteq B_1$ and $A_2 \subseteq B_2$ such that

$$\langle \mu_{A_1} \circ \mu_{A_2}, 1_S \rangle \geq 1 - \delta$$

and

$$\min\left(\frac{|A_1|}{|B_1|}, \frac{|A_2|}{|B_2|}\right) \geq \frac{1}{4}|A|^{-2p}\|1_A \circ 1_A\|_{p(\mu)}^{2p}.$$

Proof. Apply Lemma 4.1 with $f = 1_{G\setminus S}$. This produces some $A_1 \subseteq B_1$ and $A_2 \subseteq B_2$ such that

$$\langle \mu_{A_1} \circ \mu_{A_2}, 1_{G\setminus S} \rangle \leq 2\frac{\langle (1_A \circ 1_A)^p, 1_{G\setminus S} \rangle_\mu}{\|1_A \circ 1_A\|_{p(\mu)}^p}$$

and

$$\min\left(\frac{|A_1|}{|B_1|}, \frac{|A_2|}{|B_2|}\right) \geq \frac{1}{4}|A|^{-2p}\|1_A \circ 1_A\|_{p(\mu)}^{2p}.$$

It then suffices to note that

$$\langle \mu_{A_1} \circ \mu_{A_2}, 1_S \rangle = 1 - \langle \mu_{A_1} \circ \mu_{A_2}, 1_{G\setminus S} \rangle$$

and by definition of S we have

$$\langle (1_A \circ 1_A)^p, 1_{G\setminus S} \rangle_\mu \leq (1-\epsilon)^p\|1_A \circ 1_A\|_{p(\mu)}^p \sum_x \mu(x) = (1-\epsilon)^p\|1_A \circ 1_A\|_{p(\mu)}^p.$$

Now use the fact that $p \geq \epsilon^{-1}\log(2/\delta)$ together with the inequality $1 - x \leq e^{-x}$ to deduce that the right-hand side is $\leq \frac{\delta}{2}\|1_A \circ 1_A\|_{p(\mu)}^p$. $\qquad\square$

Corollary 4.3. *Let $\epsilon, \delta > 0$ and $p \geq \max(2, \epsilon^{-1}\log(2/\delta))$ be an even integer and $\mu \equiv 1/N$. If $A \subseteq G$ has density α and*

$$S = \{x : \mu_A \circ \mu_A(x) \geq (1-\epsilon)\|\mu_A \circ \mu_A\|_{p(\mu)}\}$$

then there are $A_1, A_2 \subseteq G$ such that

$$\langle \mu_{A_1} \circ \mu_{A_2}, 1_S \rangle \geq 1 - \delta$$

and both A_1 and A_2 have density

$$\geq \frac{1}{4}\alpha^{2p}.$$

Proof. We apply Lemma 4.2 with $B_1 = B_2 = G$. It remains to note that

$$\|1_A \circ 1_A\|_{p(\mu)} \geq \|1_A \circ 1_A\|_{1(\mu)} = |A|^2/N.$$

$\qquad\square$

Chapter 5

Finite field model

Theorem 5.1. *If $A_1, A_2, S \subseteq \mathbb{F}_q^n$ are such that A_1 and A_2 both have density at least α then there is a subspace V of codimension*

$$\operatorname{codim}(V) \leq 2^{27} \mathcal{L}(\alpha)^2 \mathcal{L}(\epsilon\alpha)^2 \epsilon^{-2}$$

such that

$$\left| \langle \mu_V * \mu_{A_1} * \mu_{A_2}, 1_S \rangle - \langle \mu_{A_1} * \mu_{A_2}, 1_S \rangle \right| \leq \epsilon.$$

Proof. (In this proof we write $G = \mathbb{F}_q^n$.) Let $k = \lceil \mathcal{L}(\epsilon\alpha/4) \rceil$. Note that $|A_1 + G| = |G| \leq \alpha^{-1}|A|$. Furthermore, $|A_2|/|S| \geq \alpha$. Therefore by Theorem 1.8 there exists some $T \subseteq G$ with

$$|T| \geq \exp(-2^{16} \mathcal{L}(\alpha)^2 k^2 \epsilon^{-2})|S|$$

such that

$$\| \mu_T^{(k)} * \mu_{A_1} * \mu_{A_2} \circ 1_S - \mu_{A_1} * \mu_{A_2} \circ 1_S \|_\infty \leq \epsilon/4.$$

Let $\Delta = \Delta_{1/2}(\mu_T)$ and

$$V = \{ x \in G : \gamma(x) = 1 \text{ for all } \gamma \in \Delta \}.$$

Note that

$$\langle \mu_V * \mu_{A_1} * \mu_{A_2}, 1_S \rangle = \langle \mu_V, \mu_{A_1} * \mu_{A_2} \circ 1_S \rangle = \frac{1}{|V|} \sum_{v \in V} \mu_{A_1} * \mu_{A_2} \circ 1_S(v)$$

and

$$\langle \mu_{A_1} * \mu_{A_2}, 1_S \rangle = \mu_{A_1} * \mu_{A_2} \circ 1_S(0).$$

