Graph Theory and Additive Combinatorics

Lecturer: Prof. Yufei Zhao

4
Pseudorandom graphs

The term "pseudorandom" refers to a wide range of ideas and phenomenon where non-random objects behave in certain ways like genuinely random objects. For example, while the prime numbers are not random, their distribution among the integers have many properties that resemble random sets. The famous Riemann Hypothesis is a notable conjecture about the pseudorandomness of the primes in a certain sense.

When used more precisely, we can ask whether some given objects behaves in some *specific* way similar to a typical random object? In this chapter, we examine such questions for graphs, and study ways that a non-random graph can have properties that resemble a typical random graph.

4.1 Quasirandom graphs

The next theorem is a foundational result in the subject. It lists several seemingly different pseudorandomness properties that a graph can have (with some seemingly easier to verify than others), and asserts, somewhat surprisingly, that these properties turn out to be all equivalent to each other.

Theorem 4.1. *Let $\{G_n\}$ be a sequence of graphs with G_n having n vertices and $(p+o(1))\binom{n}{2}$ edges, for fixed $0 < p < 1$. Denote G_n by G. The following properties are equivalent:*

1. ***DISC** ("discrepancy"): $|e(X,Y) - p|X||Y|| = o(n^2)$ for all $X, Y \subset V(G)$.*
2. ***DISC'**: $|e(X) - p\binom{|X|}{2}| = o(n^2)$ for all $X \subset V(G)$.*
3. ***COUNT**: For all graphs H, the number of **labeled copies** of H in G (i.e. vertices in H are distinguished) is $(p^{e(H)} + o(1))n^{v(H)}$. The $o(1)$ term may depend on H.*

Chung, Graham, and Wilson (1989)

Theorem 4.1 should be understood as a theorem about *dense graphs*, i.e., graphs with constant order edge density. Sparser graphs can have very different behavior and will be discussed in later sections.

4. ***C4****: The number of labeled copies of C_4 is at most $(p^4 + o(1))n^4$.*

5. ***CODEG*** *(codegree): If $\mathrm{codeg}(u, v)$ is the number of common neighbors of u and v, then $\sum_{u,v \in V(G)} |\mathrm{codeg}(u, v) - p^2 n| = o(n^3)$.*

6. ***EIG*** *(eigenvalue): If $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{v(G)}$ are the eigenvalues of the adjacency matrix of G, then $\lambda_1 = pn + o(n)$ and $\max_{i \neq 1} |\lambda_i| = o(n)$.*

Remark 4.2. In particular, for a d-regular graph, the largest eigenvalue is d, with corresponding eigenvector the all-1 vector, and **EIG** states that $\lambda_2, \lambda_{v(G)} = o(n)$.

We can equivalently state the conditions in the theorem in terms of some ϵ: for instance, **DISC** can be reformulated as

DISC(ϵ): For all $X, Y \subset V(G)$, $|e(X, Y) - p|X||Y|| < \epsilon n^2$.

Then we will see from the proof of Theorem 4.1 that the conditions in the theorem are equivalent up to at most polynomial change in ϵ, i.e. **Prop1**$(\epsilon) \implies$ **Prop2**(ϵ^c) for some c.

Since we will use the Cauchy–Schwarz inequality many times in this proof, let's begin with an exercise.

Lemma 4.3. *If G is a graph with n vertices, $e(G) \geq pn^2/2$, then the number of labeled copies of C_4 is $\geq (p^4 - o(1))n^4$.*

Proof. We want to count the size of $S = \mathrm{Hom}(C_4, G)$, the set of graph homomorphisms from C_4 to G. We also include in S some non-injective maps, i.e. where points in C_4 may map to the same point in G, since there are only $O(n^3)$ of them anyway. It is equal to $\sum_{u,v \in V(G)} \mathrm{codeg}(u, v)^2$, by considering reflections across a diagonal of C_4. Using Cauchy–Schwarz twice, we have

$$\begin{aligned}
|\mathrm{Hom}(C_4, G)| &= \sum_{u,v \in V(G)} \mathrm{codeg}(u, v)^2 \\
&\geq \frac{1}{n^2} \left(\sum_{u,v \in V(G)} \mathrm{codeg}(u, v) \right)^2 \\
&= \frac{1}{n^2} \left(\sum_{x \in G} \deg(x)^2 \right)^2 \\
&\geq \frac{1}{n^2} \left(\frac{1}{n} \left(\sum_{x \in G} \deg(x) \right)^2 \right)^2 \\
&= \frac{1}{n^2} \left(\frac{1}{n} (pn^2)^2 \right)^2 \\
&= p^4 n^4
\end{aligned}$$

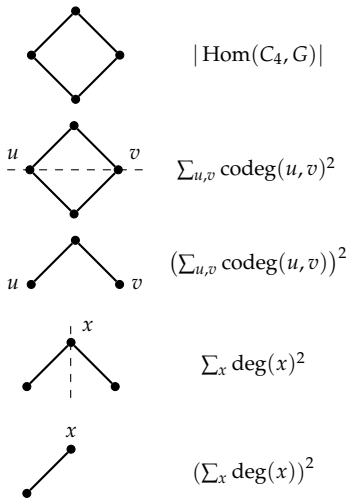


Figure 4.1: Visualization of Cauchy–Schwarz

where in the second line we have $\sum_{u,v\in V(G)} \operatorname{codeg}(u,v) = \sum_{x\in G} \deg(x)^2$ by counting the number of paths of length 2 in two ways. □

Remark 4.4. We can keep track of our Cauchy–Schwarz manipulations with a "visual anchor": see Figure 4.1. We see that Cauchy–Schwarz bounds exploit symmetries in the graph.

Now we prove the theorem.

Proof. **DISC** $\implies$ **DISC′**: Take $Y = X$ in **DISC**.

DISC′ $\implies$ **DISC**: By categorizing the types of edges counted in $e(X,Y)$ (see Figure 4.2), we can write $e(X,Y)$ in terms of the edge counts of individual vertex sets:

$$e(X,Y) = e(X\cup Y) + e(X\cap Y) - e(X\setminus Y) - e(Y\setminus X).$$

Then we can use **DISC′** to get that this is

$$\begin{aligned} &p\left(\binom{|X\cup Y|}{2} + \binom{|X\cap Y|}{2} + \binom{|X\setminus Y|}{2} + \binom{|Y\setminus X|}{2} + o(n^2)\right)\\ =&p|X||Y| + o(n^2). \end{aligned}$$

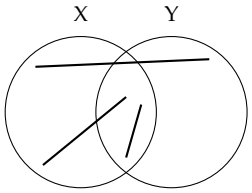


Figure 4.2: Visualization of the expression for $e(X,Y)$

DISC $\implies$ **COUNT**: This follows from the graph counting lemma (Theorem 3.27), taking $V_i = G$ for $i = 1,\dots,v(H)$.

