

# Extremal Graph Theory

Instructor: Asaf Shapira  
Scribed by Guy Rutenberg\*

Fall 2013<sup>†</sup>

## Contents

<b>1</b>	<b>First Lecture</b>	<b>2</b>
1.1	Mantel's Theorem	3
1.2	Turán's Theorem	4
1.3	Ramsey's Theorem	5
<b>2</b>	<b>Second Lecture</b>	<b>7</b>
2.1	The Erdős-Stone-Simonovits Theorem	7
2.2	The Zarankiewicz Problem for $K_{2,2}$	9
2.3	The Zarankiewicz Problem for General $K_{s,t}$	10
<b>3</b>	<b>Third Lecture</b>	<b>11</b>
3.1	Applications of the Zarankiewicz Problem	11
3.2	The Turán Problem for Trees	12
3.3	The Girth Problem and Moore's Bound	13
3.4	Application of Moore's Bound to Graph Spanners	14
<b>4</b>	<b>Fourth Lecture</b>	<b>15</b>
4.1	The Turán Problem for Long Cycles	15
4.2	Pancyclic Graphs and Bondy's Theorem	16
4.3	The Moon-Moser Inequalities	17
<b>5</b>	<b>Fifth Lecture</b>	<b>20</b>
5.1	The Hypergraph Turán Problem	20
5.2	Extremal Problems Related to Graph Minors	22
5.3	Application of Topological $K_t$ -Minors to Graph Linkage	24
<b>6</b>	<b>Sixth Lecture</b>	<b>25</b>
6.1	Introduction to Szemerédi's Regularity Lemma	25
6.2	The Triangle-Removal Lemma and Roth's Theorem	26
6.3	Proof of the Regularity Lemma	28

---

\*Submit comments to [guyrutenberg@mail.tau.ac.il](mailto:guyrutenberg@mail.tau.ac.il)

<sup>†</sup>Typeset using L<sup>A</sup>T<sub>E</sub>X on March 2, 2013.

<b>7</b>	<b>Seventh Lecture</b>	<b>30</b>
7.1	The Counting Lemma	30
7.2	Another Proof of the Erdős-Stone-Simonovits Theorem	31
7.3	Ramsey Numbers of Bounded Degree Graphs	32
<b>8</b>	<b>Eighth Lecture</b>	<b>32</b>
8.1	The Induced Ramsey Theorem	32
8.2	The (6, 3)-Problem and the Induced Matchings Problem	35
<b>9</b>	<b>Ninth Lecture</b>	<b>36</b>
9.1	Behrend's Construction	36
9.2	Lower Bound for the Triangle Removal Lemma	38
9.3	Deducing Szemerédi's Theorem from the Hypergraph Removal Lemma	38
<b>10</b>	<b>Tenth Lecture</b>	<b>41</b>
10.1	The Ramsey-Turán Problem	41
10.2	Quasi-Random Graphs	43
<b>11</b>	<b>Eleventh Lecture</b>	<b>45</b>
11.1	The Chung-Graham-Wilson Theorem	45
11.2	An Algorithmic Version of the Regularity Lemma	47
11.3	Approximating MAX-CUT using the Regularity Lemma	48
<b>12</b>	<b>Twelfth lecture</b>	<b>50</b>
12.1	Hypergraph Ramsey Numbers	50
<b>13</b>	<b>Notations</b>	<b>52</b>
	<b>References</b>	<b>53</b>

## 1 First Lecture

All the graphs considered here are finite, simple, loop-free, etc. Unless otherwise mentioned, they will also be undirected. The following (very standard) notation is used throughout the notes. For a graph  $G$  we set  $m = |E(G)|$  and  $n = |V(G)|$ . We use  $N(x)$  to denote the neighborhood of a vertex  $x$ , and  $d(x)$  to denote its degree (so  $d(x) = |N(x)|$ ). We will use  $\delta(G)$  to denote the smallest degree of a vertex in  $G$ , and  $\Delta(G)$  to denote the largest degree. We use  $e(G)$  to denote the number of edges of a graph  $G$  (so  $e(G) = |E(G)|$ ). We will sometimes use  $e(S)$  to denote the number of edges inside a vertex set  $S$  which is the subset of the vertex set of some graph. Similarly  $e(A, B)$  will denote the number of edges connecting a vertex in  $A$  to a vertex in  $B$ . The complete graph on  $n$  vertices is denoted  $K_n$ , while the complete  $r$ -partite graph with partitions of sizes  $n_1, n_2, \dots, n_r$  is denoted  $K_{n_1, \dots, n_r}$ .

We will frequently use *Jensen's inequality*, which states that if  $f$  is a convex function then for every  $x_1, \dots, x_n$  and  $\alpha_1, \dots, \alpha_n$  with  $0 \leq \alpha_i \leq 1$  and  $\sum_i \alpha_i = 1$  we have

$$f\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i f(x_i)$$

**Exercise.** Prove Jensen's Inequality (hint: use induction on  $n$ ).

## 1.1 Mantel's Theorem

**Theorem 1.1** (Mantel). *If  $G$  is a triangle-free graph, then  $m \leq \lfloor \frac{n^2}{4} \rfloor$ . Furthermore, the only triangle-free graph with  $\lfloor \frac{n^2}{4} \rfloor$  edges is the complete bipartite graph  $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ .*

**Proof** (First): It is easy to verify the proposition for  $n = 3$  and  $n = 4$ , so assume  $n > 4$ . Let  $(x, y)$  be an edge in  $G$ . Clearly  $x$  and  $y$  cannot have a common neighbor (otherwise the graph will have a triangle), hence  $d(x) + d(y) \leq n$ . By removing  $x$  and  $y$ , we get a triangle-free graph  $H = G - \{x, y\}$ , which has, by induction, at most  $\frac{(n-2)^2}{4}$  edges. Thus

$$m \leq \left\lfloor \frac{(n-2)^2}{4} \right\rfloor + n - 1 = \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Now, if  $m = \lfloor \frac{n^2}{4} \rfloor$ , then all the above inequalities must be equalities. In particular, we must have  $d(x) + d(y) = n$ . It is then easy to see that this implies that  $N(x) \cup N(y) = V(G)$ . Since  $G$  is triangle-free, we know that  $N(x)$  and  $N(y)$  are both independent sets, implying that  $G$  is bipartite (with bipartition  $N(x), N(y)$ ). Finally, note that the only bipartite graph on  $n$  vertices with  $\lfloor \frac{n^2}{4} \rfloor$  edges is  $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ .  $\square$

**Exercise.** Prove Mantel's Theorem using induction on  $n$ , but remove only a single vertex each time.

**Proof** (Second): As  $G$  is triangle-free we infer (as in the previous proof) that  $d(x) + d(y) \leq n$  for every edge  $xy \in E(G)$ . Summing this inequality over all the edges we get

$$\sum_{xy \in E} (d(x) + d(y)) \leq mn.$$

On the other hand, by Jensen's inequality

$$\sum_{xy \in E} (d(x) + d(y)) = \sum_x d^2(x) \geq n \underset{\text{Jensen}}{\left( \frac{\sum_x d(x)}{n} \right)^2} = n \underset{\sum_x d(x) = 2m}{\frac{4m^2}{n^2}} = 4 \frac{m^2}{n}$$

thus  $4 \frac{m^2}{n} \leq mn$ , giving the desired bound. Furthermore if  $m = \lfloor \frac{n^2}{4} \rfloor$ , then all the above inequalities must be equalities. In particular, we must have  $d(x) + d(y) = n$  for any edge  $xy \in E(G)$ . As in the previous proof, this means that  $G$  must be  $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ .  $\square$

**Proof** (Third): Let  $A$  be an independent set of maximal size in  $G$ . Since  $G$  is triangle-free we know that  $N(x)$  is an independent set for every  $x$ , implying that  $\Delta(G) \leq |A|$  (where we use  $\Delta(G)$  to denote the maximum degree of a vertex in  $G$ ). Setting  $B = V \setminus A$ , we see that as  $A$  is an independent set, every edge touches  $B$ . Therefore,

$$m \leq \sum_{x \in B} d(x) \leq |B| \Delta(G) \leq |B| |A| \leq \left( \frac{|A| + |B|}{2} \right)^2 \leq \frac{n^2}{4},$$

implying the desired bound on  $m$ . Finally, if  $m = \lfloor \frac{n^2}{4} \rfloor$ , then the above inequalities are equalities and so  $B$  must be an independent set, implying that  $G$  is bipartite. As in previous proofs, this means that  $G$  must be  $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ .  $\square$

**Proof (Fourth):** Let  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  be a vector satisfying  $\sum x_i = 1$  and  $0 \leq x_i \leq 1$ . Define  $f(x) = \sum_{ij \in E} x_i x_j$ . We first observe that taking  $x = (\frac{1}{n}, \dots, \frac{1}{n})$  gives an  $x$  for which  $f(x) \geq \frac{m}{n^2}$ . We now claim that  $f(x) \leq \frac{1}{4}$  for every  $x$ . Indeed, if  $ij \notin E$  then shifting the weight assigned to  $x_i$  to  $x_j$  (assuming the weight of the neighbors of  $x_j$  is at least as large as the weight assigned to the neighbors of  $x_i$ ) does not decrease the value of  $f$ . If we repeat this process we get that  $f$  is maximized when all the weight is concentrated on a clique. As  $G$  is  $K_3$ -free, the weight is concentrated on two vertices, hence  $f(x) = x_i x_j \leq \frac{1}{4}$ . Combining the above two observations, we see that  $\frac{m}{n^2} \leq \frac{1}{4}$ .  $\square$

## 1.2 Turán's Theorem

**Theorem 1.2** (Turán's Theorem (weak version)). *If  $G$  is  $K_{t+1}$ -free then  $m \leq (1 - \frac{1}{t}) \frac{n^2}{2}$ .*

**Definition.** Turán's graph, denoted  $T_r(n)$ , is the complete  $r$ -partite graph on  $n$  vertices which is the result of partitioning  $n$  vertices into  $r$  almost equally sized partitions  $(\lfloor n/r \rfloor, \lceil n/r \rceil)$  and taking all edges connecting two different partition classes (note that if  $n \leq r$  then  $T_r(n) = K_n$ ). Denote the number of edges in Turán's graph by  $t_r(n) = |E(T_r(n))|$ .

**Fact.** *Turán's graph  $T_r(n)$ , maximizes the number of edges among all  $r$ -partite graphs on  $n$  vertices, and it is the only graph which does so.*

**Exercise.** Prove the above assertions.

**Theorem 1.3** (Turán's Theorem (strong version)). *If  $G$  is  $K_{r+1}$ -free then  $m \leq t_r(n)$ . Furthermore, equality holds if and only if  $G = T_r(n)$ .*

**Proof (First):** We start with the weak version, and proceed by induction on  $n$ , noting that the assertions is trivial for  $n \leq r$ . Assume  $G$  is maximal  $K_{r+1}$ -free graph, hence it has a  $K_r$  subgraph. We denote this copy of  $K_r$  by  $X$ , and the rest of  $G$  by  $Y$ . Since  $G$  is  $K_{r+1}$ -free, the number of edges from any vertex in  $Y$  to  $X$  is at most  $r - 1$ , therefore the number of edges between  $X$  and  $Y$  is bounded by  $(n - r)(r - 1)$ . By induction the number of edges in  $Y$  is bounded by  $(1 - \frac{1}{r}) \frac{(n-r)^2}{2}$ , thus the number of edges in  $G$  is at most

$$\binom{r}{2} + (n - r)(r - 1) + (1 - \frac{1}{r}) \frac{(n - r)^2}{2} = (1 - \frac{1}{r}) \frac{n^2}{2},$$

proving the weak statement.

A close look at the above proof gives the strong statement as well (we again proceed by induction on  $n$ ). We first observe that removing a single vertex from each of the clusters of  $T_r(n)$  gives  $T_r(n - r)$ , implying that

$$\binom{r}{2} + (n - r)(r - 1) + t_r(n - r) = t_r(n).$$

Hence, if  $m = t_r(n)$ , then  $Y$  must contain  $t_r(n - r)$  edges and induction then gives that  $G$  restricted to  $Y$  must be isomorphic to  $T_r(n - r)$  and is in particular a complete  $r$ -partite graph. Furthermore, equality also means that every vertex of  $Y$  must be connected to exactly  $r - 1$  of the vertices of  $X$ . It is then easy to see that since  $G$  is  $K_{r+1}$ -free that this means that  $G$  is also  $r$ -partite (we can add each vertex  $x$  of  $X$  to the cluster of  $T_r(n - r)$  that has the vertices that are connected to all the vertices of  $X$  besides  $x$ . Convince yourself that this is well-defined). Since  $T_r(n)$  is the only  $r$ -partite graphs with  $t_r(n)$  edges, the proof is complete.  $\square$

**Proof (Second):** We will prove that if  $G$  is a  $K_{r+1}$ -free graph on  $n$  vertices with the maximum possible number of edges, then it must be a complete  $k$ -partite graph. Since  $G$  is assumed to be  $K_{r+1}$ -free this

means that  $k \leq r$ , and so  $m \leq t_r(n)$  and that equality can only hold if  $G = T_r(n)$  (again, since  $T_r(n)$  is the only  $r$ -partite graph with  $t_r(n)$  edges).

