The Combinatorics of Patterns in Subsets and Graphs

Based on a series of lectures by
Fan Chung Graham

Fall 2004

Preface

During the fall of 2004 Fan Chung Graham taught a graduate course at UCSD about the combinatorics of patterns in subsets and graphs. Class members would take turns taking notes and then write them up, this is a collection of those notes (with some light revision).

We thank Fan for teaching the class and for providing some corrections/directions on the lecture notes. We would also like to thank Van Vu for being a guest lecturer (Lecture 7).

These lecture notes were written by Blair Angle, Steven Butler, Kevin Costello, Daniel Felix, Paul Horn, Ross Richardson, D. Jacob Wildstrom and Lei Wu. This is a work in progress and is being updated from time to time. The most current version can be found by visiting the website below.

www.math.ucsd.edu/~sbutler/math262/

Contents

Lecture 1

Introduction

Steven Butler

1.1 Introduction

A recent result in additive number theory is the proof by Green and Tao [26] that the primes contain arbitrarily long arithmetic progressions. That is for any k there is an a and b so that $a, a+b, a+2b, \ldots, a+(k-1)b$ are all primes.

This is a major result in the field of additive number theory. A key component of their approach came out of combinatorics, particularly the Regularity Lemma. In the first few lectures we will work to understand what the Regularity Lemma tells us and how we can apply it.

1.1.1 Szemerédi

The Regularity Lemma can be traced to the writings of Szemerédi [40, 41] where he showed that a set of integers with positive upper density must contain arbitrarily long arithmetic progressions of integers.

One way to think of this is that once density exceeds a certain threshold then patterns must emerge (i.e., there are some patterns that can't be avoided). This is another way of saying that total disorder is impossible.

It should be noted that the primes do *not* have positive upper density (so the result of Green and Tao are definitely not trivial), nevertheless the key to Green's proof was not so much about primes but rather about the density and distribution of primes.

1.2 Graphs

A graph $G = (V, E)$ is composed of V, a vertex set which is a collection of elements (pictorially we represent these by dots), and E, an edge set which gives a collection of relations (pictorially we represent these by lines connecting the dots). An example of a graph is the telephone graph where every vertex represents a telephone number and an edge is in the graph if one number has called the other. Graphs are difficult to work with as they can have complex structure and so it is difficult to say what is happening to an "arbitrary" graph.

The integers can intuitively be thought of as points on a line and in geometry we can think of points in an array. What we would like is to simplify the way to think of a graph so that we can have better intuition about what is going on.

1.2.1 Graphs and the Regularity Lemma

Imagine that we are dealing with a graph G with n vertices (where we think of n as a large number). The idea of the Regularity Lemma is that we can partition the graph into finitely many pieces so that between any two pieces the edges behave "random"-like.

The partitioning of the graph lets us use the "divide and conquer" strategy where we break the graph into pieces where each piece is easy to work with. When we say "finitely many pieces" this means that the number of pieces is independent of n and depends only on how well we want to control some property.

One question is what we mean by random-like. Intuitively, we think that a random-like graph should somehow have the edges evenly distributed. One property that we would expect in such a graph is that between two subsets of vertices of sufficient size, if we take half of the vertices of each subset and count the number of edges that are left we should have approximately a quarter of the total edges between the two subsets. There are a wide variety of properties which turn out to be equivalent to this that we would expect a random graph to have, this has led to the formulation and widespread use of quasi-random graphs which were developed by Chung, Graham and Wilson [15].

If two pieces are random-like we can cut down the number of parameters (i.e., amount of information) tremendously, since the important information between two such pieces behaving randomly is their edge density.

While it is unrealistic to think that for an arbitrary graph that we can reduce it down to a single parameter the Regularity Lemma says we can cut it down to only finitely many (where the number depends only on the accuracy of what we want to control/measure and not the size of the original graph).

1.2.2 Drawbacks on the Regularity Lemma

The Regularity Lemma is a powerful tool but it does have some drawbacks. First, though the number of parameters is finite it can be very large (i.e., leading to towers of numbers like those below) so it is not practical to implement it algorithmically. Most results involving the Regularity Lemma will involve the phrase "for n sufficiently large."

Second, this only works when G has positive edge density, i.e., if e is the number of edges and n the number of vertices then to apply the Regularity Lemma there is some absolute positive value ε independent of n for which

$$\frac{e}{\binom{n}{2}} > \varepsilon.$$

So the number of edges behaves like cn^2 where c is some positive constant. This is a large number of edges and so the Regularity Lemma does not work well for graphs with few edges (such graphs are called sparse graphs, the telephone graph is an example of such a graph).

1.3 van der Waerden Theorem

An early result on arithmetic progressions was given by van der Waerden.

Theorem 1.1 (van der Waerden [44]). *Given integers $k, t > 0$ there exists $W(k,t)$ such that for all $n \geq W(k,t)$ then any partitioning of $\{1, 2, \ldots, n\}$ into t parts contains one part with an arithmetic progression of length k.*

One way to think of this is that for any k then for n sufficiently large, no matter how we paint $\{1, 2, \ldots, n\}$ with t different colors we can always find one arithmetic progression of length k all of one color.

In his paper van der Waerden attributed the problem to a conjecture of Baudet though it is more likely that the problem came from Schur.

The number $W(k, t)$ is known as the van der Waerden number. This number is found using double induction and tends to be unmanageable. Ron Graham, following in the spirit of Erdős, put up \$1000 to show that

$$W(k, 2) \leq 2^{2^{\cdot^{\cdot^{\cdot^{2}}}}} \text{ — a tower of } k \text{ 2's.}$$

Shelah [38] made some progress and was awarded \$500. Later Tim Gowers was able to cut the upper bound to

$$W(k, 2) \leq 2^{2^{2^{2^{2^{k+9}}}}}$$

and was subsequently awarded \$1000 by Graham (the first and to date only such \$1000 award given by Graham).

It is known, by Berlekamp [4], that for k a prime that $k2^k \leq W(k, 2)$ and Graham has put up another \$1000 to show that $W(k, 2) \leq 2^{k^2}$.

1.4 A Conjecture of Erdős

One of the most famous open problems of Erdős is to show that if $a_1, a_2, \ldots$ is a subset of the integers such that $\sum \frac{1}{a_i}$ is divergent then show that the a_i contain arbitrarily long arithmetic progressions (it is not even yet known if it must contain an arithmetic progression of length three).

Lecture 2

Szemerédi's Regularity Lemma, Part 1

Daniel Felix

2.1 Edge Density

Let $G = (V, E)$ be a graph, and let A and B be disjoint subsets of V. Let $e(A, B)$ denote the number of edges joining a vertex in A to a vertex in B and consider the quantity $\delta(A, B) := e(A, B)/|A||B|$. Note that this is precisely the fraction of edges of the complete bipartite graph on (A, B) which appear in G. For this reason we call $\delta(A, B)$ the *density* of the pair (A, B).

Definition 2.1. We say the pair (A, B) is ε-*regular* if for every $A' \subseteq A, B' \subseteq B$ satisfying $|A'| \geq \varepsilon|A|$ and $|B'| \geq \varepsilon|B|$ we have

$$|\delta(A, B) - \delta(A', B')| < \varepsilon.$$

In words, an ε-regular pair (A, B) is one whose density is nearly the same as the density of every reasonably sized subpair. Thus the edges between A and B behave almost as if they had been generated at random, each appearing independently with probability $\delta(A, B)$.

2.2 Statement of the Regularity Lemma

Theorem 2.1 (Szemerédi Regularity Lemma). *For every $\varepsilon > 0$ and $m > 0$ there exist two integers $M(\varepsilon, m)$ and $N(\varepsilon, m)$ such that the following holds: if $G = (V, E)$ is any graph on $n \geq N(\varepsilon, m)$ vertices, then we can partition V into $k+1$ classes $V_0, V_1, \ldots, V_k$ such that*

1. $m \leq k \leq M(\varepsilon, m)$,
2. $|V_0| < \varepsilon n$,
3. $|V_1| = |V_2| = \cdots = |V_k|$,
4. *all but εk^2 pairs (V_i, V_j) are ε-regular.*

We call such a partition a *Szemerédi partition*. The set V_0 is called the *exceptional class*. The pairs (V_i, V_j) which are not ε-regular will be called *irregular*, though they may be ε'-regular for some $\varepsilon' > \varepsilon$. We will refer to edges in irregular pairs, as well as those edges with both end vertices in the same class, as *irregular edges*. The *regular edges* are those edges which are not irregular.

Put very roughly, the Regularity Lemma says that all sufficiently large graphs behave approximately like a random multipartite graph (of a bounded number of pieces) with a few offending vertices and edges added into the mix. To become familiar with the statement of the lemma, we'll take some time to see why the conditions and conclusions are important.

First, the fact that k is bounded is essential; otherwise, we could take each set in our partition to be a singleton. Then *every* pair would be trivially ε-regular for every $\varepsilon > 0$.

This is also one reason why we should find it sensible that m should be allowed to be large. Without a lower bound on k, there is no way to know that the partition guaranteed by the theorem doesn't consist solely of V_1, which would trivially satisfy the conclusions of the lemma. But this isn't the entire story on m. Notice that the number of edges whose endvertices belong to the same class, together with those which have at least one endvertex in the exceptional class, is bounded from above by

$$k\binom{\lfloor n/k \rfloor}{2} + \varepsilon n(n-1) \leq \frac{n^2}{2k} + \varepsilon n^2,$$

which is less than $2\varepsilon n^2$ if $k \geq m > 1/2\varepsilon$. If we add to this the number of edges which join irregular pairs, we're still left with at most $3\varepsilon n^2$ offending, irregular edges. So a very large m will ensure that nearly all of the edges occur between regular pairs. That is, all but a small fraction of the edges can be thought of as being randomly distributed in the sense we have described.

As noted in [30], the role of the "trash can" V_0 is purely technical; it allows us just enough freedom to conclude that all of the other V_i have the same size. One can easily deduce a modified version of the Regularity Lemma from its original statement in which $|V_0| = 0$, as long $||V_i| - |V_j||$ is allowed to be 0 or 1.

Another cleaner version of the Regularity Lemma is the following:

Theorem 2.2. *Let $\varepsilon' > 0$, let $G = (V, E)$ be any graph on n vertices, and let P be a partition of V with $|P|$ an absolute constant. Then there exist two integers $M(\varepsilon', |P|)$ and $N(\varepsilon', |P|)$ such that the following holds: if $n \geq N(\varepsilon', |P|)$, then there exists a refinement P' of P, $|P'| < M(\varepsilon', |P|)$, such that all but $\varepsilon' n^2$ edges lie in ε'-regular pairs of P'.*

It's not too hard to show that the original statement implies this one.

Proposition 2.1. *Theorem 2.1 implies Theorem 2.2.*

Proof. Given P and $0 < \varepsilon' \leq 1$, we simply apply Theorem 2.1 with $\varepsilon = (\varepsilon'/4|P|)^2$ and $m = \max(|P|, 1 + 1/2\varepsilon)$. As long as $n \geq N(\varepsilon, m)$, we're left with a Szemerédi partition which has, as we've just seen, fewer than

$$3\varepsilon n^2 \leq 3\left(\frac{\varepsilon'}{4|P|}\right)^2 n^2 \leq \frac{3}{4}\varepsilon' n^2$$

irregular edges.

Denote the sets in P by $S_1, \ldots, S_{|P|}$. The refinement P' that we are looking for consists of the sets $S_i \cap V_j$, where the V_j are the classes of the Szemerédi partition. Notice that the number of sets in P' is at least $|P|$ and is no greater than $|P|M(\varepsilon, m)$. Of course, we still must show that nearly all the edges lie between ε'-regular pairs of P'.

The key fact here is that if $V_j' \subset V_j$ and $V_l' \subset V_l$ with $|V_j'| > \sqrt{\varepsilon}|V_j|$ and $|V_l'| > \sqrt{\varepsilon}|V_l|$, then (V_j', V_l') is $2\sqrt{\varepsilon}$-regular, and therefore ε'-regular as well. Thus, if (V_i, V_j) is ε-regular, $(S_i \cap V_j, S_p \cap V_l)$ will always be ε'-regular, unless at least one set in the pair is too small.

So we see that the existence of many small classes in P' will increase the number of irregular edges. All that remains to be checked is that this increase is not appreciable. If $S_i \cap V_j$ is one of these small classes, then surely

$$\begin{aligned} e(S_i \cap V_j, G) &\le |S_i \cap V_j| n \\ &\le \frac{\varepsilon'}{4|P|} |V_j| n. \end{aligned}$$

Summing this expression over all small classes gives an upper bound on the number of edges incident to small classes, namely,

$$\begin{aligned} \sum_{\substack{i,j \\ \text{small classes}}} \frac{\varepsilon'}{4|P|} |V_j| n &\le \frac{\varepsilon'}{4} n \sum_{j=0}^{k} |V_j| \\ &= \frac{\varepsilon'}{4} n^2. \end{aligned}$$

Adding this to the initial $3\varepsilon' n^2/4$ irregular edges at last yields that fewer than $\varepsilon' n^2$ edges in P' are irregular. Hence Theorem 2.1 implies Theorem 2.2. □

2.3 Cauchy-Schwarz Inequality

Our starting point in the proof of the Regularity Lemma is the Cauchy-Schwarz inequality

$$\left(\sum_{i=1}^{n} a_i^2 \right) \cdot \left(\sum_{i=1}^{n} b_i^2 \right) \ge \left(\sum_{i=1}^{n} a_i b_i \right)^2,$$

where the a_i and b_i are taken to be in $\mathbb{R}$. This follows easily from the next lemma (which is a special case of Jensen's inequality).

Lemma 2.1. *Let $\alpha_1, \ldots, \alpha_n$ and $x_1, \ldots, x_n$ be elements of $\mathbb{R}$ satisfying $\alpha_i \ge 0$ for all i and $\sum_{i=1}^{n} \alpha_i = 1$. Then*

$$\sum_{i=1}^{n} \alpha_i x_i^2 \ge \left(\sum_{i=1}^{n} \alpha_i x_i \right)^2.$$

Proof. Let $\bar{x} := \sum_{i=1}^{n} \alpha_i x_i$. Then,

$$\begin{aligned} 0 &\le \sum_{i=1}^{n} \alpha_i (x_i - \bar{x})^2 \\ &= \left(\sum_{i=1}^{n} \alpha_i x_i^2 \right) - \bar{x}^2 \end{aligned}$$

and the result follows. □

One obtains the Cauchy-Schwarz inequality by setting $\alpha_i = a_i^2/\sum_j a_j^2$ and $x_i = b_i \sum_j a_j^2/a_i$ (note that, without loss of generality, a_i is never zero).

We'll also need the following strengthening of the Cauchy-Schwarz inequality.

Lemma 2.2. *Let α_i, x_i, and $\bar{x}$ be as above, and let $\varepsilon > 0$. Suppose that for some $m < n$ we have*

$$\sum_{i=1}^{m} \alpha_i x_i \geq \left(\sum_{i=1}^{m} \alpha_i \bar{x}\right) + \varepsilon.$$

Then

$$\sum_{i=1}^{n} \alpha_i x_i^2 \geq \left(\sum_{i=1}^{n} \alpha_i x_i\right)^2 + \varepsilon^2 \Big/ \sum_{i=1}^{m} \alpha_i \sum_{j=m+1}^{n} \alpha_j.$$

Proof. The proof is straightforward.

$$\begin{aligned}
\sum_{i=1}^{n} \alpha_i x_i^2 - \bar{x}^2 &= \sum_{i=1}^{n} \alpha_i (x_i - \bar{x})^2 \\
&= \left(\sum_{i=1}^{m} \alpha_i (x_i - \bar{x})^2\right) + \left(\sum_{i=m+1}^{n} \alpha_i (x_i - \bar{x})^2\right) \\
&\geq \frac{\left(\sum_{i=1}^{m} \alpha_i (x_i - \bar{x})\right)^2}{\sum_{i=1}^{m} \alpha_i} + \frac{\left(\sum_{i=m+1}^{n} \alpha_i (x_i - \bar{x})\right)^2}{\sum_{i=m+1}^{n} \alpha_i} \\
&\geq \frac{\varepsilon^2}{\sum_{i=1}^{m} \alpha_i} + \frac{\varepsilon^2}{\sum_{i=m+1}^{n} \alpha_i} \\
&= \frac{\varepsilon^2}{\sum_{i=1}^{m} \alpha_i \sum_{j=m+1}^{n} \alpha_j},
\end{aligned}$$

where in going from the second to the third line we used the Cauchy-Schwarz inequality. The result now follows. □

Remark. The above work can be more easily digested by considering a random variable X which takes the value x_i with probability α_i. Interpreted this way, Lemma 2.1 amounts to nothing more than the statement $\mathrm{Var}[X] \geq 0$ (more specifically, $\mathrm{E}[X^2] \geq \mathrm{E}[X]^2$), while the proof of Lemma 2.2 boils down to the same fact for conditional expectation.

Lecture 3

Szemerédi's Regularity Lemma, Part 2

Kevin Costello

3.1 Quick Review of Definitions

As in the last lecture, given two subsets A and B of a graph, we define the edge density between A and B to be

$$\delta_{A,B} := \delta(A, B) = \frac{e(A, B)}{|A||B|},$$

where $e(A, B)$ denotes the number of edges between vertices in A and vertices in B. We say the pair (X, Y) is *ε-regular* if for every pair of subsets (A, B) with $|A| \geq \varepsilon|X|$ and $|B| \geq \varepsilon|Y|$ we have

$$|\delta_{A,B} - \delta_{X,Y}| < \varepsilon,$$

that is, the density between "large" subsets of X and Y is close to the density between X and Y.

We would expect a random graph on $A \cup B$ with edge density $\delta_{A,B}$ to have this property, since all sufficiently large pairs should have approximately the same edge density. This is one of many possible starting points for a theory of "quasi-randomness" in graphs, as ε-regularity can be shown to be equivalent to other properties which we would expect to be satisfied by a random graph with edge density δ.

3.2 A Version of the Regularity Lemma

We will prove a version of the Regularity Lemma which states that given any partition of the vertices of a sufficiently large graph, we can refine the partition further so that a large portion of the edges lie between ε-regular pairs of blocks of vertices. To be precise:

Theorem 3.1. *Let c be a given number and let G a graph with P a partition of $V(G)$ where $|P| \leq c$. Then for all $\varepsilon > 0$ there exists $M(\varepsilon, c)$ such that there is a partition $\mathbb{P}$ with the following properties:*

1. *$\mathbb{P}$ is a refinement of P,*
2. *$|\mathbb{P}| < M(\varepsilon, c)$,*
3. *all but at most εn^2 of the edges of G have their endpoints lying in two blocks of $\mathbb{P}$ which form an ε-regular pair.*

The problem with using this theorem in practice (as will be seen from the proof below) is that the value of M can be *extremely* large. In practice, mathematicians using the Regularity Lemma will often parallel portions of the proof of the lemma rather than apply the lemma itself, as this will often lead to more reasonable bounds.

3.3 The Index of a Partition and its Properties

For a given partition P of the subsets X and Y of vertices of the graph, define the *index* between X and Y by

$$I(P(X,Y)) = \sum_{\substack{X_i \subseteq X \\ Y_i \subseteq Y}} \frac{(e(X_i, Y_j))^2}{|X||Y||X_i||Y_j|},$$

we define $I(P) := I(P(V,V))$ so that

$$I(P) = \sum_{X_i, Y_j \in P} \frac{(e(X_i, Y_j))^2}{|V|^2|X_i||Y_j|}.$$

We now give a series of facts which will serve as the basis for an algorithm for constructing the desired $\mathbb{P}$.

Fact 3.1.

$$I(P) = \sum_{X_i, Y_j \in P} \alpha_{X_i,Y_j} \delta^2_{X_i,Y_j} \quad \text{where} \quad \alpha_{X_i,Y_j} = \frac{|X_i||Y_j|}{|V|^2}.$$

Proof. This is immediate from the definition of $I(P)$ and δ_{X_i,Y_j}. □

Fact 3.2. If P' is a refinement of P, $I(P') \geq I(P)$.

Proof. Suppose block X of P is divided into $X_1, \ldots, X_r$ and block Y of P is divided into $Y_1, \ldots, Y_s$. The total contribution to $I(P')$ from ordered pairs corresponding to portions of the pair (X, Y) in P is

$$\sum_{i,j} \frac{(e(X_i, Y_j))^2}{|V|^2 |X_i||Y_j|}.$$

Applying the Cauchy-Schwarz inequality from the last lecture we have the following.

$$\begin{aligned}
\sum_{i,j} \frac{(e(X_i, Y_j))^2}{|V|^2 |X_i||Y_j|} &= \frac{|X||Y|}{|V|^2} \sum_{i,j} \underbrace{\frac{|X_i||Y_j|}{|X||Y|}}_{=\alpha_{i,j}} \Bigg(\underbrace{\frac{e(X_i, Y_i)}{|X_i||Y_j|}}_{=a_{i,j}} \Bigg)^2 \\
&\geq \frac{|X||Y|}{|V|^2} \left(\sum_{i,j} \frac{|X_i||Y_j|}{|X||Y|} \frac{e(X_i, Y_j)}{|X_i||Y_j|} \right)^2 \\
&= \frac{(e(X,Y))^2}{|V|^2 |X||Y|}.
\end{aligned}$$

Note that when computing $I(P')$ we also include terms corresponding to edges between the X_i and also between the Y_j, but these terms are nonnegative, so removing them can only decrease the index. Since this holds true for every pair of blocks we can conclude that $I(P') \geq I(P)$. □

Fact 3.3. If (X, Y) is not an ε-regular pair, we can refine our partition further in such a way that $I(P'(X, Y)) \geq I(P(X, Y)) + \varepsilon^4 = \delta^2_{X,Y} + \varepsilon^4$.

Proof. Since our pair is not ε-regular, we can find subsets $|A|$ and $|B|$ with $|A| \geq \varepsilon|X|$, $|B| \geq \varepsilon|Y|$, and $|\delta_{A,B} - \delta_{X,Y}| \geq \varepsilon$. Without loss of generality assume that $\delta_{A,B} > \delta_{X,Y}$. We now refine our partition by splitting X into $(X - A) \cup A$ and Y into $(Y - B) \cup B$.

Before the refinement the contribution to the index from the pair (A, B) was

$$\frac{(e(A,B) + e(B, X-A) + e(A, Y-B) + e(X-A, Y-B))^2}{|X||Y|},$$

afterwards, the contribution is at least

$$\frac{(e(A,B))^2}{|A||B|} + \frac{(e(B,X-A))^2}{|B||X-A|} + \frac{(e(A,Y-B))^2}{|A||Y-B|} + \frac{(e(X-A,Y-B))^2}{|X-A||Y-B|}.$$

Again, we are leaving out the contribution from edges between A and $(X-A)$ and those between B and $(Y-B)$, but those can only increase the index. We now apply the modified Cauchy-Schwarz inequality from the previous lecture, using $m = 1$ and as our variables

$$\alpha_1 = \frac{|A||B|}{|X||Y|}, \alpha_2 = \frac{|B||X-A|}{|X||Y|},$$
$$\alpha_3 = \frac{|A||Y-B|}{|X||Y|}, \alpha_4 = \frac{|X-A||Y-B|}{|X||Y|},$$
$$x_1 = \frac{e(A,B)}{|A||B|}, x_2 = \frac{e(B,X-A)}{|B||X-A|},$$
$$x_3 = \frac{e(A,Y-B)}{|A||Y-B|}, x_4 = \frac{e(X-A,Y-B)}{|X-A||Y-B|}.$$

Note that $x_1 = \delta_{A,B}$, and that $\bar{x} = \sum x_i \alpha_i = e(X,Y)/(|X||Y|) = \delta_{X,Y}$, thus our discrepancy is

$$\varepsilon' = \alpha_1(x_1 - \bar{x}) = \frac{|A||B|}{|X||Y|}(\delta_{A,B} - \delta_{X,Y}) \geq \frac{(\varepsilon|X|)(\varepsilon|Y|)}{|X||Y|}\varepsilon = \varepsilon^3.$$

The inequality now gives that

$$I(P') \geq |X||Y|\big(\sum_{i=1}^{4} \alpha_i x_i^2\big) \geq |X||Y|\left(\big(\sum_{i=1}^{4} \alpha_i x_i\big)^2 + \frac{(\varepsilon')^2}{\alpha_1(\alpha_2+\alpha_3+\alpha_4)}\right)$$
$$\geq \frac{(e(X,Y))^2}{|X||Y|} + \frac{\varepsilon^6}{\varepsilon^2}\frac{(|X||Y|)^2}{(|X||Y|)^2 - |A||B|} \geq I(P) + \varepsilon^4.$$

□

Fact 3.4. If P is the partition of the vertex set into a single block, $I(P) = \delta^2$.

