Special Topics in Complexity Theory, Fall 2017. Instructor: Emanuele Viola

1 Lectures 16-17, Scribe: Tanay Mehta

In these lectures we prove the corners theorem for pseudorandom groups, following Austin [Aus16]. Our exposition has several non-major differences with that in [Aus16], which may make it more computer-science friendly. The instructor suspects a proof can also be obtained via certain local modifications and simplifications of Green's exposition [Gre05b, Gre05a] of an earlier proof for the abelian case. We focus on the case $G = SL_2(q)$ for simplicity, but the proof immediately extends to other pseudorandom groups.

Theorem 1. Let $G = SL_2(q)$. Every subset $A \subseteq G^2$ of density $\mu(A) \geq 1/\log^a |G|$ contains a corner, i.e., a set of the form $\{(x, y), (xz, y), (x, zy) \mid z \neq 1\}$.

1.1 Proof Overview

For intuition, suppose A is a product set, i.e., $A = B \times C$ for $B, C \subseteq G$. Let's look at the quantity

$$\mathbb{E}_{x,y,z \leftarrow G}[A(x, y)A(xz, y)A(x, zy)]$$

where $A(x, y) = 1$ iff $(x, y) \in A$. Note that the random variable in the expectation is equal to 1 exactly when x, y, z form a corner in A. We'll show that this quantity is greater than $1/|G|$, which implies that A contains a corner (where $z \neq 1$). Since we are taking $A = B \times C$, we can rewrite the above quantity as

$$\begin{aligned}
&\mathbb{E}_{x,y,z \leftarrow G}[B(x)C(y)B(xz)C(y)B(x)C(zy)]\\
&= \mathbb{E}_{x,y,z \leftarrow G}[B(x)C(y)B(xz)C(zy)]\\
&= \mathbb{E}_{x,y,z \leftarrow G}[B(x)C(y)B(z)C(x^{-1}zy)]
\end{aligned}$$

where the last line follows by replacing z with $x^{-1}z$ in the uniform distribution. If $\mu(A) \geq \delta$, then $\mu(B) \geq \delta$ and $\mu(C) \geq \delta$. Condition on $x \in B$, $y \in C$, $z \in B$. Then the distribution $x^{-1}zy$ is a product of three independent distributions, each uniform on a set of measure greater than δ. By pseudorandomness $x^{-1}zy$ is $1/|G|^{\Omega(1)}$ close to uniform in statistical distance. This implies that the above quantity equals

$$\begin{aligned}
&\mu(A) \cdot \mu(C) \cdot \mu(B) \cdot \left(\mu(C) \pm \frac{1}{|G|^{\Omega(1)}}\right) \\
&\geq \delta^3 \left(\delta - \frac{1}{|G|^{\Omega(1)}}\right) \\
&\geq \delta^4/2 \\
&> 1/|G|.
\end{aligned}$$

Given this, it is natural to try to write an arbitrary A as a combination of product sets (with some error). We will make use of a more general result.

1.2 Weak Regularity Lemma

Let U be some universe (we will take $U = G^2$). Let $f : U \to [-1, 1]$ be a function (for us, $f = 1_A$). Let $D \subseteq \{d : U \to [-1, 1]\}$ be some set of functions, which can be thought of as "easy functions" or "distinguishers."

Theorem 2.[Weak Regularity Lemma] For all $\epsilon > 0$, there exists a function $g := \sum_{i \leq s} c_i \cdot d_i$ where $d_i \in D$, $c_i \in \mathbb{R}$ and $s = 1/\epsilon^2$ such that for all $d \in D$

$$\mathbb{E}_{x \leftarrow U}[f(x) \cdot d(x)] = \mathbb{E}_{x \leftarrow U}[g(x) \cdot d(x)] \pm \epsilon.$$

The lemma is called 'weak' because it came after Szemerédi's regularity lemma, which has a stronger distinguishing conclusion. However, the lemma is also 'strong' in the sense that Szemerédi's regularity lemma has s as a tower of $1/\epsilon$ whereas here we have s polynomial in $1/\epsilon$. The weak regularity lemma is also simpler. There also exists a proof of Szemerédi's theorem (on arithmetic progressions), which uses weak regularity as opposed to the full regularity lemma used initially.

Proof. We will construct the approximation g through an iterative process producing functions $g_0, g_1, \ldots, g$. We will show that $||f - g_i||_2^2$ decreases by $\geq \epsilon^2$ each iteration.

1. **Start**: Define $g_0 = 0$ (which can be realized setting $c_0 = 0$).
2. **Iterate**: If not done, there exists $d \in D$ such that $|\mathbb{E}[(f - g) \cdot d]| > \epsilon$. Assume without loss of generality $\mathbb{E}[(f - g) \cdot d] > \epsilon$.