Showing that  $G$  is a complete  $k$ -partite graph is equivalent to showing that for any triple of vertices  $x, y, z \in V$ , if  $xy \notin E$  and  $yz \notin E$  then  $xz \notin E$ . So assume  $xy \notin E$  and  $yz \notin E$  but  $xz \in E$ . If  $d(y) < d(x)$ , we can replace  $y$  by a copy of  $x$ . The resulting graph will still be  $K_{r+1}$ -free and will contain more edges, thus  $d(y) \geq d(x)$ . Similarly,  $d(y) \geq d(z)$ . Replacing both  $x$  and  $z$  by two copies of  $y$ , it is easy to check that we get a graph that is still  $K_{r+1}$ -free and contains more edges than  $G$  which is again a contradiction.  $\square$

**Proof** (Third): We will prove by induction on  $n$  that if  $G$  has  $t_r(n)$  edges and  $G$  is  $K_{r+1}$ -free then  $G = T_r(n)$  (convince yourself that this indeed proves the theorem). The claim is clearly true when  $n \leq r$ . Let  $x$  be a vertex of minimal degree. Then  $\delta(T_r(n)) \geq d(x)$  (as  $T_r(n)$  maximizes  $\delta(G)$  in comparison to every other graph with  $t_r(n)$  edges). We thus get

$$e(G \setminus \{x\}) \geq e(G) - d(x) = t_r(n) - d(x) = t_r(n-1),$$

where the last equality follows from observing that  $T_r(n-1)$  is obtained from  $T_r(n)$  by removing a vertex from a cluster of size  $\lceil n/r \rceil$ , which amounts to removing  $d(x)$  edges. Since  $G \setminus \{x\}$  is  $K_{r+1}$ -free, we get from induction that  $G \setminus \{x\} = T_r(n-1)$ , so in particular  $G \setminus \{x\}$  is  $r$ -partite. Since  $G$  is  $K_{r+1}$ -free,  $x$  is connected to vertices in at most  $r-1$  of the partition classes of  $G \setminus \{x\}$  so  $G$  is also  $r$ -partite and so must be isomorphic to  $T_r(n)$ .  $\square$

The problem of proving existence of independent sets is of course closely related to that of finding cliques as each independent set corresponds to a clique in the complement graph and vice versa.

**Proposition 1.4.**

$$\alpha(G) \geq \sum_x \frac{1}{1+d(x)}$$

where  $\alpha(G)$  is the size of the largest independent set in  $G$ .

**Proof:** Given some arbitrary ordering on the vertices of  $G$ , say  $x_1, \dots, x_n$ , let  $I$  be a vertex set with the property that  $x_i \in I$  and  $x_i x_j \in E$  then  $i < j$  (all the neighbors of  $x_i$  appear after it in the ordering). Obviously  $I$  is an independent set.

Pick a random ordering of the vertices and define  $X_v = 1 \iff v \in I$ .

$$\mathbb{E}[|I|] = \mathbb{E}\left[\sum_v X_v\right] = \sum_v \mathbb{E}[X_v] = \sum_v P[X_v = 1] = \sum_v \frac{1}{1+d(v)}$$

Thus there is an ordering for which  $|I| \geq \sum_x \frac{1}{1+d(x)}$ , and in particular  $\alpha(G) \geq \sum_x \frac{1}{1+d(x)}$ .  $\square$

### 1.3 Ramsey's Theorem

**Theorem 1.5** (Ramsey). *For every  $t$  there exists  $N = R(t)$  such that every 2-coloring of the edges of  $K_N$  has a monochromatic  $K_t$  subgraph.*

**Proof** (First): We will show that  $R(t) \leq 4^t$ . Take a graph on  $2^{2t}$  vertices, and pick an arbitrary vertex  $v$ . At least one of the colors, say black, appears on at least half of the edges incident with  $v$ . We now set  $x_1 = v$  and color it black (if the popular color was white we color  $x_1$  white). Now, remove  $v$  and vertices  $u$  so that  $vu$  is colored white and note that we are left with at least  $2^{2t-1}$  vertices. We repeat this process with whatever vertices left, resulting in  $x_2, x_3, \dots$ . After  $i$  iterations, there are at least  $2^{2t-i}$  vertices left,

hence the process will pick  $2t - 1$  vertices. At least  $t$  of them will have the same color and it is easy to see that the way we defined them guarantees that they will span a monochromatic clique of size  $t$ .  $\square$

**Proof (Second):** We prove by induction on  $s + t$  that

$$R(s, t) \leq \binom{s+t-2}{s-1},$$

where  $R(s, t)$  is the generalization of  $R(t)$  (we want either black  $K_t$  or white  $K_s$ ). Base case: when  $s = 2$  ( $t = 2$  is symmetric) then  $R(2, t) = t$  (either all the edges black and we get a black  $K_t$ , or we have a white  $K_2$  which is a single white edge). So suppose we have established the required upper bound for both  $R(s-1, t)$  and  $R(s, t-1)$ . We will now prove that

$$R(s, t) \leq R(s-1, t) + R(s, t-1). \quad (1.1)$$

Induction will then imply that

$$R(s, t) \leq \binom{s+t-3}{s-2} + \binom{s+t-3}{s-1} = \binom{s+t-2}{s-1}.$$

Take a colored graph on  $R(s-1, t) + R(s, t-1)$  vertices. Pick an arbitrary vertex  $x$ . Denote by  $A$  the vertices connected to  $x$  by black edges and by  $B$  those connected to it by white edges. Either  $|A| \geq R(s, t-1)$  or  $|B| \geq R(s-1, t)$ , w.l.o.g. assume the former. Then, by the induction hypothesis  $A$  has a white  $K_s$  or black  $K_{t-1}$  and together with  $x$  it forms a black  $K_t$ , and anyway we are done. Hence (1.1) holds, and we are done.  $\square$

**Corollary 1.6.**

$$R(t) = R(t, t) \leq \binom{2t-2}{t-1} \leq \binom{2t}{t} = \frac{(2t)!}{(t!)^2} \underset{\text{Stirling}}{\approx} \frac{\sqrt{4\pi t} \left(\frac{2t}{e}\right)^{2t}}{\sqrt{2\pi t} \left(\frac{t}{e}\right)^t} = \frac{4^t}{\sqrt{\pi t}}$$

*Remark.* Stirling approximation states that  $n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$

**Theorem 1.7.** For every  $t \geq 3$  we have  $R(t) \geq 2^{t/2}$ .

**Proof:** We first claim that if  $\binom{n}{t} 2^{1-\binom{t}{2}} < 1$  then  $R(t) \geq n$ . We can prove this using a “probabilistic argument” by noting that if we take a random coloring of  $K_n$  then the probability to have a monochromatic  $K_t$  subgraph is at most  $\binom{n}{t} 2^{1-\binom{t}{2}}$  so if the above inequality holds then with positive probability there will be no monochromatic  $K_t$  so there must be at least one coloring with no monochromatic  $K_t$ . We can also prove this using an equivalent “counting argument”, as follows; the number of coloring of the complete graph on  $n$  vertices with a monochromatic  $K_t$  is bounded by  $\binom{n}{t} \cdot 2 \cdot 2^{\binom{n}{2}-\binom{t}{2}}$  so if the above inequality holds then this number is smaller than the total number of colorings which is  $2^{\binom{n}{2}}$ , hence there must be one coloring with no monochromatic  $K_t$ .

We now just need to work out the largest  $n$  for which the above inequality holds. Note that  $\binom{n}{t} \leq \frac{n^t}{t!}$ , so taking  $n = 2^{t/2}$  gives

$$\binom{n}{t} \cdot 2^{1-\binom{t}{2}} \leq \frac{n^t \cdot 2 \cdot 2^{t/2}}{t! \cdot 2^{t^2/2}} = \frac{2^{t^2/2} 2^{1+t/2}}{t! 2^{t^2/2}} = \frac{2^{t/2+1}}{t!} < 1$$

for every  $t \geq 3$ .  $\square$

## 2 Second Lecture

In the last lecture, we proved Turán's Theorem, which says that  $ex(n, K_{r+1}) \sim (1 - \frac{1}{r})\frac{n^2}{2}$ . Today we will generalize it to arbitrary graphs.

### 2.1 The Erdős-Stone-Simonovits Theorem

**Theorem 2.1** (Erdős-Stone-Simonovits).

$$ex(n, H) = \left(1 - \frac{1}{\chi(H) - 1} \pm o(1)\right) \frac{n^2}{2}.$$

In other words for every  $H, r$  and  $\varepsilon$ , such that  $\chi(H) = r + 1$ , there exists  $n_0 = n_0(r, H, \varepsilon)$  such that for every  $n > n_0$

$$\left(1 - \frac{1}{\chi(H) - 1} - \varepsilon\right) \frac{n^2}{2} \leq ex(n, H) \leq \left(1 - \frac{1}{\chi(H) - 1} + \varepsilon\right) \frac{n^2}{2}$$

The lower bound is very easy; the Turán graph  $T_r(n)$  is  $r$ -colorable, hence it is  $H$ -free (recall that we assume  $\chi(H) = r + 1$ ). It is now easy to check (do this) that

$$e(T_r(n)) = t_r(n) \geq \left(1 - \frac{1}{r} - o(1)\right) \frac{n^2}{2},$$

so we are just left with proving the upper bound. We note that it is clearly enough to prove the upper bound only for  $K_{r+1}^v$  (which is the  $v$ -blowup<sup>1</sup> of  $K_{r+1}$ ) as if  $\chi(H) = r + 1$ , then  $H$  is a subgraph of  $K_{r+1}^v$  for  $|v| = |V(H)|$ . So our goal is now to prove the following

**Theorem 2.2** (Erdős-Stone-Simonovits (upper bound)). *For every  $r, v$  and  $\varepsilon > 0$ , there exists  $n_0 = n_0(r, \varepsilon, v)$  such that every graph on  $n \geq n_0$  vertices with  $m \geq (1 - \frac{1}{r} + \varepsilon)\frac{n^2}{2}$  edges contains  $K_{r+1}^v$ .*

It will be easier to prove Theorem 2.2 by assuming that the graph is not only dense "globally" but also "locally". This is achieved by the following claim.

**Claim 2.3.** *For every  $\varepsilon > 0$  there is  $n_2 = n_2(\varepsilon)$  so that for every  $n \geq n_2$  if  $G$  has  $n$  vertices and at least  $(1 - \frac{1}{r} + \varepsilon)\frac{n^2}{2}$  edges, then it has a subgraph  $G'$  on  $n' \geq \sqrt{\varepsilon}n/4$  vertices satisfying  $\delta(G') \geq (1 - \frac{1}{r} + \frac{\varepsilon}{2})n'$ .*

Given the claim above, it will be sufficient to prove the following variant of Theorem 2.2.

**Theorem 2.4.** *For every  $r, v$  and  $\varepsilon > 0$ , there exists  $n_1 = n_1(r, \varepsilon, v)$  such that every graph  $G$  on  $n \geq n_1$  vertices satisfying  $\delta(G) \geq (1 - \frac{1}{r} + \varepsilon)n$  contains  $K_{r+1}^v$ .*

Let us first show how to derive Theorem 2.2 from Claim 2.3 and Theorem 2.4.

**Proof** (Theorem 2.2): By Claim 2.3 if we start with  $G$  on  $n \geq n_2(\varepsilon)$  vertices we can pass to a subgraph  $G'$  on  $n' \geq \sqrt{\varepsilon}n/4$  vertices for which  $\delta(G') \geq (1 - \frac{1}{r} + \frac{\varepsilon}{2})n'$ . In order to be able to use Theorem 2.4 on  $G'$  we pick  $n_0$  such that  $n' \geq \sqrt{\varepsilon}n_0/4 \geq n_1(r, \frac{\varepsilon}{2}, v)$ , so we set  $n_0(r, \varepsilon, v) = \max\{4n_1(r, \varepsilon/2, v)/\varepsilon, n_2(\varepsilon)\}$ .  $\square$

To complete the proof of Theorem 2.2 we now turn to prove Claim 2.3 and Theorem 2.2.

<sup>1</sup>The  $v$ -blowup of a graph  $K$  is obtained by replacing every vertex  $x$  of  $K$  by an independent set  $I_x$  of size  $v$ , and replacing every edge  $(x, y)$  by a complete bipartite graph connecting  $I_x$  to  $I_y$ .