Proof. There is only a single term in the sum defining the index, and that term is $e(V,V)^2/|V|^4$. □

Fact 3.5. If P is the trivial partition (every block consists only of a single vertex), $I(P) = \delta$.

Proof. Each edge of G contributes a $1/n^2$ to the sum, so the total sum is $e(V,V)/|V|^2 = \delta$. □

3.4 The Proof of Regularity Lemma

Proof of the Regularity Lemma. The proof is algorithmic. Let $P_0 = P$ be the initial partition which combining Fact 3.2 and Fact 3.4 satisfies $I(P_0) \geq \delta^2$. We then perform the following iterated algorithm, if P_t has at most εn^2 edges then P_t is the desired partition and we stop. Otherwise, for each irregular pair in P_t we use Fact 3.3 to construct a refinement of P_t which increases the index. Let P_{t+1} be the coarsest partition which is still a refinement of P_t and includes **all** of the $(A \cup (X - A), B \cup (Y - B))$ refinements from Fact 3.3.

Let ${}_rX_i$ denote the terms of the refinement of X_i and ${}_sY_j$ similarly denote the terms of the refinement of Y_j. Then using the algorithm we have that

$$I(P_{t+1}) = \sum_{\substack{i,j \\ (X_i,Y_j)\in P_t \\ \varepsilon\text{-regular}}} \alpha_{X_i,Y_j}\delta^2_{X_i,Y_j} + \sum_{\substack{i,j \\ (X_i,Y_j)\in P_t \\ \text{not } \varepsilon\text{-regular}}} \alpha_{X_i,Y_j} \underbrace{\sum_{\substack{r,s \\ {}_rX_i\subseteq X_i \\ {}_sY_j\subseteq Y_j}} \frac{(e({}_rX_i, {}_sY_j))^2}{|X_i||Y_j|||{}_rX_i||{}_sY_j|}}_{q_{i,j}}.$$

Examining the term $q_{i,j}$ this corresponds to the index of a (possibly trivial) refinement of the refinement that was done by the algorithm on the ε-irregular pair (X_i, Y_j). The important part of this is that by Fact 3.2 and Fact 3.3 we have

$$q_{i,j} \geq I(P'(X_i,Y_j)) \geq I(P(X_i,Y_j)) + \varepsilon^4 = \delta^2_{X_i,Y_j} + \varepsilon^4.$$

Putting this in we have that

$$I(P_{t+1}) \geq I(P_t) + \sum_{\substack{i,j \\ (X_i,Y_j)\in P_t \\ \text{not } \varepsilon\text{-regular}}} \alpha_{X_i,Y_j}\varepsilon^4 \geq I(P_t) + \varepsilon^5.$$

Applying Fact 3.2 we know that the index of a partition can never be larger than that of the trivial partition, which by Fact 3.5 is δ. Since we started with an index of at least δ^2, we can refine our partition at most $(\delta - \delta^2)/\varepsilon^5$ times before our process must halt. But by construction our process only halts when we have an ε-regular partition, so after finitely many refinements we have a partition with at most εn^2 edges in ε-irregular pairs.

Any block in $|P_t|$ is involved in at most $|P_t|$ irregular pairs. In the worst case, the refinements necessary for each of these pairs are possibly distinct, and together divide the block into at most $2^{|P_t|}$ parts. Therefore we have that $|P_{t+1}| \leq |P_t| 2^{|P_t|}$. Since we know that we will have our ε-regular partition after at most $(\delta - \delta^2)/\varepsilon^5$ steps, letting $f(x) = x2^x$, we have that

$$M(\varepsilon, c) = f \circ f \circ f \circ \cdots \circ f(c) \approx 2^{2^{2^{\cdot^{\cdot^{\cdot^{2^c}}}}}}$$

as our bound, where the height of the tower and the number of times f is composed are both $(\delta - \delta^2)/\varepsilon^5 \leq 1/4\varepsilon^5$. □

For some problems, the value $M(\varepsilon, c)$ is still the best known upper bound.

3.5 A Problem on Induced Matchings

Let G be a graph on n vertices, and suppose that $E(G)$ is the union of n induced matchings (here a "matching" is a set of vertex-disjoint edges, we call a matching "induced" if there are no edges in G connecting vertices belonging to two different edges in the matching). How large can $E(G)$ be?

It is known that $|E(G)| = o(n^2)$, and a rough outline of the proof follows (a full proof is contained in the next lecture). Assume our graph has αn^2 vertices. Starting with a regular partition of our original graph, we construct a subgraph which still has many edges, but in which all the edges lie in regular pairs which have large intersections with some matching. The matchings themselves (or most of them, at least) have near 0 density since they only have $n/2$ edges, while at least some of the regular pairs should have positive edge density since the graph itself has positive density, which is a contradiction.

Use of the Regularity Lemma allows us to conclude that $E(G) = O(\frac{n^2}{\log^* n})$, where $\log^* n$ is the number of successive logs one must apply to n to get below 1 (a sort of inverse to the "tower of exponents" function). On the other hand, it is known [24] that a bound of n^α for any $\alpha < 2$ cannot be true.

Lecture 4

Applications of the Regularity Lemma

Ross Richardson

4.1 Induced Matching Problem

Our first application will prove to be the building block for the remaining two. For more detail, refer to [30] and [37]. Here we use the *Hungarian* method, that is, we throw away structure which is insignificant, thus allowing us greater control over the remaining controlling structure.

A subgraph of a graph G is a *matching* if every vertex has degree one. A matching M is an *induced matching* if all edges of G between the vertices of M are edges in M.

Theorem 4.1. *Let G be a graph on n vertices. If $E(G)$ is the union of n induced matchings, then*

$$e(G) = o(n^2).$$

Proof. We shall assume that $e(G) > \alpha n^2$, and derive a contradiction. Select $\varepsilon < \alpha/8$, and apply the Regularity Lemma (see Theorem 2.1). We then obtain a Szemerédi partition,

$$P = V_1 + \ldots + V_k,$$

and further we know the number of edges between $\varepsilon-$regular pairs is at least $\alpha n^2/4$.

Now, let $M_1, \dots, M_n$ be the matchings composing $E(G)$. We remove matchings of small size as follows: for each i, if $E(M_i) < \alpha n/2$, then we delete M_i. In this way we delete at most $\alpha n^2/2$ edges. Denote by G' the graph on the remaining edges. We note that $e(G') \geq \alpha n^2/4$.

We now modify our matchings. For each $1 \leq i \leq k, 1 \leq j \leq n$, if $|V_i \cap V(M_j)| \leq \alpha |V_i|/8$, then delete those edges in M_j incident to V_i. Observe that we delete at most $\sum_i \frac{\alpha}{8}|V_i| = \alpha n/8$ edges in M_j.

After this modification, we note that each (modified) matching has at least half of its original edges. Denote by G'' the graph formed by these modified matchings. Clearly, $e(G'') \geq \alpha n^2/8$.

For each edge remaining let (V_i, V_j) be the ε-regular pair containing it. Say M_l contains this edge. Put $A = V_i \cap V(M_l)$ and $B = V_j \cap V(M_l)$. Then we have that $|A| \geq \alpha|V_i|/8$ and similarly $|B| \geq \alpha|V_j|/8$. Now, since our pair (V_i, V_j) is regular, this means that in G

$$|\delta(V_i, V_j) - \delta(A, B)| \leq \varepsilon.$$

Now,

$$\delta(A, B) = \frac{e(A, B)}{|A||B|},$$

and since A and B are both in *the induced matching* M_l, then $e(A, B) \leq \min(|A|, |B|)$ (without loss of generality suppose $|A|$ is smallest). Hence,

$$\delta(A, B) \leq \frac{\min(|A|, |B|)}{|A||B|} \leq 1/|B| \leq (\alpha|V_j|/8)^{-1}.$$

But $|V_j|$ becomes arbitrarily large as n grows (follows from the regularity lemma), so this quantity is arbitrarily small. Thus, $|\delta(V_i, V_j)| \leq 2\varepsilon$ for sufficiently large n.

From this we can compute:

$$\begin{aligned} e(G'') &\leq \left\{ \begin{matrix} \#\text{ edges in} \\ \text{irregular pairs} \end{matrix} \right\} + \sum_{\substack{\varepsilon-\text{regular pairs} \\ (V_i, V_j)}} e(V_i, V_j) \\ &\leq \varepsilon n^2 + 2\varepsilon \sum_{1 \leq i < j \leq k} |V_i||V_j| < 3\varepsilon n^2. \end{aligned}$$

As $e(G'') \geq \alpha n^2/8$, this shows that α is arbitrarily small, hence our contradiction. □

Note that by "throwing away" all of the unnecessary parts the Szemerédi partition let us control what is left. The best that we can conclude about the number of edges from this approach is that it is bounded by $n^2/\log^*(n)$.

4.2 The $(6,3)$ problem

A *hypergraph* is denoted by $H = (V, \mathcal{E})$, where $\mathcal{E} \subset 2^V$. We call members of $\mathcal{E}$ *hyperedges*. A hypergraph is k-uniform if every $e \in \mathcal{E}$ has $|e| = k$. For a 3-uniform hypergraph (or 3-*graph*), we use the terminology *triples* as well as hyperedges.

The $(6,3)$ problem, conjectured by Erdős, demonstrates a simple local condition on 3-graphs which forces $e(H) = o(n^2)$, a significantly fewer number than for general 3-graphs.

Theorem 4.2. *If a 3-graph $H = (V, \mathcal{E})$ on n vertices contains no 6 points with at least 3 triples, then $e(H) = o(n^2)$.*

Proof. Note that with a 3-graph H there is an associated 2-graph where we include an edge $\{u, v\}$ if and only if there is some hyperedge e so that $\{u, v\} \subset e$. We will show that this associated 2-graph does not have many edges and so neither can H.

Fix a vertex v of degree at least 3. We create a matching M_v in our associated 2-graph by:

$$M_v := \{e - \{v\} \,|e \in \mathcal{E} \text{ and } v \in e\},$$

(note that $e - \{v\}$ is a 2-edge). If M_v has degree less than 3 let M_v be empty.

First,we show that this forms a matching. For vertices of degree at least 3, we claim that any two edges adjacent to v intersect only at v. If this were not the case, we would have the following situation (see Figure 4.1) which violates our $(6,3)$ condition. This then forces the edges of M_v to be disjoint (hence a matching).

This matching is also induced. To see this, say that it were not. Then there would be an edge in G between two endpoints in our matching. But because all edges come from triangles, we have the situation depicted in Figure 4.2. Again, we see that this violates the $(6,3)$ condition.

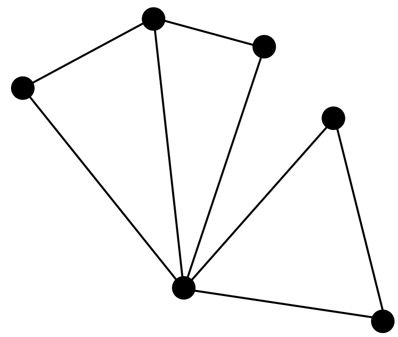

Figure 4.1: The $(6,3)$ configuration results if a degree 3 or greater vertex has hyperedges that intersect at two points.

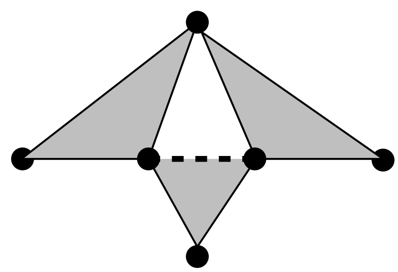

Figure 4.2: If our matching is not induced, then there will be an edge (dashed) between two vertices in our matching. Since that edge had to come from another triangle, we get the above configuration, which is $(6,3)$.

Let G be the graph formed by the union of these matchings. By our previous theorem, we know that $e(G) = o(n^2)$. But $e(H) \leq e(G) + 2n$ since every edge in G comes from an edge in H, and only the edges attached to degree at most two vertices in H don't (possibly) contribute edges. Thus $e(H) = o(n^2)$. □

4.3 Solymosi's proof of a result of Ajtai and Szemerédi

We give a simple proof due to Solymosi [39] of the following theorem of Ajtai and Szemerédi. The next set of notes show its connection to Roth's theorem.

Theorem 4.3. *Let $R \subset [N]^2$. For $\delta > 0$, there exists an N_0 such that if $|R| \geq \delta N^2$ and $N \geq N_0$, then there exists $(x, y), (x + d, y), (x, y + d) \in R$ for some integer $d \neq 0$.*

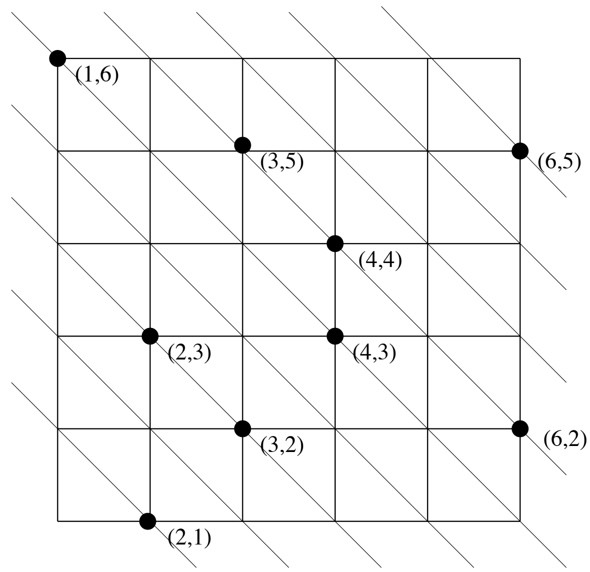

Figure 4.3: Viewing R as a subset of $[N]^2$. Note the diagonal lines denote equivalence sets (hence matchings). We see here $(3,2),(6,2),(3,5)$ form our desired triple.

Proof. We construct a bipartite graph G on the vertices $\{v_1,\ldots,v_n\}$ and $\{w_1,\ldots,w_n\}$. We set $(v_i,w_j)\in E(G)$ if and only if $(i,j)\in R$.

Now, create an equivalence class of edges where $(v_i,w_j)\sim(v_{i'},w_{j'})$ if and only if $i+j=i'+j'$. In other words, if we view the R as a subset of the grid $[N]^2$, then the equivalence classes are those members which lie on the same NW-SE diagonal (see Figure 4.3).

It is clear that each class is a matching. If N is large enough, the fact that $|R|$ is a positive fraction of N^2 implies that one of our matchings cannot be induced. Hence, we have two edges (v_i,w_j) and $(v_{i'},w_{j'})$ in the same matching with an edge connecting a vertex in each, say $(v_i,w_{j'})$. Now, $j'=j+d$, hence,

$$i+j=i'+(j+d)\to i'=i-d.$$

Hence, in R we have:

$$\begin{aligned}(i,j) &= (i,j)\\ (i-d,j+d) &= (i',j')\\ (i,j+d) &= (i,j').\end{aligned}$$

This is the configuration desired. ☐

Lecture 5

More Applications of the Regularity Lemma

D. Jacob Wildstrom

5.1 More Szemerédi Regularity Lemma Applications

5.1.1 Nonintersecting triangles

Theorem 5.1. *Given a graph G with n vertices, if each edge of G lies in a unique triangle (i.e., there are no edges which lie in 2 or more distinct K_3 subgraphs of G), then $e(G) = o(n^2)$.*

Exercise 5.1. Prove Theorem 5.1.

5.1.2 3-Term Arithmetic Progressions

Theorem 5.2 (Roth's Theorem [36]). *For every $\delta > 0$ there is an $N(\delta)$ so that for all $n > N(\delta)$ then if $A \subseteq [n]$ with $|A| > \delta n$, A contains 3 terms in arithmetic progression.*

Proof. We will make use of a result of Ajtai and Szemerédi (Theorem 4.3 from the previous lecture).

Let us define $R \subseteq [n]^2$ as follows: $(x, y) \in R$ if $x - y \in A$. Note that the elements of A correspond to the diagonals in the lower-left of R and so each element of A corresponds to between 1 and n elements of R. Even minimizing the number of elements in R, we get that $|R| > \frac{1}{2}(\delta n)^2 = \varepsilon n^2$. Thus, by the Ajtai-Szemerédi result using $\varepsilon = \frac{1}{2}\delta^2$, we know that for sufficiently large n that R contains a set of three points (x, y), $(x + d, y)$ and $(x, y + d)$. Since these points are in R, it follows that $x - y - d$, $x - y$ and $x + d - y$ are in A, and these three elements form an arithmetic progression of length three in A. □

By now, we've seen several $o(n)$ and $o(n^2)$ upper bounds on system sizes, but what about lower bounds? For most of these problems, we have the quite good lower bound of $\frac{n^2}{e^{c\sqrt{\log n}}}$, which exceeds $n^{2-\varepsilon}$ for all positive ε.

We shall construct such a lower bound for the previous question.

Theorem 5.3 (Behrend's Theorem [3]). *Let $r_3(n)$ be the maximum size of a subset of $[n]$ not containing a 3-term arithmetic progression. Then $r_3(n) > \frac{n}{e^{c\sqrt{\log n}}}$.*

Proof. We will construct such a system based on parameters m and d which we shall determine later. Let us consider a set $A \subseteq \{0, 1, \ldots, m-1\}^d$, with an addition operation defined termwise. There is no 3-term arithmetic progression in A if there is no solution to the equation $x + y = 2z$ for $x, y, z \in A$.

For a fixed t let $A_t = \{(x_1, x_2, \ldots, x_d) : x_1^2 + x_2^2 + \cdots + x_d^2 = t\}$, so A_t consists of the non-negative integer coordinates on the surface of a d-dimensional sphere of radius t. Note that every element of $\{0, 1, \ldots, m-1\}^d$ lies in some A_t for $0 \leq t \leq m^2 d$, so some A_t contains at least $m^d/m^2 d = m^{d-2}/d$ elements. From geometry we know that no three points on a sphere are colinear, so A_t has no three-term arithmetic progression.

Lastly, let $\phi : A_t \to \mathbb{Z}$ be given by $\phi(x) = \sum_{i=1}^{d} x_i (2m)^{i-1}$; essentially, $\phi(x)$ is the representation of x as a digit-string base $2m$, so $\phi(A_t) \subseteq \{0, \ldots, (2m)^d\}$. From the consideration of $\phi(x)$ as a digit-string, it is clear that $\phi(x) + \phi(y) = 2\phi(z)$ implies that $x + y = 2z$, from which it follows that $\phi(A_t)$ does not contain a 3-term arithmetic progression since A_t does not.

By construction $\phi(A_t)$ is a set of size m^{d-2}/d without a 3-term arithmetic progression drawn from a set of $(2m)^d$ consecutive integers; our lower bound is thus attained by maximizing m^{d-2}/d subject to the constraint $(2m)^d = n$, which we may rewrite as $\ln m = \frac{\ln n}{d} - \ln 2$. So we seek to maximize

$$\ln\left(\frac{m^{d-2}}{d}\right) = (d-2)\left(\frac{\ln n}{d} - \ln 2\right) - \ln d$$

whose derivative with respect to d is

$$\frac{2\ln n - d - d^2 \ln 2}{d^2},$$

so the maximizing value of d is a solution to the quadratic $d^2 \ln 2 + d - 2\ln n = 0$. An approximate solution (and thus an approximate maximization) is $d = \sqrt{\frac{2\log n}{\log 2}} = \sqrt{2\log_2 n}$. Then $m = n^{\frac{1}{d}}/2$, so

$$\frac{m^{d-2}}{d} = \frac{n^{\frac{d-2}{d}}}{2^{d-2}d} = n^{1-\frac{2}{n}-\frac{\log d}{\log n}-\frac{(d-2)\log 2}{\log n}},$$

which, for vary large n, can be bounded by $n^{1-\frac{2\sqrt{2\log 2}+\varepsilon}{\sqrt{\log n}}}$ for arbitrarily small ε. □

5.2 Regularity Generalized

Our bounds on sizes of sets avoiding three-term arithmetic progresion relied essentially on properties concerning pairs; a similar bound on avoiding four-term arithmetic progressions would require properies of triples. To generalize our results, we must generalize the Szemerédi Regularity Lemma. Doing so is not entirely straightforward. We proceed from considering graphs, that is, structures of vertices and vertex pairs, to hypergraphs, which are structures of vertices and vertex *sets*; i.e. hyper-edges are correlations among any number of vertices, instead of simply 2. We call a hypergraph k-uniform if each edge consists of k vertices. A graph would be considered a 2-graph. In order to extend our results on arithmetic progressions, we shall explore variations on the Regularity Lemma for 3-graphs and beyond.

A starting point is determining how the concepts of the Regularity Lemma extend to triples, and in so doing we immediately encounter the complexities added by our attempt to generalize. For instance, ε-regularity does not have an extension which is both intuitive and useful: while we might naïvely define it as a requirement on vertex-sets X, Y, and Z that the densities of triples among subsets A, B, and C differs only by ε from the overall density, this is not a property which always resembles randomness.

As an example, if we construct a random 2-graph G and induce the 3-graph H to have as edges all triples among which there are an even number of 2-edges in G (or alternatively among which there are 0 or 1 edges in G), we get hypergraphs with distributions unlike a random hypergraph generated by including or excluding each triple with a particular probability. We seek a property of graphs which will be true for a wide variety of random-hypergraph generation methods.

Lecture 6

Regularity Lemma for Hypergraphs

Steven Butler

In the previous lecture we saw how to use results from the regularity lemma to show that given $\delta > 0$ then for n sufficiently large any subset of $[n]$ containing at least δn elements has an arithmetic progression of length 3. To show a similar result holds for arithmetic progressions of length 4 we will need to generalize our methods to hypergraphs.

6.1 Hypergraphs

A hypergraph $H = (V, \mathcal{E})$ is is composed of V, a vertex set which is a collection of elements, and $\mathcal{E}$ which corresponds to our "edges". In a hypergraph, edges are subsets of the vertices. In this lecture our focus will be on 3-uniform hypergraphs (which we will sometimes denote as 3-graphs) in which all the edges are subsets of order 3, i.e.,

$$\mathcal{E} \subseteq \big\{\{x, y, z\} : x, y, z \in V \text{ and } x, y, z \text{ distinct}\big\}.$$

The graphs that we were introduced to in Section 1.2 would correspond to 2-graphs.

The hardest part of working with hypergraphs is developing an intuition about them. Our approach will be to define through analogy the terms which were used in the Regularity Lemma for 2-graphs, if done properly then our proof for the 2-case will apply to the more general case with hypergraphs with minimal change. We note that our approach is based on the work of Chung [13].

6.2 Density in Hypergraphs

In Section 2.1 we defined the edge density between two subsets of vertices. Because of the added structure in a hypergraph we have more flexibility in how to define density and we will examine two definitions of density for the 3-graphs.