COUNT $\implies$ **C4**: **C4** is just a special case of **COUNT**.

C4 $\implies$ **CODEG**: Given **C4**, we have

$$\sum_{u,v\in G} \operatorname{codeg}(u,v) = \sum_{x\in G} \deg(x)^2 \ge n\left(\frac{2e(G)}{n}\right)^2 = \left(p^2+o(1)\right)n^3.$$

We also have

$$\begin{aligned} \sum_{u,v} \operatorname{codeg}(u,v)^2 &= \text{Number of labeled copies of } C_4 + o(n^4)\\ &\le \left(p^4+o(1)\right)n^4. \end{aligned}$$

Therefore, we can use Cauchy–Schwarz to find

$$\begin{aligned} \sum_{u,v\in G} |\operatorname{codeg}(u,v) - p^2n| &\le n\left(\sum_{u,v\in G}\left(\operatorname{codeg}(u,v)-p^2n\right)^2\right)^{1/2}\\ &= n\left(\sum_{u,v\in G}\operatorname{codeg}(u,v)^2 - 2p^2n\sum_{u,v\in G}\operatorname{codeg}(u,v) + p^4n^4\right)^{1/2}\\ &\le n\left(p^4n^4 - 2p^2n\cdot p^2n^3 + p^4n^4 + o(n^4)\right)^{1/2}\\ &= o(n^3), \end{aligned}$$

as desired.

Remark 4.5. This technique is similar to the ***second moment method*** in probabilistic combinatorics: we want to show that the variance of $\operatorname{codeg}(u, v)$ is not too large.

CODEG $\implies$ **DISC**: First, note that we have

$$\begin{aligned}
\sum_{u \in G} |\deg u - pn| &\le n^{1/2} \left(\sum_{u \in G} (\deg u - pn)^2 \right)^{1/2} \\
&= n^{1/2} \left(\sum_{u \in G} (\deg u)^2 - 2pn \sum_{u \in G} \deg u + p^2 n^3 \right)^{1/2} \\
&= n^{1/2} \left(\sum_{u,v \in G} \operatorname{codeg}(u, v) - 4pn \cdot e(G) + p^2 n^3 \right)^{1/2} \\
&= n^{1/2} \left(p^2 n^3 - 2p^2 n^3 + p^2 n^3 + o(n^3) \right)^{1/2} \\
&= o(n^2).
\end{aligned}$$

Then we can write

$$\begin{aligned}
|e(X, Y) - p|X||Y|| &= \left| \sum_{x \in X} (\deg(x, Y) - p|Y|) \right| \\
&\le n^{1/2} \left(\sum_{x \in X} (\deg(x, Y) - p|Y|)^2 \right)^{1/2}.
\end{aligned}$$

Since the summand is nonnegative, we can even enlarge the domain of summation from X to $V(G)$. So we have

$$\begin{aligned}
|e(X, Y) - p|X||Y|| &\le n^{1/2} \left(\sum_{x \in V} \deg(x, Y)^2 - 2p|Y| \sum_{x \in V} \deg(x, Y) + p^2 n |Y|^2 \right)^{1/2} \\
&= n^{1/2} \left(\sum_{y, y' \in Y} \operatorname{codeg}(y, y') - 2p|Y| \sum_{y \in Y} \deg y + p^2 n |Y|^2 \right)^{1/2} \\
&= n^{1/2} \left(|Y|^2 p^2 n - 2p|Y| \cdot |Y| pn + p^2 n |Y|^2 + o(n^3) \right)^{1/2} \\
&= o(n^2).
\end{aligned}$$

Now that we have proven the "C_4" between the statements **DISC** $\implies$ **COUNT** $\implies$ **C4** $\implies$ **CODEG** $\implies$ **DISC**, we relate the final condition, **EIG**, to the **C4** condition.

EIG $\implies$ **C4**: The number of labeled C_4s is within $O(n^3)$ of the number of closed walks of length 4, which is $\operatorname{tr}(A_G^4)$, where A_G is the adjacency matrix of G. From linear algebra, $\operatorname{tr}(A_G^4) = \sum_{i=1}^n \lambda_i^4$. The main term is λ_1: by assumption, $\lambda_1^4 = p^4 n^4 + o(n^4)$. Then we want to make sure that the sum of the other λ_i^4s is not too big. If you bound them individually, you just get $o(n^5)$, which is not enough. Instead,

we can write

$$\sum_{i\geq 2} \lambda_i^4 \leq \max_{i\neq 2} |\lambda_i|^2 \sum_{i\geq 1} \lambda_i^2$$

and note that $\sum_{i\geq 1} \lambda_i^2 = \operatorname{tr}(A_G^2) = 2e(G)$, so

$$\sum_{i=1}^{n} \lambda_i^4 = p^4 n^4 + o(n^4) + o(n^2)n^2 = p^4 n^4 + o(n^4).$$

C4 $\implies$ **EIG**: We use the *Courant–Fischer theorem* (also called the *min-max theorem*): for a real symmetric matrix A, the largest eigenvalue is

$$\lambda_1 = \sup_{x\neq 0} \frac{x^T A x}{x^T x}.$$

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of A_G, and let $\mathbb{1}$ be the all-1 vector in $\mathbb{R}^{V(G)}$. Then we have

$$\lambda_1 \geq \frac{\mathbb{1}^T A_G \mathbb{1}}{\mathbb{1}^T \mathbb{1}} = \frac{2e(G)}{n} = (p + o(1))\, n.$$

But from **C4**, we have

$$\lambda_1^4 \leq \sum_{i=1}^{n} \lambda_i^4 = \operatorname{tr} A_G^4 \leq p^4 n^4 + o(n^4),$$

which implies $\lambda_1 \leq pn + o(n)$. Hence, $\lambda_1 = pn + o(n)$.

We also have

$$\max_{i\neq 1} |\lambda_i|^4 \leq \operatorname{tr}(A_G^4) - \lambda_1^4 \leq p^4 n^4 - p^4 n^4 + o(n^4) = o(n^4),$$

as desired. □

What is most remarkable about Theorem 4.1 that the **C4** condition, seemingly the weakest of all the conditions, actually implies all the other conditions.