**Proof** (Claim 2.3): We prune the vertices of  $G$  until we get the appropriate subgraph. As long as  $G$  has a vertex with a degree  $\leq (1 - \frac{1}{r} + \varepsilon/2)n'$  we remove it from  $G$  ( $n'$  is the number of vertices currently left). Assume the process stopped with  $n'$  vertices. Thus

$$\begin{aligned} m &\leq n \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right) + (n-1) \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right) + \cdots + n' \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right) + \binom{n'}{2} \\ &= \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right) \binom{n+n'}{2} (n-n'+1) + \binom{n'}{2} \\ &\leq \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right) \left(\frac{n^2 - (n')^2}{2}\right) + 2n + \frac{(n')^2}{2} \\ &\leq \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right) \frac{n^2}{2} + (n')^2 + 2n. \end{aligned}$$

Since the claim's assumption is that  $m \geq (1 - \frac{1}{r} + \varepsilon) \frac{n^2}{2}$ , we can combine these lower/upper bounds on  $m$  to deduce that  $(n')^2 \geq \frac{\varepsilon n^2}{4} - 2n$ . If we add the assumption that  $n \geq \frac{16}{\varepsilon}$  then the right hand side in this last inequality is at least  $\frac{\varepsilon n^2}{8}$  implying that  $n' \geq \sqrt{\varepsilon}n/4$ , and the proof is complete (and we set  $n_1(\varepsilon) = 16/\varepsilon$  as the function in the claim's statement).  $\square$

**Proof** (Theorem 2.4): Assume the statement is true for  $r-1$ , namely that  $n_1(r-1, \varepsilon, v)$  was defined for every  $v$  and  $\varepsilon > 0$ . Given  $v$  and  $\varepsilon > 0$  we now wish to define  $n_1(r, \varepsilon, v)$ . We first ask that  $n_1(r, \varepsilon, v) \geq n_1(r-1, \varepsilon, \bar{v})$ , for a  $\bar{v}$  that we will determine later. This means that if  $G$  has at least  $n_1(r, \varepsilon, n)$  vertices and  $\delta(G) \geq (1 - \frac{1}{r} + \varepsilon)n \geq (1 - \frac{1}{r-1} + \varepsilon)n$  then  $G$  contains a copy of  $K_r^{\bar{v}}$ . Let  $B_1, \dots, B_r$  denote the  $r$  partition classes of this copy of  $K_r^{\bar{v}}$ . Our plan is to show that we can find in  $U = V \setminus (B_1 \cup \dots \cup B_r)$  a set  $\bar{B}_0$  of  $v$  vertices so that for every  $1 \leq i \leq r$ , there is a set  $\bar{B}_i \subseteq B_i$  of size  $v$ , so that  $\bar{B}_0, \bar{B}_1, \dots, \bar{B}_r$  form a copy of  $K_r^v$ . Note that no matter how we pick  $\bar{B}_1 \subseteq B_1, \dots, \bar{B}_r \subseteq B_r$  these sets will form a complete  $r$ -partite graph, so we just need to make sure that the  $v$  vertices of  $\bar{B}_0$  will be connected to all the vertices of  $\bar{B}_1, \dots, \bar{B}_r$ .

Let  $W \subseteq U$  contain all the vertices  $w$  such that  $w$  is connected to at least  $v$  vertices in each of the sets  $B_i$ . We claim that  $|W| \geq \frac{\varepsilon n}{2}$ . To prove this we bound  $e(U, V \setminus U)$  from above and from below (recall that  $V \setminus U = (B_1 \cup \dots \cup B_r)$ ). Since  $|B_1 \cup \dots \cup B_r| = \bar{v}r$  we have as a lower bound

$$e(U, U^c) \geq \bar{v}v\delta(G) - \underset{\substack{\uparrow \\ \# \text{ edges in } U}}{(r\bar{v})^2} \geq \bar{v}vn \left(1 - \frac{1}{r} + \varepsilon\right) - (r\bar{v})^2 \geq \bar{v}vn - \bar{v}n + \varepsilon r\bar{v}n - (r\bar{v})^2.$$

Since each vertex  $u \in U \setminus W$  has some  $B_i$  where it has less than  $v$  neighbors, we have as an upper bound

$$\begin{aligned} e(U, U^c) &\leq |W|r\bar{v} + (|U| - |W|)((v-1) + (r-1)\bar{v}) \\ &\leq |W|r\bar{v} + (n - |W|)(v + (r-1)\bar{v}) \\ &= |W|(\bar{v} - v) + vn + r\bar{v}n - \bar{v}n. \end{aligned}$$

Combining the above estimates for  $e(U, U^c)$  we infer that

$$|W|\bar{v} \geq \varepsilon r\bar{v}n - (r\bar{v})^2 - vn.$$

If we now add the requirement that  $n \geq 4r\bar{v}/\varepsilon$  then we are guaranteed that  $(r\bar{v})^2 \leq \varepsilon r\bar{v}n/4$ , and if we assume that  $\bar{v} \geq 4v/\varepsilon$  then we are guaranteed that  $vn \leq \varepsilon r\bar{v}n/4$ . So if both conditions hold then we get

$$|W|\bar{v} \geq \frac{\varepsilon r\bar{v}n}{2},$$



implying that  $|W| \geq \frac{\varepsilon n}{2}$ , as needed. So to make the above work we set  $\bar{v} = 4v/\varepsilon$  and add the further assumption that  $n \geq 4r\bar{v}/\varepsilon = 16rv/\varepsilon^2$ . Note that up to now we assume that  $n \geq \max\{n_1(r-1, \varepsilon, 4v/\varepsilon), 16rv/\varepsilon^2\}$ .

Now, each of the vertices in  $W$  is connected to  $r$  vertices in each of the sets  $B_1, \dots, B_r$ . There are  $\binom{\bar{v}}{v}^r$  ways to choose  $r$  subsets  $B'_1 \subseteq B_1, \dots, B'_r \subseteq B_r$  of size  $v$  each. Therefore, there is a  $W' \subseteq W$  of size at least  $|W|/\binom{\bar{v}}{v}^r$  so that all the vertices in  $W'$  are connected to the same collection of subsets in  $B_1, \dots, B_r$ . We can finally set this collection of subsets to be  $\bar{B}_1, \dots, \bar{B}_r$  and take  $\bar{B}_0$  to be  $v$  of the vertices of  $W'$ . Of course we need to make sure that  $|W'| \geq |W|/\binom{\bar{v}}{v}^r \geq \frac{\varepsilon n}{2} \binom{\bar{v}}{v}^r \geq v$  so we just need to further require that  $n \geq \bar{v}^{rv}/\varepsilon$ . We can finally set  $n_1(r, \varepsilon, v) = \max\{n_1(r-1, \varepsilon, 4v/\varepsilon), 16rv/\varepsilon^2, (4v/\varepsilon)^{rv}/\varepsilon\}$ .  $\square$

## 2.2 The Zarankiewicz Problem for $K_{2,2}$

For  $\chi(H) \geq 3$ , the **Erdős-Stone-Simonovits** Theorem gives us a pretty good estimate for

$$ex(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) \frac{n^2}{2} = \Theta(n^2).$$

But when  $\chi(H) = 2$  we get that  $ex(n, H) = o(n^2)$ , and wish to know exactly how “small” is this  $o(n^2)$ . We start by exploring complete bipartite graphs  $K_{s,t}$  and specifically  $K_{2,2}$ .

**Theorem 2.5.**

$$ex(n, K_{2,2}) \leq \frac{n}{4} (1 + \sqrt{4n - 3}) \sim \frac{n^{3/2}}{2}$$

**Proof:** Let us count the number of triples of vertices  $(x, y, z)$  so that  $x, y, z$  form a copy of  $K_{1,2}$  with  $y$  as the middle vertex. Denote by  $\#K_{1,2}$  the number of such triples. Then

$$\#K_{1,2} = \sum_v \binom{d(v)}{2} \underset{\text{Jensen}}{\geq} n \binom{\frac{1}{n} \sum d(v)}{2} = n \binom{2m/n}{2} = \frac{2m^2}{n} - m.$$

If  $G$  is  $K_{2,2}$ -free then  $\#K_{1,2} \leq \binom{n}{2}$ , otherwise two of the  $K_{1,2}$  share the same  $x, z$  coordinates forming a  $K_{2,2}$ . Combining the above two estimates on  $\#K_{1,2}$  we infer that  $\frac{2m^2}{n} - m \leq \binom{n}{2}$  or equivalently that  $4m^2 - 2mn - n^3 - n^2 \leq 0$ , implying the desired upper bound on  $m$ .  $\square$

A lower bound (nearly) matching the one given in Theorem 2.5 is given in the next theorem.

**Theorem 2.6.**

$$ex(n, K_{2,2}) \geq \left(\frac{1}{2} - o(1)\right) n^{3/2}$$

**Proof:** We first show that for every prime  $p$ , we can construct a  $K_{2,2}$ -free graph on  $n = p^2 - 1$  vertices with  $m \geq (p^2 - 1)(p - 1)/2 = (\frac{1}{2} - o(1)) n^{3/2}$  edges. We will then explain why this implies the lower bound for all  $n$ . So given a prime  $p$  we define a graph on  $p^2 - 1$  vertices where each vertex is a pair  $(a, b) \in F_p \times F_p$ ,  $(a, b) \neq (0, 0)$ . We “connect” vertex  $(a, b)$  to vertex  $(x, y)$  iff  $ax + by = 1$  (over  $F_p$ ).

Assume  $v = (a, b) \neq (0, 0)$ . Then, it is easy to check that in all cases (i.e. if  $v$  is either  $(a, 0), (0, b)$  or  $(a, b)$  with  $a, b \neq 0$ ), that there are exactly  $p$  solutions to  $ax + by = 1$ . This means that we always have  $d(v) \geq p - 1$  (we omit the possible solution satisfying  $x = a, y = b$  since we do not allow loops), implying that  $m \geq (p^2 - 1)(p - 1)/2$  as needed.

To show that the graph is indeed  $K_{2,2}$ -free, take any  $v = (a', b'), u = (a, b)$ ,  $u \neq v$ . Then the equations

$$\begin{aligned} ax + by &= 1 \\ a'x + b'y &= 1 \end{aligned}$$

have at most one solution implying that  $v$  and  $u$  have at most one common neighbor, so the graph is indeed  $K_{2,2}$ -free.

The above construction works when  $n$  is a prime. But since it is known that for every integer  $n$  there exists a prime  $p$  satisfying  $n \leq p \leq (1 + o(1))n$  the above lower bound applies to all values of  $n$ .  $\square$

### 2.3 The Zarankiewicz Problem for General $K_{s,t}$

**Theorem 2.7.** *One can extend the argument we used in order to prove the upper bound for  $ex(n, K_{2,2})$  to obtain the following general bound*

*For every  $s \leq t$  we have*

$$ex(n, K_{s,t}) \leq \left( \frac{1}{2} + o(1) \right) (t-1)^{1/s} n^{2 - \frac{1}{s}}$$

where  $o(1) \rightarrow 0$  when  $n \rightarrow \infty$ .

*Remark.* Prove the above theorem. A slightly weaker bound (with  $\frac{1}{2} + o(1)$  replaced by  $e/2$  follows by using the inequalities  $\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k$ .

We now give a general lower bound for  $ex(n, K_{s,t})$ .

**Theorem 2.8.** *For every  $s \leq t$  we have*

$$ex(n, K_{s,t}) \geq \frac{1}{16} n^{2 - \frac{s+t-2}{st-1}}$$

**Proof:** Consider the random graph  $G(n, p)$ . The expected number of edges and copies of  $K_{s,t}$  are given by

$$\begin{aligned} \mathbb{E}[m] &= p \binom{n}{2} > pn^2/4 \\ \mathbb{E}[\#K_{s,t}] &= \binom{n}{s} \binom{n}{t} p^{st} \leq n^{s+t} p^{st} \end{aligned}$$

We now want to pick  $p$  so that we will have the relation

$$\mathbb{E}[\#K_{s,t}] \leq n^{s+t} p^{st} \leq \frac{1}{8} pn^2 \leq \frac{1}{2} \mathbb{E}[m],$$

so we set  $p = 1/2n^{\frac{s+t-2}{st-1}}$ . We then get that

$$\mathbb{E}[m - \#K_{s,t}] = \mathbb{E}[m] - \mathbb{E}[\#K_{s,t}] \geq \frac{1}{2} \mathbb{E}[m] \geq \frac{1}{8} pn^2 = \frac{1}{16} n^{2 - \frac{s+t-2}{st-1}}.$$

So there exists a graph  $G$  in which  $m - \#K_{s,t} \geq \frac{1}{16} n^{2 - \frac{s+t-2}{st-1}}$ , and by removing one edge from each  $K_{s,t}$  we get a  $K_{s,t}$ -free graph with at least  $\frac{1}{16} n^{2 - \frac{s+t-2}{st-1}}$  edges.  $\square$

### 3 Third Lecture

#### 3.1 Applications of the Zarankiewicz Problem

Recall that last lecture we posed as an exercise to show that  $ex(n, K_{s,t}) \leq 2(t-1)^{1/s} n^{2-1/s}$  for every  $s \leq t$ . We now give two applications of this bound.

Let  $S$  be a set of  $n$  points in the plane. How many pairs can be at distance 1? It is conjectured that the maximum number of such pairs is  $n^{1+o(1)}$ . Currently, the best known bound is  $O(n^{4/3})$ . Let us quickly observe a weaker upper bound of  $O(n^{3/2})$ . Given  $n$  points, we define a graph over these  $n$  points, where  $(u, v)$  is an edge iff  $d(u, v) = 1$ . It follows from elementary Geometry that this graph is  $K_{2,3}$ -free<sup>2</sup>, hence by the above bound for  $ex(n, K_{2,3})$  we get that the graph has  $O(n^{3/2})$  edges, implying the desired bound on the number of pairs at distance 1.

We now move to a more complicated application.