6.2.1 $(3, 1)$ Density in Hypergraphs

Suppose that $X, Y, Z \subseteq V$ then we will let $e_H(X, Y, Z)$ denote the number of edges spanned by X, Y, Z, that is

$$e_H(X,Y,Z) = \big|\{(x,y,z) : \{x,y,z\} \in \mathcal{E}(H), x \in X, y \in Y, z \in Z\}\big|.$$

Note that by this convention if there is any triple (i.e., "edge" in our 3-graph) with vertices lying in $X \cap Y \cap Z$ that this will get counted 6 times.

Intuitively, the edge density should be the ratio of the number of edges present to the total number of edges possible. So let $\mathcal{K}$ denote the complete 3-graph (i.e., it has every possible triple in the graph). In the case X, Y, Z are all disjoint then $e_{\mathcal{K}}(X, Y, Z) = |X||Y||Z|$. We now define the density

$$\delta_{3,1}(X,Y,Z) = \frac{e_H(X,Y,Z)}{e_{\mathcal{K}}(X,Y,Z)}.$$

the $3, 1$ in the subscript tells us that we are working with 3-graphs and we are looking at the density from the 1-"dimensional" structure of the hypergraph (i.e., the density from the vertices).

With a definition of density we can now define what it means to be ε-regular. In Section 2.1 we considered subsets $A \subseteq X$ and $B \subseteq Y$ so that $|A| \geq \varepsilon|X|$ and $|B| \geq \varepsilon|Y|$. But looking through the material that followed the only time we used this condition was when we had $|A||B| \geq \varepsilon^2|X||Y|$, so we could have used this as the definition and then only have been off by an ε. This will be our approach for hypergraphs.

Definition 6.1. Given $X, Y, Z \subseteq V$. We say that (X, Y, Z) is $(3, 1)$-ε-regular if for all $A \subseteq X, B \subseteq Y, C \subseteq Z$ for which

$$e_{\mathcal{K}}(A,B,C) \geq \varepsilon e_{\mathcal{K}}(X,Y,Z),$$

satisfies

$$|\delta_{3,1}(A, B, C) - \delta_{3,1}(X, Y, Z)| < \varepsilon.$$

6.2.2 $(3, 2)$ Density in Hypergraphs

If $(3, 1)$ density used the one dimensional vertices then intuitively we would want $(3, 2)$ to use the two dimensional edges (i.e., 2-graphs) to define density. That is previously we used the natural intuition of induced triples on a vertex set (1 element subsets of V) and now we want to think about induced triples on pairs (2 element subsets of V).

So given G_1, G_2, G_3 (not necessarily disjoint) which are a set of 2-graphs on V then we will let $e_H(G_1, G_2, G_3)$ denote the number of edges induced by G_1, G_2, G_3 and define it as follows,

$$e_H(G_1, G_2, G_3) = \big|\{(x, y, z) : \{x, y, z\} \in \mathcal{E}(H), \\ \{x, y\} \in G_1, \{y, z\} \in G_2, \{z, x\} \in G_3\}\big|.$$

To count density we again want to consider the ratio of number of edges present over the number of edges possible. Again letting $\mathcal{K}$ denote the complete 3-graph let

$$e_{\mathcal{K}}(G_1, G_2, G_3) = \big|\{(x, y, z) : \{x, y, z\} \in \mathcal{E}(\mathcal{K}), \\ \{x, y\} \in G_1, \{y, z\} \in G_2, \{z, x\} \in G_3\}\big|.$$

Then we will define our $(3, 2)$ density by

$$\delta_{3,2}(G_1, G_2, G_3) = \frac{e_H(G_1, G_2, G_3)}{e_{\mathcal{K}}(G_1, G_2, G_3)}.$$

This leads to the definition of being $(3, 2)$-ε-regular.

Definition 6.2. Given the 2-graphs G_1, G_2, G_3. We say that (G_1, G_2, G_3) is $(3, 2)$-ε-regular if for all $T_1 \subseteq G_1, T_2 \subseteq G_2, T_3 \subseteq G_3$ for which

$$e_{\mathcal{K}}(T_1, T_2, T_3) \geq \varepsilon e_{\mathcal{K}}(G_1, G_2, G_3),$$

satisfies

$$|\delta_{3,2}(T_1, T_2, T_3) - \delta_{3,2}(G_1, G_2, G_3)| < \varepsilon.$$

6.2.3 Comparing $(3,1)$ with $(3,2)$

While the definitions for $(3,1)$-ε-regular and $(3,2)$-ε-regular are very similar to each other we note that a special case of $(3,2)$-ε-regular gives $(3,1)$-ε-regular, as shown in the following exercise.

Exercise 6.1. Show that for a given graph G, (X,Y,Z) is $(3,1)$-ε-regular if and only if $(X \times Y, Y \times Z, Z \times X)$ is $(3,2)$-ε-regular. Where by $X \times Y$ we mean the complete bipartite 2-graph which joining vertices of X to all vertices of Y.

It is possible for a graph to be ε-regular in the $(3,1)$ sense but not ε-regular in the $(3,2)$ sense. To see this, let G be a random 2-graph with the probability of including an edge as $1/2$, and construct a 3-graph, H, by including a triple in H if and only if 0 or 2 of the possible 2-subsets of the triple are in G.

Exercise 6.2. Show that the graph H constructed as above is ε-regular in the $(3,1)$ sense, i.e., show that every triple is in H with probability of $1/2$. Then show that this graph is not ε-regular in the $(3,2)$ sense. (Hint: let $G_1 = G_2 = G_3 = G$ when considering the $(3,2)$ case.)

6.2.4 Generalization to k-graphs

The move from $(3,1)$ to $(3,2)$ is a big leap. Once we get used to it the definitions will come very naturally and everything will fall into place.

For k-graphs we can similarly construct a chain of different measures of density, namely we can talk about a k-graph as being (k,r)-ε-regular for any value $1 \le r \le k-1$. Generally speaking these different measure of ε-regularity are not equivalent and the higher the value of r the "stronger" the measure of ε-regularity.

The different measures of ε-regularity give a grading to our space. This allows one to use the language of homology when talking about the Regularity Lemma for k-graphs. For our purposes now we will be satisfied with 3-graphs.

6.3 Regularity Lemma for Hypergraphs

By $\begin{bmatrix} V \\ k \end{bmatrix}$ we mean the collection of all k-subsets of V. As a special case, $\begin{bmatrix} V \\ 2 \end{bmatrix}$ will denote the collection of all 2-subsets, or edges in a complete 2-graph.

Definition 6.3. Given $H = (V, \mathcal{E})$, we say that a partition P of $\begin{bmatrix} V \\ 2 \end{bmatrix}$ is $(3,2)$-ε-regular if

$$\sum_{\substack{\mathcal{X}_i,\mathcal{X}_j,\mathcal{X}_k \in P \\ \text{not regular}}} e_{\mathcal{K}}(\mathcal{X}_i, \mathcal{X}_j, \mathcal{X}_k) < \varepsilon e_{\mathcal{K}}\left(\begin{bmatrix} V \\ 2 \end{bmatrix}, \begin{bmatrix} V \\ 2 \end{bmatrix}, \begin{bmatrix} V \\ 2 \end{bmatrix}\right).$$

That is, we want the irregular part to be small. Note we use $\mathcal{X}$ to emphasize that we are working with a partition of edges and not of vertices. If we wanted to talk about a partition being $(3,1)$-ε-regular we would use partitions of $\begin{bmatrix} V \\ 1 \end{bmatrix}$, and modify the rest of the definition similarly.

With this definition in hand we are ready to state the Regularity Lemma for hypergraphs, which is a direct analog of Theorem 3.1 in Lecture 3.

Theorem 6.1 (Regularity Lemma for Hypergraphs). *Given $\varepsilon > 0$ and $B > 1$. Then for a 3-graph H with a partition P of $\begin{bmatrix} V \\ 2 \end{bmatrix}$ where $|P| \leq B$, there is a refinement Q of P with is $(3,2)$-ε-regular and $|Q|$ is bounded above by a constant depending only on ε and B.*

The proof is analogous to that done for the Regularity Lemma for 2-graphs. Since we have introduced terminology in a way that directly corresponds to 2-graphs our proof will only be a short outline. The details of any omitted step can be found by examining the proof in Sections 3.3 and 3.4 by inserting the proper language.

First for a given partition P of the edge sets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ we define the index at $(\mathcal{X}, \mathcal{Y}, \mathcal{Z})$ by

$$I(P(\mathcal{X}, \mathcal{Y}, \mathcal{Z})) = \sum_{\substack{\mathcal{X}_i \subseteq \mathcal{X} \\ \mathcal{Y}_j \subseteq \mathcal{Y} \\ \mathcal{Z}_k \subseteq \mathcal{Z}}} \frac{e_{\mathcal{K}}(\mathcal{X}_i, \mathcal{Y}_j, \mathcal{Z}_k)}{e_{\mathcal{K}}(\mathcal{X}, \mathcal{Y}, \mathcal{Z})} \big(\delta_{3,2}(\mathcal{X}_i, \mathcal{Y}_j, \mathcal{Z}_k)\big)^2,$$

we then define $I(P) := I(P(\begin{bmatrix} V \\ 2 \end{bmatrix}, \begin{bmatrix} V \\ 2 \end{bmatrix}, \begin{bmatrix} V \\ 2 \end{bmatrix}))$ so that

$$I(P) = \sum_{\mathcal{X}_i, \mathcal{Y}_j \mathcal{Z}_k \subseteq \begin{bmatrix} V \\ 2 \end{bmatrix}} \frac{e_{\mathcal{K}}(\mathcal{X}_i, \mathcal{Y}_j, \mathcal{Z}_k)}{e_{\mathcal{K}}(\begin{bmatrix} V \\ 2 \end{bmatrix}, \begin{bmatrix} V \\ 2 \end{bmatrix}, \begin{bmatrix} V \\ 2 \end{bmatrix})} \big(\delta_{3,2}(\mathcal{X}_i, \mathcal{Y}_j, \mathcal{Z}_k)\big)^2.$$

Note that the first term $(e_{\mathcal{K}}(\mathcal{X}_i, \mathcal{Y}_j, \mathcal{Z}_k)/e_{\mathcal{K}}(\left[{V \atop 2}\right], \left[{V \atop 2}\right], \left[{V \atop 2}\right]))$ is the portion of the edges and sums to 1. The second term is the density, since that term is squared we see that we are in a perfect position to apply Cauchy-Schwarz, doing so gives us the following.

Fact 6.1. If P' is a refinement of P then $I(P') \geq I(P)$.

We also have the following facts.

Fact 6.2. If $(\mathcal{X}, \mathcal{Y}, \mathcal{Z})$ is not $(3, 2)$-ε-regular then there is a refinement P' (achieved by subdividing $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ into two pieces each) such that

$$I(P'(\mathcal{X}, \mathcal{Y}, \mathcal{Z})) \geq I(P(\mathcal{X}, \mathcal{Y}, \mathcal{Z})) + \varepsilon^3 = \delta_{3,2}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}) + \varepsilon^3.$$

Note when comparing this fact to the form given for the 2-graph version we see that the ε is now cubed instead of being raised to the fourth power. The difference is in how we defined ε-regular, as alluded to earlier.

Fact 6.3. If the partition P is just one block then $I(P) = \delta_{3,2}^2$. Alternatively, if the partition P is the individual edges then $I(P) = \delta_{3,2}$.

With these facts the proof of the hypergraph version of the Regularity Lemma proceeds as before. Namely we start with our partition of edges, if it is $(3, 2)$-ε-regular then we are done, otherwise using our Fact 6.2 we partition each of the triples of edge sets which are not $(3, 2)$-ε-regular and then repeat. Because at each stage we will increase the index by at least ε^4 then we can perform at most $(\delta_{3,2} - \delta_{3,2}^2)/\varepsilon^4$ refinements before having a $(3, 2)$-ε-regular partition. Finally we can count the maximum elements in the resultant partition by checking at each stage how much we have subdivided the elements. Giving us a tower of 2's similar to what we had before.

Lecture 7

Results on Sumsets
Lei Wu

This lecture was given by our guest lecturer Van Vu.

In this lecture, we will discuss many well-known results on sumsets.

7.1 Direct Problems

Definition 7.1. Let A be a finite set of integers, define the *2-fold sumset* $A + A = \{a + a'|a, a' \in A\}$.

Intuitively, we want to say that if the set has "nice" structure, then the sumset should be small:

Example 7.1. Let $A = \{1, \ldots, n\}$, then $A + A = \{2, \ldots, 2n\}$. Also, one can take an arithmetic progression instead of an interval as set A, and still obtain a small sumset (see Theorem 7.5). If $A \subseteq \{1, \ldots, n\}$ and $|A| \geq \delta n$, then $A + A \subseteq \{2, \ldots, 2n\}$ and trivially we will have

$$|A + A| < 2n \leq \frac{2}{\delta}|A|.$$

In fact, one can extend this observation to the d-dimensional case. For example, consider $A \subseteq \mathbb{Z}^d$, $A \subseteq Q$, where Q is a d-dimensional integer parallelepiped, then we have similar results as above. Now what if we consider mapping the parallelepiped to a line? The result is what we consider as the generalized arithmetic progression.

Definition 7.2. Let $a, q_1, \ldots, q_d$ be elements of an abelian group G, and let $l_1, \ldots, l_d$ be positive integers. The set

$$\begin{aligned} Q &= Q(a; q_1, \ldots, q_d; l_1, \ldots, l_d) \\ &= \{a + x_1 q_1 + \cdots + x_d q_d : 0 \leq x_i < l_i \quad \text{for} \quad i = 1, \ldots, d\} \end{aligned}$$

is called a *d-dimensional (Generalized) Arithmetic Progression (GAP for short)* in the group G. The *length* of Q is $l(Q) = l_1 \cdots l_d$. Clearly, $|Q| \leq l(Q)$, and Q is called *proper* if $|Q| = l(Q)$.

We can see that if the map from a d dimensional lattice to 1 dimension is one-to-one, then the GAP is proper. And this map is called Freiman isomorphism.

7.2 Inverse Problems

Now we consider problems in the other direction, namely, if we know the properties of sumset $A + A$, can we deduce anything about the structure of A? This is what we call the "inverse problem".

Theorem 7.1 (Freiman). *Let A be a finite set of integers such that $|A+A| \leq c|A|$. Then A is a large subset of some d-dimensional arithmetic progressions, i.e. there exist integers $a, q_1, \ldots, q_d, l_1, \ldots, l_d$ such that*

$$A \subseteq Q = \{a + x_1 q_1 + \cdots + x_d q_d : 0 \leq x_i < l_i \text{ for } i = 1, \ldots, d\},$$

where $|Q| \leq \delta|A|$, and d and δ depend only on c.

A generalization of this proof is given by Ruzsa.

Theorem 7.2 (Ruzsa). *Let c, c_1, c_2 be positive real numbers and $n \geq 1$. If A and B are finite subsets of a torsion-free abelian group such that*

$$c_1 n \leq |A|, |B| \leq c_2 n$$

and

$$|A + B| \leq cn,$$

then A is a subset of a d-dimensional arithmetic progression of length at most l, where d and l depend only on c, c_1, and c_2.

In fact, it was later proved that the GAP containing A can be made proper (e.g., see Chang [8] and Bilu [5]). Chang also improved the bounds of the dependency on c in Freiman's Theorem, namely to

$$d(c) \leq \lfloor c - 1 \rfloor, \qquad \delta(c) \leq \exp\{Kc^2(\ln c)^3\},$$

where K is an absolute constant.

Bilu[5] also gave a proof of Freiman's Theorem, which is close to Freiman's proof and has more combinatorial flavor.

Theorem 7.3. *Let $A \subseteq \mathbb{Z}$ and $|A+A| \leq c|A|$, then there exists a d-dimensional arithmetic progression Q such that*

$$d = \lfloor \log_2 c \rfloor,$$

and $A' \subseteq A$ such that

$$|A'| \geq \alpha|A|, \quad A' \subseteq Q,$$

and

$$|A'| \geq \delta|Q|,$$

where $\alpha = \alpha(c)$ and $\delta = \delta(c) = \exp\{-\exp(c)\}$.

7.3 Balog-Szemerédi

Suppose we only have partial information about the sumset, say, if a "large portion" of it is "under controll", then we can still deduce information about a large subset of the original set.

Theorem 7.4 (Balog-Szemerédi). *Suppose A is a subset of a finite abelian group with $|A| = n$, and G is a bipartite graph with color classes A and A, and*

$$|E(G)| \geq \epsilon n^2.$$

Let $A +_G A = \{a + a' | (a, a') \in E(G)\}$, and assume

$$|A +_G A| \leq c|A|,$$

then there exists a set $A' \subseteq A$ such that

$$|A'| \geq \alpha|A|,$$

and

$$|A' + A'| \leq k|A'|,$$

where $\alpha = \alpha(c, \epsilon)$ and $k = k(c, \epsilon)$.

The original proof of this theorem used the Regularity Lemma, hence the bound is a tower function in c. Tim Gowers improved this bound by obtaining a polynomial function in c in his proof of Szemerédi's Theorem for arithmetic progression of length 4. Just as the generalization of Freiman's Theorem by Ruzsa, this theorem could also be generalized to the case when we take two different sets A and B of roughly the same size. An open problem is that if $|A|$ is significantly smaller than $|B|$, say $|A| = n^{1/3}$ and $|B| = n$, $|A +_G B| \leq c|B|$, and $|E(G)| \geq \epsilon n^{4/3}$, then what can we say about the structure of (a large subset of) A and B?

7.4 Cauchy-Davenport

There are also a few problems concerning sumsets in congruence classes. First, recall two well-known theorems:

Theorem 7.5. *Let $A \subset \mathbb{Z}$, and $|A| = n$. Then*

$$|A + A| \geq 2n - 1,$$

and equality holds if and only if A is an arithmetic progression.

Theorem 7.6 (Cauchy-Davenport). *Let p be a prime number, and A, B be nonempty subsets of $\mathbb{Z}_p$, then*

$$|A + B| \geq \min\{p, |A| + |B| - 1\}.$$

Similar questions can be raised on distinct sumsets.

Definition 7.3. Let $A \subseteq \mathbb{Z}$, define *distinct 2-fold sumset*

$$A \overset{*}{+} A = \{a + a' | a, a' \in A, a \neq a'\}.$$

And we have the following counterparts of Theorem 7.5 and Theorem 7.6.

Theorem 7.7. *Let A be defined as in Theorem 7.5, then*

$$|A \overset{*}{+} A| \geq 2n - 3,$$

and equality holds if and only if A is an arithmetic progression.

Theorem 7.8 (Erdös-Heilbronn Conjecture). *Let A be as in Theorem 7.6, then*

$$|A \overset{*}{+} A| \geq \min\{2n - 3, p\}.$$

This conjecture was proved in the 90's by Dias da Silva and Hamidoune using representation theory and linear algebra, and later by Alon, Nathanson and Ruzsa through polynomial methods.

Note that all the problems in this section can be generalized to h-fold sumsets for all $h \geq 2$.

Lecture 8

Quasi-Random Hypergraphs

Paul Horn

The last lecture dealt primarily with introducing various results on sumsets. It is interesting to note that we can view sumsets in a graph theoretic light. In particular we can form a cyclic-like graph as follows. Let $A \subseteq [n]$. Then define G such that

$$\begin{aligned} V(G) &= [n], \\ E(G) &= \{\{i, j\} : i + j \pmod{n} \in A\}. \end{aligned}$$

A cyclic graph can be generated by a matrix of the form

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 1 & & & & 0 \end{bmatrix}$$

While the matrix representation of graphs that result from our above construction are not quite cyclic, they are of a similar form, namely

$$\begin{bmatrix} b_0 & b_1 & b_2 & \dots \\ b_1 & b_2 & \iddots & \\ b_2 & \iddots & & \\ \vdots & & & \end{bmatrix}$$

One advantage to these cases is that the structure provides a much shorter representation, instead of $\binom{n}{2}$ entries needed to represent a graph, only the first row in necessary. Also, these graphs have very nice eigenvalues, a nice fourier analysis, etc. So these sumsets on n-sets are much cleaner than arbitrary graphs.

However, when we restrict ourselves to these graphs, we lose some freedom, and thus it is necessary to look more generally at things.

We eventually want to return to a proof of the existence of a 4-term AP as an application of the hypergraph regularity lemma, but first we will turn to quasi-random hypergraphs.

8.1 Quasi-Random Hypergraphs

We begin by looking at quasi-random properties of 2-graphs, and we look to see which properties we can 'push up' to the more general hypergraph case. It is interesting to note that that some properties can be pushed up, but others cannot. As an example of this, consider the eigenvalue property. Eigenvalues control a lot of properties of graphs, but there is no analogue for hypergraphs. For a 2-graph, we can find a matrix representation of the graph (the adjacency matrix for instance) and find the characteristic polynomial in order to find the eigenvaules.

What would the analog be for a 3-uniform hypergraph? We can replace an $n \times n$ matrix with an $n \times n \times n$ matrix, where the (x, y, z) entry of our matrix is 1 if $(x, y, z) \in G$ and is 0 otherwise. While this is a nice generalazation of an adjacency matrix there is no known analogue to an eigenvalue in the $n \times n \times n$ case.

8.1.1 Properties of Quasi-Random Graphs

It is important to understand that the 'quasi-random' properties that we look at are not random at all, in fact they are completely deterministic. They are, however, *motivated* by randomness, and can be thought of as a measure of how random a particular graph is. Quasi-randomness is not a phenomenon that is exclusive to graph theory, however. For instance Chung and Graham [12] treat the subject of quasi-random subsets and sequences.

The following are some properties of quasi-random graphs. Let G be a graph on n vertices.
$\mathbf{P}_1$: G has $\geq (1+o(1))\frac{n^2}{4}$ edges, and $\leq (1+o(1))\frac{n^4}{16}$ C_4's.
$\mathbf{P}_2(s)$ (for a fixed $s \geq 4$): Each graph M on S vertices occurs in $(1+o(1))\frac{n^s}{2^{\binom{s}{2}}}$ times as an individual subgraph.
$\mathbf{P}_3$: For any subset $S \subseteq V(G)$, the number of edges spanned by S is $e(S) = \frac{|S|^2}{4} + o(n^2)$.
$\mathbf{P}_4$: For all but $o(n^2)$ pairs of vertices u and v, the number, $s(u,v)$ of vertices adjacent to both u and v is $s(u,v) = (1+o(1))\frac{n}{2}$.
$\mathbf{P}_5$: If A is the adjacency matrix of G, with eigenvalues $\lambda_1 \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$., then

$$\begin{aligned} \lambda_1 &= (1+o(1))\frac{n}{2}, \\ \lambda_2 &= o(n). \end{aligned}$$

Exercise 8.1. 1.) Show that $\mathbf{P}_2(s) \Leftarrow \mathbf{P}_2(s+1)$
2.) Show that $\mathbf{P}_5 \Rightarrow \mathbf{P}_3$
3.) Show that $\mathbf{P}_1 \Rightarrow \mathbf{P}_5$.

8.1.2 Notes on Properties of Quasi-Random Graphs

It is interesting to note that we have the following relations regarding $\mathbf{P}_2(S)$.

$$\begin{array}{ccccccccccc} \mathbf{P}_2(2) & \Leftarrow & \mathbf{P}_2(3) & \Leftarrow & \mathbf{P}_2(4) & \Leftarrow & \mathbf{P}_2(5) & \Leftarrow & \ldots & \mathbf{P}_2(s) \\ & \not\Rightarrow & & \not\Rightarrow & & \Rightarrow & & \Rightarrow & \Rightarrow & \end{array}$$

One can, for instance, see that $\mathbf{P}_2(2) \not\Rightarrow \mathbf{P}_2(3)$ by considering G being the disjoint union of two copies of $K_{n/2}$, which has the expected number of edges, but not the expected number of triangles.