Remember that this theorem is about dense graphs (i.e. p is constant). We can write some analogs of the conditions for sparse graphs, where $p = p_n \to 0$ as $n \to \infty$. For example, in **DISC**, we need to change the $o(n^2)$ to $o(pn^2)$ to capture the idea that the number of edges of the quasirandom graph should be close to the expected number of edges of a truly random graph. Analogously, in **COUNT**, the number of labeled copies of H is $(1 + o(1))p^{e(H)}n^{v(H)}$. However, these conditions are *not* equivalent for sparse graphs. In particular, the counting lemma fails. For instance, here is a graph that satisfies the sparse analog of **DISC**, but does not even have any C_3.

Example 4.6. Take $p = o(n^{-1/2})$. The number of C_3s should be around $\binom{n}{3}p^3$, and the number of edges is $\binom{n}{2}p$. But by choice of p, the number of C_3s is now asymptotically smaller than the number

of edges, so we can just remove an edge from each triangle in this $G(n,p)$. We will those have removed $o(n^2p)$ edges, so the sparse analog of **DISC** still holds, but now the graph is triangle-free. This graph is pseudorandom in one sense, in that it still satisfies the discrepancy condition, but not in another sense, in that it has zero triangles.

4.2 *Expander mixing lemma*

Now we talk about a certain class of graphs, *expander graphs*, with a particularly strong discrepancy property.

Theorem 4.7 (Expander mixing lemma)**.** *Let G be an n-vertex, d-regular graph, with adjacency matrix having eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Let $\lambda = \max\{|\lambda_2|, |\lambda_n|\}$. Then for all $X, Y \subset V(G)$,*

$$\left| e(X,Y) - \frac{d}{n}|X||Y| \right| \leq \lambda \sqrt{|X||Y|}.$$

Proof. Let J be the all-1 matrix. We have

$$\begin{aligned} \left| e(X,Y) - \frac{d}{n}|X||Y| \right| &= \left| 1_X^T \left(A_G - \frac{d}{n} J \right) 1_Y \right| \\ &\leq \left\| A_G - \frac{d}{n} J \right\| |1_X|\,|1_Y| \\ &= \left\| A_G - \frac{d}{n} J \right\| \sqrt{|X||Y|}. \end{aligned}$$

It suffices to prove that the largest eigenvalue of $A_G - \frac{d}{n}J$ is at most λ.

Let v be an eigenvector of A_G. Since G is d-regular, one possibility for $v = (v_1, \ldots, v_n)$ is $\mathbb{1}$, which has corresponding eigenvalue d in A_G. Then $\mathbb{1}$ is also an eigenvector of $A_G - \frac{d}{n}J$, with corresponding eigenvalue 0. If $v \neq \mathbb{1}$, then it is orthogonal to $\mathbb{1}$, i.e. $v \cdot \mathbb{1} = \sum_{i=1}^n v_i = 0$. Therefore, $Jv = 0$, so v is also an eigenvector of $A_G - \frac{d}{n}J$ with same eigenvalue as in A_G. Thus, $A_G - \frac{d}{n}J$ has eigenvalues $0, \lambda_2, \lambda_3, \ldots, \lambda_n$, so its largest eigenvalue is λ, as desired. □

Expanders are related to pseudorandom graphs: when you have some small subset of vertices, you can expect them to have many neighbors. These kinds of graphs are called expanders because many vertices of the graph can be quickly reached via neighbors.

We now restrict our attention to a special class of graphs.

Definition 4.8. An ***(n,d,λ)-graph*** is an n-vertex, d-regular graph whose adjacency matrix has eigenvalues $d = \lambda_1 \geq \cdots \geq \lambda_n$ satisfying $\max\{|\lambda_2|, |\lambda_n|\} \leq \lambda$.

The expander mixing lemma (Theorem 4.7) can be rephrased as saying that if G is an (n, d, λ)-graph, then

$$\left| e(X, Y) - \frac{d}{n}|X||Y| \right| \leq \lambda \sqrt{|X||Y|}$$

for all $X, Y \subseteq V(G)$.

A random graph is pseudorandom with high probability. However, we would like to give deterministic constructions that have pseudorandom properties. The following is an example of such a construction.

Definition 4.9. Let Γ be a finite group, and let $S \subseteq \Gamma$ be a subset with $S = S^{-1}$. The ***Cayley graph*** $\mathrm{Cay}(\Gamma, S) = (V, E)$ is defined by $V = \Gamma$ and

$$E = \{(g, gs) \colon g \in \Gamma, s \in S\}.$$

Example 4.10. The ***Paley graph*** is a graph $\mathrm{Cay}(\mathbb{Z}/p\mathbb{Z}, S)$ for $p \equiv 1 \pmod 4$ a prime, and S the set of nonzero quadratic residues in $\mathbb{Z}/p\mathbb{Z}$.

Unfortunately, Raymond Paley was killed by an avalanche at the age of 26. His contributions include Paley graphs, the Paley–Wiener theorem, and Littlewood–Paley theory.

Proposition 4.11. *The Paley graph $G = \mathrm{Cay}(\mathbb{Z}/p\mathbb{Z}, S)$ satisfies $|\lambda_2|, |\lambda_p| \leq \frac{\sqrt{p}+1}{2}$, where $\lambda_1, \ldots, \lambda_p$ are the eigenvalues of its adjacency matrix.*

Proof. We simply write down a list of eigenvectors. Let the vertex 0 correspond to the first coordinate, the vertex 1 correspond to the second coordinate, etc. Let

$$\begin{aligned} v_1 &= (1, \ldots, 1) \\ v_2 &= (1, \omega, \omega^2, \ldots, \omega^{p-1}) \\ v_3 &= (1, \omega^2, \omega^4, \ldots, \omega^{2(p-1)}) \\ &\vdots \\ v_p &= (1, \omega^{p-1}, \ldots, \omega^{(p-1)(p-1)}), \end{aligned}$$

where ω is a primitive p-th root of unity.