3. **Update**: $g' := g + \lambda d$ where $\lambda \in \mathbb{R}$ shall be picked later.

Let us analyze the progress made by the algorithm.

$$\begin{aligned}
||f - g'||_2^2 &= \mathbb{E}_x[(f - g')^2(x)] \\
&= \mathbb{E}_x[(f - g - \lambda d)^2(x)] \\
&= \mathbb{E}_x[(f - g)^2] + \mathbb{E}_x[\lambda^2 d^2(x)] - 2\mathbb{E}_x[(f - g) \cdot \lambda d(x)] \\
&\leq ||f - g||_2^2 + \lambda^2 - 2\lambda\mathbb{E}_x[(f - g)d(x)] \\
&\leq ||f - g||_2^2 + \lambda^2 - 2\lambda\epsilon \\
&\leq ||f - g||_2^2 - \epsilon^2
\end{aligned}$$

where the last line follows by taking $\lambda = \epsilon$. Therefore, there can only be $1/\epsilon^2$ iterations because $||f - g_0||_2^2 = ||f||_2^2 \leq 1$. □

1.3 Getting more for rectangles

Returning to the lower bound proof, we will use the weak regularity lemma to approximate the indicator function for arbitrary A by rectangles. That is, we take D to be the collection of indicator functions for all sets of the form $S \times T$ for $S, T \subseteq G$. The weak regularity lemma gives us A as a linear combination of rectangles. These rectangles may overlap. However, we ideally want A to be a linear combination of *non-overlapping* rectangles.

Claim 3. Given a decomposition of A into rectangles from the weak regularity lemma with s functions, there exists a decomposition with $2^{O(s)}$ rectangles which don't overlap.

Proof. Exercise. □

In the above decomposition, note that it is natural to take the coefficients of rectangles to be the density of points in A that are in the rectangle. This gives rise to the following claim.

Claim 4. The weights of the rectangles in the above claim can be the average of f in the rectangle, at the cost of doubling the distinguisher error.

Consequently, we have that $f = g + h$, where g is the sum of $2^{O(s)}$ non-overlapping rectangles $S \times T$ with coefficients $\Pr_{(x,y) \in S \times T}[f(x, y) = 1]$.

Proof. Let g be a partition decomposition with arbitrary weights. Let g' be a partition decomposition with weights being the average of f. It is enough to show that for all rectangle distinguishers $d \in D$

$$|\mathbb{E}[(f - g')d]| \leq |\mathbb{E}[(f - g)d]|.$$

By the triangle inequality, we have that

$$|\mathbb{E}[(f - g')d]| \leq |\mathbb{E}[(f - g)d]| + |\mathbb{E}[(g - g')d]|.$$

To bound $\mathbb{E}[(g - g')d]|$, note that the error is maximized for a d that respects the decomposition in non-overlapping rectangles, i.e., d is the union of some non-overlapping rectangles from the decomposition. This can be argues using that, unlike f, the value of g and g' on a rectangle $S \times T$ from the decomposition is fixed. But, for such d, $g' = f$! More formally, $\mathbb{E}[(g - g')d] = \mathbb{E}[(g - f)d]$. $\square$

We need to get a little more from this decomposition. The conclusion of the regularity lemma holds with respect to distinguishers that can be written as $U(x) \cdot V(y)$ where U and V map $G \to \{0, 1\}$. We need the same guarantee for U and V with range $[-1, 1]$. This can be accomplished paying only a constant factor in the error, as follows. Let U and V have range $[-1, 1]$. Write $U = U_+ - U_-$ where U_+ and U_- have range $[0, 1]$, and the same for V. The error for distinguisher $U \cdot V$ is at most the sum of the errors for distinguishers $U_+ \cdot V_+$, $U_+ \cdot V_-$, $U_- \cdot V_+$, and $U_- \cdot V_-$. So we can restrict our attention to distinguishers $U(x) \cdot V(y)$ where U and V have range $[0, 1]$. In turn, a function $U(x)$ with range $[0, 1]$ can be written as an expectation $\mathbb{E}_a U_a(x)$ for functions U_a with range $\{0, 1\}$, and the same for V. We conclude by observing that

$$\mathbb{E}_{x,y}[(f - g)(x, y)\mathbb{E}_a U_a(x) \cdot \mathbb{E}_b V_b(y)] \leq \max_{a,b} \mathbb{E}_{x,y}[(f - g)(x, y)U_a(x) \cdot V_b(y)].$$

1.4 Proof

Let us now finish the proof by showing a corner exists for sufficiently dense sets $A \subseteq G^2$. We'll use three types of decompositions for $f : G^2 \to \{0, 1\}$, with respect to the following three types of distinguishers, where U_i and V_i have range $\{0, 1\}$:

1. $U_1(x) \cdot V_1(y)$,

2. $U_2(xy) \cdot V_2(y)$,

3. $U_3(x) \cdot V_3(xy)$.