With $\mathbf{P}_3$, note that for a small subset, the guess could be very wrong, hence the $o(n^2)$ error term. For instance, Ramsey theorey implies that G will have either an independent set or a clique of size $\log(n)$. Thus this error term is necessary for the property to be valid.

Note that for $\mathbf{P}_5$, we listed the eigenvalues of G as $\lambda_1 \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. This is a consequence of Perron and Frobenius [27], which implies that λ_1, the largest eigenvalue in absolute value, is positive. Also as noted above, since these are eigenvalue properties, they do not push through to the hypergraph case.

It is also interesting to note that although these conditions have certain relationships (see, for instance, the Exercise 8.1) they are in some sense very different. For instance, $\mathbf{P}_3$ is a global property, but $\mathbf{P}_4$ can be thought of as a neighborhood property. Also, some of the properties are more difficult to check than others. $\mathbf{P}_2$ is hard to verify, for instance, but $\mathbf{P}_5$ can be checked by computing the characteristic equation, which takes polynomial time in n. Likewise, $\mathbf{P}_1$ can be verified in n^4 steps. Thus the game is to find properties that are easy to verify, and estimate others, and use the relationships between properties for information on the graphs.

Finally, we note that these same properties can make sense in other configurations. For instance, these same basic properties make sense if we restrict to the family of cyclic graphs, or equivalently subsets of an n set. More details on this can be found in [14].

8.1.3 Properties for Quasi-Random Hypergraphs

We now work to see which properties can be generalized to k-graphs. Chung and Graham [15] examine this generalization.

First we note that $\mathbf{P}_2(s)$ is easy to generalize, as we can talk about hypergraphs on s vertices, and inspect a hypergraph for copies of a certain induced subhypergraph. Thus we can generalize it as the following k-graph property for a k-regular hypergraph on n vertices H.
$\mathbf{Q}_2(s)$: Each k-graph M on s vertices occurs in H

$$(1 + o(1))\frac{n^s}{2^{\binom{s}{k}}}$$

times.

We next set our sights at generalizing $\mathbf{P}_3$. We begin by considering the $k = 3$

case. In a previous lecture we looked at 3-regular hypergraphs induced on 2-graphs. Given a 3-graph H, and a 2-graph $S \subseteq \left[{n \atop 2}\right]$ we are concerned with looking for copies of K_3 within S. We would expect that there are about $\frac{1}{2}\left(\frac{|S|}{n^2}\right)\frac{n^3}{6}$ unordered triples in the graph, where $\frac{|S|}{n^2}$ is the probability that a particular edge is in S, and the constant $\frac{1}{2}$ comes from our view that H should look like a hypergraph with edge probability $\frac{1}{2}$.

Generalizing this to a k-graph we get that the expected number of hyperedges spanned by a $(k-1)$-graph is $\frac{1}{2}\frac{n^k}{k!}\left(\frac{|S|}{n^{k-1}}\right)^k$, and hence we get our generalized $\mathbf{P}_3$.

$\mathbf{Q}_3$: For any $(k-1)$-graph S, the number of edges of H spanned by S is

$$(1+o(1))\frac{n^k}{k!}\left(\frac{|S|}{n^{k-1}}\right)^k.$$

Lecture 9

Quasirandomness, Ramsey Graphs, and Explicit Constructions

Ross Richardson

We discuss graphs (specifically, graph families) with given properties, and then discuss the problem of explicitly describing graphs in such families.

9.1 Quasirandom Graphs

We begin by stating a few properties for a graph. We remark that these properties are satisfied by random graphs from $G(n, 1/2)$. Nonetheless, the graph properties that we describe do not depend on any input of "randomness". For the following, let G be a graph on n vertices.

- $\mathbf{P(s)}$ (s fixed and at least 4): For any fixed graph M on s vertices, its occurrence in G is $(1 + o(1))\frac{n^s}{2^{\binom{s}{2}}}$ as a labeled induced subgraph.
- $\mathbf{P_1}$: G has at least $(1 + o(1))\frac{n^2}{4}$ edges and less than $(1 + o(1))\frac{n^4}{16}$ labeled induced C_4's.
- $\mathbf{P_2}$: For any $S \subset V(G)$, $e(S) = \frac{|S|^2}{4} + o(n^2)$.

- $\mathbf{P_3}$: Let $S(u, v) = \{w : (w \sim u \text{ and } w \sim v) \text{ or } (w \nsim u \text{ and } w \nsim v)\}$. For almost all pairs $|S(u, v)| = (1 + o(1))\frac{n}{2}$.

As elaborated in the last notes, these (and other) properties are equivalent and form a family of checkable properties for a graph to be quasi-random. It is helpful to consider examples of graph families that satisfy the above properties. First of all, it is not hard to show random graphs (or almost all graphs) are quasi-random. It is often desirable to have "explicit constructions" instead of only existence proofs via probabilistic methods. By explicit constructions, we mean a succinct description (a polynomial algorithm, for example) for an infinite family of graphs which satisfy the desired property (properties).

Happily, there have been such explicit constructions of quasi-random graphs.

Example 9.1 (Paley Graphs). Let p be prime. Define the graph Q_p on vertices $\{1, \dots, p-1\}$ by the relation that $u \sim v$ if and only if $u - v$ is a quadratic residue modulo p.

We show that this graph satisfies property $\mathbf{P_3}$.

We shall need two facts about quadratic residues.

Lemma 9.1. *The number of quadratic residues in $\{0, \dots, p-1\}$ modulo p is $\frac{p-1}{2}$.*

Proof. This follows from the fact that for $x, y \in [p-1]$, $x^2 = y^2$ exactly when p divides one of $x + y$ or $x - y$. Hence, every quadratic residue corresponds to two distinct elements of $[p-1]$, yielding our equality. □

Lemma 9.2. *For $x, y \in [p-1]$, $\frac{x}{y}$ is a quadratic residue precisely when both x and y are or neither are.*

Proof. Observe that for a fixed quadratic residue x, $\frac{x}{y}$ is a quadratic residue exactly when y is. As the quotients $\frac{x}{y}$ are distinct for all $y \in [p-1]$, we use the above lemma to see that the number of $\frac{x}{y}$ which are not residues is $\frac{p-1}{2}$ (corresponding to y a non-residue).

Thus, if x is a non-residue, we know that $\frac{p-1}{2}$ of the quotients $\frac{x}{y}$ are non-residues, corresponding to the case when y is a residue. Thus, the remaining quotients are all residues, corresponding to the case when both x and y are non-residues. □

Thus, in particular, we see that Q_p is a $(p-1)/2$ regular graph. Further, for $u, v, w \in V(Q_p)$, $w \in S(u, v)$ if and only if $u - w$ and $v - w$ are both residues or non-residues. Hence, by our theorem $w \in S(u, v)$ if and only if $\frac{u-w}{v-w}$ is a residue. Thus, we need only count the number of times the quotient $\frac{u-w}{v-w}$ is a residue for w not equal to u or v, as this will yield the size of $S(u, v)$. As this quotient is unique for each w, and ranges over all values except 0 and 1, we see that $|S(u, v)| = \frac{p-1}{2} - 1$, corresponding to the obtained residues. But this is enough to see that the family Q_p satisfies $\mathbf{P_3}$.

We mention another construction for a quasi-random graph, which will generalize to the hypergraph case.

Example 9.2. Let G be a graph on $2n$ vertices such that $V(G) = \left[{2n \atop n}\right]$, that is, subsets of size n of the set $[2n]$. Let $U \sim V$ if and only if $|U \cap V| \equiv 0 \pmod 2$.

9.2 Ramsey Graphs

We can define $r(k)$ to be the number of vertices required for a graph to have either a k clique or a k independent set. Ramsey theory gives bounds

$$c\sqrt{2}^k \leq r(k) \leq c'4^k,$$

for constants c and c'.

Thus, we are assured that for n large there exist graphs of with no clique of size $c'' \log n$ and no independent set of size $c'' \log n$. We say such graphs have the *Ramsey Property*. It is of interest to find explicit constructions of such graphs (Erdős offered \$100 for such constructions). Such a construction however, has been elusive. Constructions such as the Paley graph, which provide concrete examples of many types of conditions (quasirandom, expander, etc.), fail to have this Ramsey Property. Indeed, though one can see that the Paley graph has clique size at most $c\sqrt{p}$, it was shown to contain cliques of size $c \log p \log \log \log p$ infinitely often.

The search for constructions of graphs with cliques and independent sets of size $n^{1/k}$ remains difficult. The current state of the art rests at $e^{\sqrt{\log n}}$.

Here, we list a graph shown by Frankl and Wilson to have clique and independent size at most $e^{c(\log n \log \log n)^{1/2}}$ (which forces $r(k) > k^{c \log k / \log \log k}$).

Example 9.3. Let q be a prime power. The graph G will have vertex set $V = \{F \subset [m] : |F| = q^2 - 1\}$ and edge set $E = \{(F, F') : |F \cap F'| \not\equiv -1$

(mod q)}. This graph contains no clique of size $\binom{m}{q-1}$. By setting $m = q^3$, we find a graph on $n = \binom{m}{q^2-1}$ vertices with no clique or independent set of the size claimed.

The point here is that such Ramsey graphs are somehow stricter measures of randomness than properties such as quasi-randomness, and thus it is not surprising that explicit construction of Ramsey graphs are harder to come by.

9.3 Hypergraphs

In the last lecture, we saw how to generalize the property $\mathbf{P}(s)$ to hypergraphs. We note that much as in the graph case, infinitely many of these properties are equivalent. More precisely, for the $k-$uniform case, we have the following implications:

$$\begin{array}{ccccccccccc} \mathbf{P}(k) & \Leftarrow & \mathbf{P}(k+1) & \Leftarrow & \dots & \Leftarrow & \mathbf{P}(2k) & \Leftarrow & \mathbf{P}(2k+1) & \dots \\ & \not\Rightarrow & & \not\Rightarrow & & \not\Rightarrow & & \Rightarrow & & \Rightarrow \end{array}$$

One question is whether we might generalize property $\mathbf{P_1}$, that is, whether we might find some hypergraph analogue for C_4. The answer to our question is yes, and such a graph is as follows.

We define the $k-$uniform hypergraph $\mathcal{O}_k$ by $V(\mathcal{O}_k) = \{v_{i,j} : 1 \leq i \leq k, 0 \leq j \leq 1\}$ with edges $\mathcal{E}(\mathcal{O}_k) = \{\{v_{1,i_1}, v_{2,i_2}, \dots, v_{k,i_k}\} : i_1, \dots, i_k \in \{0,1\}\}$. When $k = 3$ the resulting graph corresponds to the faces of an octahedron, hence its name.

Property $\mathbf{P_1}$ is then readily generalized using this graph.

9.3.1 Quasi-random Hypergraphs

We also demonstrate two quasi-random hypergraphs. The first is a generalization of the subset intersection graph we saw earlier.

Example 9.4 (Even Intersection Graph). Construct a $k-$uniform hypergraph with vertex set $V(H) = 2^{[n]}$ (that is the vertices consist of all possible subsets of $[n]$). Then the subsets $X_1, \dots, X_k$ form an edge if and only if $|X_1 \cap \dots \cap X_k| \equiv 0 \pmod 2$.

We may also generalize the Paley graph and obtain a quasi-random hypergraph, as follows:

Example 9.5. For p a prime, let $\{i_1, \ldots, i_k\}$ be an edge if and only if $i_1 + \ldots + i_k$ is a quadratic residue.

Lecture 10

Discrepancy of Graphs

Blair Angle

10.1 Some Preliminary Definitions

Definition 10.1. Given a k-uniform hypergraph $H = (V, \mathcal{E})$, define the indicator function $\mu_k : \binom{V}{k} \to \{-1, 1\}$ by

$$\mu_H(v_1, v_2, \ldots, v_k) := \begin{cases} -1 & \text{if } \{v_1, v_2, \ldots, v_k\} \in \mathcal{E}, \\ 1 & \text{otherwise.} \end{cases}$$

Note that in this definition, we are assuming that the v_i's are distinct. We will slightly modify this definition where we now allow $v_i = v_j$ for $i \neq j$.

Definition 10.2. Given a k-uniform hypergraph $H = (V, E)$, we define the function $\bar{\mu}_H : V^k \to \{-1, 1\}$ by

$$\bar{\mu}_H(v_1, v_2, \ldots, v_k) := \begin{cases} \mu_H(v_1, v_2, \ldots, v_k) & \text{if the } v_i \text{ are all distinct,} \\ 1 & \text{if } v_i = v_j \text{ for some } i \neq j. \end{cases}$$

Note that edges in H is given by $\mu_H^{-1}(-1) = \bar{\mu}_H^{-1}(-1)$, and so the number of edges is $|\mu_H^{-1}(-1)|$.

Definition 10.3. Given a k-uniform hypergraph $H = (V, \mathcal{E})$, let $\bar{H}$ denote the complement of H, the the k-uniform hypergraph with vertex set V and edge set $\binom{V}{k} - \mathcal{E}(H)$.

We remind the reader of the definition of an octahedron.

Definition 10.4. An octahedron, $\mathcal{O}_{2k}$, is a k-hypergraph with vertex set $V = \{v_i^{\varepsilon_i} : 1 \le i \le k, \varepsilon_i = 0, 1\}$ and edge set $\mathcal{E} = \big\{\{v_1^{\varepsilon_1}, \dots, v_k^{\varepsilon_k}\} : \varepsilon_i = 0 \text{ or } 1\big\}$.

When $k = 2$, $\mathcal{O}_4$ is C_4, the cycle on 4 vertices. The reader may want to verify many of the facts presented in the next section of this chapter in this special case (i.e., $k = 2$) before tackling general k.

10.2 The Deviation of a Hypergraph

10.2.1 Some Facts

We now define the deviation of a hypergraph.

Definition 10.5. Given a k-uniform hypergraph $H = (V, \mathcal{E})$, the deviation of H is given by

$$\operatorname{dev} H := \frac{1}{n^{2k}} \sum_{\substack{1 \le i \le k \\ v_i^0, v_i^1 \in V}} \prod_{\substack{\varepsilon_j \in \{0,1\} \\ 1 \le j \le k}} \bar{\mu}_H\big(v_1^{\varepsilon_1}, \dots, v_k^{\varepsilon_k}\big).$$

Given a set of $2k$ vertices we have that a term in the sum above is 1 if an even number of edges in $\mathcal{O}_{2k}$ is in H and gives -1 if an odd number of edges in $\mathcal{O}_{2k}$ is in H. So another way to calculate $\operatorname{dev} H$ is to take the number of "even octahedron" then take away the number of "odd octahedron", take the resulting value and divide by n^{2k}.

While the formula for deviation is somewhat ugly, its redeeming property is that it is straightforward to compute. One application of the deviation function is given by the following theorem:

Theorem 10.1. *Given a k-uniform hypergraph $H(n)$ on n vertices, for any fixed $G(t)$ (a k-hypergraph on t vertices), a bound on the number of labelled occurrences of $G(t)$ in $H(n)$ is given by:*

$$\left| \#\{G(t) < H(n)\} - \frac{n^{n(n-1)\cdots(n-t+1)}}{2^{\binom{t}{k}}} \right| \le 5n^t (devH(n))^{2^{-k}}.$$

Now we present several facts about the deviation of a k-uniform hypergraph H.

Fact 10.1. $\operatorname{dev} H = \operatorname{dev} \bar{H}$.

This follows immediately from the definition of deviation. Since an edge is in H if and only if it is not in $\bar{H}$. In changing from $\operatorname{dev} H$ to $\operatorname{dev} \bar{H}$ we are flipping $2k$ signs and so overall nothing changes (i.e, each term in the sum is a product of $2k$ values which go from 1 to -1, or go from -1 to 1). Another way to see this is to note that the number of faces in $\mathcal{O}_{2k}$ is even.

Fact 10.2. $0 \leq \operatorname{dev} H \leq 1$.

There are exactly n^{2k} ordered $2k$-tuples of the vertices of H (allowing repeats). So clearly

$$\left| \sum_{\substack{1 \leq i \leq k \\ v_i^0, v_i^1 \in V}} \prod_{\substack{\varepsilon_j \in \{0,1\} \\ 1 \leq j \leq k}} \bar{\mu}_H\big(v_1^{\varepsilon_1}, \ldots, v_k^{\varepsilon_k}\big) \right| \leq n^{2k}.$$

So with the normalization factor of $\frac{1}{n^{2k}}$ in the definition of deviation, it is clear that $\operatorname{dev} H \leq 1$. The other inequality is left as an exercise. One way this can be achieved is to show that the deviation of H can be expressed as

$$\frac{1}{n^{2k}} \sum_{\substack{1 \leq i \leq k-1 \\ v_i^0, v_i^1 \in V}} \left(\sum_{w \in V} \prod_{\varepsilon_1, \ldots, \varepsilon_{k-1}} \bar{\mu}_H\big(v_1^{\varepsilon_1}, \ldots, v_{k-1}^{\varepsilon_{k-1}}, w\big) \right)^2.$$

(Note: in deriving this expression, the reader may want to first start with the simplest octahedron, C_4.)

The neighborhood of a vertex in a k-hypergraph is a $(k-1)$-hypergraph. Given any vertex in a k-hypergraph H, define

$$H(v) := \text{ the neighborhood of } v.$$

Definition 10.6. The *sameness graph* of u and v denoted by $H(u) \nabla H(v)$ is a $(k-1)$-hypergraph with vertex set $H(u) \cup H(v)$ where e is an edge of the sameness graph if and only if e is either in both of $H(u)$ and $H(v)$, or e is in neither.

This leads us to our next fact:

Fact 10.3.

$$\operatorname{dev} H = \frac{1}{n^2} \sum_{u,v \in V} \operatorname{dev} \Big(H(u) \nabla H(v) \Big).$$

This is done by using the definition of deviation and the observation that $\mu_{H \triangledown H'} = -\mu_H \mu_{H'}$.

Definition 10.7. Given any 2-graphs G and H, the Cartesian Product $G \square H$ is the 2-graph with vertex set $V(G) \times V(H)$ and $(u, v) \sim (a, b)$ if and only if $u = a$ and $v \sim b$ in H, or $v = b$ and $u \sim a$ in G.

Note there are several ways to define the product of a graph, you only need to determine what happens on the four cases of (edge/no edge)×(edge/no edge). The advantage of the definition above is that it makes grids which simulate multidimensional space.

Fact 10.4. $\operatorname{dev} G \square H = (\operatorname{dev} G)(\operatorname{dev} H)$.

The proof of this is left as an exercise.

We can define the Cartesian product of 3-graphs similarly. That is we have vertex set $V(G) \times V(H)$ and $\{(u_1, v_1), (u_2, v_2), (u_3, v_3)\}$ is an edge if and only if $\big\{u_1 = u_2 = u_3$ and $\{v_1, v_2, v_3\} \in \mathcal{E}(H)\big\}$ or $\big\{v_1 = v_2 = v_3$ and $\{u_1, u_2, u_3\} \in \mathcal{E}(G)\big\}$.

10.2.2 Deviation and Discrepancy

A quantity related to discrepancy is deviation, which though deviation has a nicer looking formula, it is actually more difficult to compute than the deviation. In the next chapter, we will look at relationships between the deviation and discrepancy. We will prove the following theorem:

Theorem 10.2. *For any k-hypergraph H, $\mathit{disc}\, H \leq (\mathit{dev}\ H)^{2^{-k}}$ and $\mathit{dev}\ H \leq 4^k (\mathit{disc}\ H)^{2^{-k}}$.*

Lecture 11

Relating Deviation and Discrepancy for Hypergraphs, Part 1

Kevin Costello

11.1 Deviation

Let H be a k-regular hypergraph on n vertices (with k small relative to n). Equivalently, we can think of H as a function $\mu = \mu_H$ from the $\binom{n}{k}$ different k-element subsets of the vertex set V to the set $\{1, -1\}$, where $\mu_H(v_1, v_2, \ldots, v_k) = 1$ if and only if $\{v_1, \ldots, v_k\}$ is *not* an edge of H. We can extend μ to a function $\bar{\mu}$ on all ordered k-tuples of vertices by defining $\bar{\mu}(v_1, \ldots, v_k)$ to be 1 if any two of $v_1, \ldots, v_k$ are equal.

We define the **deviation** of H to be

$$\frac{1}{n^{2k}} \sum_{\substack{v_i^0, v_i^1 \\ i=1,\ldots,k}} \prod_{\substack{\varepsilon_i \in \{0,1\} \\ i=1,\ldots,k}} \bar{\mu}(v_1^{\varepsilon_1}, \ldots, v_k^{\varepsilon_k}).$$

11.2 Discrepancy

Let H be a k-regular hypergraph and G a $(k-1)$-regular hypergraph on the same set of vertices. Define $E(H; G)$ to be those edges in H "induced" by G, that is all edges in H such that dropping any vertex gives an edge of G. For example, in a standard simplex every element of the k-skeleton is induced by the $(k-1)$-skeleton. Let $e(H;G) = k!|E(H;G)|$ be the number of (labelled) $k-$edges in H induced from G.

We define the **discrepancy** on H to be

$$\frac{1}{n^k} \max_G |e(H;G) - e(\bar{H};G)|,$$

where the maximum is taken over all $(k-1)$-graphs on the vertex set of H. Here we are assuming that H has edge density approximately $1/2$, so we expect $e(H;G)$ and $e(\bar{H};G)$ to be roughly equal.

If we take G to be the complete $(k-1)-$graph on the vertices of H, we get that the discrepancy is bounded below by $2|\delta(H) - 1/2|$ (where $\delta(H)$ is the edge density of H). Similarly, we can use the discrepancy to get a bound on the edge density of any large subset of the vertex set.

Although the discrepancy of H has many useful properties, it is difficult to compute since there are many possible choices for G. Our eventual goal is to relate the discrepancy to the (more computable) deviation. Specifically, we aim to show that for any H, $\operatorname{disc} H \leq (\operatorname{dev} H)^{2^{-k}}$ and $\operatorname{dev} H \leq 4^k (\operatorname{disc} H)^{2^{-k}}$.

A corollary to this result is that if either the deviation or the discrepancy of H is $o(1)$, the other must be $o(1)$ as well.

11.3 The l-deviation of a Graph and its properties

Define the l-deviation of H as follows:

$$\operatorname{dev}_l H = \frac{1}{n^{k+l}} \sum_{\substack{v_i^0, v_i^1 \\ i=1,\dots,l}} \sum_{\substack{v_j \\ j=l+1,\dots,k}} \prod_{\substack{\varepsilon_1,\dots,\varepsilon_l \\ \varepsilon_i \in \{0,1\}}} \bar{\mu}(v_1^{\varepsilon_1}, \dots, v_l^{\varepsilon_l}, v_{l+1}, \dots, v_k).$$

Note that we have $\operatorname{dev}_k H = \operatorname{dev} H$ and $\operatorname{dev}_0 H = \frac{1}{n^k} e(V, V, \dots, V) = \delta(H)$.