We first check that these are eigenvectors. The all 1's vector v_1 has eigenvalue $d = \lambda_1$. We compute that the j-th coordinate of $A_G v_2$ is

$$\sum_{s \in S} \omega^{j+s} = \omega^j \sum_{s \in S} \omega^s.$$

Since ω_j is the j-th coordinate of v_2, and this holds for all j, the sum is the eigenvalue. In general, for $0 \leq k \leq p-1$,

$$\lambda_{k+1} = \sum_{s \in S} \omega^{ks}.$$

Note that this is a generic fact about Cayley graphs on $\mathbb{Z}/p\mathbb{Z}$, and the eigenvectors do not depend on S. Now we compute the sizes of the λ_i. For $k > 0$, we have

$$2\lambda_{k+1} + 1 = \sum_{a \in \mathbb{Z}/p\mathbb{Z}} \omega^{ka^2}.$$

Here, we used that S is the set of nonzero quadratic residues. The sum on the right is known as a ***Gauss sum***. It is evaluated as follows. We square the sum to get

$$\left| \sum_{a \in \mathbb{Z}/p\mathbb{Z}} \omega^{ka^2} \right|^2 = \sum_{a,b \in \mathbb{Z}/p\mathbb{Z}} \omega^{k((a+b)^2 - a^2)} = \sum_{a,b \in \mathbb{Z}/p\mathbb{Z}} \omega^{k(2ab+b^2)}.$$

For $b \neq 0$, the sum

$$\sum_{a \in \mathbb{Z}/p\mathbb{Z}} \omega^{k(2ab+b^2)} = 0,$$

since $k(2ab + b^2)$ for $a \in \mathbb{Z}/p\mathbb{Z}$ is a permutation of $\mathbb{Z}/p\mathbb{Z}$. For $b = 0$,

$$\sum_{a} \omega^{k(2ab+b^2)} = p.$$

Thus, the square of the Gauss sum is equal to p, so $\lambda_{k+1} = \frac{\pm\sqrt{p}-1}{2}$ for all $k > 0$. □

You might recognize $\sum_{s \in S} \omega^{ks}$ as a Fourier coefficient of the indicator function of S, viewed as a function on $\mathbb{Z}/p\mathbb{Z}$. Indeed, there is an intimate connection between the eigenvalues of a Cayley graph of an abelian group and the Fourier transform of a function on the group. In fact, the two spectra are identical up to scaling (partly the reason why we use the name "spectrum" for both eigenvalues and Fourier). There is a similar story for non-abelian groups, though Fourier analysis on non-abelian groups involves representation theory.

4.3 *Quasirandom Cayley graphs*

We saw that the Chung–Graham–Wilson theorem fails to hold for sparse analogs of the pseudorandomness conditions. However, it turns out, somewhat surprisingly, that if we restrict to Cayley graphs of groups (including non-abelian), no matter at what edge-density, the sparse analogs of DISC and EIG are equivalent.

For sparse graphs in general, the sparse analog of DISC does not imply the sparse analog of EIG. Consider the disjoint union of a large random d-regular graph and a K_{d+1}. This graph satisfies the sparse analog of DISC because the large random d-regular graph does. However, the top two eigenvalues are both $\lambda_1 = \lambda_2 = d$, because the all 1's vectors on each of the components is an eigenvector

with eigenvalue d, where as the sparse analog of EIG would give $\lambda_2 = o(d)$.

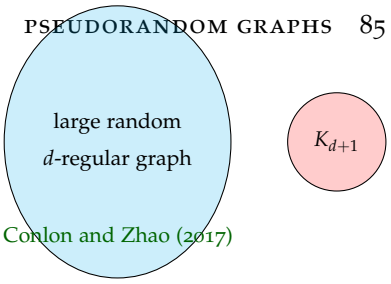


Figure 4.3: DISC does not imply EIG for a general graph.

Theorem 4.12 (Conlon–Zhao). *Let Γ be a finite group and $S \subseteq \Gamma$ a subset with $S = S^{-1}$. Let $G = \mathrm{Cay}(\Gamma, S)$. Let $n = |\Gamma|$ and $d = |S|$. For $\epsilon > 0$, we say that G has the property*

- *$DISC(\epsilon)$ if for all $X, Y \subseteq G$, we have $|e(X,Y) - \frac{d}{n}|X||Y|| \leq \epsilon dn$, and*
- *$EIG(\epsilon)$ if G is an (n,d,λ)-graph with $\lambda \leq \epsilon d$.*

Then if G satisfies $EIG(\epsilon)$, it also satisfies $DISC(\epsilon)$, and if it satisfies $DISC(\epsilon)$, then it also satisfies $EIG(8\epsilon)$.

Conlon and Zhao (2017)

The proof of Theorem 4.12 uses ***Grothendieck's inequality***.

Theorem 4.13 (Grothendieck's inequality). *There exists an absolute constant $K > 0$ such that for all matrices $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$,*

$$\sup_{\substack{x_i \in B \\ y_i \in B}} \sum_{i,j} a_{i,j} \langle x_i, y_i \rangle \leq K \sup_{\substack{x_i \in \{\pm 1\} \\ y_i \in \{\pm 1\}}} \sum_{i,j} a_{i,j} x_i y_j.$$

In the left hand side, the supremum is taken over all unit balls B in some $\mathbb{R}^m$.

Grothendieck (1953)

The right hand side of Grothendieck's inequality is the supremum of the bilinear form $\langle x, Ay \rangle$ over a discrete set. It is important combinatorially, but hard to evaluate. The left hand side is a "semidefinite relaxation" of the right hand side. There exist efficient methods to evaluate it, it is always at least the right hand side, and Grothiendieck's inequality tells us that we don't lose more than a constant factor when using it as an approximation for the right hand side.

Remark 4.14. It is known that $K = 1.78$ works. The optimal value, known as the "real Grothendieck constant," is unknown.

Krivine (1979)

Proof of Theorem 4.12. The fact that $EIG(\epsilon)$ implies $DISC(\epsilon)$ follows from the expander mixing lemma. Specifically, it tells us that

$$\left| e(X,Y) - \frac{d}{n}|X||Y| \right| \leq \lambda \sqrt{|X||Y|} \leq \epsilon dn$$

for any $X, Y \subseteq G$, which is what we want.

To prove the other implication, suppose $DISC(\epsilon)$ holds. For all $x, y \in \{\pm 1\}^\Gamma$, let $x^+, x^-, y^+, y^- \in \{0,1\}^\Gamma$ be such that

$$x_g^+ = \begin{cases} 1 & \text{if } x_g = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad x_g^- = \begin{cases} 1 & \text{if } x_g = -1 \\ 0 & \text{otherwise.} \end{cases}$$

Then $x = x^+ - x^-$. Similarly define y^+ and y^-.