Therefore

$$\left| \langle \mu_V * \mu_{A_1} * \mu_{A_2}, 1_S \rangle - \langle \mu_{A_1} * \mu_{A_2}, 1_S \rangle \right| \leq \frac{1}{|V|} \sum_{v \in V} \left| \mu_{A_1} * \mu_{A_2} \circ 1_S(v) - \mu_{A_1} * \mu_{A_2} \circ 1_S(0) \right|.$$

In particular it suffices to show that, for any $v \in V$,

$$\left| \mu_{A_1} * \mu_{A_2} \circ 1_S(v) - \mu_{A_1} * \mu_{A_2} \circ 1_S(0) \right| \leq \epsilon.$$

By the triangle inequality and construction of T, it suffices to show that

$$\left| \mu_T^{(k)} * \mu_{A_1} * \mu_{A_2} \circ 1_S(v) - \mu_T^{(k)} * \mu_{A_1} * \mu_{A_2} \circ 1_S(0) \right| \le \epsilon/2.$$

By the Fourier transform we have, for any $x \in G$,

$$\mu_T^{(k)} * \mu_{A_1} * \mu_{A_2} \circ 1_S(x) = \frac{1}{N} \sum_{\gamma \in \widehat{G}} \widehat{\mu_T}(\gamma)^k \widehat{\mu_{A_1}}(\gamma) \widehat{\mu_{A_2}}(\gamma) \widehat{1_{-S}}(\gamma) \gamma(x).$$

Therefore the left-hand side of the desired inequality is, by the triangle inequality, at most

$$\frac{1}{N} \sum_{\gamma \in \widehat{G}} |\widehat{\mu_T}(\gamma)|^k \left| \widehat{\mu_{A_1}}(\gamma) \widehat{\mu_{A_2}}(\gamma) \widehat{1_{-S}}(\gamma) \right| |\gamma(v) - 1|.$$

By choice of $v \in V$ the summand vanishes when $\gamma \in \Delta$. When $\gamma \notin \Delta$ the summand is bounded above by

$$2^{1-k} \left| \widehat{\mu_{A_1}}(\gamma) \widehat{\mu_{A_2}}(\gamma) \widehat{1_{-S}}(\gamma) \right|.$$

Therefore the left-hand side of the desired inequality is at most

$$2^{1-k} \frac{1}{N} \sum_{\gamma} \left| \widehat{\mu_{A_1}}(\gamma) \widehat{\mu_{A_2}}(\gamma) \widehat{1_{-S}}(\gamma) \right| \le 2^{1-k} |S| \frac{1}{N} \sum_{\gamma} \left| \widehat{\mu_{A_1}}(\gamma) \widehat{\mu_{A_2}}(\gamma) \right|$$

using the trivial bound $|\widehat{1_S}| \le |S|$. By the Cauchy-Schwarz inequality the sum on the right is at most

$$\left(\sum_{\gamma} \left| \widehat{\mu_{A_1}} \right|^2 \right)^{1/2} \left(\sum_{\gamma} \left| \widehat{\mu_{A_2}} \right|^2 \right)^{1/2}.$$

By Parseval's identity this is at most α^{-1}. Therefore the desired inequality follows from

$$2^{1-k} |S| \frac{1}{N} \alpha^{-1} \le 2^{1-k} \alpha^{-1} \le \epsilon/2.$$

It remains to check the codimension of V. For this, let $\Delta' \subseteq \Delta$ be as provided by Chang's lemma, Lemma 2.10, so that

$$\Delta \subseteq \left\{ \sum_{\gamma \in \Delta'} c_\gamma \gamma : c_\gamma \in \{-1, 0, 1\} \right\}.$$

Let

$$W = \{x \in G : \gamma(x) = 1 \text{ for all } \gamma \in \Delta'\}.$$

It follows that $W \le V$, so it suffices to bound the codimension of W. This we can bound trivially using the bound from Chang's lemma and the fact that $\mathcal{L}(\delta) = \log(e^2/\delta) \le 2 + \log(1/\delta) \le 4 \log(1/\delta)$, provided $\log(1/\delta) \ge 1$, so

$$|\Delta'| \le \lceil 4e \mathcal{L}(\delta) \rceil \le 2^7 \log(1/\delta),$$

where

$$\delta = |T|/N \ge \exp(-2^{16} \mathcal{L}(\alpha)^2 k^2 \epsilon^{-2}),$$

so

$$\operatorname{codim}(V) \le |\Delta'| \le 2^{23} \mathcal{L}(\alpha)^2 k^2 \epsilon^{-2} \le 2^{25} \mathcal{L}(\alpha)^2 \mathcal{L}(\epsilon\alpha/4)^2 \epsilon^{-2},$$

and now use $\mathcal{L}(\epsilon\alpha/4) \le 2\mathcal{L}(\epsilon\alpha)$, say. $\qquad\square$