Theorem 11.1. *We have* $\operatorname{dev}_l H \geq (\operatorname{dev}_{l-1} H)^2$. *In particular,* $\operatorname{dev}_l H$, *(and thus* $\operatorname{dev} H$*) is always nonnegative.*

Proof. First, we split up the innermost product according to whether $\varepsilon_l = 0$ or $\varepsilon_l = 1$, as shown below.

$$\prod_{\substack{\varepsilon_1,\ldots,\varepsilon_l \\ \varepsilon_i \in \{0,1\}}} \bar{\mu}(v_1^{\varepsilon_1}, \ldots, v_{l-1}^{\varepsilon_{l-1}}, v_l^{\varepsilon_l}, v_{l=1}, \ldots, v_k) =$$

$$\Big(\prod \bar{\mu}(v_1^{\varepsilon_1}, \ldots, v_{l-1}^{\varepsilon_{l-1}}, v_l^0, v_{l+1}, \ldots, v_k)\Big)\times$$
$$\Big(\prod \bar{\mu}(v_1^{\varepsilon_1}, \ldots, v_{l-1}^{\varepsilon_{l-1}}, v_l^1, v_{l+1}, \ldots, v_k)\Big)$$

Using this we can rewrite our sum for $\operatorname{dev}_l H$ as

$$\sum_{\substack{v_i^0, v_i^1 \\ i=1,\ldots,l-1}} \sum_{\substack{v_j \\ j=l+1,\ldots,k}} \left(\sum_{v_l^0} \sum_{v_l^1} \left(\begin{array}{c} \prod \bar{\mu}(v_1^{\varepsilon_1}, \ldots, v_{l-1}^{\varepsilon_{l-1}}, v_l^0, v_{l+1}, \ldots, v_k) \times \\ \prod \bar{\mu}(v_1^{\varepsilon_1}, \ldots, v_l^{\varepsilon_{l-1}}, v_l^1, v_{l+1}, \ldots, v_k) \end{array} \right) \right).$$

The innermost two sums are independent and we can factor according to

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j = (\sum_{i=1}^{n} a_i)(\sum_{j=1}^{m} b_j).$$

The two terms in the product of the right side of this identity correspond to taking $\varepsilon_l = 0$ and $\varepsilon_l = 1$, which are equal by symmetry. Thus we are left with:

$$\sum_{\substack{v_i^0, v_i^1 \\ i=1,\ldots,l-1}} \sum_{\substack{v_j \\ j=l+1,\ldots,k}} \left(\sum_{v_l} \Big(\prod_{\substack{\varepsilon_1,\ldots,\varepsilon_{l-1} \\ \varepsilon_i \in \{0,1\}}} \bar{\mu}(v_1^{\varepsilon_1}, \ldots, v_{l-1}^{\varepsilon_{l-1}}, v_l, v_{l+1}, \ldots, v_k)\Big) \right)^2.$$

We now apply the Cauchy-Schwartz inequality ($\sum_{i=1}^m a_i^2 \geq (\sum_{i=1}^m a_i)^2/m$) with a_j equal to the innermost sum in the above expression. Here m is equal to the number of summands, $n^{2(l-1)+(k-l)}$. Doing so gives us that

$$\begin{aligned} \operatorname{dev}_l H &\geq \frac{\left(\sum\limits_{\substack{v_i^0, v_i^1 \\ i=1,\ldots,l-1}} \sum\limits_{\substack{v_j \\ j=l,\ldots,k}} \prod\limits_{\substack{\varepsilon_1,\ldots,\varepsilon_{l-1} \\ \varepsilon_i \in \{0,1\}}} \bar{\mu}(v_1^{\varepsilon_1}, \ldots, v_{l-1}^{\varepsilon_{l-1}}, v_l, \ldots, v_k) \right)^2}{n^{k+l} n^{2(l-1)+(k-1)}} \\ &= \frac{(n^{k+l-1} \operatorname{dev}_{l-1} H)^2}{n^{k+l+2(l-1)+(k-l)}} = (\operatorname{dev}_{l-1} H)^2. \end{aligned}$$

□

Lecture 12

Relating Deviation and Discrepancy for Hypergraphs, Part 2

Daniel Felix

12.1 Discrepancy and Deviation

Let $H = (V, \mu_{_H})$ be a k-uniform hypergraph. Recall the following two definitions.

Definition 12.1. The *discrepancy* of H is defined as

$$\operatorname{disc} H = n^{-k} \max_{\substack{G \\ (k-1)-\text{graph}}} |e(H;G) - e(\bar{H};G)|.$$

It is understood that each of the $(k-1)$-uniform hypergraphs satisfies $V(G) \subset V$. By convention, no edge of G contains a repeated vertex.

Definition 12.2. The *deviation* of H is defined as

$$\operatorname{dev} H = n^{-2k} \sum_{\substack{v_i(0), v_i(1) \\ 1 \le i \le k}} \prod_{\substack{\varepsilon_j \in \{0,1\} \\ 1 \le j \le k}} \bar{\mu}_H(v_1(\varepsilon_1), \ldots, v_k(\varepsilon_k)).$$

(In previous lectures we wrote this as $(v_1^{\varepsilon_1}, \ldots, v_k^{\varepsilon_k})$.) We will usually express this more compactly as

$$\operatorname{dev} H = n^{-2k} \sum_{\vec{v}} \prod_{\vec{\varepsilon}} \bar{\mu}_H(\vec{v}(\vec{\varepsilon})), \tag{12.1}$$

where the sum is taken over all $\vec{v} = (v_1(0), v_1(1), \ldots, v_k(0), v_k(1))$ in V^{2k} and the product is taken over all $\vec{\varepsilon} \in \{0,1\}^k$.

We will show that these two quantities are closely related, and that one can be small only if the other is as well. We state this more precisely as a theorem.

12.2 Connection Between Discrepancy and Deviation

Theorem 12.1. *For any k-graph $H = (V, \mu_H)$, the following inequalities always hold:*

(i) $\operatorname{disc} H \leq (\operatorname{dev} H)^{2^{-k}}$,

(ii) $\operatorname{dev} H \leq 4^k (\operatorname{disc} H)^{2^{-k}}$.

This is theorem 8.1 in [12], which gives a thorough treatment of many aspects of deviation.

We will only prove the first inequality in this lecture, and before doing so it will be convenient to introduce some notation. We will use $\vec{v}(i)$ to denote the vector in which the $v_i(0)$ and $v_i(1)$ coordinates have been removed from $\vec{v}$:

$$v_i(i) = (v_1(0), v_1(1), \ldots, v_{i-1}(0), v_{i-1}(1), v_{i+1}(0), v_{i+1}(1), \ldots, v_k(0), v_k(1)).$$

Similarly, $\vec{\varepsilon}(i)$ will denote the vector

$$(\varepsilon_1, \ldots, \varepsilon_{i-1}, \varepsilon_{i+1}, \ldots, \varepsilon_k).$$

We let $\vec{v}(\vec{\varepsilon}(i); u)$ denote the vector acquired by removing the ith entry of $\vec{v}(\vec{\varepsilon})$ and replacing it with the vertex u. That is,

$$\vec{v}(\vec{\varepsilon}(i); u) = (v_1(\varepsilon_1), \ldots, v_{i-1}(\varepsilon_{i-1}), u, v_{i+1}(\varepsilon_{i+1}), \ldots, v_k(\varepsilon_k)).$$

Lastly, $\xi(G)$ will denote the edge set of the hypergraph G.

Proof of (i). First consider the case $k \geq 3$, and let G be a $(k-1)$-uniform hypergraph with no "loops" (edges containing a repeated vertex) and which satisfies $V(G) \subset V$. Notice that by definition

$$
\begin{aligned}
n^{2k}\text{dev}\, H &= \sum_{\vec{v}} \prod_{\vec{\varepsilon}} \bar{\mu}_{H}(\vec{v}(\vec{\varepsilon})) \\
&= \sum_{\vec{v}(1)} \left(\sum_{v_1(0), v_1(1)} \prod_{\vec{\varepsilon}(1)} \bar{\mu}_H(\vec{v}(\vec{\varepsilon}(1); v_1(0))) \bar{\mu}_H(\vec{v}(\vec{\varepsilon}(1); v_1(1))) \right) \\
&= \sum_{\vec{v}(1)} \left(\sum_{u} \prod_{\vec{\varepsilon}(1)} \bar{\mu}_H(\vec{v}(\vec{\varepsilon}(1); u)) \right)^2 \\
&\geq \sum_{\vec{v}(1) \subset \xi(G)} \left(\sum_{u} \prod_{\vec{\varepsilon}(1)} \bar{\mu}_H(\vec{v}(\vec{\varepsilon}(1); u)) \right)^2 . \qquad (12.2)
\end{aligned}
$$

We pause here for a clarification. Clearly $\vec{v}(1)$ is an element of V^{2k-2}, which itself is not a subset of $\xi(G)$. When we write $\vec{v}(1) \subset \xi(G)$ we mean that the above sum is to be taken over all vectors $\vec{v}(1)$ such that $\vec{v}(1)(\vec{\varepsilon}(1))$ is an edge of G for all $\vec{\varepsilon}(1) \in \{0,1\}^{k-1}$. The expression $\vec{v}(1)(\vec{\varepsilon}(1))$ should be interpreted in the same way as $\vec{v}(\vec{\varepsilon})$ in equation (12.1); that is,

$$
\vec{v}(1)(\vec{\varepsilon}(1)) = (v_2(\varepsilon_2), \ldots, v_k(\varepsilon_k)).
$$

By expanding the squared term of (12.2), we find that

$$
n^{2k}\text{dev}\, H \geq \sum_{\substack{\vec{v} \\ \vec{v}(1) \subset \xi(G)}} \prod_{\vec{\varepsilon}} \bar{\mu}_{H}(\vec{v}(\vec{\varepsilon})).
$$

Assuming that

$$
n^{2k}\text{dev}\, H \geq \sum_{\substack{\vec{v} \\ \vec{v}(i) \subset \xi(G) \\ 1 \leq i \leq j}} \prod_{\vec{\varepsilon}} \bar{\mu}_{H}(\vec{v}(\vec{\varepsilon})),
$$

where j is fixed, we have

$$
\begin{aligned}
n^{2k}\text{dev}\, H &\geq \sum_{\substack{\vec{v} \\ \vec{v}(i) \subset \xi(G) \\ 1 \leq i \leq j}} \prod_{\vec{\varepsilon}} \bar{\mu}_{H}(\vec{v}(\vec{\varepsilon})) \\
&= \sum_{\vec{v}(j+1)} \sum_{v_{j+1}(0) \in S_0} \sum_{v_{j+1}(1) \in S_1} \prod_{\substack{\vec{\varepsilon}(j+1) \\ l=0,1}} \bar{\mu}_H(\vec{v}(\vec{\varepsilon}(j+1); v_{j+1}(l))). \qquad (12.3)
\end{aligned}
$$

The sets S_0 and S_1 are another piece of notational shorthand. Our sum must contain the information that $\vec{v}(i) \subset \xi(G)$ for all $1 \leq i \leq j$, so we use S_0 and S_1 to represent that the choices of $v_{j+1}(0)$ and $v_{j+1}(1)$ are restricted. If, for example, no choice of $v_{j+1}(1)$ can be made that fulfills the requirements, then S_1 is empty. Notice that in principle S_0 does depend on the previously chosen vertices $v_1(0), \dots, v_j(1), v_{j+2}(0), \dots, v_k(1)$, while S_1 depends on these vertices as well as $v_{j+1}(0)$.

Our first observation is that S_1 does not actually depend on $v_{j+1}(0)$ because in no choice of $\vec{\varepsilon}$ does the vector $\vec{v}(\vec{\varepsilon})$ contain both of the vertices $v_{j+1}(0)$ and $v_{j+1}(1)$. Secondly, if u and v are valid choices for $v_{j+1}(0)$ and $v_{j+1}(1)$, then surely we can interchange the two vertices to obtain another valid pair. Hence $S_0 = S_1$, and $v_{j+1}(0)$ and $v_{j+1}(1)$ can be chosen arbitrarily from this set. We may then rewrite (12.3) as follows:

$$\begin{aligned} n^{2k} \operatorname{dev} H &\geq \sum_{\vec{v}(j+1)} \left(\sum_{u \in S_0} \prod_{\vec{\varepsilon}(j+1)} \bar{\mu}_H(\vec{v}(\vec{\varepsilon}(j+1); u)) \right)^2 \\ &\geq \sum_{\vec{v}(j+1) \in \xi(G)} \left(\sum_{u \in S_0} \prod_{\vec{\varepsilon}(j+1)} \bar{\mu}_H(\vec{v}(\vec{\varepsilon}(j+1); u)) \right)^2 . \end{aligned}$$

Expanding the resulting sum yields

$$n^{2k} \operatorname{dev} H \geq \sum_{\substack{\vec{v} \\ \vec{v}(i) \subset \xi(G)\ 1 \leq i \leq j+1}} \prod_{\vec{\varepsilon}} \bar{\mu}_H(\vec{v}(\vec{\varepsilon})).$$

Therefore, by induction,

$$n^{2k} \operatorname{dev} H \geq \sum_{\substack{\vec{v} \\ \vec{v}(i) \subset \xi(G)\ 1 \leq i \leq k}} \prod_{\vec{\varepsilon}} \bar{\mu}_H(\vec{v}(\vec{\varepsilon})). \tag{12.4}$$

Note that we may rewrite the condition $\vec{v}(i) \subset \xi(G)$ for all i as $\vec{v} \subset \xi(K, G)$, where K denotes the complete k-uniform hypergraph on V.

The next step is to apply the Cauchy-Schwarz inequality to (12.4) k times. We prove by induction on t that

$$n^{2k} \operatorname{dev} H \geq \prod_{l=1}^{t} \left(\frac{1}{n^{2k-l-1}} \right)^{2^{t-1}} \left(\sum_{u_1, \dots, u_t} \sum_{\substack{v_i(0), v_i(1) \\ i > t \\ \vec{v} \subset \xi(K,G)}} \prod_{\substack{\varepsilon_j \\ j > t}} \bar{\mu}_H(\vec{u}, \vec{v}(\vec{\varepsilon})) \right)^{2^t} \tag{12.5}$$

for all $1 \leq t \leq k$. The expression $\bar{\mu}_H(\vec{u}, \vec{v}(\vec{\varepsilon}))$ is used in place of the more cumbersome

$$\bar{\mu}_H(u_1, \dots, u_t, v_{t+1}(\varepsilon_{t+1}), \dots, v_k(\varepsilon_k)).$$

Our intention is for the value of t in such an expression to be inferred from the ranges of the above sums and products.

Rewriting (12.3) yields

$$n^{2k}\text{dev}\, H \geq \sum_{\substack{v_i(0),v_i(1)\\ i>1}} \left(\sum_{\substack{u_1\\ \vec{v}\subset\xi(K,G)}} \prod_{\substack{\varepsilon_j\\ j>1}} \bar{\mu}_H(u_1, v_2(\varepsilon_2), \ldots, v_k(\varepsilon_k)) \right)^2$$

Compare this to (12.2). The two are nearly identical; only the condition $\vec{v} \subset \xi(K,G)$ separates them. We now apply a version of the Cauchy-Schwarz inequality

$$\sum_{i+1}^{n} a_i^2 \geq \frac{1}{n}\left(\sum_{i+1}^{n} a_i\right)^2$$

to obtain that

$$n^{2k}\text{dev}\, H \geq \frac{1}{n^{2k-2}}\left(\sum_{\substack{v_i(0),v_i(1)\\ i>1}} \sum_{\substack{u_1\\ \vec{v}\subset\xi(K,G)}} \prod_{\substack{\varepsilon_j\\ j>1}} \bar{\mu}_H(u_1, v_2(\varepsilon_2), \ldots, v_k(\varepsilon_k)) \right)^2$$

which is exactly (12.5) for $t = 1$. Now assume that (12.5) is true for $t = 1, \ldots, t_0$ and consider the sum

$$\begin{aligned}
&\sum_{u_1,\ldots,u_{t_0}} \sum_{\substack{v_i(0),v_i(1)\\ i>t_0\\ \vec{v}\subset\xi(K,G)}} \prod_{\substack{\varepsilon_j\\ j>t_0}} \bar{\mu}_H(\vec{u}, \vec{v}(\vec{\varepsilon}))\\
&= \sum_{u_1,\ldots,u_{t_0}} \sum_{\substack{v_i(0),v_i(1)\\ i>t_0\\ \vec{v}\subset\xi(K,G)}} \prod_{\substack{\varepsilon_j\\ j>t_0}} \bar{\mu}_H(u_1, \ldots, u_{t_0}, v_{t_0+1}(\varepsilon_{t_0+1}), \ldots, v_k(\varepsilon_k))\\
&= \sum_{u_1,\ldots,u_{t_0}} \sum_{\substack{v_i(0),v_i(1)\\ i>t_0+1}} \left(\sum_{\substack{u_{t_0+1}\\ \vec{v}\subset\xi(K,G)}} \prod_{\substack{\varepsilon_j\\ j>t_0+1}} \bar{\mu}_H(u_1, \ldots, u_{t_0+1}, v_{t_0+2}(\varepsilon_{t_0+2}), \ldots, v_k(\varepsilon_k)) \right)^2\\
&\geq \frac{1}{n^{2k-(t_0+1)-1}}\left(\sum_{u_1,\ldots,u_{t_0+1}} \sum_{\substack{v_i(0),v_i(1)\\ i>t_0+1\\ \vec{v}\subset\xi(K,G)}} \prod_{\substack{\varepsilon_j\\ j>t_0+1}} \bar{\mu}_H(u_1, \ldots, u_{t_0+1}, v_{t_0+2}(\varepsilon_{t_0+2}), \ldots, v_k(\varepsilon_k)) \right)^2\\
&= \frac{1}{n^{2k-(t_0+1)-1}}\left(\sum_{u_1,\ldots,u_{t_0+1}} \sum_{\substack{v_i(0),v_i(1)\\ i>t_0+1\\ \vec{v}\subset\xi(K,G)}} \prod_{\substack{\varepsilon_j\\ j>t_0+1}} \bar{\mu}_H(\vec{u}, \vec{v}(\vec{\varepsilon})) \right)^2.
\end{aligned}$$

Substituting this expression into (12.5) when $t = t_0$ gives

$$n^{2k}\text{dev}\, H \geq \prod_{l=1}^{t_0+1}\left(\frac{1}{n^{2k-l-1}}\right)^{2^{(t_0+1)-1}} \left(\sum_{u_1,\ldots,u_{t_0+1}} \sum_{\substack{v_i(0),v_i(1)\\ i>t_0+1\\ \vec{v}\subset\xi(K,G)}} \prod_{\substack{\varepsilon_j\\ j>t_0+1}} \bar{\mu}_H(\vec{u}, \vec{v}(\vec{\varepsilon})) \right)^{2^{t_0+1}}.$$

Hence, by induction, (12.5) is true for all $1 \le t \le k$. For $t = k$ this yields

$$n^{2k} \operatorname{dev} H \ge \frac{1}{n^{k(2^k-2)}} \Bigg(\sum_{\substack{u_1,\ldots,u_k \\ \vec{v} \subset \xi(K,G)}} \bar{\mu}_H(u_1, \ldots, u_k) \Bigg)^{2^k},$$

or, equivalently,

$$\operatorname{dev} H \ge \Bigg(\frac{1}{n^k} \sum_{\substack{u_1,\ldots,u_k \\ \vec{v} \subset \xi(K;G)}} \bar{\mu}_H(u_1, \ldots, u_k) \Bigg)^{2^k}. \tag{12.6}$$

Since no edge of G contains a repeated vertex, we know that no $(k-1)$-subset of $\{u_1, \ldots, u_k\}$ can contain a repeated vertex for any such vector in the index of summation of (12.6). Therefore, since $k \ge 3$, no two vertices u_i and u_j in such a vector $(u_1, \ldots, u_k)$ can be equal since they are both contained in some $(k-1)$-subset of $\{u_1, \ldots, u_k\}$. Hence

$$\begin{aligned} \Bigg(\frac{1}{n^k} \sum_{\substack{u_1,\ldots,u_k \\ \vec{v} \subset \xi(K,G)}} \bar{\mu}_H(u_1, \ldots, u_k) \Bigg)^{2^k} &= \Bigg(\frac{1}{n^k} \sum_{\substack{u_1,\ldots,u_k \\ \vec{v} \subset \xi(K,G)}} \mu_H(u_1, \ldots, u_k) \Bigg)^{2^k} \\ &= \Bigg(\frac{1}{n^k} |e(H;G) - e(\bar{H};G)| \Bigg)^{2^k}. \end{aligned}$$

Since this is true for an arbitrary G, the inequality still holds when we take the maximum of the right hand side over all $(k-1)$-hypergraphs. Taking 2^kth roots of both sides finally gives

$$(\operatorname{dev} H)^{2^{-k}} \ge \operatorname{disc} H.$$

For $k = 2$ most of our above work still holds so that

$$\operatorname{dev} H \ge \Bigg(\frac{1}{n^2} \sum_{\substack{u_1,u_2 \\ u_1,u_2 \in G}} \bar{\mu}_H(u_1, u_2) \Bigg)^4.$$

Unfortunately, we cannot conclude that u_1 and u_2 must be distinct, as in the $k \ge 3$ case. Since $|V| = n$ the most we can currently say is

$$\operatorname{disc} H \le (\operatorname{dev} H)^{\frac{1}{4}} + O(1/n).$$

□

Lecture 13

4-term Arithmetic Progressions

Paul Horn

We now present a proof of the existence of a 4-term arithmetic progression using the Hypergraph Regularity Lemma. This version is a cleaned up version of a proof of Gowers [23]. Note that this proof can be generalized, but the 4-term proof gives a good representation of the complexity and workings of the proof.

Recall that we previously proved Roth's theorem [36] using the regularity lemma for 2-graphs, and in particular using a result of Ajtai and Szemerédi.

13.1 A generalization of a result of Ajtai and Szemerédi

Our proof of Roth's Theorem used a variation on the result of Ajtai and Szemerédi [39] that states that given a subset of $[N]^2$ with positive density, there exists an isosceles right triangle amongst the points, that is there exists points (x, y), $(x, y + d)$ and $(x + d, y)$ in our set.

In that case, we were looking at a subset of $[N]^2$. Now, we are looking at things in one higher dimension, and hence we are interested in looking at subsets of $[N]^3$. We are interested in proving the following generalization of Ajtai and Szemerédi's result.

Theorem 13.1. *Let $R \subseteq [N]^3$. For $\varepsilon > 0$, there exists an $N_0(\varepsilon)$ such that if*

$|R| \geq \varepsilon N^3$ and $N \geq N_0$, then there exists $(x, y, z), (x+d, y, z), (x, y+d, z), (x+d, y+d, z+d) \in R$, for some $d \neq 0$.

Note that geometrically what this theorem is saying is that as long as N is large enough, and R is a positive fraction of $[N]^3$ we can't avoid having four corner vertices to a cube, such that three lie on the same z plane and the fourth is the vertex above the remaining vertex on the z plane. Assuming we can prove this, the four term arithmetic progression easily follows.

13.2 A proof of the existence of a 4-term arithmetic progression

Suppose we have the above theorem, then the proof of the 4-term arithmetic progression is trivial.

Theorem 13.2. *Let $A \subseteq [N]$ with $|A| \geq \varepsilon N$. Then if $N \geq N_0(\varepsilon)$ then there is a 4-term arithmetic progression within A.*

Proof. We have $A \subseteq [N]$, with $|A| \geq \varepsilon N$. Then consider the set $R = \{(x, y, z) : 2x - y \in A\}$. Since A is a positive fraction of $[N]$, we have that R is a positive fraction of $[N]^3$. Therefore by the above theorem, there exists $(x, y, z), (x + d, y, z), (x, y + d, z), (x + d, y + d, z + d) \in R$. Therefore we have

$$\begin{aligned} 2x - y &\in A, \\ 2x + 2d - y &\in A, \\ 2x - y - d &\in A, \\ 2x + 2d - y - d = 2x + d - y &\in A, \end{aligned}$$

and these four elements form our desired arithmetic progression starting at $2x - y - d$ and changing by steps of d. □

Therefore the proof of a 4-term arithmetic progression follows directly from the previous theorem. In order to prove that theorem, we need the following theorem of Frankl and Rödl[34]. We state the theorem for 3-graphs, however it should be noted that the generalization of this to k graphs is also true.