Consider the matrix $A \in \mathbb{R}^{\Gamma\times\Gamma}$ with $A_{g,h} = 1_S(g^{-1}h) - \frac{d}{n}$ (here 1_S is the indicator function of S). Then

$$\langle x, Ay\rangle = \langle x^+, Ay^+\rangle - \langle x^-, Ay^+\rangle - \langle x^+, Ay^-\rangle + \langle x^-, Ay^-\rangle.$$

Each term in this sum is controlled by DISC. For example,

$$\langle x^+, Ay^+\rangle = e(X^+, Y^+) - \frac{d}{n}|X^+||Y^+|,$$

where $X^+ = \{g \in \Gamma : x_g = 1\}$, and $Y^+ = \{g \in \Gamma : y_g = 1\}$. Thus, $|\langle x^+, Ay^+\rangle| \le \epsilon dn$. This holds for the other terms as well, so

$$|\langle x, Ay\rangle| \le 4\epsilon dn \quad \text{for all} \quad x, y, \in \{\pm 1\}^\Gamma. \tag{4.1}$$

By the min-max characterization of the eigenvalue,

$$\max\{|\lambda_2|, |\lambda_n|\} = \sup_{\substack{|x|,|y|=1\\ x,y\in\mathbb{R}^\Gamma}} \langle x, Ay\rangle.$$

For all $x \in \mathbb{R}^\Gamma$, define $x^g \in \mathbb{R}^\Gamma$ by setting the coordinate $x^g_s = x_{sg}$ for all $s \in \Gamma$. Then $|x| = |x^g|$ since x^g simply permutes the coordinates of x. Then for all $x, y \in \mathbb{R}^\Gamma$ with $|x|, |y| = 1$,

$$\begin{aligned}
\langle x, Ay\rangle &= \sum_{g,h} A_{g,h} x_g y_h \\
&= \frac{1}{n} \sum_{g,h,s} A_{sg,sh} x_{sg} y_{sh} \\
&= \frac{1}{n} \sum_{g,h,s} A_{g,h} x_{sg} y_{sh} \\
&= \frac{1}{n} \sum_{g,h} A_{g,h} \langle x^g, y^h\rangle \le 8\epsilon d.
\end{aligned}$$

The inequality comes from Grothendieck's inequality with $K < 2$ combined with (4.1). Thus, $EIG(8\epsilon)$ is true. □

4.4 *Alon–Boppana bound*

In an (n, d, λ) graph, the smaller λ is, the more pseudorandom the graph is. A natural question to ask is, for fixed d, how small can λ be? We have the ***Alon–Boppana bound***.

Theorem 4.15 (Alon–Boppana bound). *Fix d. If G is an n-vertex graph whose adjacency matrix A_G has eigenvalues $\lambda_1 \ge \cdots \ge \lambda_n$, then* Alon (1986)

$$\lambda_2 \ge 2\sqrt{d-1} - o(1),$$

where $o(1) \to 0$ as $n \to \infty$.

Proof. Let $V = V(G)$. By Courant–Fischer, it suffices to exhibit a vector $z \in \mathbb{R}^V - \{0\}$ such that $\langle z, 1\rangle = 0$ and

Nilli (1991)

$$\frac{z^T A z}{z^T z} \geq 2\sqrt{d-1}.$$

Let $r \in \mathbb{N}$. Pick $v \in V$, and let V_i be the set of vertices at distance i from v. For example, $V_0 = \{v\}$ and $V_1 = N(v)$. Let $x \in \mathbb{R}^V$ be the vector with

$$x_u = w_i := (d-1)^{-i/2} \quad \text{for} \quad u \in V_i, 0 \leq i \leq r-1,$$

and $x_u = 0$ for all u such that $\operatorname{dist}(u, v) \geq r$. We claim that

$$\frac{x^T A x}{x^T x} \geq 2\sqrt{d-1}\left(1 - \frac{1}{2r}\right). \tag{4.2}$$

To show this, we compute

$$x^T x = \sum_{i=0}^{r-1} |V_i| w_i^2,$$

and

$$\begin{aligned} x^T A x &= \sum_{u \in V} x_u \sum_{u' \in N(u)} x_{u'} \\ &\geq \sum_{i=0}^{r-1} |V_i| w_i (w_{i-1} + (d-1) w_{i+1}) - (d-1)|V_{r-1}| w_{r-1} w_r \\ &= 2\sqrt{d-1}\left(\sum_{i=0}^{r-1} |V_i| w_i^2 - \frac{1}{2}|V_{r-1}| w_r^2\right). \end{aligned}$$

The inequality comes from the fact that each neighbor of $u \in V_i$ has distance at most $i+1$ from v and at least one neighbor has distance $i-1$ (note that the w_i are decreasing). However, since $x_u = 0$ for $\operatorname{dist}(u, w) \geq r$, so we must subtract off $(d-1)|V_{r-1}| w_{r-1} w_r$. Note that $|V_{i+1}| \leq (d-1)|V_i|$, so the above expression is

$$\geq 2\sqrt{d-1}\left(\sum_{i=1}^{r-1} |V_i| w_i^2\right)\left(1 - \frac{1}{2r}\right).$$

This proves (4.2). But we need $\langle z, 1\rangle = 0$. If $n > 1 + (d-1) + (d-1)^2 + \cdots + (d-1)^{2r-1}$, then there exist vertices $u, v \in V(G)$ at distance at least $2r$ from each other. Let $x \in \mathbb{R}^V$ be the vector obtained from the above construction centered at v. Let $y \in \mathbb{R}^V$ be the vector obtained from the above construction centered at u. Then x and y are supported on disjoint vertex sets with no edges between them. Thus, $x^T A y = 0$.

Choose a constant $c \in \mathbb{R}$ such that $z = x - cy$ has $\langle z, 1\rangle = 0$. Then

$$z^T z = x^T x + c^2 y^T y$$

and

$$z^T Az = x^T Ax + c^2 y^T Ay \geq 2\sqrt{d-1}\left(1 - \frac{1}{2r}\right) z^T z.$$

Taking $r \to \infty$ as $n \to \infty$ gives the theorem. □

We give a second proof of a slightly weaker result, but which is still in the spirit of Theorem 4.15.