**Definition 3.1.** Given a set of integers  $A$ , we denote by  $A + A = \{a + a' \mid a, a' \in A\}$  its sumset.

We now ask how small can be a set  $A \subset \mathbb{N}$  if  $\{1, 2^2, 3^2, \dots, m^2\} \subseteq A + A$ ? It is obvious that  $|A| \geq \sqrt{m}$  (make sure you see this), but this bound can be improved as follows:

**Theorem 3.2** (Erdős-Newman). *If  $\{1, 2^2, 3^2, \dots, m^2\} \subseteq A + A$  then  $A \geq m^{2/3-o(1)}$ .*

*Remark.* The best known upper bound is a construction of a set of size  $|A| \leq \frac{m}{\log^{\omega(1)}(m)}$ , so there is still a huge gap between the best upper/lower bounds.

**Proof:** Suppose  $\{1, 2^2, 3^2, \dots, m^2\} \subseteq A + A$ , and set  $n = |A|$ . Define a graph  $G$  whose vertices are the  $n$  members of  $A$  and where  $(a_1, a_2)$  is an edge iff  $a_1 + a_2 = x^2$  for some  $1 \leq x \leq m$ . Note that the assumption on  $A$  implies that  $G$  has at least  $m$  edges. We will now show that  $G$  has no copy of  $K_{2,t}$  where  $t = n^{o(1)}$ . The above bound on  $ex(n, K_{s,t})$  then gives  $m \leq |E(G)| \leq 2(n^{o(1)})^{1/2} n^{2-1/2} = n^{o(1)+3/2}$  implying that  $n \geq m^{2/3-o(1)}$ . Observe that showing that  $G$  has no copy of  $K_{2,t}$  is equivalent to showing that every pair of vertices have at most  $t$  common neighbors. So our goal now is to show that every pair of vertices have  $n^{o(1)}$  common neighbors.

Fix a pair of vertices  $a_1, a_2$  and let  $b$  be one of their common neighbors. Then there are  $x, y$  satisfying  $a_1 + b = x^2$  and  $a_2 + b = y^2$ , implying that  $a_1 - a_2 = x^2 - y^2$ . For each such common neighbor  $b$  we define a label  $L_b = (x + y, x - y)$ . Note that if  $b, b'$  are two distinct common neighbors of  $a_1, a_2$  then  $L_b \neq L_{b'}$  since given a label  $(p, q)$  we can recover  $x$  and  $y$  (by solving  $x + y = p$  and  $x - y = q$ ), and therefore also  $b$  itself. Now note that for any label  $(p, q)$  both  $p$  and  $q$  divide  $a_1 - a_2$ , hence the the number of common neighbors of  $a_1$  and  $a_2$  is at most  $(D(|a_1 - a_2|))^2$  where  $D(x)$  denotes the number of integers that divide  $x$ . Hence all that is left is to prove that  $D(n) = n^{o(1)}$ .

Suppose  $n = \prod_{i=1}^r p_i^{\alpha_i}$  is the prime factorization of  $n$ . Then  $D(n) = \prod_{i=1}^r (1 + \alpha_i)$  and hence we need to prove that for every  $\varepsilon > 0$  and large enough  $n$ , we have  $\sum_{i=1}^r \log(1 + \alpha_i) \leq \varepsilon \log n$  or equivalently that

$$\frac{\sum_{i=1}^r \log(1 + \alpha_i)}{\sum_{i=1}^r \alpha_i \log p_i} \leq \varepsilon.$$

Observe that  $\frac{\log(1+\alpha_i)}{\alpha_i \log p_i} \leq \varepsilon/2$  if either  $\alpha_i \geq n_0(\varepsilon)$  or if  $p_i \geq n_1(\varepsilon)$ . So suppose there is  $i_0 = i_0(\varepsilon)$  so that

---

<sup>2</sup>There is at most one circle of radius  $r$  passing through any 3 points, and a point connecting to those three points must be the center of that circle

$\log(1 + \alpha_i) \leq \frac{\varepsilon}{2} \alpha_i \log p_i$  for all  $i > i_0$ . Examining the numerator, we have that for some  $m(\varepsilon)$  we have

$$\begin{aligned} \sum_{i=1}^r \log(1 + \alpha_i) &= \sum_{i \leq i_0} \log(1 + \alpha_i) + \sum_{i > i_0} \log(1 + \alpha_i) \\ &\leq m(\varepsilon) + \frac{\varepsilon}{2} \sum_i \alpha_i \log p_i \\ &\leq m(\varepsilon) + \frac{\varepsilon}{2} \log n. \end{aligned}$$

We now conclude that

$$\frac{\sum_{i=1}^r \log(1 + \alpha_i)}{\sum_{i=1}^r \alpha_i \log p_i} \leq \frac{m(\varepsilon) + \frac{\varepsilon}{2} \log n}{\log n} \leq \frac{m(\varepsilon)}{\log n} + \frac{\varepsilon}{2} \leq \varepsilon,$$

for all large enough  $n$ , thus completing the proof.  $\square$

### 3.2 The Turán Problem for Trees

We will now try to estimate  $ex(n, T)$  where  $T$  is a tree on  $t + 1$  vertices (we will always assume that  $t \geq 1$ ). A simple lower bound can be achieved by taking a disjoint union of cliques of size  $t$ , implying that  $ex(n, T) \geq \frac{n(t-1)}{2}$ . The following was a well known conjecture.

**Conjecture** (Erdős-Sos). *The above lower bound,  $ex(n, T) \geq \frac{n(t-1)}{2}$ , is tight.*

**Theorem 3.3** (Ajtai-Komlos-Simonovits-Szemerédi). *The Erdős-Sos Conjecture is correct.*

We will not prove the above theorem (whose proof actually has not appeared yet), but rather focus on some special cases and some weaker bounds. We start with the following simple but very useful fact.

**Claim.** *If  $G$  has  $m$  edges, then  $G$  has a subgraph  $H$  with min-degree  $\delta(H) > \frac{m}{n}$ .*

**Proof:** If  $\delta(G) > \frac{m}{n}$  we are done. Else, pick  $v \in V$  where  $d(v) \leq \frac{m}{n}$ , and remove it from the graph. The new average degree is at least

$$\frac{\sum_{i=1}^n d(i) - \frac{2m}{n}}{n-1} = \frac{2m(1 - \frac{1}{n})}{n(1 - \frac{1}{n})} = \frac{2m}{n}$$

hence the average degree never decreases. Continuing in this way we eventually end up with the required subgraph.  $\square$

We now prove a weak version of the Erdős-Sos Conjecture.

**Claim.** *For every tree  $T$  on  $t + 1$  vertices, we have  $ex(n, T) < (t - 1)n$ .*

**Proof:** If  $G$  has  $(t - 1)n$  edges, then by the above claim it has a subgraph  $H$  with min-degree at least  $t$ . So all that is left is to prove that if  $\delta(H) \geq t$  then  $H$  has a copy of every tree of size  $t + 1$ . Recall that every tree has a leaf, hence we can order the vertices  $v_1, v_2, \dots, v_{t+1}$  such that every  $v_i$  is connected to exactly one of the vertices  $v_1, \dots, v_{i-1}$ . So assuming we have already found in  $H$  a copy of the tree spanned by  $v_1, \dots, v_{i-1}$  we now wish to find a vertex to represent  $v_i$  and thus obtain a copy of the tree spanned by  $v_1, \dots, v_i$ . Assume  $v_i$  is connected to  $v_j$ ,  $1 \leq j \leq i - 1$ , thus we need to find a neighbor of  $v_j$  that does not belong to  $v_1, \dots, v_{i-1}$ . As  $d(v_j) \geq t$  and  $v_j$  has at most  $t - 1$  neighbors among the vertices  $v_1, \dots, v_{i-1}$  we can find such a vertex which we will take to be  $v_i$ .  $\square$

**Claim 3.4.** *Every graph has a cycle/path of length at least  $\delta(G) + 1$ .*

**Proof:** Take the longest path in the graph,  $x_1, \dots, x_m$ . All the neighbors of  $x_m$  must be on that path. One of them, together with  $x_m$ , closes a cycle of length at least  $d(x_m) + 1 \geq \delta(G) + 1$ .  $\square$

**Theorem 3.5** (Dirac). *If  $\delta(G) \geq \frac{n}{2}$  then  $G$  has a Hamiltonian cycle.*

**Proof:** Consider a path,  $x_1, \dots, x_m$ , of maximum length. All the neighbors of  $x_1$  and  $x_m$  are on that path as well. We claim that there exists  $1 \leq i \leq m - 1$  such that  $(x_1, x_{i+1}) \in E$  and  $(x_i, x_m) \in E$ . At least  $n/2$  of the vertices  $x_1, \dots, x_{m-1}$  (which are at most  $n - 1$  vertices) are connected to  $x_m$ , and at least  $n/2$   $x_i$  vertices such that  $x_{i+1}$  is connected to  $x_1$ . By the pigeon-hole-principle, there is a  $x_i$  which fulfills both conditions.

Now,  $x_1, \dots, x_i, x_m, \dots, x_{i+1}, x_1$  is a cycle of length  $m$ . If  $n = m$  we are done. Otherwise, it is enough to show a vertex outside the cycle which is connected to it (which forms a path of length  $m + 1$ ). This holds if  $G$  is connected, and this follows directly from  $\delta(G) \geq \frac{n}{2}$  (as every two vertices are either connected or share a common neighbor).  $\square$

It is easy to see that the lower bound on the min-degree in Dirac's Theorem is tight. For example, we can take two copies of  $K_n$  sharing a single vertex.

The next claim shows that we can improve the previous bound of  $\delta(G) + 1$  on the length of the longest path in a graph if we further assume the graph to be connected (note that without this assumption the bound  $\delta(G) + 1$  is tight).

**Claim 3.6.** *If  $G$  is connected, then  $G$  has a path of length  $\min\{n, 2\delta(G) + 1\}$ .*

**Proof:** The proof is very similar to the proof of Dirac's Theorem and is thus left as an exercise. It is also easy to see that the bound is tight.  $\square$

**Theorem 3.7** (Erdős-Gallai). *If  $G$  is  $P_{t+1}$ -free then  $m \leq \frac{(t-1)n}{2}$ . Furthermore, equality holds only when  $G$  is a disjoint union of  $K_t$ 's.*

**Proof:** By induction on  $n$ . For  $n \leq t$ , this is trivial. If  $G$  is not connected, we can apply induction on each of the connected component of  $G$ .

Assume that  $G$  is connected. Our goal is then to show that  $m < \frac{(t-1)n}{2}$ . Since  $G$  is connected and  $n > t$  we infer that  $G$  must be  $K_t$ -free. Also, if  $\delta(G) \geq \frac{t}{2}$ , then by Claim 3.6  $G$  has a path of length  $t + 1$ , a contradiction. So assume  $G$  has a vertex with degree  $\leq \frac{t-1}{2}$ . Removing this vertex, we get from induction (using the assumptions that  $G$  is  $K_t$ -free and  $P_{t+1}$ -free) that the number of edges in the new graph is strictly smaller than  $\frac{(t-1)(n-1)}{2}$ . As we removed a vertex with degree  $\leq \frac{t-1}{2}$ , we get that  $m < \frac{t-1}{2} + \frac{(t-1)(n-1)}{2} = \frac{(t-1)n}{2}$ .  $\square$

### 3.3 The Girth Problem and Moore's Bound

**Definition 3.8.** The girth of a graph  $G$ , denoted  $g(G)$ , is the length of the shortest cycle in  $G$ .

We are interested in how many edges are needed to guarantee that  $g(G)$  is small. We start with a lower bound.

**Theorem 3.9.** *There exists a graph on  $n$  vertices and  $c'n^{1+\frac{1}{2k-1}}$  satisfying  $g(G) > 2k$ .*

**Proof:** Let  $p = \frac{cn^{1/(2k-1)}}{n}$ , for a constant  $c$  that will be chosen later, and consider  $G(n, p)$ . Then the expected number of edges and copies of  $C_t$  are given by

$$\begin{aligned}\mathbb{E}[m] &\approx \frac{n^2}{2}p = \frac{c}{2}n^{1+\frac{1}{2k-1}} \\ \mathbb{E}[\#C_t] &\leq n^t p^t \leq c^t n^{t/(2k-1)}\end{aligned}$$

$$\mathbb{E}[\#C_3 + \#C_4 + \dots + \#C_{2k}] = \sum_{t=3}^{2k} c^t n^{t/(2k-1)} = c^{2k} n^{1+\frac{1}{2k-1}} + O(n)$$

Therefore

$$\mathbb{E}[m - \#C_{\leq 2k}] = n^{1+\frac{1}{2k-1}} \left( \frac{c}{2} - c^{2k} - o(1) \right) \sim c' n^{1+\frac{1}{2k-1}}$$

for  $c \leq \frac{1}{2}$ . By taking a graph for which  $m - \#C_{\leq 2k}$  is at least the expected value, and removing an edge from each short cycle, we get the desired graph.  $\square$

We now turn to prove an upper bound, which is usually referred to as the Moore Bound.