Theorem 13.3. *Let H be a 3-uniform hypergraph such that each edge (triple) in H is in at most one $K_4^{(3)}$. Then we can remove $o(n^3)$ edges such that the remaining graph is $K_4^{(3)}$ free.*

This is a direct generalization of one of our applications of the Regularity Lemma, which looked at the situation where each edge in a (2-)graph occurs in at most one triangle. To prove this in the 2-graph case, we used the induced matching theorem. This also aids in seeing the generalization of this theorem. Note that in the $k = 2$ case, we were interested in the triangles, hence $K_3^{(2)}$, and in the $k = 3$ case we are interested in $K_4^{(3)}$. In the general case of the the theorem we are interested in occurrences of $K_{k+1}^{(k)}$.

We return to a proof of this theorem after we prove the generalization of Ajtai and Szemerédi's result.

13.3 Relating points in a cube to $K_4^{(3)}$

We assume the immediately proceeding theorem. In order to use it we would like to somehow relate the points in a cube to a $K_4^{(3)}$ so that we can apply Theorem 13.3.

In the two dimensional case, we have that two lines determine a point. Our proof of Ajtai and Szermédi's result was effectively finding lines of the form $x = i, y = j$ and $x - y = k$ such that they pairwise have solutions in A. Now we have 4 points to look for, and it takes 3 hypergraph equations to specify a point.

We define the following 4 groups of equations (which we can think of as subsets of $\mathbb{Z}[x, y, z]$):

$$\begin{aligned} A &= \{\text{Equations of the form } -y+z=i_1\}, \\ B &= \{\text{Equations of the form } x+y-z=i_2\}, \\ C &= \{\text{Equations of the form } z=i_3\}, \\ D &= \{\text{Equations of the form } -x+z=i_4\}. \end{aligned}$$

Note that if we pick one equation from each of A, B, and C it uniquely defines a point $(a, b, c) = (i_1 + i_2, i_3 - i_1, i_3)$. Likewise if we pick one equation from each of B, C and D we get $(i_3 - i_4, i_2 + i_4, i_3)$, one from each of C, D, A we get $(i_3 - i_4, i_3 - i_4, i_3)$, and one from each of A, B, D we get $(i_1 + i_2, i_2 + i_4, i_1 + i_2 + i_4)$. If each of these solutions were in R, then we would have our four points $(x, y, z), (x+d, y, z), (x, y+d, z), (x+d, y+d, z+d)$, with $x = i_3 - i_4$, $y = i_3 - i_1$, $z = i_3$ and $d = i_2 + i_1 + i_4 - i_3$. Thus the proof of this theorem boils down to showing that such a set of equations exist. We can now prove Theorem 16.1.

Proof of Theorem 16.1. Let $R \subseteq [N]^3$ with $|R| \geq \varepsilon N^3$. We define a hypergraph H on the set of vertices $V = A \cup B \cup C \cup D$, with $(f, g, h) \in \mathcal{E}$ if no two of f, g, h are in A, B, C or D, and if the point that is their common solution is in R. Let $(a, b, c) \in R$. Then note that (a, b, c) is the solution of the three equations $z = c \in A$, $-y + z = -b + c \in B$ and $x + y - z = a + b - c \in C$, and hence $(z = c, -y + z = -b + c, x + y - z = a + b - c) \in \mathcal{E}$, so $\mathcal{E}$ is not empty, and in fact $|\mathcal{E}| \geq |R| \geq \varepsilon N^3$. Let $(f, g, h) \in \mathcal{E}$. We show that it is in at least one $K_4^{(3)}$. Note that the three functions f, g, h define a point as above. For instance if $f \in A, g \in B, h \in D$, we have a point $(i_1 + i_2, i_3 - i_4, i_3)$. Consider then that $-x + z = i_3 - i_1 - i_2 \in D$, and any three of the equations $f, g, h, -x + z = i_3 - i_1 - i_2$ define the same point $(i_1 + i_2, i + 3 - i_4, i_3) \in R$, so this is a $K_4^{(3)}$. Likewise, if f, g, h were in any three of the sets A, B, C, D, an equation can be found in the fourth set such that any two of f, g, h and the new equation all define the same point in R, and hence any $(f, g, h) \in \mathcal{E}$ is in at least one $K_4^{(3)}$.

If each edge were in at most one $K_4^{(3)}$, by Theorem 13.3 we could remove $o(n^3)$ edges and have a $K_4^{(3)}$ free graph. But $|\mathcal{E}| > \varepsilon N^3$, and in fact, there are εN^3 edges (f, g, h) where $f \in A$, $g \in B$ and $h \in C$, each of which in at least one $K_4^{(3)}$ so after removing $o(n^3)$ edges, one of these must remain with it's $K_4^{(3)}$. That means for some edge $(f, g, h) \in \mathcal{E}$ it is involved with at least two $K_4^{(3)}$'s. Thus there exists some $-y + z = i_1$, $x + y - z = i_2$, $z = i_3$ and $-x + z = i_4$ such that $i_4 \neq i_3 - i_1 - i_2$. Note that by the discussion above, we have that with $i_3 - i_4 = x$, and $i_3 - i_1 = y$ and $i_2 + i_1 + i_4 - i_3 = d$ (with $d \neq 0$ as $i_4 \neq i_3 - i_1 - i_2$) that $(x, y, z), (x + d, y, z), (x, y + d, z), (x + d, y + d, z + d) \in R$ as desired. □

It is perhaps interesting before we move on to note that this all has a geometric interpretation as the intersection of hyperplanes, instead of using sets of equations.

13.4 Proving Theorem 13.3

We now turn back to Theorem 13.3, which we used to prove the previous result, that is if each edge in a 3-graph is in at most one $K_4^{(3)}$ we can destroy all of the $K_4^{(3)}$'s in the graph by removing $o(n^3)$ edges.

Recall that by the Hypergraph Regularity Lemma, given $\varepsilon > 0$, a 3-uniform hypergraph H on n vertices $K_n^{(2)}$, can be partitioned into finitely many parts

$$K_n^{(2)} = B_1 \cup \cdots \cup B_k,$$

where $k = k(\varepsilon)$, such that almost all edges fall into ε-regular triples.

Further recall that a triple (X, Y, Z) is ε-regular if for $A \subseteq X$, $B \subseteq Y$ and $C \subseteq Z$ with $|A| \geq \varepsilon|X|$, $|B| \geq \varepsilon|Y|$ and $|C| \geq \varepsilon|Z|$ we have that

$$\left|\frac{e_H(A,B,C)}{e_K(A,B,C)} - \frac{e_H(X,Y,Z)}{e_K(X,Y,Z)}\right| < \varepsilon$$

where $e_H(A, B, C)$ is the number of edges in H induced on A, B, C, and $e_K(A, B, C)$ is the number of edges of a complete hypergraph induced on A, B, C.

The idea of this proof is to use the Hungarian method, that is throw away some insignificant portion of the edges so that the structure of the graph is made more clear but the remaining graph still has a positive fraction of the edges. Given a graph $H(V, \mathcal{E})$, if $\mathcal{E} < \varepsilon n^3$ we are done, so we can take H to have αn^3 edges where α is large compared to ε (say, $\alpha > 100\varepsilon$). Note that each vertex in $v \in \mathcal{E}$ induces a 2 graph on V where (s, t) is an edge if $(v, s, t) \in \mathcal{E}(H)$. We call this graph G_v. Note that a $K_4^{(3)}$ involving a vertex v corresponds to a triangle in G_v that is induced by an edge in H.

Since H has a large number of edges, we have that most G_v's will have a large number of edges, cn^2 for instance. Those that do not, we discard (and in the process since there are not many of them, and they don't contain many edges, we do not remove many edges from H). Consider a regular partition of the vertices, $K_n^{(2)} = B_1 \cup B_2 \cup \cdots \cup B_n$, which we get from the regularity lemma for 3-graphs. Now we consider the sets $B_i \cap G_v$. We would like if $|B_i \cap G_v| \neq \emptyset$ for $|B_i \cap G_v|$ to be large, and hence be a positive fraction of B_i (say $|B_i \cap G_v| > 5\varepsilon|B_i|$. Now we can apply the Regularity Lemma if (B_i, B_j, B_k)

has positive density (i.e $\frac{e_H(B_i,B_j,B_k)}{e_K(B_i,B_j,B_k)} > 2\varepsilon$), then $(B_i \cap G_v, B_j \cap G_v, B_k \cap G_v)$ has positive density, and hence we can find a triangle in G_v. We want to show that even after removing $o(n^3)$ edges, there still must be triangles in H – so our assumption that $|\mathcal{E}(H)|$ was large must be incorrect.

Proof of Theorem 13.3. We delete all edges not in a $K_3^{(4)}$, and it is enough to show that the remaining graph H has at most $o(n^3)$ edges. Start with H, with $e(H) > 100\varepsilon n^3$. We delete the following edges from H by iteratively applying the following operations regardless of the order:

1. Delete all edges of H incident to a vertex v if $|G_v| < 10\varepsilon n^2$.
2. Delete edges of H induced by B_i, B_j, B_k if $\delta(B_i, B_j, B_k) < 2\varepsilon$.
3. Delete edges of H in ε-irregular triples (B_i, B_j, B_k)

The procedure stops before we delete $20\varepsilon n^3$ edges from H. In the remaining graph, any edge is in (B_i, B_j, B_k) with $\delta(B_i, B_j, B_k) > 2\varepsilon$. Thus $\delta(B_i \cap G_v, B_j \cap G_v, B_k \cap G_v) > \varepsilon$, and there is an edge x of H induced by $(B_i \cap G_v, B_j \cap G_v, B_j \cap G_v)$. Together with the vertex v, and the three triples containing v and a pair in x we have a $K_4^{(3)}$. □

Lecture 14

More on Patterns in Subsets

Jake Wildstrom

14.1 Ramsey- and Turán-type problems

14.1.1 Turán numbers

Ramsey problems we are all familiar with. Turán's Theorem approaches the question of inevitability from a different viewpoint: instead of establishing that a certain number of vertices suffice to induce monochromatic patterns in colorings of edges, we explore the edge density necessary to induce patterns in arbitrary-size graphs.

Theorem 14.1 (Turán's Theorem [42]). *For G a graph on n vertices, if $e(G) > (1 - \frac{1}{k})\frac{n^2}{2}$, then G contains a K_{k+1}.*

We denote by $t(n, H)$ the maximum number of edges on an H-avoiding graph on n vertices. So Theorem 14.1 gives an upper bound on $t(n, K_{k+1})$, and in fact a very nearly sharp bound. As an example, $t(n, K_3) = \lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil$. That it must be at least this large is easily illustrated by considering a complete bipartite graph on n vertices, with $\lfloor \frac{n}{2} \rfloor$ vertices in one part and $\lceil \frac{n}{2} \rceil$ in the other; this graph can easily be shown not to contain a K_3 subgraph. In fact, generalizing this method to k-partite complete graphs, in which each part has as close to the same size as possible, we find a lower bound of approximately $(1 - \frac{1}{k})\frac{n^2}{2}$ on $t(n, K_{k+1})$. The actual value of $t(n, K_{k+1})$ derived from such a division is highly inelegant, due to the necessity to split the parts slightly unevenly when k does not divide n, but this approximation is quite close.

Other Turán numbers are known as well. Erdős [20] accidentally determined that $t(n, C_4) \approx n^{\frac{3}{2}}$ in pursuit of a different result, before Turán even posed the problem, but failed to generalize to the wider concept of Turán-type problems. $t(n, C_6)$ is of order $n^{\frac{4}{3}}$ and $t(n, C_10)$ is of order $n^{\frac{6}{5}}$. Erdős [19] proposed an open problem to show that $t(n, C_{2k}) \geq cn^{1+\frac{1}{k}}$ for $k = 4$ and $k \geq 6$. An upper bound for the Turán number of a general even cycle is known due to work of Erdős, Bondy, and Simonovits [19, 6]: $t(n, C_{2k}) \leq ckn^{1+\frac{1}{k}}$.

Another class of open Turán-type problems exists in hypergraphs. Very little is known about these problems: even the simple question of asymptotic extremal density of $K_4^{(3)}$-free 3-graphs is unknown. There is a straightforwardly constructed lower bound thought to be sharp:

Theorem 14.2. $t_3(n, K_4^{(3)}) \gtrsim \frac{5}{9}\binom{n}{3}$.

Proof. Let $n = 3m$ for simplicity; the proof is easily generalized when we may partition the vertex-set into three nearly-equal rather than equal sets. Let V_1, V_2, and V_3 be three disjoint vertex-sets of size m each, and let G be a 3-graph on the union of these sets such that (u_1, u_2, u_3) is an edge in G if and only if either $u_1, u_2 \in V_i$ and $u_3 \in V_{i+1}$ letting $V_4 = V_1$ conventionally, or if $u_i \in V_i$ for each i. We shall now count the edges in this graph: an edge with vertices only in V_i and V_{i+1} may have $\binom{m}{2}$ choices of vertices from V_i and m choices of vertex from V_{i+1}, so there are $\frac{m^2(m-1)}{2}$ such edges among each pair, and since there are 3 such pairs of parts (V_1, V_2), (V_2, V_3), and (V_3, V_1), there are $3\frac{m^3-m^2}{2}$ edges of the first type. The second type of edge draws one vertex from each part, so there are m^3 possible edges. Thus the total number of vertices in this graph is $\frac{5m^3-3m^2}{2} = \frac{5n^3-9n^2}{54} \approx \frac{5}{9}\binom{n}{3}$, as expected.

It now remains simply to show that this graph is $K_4^{(3)}$-free. Let us consider all possible ways 4 vertices u_1, u_2, u_3, u_4 can be distributed among the 3 parts of the vertex set: $(4, 0, 0)$, $(3, 1, 0)$, $(2, 2, 0)$, and $(2, 1, 1)$. The first two possibilities are guaranteed not to span a $K_4^{(3)}$ subgraph, since that would require an edge with all three vertices in a single part. For the third possibility, we may without loss of generality assume $u_1, u_2 \in V_1$ and $u_3, u_4 \in V_2$, in which case (u_1, u_3, u_4) and (u_2, u_3, u_4) are not edges in G; and for the last case, we without loss of generality assume $u_1, u_2 \in V_1$, $u_3 \in V_2$, and $u_4 \in V_3$, in which case (u_1, u_2, u_4) is not an edge in G. Thus no 4 points of this graph induce a $K_4^{(3)}$. □

This bound is thought to be sharp, but the best upper bound on the extremal density, given in [16], is $\frac{3+\sqrt{17}}{12}$. There is a \$1000 Erdős prize for the solution of a 1941 conjecture of Turán [42] as to the exact solution to this problem.

14.1.2 Relating Regularity to Turán's Theorem

A singular example of how to use regularity in Turán problems is the following result, which was a breakthrough of great complexity when it was proven in 1948, but is rendered far more tractable by use of the Regularity Lemma:

Theorem 14.3 (Erdős-Stone Theorem[22]). *Let $K_{r+1}(s, s, \ldots, s)$ denote the $(r+1)$-partite complete graph with s vertices in each part. Then*

$$t(n, K_{r+1}(s, s, \ldots, s)) = (1 - \frac{1}{r})\binom{n}{2} + o(n^2)$$

for sufficiently large n.

The proof will be given in the next lecture.

Lecture 15

Regularity Lemma and Turán Type Problems

Steven Butler

15.1 The Erdős-Stone Theorem

As the course is progressing we continue to look for ways to use the Regularity Lemma to give new proofs of old results that are more easily understood. In the last lecture we looked at the Turán Theorem which states that if a graph has enough edges (i.e., density is sufficiently large) then the graph must contain a clique. Specifically, it is stated as follows.

Theorem 15.1 (Turán's Theorem). *Let G_n denote a graph on n vertices and $e(G_n)$ the number of edges of G_n. Then if*

$$e(G_n) > \big(1 - \frac{1}{k}\big)\frac{n^2}{2},$$

then G_n must contain K_{k+1} (a $k+1$ clique or a complete graph on $k+1$ vertices).

Turán type problems are problems that deal with the edge density of a graph. So natural questions that arise are what graph invariants are sensitive to density. A major result along these lines is the following.

Theorem 15.2 (Erdős-Stone Theorem). *Let H be a fixed graph with chromatic number $p+1$. Then for n sufficiently large, i.e., larger than some* *constant depending only on H, we have that if*

$$e(G_n) > \big(1 - \frac{1}{p}\big)\frac{n^2}{2},$$

then G_n contains a copy of H.

To say that G_n contains a copy of H is to say that there is an injective map $\phi : V(H) \to V(G_n)$ such that if u and v are adjacent then $\phi(u)$ and $\phi(v)$ are adjacent. Or using graph terminology H is a subgraph of G_n.

The original proof predates the Regularity Lemma and so is much more complicated than the argument we will present here. As an initial step we introduce the notion of a reduced graph. The idea is to extract the significant part of the graph to see what patterns emerge.

15.1.1 The Reduced Graph

Given a graph G with at least cn^2 edges, where $c > 0$ (which is to say that the graph has positive edge density), and $\varepsilon > 0$ we can apply the Regularity Lemma and get a partition of the vertices. In particular, in this case we will use the form of the Regularity Lemma that allows us to have our partition $P = V_0 + V_1 + \cdots + V_k$ have $|V_i| = |V_j|$ for $i, j > 0$ and $|V_0| \leq \varepsilon n$. This form can be derived from the Regularity Lemma that we have been using by adding one final step in our construction of our partition in which we cut up the pieces of our final partition into equally sized blocks with a few "junk" vertices.

The reduced graph of G, which we will denote by R has vertex set $V(R) = \{V_1, \ldots, V_k\}$ with V_i adjacent to V_j if and only if the pair (V_i, V_j) is a ε-regular pair. Also related to this is the notion of the fat reduced graph, this can be thought of as either a pumped up version of the reduced graph or as a subgraph of G. Namely, the fat reduced graph is found by deleting V_0 and all edges adjacent to it along with all edges not in ε-regular pairs. This process will at most delete $2\varepsilon n^2$ edges and so we are left with essentially the same number of edges as before but now we have the "soul of the graph". Grouping the vertices then according to the partition gives us our fat reduced graph.

15.1.2 Throwing Away the Noise

We recall a result from the beginning of the lectures which will prove useful in our proof. Namely, when working with a ε-regular pair (V, W), each with m

vertices and with edge density δ between them, we have that for most vertices in V its neighborhood in W is roughly the right size. Stated less vaguely it is as follows.

Lemma 15.1. *For V, W as given above we have that at least $(1-2\varepsilon)m$ of the vertices $v \in V$ satisfy*

$$(\delta - \varepsilon)m \leq |\{\#w \in W | v \sim w\}| \leq (\delta + \varepsilon)m.$$

Proof. Let $V_1 = \{v \in V : |\{\#w \in W|v \sim w\}| > (\delta+\varepsilon)m\}$. Then note that the edge density between V_1 and W varies from δ by more than ε. But by definition of ε-regular this can only happen if $|V_1| < \varepsilon|V| = \varepsilon m$. By a similar argument we have that $V_2 = \{v \in V : |\{\#w \in W|v \sim w\}| < (\delta-\varepsilon)m\}$ satisfies $|V_2| < \varepsilon|V| = \varepsilon m$. It follows then that the V satisfying the desired relationship lie in $V - (V_1 + V_2)$ has at least $(1-2\varepsilon)m$ vertices. □

Another useful construction that will come in handy is that for any ε-regular pair (V, W), each with m vertices and with edge density δ between them we can construct a bipartite graph $K_{t,t}$ between V and W. Moreover, we can say something about the size of t in terms of δ and ε, in particular $t \approx \log(\varepsilon)/\log(\delta)$.

To see this construction we start by picking vertices in $v_1 \in V$ and $w_1 \in W$ so that the neighborhood of v in W and the neighborhood of w in V are both of size $\approx \delta n$, call these neighborhoods V_1 and W_1, if $\delta n \geq \varepsilon n$ then the pair (V_1, W_1) is also ε-regular and we repeat the process by selecting points v_2 and w_2 which have neighborhoods V_2 and W_2 respectively, each of these are both of size $\approx \delta^2 n$. We can continue this process until the sets V_k and W_k are too small to be ε-regular, i.e., if $\delta^k < \varepsilon$, hence the reason we can do this approximately $\log(\varepsilon)/\log(\delta)$ times. Now the $K_{t,t}$ is formed by $\{v_1, \ldots, v_t, w_1, \ldots, w_t\}$.

The idea in this construction is that we pick vertices $v \in V$ and $w \in W$ and their corresponding neighborhoods V_1 and W_1, we then construct a $K_{t-1,t-1}$ between V_1 and W_1 and then add in v and w to form the $K_{t,t}$. Looking ahead when we construct subgraphs we will first start with a vertex in part of our ε-regular partition, look at the neighborhoods of that vertex inside of the ε-regular pairs and then construct the corresponding smaller graph among these neighborhoods.

The above proof can be generalized to the following key lemma: Suppose the reduced graph R of G contains K_s with all edge densities for the ε-regular pairs in the K_s being greater then some positive constant η. Then G contains $K(t, t, \ldots, t)$ (where t is repeated s times) provided that n is sufficiently large.

15.1.3 Proof of Erdős-Stone

Using the reduced graph we are now in a position to prove the Erdős-Stone Theorem.

Proof of Erdős-Stone. Given a graph G_n let R denote the reduced graph with respect to some ε, to be determined, and G' the fat reduced graph. We note it suffices to show that G' contains a copy of H. Since

$$\frac{e(R)}{k^2/2} \geq \frac{e(G')}{n^2/2} > \big(1 - \frac{1}{p}\big),$$

it then follows by Turán's theorem that R contains a K_{p+1}. Using the method of construction above we have that for n sufficiently large that G' contains $K_{p+1}(t, \ldots, t)$ for some t. Further by choosing ε sufficiently small we can get t to be arbitrarily large, in this case pick $t = |H|$. It follows then that H is a subgraph of $K_{p+1}(t, \ldots, t)$ which is a subgraph of G' which is in turn a subgraph of G. Concluding the proof. □

15.2 Some Extremal Results

The Erdős-Stone Theorem gives conditions to guarantee the existence of a small fixed graph H in some large graph G. We can extend it to show that in some sense it not only shows up but shows up often, as done in the following theorem.

Theorem 15.3. *For a given $\beta > 0$ suppose that*

$$e(G_n) > \big(1 - \frac{1}{p} + \beta\big)\frac{n^2}{2}.$$

Then for H a fixed graph with chromatic number $p+1$, and $|V(H)| = h$ then

$$\big|\{\textit{copies of } H \textit{ in } G_n\}\big| \geq cn^h,$$

for some absolute constant depending on H, β, p but not on n.

We note that c is a *very* small constant, and further it can be shown that c has the form $c'\beta^h$. Note that the number of copies of H in G_n is the number of maps of vertices of H into vertices of G_n which preserve adjacency (as discussed above). The number of such mappings is n^h and so this theorem says that some positive fraction of the mappings will contain H.

The idea of the proof is as follows, for n sufficiently large we form an ε-regular partition and its corresponding reduced graph, which will contain a copy of H, for which we can order the vertices in some arbitrary manner, say $W_1, \ldots, W_h$. Now look at this in the fat reduced graph and we start picking out vertices, since the number of vertices in the reduced graph is independent of n, the block that is the blow up of the first vertex is some positive fraction $c_1 n$ of the total number of vertices. So for our first choice we get $c_1 n$ choices. Now for every $W_i \sim W_j$ in H we take the neighborhood of W_j adjacent to our choice for the first vertex while if $W_i \not\sim W_j$ in H we take all of W_j. The key is that regularity guarantees that we still have a positive fraction of the vertices, so our second choice is from some $c_2 n$ different vertices. Continuing in this manner establishes the proof.