Proof 2 (slightly weaker result). We'll show that $\max\{|\lambda_2|, |\lambda_n|\} \geq 2\sqrt{d-1} - o(1)$. This is an illustration of the trace method, also called the moment method. We have

$$\sum_{i=1}^{n} \lambda_i^{2k} = \operatorname{tr}(A^{2k}).$$

The right hand side is the number of closed walks of length $2k$ on G. Now, the number of closed walks of length $2k$ starting at a fixed vertex v in a d-regular graph is at least the number of closed walks of length $2k$ starting at a fixed v in an infinite d-regular tree. To see why this is true, note that given any walk on the infinite d-regular tree, we can walk in the same way on G by assigning an orientation to each vertex. But G may have more walks if it has cycles.

There are at least $C_k(d-1)^k$ closed walks of length $2k$ starting at a fixed v in an infinite d-regular tree, where $C_k = \frac{1}{k+1}\binom{2k}{k}$ is the k-th Catalan number. Thus, the number of walks of length $2k$ on G is at least $\frac{n}{k+1}\binom{2k}{k}(d-1)^k$. On the other hand,

$$d^{2k} + (n-1)\lambda^{2k} \geq \sum_{i=1}^{n} \lambda_i^{2k}.$$

Thus,

$$\lambda^{2k} \geq \frac{1}{k+1}\binom{2k}{k}(d-1)^k - \frac{d^{2k}}{n}.$$

The term $\frac{1}{k+1}\binom{2k}{k}$ is $(2 - o(1))^{2k}$ as $k \to \infty$. Letting $k \to \infty$ and $k = o(\log n)$ as $n \to \infty$ gives us $\lambda \geq 2\sqrt{d-1} - o(1)$. □

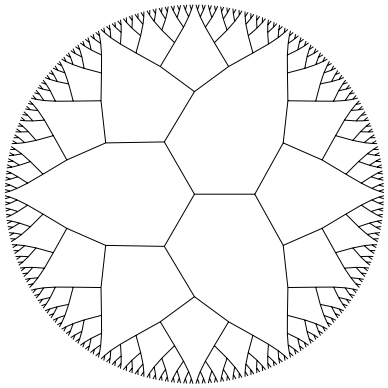


Figure 4.4: Infinite 3-regular tree. Image taken from the excellent survey on expander graphs: Shlomo, Linial, and Wigderson (2006)

Remark 4.16. Note that $2\sqrt{d-1}$ is the spectral radius of the infinite d-regular tree.

4.5 *Ramanujan graphs*

Definition 4.17. A ***Ramanujan graph*** is a d-regular graph whose adjacency matrix has eigenvalues $d = \lambda_1 \geq \cdots \geq \lambda_n$ so that $|\lambda_2|, |\lambda_n| \leq 2\sqrt{d-1}$, i.e. an (n, d, λ)-graph with $\lambda \leq 2\sqrt{d-1}$.

One example of a Ramanujan graph is K_{d+1}, as $\lambda_2 = \cdots = \lambda_n = -1$, but we are more interested in fixing d. For fixed d, do there exist infinitely many d-regular Ramanujan graphs?

Conjecture 4.18. *For all $d \geq 3$, there exist infinitely many d-regular Ramanujan graphs.*

We will discuss some partial results towards this conjecture.

Theorem 4.19 (Lubotzky–Phillips–Sarnak, Margulis)**.** *The above conjecture is true for all d with $d-1$ prime.*

Lubozsky, Phillips, and Sarnak (1988)
Margulis (1988)

Theorem 4.19 is proven by explicitly constructing a Cayley graph on the group $PSL(2,q)$ by invoking deep results from number theory relating to conjectures of Ramanujan, which is where the name comes from. In 1994, Morgenstern strengthened Theorem 4.19 result to all d for which $d-1$ is a prime power. This is essentially all that is known. In particular, Conjecture 4.18 is open for $d = 7$.

Morgenstern (1994)

It is interesting to consider the case of random graphs. What is the distribution of the largest non-λ_1 eigenvalue?

Theorem 4.20 (Friedman)**.** *Fix $d \geq 3$. A random n-vertex d-regular graph is, with prability $1 - o(1)$, a nearly-Ramanujan graph in the sense that*

$$\max\{|\lambda_2|, |\lambda_n|\} \leq 2\sqrt{d-1} + o(1)$$

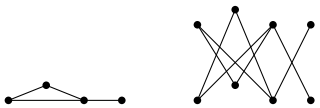
where the $o(1)$ term goes to 0 as $n \to \infty$.

Friedman (2004)

Experimental evidence suggests that, for all fixed d, a fixed proportion (between 0 and 1) of graphs on n vertices should be Ramanujan as $n \to \infty$. However, no rigorous results are known in this vein.

Recently, there has been some important progress on a bipartite analogue of this problem:

Note that for all bipartite graphs, $\lambda_i = -\lambda_{n+1-i}$. To see this, let the parts be A and B and take an eigenvector v with eigenvalue λ. Let v consist of v_A on A and v_B on B. Then negating v_B gives an eigenvector v' with eigenvalue $-\lambda$. So, a bipartite graph is called ***bipartite Ramanujan*** if $\lambda_2 \leq 2\sqrt{d-1}$.



G $\quad$ $G \times K_2$

An example of a graph G and its corresponding graph $G \times K_2$

Every Ramanujan graph G has an associated bipartite Ramanujan graph: we can construct $G \times K_2$; if G has eigenvalues $\{\lambda_i\}$ then $G \times K_2$ has eigenvalues $\{\lambda_i\} \cup \{-\lambda_i\}$, so the d-regular bipartite Ramanujan graph problem is a weakening of the original problem.

Theorem 4.21 (Marcus–Spielman–Srivastava)**.** *For all d, there exist infinitely many d-regular bipartite Ramanujan graphs.*

Marcus, Spielman, and Srivastava (2015)

Theorem 4.21 uses a particularly clever construction of randomized graphs.

4.6 *Sparse graph regularity and the Green–Tao theorem*

We will now combine the concepts of pseudorandom graphs with regularity involving sparse graphs. ***Sparse*** means edge density $o(1)$

— here we always consider a sequence of graphs on n vertices as $n \to \infty$, and $o(1)$ is with respect to n. The naïve analogue of the triangle removal lemma in a sparse setting is not true; we need an additional constraint:

Meta-Theorem 4.22 (Sparse triangle removal lemma)**.** For all $\epsilon > 0$, there exists $\delta > 0$ so that, if Γ is a sufficiently pseudorandom graph on n vertices with edge density p and G is a subgraph of Γ with fewer than $\delta n^3 p^3$ triangles, then G can be made triangle-free by deleting $\epsilon n^2 p$ edges.