**Theorem 3.10** (Moore's Bound). *Every graph with at least  $n^{1+\frac{1}{k}}$  edges satisfies  $g(G) \leq 2k$ .*

**Proof:** We claim that every graph  $G$  on  $m \geq 2n$  edges satisfies  $g(G) \leq 2 \log(n)/\log(m/n)$ . To see this, we first use Claim 3.2 in order to find a subgraph  $G'$  of  $G$  on  $n'$  vertices with minimal degree  $d > \frac{m}{n} \geq 2$ . Set  $g = g(G)$ . Pick a vertex  $x$  in  $G'$  and grow a BFS tree of depth  $g/2$  (assume  $g$  is even) with  $x$  as its root. Since  $g(G') \geq g(G) = g$ , we see that all vertices of this tree must be distinct, so

$$n' \geq 1 + d + d(d-1) + \dots + d(d-1)^{g/2-1} \geq (d-1)^{g/2},$$

implying that

$$g \leq \frac{2 \log n'}{\log(d-1)} \leq \frac{2 \log n}{\log(m/n)},$$

as needed. Finally, if  $m \geq n^{1+\frac{1}{k}}$  then the above bound implies that  $g \leq \frac{2 \log n}{\log(n^{1/k})} = 2k$ .  $\square$

It is conjectured that the Moore bound gives the correct answer for the girth problem, namely, that there are graphs with  $\Omega(n^{1+\frac{1}{k}})$  edges and no cycle of length at most  $2k$ .

### 3.4 Application of Moore's Bound to Graph Spanners

Given a graph  $G$ , we wish to find a sparse subgraph  $H$  which will satisfy the condition

$$\delta_G(u, v) \leq \delta_H(u, v) \leq (2k-1)\delta_G(u, v)$$

for every pair of vertices  $u, v$ . That is, we want to remove edges but to keep the distance between any pair of vertices close to the original distance. A graph satisfying the above condition is called a  $(2k-1)$ -spanner.

**Theorem.** *Every graph has a subgraph which is a  $(2k-1)$ -spanner with at most  $n^{1+\frac{1}{k}}$  edges.*

**Proof:** Add the edges of  $G$  one by one where an edge is put into  $H$  only if the edge does not close a cycle of length  $\leq 2k$ . Then  $H$  has no cycle of length at most  $2k$ , hence by Moore's bound it has at most  $n^{1+1/k}$  edges. It is easy to see that  $H$  is indeed a  $(2k-1)$ -spanner.  $\square$

**Exercise.** Show that if Moore's Bound gives the correct bound for the girth problem then there exist graphs, whose every  $(2k-1)$ -spanner has  $\Omega(n^{1+\frac{1}{k}})$  edges.

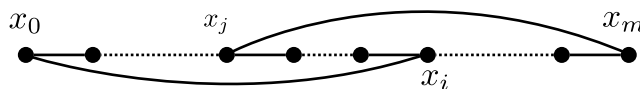


Figure 4.1: Cycle  $C = x_0 \dots x_j x_m \dots x_i x_0$

## 4 Fourth Lecture

### 4.1 The Turán Problem for Long Cycles

**Definition 4.1.** The circumference of a graph  $G$ , denoted  $c(G)$ , is the length of the longest cycle in  $G$ .

We now ask how many edges are needed to guarantee that  $c(G) > t$ . The following theorem gives an exact answer.

**Theorem 4.2** (Erdős-Gallai). *If  $G$  has more than  $\frac{t(n-1)}{2}$  edges then  $c(G) \geq t + 1$ . Furthermore, this bound is tight when  $n = k(t - 1) + 1$  (consider  $k$  copies of  $K_t$  all sharing a single vertex).*

To prove the above theorem, we will need the following lemma.

**Lemma 4.3.** *If  $G$  is 2-connected, then  $G$  has a cycle of length at least  $\min\{n, 2\delta(G)\}$  (tight when  $G$  consists of several  $K_t$  sharing two vertices).*

**Proof** (Theorem 4.2): By induction on  $n$ . Since there is nothing to prove when  $n \leq t$  assume  $n > t$ . If the graph has a vertex  $x$  where  $\deg(x) \leq t/2$ , then

$$e(G - \{x\}) \geq e(G) - \deg(x) \geq \frac{t(n-1)}{2} - \frac{t}{2} \geq \frac{t(n-2)}{2}$$

and by the induction hypothesis, the statement holds.

Otherwise,  $\delta(G) \geq \frac{t+1}{2}$ . If  $G$  is 2-connected, then by the lemma  $G$  has a cycle of length  $\geq t + 1$  (here we are also using the assumption that  $n > t$ ). If  $G$  is not 2-connected, then it has a cut-vertex  $x$  which separates it into two otherwise disjoint components  $G_1$  and  $G_2$  which have  $n_1$  and  $n_2$  vertices. As  $G$  is assumed to contain at least  $\frac{t(n-1)}{2} = \frac{t(n_1-1)}{2} + \frac{t(n_2-1)}{2}$  edges, we can apply the induction to at least one of the components, completing the proof.  $\square$

We now turn to prove the lemma we used.

**Proof** (Lemma 4.3): Let  $P = x_0, \dots, x_m$  be the longest path in  $G$ . There are several possible cases:

1. If there is a vertex  $x_i$ , where  $0 < i < m$  such as  $x_0$  is connected to  $x_i$  and  $x_m$  to  $x_{i-1}$ , then we can turn this path into a cycle (as in the proof of Dirac's Theorem). Furthermore, since  $G$  is assumed to be connected, this cycle must be a Hamilton cycle, as otherwise some vertex of the cycle has to be connected to a vertex not belonging to it, which gives a longer path.
2. If there is no vertex  $x_i$  as above, then assume there is  $j < i - 1$  such that  $x_0$  is connected to  $x_i$  and  $x_m$  to  $x_j$  and take  $i, j$  such as  $|i - j|$  is minimal (see Figure 4.1). Consider the cycle  $C$  that starts with  $x_0$  goes to  $x_j$  then jumps to  $x_m$ , then goes to  $x_i$  and then jumps back to  $x_0$ . We claim that this cycle is of length at least  $2\delta(G)$ . By definition,  $C$  has  $x_m$  and all its neighbors (as  $P$  has all of  $x_m$ 's neighbors, and none of them is between  $i$  and  $j$ ), which are at least  $\delta(G)$  vertices. Hence  $C$  has at least  $\delta(G) + 1$  vertices. Furthermore, if  $x_l$  ( $l \neq i$ ) is a neighbor of  $x_0$ , then  $C$  contains  $x_{l-1}$  (as either  $l > i$  or  $l < j$ ). As we assume that the first case does not hold<sup>3</sup>, we get that  $C$  has at least  $\delta(G) + 1 + \delta(G) - 1 = 2\delta(G)$  vertices.

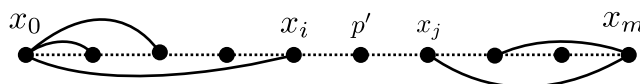


Figure 4.2: Third case in the proof of the lemma. The illustration assumes that there is no common neighbor, but this is not necessary for the proof.

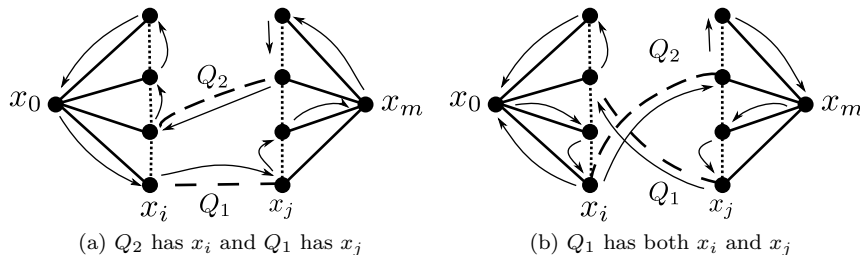


Figure 4.3: Third case in the proof of the lemma. Dotted and dashed lines denote disjoint paths.

3. All the neighbor of  $x_0$  are before the neighbors of  $x_m$  (see Figure 4.2). Let  $A$  be the set of  $x_0$  and all its neighbors and  $B$  the set of  $x_m$  and its neighbors. As  $G$  is 2-connected there are two vertex disjoint paths  $Q_1, Q_2$  which connect  $A$  and  $B$  and w.l.o.g. we may assume  $Q_2$  starts in  $x_i$ . (as there is a path which connects  $x_i$  to  $x_j$ ). If  $Q_1$  and  $Q_2$  do not start in  $x_i$  we begin traversing from  $x_i$  through  $p'$  until we encounter  $Q_2$  (or  $Q_1$ ) that results in path that starts in  $x_i$ . Similarly, we may assume a path that ends in  $x_j$  (possibly the same path). Therefore, it is possible to construct a cycle on all the vertices of  $A$  and  $B$  (see Figure 4.3) which has

$$\deg(x_0) + 1 + \deg(x_m) + \underset{\substack{\uparrow \\ \text{if } x_i=x_j}}{1-1} \geq 2\delta(G)$$

These are all the possible cases, thus proving the lemma.  $\square$

## 4.2 Pancyclic Graphs and Bondy's Theorem

**Definition 4.4.** A graph is pancyclic if it contains a cycle of every length  $3 \leq l \leq n$ .

**Theorem 4.5** (Bondy). *If  $G$  is a Hamiltonian graph and  $m \geq \frac{n^2}{4}$  then it is pancyclic unless  $m = \frac{n^2}{4}$  and  $G = K_{\frac{n}{2}, \frac{n}{2}}$ . [Bondy, 1971]*

**Corollary.** *If  $m > \frac{n^2}{4}$  and  $G$  is Hamiltonian then  $G$  is pancyclic.*

**Corollary.** *If  $\delta(G) \geq \frac{n}{2}$  then  $G$  is pancyclic (stronger than Dirac's theorem).*

**Proof (Bondy):** Assume  $G$  has a cycle of length  $n - 1$ , and let  $x \notin C_{n-1}$ . If  $d(x) \leq \frac{n-1}{2}$  then in the induced subgraph on the cycle there are at least  $\frac{n^2}{4} - \frac{n-1}{2} > \frac{(n-1)^2}{4}$  edges. By induction, we get that the graph induced by this  $C_{n-1}$  has a cycles of every length  $\leq n - 1$  and we are done. Assume then that  $d(x) > \frac{n}{2}$ . By the pigeonhole principle we get that for every  $3 \leq l \leq n - 2$  there are two neighbors of  $x$  that together with it form a cycle of length  $l$  (make sure you see this point) and we are done again.

<sup>3</sup>This means that  $x_{l-1}$  is never a neighbor of  $x_m$ .



So assume that  $G$  has no cycle of length  $n - 1$ . We need to prove that  $G = K_{\frac{n}{2}, \frac{n}{2}}$ . Note that in the Hamilton cycle, it is impossible that two adjacent vertices will be connected to two vertices of distance 2 (as this would result in a cycle of length  $n - 1$ ). Hence,  $G$  has no vertices  $i, j$  such that  $i$  is connected to  $j + 2$  and  $i - 1$  to  $j$ . This means that  $d(x_i) + d(x_{i-1}) \leq n$  (as for every  $j$ ,  $(x_{i-1}, x_j)$  and  $(x_i, x_{j+2})$  cannot be both edges). Therefore

$$4m = \sum_{i=1}^n d(x_i) + d(x_{i-1}) \leq n^2,$$

implying that  $m \leq \frac{n^2}{4}$ . But we assume that  $m \geq \frac{n^2}{4}$  so  $m = \frac{n^2}{4}$ , hence  $m$  must be even and for every  $i, j$  the graph  $G$  contains exactly one of the edges  $(x_{i-1}, x_j)$  or  $(x_i, x_{j+2})$ . We are left with showing that  $G$  is bipartite (and then from  $m = \frac{n^2}{4}$  it follows that  $G = K_{\frac{n}{2}, \frac{n}{2}}$ ). We color the vertices along the cycle. From the constrain that only one edge from  $(x_{i-1}, x_j)$  and  $(x_i, x_{j+2})$  is in the graph, we get that  $G$  is bipartite (edge with different colors cannot exist as otherwise we would have a cycle of length  $n - 1$ ).  $\square$

### 4.3 The Moon-Moser Inequalities

**Theorem 4.6.** *For every graph  $G$  we have*

$$N_3 \geq \frac{4N_2}{3} \left( \frac{N_2}{N_1} - \frac{N_1}{4} \right)$$

where  $N_i$  is the number of copies of  $K_i$  in  $G$  (so  $N_1 = n$  and  $N_2 = m$ ).

**Proof:** For an edge  $e$  denote by  $d(e)$  the number of triangles that contain  $e$ . Then we have

$$\begin{aligned} 3N_3 &= \sum_e d(e) \geq \sum_{(x,y)=e} (d(x) + d(y) - n) \\ &= \sum_x d^2(x) - mn \geq n \left( \frac{\sum d(x)}{n} \right)^2 - mn \\ &= \frac{4m^2}{n} - mn, \end{aligned}$$

and rearranging terms gives the required lower bound on  $N_3$ .  $\square$

Note that the above theorem implies that if a graph has  $\frac{n^2}{4} + 1$  edges then it contains at least  $\frac{n}{3}$  triangles. So in particular this implies Mantel's Theorem. Note that the  $\frac{n}{3}$  bound on the number of triangles in graphs with  $\frac{n^2}{4} + 1$  edges is weaker than the (tight)  $\lfloor n/2 \rfloor$  bound you were asked to prove in the home assignments. However, the theorem here is more general since it gives a lower bound for  $N_3$  for any  $m > n^2$ . One can generalize the above theorem to the case of  $K_s$ .

**Theorem 4.7** (Moon-Moser). *(Generalization of the previous theorem) If  $N_{s-1} \neq 0$ :*

$$\frac{N_{s+1}}{N_s} \geq \frac{s^2}{s^2 - 1} \left( \frac{N_s}{N_{s-1}} - \frac{n}{s^2} \right)$$

Instead of proving the above theorem, we will prove the following more general theorem that deals with copies of  $K_s$  in  $r$ -uniform hypergraphs.