The argument given in the proceeding paragraph is only an outline, were we to be more careful we would also note that at each stage only a small (i.e., ε-)portion in every cluster is bad, and we can safely ignore these.

An immediate consequence of the preceding theorem is that if a pattern does not appear frequently enough we can remove relatively few edges and destroy all such patterns that are present.

Theorem 15.4. *For every $\beta > 0$ and H, a fixed graph, there is a $\gamma = \gamma(B, H)$ so that for all n sufficiently large that whenever G_n has at most $\gamma m^{|V(H)|}$ copies of H then by deleting βn^2 edges of G_n the remaining graph is H-free.*

The "remaining graph" we will end up with is G', the fat reduced graph. The idea is to look at the reduced graph, if it contains H then by construction above we get too many copies of H, a contradiction. It follows that G' is H free, and as noted earlier can be gotten by taking out a small number of edges.

15.3 Theorem of Szemerédi

A related result using arguments along the same lines is the following. Note that the independence number of a graph G_n, denoted by $\alpha(G_n)$, is the maximal number of mutually disjoint vertices in the graph, or equivalently the largest clique size in the complement of the graph.

Theorem 15.5 (Szemerédi). *If G_n does not contain K_4 and $\alpha(G_n) = o(n)$ then $e(G_n) < n^2/8 + o(n^2)$.*

Before we outline the proof we pause to note that the fact that G_n does not contain K_4 implies by the Turán Theorem that $e(G_n) \leq n^2/3$ edges. So the added assumption on the independence number helps to significantly lower the number of edges.

First, we may assume that G_n has some positive density of edges, otherwise we are already done. In particular, we may assume that $e(G_n) > n^2/8 + 100\varepsilon n^2$. The idea is to consider the reduced graph R and break it up into two cases. In each case we will get a contradiction by constructing a K_4.

Case 1: Suppose that R has more than $k^2/4$ edges, where k is the number of vertices of R. Then Turán's Theorem implies that R contains a triangle. Suppose that U, V, W are the three vertices of R that contain a triangle and look at the three vertex sets U, V, W in the fat reduced graph. First we can remove the "junk" vertices, i.e., those vertices in U, V, W that do not have neighborhoods the right size. From Lemma 15.1 we know that removing such vertices will leave us with most of U, V, W. Now we start with some vertex p left over in U, and consider the neighborhood of p in V and W i.e., $N_p \cap V$ and $N_p \cap W$. Now let

$$V_1 = \{q \in V \cap N_p : \big| |N_q \cap N_p \cap W| - \delta_2\delta_3|W| \big| < \varepsilon|W|\}$$

where δ_2, δ_3 are the edge densities between V, W and U, W respectively. Note that the term $\delta_2\delta_3|W|$ is the expected size of $|N_q \cap N_p \cap W|$. It follows from ε-regularity that since $V \cap N_p$ is sufficiently large (of order $\delta_1|V|$ where δ_1 is the edge density between U, V) that $|V_1| > 0$. Pick some arbitrary point $q \in V_1$ then since $|N_q \cap N_p \cap W|$ is some positive fraction of n and the independence number is small (i.e., $o(n)$) there must be an edge in $N_q \cap N_p \cap W$. Combining this edge with p and q gives the desired K_4.

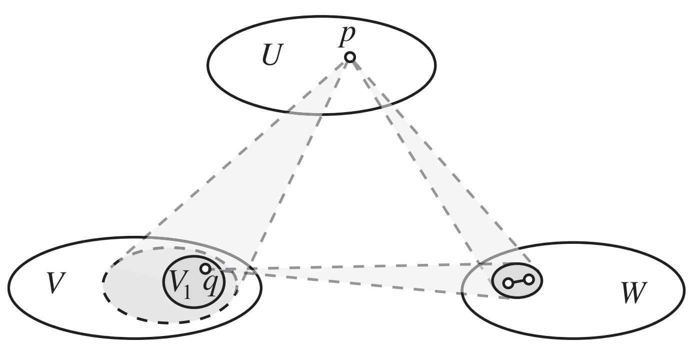

Figure 15.1: The visual interpretation of case 1

Case 2: Suppose that R has less than or equal to $k^2/4$ edges. Then note that

$$\sum_{\substack{V_i,V_j) \\ \varepsilon\text{-regular} \\ \text{pairs}}} \delta(V_i, V_j) = \frac{e(G')}{m^2} \geq \frac{n^2/8 + 90\varepsilon k^2}{k^2/4} = \frac{k^2}{8} + 90\varepsilon k^2,$$

so calculating we have that the average density between ε-regular pairs is at least the above divided by $k^2/4$. in particular, the average density is at least $1/2 + c\varepsilon$. So pick now a (U, V) so that the density between then is at least $1/2 + c\varepsilon$. Then pick two vertices p, q of U that make an edge and look at the intersection of N_p, N_q in W. Because the edge density is so large the intersection must be at least $c\varepsilon|W|$, but then this must contain an edge because the independence number is $o(n)$. This gives a K_4.

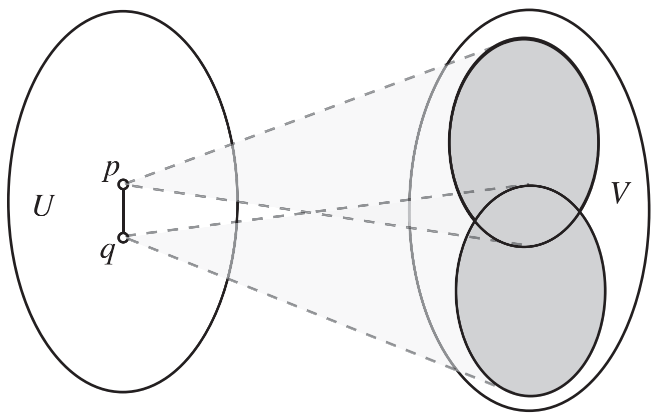

Figure 15.2: The visual interpretation of case 2

Lecture 16

Ramsey Theory Applications of the Regularity Lemma

Paul Horn

During the last lecture we examined several applications of the Regularity Lemma, including proving the Erdős-Stone theorem, and the Ramsey-Turán theorem. We now turn out attention to applications of the Regularity Lemma in Ramsey theory. Recall that we define the Ramsey number

$$r(k,k) = \min\left\{n : \begin{array}{c}\text{any 2-coloring of the edges of}\\ K_n \text{ contains a monochromatic } K_k\end{array}\right\}.$$

We are interested in a generalization of this, namely given a graph H we define

$$r(H) = \min\left\{n : \begin{array}{c}\text{any 2-coloring of the edges of}\\ K_n \text{ contains a monochromatic } H\end{array}\right\}.$$

What types of patterns, H, should have 'large' Ramsey numbers, and what characterizations give relatively small Ramsey numbers? Here we examine the interesting cases which yield linear Ramsey numbers, (i.e. $r(H) < c|V(H)|$). Some of these results follow relatively easily from the Regularity Lemma. We now look at some of these results, and also at some conjectures.

16.1 A result on $r(H)$ for graphs with bounded degree

In 1975 Paul Erdős made the following conjecture: if H is a graph on n vertices with bounded degree Δ then $r(H) \leq cn$, where c is a constant that depends only on Δ. This result was proved in 1983 by Chvatál, Szemerédi, Trotter, and Rödl [17]. We will return to prove this later.

Theorem 16.1. *Suppose H has bounded degree δ, $|V(H)| = n$. Then $r(H) \leq cn$ where $c = c(\Delta)$*

Note that last time, we showed that if H is a fixed graph (and hence $V(H)$ constant) then H will appear cn^h times if H appears in the reduced graph of a large graph G. However now we are talking about H with a constant proportion of the vertices in G and the result is inherently similar concerning the frequency of appearance of H in G, so long as the degree is bounded. This theorem states that $r(H) \leq cn$ for graphs where the H with bounded degree, however c is a very large constant. The regularity lemma gives this constant a tower type $2^{2^{\cdot^{\cdot^{2}}}}$ with the height of the tower a function of Δ. Subsequent results have improved this. Nancy Eaton [18], as part of her Ph.D. thesis reduced this to

$$c(\Delta) < 2^{2^{c\Delta}}.$$

In 2000, R.L. Graham, Rödl, and Ruciński [25] reduced this further still to

$$c(\Delta) < 2^{c\Delta(\log \Delta)^2}.$$

These proofs improved the bound by using the idea of the Regularity Lemma without applying it directly.

16.2 More results on $r(H)$ for different classes of graphs

There are many classes of graphs where, while the degree is not bounded, it acts in some sense like the degree is bounded. In some of these cases we can also get linear Ramsey numbers. Chen and Schelp [9] proved the following

Theorem 16.2. *If H is a planar graph on n vertices, then $r(H) \leq cn$.*

A planar graph does not necessarily have bounded degree. For instance, a star graph is planar but does not have bounded degree. The average degree, however, for a planar graph is small. Planar graphs have some other nice properties as well, for instance a planar graph has to have a vertex with degree at most 5, and a planar graph can be built by iteratively adding vertices with degree at most 5. This turns out not to be enough to show the theorem, so a stronger property, called c-arrangeability, is needed. A graph G is *c-arrangeable* if all vertices of H can be order $v_1, v_2, \dots, v_n$ such that for all i we have

$$|\{j : v_i \sim v_j, k > i, v_k \sim j, j < i\}| \leq c$$

It is not immediately obvious that planar graphs satisfy this condition, however this is the condition that is key to showing that planar graphs have a linear Ramsey number. c-arrangeability tells us the order to 'grow' our graphs using the the regularity lemma. Here is a rough sketch of the proof in [9].

At any point, the choices for v_i are restricted only by the v_j that came before it, and will influence the future neighbors of v_i. There are only a finite number of these, and each has a 'bag' of possible choices. That is for each v_i there is a positive fraction number of vertices in the regular partition that it is being embedded into that are 'candidate' v_i, and each of these 'bags' restrict the potential v_i. However, since there are only finitely many restrictions, then ε-regularity gives us that there will be a positive-fraction number of candidates for v_i, so we are in some sense setup to find the next vertex.

Originally when it was shown that planar graphs were c-arrangeable, it was for a very high c, somewhere in the neighborhood of $c = 761$. However Kierstead and Trotter [28] showed that planar graphs are 10-arrangeable.

Actually we get a stronger theorem than Theorem 16.2, also due to Chen and Schelp.

Theorem 16.3. *If H is c-arrangeable, $|V(H)| = n$ then $r(H) \leq c'n$*

Rödl and Thomas [35] further extended this to graphs with bounded genus.

Theorem 16.4. *If H is of bounded genus, $|V(H)| = n$, then $r(H) \leq cn$.*

Despite many nice results that are known concerning graphs with linearly bounded Ramsey numbers, there are still a number of related open conjectures.

16.3 Conjectures regarding graphs with linearly bounded Ramsey numbers.

Erdős and Burr [7] conjectured the following:

Conjecture 16.1. If H has average degree Δ, with $|V(H)| = n$, then

$$r(H) \leq cn$$

where $c = c(\Delta)$.

They also proposed the following, somewhat stronger version.

Conjecture 16.2. Suppose H is a graph, with $|V(H)| = n$. If, for all subgraphs $H' \subset H$, H' has minimum degree less than Δ then

$$r(H) \leq cn$$

where $c = c(\Delta)$.

Note that both of these are very closely related to the result of Chvátal, Rödl, Szemerédi, and Trotter given above as Theorem 16.1, which states that if the degree is bounded by Δ then the Ramsey numbers are linearly bounded. In another direction, Erdős and Burr also made the following conjecture

Conjecture 16.3. If H is the union of t forests, with $|V(H)| = n$, then

$$r(H) \leq cn$$

where $c = c(t)$.

Recall that a forest is an acyclic graph, or a collection of trees, not necessarily connected. A somewhat related conjecture on trees is the following 1963 conjecture of Erdős and Sós.

Conjecture 16.4. Every graph on n vertices having at least $n(k-1)/2+1$ edges must contain as a subgraph every tree of k vertices.

Note that if this last conjecture is true, it is the best possible. As the star is a tree on k vertices, there must be some vertex with degree n, but there also must be a path of length k in the graph. The best effort so far towards proving this theorem is an approximation version using the regularity lemma. An obvious way of attacking this problem is by mathematical induction, and this can be proven inductively without the factor of 2 (i.e., assuming that G has more than $n(k-1)$ edges).

16.4 A proof of Theorem 16.1

We return back to give a proof of Theorem 16.1 using the Regularity Lemma. Recall that Theorem 16.1 states that if H is a graph on n vertices with bounded degree Δ, then $r(H) \leq cn$ where $c = c(\Delta)$.

Proof of Theorem 16.1. If H has bounded degree Δ, then we have that the chromatic number $k = \chi(H) \leq \Delta + 1$. Let G be a graph on $N = cn$ vertices. We think of G as the subgraph induced by the blue edges in a 2-coloring of the edges K_N. We want to show that either G or $\overline{G}$ contains H.

Consider a regular partition for some $\varepsilon > 0$. We can assume that all classes in our partition are the same size, save one small exceptional class. Let R be the reduced graph of our regular partition. We color the edges of R as follows; we color an edge of R blue if the density of the ε-regular pair it connects is $\geq \frac{1}{2}$, and an edge of R red if the density of the ε-regular pair it connects is $< \frac{1}{2}$.

Note that by the regularity lemma, $|R|$ is finite, but we can increase the size of $|R|$ by taking a finer partition. We assume that $|R| \geq r'(k,k)$, where $r'(k,k)$ is the Ramsey number that guarantees that a graph of the form K_n with $o(n^2)$ edges removed contains either a red or blue K_k. Thus R contains either a red or a blue K_k. Erdős-Stone implies that, if R contains a blue K_k, that G contains H or, if R contains a red K_k, that $\overline{G}$ contains H. Thus as $|G| = N = cn$, we have that $r(H) \leq cn$ as desired. □

Lecture 17

Extremal Conjectures Related to the Regularity Lemma and an Introduction to Expanders

Dan Felix

17.1 Open Conjectures in Extremal Graph Theory

Before beginning the treatment of expander graphs we will mention some open conjectures in extremal graph theory. In most cases we do have partial results, and the proofs make use of various versions of the Regularity Lemma.

Conjecture 17.1 (Erdős-Sós Conjecture). If a graph G on n vertices has more than $(k-1)n/2$ edges, then it contains all trees having k edges.

Ajtai, Komlós, and Szemerédi [1] have proven the following approximate version of this conjecture.

Theorem 17.1. *For every $\epsilon > 0$ there exists an $n_0 = n_0(\epsilon)$ such that if a graph G on $n > n_0$ vertices has at least $(1+\epsilon)(k-1)n/2$ edges, then it contains all trees having k edges.*

Our next conjecture is also known as Loebl's Conjecture.

Conjecture 17.2 ($n/2$-$n/2$-$n/2$ Conjecture). Let G be a graph on n vertices. If G has at least $n/2$ vertices of degree at least $n/2$, then G contains all trees on at most $n/2$ vertices.

Ajtai, Komlós, and Szemerédi [2] have proven an approximate form of this conjecture as well.

Theorem 17.2. *Let G be a graph on n vertices. For every $\epsilon > 0$ there exists an $n_0 = n_0(\epsilon)$ with the following property. If G has at least $(1+\epsilon)n/2$ vertices of degree at least $(1+\epsilon)n/2$, and if $n > n_0$, then G contains all trees on at most $n/2$ vertices.*

Yi Zhao may have recently proven this conjecture for all sufficiently large graphs [45]. A stronger version, conjectured by Komlós and Sós [21], is still open.

Conjecture 17.3. Let G be a graph on n vertices. If G has at least $n/2$ vertices of degree at least k, then G contains every tree on at most k vertices.

Lastly, the following strengthening of Dirac's theorem is still open. Note that the *kth power of a graph G* is the graph obtained from G by adding the edge $\{u, v\}$ if the distance between the vertices u and v is at most k in G. It is so named because the adjacency matrix of the kth power of G is simply the kth power of the adjacency matrix of G, with all nondiagonal nonzero entries replaced by 1, and with all diagonal entries replaced by 0.

Conjecture 17.4 (Posa-Seymour Conjecture). Let G be a graph on n vertices with minimum degree at least $kn/(k+1)$. Then G contains the kth power of a Hamiltonian cycle.

Komlós, Sárközy, and Szemerédi have nearly settled this problem [29].

Theorem 17.3. *Let G be a graph on n vertices. For every $\epsilon > 0$ there exists an $n_0 = n_0(\epsilon)$ with the following property. If G has minimum degree at least $(1+\epsilon)kn/(k+1)$, and if $n > n_0$, then G contains the kth power of a Hamiltonian cycle.*

17.2 Network Routing

We now shift our attention to expander graphs and related concepts. The study of these types of graphs, which we will define later, has risen out of network routing problems. Take, for example, the problem of constructing a permutation network.

Imagine that we are given a set of n inputs (say, electrical signals) and a set of n destinations. We are also given a large supply of two-state switches, whose states are represented schematically in the following figure. Each switch can

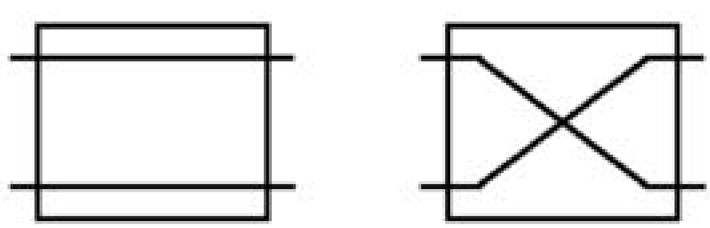

Figure 17.1: The two states of a switch.

take two signals as inputs, and it will output them unchanged or transposed, depending on the state of the switch. We will refer to the unchanged state as the *off* state, and its counterpart will be the *on* state. Our goal is to use these switches to connect the initial n inputs to the n destinations in such a way that any desired configuration (that is, any set of pairings of inputs with destinations) can be achieved by setting the switches to their appropriate states. Such a network is called a *permutation network*. Figure 2 is an example of a permutation network having 8 inputs, though we will not prove this fact. Note that each box in this figure represents a switch of the type depicted in figure 1.

Of course, we would like to use as few switches as possible. We can establish a lower bound on the fewest number of switches needed (as a function of the number of inputs) rather easily. We identify the initial inputs and the destinations with the set $\{1, \ldots, n\}$. Hence a particular choice of states for our switches can be identified with a permutation in S_n. Let k be the number of switches used in our permutation network. There are clearly 2^k ways to set our switches to their on or off states, while we must be able to produce $n!$ distinct permutations with these settings. Therefore $2^k > n!$ is a necessary condition, which implies that

$$k \gtrsim n \log_2(n)$$

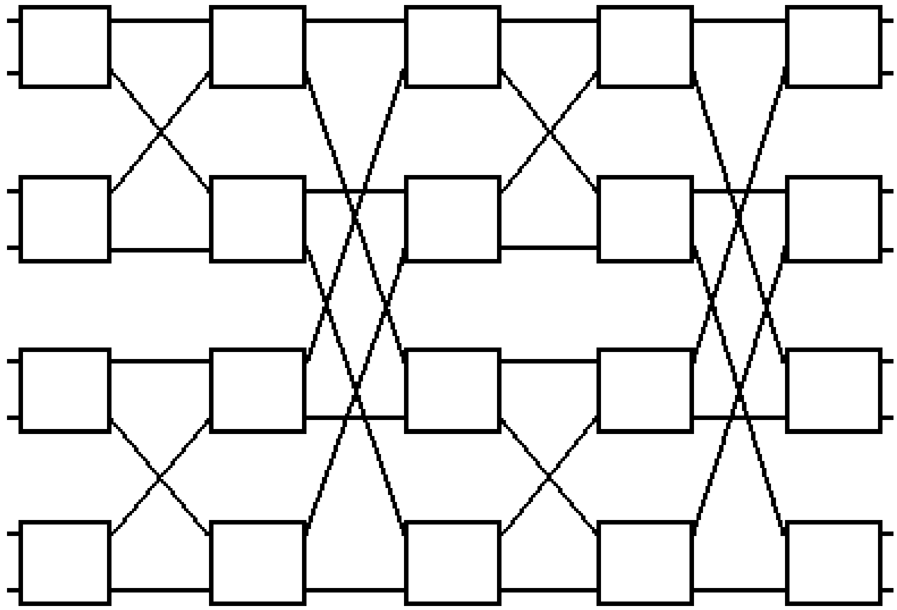

Figure 17.2: A permutation network having 8 inputs.

for all permutation networks.

17.2.1 Expander Graphs

An upper bound on the minimum number of switches needed calls for an existence proof. Expander graphs help provide the necessary constructions for upper bounds for many related problems. We first present one of the original definitions of an expander.

Definition 17.1. Let $G = (V, E)$ be a bipartite graph with bipartition $V = A \bigcup B$ such that $|A| = |B| = n$. We say that G is an *ϵ-expander* if for every $X \subset A$, $Y \subset B$ with $|X| > \epsilon n$ and $|Y| > \epsilon n$, there exists an edge e with one endvertex in X and the other in Y.

Notice that this innocent definition is actually quite strong. Letting $Y \subset B$ be the set of non-neighbors of $X \subset A$, we see that $|Y| \leq \epsilon n$ if $|X| > \epsilon n$. Hence, letting $\Gamma(X)$ denote the set of neighbors of vertices in X, we find that

$$|\Gamma(X)| \geq (1 - \epsilon)n$$

for *every* $X \subset A$ satisfying $|X| > \epsilon n$. Because of the restrictive nature of such a strong definition, we often use the following definition instead.

Definition 17.2. Let $G = (V, E)$ be a bipartite graph with bipartition $V = A \bigcup B$ such that $|A| = |B| = n$. Let $0 < \epsilon < 1$ be any fixed real number. We say that G is an ϵ-*expander* if

$$|\Gamma(X)| > (1 + \epsilon)|X|$$

for every $X \subset A$ satisfying $|X| \leq n/2$.

Lecture 18

Expander Graphs and Superconcentrators

Steven Butler

18.1 Expander graphs

In this lecture we will introduce expander graphs and use them to construct superconcentrators. Intuitively an expander graph is one in which every subset, which is not too large, has a big neighborhood. More particularly, we say that a graph on n vertices is an expander if there are constants $c_1, c_2 > 0$ so that if S is a subset of the vertices with $|S| \leq c_1 n$ then $N(S)$, the neighborhood of S which at this time we will define to be vertices adjacent to at least one vertex in S, satisfies $|N(S)| \geq (1 + c_2)|S|$.

The constant c_1 tells us what we mean by a subset which is not too large, in this case we can have up to some positive fraction of the vertices. The constant c_2 is a measure of how much we are guaranteed to grow by as we look at neighborhoods. For expander graphs we usually try to make both c_1 and c_2 large enough to suit our purposes, an example of what this means will be given below.

We also note that the definition of expander graphs can change depending on how we define $N(S)$, so for example, we could talk about vertex neighborhoods of a set consisting of vertices adjacent to some vertices in a set, or edge neighborhoods of a set consisting of edges incident to a set.

18.2 Superconcentrator

A superconcentrator, $S(n)$, is a graph which efficiently connects inputs and outputs together as explained as follows. Pictorially, we think of what is shown in Figure 18.1.