We call this a ***meta-theorem*** as the condition "sufficiently pseudorandom" is not made explicit: the result is precisely true for some pseudorandomness conditions on which we will elaborate later. We can consider the traditional triangle removal lemma to be a special case of this where Γ is a complete graph.

Remark 4.23. Meta-Theorem 4.22 is not true without the hypothesis of Γ: take G as in Corollary 3.18 to have n vertices and $n^{2-o(1)}$ edges, where every edge belongs to exactly one triangle.

Remark 4.24. If $\Gamma = G(n,p)$ is an Erdős–Rényi graph with $p \geq \frac{C}{\sqrt{n}}$, then the conclusion of Meta-Theorem 4.22 holds. Conlon and Gowers (2014)

The motivation for the above is the Green–Tao Theorem:

Theorem 4.25 (Green–Tao)**.** *The primes contain arbitrarily long arithmetic progressions.* Green and Tao (2008)

This is in some sense a sparse extension of Szemerédi's Theorem: the density of the primes up to n decays like $\frac{1}{\log n}$ by the Prime Number Theorem.

The strategy for proving the Theorem 4.25 is to start with the primes and embed them (with high relative density) in what we will call *pseudoprimes*: numbers with no small prime divisors. This set is easier to analyze with analytic number theory, specifically using sieve methods. In particular, we can more easily show that the pseudoprimes are sufficiently pseudorandom, allowing the use of sparse hypergraph removal lemmas.

Recall the three main steps of using regularity: partitioning, cleaning, and counting. Naïve attempts to apply this approach to prove the sparse triangle removal lemma result in serious difficulties, and new ideas are needed. We require a sparse notion of regularity separate from the standard notion:

Definition 4.26. Given a graph G, a pair $(A,B) \subset V(G)^2$ is called (ϵ,p)***-regular*** if, for all $U \subset A, W \subset B$ with $|U| \geq \epsilon|A|, |W| \geq \epsilon|B|$, then

$$|d(U,W) - d(A,B)| < \epsilon p.$$

An equitable partition $V(G) = V_1 \sqcup \cdots \sqcup V_k$ is said to be (ϵ, p)-***regular*** if all but at most ϵ proportion of pairs are (ϵ, p)-regular.

Theorem 4.27 (Sparse regularity lemma)**.** *For all $\epsilon > 0$ there exists some $M \in \mathbb{N}$ for which every graph with edge density at most p has an (ϵ, p)-regular partition into at most M parts.*

Scott (2010)

Sparse objects have in some sense more freedom of structure, which is why statements like the sparse regularity lemma are much more intricate than the dense regularity lemma.

Theorem 4.27 is true but quite misleading: it could be true that most edges are inside irregular pairs. This makes the cleaning step more difficult as it might clean away too many of your edges. One example of this is a clique on $o(n)$ vertices.

In practice, G is often assumed to satisfy some "upper-regularity" hypothesis. For example, a graph is said to have ***no dense spots*** if there exists $\eta = o(1)$ and a constant $C > 0$ such that, for all $X, Y \subseteq V(G)$, if $|X|, |Y| \geq \eta |V|$, then

$$d(X, Y) \leq Cp.$$

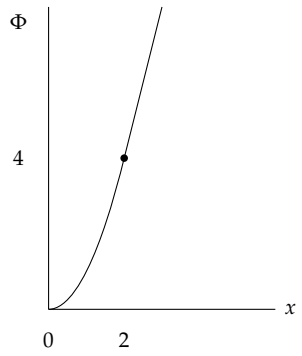
We will now prove Theorem 4.27 with the "no dense spots" hypothesis:

Proof sketch of Theorem 4.27 under the "no dense spots" hypothesis. This is essentially the same proof as in Szemerédi's Regularity Lemma. The key property we used in the energy increment argument was that the energy was bounded above by 1 and increased by ϵ^5. Now the energy increases by $\epsilon^5 p^2$. This depends on p, which could break the proof. However, as there are no dense spots, the final energy is at most $O(C^2 p^2)$, so the number of bad steps is bounded (depending on ϵ). □

Theorem 4.27 is still true without the condition "no dense spots," however:

Proof sketch of Theorem 4.27 in generality. We repeat the proof of Theorem 3.5 and instead of using x^2 as the energy, consider

$$\Phi(x) = \begin{cases} x^2 & \text{if } 0 \leq x \leq 2 \\ 4x - 4 & \text{if } x > 2. \end{cases}$$



Scott's energy function $\Phi(x)$.

This function has the boosting step: for all random variables $X \geq 0$, if $\mathbb{E}[X] \leq 1$,

$$\mathbb{E}\Phi(X) - \Phi(\mathbb{E}X) \geq \frac{1}{4} \operatorname{Var} X.$$

Furthermore, the inequality

$$\mathbb{E}\Phi(X) \leq 4\mathbb{E}X$$

allows us to bound the total energy of a partition by $O(1)$. □

Theorem 4.27 shows that the hard part of Meta-Theorem 4.22 is not the regularity lemma but the counting step. There is no counting lemma for sparse regular graphs. However, given our hypothesis that G is a subgraph of a pseudorandom graph Γ, we can construct a counting lemma which will allow us to prove the sparse triangle removal lemma.

We want something like the following to be true:

If you have three sets V_1, V_2, V_3 so that (V_i, V_j) are (ϵ, p)-regular $\forall i \neq j$ with edge density $d_{i,j}$, the number of triangles with one vertex in each part is

$$(d_{12}d_{23}d_{31} + O(\epsilon^c))\, p^3 |V_1||V_2||V_3|.$$

However, no such statement holds; take $G(n,p)$ with $p \ll \frac{1}{\sqrt{n}}$ and remove an edge from each triangle.

There is another example, due to Alon:

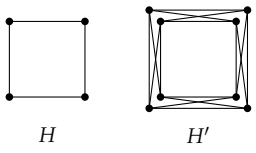
Example 4.28. There exists a triangle-free pseudorandom d-regular graph Γ with $d = \Theta\left(n^{2/3}\right)$ that is a (n, d, λ)-graph with $\lambda = \Theta\left(\sqrt{d}\right)$. Alon (1995)

To fix the issues with the above attempt, we have the following "meta-theorem:"

Meta-Theorem 4.29. Given three sets V_1, V_2, V_3 in G where G is a subgraph of a sufficiently pseudorandom graph with edge density p so that (V_i, V_j) are (ϵ, p)-regular for all $i \neq j$ with edge density $d_{i,j}$, the number of triangles with one vertex in each part is

$$(d_{12}d_{23}d_{31} + O(\epsilon^c))\, p^3 |V_1||V_2||V_3|.$$

We will now create a precise "sufficiently pseudorandom" condition for Meta-Theorem 4.22 and Meta-Theorem 4.29. We say that, given a graph H, a graph Γ is *pseudorandom with respect to H-density* if it has H-density $(1 + o(1))p^{e(H)}$. It turns out that the sparse triangle counting lemma Meta-Theorem 4.22 holds if Γ is pseudorandom with respect to H-density for every subgraph H of $K_{2,2,2}$.