**Theorem 4.8** (de-Caen). *Moon-Moser Theorem for  $r$ -uniform hypergraphs. If  $N_{s-1} \neq 0$ :*

$$\frac{N_{s+1}}{N_s} \geq \frac{s^2}{(s-r+1)(s+1)} \left( \frac{N_s}{N_{s-1}} - \frac{(r-1)(n-s)+s}{s^2} \right)$$

**Proof:** We try to imitate the proof of Theorem 4.6. In what follows  $e$  denotes some copy of  $K_s$ . We start with the equation

$$(s+1)N_{s+1} = \sum_e d(e),$$

where  $d(e)$  is the number of copies of  $K_{s+1}$  contain the vertex set of  $e$ . For each copy  $e$  of  $K_s$ , denote  $e_1, \dots, e_s$  the  $s$  copies of  $K_{s-1}$  contained in it, where  $e_i$  the result of removing vertex  $i$ . We will shortly prove that

$$d(e) \geq \frac{\sum_i d(e_i) - (n-s)(r-1) + s}{s-r+1},$$

but let us first show how to finish the proof using the above inequality. Using  $e'$  to denote copies of  $K_s$  we get

$$\begin{aligned} (s-r+1)(s+1)N_{s+1} &= (s-r+1) \sum_e d(e) \\ &\geq \sum_e \left( \sum_i d(e_i) - (n-s)(r-1) + s \right) \\ &= \sum_{e'} d^2(e') - N_s ((n-s)(r-1) + s) \\ &\geq N_{s-1} \left( \frac{\sum_{e'} d(e')}{N_{s-1}} \right)^2 - N_s ((n-s)(r-1) + s) \\ &= \frac{s^2 N_s^2}{N_{s-1}} - N_s ((n-s)(r-1) + s) \end{aligned}$$

and rearranging the terms gives the theorem.

And now to the proof of the claim. We count how many copies of  $K_s$  there are with  $s-1$  vertices in  $e$  (recall that  $e$  contains  $s$  vertices) and 1 in  $V \setminus e$ . Denote that number by  $p$ . We first observe that

$$p = \sum_{i=1}^s (d(e_i) - 1) = \sum_{i=1}^s d(e_i) - s.$$

Next, we note that in an  $r$ -uniform hypergraph, if a vertex  $u$  does not form a copy of  $K_{s+1}$  with a set of  $s$  vertices in  $e$ , then  $u$  forms a copy of  $K_s$  with at most  $r-1$  of the subsets of  $e$  of size  $s-1$ . Now, there are  $d(e)$  vertices that form a copy of  $K_{s+1}$  with  $e$ , so each such vertex forms a copy of  $K_s$  with the  $s$  subsets of  $e$  of size  $s-1$ . By the above observation, the  $n-s-d(e)$  vertices that do not form a copy of  $K_{s+1}$ , form a copy of  $K_s$  with at most  $r-1$  of the subsets of  $e$  of size  $s-1$ . So we have

$$p \leq s \cdot d(e) + (n-s-d(e))(r-1),$$

and combining the above two estimates for  $p$  proves the claim.  $\square$

We now want to prove that if a graph has  $m > \left(1 - \frac{1}{t}\right) \frac{n^2}{2}$  edges, then it does not contain only one copy of  $K_t$ , but actually many of them. To this end

**Claim 4.9.** Suppose  $G$  has  $m$  edges and set  $x$  so that  $m = (1 - \frac{1}{x}) \frac{n^2}{2}$ . Then

$$\frac{N_{s+1}}{N_s} \geq \frac{n(x-s)}{x(s+1)}$$

**Proof:** Induction on  $s$ . For  $s = 1$  by the definition of  $m$  we have

$$\frac{N_2}{N_1} \geq \frac{n(x-1)}{2x}.$$

Assuming the claim holds for  $s$  we will prove it for  $s + 1$ .

$$\begin{aligned} \frac{N_{s+2}}{N_{s+1}} &\stackrel{\text{Theorem 4.7}}{\geq} \frac{(s+1)^2}{(s+1)^2 - 1} \left( \frac{N_{s+1}}{N_s} - \frac{n}{(s+1)^2} \right) \\ &\stackrel{\text{Induction}}{\geq} \frac{(s+1)^2}{(s+1)^2 - 1} \left( \frac{n(x-s)}{x(s+1)} - \frac{n}{(s+1)^2} \right) \\ &= \frac{n(x-s-1)}{x(s+2)} \end{aligned}$$

□

**Corollary 4.10.** If  $m > (1 - \frac{1}{s}) \frac{n^2}{2}$  then  $x > s$  and hence from the claim we get

$$N_{s+1} > 0, \dots, N_4 > 0, N_3 > 0$$

which is the weak version of Turán's Theorem.

**Corollary.** If  $m \geq (1 - \frac{1}{s} + c) \frac{n^2}{2}$  then  $G$  has at least  $\frac{c}{s} n^{s+1}$  copies of  $K_{s+1}$ <sup>4</sup>.

**Proof:**

$$\begin{aligned} \frac{N_{s+1}}{N_1} &= \frac{N_{s+1}}{N_s} \cdot \frac{N_s}{N_{s-1}} \cdot \dots \cdot \frac{N_2}{N_1} \geq \prod_{i=1}^s \frac{n(x-i)}{x(i+1)} \\ &= \left(\frac{n}{x}\right)^s \frac{(x-1)(x-2) \cdots (x-s)}{1 \cdot 2 \cdots s} \end{aligned}$$

Hence

$$N_{s+1} \geq \left(\frac{n}{x}\right)^{s+1} \binom{x}{s}$$

Now if  $m = (1 - \frac{1}{s} + c) \frac{n^2}{2}$  then  $x \sim s + c$  and picking small enough  $c$  we get

$$N_{s+1} \geq c \cdot \left(\frac{n}{s+1}\right)^{s+1}$$

□

---

<sup>4</sup>Follows from problem 3 in home assignment 1.

## 5 Fifth Lecture

### 5.1 The Hypergraph Turán Problem

It is natural to consider the Turán Problem in the setting of  $r$ -uniform hypergraphs ( $r$ -graphs for short). Surprisingly, while we have rather exact results for the graph Turán problem, already for 3-graphs the problem becomes much harder. In fact, it becomes much harder even in the “first” non-trivial case, when we forbid  $K_4^3$ , which is the complete 3-graph on 4 vertices. Let  $ex(n, K_4^3)$  denote the maximum number of edges a 3-graph can contain if it does not contain a copy of  $K_4^3$ . It is not hard to show that  $ex(n, K_4^3) \geq (\frac{5}{9} - o(1))\binom{n}{3}$ ; to this end we do something very similar to what we did in the graph case. We take  $n$  vertices and partition them into 3 sets  $V_1, V_2, V_3$  of almost equal size. We then take as edges all triples of vertices  $(x, y, z)$  if they are of the form  $x \in V_1, y \in V_2, z \in V_3$ , or of the form  $x, y \in V_i, z \in V_{i+1}$ , where  $i \in [3]$  (and addition is modulo 3). Let’s call this graph  $T^{3,4}$ . A well known conjecture is that  $ex(n, K_4^3) \leq (\frac{5}{9} + o(1))\binom{n}{3}$ . To date, the best known upper bound is  $ex(n, K_4^3) \leq 0.561\binom{n}{3}$ .

A possible explanation for the hardness of proving tight bounds for hypergraph Turán Problem is the following. Recall that most proofs we gave in the graph case actually proved not only a tight upper bound but that there is a unique graph (the Turán graph) which attains the maximum. As it turns out, the 3-graph  $T^{3,4}$  we described above is not the unique  $K_4^3$ -free 3-graph with this many edges. In fact, for every  $n$  there are exponentially many non-isomorphic 3-graphs that are  $K_4^3$ -free and have the same number of edges as  $T^{3,4}$ . So if indeed  $ex(n, K_4^3) = |E(T^{3,4})|$  then any proof would have to “avoid” proving that  $T^{3,4}$  is the unique maximum.

Looking back at our previous proofs of the graph Turán Problem, we can see that the one using the Moon-Moser inequality (see Corollary 4.10) did not prove uniqueness of the graph maximizing the number of edges. As it turns out one can also use the Moon-Moser inequality for  $r$ -graphs (see Theorem 4.8), to prove the following general upper bound.

**Theorem 5.1.** *Let  $K_s^r$  denote the complete  $r$ -graph on  $s$  vertices. Then*

$$ex(n, K_s^r) \leq \left(1 - \frac{1}{\binom{s-1}{r-1}} + o(1)\right) \binom{n}{r}$$

Observe that given the discussion above, this upper bound is not tight already for  $ex(n, K_4^3)$ . The proof of this theorem proceeds along the lines of the proof of Claim 4.9, but is much more tedious so we will skip it. Instead we will prove the following general lower bound

**Theorem 5.2.** *Let  $K_s^r$  denote the complete  $r$ -graph on  $s$  vertices. Then*

$$ex(n, K_s^r) \geq \left(1 - \left(\frac{r-1}{s-1}\right)^{r-1} - o(1)\right) \binom{n}{r}$$

**Proof:** It will be easier to consider the complementary problem of constructing an  $r$ -graph with at most  $\left(\left(\frac{r-1}{s-1}\right)^{r-1} + o(1)\right) \binom{n}{r}$  edges so that every set of  $s$  vertices contains an edge (taking the complement of such an  $r$ -graph then proves the theorem). So consider the following  $r$ -graph; partition the  $n$  vertices to  $s-1$  almost equal sized sets  $V_1, \dots, V_{s-1}$ . Let  $e$  be a set of  $r$  vertices. Then  $e$  is an edge iff there exists an index  $j$  so that

$$\sum_{i=0}^{k-1} |e \cap V_{j+i}| \geq k+1, \quad \forall 1 \leq k \leq r-1, \quad (5.1)$$

where the subscripts of  $V_{j+i}$  are taken modulo  $s-1$ . Note that showing that every set of  $s$  vertices contains an edge follows from the fact that if we place  $s$  balls into  $s-1$  buckets  $V_1, V_2, \dots, V_{s-1}$  that are arranged

on a cycle, then there is an index  $j$  so that for every  $1 \leq k \leq s-1$  the buckets  $V_j, V_{j+1}, \dots, V_{j+k-1}$  have together at least  $k+1$  balls. We leave the solution of this “riddle” as an exercise.

So we are left with proving an upper bound on the number of edges. It will turn out to be easier to count the number of ordered  $r$ -tuples  $(x_1, \dots, x_r)$  such that  $\{x_1, \dots, x_r\}$  is an edge. Let  $K$  denote the number of such  $r$ -tuples. Since  $K = r!|E|$  (we are counting each edge exactly  $r!$  times), what we need to show is that

$$K \leq r! \left( \left( \frac{r-1}{s-1} \right)^{r-1} + o(1) \right) \binom{n}{r} \leq (1 + o(1)) n^r \left( \frac{r-1}{s-1} \right)^{r-1}.$$

Recall that an  $r$ -tuple of vertices  $(x_1, \dots, x_r)$  is an edge iff there is an index  $j$  for which condition (5.1) holds. Let us assign to each ordered  $r$ -tuple of vertices  $(x_1, \dots, x_r)$ , for which  $\{x_1, \dots, x_r\}$  is an edge, a signature  $(j, c)$ , where  $j$  is an index  $j$  for which (5.1) holds (if there are several such  $j$  pick an arbitrary one), and  $c = (c_1, \dots, c_r)$  is an  $r$ -tuple of non-negative integers where  $c_i = t$  if  $x_i \in V_{j+t}$ . Now, observe that for every pair  $(j, c)$  there are  $(1 + o(1)) \left( \frac{n}{s-1} \right)^r$  ways to pick an  $r$ -tuple  $(x_1, \dots, x_r)$  whose signature is  $(j, c)$ . Furthermore, note that the signature of an  $r$ -tuple  $(x_1, \dots, x_r)$  determines whether it forms an edge. Let us say that  $(c_1, \dots, c_r)$  is legal if any  $r$ -tuple  $(x_1, \dots, x_r)$  whose signature is  $(j, c)$  forms an edge<sup>5</sup>. Therefore, we are down to proving that there are at most  $(r-1)^{r-1}$  ways to pick a legal  $(c_1, \dots, c_r)$ . Indeed, this will give that there are at most  $(s-1)(r-1)^{r-1}$  signatures that define an edge, and since each signature defines  $(1 + o(1)) \left( \frac{n}{s-1} \right)^r$  ordered  $r$ -tuples  $(x_1, \dots, x_r)$ , this will give the required upper bound on  $K$ .