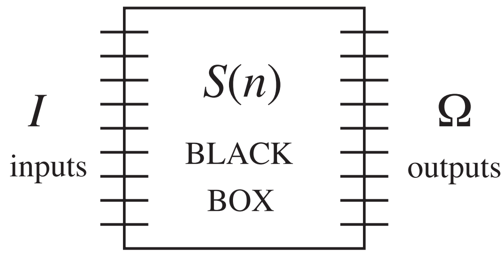

Figure 18.1: A pictorial interpretation of superconcentrator $S(n)$

What we have is a set I of inputs, a set Ω of outputs with $|I| = |\Omega| = n$. Between the inputs and outputs is the superconcentrator graph (our "black box") for which given $A \subseteq I$, $B \subseteq \Omega$ with $|A| = |B|$ there exist vertex disjoint paths joining the vertices in A to the vertices of B. Here we do not care what vertex in A goes to what vertex in B, i.e., we do not care who is talking to who if we think of this as a communications network.

One trivial way to be able to construct our black box is to connect all of the vertices in I to those of Ω. This requires n^2 edges, which as n gets large grows very quickly. We would expect that we can do better than this. Clearly we need at least n edges, and so the best possible that we can hope for is dn, for some constant d. This turns out to be possible and we will "open" such a black box for which we can use $d = 76$. The best lower bound for d is 5 due to Lev and Valiant [31, 43], the best known upper bound for d is 36. We now outline an inductive construction for our superconcentrator.

18.2.1 Our first connections

As a first step in our construction we include n edges that connects each vertex in I to a unique vertex in Ω. This is shown in Figure 18.2.

In the same figure we have outlined some sets A and B that we want our superconcentrator to connect. Note that it might happen that there are vertices in A connected directly to vertices in B, for such vertices we make the connection. After making such connections the size of what remains in A and B is at most $n/2$.

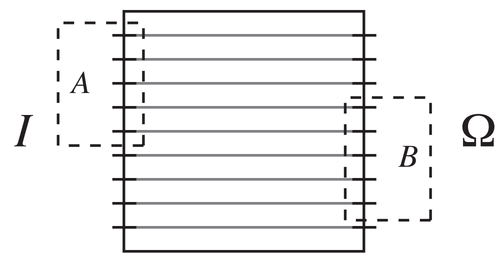

Figure 18.2: First step in opening our black box

So we are now left with the problem of connecting a set A in I to a set B in Ω where $|A| = |B| \leq n/2$. While initially it might not seem that we have done much, notice that since we are restricting ourselves only to sets of size $n/2$ (i.e., sets which are not too large) we are now putting ourselves in the position where we can use expander graphs.

18.2.2 Reducing to a smaller superconcentrator

We can now make the remaining connections by using a smaller superconcentrator (this is the inductive part of our argument, once we have shown how to connect it back to a smaller superconcentrator then we can repeat this as many times as needed to get down to a case involving a trivial superconcentrator thus giving us what we want). The best way to think of the construction is by picture, so consider Figure 18.3

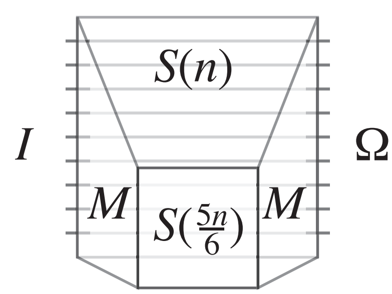

Figure 18.3: Connecting to a smaller superconcentrator

We still have the edges we initially included. We then connect the n vertices of the inputs in $S(n)$ to the $\frac{5}{6}n$ vertices of the inputs in $S(\frac{5}{6}n)$ by a bipartite graph M, and do a similar process for the outputs. Assuming we can construct $S(\frac{5}{6}n)$ then all that remains is to describe what properties M needs to satisfy, and show that such an M exists.

The defining characteristic of our bipartite graph M will be that for all sets S

with $|S| \leq \frac{1}{2}n$ that $|N(S)| \geq |S|$, as seen in Figure 18.4.

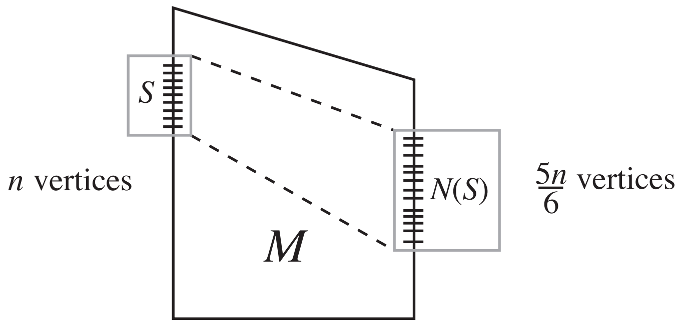

Figure 18.4: Our graph M

To do this we split up the vertices on the left hand side as shown in Figure 18.5.

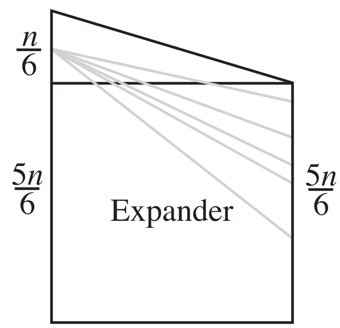

Figure 18.5: (De)constructing M

The key to constructing M is to do it in two parts. We split the n vertices on the left of M into two sets of size $\frac{1}{6}n$ and $\frac{5}{6}n$ respectively. For the set of $\frac{1}{6}n$ we connect each such vertex to five distinct vertices on the right (i.e., no overlap of vertices on the left between two such vertices). For the other $\frac{5}{6}n$ vertices we will put in an expander graph. Note that we cannot just put in a straight connection as we did when constructing our black box. The first question to ask is what parameters (i.e. c_1 and c_2 as mentioned earlier) do we want to have for our expander graph.

Since the $\frac{1}{6}n$ vertices expand to five times their original size we see that if we start with at least $\frac{1}{10}n$ vertices in S in these vertices then we will have no problems expanding. So our extreme case we need to consider for our expander graph is when we have $c_1 = (\frac{1}{2} - \frac{1}{10}) = \frac{2}{5}$ of the vertices on the left, and we need this to expand to at least $\frac{1}{2}$ of the vertices on the right. Putting this in terms of $N = \frac{5}{6}n$ we need to have for $|S| \leq \frac{12}{25}N$ on the left hand side of the expander portion grow to size at least $\frac{3}{5}N$. From this it follows that we need to have $c_2 = \frac{1}{4}$.

To build the desired expander we will use a Ramanujan graph. This is a graph that can be built using number theory properties which will have degree a prime number $+1$. We will not describe the construction of Ramanujan graphs here, just assume their existence, and that their key property is that they are good expanders. In our particular case, we will want to use a Ramanujan graph with degree 6.

Counting the number of edges of M we will have 5 for every vertex in the $\frac{1}{6}n$ portion and 6 for every vertex in the $\frac{5}{6}n$ portion. Tallying, this gives a total of $\frac{35}{6}n$ edges in M. Using this, we can calculate the appropriate value of d, note that we have n edges from our first step then twice the number of edges of M along with the number of edges in our smaller superconcentrator. So d should satisfy the following,

$$dn = n + d\frac{5}{6}n + 2\frac{35}{6}n \quad \Rightarrow \quad d = 76.$$

18.2.3 Random graphs as expanders

The key to our construction above is that we have an expander which only has degree 6. For those more familiar with random graphs a good exercise is to prove that a random degree 6 graph has good expansion. The difficulty in doing this is first establishing a random graph model.

There are several approaches, one is to first construct random matchings (random graphs of degree 1), then to construct a random graph of degree 6 combine 6 of these random matchings. In this situation it might be possible that we get multiedges, but it can be shown that this is a rare occurrence.

We can also try to analyze all of the graphs which have degree 6, the difficulty in this (compared to most random graph models) is that there is some interdependece among wether or not to include edges, making it difficult to work with.

Another choice is to work with $G_{n,p}$ model with the probability of including any particular edge as $p = \frac{6}{n}$. Intuitively these are easier to work with but hard to control. For example it can be shown that the size of the maximal is $\sim c \log n$, which will get bigger than 6. It is still easy to show that this is an expander graph using traditional tools of random graphs.

18.3 Eigenvalue certification of expander graphs

We want to find a way to easily test if a graph is an expander graph. We will develop a tool which will allow us to use eigenvalues to certify the expansion property. This makes random constructions just as good as explicit constructions.

We will develop a bipartite version of the result. So suppose that G is a bipartite graph with $V = X \cup Y$ and $|X| = |Y| = n$, further suppose that G is k-regular. Note that the adjacency matrix of G has the following form, where X and Y are marked to indicate how we index the rows/columns of the adjacency matrix.

$$A = \begin{array}{c} \\ X \\ Y \end{array} \begin{array}{c} \begin{array}{cc} X & Y \end{array} \\ \left| \begin{array}{cc} O & M \\ M^T & O \end{array} \right| \end{array} .$$

In particular the adjacency matrix M whose rows are indexed by X an whose columns are indexed by Y contains all of the information about G. It is this M that we will work with.

The matrix MM^* is a symmetric positive semi-definite matrix, in particular all of the eigenvalues of MM^* are nonnegative. It follows from the Peron-Frobenius Theorem that the largest eigenvalue of MM^* is k^2, let $\rho^2 k^2$ denote the second largest eigenvalue. Then we have that $\rho \leq 1$ (note it is also possible to show that $\rho \geq 1/\sqrt{k}$).

The key result is that for all S that

$$\frac{|N(S)|}{|S|} \geq \frac{1}{\rho^2 + (1-\rho^2)\frac{|S|}{n}}.$$

Note that when $|S|$ is small compared to n the right hand side is approximately $1/\rho^2$. In the other extreme when $|S| = n$ there can be no expansion and note that the right hand side is 1 as should be predicted. The way to think of the $(1-\rho^2)\frac{|S|}{n}$ term is that it has a tapering effect on the size of our expansion as $|S|$ increases. This result will be proved in the next lecture. For Ramanujan graphs it can be shown that $\rho = \frac{2\sqrt{k-1}}{k}$. Using this along with the formula above shows that the Ramanujan graph we used as our expander graph satisfies our desired criteria.

Lecture 19

Spectral Analysis of Expander Graphs

Kevin Costello

19.1 An Eigenvalue Bound on Expansion for Bipartite Graphs

Although it is possible to check directly whether a graph is an expander graph, doing so is computationally very expensive since the expansion property must be checked for every subset of size at most $c_1 n$. Our goal is to prove a theorem which will allow us to certify that a bipartite graph is an expander using a bound on a single numerical value.

Theorem 19.1. *Let G be a k-regular bipartite graph between X and Y with $|X| = |Y| = n$. Let M be the $n \times n$ adjacency matrix of X and Y, and assume the second largest eigenvalue of MM^* is at most $\rho^2 k^2$. Then for all $S \subseteq X$,*

$$\frac{|N(S)|}{|S|} \geq \frac{1}{\rho^2 + (1-\rho^2)\frac{|S|}{n}}.$$

This enables us to feasibly construct expanders using probabilistic algorithms. As mentioned in the previous lecture, a random k-regular bipartite graph is an expander with high probability. Assuming the slightly stronger result that random k-regular bipartite graphs will usually also have small ρ, we can construct a expander by repeatedly taking random graphs until we find one with sufficiently small ρ. Alternatively, this theorem can be used to verify that an explicit construction gives an expander graph by bounding the eigenvalues of that construction.

Proof of Theorem 19.1. Fix $S \subseteq X$, and let χ_S be the indicator function of S, i.e.

$$\chi_S(x) = \begin{cases} 1 & \text{if } x \in S, \\ 0 & \text{if } x \notin S. \end{cases}$$

As MM^* is a symmetric matrix, it has a complete orthonormal set of (left) eigenvectors ϕ_i such that $\phi_i MM^* = \lambda_i^2 \phi_i$. Without loss of generality assume the eigenvalues are arranged in decreasing order, so $\lambda_0 = k$ and $\phi_0 = \langle \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}} \rangle$.

Consider the expression $f(S) = \frac{1}{k^2} \chi_S MM^* \chi_S^*$. On the one hand, we have

$$\frac{1}{k^2}\chi_S(MM^*\chi_S^*) = \frac{1}{k^2}\Big(\sum_{i=0}^{n-1}\langle \chi_S, \phi_i\rangle \phi_i\Big)\Big(\sum_{j=0}^{n-1}\langle \chi_S, \phi_j\rangle \lambda_i^2 \phi_j^*\Big) = \frac{1}{k^2}\sum_{i=0}^{n-1}\langle \chi_S, \phi_i\rangle^2 \lambda_i^2,$$

where the latter equality comes from the orthonormality of the ϕ_i. But for $i > 0$ we have that $\lambda_i^2 \leq \rho^2 k^2$, and we can explicitly compute

$$\langle \chi_S, \phi_0 \rangle = \sum_{x \in S} \frac{1}{\sqrt{n}} = \frac{|S|}{\sqrt{n}}.$$

Plugging these bounds in, we get that

$$f(S) = \frac{1}{k^2}\left(\frac{k^2|S|^2}{n} + \sum_{i=1}^{n-1}\langle \chi_S, \phi_i\rangle^2 \lambda_i^2\right) \leq \frac{|S|^2}{n} + \rho^2 \sum_{i=1}^{n-1}\langle \chi_S, \phi_i\rangle^2.$$

Since by the Pythagorean Theorem $\sum \langle \chi_S, \phi_i \rangle^2 = \langle \chi_S, \chi_S \rangle = |S|$, we can simplify the right hand side down further to get

$$f(S) \leq \frac{|S|^2}{n} + \rho^2\Big(|S| - \frac{|S|^2}{n}\Big) = \rho^2|S| + (1-\rho^2)\frac{|S|^2}{n}$$

On the other hand,

$$f(S) = \frac{1}{k^2}\sum_{u \in S}\sum_{v \in S} |\{w \in T | u \sim w \text{ and } v \sim w\}|.$$

Equivalently, $f(S)$ is the sum over $w \in N(S)$ of the number of pairs of neighbors of w in S, so

$$f(S) = \frac{1}{k^2} \sum_{w \in N(S)} |N(w) \cap S|^2.$$

Applying the Cauchy-Schwarz inequality with $a_i = |N(w) \cap S|$ and $b_i = 1$, we get

$$f(S) \geq \frac{(\sum_{w \in N(S)} |N(w) \cap S|)^2}{k^2 |N(S)|} = \frac{(\sum_{u \in S} deg(u))^2}{k^2 |N(S)|} = \frac{|S|^2}{|N(S)|^2}.$$

Combining the upper and lower bounds on $f(S)$, we have that

$$\frac{|S|^2}{|N(S)|^2} \leq \rho^2 |S| + (1 - \rho^2) \frac{|S|^2}{n}.$$

Dividing by $|S|$ and flipping both sides over, we get

$$\frac{|N(S)|}{|S|} \geq \frac{1}{\rho^2 + (1 - \rho^2) \frac{|S|}{n}},$$

which is the desired result. □

19.2 Explicit expanders via Ramanujan Graphs

In the construction of our superconcentrator what we needed to have was a linear expander, one such that the number of edges grows linearly with the number of vertices. If we assume that our expander is a regular bipartite graph, this corresponds to an infinite family of graphs such that n tends to infinity, but k remains constant.

Theorem 19.1 says that any such family whose second largest eigenvalue is sufficiently small will have an expansion property, and the smaller the bound on the eigenvalue the better the expansion property we can prove. Therefore it is desirable to construct families of k-regular graphs whose second eigenvalue is as small as possible. It is a corollary of a result of Alon and Boppara [32] that in any such family

$$\liminf_{n \to \infty} \lambda_1 \geq 2\sqrt{k-1},$$

so the best uniform bound on ρ that we can hope for is $\rho = \frac{2\sqrt{k-1}}{k}$. Individual graphs satisfying $\lambda_1 \leq 2\sqrt{k-1}$ are known as **Ramanujan Graphs**. Families of sparse Ramanujan graphs were first constructed using number theoretic methods [32], but more examples have been constructed, such as the coset graph for finite fields [10].

It should be noted that although the Ramanujan Graphs give us the best possible eigenvalue bound for expansion, that bound may not itself be optimal, especially for small graphs. For example, our bounds give that a k-regular Ramanujan graph will have for any S that

$$\frac{|N(S)|}{|S|} \geq \frac{k^2}{4(k-1) + \frac{(k^2-4k+4)|S|}{n}}.$$

Taking $|S|$ to be small relative to n, this gives a bound of about $k|S|/4$ for $|N(S)|$. But for small $|S|/n$ probabilistic methods give that $|N(S)| \approx k|S|$ (it is unlikely for any one vertex to be adjacent to two members of S). Using spectral theory alone "cheated" us out of a factor of 4 for small $|S|/n$.

Bibliography

[1] M. Ajtai, J. Komlós, and E. Szemerédi, "On the Erdős-Sós conjecture - an approximate version in the dense case", manuscript (1991).

[2] M. Ajtai, J. Komlós, and E. Szemerédi, "On a conjecture of Loebl", Proceedings of the 7th International Conference on Graph Theory, Combinatorics, and Algorithms, (Kalamazoo, Michigan, 1992), 1135-1146, Wiley, New York, 1995.

[3] F.A. Behrend, "On sets of integers which contain no three terms in arithmetical progression.", *Proc. Nat. Acad. Sci. U.S.A.* **32** (1946), 331–332.

[4] E.L. Berlekamp, "A construction for partitions which avoid long arithmetic progressions", Canad. Math. Bull., 11 (1968), 409–414.

[5] Y. Bilu, "Structure of sets with small sumsets", Structure Theory of Set Addition, Asterisque 258, Soc. Math. France, Montrouge, 1999, 77–108.

[6] J.A. Bondy and M. Simonovits, "Cycles of even length in graphs", J. Comb. Theory (B) 16 (1974), 97–105.

[7] S.A. Burr and P. Erdős, "On the magnitude of generalized Ramsey numbers", *Infinite and Finite Sets*, Vol. 1, Colloq. Math. Soc., János Bolyai, North Holland, Amsterdam/London, 1975.

[8] M. Chang, "A polynomial bound in Freiman's Theorem", Duke Mathematics Journal, Vol. 113, No.3, 2002, 399–419.

[9] G. Chen, R.H. Schelp, "Graphs with linearly bounded Ramsey numbers", J. Combin Theory Ser. B 57 (1993) 193–206.

[10] F.R.K. Chung, "Diameters and Eigenvalues", J. Amer. Math. Soc. 2 (1989), 187–196.

[11] F.R.K. Chung, *Spectral Graph Theory.* CBMS No. 92, American Mathematical Society, 1997.

[12] F.R.K. Chung and R.L. Graham, "Quasi-random set systems", J. Amer. Math. Soc. 4 (1991), 151–196

[13] F.R.K. Chung, "Regularity lemmas for hypergraphs and quasi-randomness", Random Structures Algorithms 2 (1991), no. 2, 241–252.

[14] F.R.K. Chung and R.L. Graham, R "Quasi-random subsets of Z_n", J. Comb. Th. (A) 61 (1992), 64–86.

[15] F.R.K. Chung, R.L. Graham and R.M. Wilson, "Quasi-random graphs", Combinatorica, 9 (1989), 345–362.

[16] F.R.K. Chung and L. Lu, "An upper bound for the Turan number $t_3(n, 4)$", J. Comb. Theory(A) 87 (1999), 381–389.

[17] C. Chvatál, E. Szemerédi, W.T. Trotter, Jr. and V. Rödl, "The Ramsey number of a graph with bounded maximum degree", J. Combin Theory Ser. B 34 (1983), no. 3, 239–243.

[18] N. Eaton, "Ramsey number for sparse graphs", Discrete Math, 185, 63–75 (1998).

[19] P. Erdős, "Extremal problems in graph theory", Theory of Graphs and its Applications, Proc. Symp. (Smolenice, 1963) (M. Fielder, ed.), 29–36. New York: Academic Press, 1965.

[20] P. Erdős, "On sequences of integers no one of which divides the product of two others and on some related problems", *Mitt. Forsch.-Inst. Math. Mech. Univ. Tomsk* 2 (1938), 74–82.

[21] P. Erdős, Z. Füredi, M. Loebl, and V.T. Sós, "Discrepancy of trees", Studia Sci. Math. Hungar. 30 (1995) no. 1-2, 47-57.

[22] P. Erdős and A.H. Stone, "On the structure of linear graphs", Bull. Amer. Math. Soc. 52 (1946), 1087–1091.

[23] W.T. Gowers, "A new proof of Szemerédi's theorem", Geom. Funct. Anal., 11 (2001), no. 3, 465–588.

[24] W.T. Gowers, "Some Unsolved Problems in Additive/Combinatorial Number Theory", www.dpmms.cam.ac.uk/˜wtg10/addnoth.survey.dvi.

[25] R. Graham, V. Rödl and A. Rucinski, "On graphs with linear Ramsey numbers", J. Graph Theory 35 (2000), no. 3, 176–192.

[26] B. Green and T. Tao, "The primes contain arbitrarily long arithmetic progressions", front.math.ucdavis.edu/math.NT/0404188.

[27] R. Horn and C. Johnson, “Matrix Analysis”, Cambridge University Press, 1985.

[28] H.A. Kierstead and W.T. Trotter, “Planar graph coloring with an uncooperative partner”, J. Graph Theory 18 (1994) 569–584.

[29] J. Komlós, G.N. Sárközy, and E. Szemerédi, “On the Pósa-Seymour conjecture”, Journal of Graph Theory 29 (1998), 167-176.

[30] J. Komlós and M. Simonovits, *Szemerédi's Regularity Lemma and its applications in graph theory*, Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993), 295–352, Bolyai Soc. Math. Stud., **2**, János Bolyai Math. Soc., Budapest, 1996.

[31] G. Lev, “Size bounds for parallel algorithms for networks”, Ph.D. Thesis, Department of Computer Science, University of Edinburg.

[32] A. Lubotzky, R. Philips, and P. Sarnak, “Ramanujan Graphs”, Combinatorica, 9 (1988), 261–277.

[33] M.B. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets* Graduate Texts in Mathematics 165, Springer-Verlag, 1996.

[34] V. Rödl, P. Frankl, “The uniformity lemma for hypergraphs”, Graphs Combin. 8 (1992) 309–312.

[35] V. Rödl and R. Thomas, “Arrangeability and clique subdivisions”, The mathematics of Paul Erdős, II, 236–239, Algorithms Combin., 14, Springer Berlin, 1997.

[36] K.F. Roth, “On certain sets of integers.”, *J. London Math. Soc.* **28** (1953), 104–109.

[37] I.Z. Ruzsa and E. Szemerédi. “Triple systems with no six points carrying three triangles.” Colloqui Mathematica Societatis János Bolyai, 18. Combinatorics, Keszthely (Hungary), 1976, 199–245.

[38] S. Shelah, “Primitive recursive bounds for van der Waerden numbers”, J. Amer. Math. Soc. 1 (1988), no. 3, 683–697.

[39] J. Solymosi. “Note on a Generalization of Roth's Theorem”. Discrete and Computational Geometry: the Goodman-Pollack Festschrift. Ed. Aranov et. al. Springer, Berlin, 2003.

[40] E. Szemerédi, “On sets of integers containing no four elements in arithmetic progression”, Acta Math. Acad. Sci. Hungar. 20 (1969), 89-104

[41] E. Szemerédi, "On sets of integers containing no k elements in arithmetic progression", Acta Arith. 27 (1975), 299-345.

[42] P. Turán, "Eine Extremalaufgabe aus der Graphentheorie", Mat. Fiz. Lapok 48 (1941), 436–452.

[43] L.G. Valiant, "A scheme for fast parallel communication", SIAM J. Comput. 11 (1982), 350–361.

[44] B.L. van der Waerden, "Beweis einer Baudetschen Vermutung", Nieuw Arch. Wisk. 15 (1927) 212–216.

[45] Y. Zhao, "Proof of the (n/2-n/2-n/2) conjecture for large n", www2.math.uic.edu/˜zhao/publications/lobel_conj.pdf.

Index