Lemma 5.2. *For any function $f : G \to \mathbb{C}$ and integer $k \geq 1$*

$$\|f * f\|_{2k} \leq \|f \circ f\|_{2k}.$$

Proof. To finish, similar trick to unbalancing. $\qquad\square$

Lemma 5.3. *For any function f with $\sum f(x) = 1$*

$$f * f - 1/N = (f - 1/N) * (f - 1/N).$$

Proof. Expand everything out. $\qquad\square$

Lemma 5.4. *For any function f with $\sum f(x) = 1$*

$$f \circ f - 1/N = (f - 1/N) \circ (f - 1/N).$$

Proof. Expand everything out. $\qquad\square$

Lemma 5.5. *Let $\epsilon > 0$ and $\mu \equiv 1/N$. If $A, C \subseteq G$, where C has density at least γ, and*

$$|N\langle \mu_A * \mu_A, \mu_C \rangle - 1| > \epsilon$$

then, if $f = (\mu_A - 1/N)$, $\|f \circ f\|_{p(\mu)} \geq \epsilon/2N$ for $p = 2\lceil \mathcal{L}(\gamma) \rceil$.

Proof. By Hölder's inequality, for any $p \geq 1$

$$\epsilon < |N\langle \mu_A * \mu_A - 1/N, \mu_C \rangle| \leq \|\mu_A * \mu_A - 1/N\|_p \gamma^{-1/p} N^{1-1/p}.$$

In particular if we choose $p = 2\lceil \mathcal{L}(\gamma) \rceil$ then $\gamma^{-1/p} \leq e^{1/2} \leq 2$ and so we deduce that, by Lemma 5.3,

$$\|f * f\|_p \geq \tfrac{1}{2}\epsilon N^{1/p-1}.$$

It remains to use Lemmas 5.3 and 5.4 and apply Lemma 5.2, and note that we can pass from the L^p norm to the $L^p(\mu)$ norm losing a factor of $N^{1/p}$. $\qquad\square$

Proposition 5.6. *Let $\epsilon \in (0, 1)$. If $A, C \subseteq \mathbb{F}_q^n$, where C has density at least γ, and*

$$|N\langle \mu_A * \mu_A, \mu_C \rangle - 1| > \epsilon$$

then there is a subspace V of codimension

$$\leq 2^{171} \epsilon^{-24} \mathcal{L}(\alpha)^4 \mathcal{L}(\gamma)^4.$$

such that $\|1_A * \mu_V\|_\infty \geq (1 + \epsilon/32)\alpha$.

Proof. By Lemma 5.5, if $f = \mu_A - 1/N$,

$$\|f \circ f\|_{p(\mu)} \geq \epsilon/2N,$$

where $p = 2\lceil \mathcal{L}(\gamma) \rceil \leq 4\mathcal{L}(\gamma)$. By Lemma 3.2 there exists some p' such that

$$p' \leq 128\epsilon^{-1} \log(96/\epsilon)\mathcal{L}(\gamma)$$

such that

$$\|f \circ f + 1/N\|_{p'(\mu)} \geq (1 + \epsilon/4)/N.$$

By Lemma 5.4 $f \circ f + 1/N = \mu_A \circ \mu_A$.

Let $q = 2\lceil p' + 2^8 \epsilon^{-2} \log(64/\epsilon) \rceil$. By Corollary 4.3, there are A_1, A_2, both of density $\geq \alpha^{2q}$ such that

$$\langle \mu_{A_1} \circ \mu_{A_2}, 1_S \rangle \geq 1 - \epsilon/32$$

where

$$S = \{x : \mu_A \circ \mu_A(x) \geq (1 - \epsilon/16)\|\mu_A \circ \mu_A\|_{q(\mu)}\}.$$

Since

$$\|\mu_A \circ \mu_A\|_{q(\mu)} \geq \|\mu_A \circ \mu_A\|_{p'(\mu)} \geq (1 + \epsilon/4)/N$$

we know

$$S \subseteq S' = \{x : \mu_A \circ \mu_A(x) \geq (1 + \epsilon/8)/N\}.$$

By Theorem 5.1 (applied with ϵ replaced by $\epsilon/32$) there is a subspace V of codimension

$$\leq 2^{37} \mathcal{L}(\alpha^{2q})^2 \mathcal{L}(\epsilon \alpha^{2q}/32)^2 \epsilon^{-2}$$

such that

$$\langle \mu_V * \mu_{A_1} \circ \mu_{A_2}, 1_{S'} \rangle \geq 1 - \tfrac{1}{16}\epsilon.$$