To prove the upper bound on the number of legal  $(c_1, \dots, c_r)$ , first note that condition (5.1) forces all  $c_i \in \{0, 1, \dots, r-2\}$ , as otherwise (5.1) will fail for  $k = r-1$ . This means that all legal  $c$  belong to  $\{0, \dots, r-2\}^r$ . So we need to prove that only  $1/(r-1)$  of the strings in  $c \in \{0, \dots, r-2\}^r$  define a legal  $c$ . To this end we will group these strings into groups of size  $r-1$  and show that each group contains at most one legal  $c$ . Define  $shift(c, \ell)$  to be the string  $c'$  satisfying  $c'_i = c_i + \ell$  (where addition is again modulo  $r-1$ ). Then each group contains  $r-1$  strings of the form  $shift(c, 0), shift(c, 1), \dots, shift(c, r-2)$ . We claim that for any  $c$  at most one of the strings  $shift(c, 0), \dots, shift(c, r-2)$  is legal. Indeed, note that if  $(j, c)$  is the signature of some  $r$ -tuple of vertices  $(x_1, \dots, x_r)$ , then  $(j, shift(c, \ell))$  is the signature of the  $r$ -tuple that results by cyclically “shifting” the vertices along the “arc” of clusters  $V_j, \dots, V_{j+r-2}$ . This means that if  $shift(c, 0)$  is legal then we must have at least  $\ell+1$  vertices in the clusters  $V_j, \dots, V_{j+\ell-1}$  and if  $shift(c, \ell)$  is legal then we must have at least  $r-\ell$  vertices in the clusters  $V_{j+\ell}, \dots, V_{j+r-2}$ . The total is at least  $r+1$  vertices, which is a contradiction.  $\square$

We note that the upper bound in the above argument is tight in the sense that the hypergraph indeed contains the stated number of edges. To see this we consider the two points where we obtained an upper bound. The first is when we claimed that each  $r$ -tuple  $(x_1, \dots, x_r)$  that defines an edge has at least one index  $j$  for which condition (5.1) holds. It is not hard to see that there can never be two such indices, meaning that given the number of legal strings  $c$ , we were not “over counting” the number  $r$ -tuples  $(x_1, \dots, x_r)$  that form an edge. To see this, suppose there are two such indices  $1 \leq j < j' \leq s-1$ ; if  $j' - j \geq r-1$  then there must be  $r$  vertices in the clusters  $B_j, \dots, B_{j+r-2}$  and at least 2 more vertices in  $B_{j'}$  for a total of at least  $r+2$ . A similar conclusion holds if the distance from  $j'$  to  $j$  is at least  $r-1$ , i.e. if  $s-1 - (j' - j) \geq r-1$ . If none of the above conditions holds, then by a similar reasoning we must have at least  $s+1 \geq r+2$  vertices. In any case we get a contradiction. A second point where the analysis is tight is where we claim that for any collection of  $r-1$  strings  $shift(c, 0), \dots, shift(c, r-2)$ , at most 1 of them is legal. Again, it is not hard to see that one of them is always legal. The proof is very similar to the solution of the “riddle” we posed at the beginning of the proof.

<sup>5</sup>Here we are using the simple observation that for every  $c$  either all signatures  $(j, c)$  define an edge or none of them do.

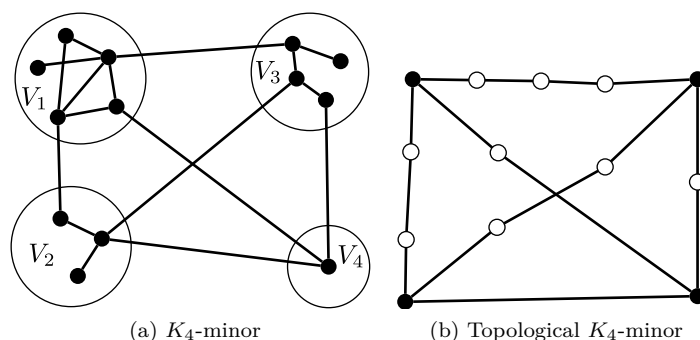


Figure 5.1

## 5.2 Extremal Problems Related to Graph Minors

We now turn to study a problem similar to the graph Turán problem, but instead of asking how many edges are needed to force the appearance of some graph  $H$  as a subgraph, we consider two relaxed notions of subgraphs, known as  $H$ -minor and topological  $H$ -minor. We start with the latter notion.

**Definition 5.3.** A graph  $G$  has a topological  $H$ -minor, where  $H = (V_H, E_H)$  with  $V_H = \{1, \dots, h\}$ , if  $G$  has  $h$  distinct vertices  $u_1, \dots, u_h$ , and a collection of internally <sup>6</sup> vertex disjoint paths  $\{P_{i,j} : (i, j) \in E_H\}$ , so that  $P_{i,j}$  connects  $u_i$  to  $u_j$ .

**Definition 5.4.** A graph  $G$  has an  $H$ -minor, where  $H = (V_H, E_H)$  with  $V_H = \{1, \dots, h\}$ , if  $G$  has  $h$  vertex disjoint connected subgraphs  $U_1, \dots, U_h$  and a collection of internally vertex disjoint paths  $\{P_{i,j} : (i, j) \in E_H\}$ , so that  $P_{i,j}$  connects a vertex from  $U_i$  to a vertex of  $U_j$ .

Note that we could have defined the notion of a graph minor using single edges (i.e. paths of length 1) instead of the arbitrarily long paths we allowed in the definition. We just wanted to draw the similarity to the notion of topological minors. A more interesting equivalent notion of graph minor is the following.

**Definition 5.5.** A graph  $G$  has an  $H$ -minor if  $H$  can be obtained from  $G$  by a sequence of:

1. Vertex deletions
2. Edge deletions
3. Edge contractions (take an edge  $(x, y)$ , and replace  $x, y$  with a single vertex connected to every neighbor of the vertices it replaced (removing loops and parallel edges)).

One of the most interesting aspects of the notion of graph minor is its relation to the conjecture of Hadwiger which states the following.

**Conjecture 5.6 (Hadwiger).** *If  $\chi(G) \geq t$  then  $G$  has a  $K_t$ -minor.*

For some time the following conjecture was open.

**Conjecture 5.7 (Hajos).** *If  $\chi(G) \geq t$  then  $G$  has a topological  $K_t$ -minor.*

The following exercise guides you into disproving this conjecture Hajos's conjecture.

**Exercise.** Show that with high probability  $G(n, 1/2)$  satisfies the following two properties:

<sup>6</sup>By this we mean that each path can meet other paths only at its first/last vertices. All the other vertices on the path are disjoint from  $v_1, \dots, v_h$  and from the other paths.

1. It does not contain a topological  $K_t$ -minor with  $t = 10\sqrt{n}$ .
2. It is not  $\frac{n}{2^{\log n}}$ -colorable.

Conclude that Hajos's conjecture is false. While you are at it, prove that with high probability,  $G(n, 1/2)$  does contain a topological  $K_t$ -minor with  $t = \sqrt{n}/10$ .

Hadwiger conjecture, on the other hand, is still open. For  $t = 2, 3$  it is trivial, for  $t = 4$ , it is an exercise. For  $t = 5$  it turns out to be equivalent to the four color theorem. It is known to hold for  $t = 6$  via a reduction to the four color theorem. For  $t > 6$ , the conjecture is open.

Let us get back to the extremal problems we are interested in. We first ask, how many edges are needed to guarantee that a graph has a  $K_t$ -minor. We will prove that surprisingly, a linear number of edges is sufficient.

**Theorem 5.8.** *Every graph with at least  $2^{t-3}n$  edges has a  $K_t$ -minor.*

We also ask, how many edges are needed to guarantee that a graph has a topological  $K_t$ -minor. Again, we show that a linear number of edges is sufficient.

**Theorem 5.9.** *Every graph with at least  $2^{\binom{t}{2}}n$  edges has a topological  $K_t$ -minor.*

The key fact that will enable us to prove these two theorems is Claim 5.11 stated below. For its proof we will need the following definition.

**Definition 5.10.** Given a graph  $G$  and a connected subgraph spanned by a vertex set  $S$ , denote by  $G/S$  the graph obtained from  $G$  by contracting  $S$  into a single vertex.

**Claim 5.11.** *Every graph  $G$  has a connected subgraph  $H$ , so that  $\delta(G[N(H)]) > \frac{m}{n} - 1$ , where  $N(H)$  are the vertices that are adjacent to some vertex of  $H$ , and  $G[N(H)]$  is the graph induced by these vertices.*

**Proof:** We describe a process for constructing  $H$ . Assume  $G$  is connected<sup>7</sup> and set  $H_1 = \{v_0\}$  where  $v_0$  is an arbitrary vertex of  $G$ . We now do the following; If  $\delta(G[N(H_1)]) > \frac{m}{n} - 1$  we done. Otherwise, take a vertex  $v_1 \in N(H_1)$ , whose degree in  $G[N(H_1)]$  is at most  $\frac{m}{n} - 1$  and define  $H_2 = H_1 + v_1$ . We continue in this manner, where at iteration  $i$  we add to  $H_i$  a vertex  $v_i \in N(H_i)$ , whose degree in  $G[N(H_i)]$  is at most  $\frac{m}{n} - 1$  (if one exists, otherwise we are done) and define  $H_{i+1} = H_i + v_i$ . For what follows, observe that we always maintain that  $H_i$  spans a connected subgraph in  $G$ . Clearly when this process ends we get a set  $H$  satisfying the condition of the claim; the only thing that can go wrong is if  $N(H) = \emptyset$ . Since  $G$  is connected, this can only happen if  $H_i = G$ . To show that this cannot happen, we will show that if  $d(G/H_i) \geq m/n$  then  $d(G/H_{i+1}) \geq m/n$ , where for a graph  $G$  we set  $d(G) = m/n$ . Since  $G/H_1 = G$  we have  $d(G/H_1) = d(G) = m/n$  implying that  $d(G/H_i) \geq m/n$  so  $H_i \neq G$ .

Set  $n_i/m_i$  to be the number of vertices/edges in  $G/H_i$ , and  $n_{i+1}/m_{i+1}$  to be the number of vertices/edges in  $G/H_{i+1}$ . Then  $n_{i+1} = n_i - 1$ . As to  $m_{i+1}$  note that when adding  $v_i$  to  $H_i$  we lose the edge that connected  $v_i$  to the vertex resulting from contracting  $H_i$  as well as the at most  $\frac{m}{n} - 1$  edges that connected  $v_i$  to its neighbors in  $G[N(H_i)]$  (all these edges will be lost since  $H_i$  is already connected to them). Altogether we get  $m_{i+1} \geq m_i - m/n$ . Thus

$$\frac{m_{i+1}}{n_{i+1}} \geq \frac{m_i - m/n}{n_i - 1} \geq \frac{m}{n},$$

where in the second inequality we use the assumption that  $m_i/n_i \geq m/n$ . □

<sup>7</sup>Why is it fine to assume so? If  $G$  was not connected, at least one of its connected components must be at least as dense as  $G$  and we could operate on it instead of  $G$ .

The proof of Theorems 5.8 and 5.9 are now easy corollaries of the claim above.

**Proof** (Theorem 5.8): Induction on  $t \geq 3$ . For  $t = 3$  it is trivial<sup>8</sup>. Assume now that  $G$  has  $m = 2^{t-3}n$  edges where  $t \geq 4$ . Then from Claim 5.11 we get a subset  $H$  such that  $\delta(G[N(H)]) > 2^{t-3} - 1$  that is  $\delta(G[N(H)]) \geq 2^{t-3}$ . Since  $\delta(G[N(H)]) \geq 2^{t-3}$  we get that if  $G[N(H)]$  has  $n'$  vertices then it has at least  $2^{t-4}n'$  edges. By induction,  $G[N(H)]$  has  $K_{t-1}$ -minor and together with  $H$  we get a  $K_t$ -minor.  $\square$

**Proof** (Theorem 5.9): We prove by induction on  $k$  that  $2^{t+k}n$  edges give a topological  $K$ -minor where  $K$  is a cycle of length  $t$  with  $k$  chords. If  $k = 0$ , then we have a graph with  $2^t n$  edges. Claim 3.2 then gives a subgraph with minimal degree at least  $2^t + 1$  and Claim 3.4 then a cycle of length  $2^t + 2$  which is clearly a  $C_t$ -minor (any cycle of length at least  $t$  is a  $C_t$  minor). Given a graph with  $2^{t+k}n$  edges we get (like in the previous proof) a connected subgraph  $H$ , so that  $N(H)$  has average degree  $2^{t+k}$ . By induction we can find in  $N(H)$  a topological  $K'$ -minor where  $K'$  is a  $C_t$  plus  $k - 1$  chords. We can now pick two vertices in the topological  $K'$ -minor that are not yet connected by a path, and connect them by a path that goes through  $H$  and avoids all other vertices/paths that belong to  $N(H)$ .  $\square$

It is of course natural to ask whether the constant  $2^{t-3}$  in Theorem 5.8 can be improved? As it turns out  $O(t\sqrt{\log tn})$  are sufficient and necessary for forcing a  $K_t$ -minor. Proving that this many edges suffice is somewhat harder than the proof of Theorem 5.8, but the proof of the necessity of this many edges follows from the following exercise

**Exercise 5.12.** Show that with high probability,  $G(n, 1/2)$  does not contain a  $K_t$ -minor with  $t = 4n/\sqrt{\log n}$ .

Taking an appropriate  $n = n(t)$  one gets from the above exercise an  $n$ -vertex graph with average degree  $\Omega(t\sqrt{\log t})$  and no  $K_t$ -minor. Taking disjoint copies of such a graph, gives arbitrarily large graphs with  $\Omega(t\sqrt{\log tn})$  edges and no  $K_t$ -minor. Let us note that with high probability  $G(n, 1/2)$  is  $n/2 \log n$ -colorable and contains a  $K_t$ -minor with  $t = n/2\sqrt{\log n}$ , so as opposed to Hajos's Conjecture,  $G(n, 1/2)$  satisfies Hadwiger's Conjecture.