The last two distinguishers can be visualized as parallelograms with a 45-degree angle between two segments. The same extra properties we discussed for rectangles hold for them too.

Recall that we want to show

$$\mathbb{E}_{x,y,g}[f(x,y)f(xg,y)f(x,gy)] > \frac{1}{|G|}.$$

We'll decompose the i-th occurrence of f via the i-th decomposition listed above. We'll write this decomposition as $f = g_i + h_i$. We do this in the following order:

$$\begin{aligned} & f(x,y) \cdot f(xg,y) \cdot f(x,gy) \\ = & f(x,y)f(xg,y)g_3(x,gy) + f(x,y)f(xg,y)h_3(x,gy) \\ & \vdots \\ = & g_1g_2g_3 + h_1g_2g_3 + fh_2g_3 + ffh_3 \end{aligned}$$

We first show that $\mathbb{E}[g_1g_2g_3]$ is big (i.e., inverse polylogarithmic in expectation) in the next two claims. Then we show that the expectations of the other terms are small.

Claim 5. For all $g \in G$, the values $\mathbb{E}_{x,y}[g_1(x,y)g_2(xg,y)g_3(x,gy)]$ are the same (over g) up to an error of $2^{O(s)} \cdot 1/|G|^{\Omega(1)}$.

Proof. We just need to get error $1/|G|^{\Omega(1)}$ for any product of three functions for the three decomposition types. By the standard pseudorandomness argument we saw in previous lectures,

$$\begin{aligned} & \mathbb{E}_{x,y}[c_1U_1(x)V_1(y) \cdot c_2U_2(xgy)V_2(y) \cdot c_3U_3(x)V_3(xgy)] \\ & = c_1c_2c_3\mathbb{E}_{x,y}[(U_1 \cdot U_3)(x)(V_1 \cdot V_2)(y)(U_2 \cdot V_3)(xgy)] \\ & = c_1c_2c_3 \cdot \mu(U_1 \cdot U_3)\mu(V_1 \cdot V_2)\mu(U_2 \cdot V_3) \pm \frac{1}{|G|^{\Omega(1)}}. \end{aligned}$$

□

Recall that we start with a set of density $\geq 1/\log^a |G|$.

Claim 6. $\mathbb{E}_{g,x,y}[g_1 g_2 g_3] > \Omega(1/\log^{4a} |G|)$.

Proof. By the previous claim, we can fix $g = 1_G$. We will relate the expectation over x, y to f by a trick using the Hölder inequality: For random variables $X_1, X_2, \ldots, X_k$,

$$\mathbb{E}[X_1 \ldots X_k] \leq \prod_{i=1}^{k} \mathbb{E}[X_i^{c_i}]^{1/c_i} \text{ such that } \sum 1/c_i = 1.$$

To apply this inequality in our setting, write

$$\mathbb{E}[f] = \mathbb{E}\left[(f \cdot g_1 g_2 g_3)^{1/4} \cdot \left(\frac{f}{g_1}\right)^{1/4} \cdot \left(\frac{f}{g_2}\right)^{1/4} \cdot \left(\frac{f}{g_3}\right)^{1/4}\right].$$

By the Hölder inequality, we get that

$$\mathbb{E}[f] \leq \mathbb{E}[f \cdot g_1 g_2 g_3]^{1/4} \mathbb{E}\left[\frac{f}{g_1}\right]^{1/4} \mathbb{E}\left[\frac{f}{g_2}\right]^{1/4} \mathbb{E}\left[\frac{f}{g_3}\right]^{1/4}.$$

Note that

$$\begin{aligned}
\mathbb{E}_{x,y} \frac{f(x,y)}{g_1(x,y)} &= \mathbb{E}_{x,y} \frac{f(x,y)}{\mathbb{E}_{x',y' \in Cell(x,y)}[f(x',y')]} \\
&= \mathbb{E}_{x,y} \frac{\mathbb{E}_{x',y' \in Cell(x,y)}[f(x',y')]}{\mathbb{E}_{x',y' \in Cell(x,y)}[f(x',y')]} \\
&= 1
\end{aligned}$$

where $Cell(x, y)$ is the set in the partition that contains (x, y). Finally, by non-negativity of f, we have that $\mathbb{E}[f \cdot g_1 g_2 g_3]^{1/4} \leq \mathbb{E}[g_1 g_2 g_3]$. This concludes the proof. □

We've shown that the $g_1 g_2 g_3$ term is big. It remains to show the other terms are small. Let ϵ be the error in the weak regularity lemma with respect to distinguishers with range $[-1, 1]$.