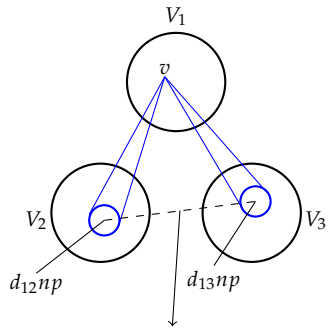


H and its 2-blowup H'.

Remark 4.30. This condition cannot necessarily be replaced by any of the other conditions given in Theorem 4.1 as our implication chain does not hold in a sparse setting.

This plays an analogous role to the C_4 condition in Theorem 4.1; C_4 was the ***2-blowup*** of an edge, while $K_{2,2,2}$ is a 2-blowup of a triangle. This acts somewhat like a graph-theoretic analogue of a second-moment: controlling copies of a graph H's second moment allows us to control copies of H in a subset of $V(G)$.

The proof Theorem 3.13 no longer works in the sparse case. Given three parts V_1, V_2, and V_3, that are pairwise (ϵ, p)-regular, we can no



There are not enough vertices to use (ϵ, p)-regularity.

longer take the neighbors of a vertex in V_1 that are in V_2 and V_3 and say that, as there are enough of them, they have enough overlap. This fails due to the extra factor of p in the sparse case.

Theorem 4.31 (Sparse counting lemma). *There exists a sparse counting lemma for counting H in $G \subset \Gamma$ if Γ is pseudorandom with respect to the density of every subgraph of the 2-blowup of H.*

Conlon, Fox, and Zhao (2015)

With this sparse counting lemma, one can prove Meta-Theorem 4.22 with the same proof structure as that of Theorem 3.15, using this pseudorandom property as our "sufficiently pseudorandom" condition on Γ.

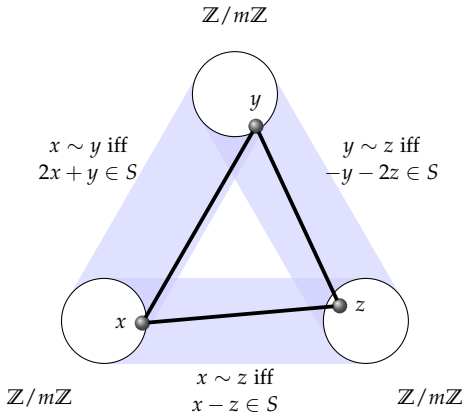
We state a equivalent version of Roth's theorem (Theorem 3.19):

Theorem 4.32 (Density Roth's Theorem). *If $A \subset \mathbb{Z}/n\mathbb{Z}$ with $|A| = \delta n$, then A contains at least $c(\delta)n^2$ 3-APs where $c(\delta) > 0$ is a constant depending only on δ.*

This can be proven by applying the proof structure from the proof of Theorem 3.19 using Theorem 3.15 (alternatively, we can use a supersaturation argument). Similarly to this, we can use Meta-Theorem 4.22 to prove a sparse analogue of Roth's Theorem:

Meta-Theorem 4.33 (Relative Roth's Theorem). If $S \subset \mathbb{Z}/n\mathbb{Z}$ is sufficiently pseudorandom with $|S| = pn$, and $A \subset S$ with $|A| \geq \delta|S|$, then A contains at least $c(\delta)n^2p^3$ 3-APs where $c(\delta) > 0$ is a constant depending only on δ.

What should "pseudorandom" mean here? Recall our proof of Roth's Theorem: creating three copies X, Y, Z of $\mathbb{Z}/n\mathbb{Z}$ and putting edges among $x \in X, y \in Y, z \in Z$ if $2x + y \in S$, $x - z \in S$, $-y - 2z \in S$. From this construction, we can read out the pseudorandom properties we want this graph Γ_S to have from our counting lemma.



Definition 4.34. We say that $S \subset \mathbb{Z}/n\mathbb{Z}$ satisfies a ***3-linear-forms condition*** if, for uniformly randomly chosen $x_0, x_1, y_0, y_1, z_0, z_1 \in \mathbb{Z}/n\mathbb{Z}$, the probability that the twelve numbers formed by the linear forms corresponding to those above:

$$\left\{ \begin{array}{lll} -y_0 - 2z_0, & x_0 - z_0, & 2x_0 + y_0, \\ -y_1 - 2z_0, & x_1 - z_0, & 2x_1 + y_0, \\ -y_0 - 2z_1, & x_0 - z_1, & 2x_0 + y_1, \\ -y_1 - 2z_1, & x_1 - z_1, & 2x_1 + y_1 \end{array} \right\}$$

are all in S is within a $1 + o(1)$ factor of the expectation if $S \subset \mathbb{Z}/n\mathbb{Z}$ were random with density p, and the same holds for any subset of these 12 expressions.

We also have a corresponding theorem, a simplification of the Relative Szemerédi Theorem used by Green–Tao:

Green and Tao (2008)

Theorem 4.35 (Relative Szemerédi Theorem). *Fix $k \geq 3$. If $S \subset \mathbb{Z}/n\mathbb{Z}$ satisfies the k-linear-forms condition then any $A \subset S$ with $|A| \geq \delta|S|$ has a lot of k-APs.* Conlon, Fox, and Zhao (2015)

There are still interesting open problems involving sparse regularity, particularly involving what sorts of pseudorandomness hypotheses are required to get counting lemmas.

Remark 4.36. Theorems like Theorem 4.35 can also be proven without the use of regularity, in particular by using the technique of ***transference***: Szemerédi's Theorem can be treated as a black box, and applied directly to the sparse setting. For more about this, see "Green–Tao theorem: an exposition" by Conlon, Fox, Zhao. Conlon, Fox, and Zhao (2014)

18.217 Graph Theory and Additive Combinatorics

Fall 2019