Using $\mathcal{L}(xy) \leq x^{-1}\mathcal{L}(y)$ we have

$$\mathcal{L}(\epsilon \alpha^{2q}/32) \leq 32\epsilon^{-1}\mathcal{L}(\alpha^{2q}),$$

and we also use $\mathcal{L}(x^y) \leq y\mathcal{L}(x)$ to simplify the codimension bound to

$$\leq 2^{51} q^4 \mathcal{L}(\alpha)^4 \epsilon^{-4}.$$

We further note that (using $\log x \leq x$ say)

$$q \leq 2^{10} p' \epsilon^{-2} \log(64/\epsilon) \leq 2^{30} \epsilon^{-5} \mathcal{L}(\gamma).$$

Therefore the desired codimension bound follows. Finally, by definition of S', it follows that

$$
\begin{aligned}
(1 + \epsilon/32)/N &\leq ((1 + \epsilon/8)/N)(1 - \epsilon/16) \\
&\leq \langle \mu_V * \mu_{A_1} \circ \mu_{A_2}, \mu_A \circ \mu_A \rangle \\
&\leq \|\mu_V * \mu_A\|_\infty \|\mu_A * \mu_{A_2} \circ \mu_{A_1}\|_1 \\
&= \|\mu_V * 1_A\|_\infty |A|^{-1},
\end{aligned}
$$

and the proof is complete. $\qquad\square$

Lemma 5.7. *If $A \subseteq G$ has no non-trivial three-term arithmetic progressions and G has odd order then*

$$\langle \mu_A * \mu_A, \mu_{2 \cdot A} \rangle = 1/|A|^2.$$

Proof. Expand out using definitions. $\qquad\square$

Theorem 5.8. *Let q be an odd prime power. If $A \subseteq \mathbb{F}_q^n$ with $\alpha = |A|/q^n$ has no non-trivial three-term arithmetic progressions then*

$$n \ll \mathcal{L}(\alpha)^9.$$

Proof. Let $t \geq 0$ be maximal such that there is a sequence of subspaces $\mathbb{F}_q^n = V_0 \geq \cdots \geq V_t$ and associated $A_i \subseteq V_i$ with $A_0 = A$ such that

16

1. for $0 \leq i \leq t$ there exists x_i such that $A_i \subseteq A - x_i$,

2. with $\alpha_i = |A_i| \, / \, |V_i|$ we have

$$\alpha_{i+1} \geq \frac{65}{64}\alpha_i$$

   for $0 \leq i < t$, and

3.
$$\operatorname{codim}(V_{i+1}) \leq \operatorname{codim}(V_i) + O(\mathcal{L}(\alpha)^8)$$

   for $0 \leq i < t$. (here the second summand should be replaced with whatever explicit codimension bound we get from the above).

Note this is well-defined since $t = 0$ meets the requirements, and this process is finite and $t \ll \mathcal{L}(\alpha)$, since $\alpha_i \leq 1$ for all i. Therefore

$$\operatorname{codim}(V_t) \ll \mathcal{L}(\alpha)^9.$$

Suppose first that

$$|V_t|\langle \mu_{A_t} * \mu_{A_t}, \mu_{2 \cdot A_t}\rangle < 1/2.$$

In this case we now apply Proposition 5.6 to $A_t \subseteq V_t$ with $\epsilon = 1/2$ (note that $N = |V_t|$ and all inner product, μ etc, are relative to the ambient group V_t now). Therefore there is a subspace $V \leq V_t$ of codimension (relative to V_t) of $\ll \mathcal{L}(\alpha)^8$ such that there exists some $x \in V_t$ with

$$\frac{|(A_t - x) \cap V|}{|V|} = 1_{A_t} * \mu_V(x) = \|1_{A_t} * \mu_V\|_\infty \geq (1 + 1/64)\alpha_t,$$

which contradicts the maximality of t, letting $V_{t+1} = V$ and $A_{t+1} = (A_t - x) \cap V_t$.

Therefore

$$|V_t|\langle \mu_{A_t} * \mu_{A_t}, \mu_{2 \cdot A_t}\rangle \geq 1/2.$$

By Lemma 5.7 the left-hand side is equal to $|V_t|/|A_t|^2$, and therefore

$$\alpha^2 \leq \alpha_t^2 \leq 2/|V_t|.$$

By the codimension bound the right-hand side is at most

$$2q^{O(\mathcal{L}(\alpha)^9)-n}.$$

If $\alpha^2 \leq 2q^{-n/2}$ we are done, otherwise we deduce that $\mathcal{L}(\alpha)^9 \gg n$ as required. $\qquad\square$