Claim 7. $|\mathbb{E}[f f h_3]| \leq \epsilon^{1/4}$.

Proof. Replace g with gy^{-1} in the uniform distribution to get

$$\begin{aligned}
&\mathbb{E}^4_{x,y,g}[f(x,y)f(xg,y)h_3(x,gy)] \\
&= \mathbb{E}^4_{x,y,g}[f(x,y)f(xgy^{-1},y)h_3(x,g)] \\
&= \mathbb{E}^4_{x,y}[f(x,y)\mathbb{E}_g[f(xgy^{-1},y)h_3(x,g)]] \\
&\leq \mathbb{E}^2_{x,y}[f^2(x,y)]\mathbb{E}^2_{x,y}\mathbb{E}^2_g[f(xgy^{-1},y)h_3(x,g)] \\
&\leq \mathbb{E}^2_{x,y}\mathbb{E}^2_g[f(xgy^{-1},y)h_3(x,g)] \\
&= \mathbb{E}^2_{x,y,g,g'}[f(xgy^{-1},y)h_3(x,g)f(xg'y^{-1},y)h_3(x,g')],
\end{aligned}$$

where the first inequality is by Cauchy-Schwarz.

Now replace $g \to x^{-1}g, g' \to x^{-1}g$ and reason in the same way:

$$\begin{aligned}
&= \mathbb{E}^2_{x,y,g,g'}[f(gy^{-1},y)h_3(x,x^{-1}g)f(g'y^{-1},y)h_3(x,x^{-1}g')] \\
&= \mathbb{E}^2_{g,g',y}[f(gy^{-1},y)\cdot f(g'y^{-1},y)\mathbb{E}_x[h_3(x,x^{-1}g)\cdot h_3(x,x^{-1}g')]] \\
&\leq \mathbb{E}_{x,x',g,g'}[h_3(x,x^{-1}g)h_3(x,x^{-1}g')h_3(x',x'^{-1}g)h_3(x',x'^{-1}g')].
\end{aligned}$$

Replace $g \to xg$ to rewrite the expectation as

$$\mathbb{E}[h_3(x,g)h_3(x,x^{-1}g')h_3(x',x'^{-1}xg)h_3(x',x'^{-1}g')].$$

We want to view the last three terms as a distinguisher $U(x)\cdot V(xg)$. First, note that h_3 has range $[-1,1]$. This is because $h_3(x,y) = f(x,y) - \mathbb{E}_{x',y'\in Cell(x,y)}f(x',y')$ and f has range $\{0,1\}$.

Fix x', g'. The last term in the expectation becomes a constant $c \in [-1,1]$. The second term only depends on x, and the third only on xg. Hence for appropriate functions U and V with range $[-1,1]$ this expectation can be rewritten as

$$\mathbb{E}[h_3(x,g)U(x)V(xg)],$$

which concludes the proof. □

There are similar proofs to show the remaining terms are small. For fh_2g_3, we can perform simple manipulations and then reduce to the above case. For $h_1g_2g_3$, we have a slightly easier proof than above.

1.4.1 Parameters

Suppose our set has density $\delta \geq 1/\log^a |G|$. We apply the weak regularity lemma for error $\epsilon = 1/\log^c |G|$. This yields the number of functions $s = 2^{O(1/\epsilon^2)} = 2^{O(\log^{2c}|G|)}$. For say $c = 1/3$, we can bound $\mathbb{E}_{x,y,g}[g_1 g_2 g_3]$ from below by the same expectation with g fixed to 1, up to an error $1/|G|^{\Omega(1)}$. Then, $\mathbb{E}_{x,y,g=1}[g_1 g_2 g_3] \geq \mathbb{E}[f]^4 = 1/\log^{4a}|G|$. The expectation of terms with h is less than $1/\log^{c/4}|G|$. So the proof can be completed for all sufficiently small a.

References

[Aus16] Tim Austin. Ajtai-Szemerédi theorems over quasirandom groups. In *Recent trends in combinatorics*, volume 159 of *IMA Vol. Math. Appl.*, pages 453–484. Springer, [Cham], 2016.

[Gre05a] Ben Green. An argument of Shkredov in the finite field setting, 2005. Available at people.maths.ox.ac.uk/greenbj/papers/corners.pdf.

[Gre05b] Ben Green. Finite field models in additive combinatorics. *Surveys in Combinatorics, London Math. Soc. Lecture Notes 327, 1-27*, 2005.