As to Theorem 5.9, it turns out that  $O(t^2 n)$  are sufficient and necessary for forcing a  $K_t$ -minor. Again, the upper bound is hard while the lower bound follows from one of the items in Exercise 5.2 regarding  $G(n, 1/2)$ . Actually, as the following exercise suggests, one does not even need to consider random graphs.

**Exercise 5.13.** Show that  $K_{t^2/10, t^2/10}$  does not contain a topological  $K_t$ -minor.

As in the case of  $K_t$ -minors, taking several disjoint copies of  $K_{t^2/9, t^2/9}$  we get arbitrarily large graphs with  $\Omega(t^2 n)$  edges and no topological  $K_t$ -minor.

### 5.3 Application of Topological $K_t$ -Minors to Graph Linkage

**Definition.** A graph  $G$  is called  $k$ -linked if for every collection of  $k$  pairs of vertices  $(s_1, t_1), \dots, (s_k, t_k)$ , with all  $2k$  vertices being distinct, one can find in  $G$  a collection of  $k$  vertex-disjoint paths  $P_1, \dots, P_k$  such that  $P_i$  connects  $s_i$  to  $t_i$ .

Note that being  $k$ -linked is a stronger property than being  $k$ -connected, that is, every  $k$ -linked graph is also  $k$ -connected. It is not clear a-priori that any measure of connectivity guarantees that a graph is  $k$ -linked. The following theorem shows that this is actually the case.

**Theorem 5.14** (Bollobas-Thomasson). *If  $G$  is  $22k$ -connected then  $G$  is  $k$ -linked [Bollobas and Thomason, 1996].*

---

<sup>8</sup>The graph has at least  $n$  edges  $\implies$  it must contain a cycle.



**Exercise 5.15.** Use the above theorem to show that for large enough  $C$ , every graph with  $Ct^2n$  edges contains a topological  $K_t$ -minor. You might find it useful to use Mader's Theorem that states that a graph with average degree  $4k$  contains a  $k$ -connected subgraph.

We now prove a weaker version of Theorem 5.14.

**Theorem 5.16.** *If  $G$  is  $2 \cdot 2^{\binom{3k}{2}}$ -connected then  $G$  is  $k$ -linked.*

**Proof:** We prove a stronger statement, that if  $h(t)$  is such that every graph with average degree  $h(t)$  contains a topological  $K_t$ -minor, then every graph which is  $(h(3k) + 2k)$ -connected is  $k$ -linked. Recall that we proved in Theorem 5.9 that  $h(t) \leq 2^{\binom{t}{2}}$  thus giving the result in the statement. Note further, that if instead we plug in the fact that  $h(t) = O(t^2)$  (which we did not prove) we get that every  $O(k^2)$ -connected graph is  $k$ -linked, which comes close to the linear bound of Theorem 5.14. Combining this with the exercise above we get that bounding the function  $h(t)$  is essentially equivalent to proving Theorem 5.14.

Let us turn to proving the statement mentioned above. If  $G$  is  $(h(3k) + 2k)$ -connected then  $\delta(G) \geq h(3k)$  hence by definition of  $h(t)$  we get that  $G$  contains a topological  $K_{3k}$ -minor. By Menger's theorem,  $G$  has  $2k$  vertex disjoint paths from  $s_1, t_2, \dots, s_k, t_k$  to  $2k$  of the "vertices" of the topological  $K_{3k}$ -minor. If exists  $u_1, \dots, u_k$  in the  $K_{3k}$  which are not part of these paths, then we can connect  $s_i$  to some vertex in the  $K_{3k}$  which can be connected via  $u_i$  to the vertex in  $K_{3k}$  connected to  $t_i$  by the Menger paths. Thus we construct  $k$  vertex disjoint paths between the pairs  $s_i, t_i$  for  $1 \leq i \leq k$ .  $\square$

## 6 Sixth Lecture

### 6.1 Introduction to Szemerédi's Regularity Lemma

**What is the Regularity Lemma?**

**Definition 6.1.** A bipartite graph  $(A, B)$  is called  $\varepsilon$ -regular if for every  $A' \subseteq A$  and  $B' \subseteq B$  such that  $|A'| \geq \varepsilon|A|$  and  $|B'| \geq \varepsilon|B|$ , we have

$$|d(A', B') - d(A, B)| \leq \varepsilon$$

where  $d(X, Y) = e(X, Y)/|X||Y|$  is the edge density between  $X$  and  $Y$ .

Note that an  $\varepsilon$ -regular has certain properties one expects to find<sup>9</sup> in a random bipartite graph of edge density  $d$ . So one can think of being  $\varepsilon$ -regularity as being  $\varepsilon$ -pseudo-random.

**Definition 6.2.** A partition of  $V(G)$  into sets  $V_1, \dots, V_k$  is called an equipartition if the sizes all the partitions differ by 1 at most.

**Definition 6.3.** An equipartition  $V_1, \dots, V_k$  of  $V(G)$  is called  $\varepsilon$ -regular if all but at most  $\varepsilon k^2$  of the pairs  $(V_i, V_j)$  form  $\varepsilon$ -regular graph.

**Theorem 6.4 (Regularity Lemma).** *For every  $\varepsilon > 0$  exists  $T = T(\varepsilon)$  such as every graph has an  $\varepsilon$ -regular equipartition of order  $k$  where  $\frac{1}{\varepsilon} \leq k \leq T(\varepsilon)$  [Szemerédi, 1975].*

<sup>9</sup>To be more precise, we would have expected such behavior in a random graph even from subsets of size  $\frac{\log n}{\varepsilon}$ .

## 6.2 The Triangle-Removal Lemma and Roth's Theorem

Let us say that a graph  $G$  is  $\varepsilon$ -far from being triangle free if one has to remove at least  $\varepsilon n^2$  edges from  $G$  in order to make it triangle-free. We now consider the following problem; what is the relation between how far  $G$  is from being triangle-free and the number of triangles in  $G$ ? One direction is very easy.

**Exercise.** Show that if  $G$  contains  $\varepsilon n^3$  triangles then  $G$  is  $\varepsilon$ -far from being triangle free.

The following famous lemma states that the reverse implication also holds.

**Theorem 6.5** (Triangle Removal Lemma). *For every  $\varepsilon > 0$  there exists  $\delta = \delta(\varepsilon) > 0$  such as if  $G$  is a  $\varepsilon$ -far from being triangle-free then  $G$  has at least  $\delta n^3$  triangles [Ruzsa and Szemerédi, 1978].*

For the proof of the triangle-removal lemma we will use the following lemma. Note that we can think of its statement as saying that if a 3-partite graph is dense/regular enough then it contains approximately the same number of triangles one expects to find in a “truly” random 3-partite graph of the same density.

**Lemma 6.6.** *If  $A, B, C$  are vertex sets such as  $d(A, B) = b, d(A, C) = c, d(B, C) = a$  and  $a, b, c \geq 2\varepsilon$  and all three bipartite graphs are  $\varepsilon$ -regular then the graph  $(A, B, C)$  has at least*

$$(1 - 2\varepsilon)(a - \varepsilon)(b - \varepsilon)(c - \varepsilon)|A||B||C|$$

triangles.

**Proof:** Note that  $A$  has at most  $\varepsilon|A|$  vertices whose number of neighbors in  $B$  is smaller than  $(b - \varepsilon)|B|$ . Otherwise, this set of vertices together with  $B$  would have given a pair of subsets contradicting the  $\varepsilon$ -regularity of  $(A, B)$ . The same goes for  $A$  and  $C$ . All in all, all the vertices of  $A$ , apart from  $\leq 2\varepsilon|A|$ , have at least  $(b - \varepsilon)|B|$  neighbors in  $B$  and  $(c - \varepsilon)|C|$  neighbors in  $C$ . Let  $v \in A$  be one of the “good” vertices. As  $b, c \geq 2\varepsilon$  and  $|N_B(v)| \geq \varepsilon|B|, |N_C(v)| \geq \varepsilon|C|$  we can apply the Regularity Lemma on  $N_B(v), N_C(v)$  and get that the number of edges between them is at least

$$(a - \varepsilon)|N_B(v)||N_C(v)| \geq (a - \varepsilon)(b - \varepsilon)(c - \varepsilon)|B||C|$$

thus every “good” vertex takes part in at least that number of triangles. As at least  $(1 - 2\varepsilon)|A|$  are “good” vertices we get that there are at least

$$(1 - 2\varepsilon)(a - \varepsilon)(b - \varepsilon)(c - \varepsilon)|A||B||C|$$

triangles. □

**Proof (Triangle Removal Lemma):** Given a graph  $G$  we apply the Regularity Lemma with  $\varepsilon/4$ . We get an  $\varepsilon/4$ -regular partition  $V_1, \dots, V_k$  with  $\frac{4}{\varepsilon} \leq k \leq T\left(\frac{\varepsilon}{4}\right)$ . We remove from the graph the following edges:

1. Edges that belong to one of the clusters  $V_i$ . There are at most  $k \frac{\binom{n/k}{2}}{2} \leq \frac{n^2}{2k} \leq \frac{\varepsilon n^2}{8}$  such edges.
2. Edges between non  $\varepsilon$ -regular  $(V_i, V_j)$  pairs. There are most  $\frac{\varepsilon}{4} k^2 (n/k)^2 = \frac{\varepsilon}{4} n^2$  such edges.
3. Edges between  $(V_i, V_j)$  such that  $d(V_i, V_j) \leq \frac{\varepsilon}{2}$ . There are at most  $\binom{k}{2} \frac{\varepsilon}{2} (n/k)^2 \leq \frac{\varepsilon}{4} n^2$  such edges.

As we removed at most  $\frac{5}{8}\varepsilon n^2$  edges, the new graph, call it  $G'$ , still has a triangle. From the type of edges we removed from  $G$ , this means that there exist three vertex sets  $V_1, V_2, V_3$  such that this triangle has a vertex in each set and so that each of the bipartite graphs  $(V_1, V_2), (V_2, V_3), (V_1, V_3)$  is  $\varepsilon/4$ -regular and

has density at least  $\varepsilon/2$ . Lemma 6.6 now tells us that the number of triangles spanned by  $V_1, V_2, V_3$  is at least

$$\frac{1}{2} \left(\frac{\varepsilon}{4}\right)^3 \left(\frac{n}{T(\frac{\varepsilon}{4})}\right)^3.$$

Since each of these triangles is also a triangle of  $G$ , this proves the lemma with  $\delta(\varepsilon) = \frac{\varepsilon^3}{128T^3(\frac{\varepsilon}{4})}$ .  $\square$

We will now use the triangle-removal lemma to prove the following famous theorem.

**Theorem 6.7** (Roth’s Theorem). *If  $S \subseteq [n]$  does not contain 3-term arithmetic progression then  $|S| = o(n)$ . [Roth, 1953]*

**Proof** (by Ruzsa-Szemerédi): Given  $S$  we construct the following graph:  $G$  has  $6n$  vertices in three sets  $A, B, C$ , where  $|A| = n$ ,  $|B| = 2n$  and  $|C| = 3n$ . We treat the vertices as integers. For every  $s \in S$  and every  $x \in [n]$  we put in  $G$  a triangle, denoted  $T_{x,s}$  on the vertices  $x \in A, x + s \in B, x + 2s \in C$ . For ease of reference, let us label the 3 edges of each of these  $n|S|$  triangles  $T_{x,s}$  with the integer  $s$  used to define them. We claim that the  $n|S| = \varepsilon n^2$  triangles we put into  $G$  are edge disjoint. To see this, note that given any edge of the graph, we can uniquely recover the integers  $x$  and  $s$  that were used when placing it as part of the triangle  $T_{x,s}$ , implying that it belongs to a unique triangle. As the graph has  $\varepsilon n^2$  edge-disjoint triangles, it is clear that at least this many edges need to be removed in order to make it triangle free. By the Triangle Removal Lemma, we get that  $G$  has at least  $c(\varepsilon)n^3$  triangles. For  $n$  sufficiently large  $c(\varepsilon)n^3 > \varepsilon n^2$  hence we get a “new” triangle which is not one of those we explicitly constructed. In other words, we get a triangle, so that the labels on its three edges are  $s_1, s_2, s_3$  and these three edges are not identical (if they were this would have been one of the triangles  $T_{x,s}$  we placed in  $G$ ). Suppose  $x$  is the vertex of this triangle in  $A$ . Then walking along the edges of this triangle, we see that  $x + s_1 + s_2 - 2s_3 = x$ , implying that  $s_1 + s_2 = 2s_3$ , that is,  $s_1, s_3, s_2$  form a (non-trivial) 3-term arithmetic progression.

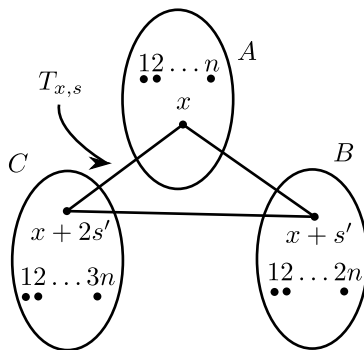


Figure 6.1: The graph with the  $T_{x,s}$  triangle.

Let  $Tower(x)$  be a tower of exponents of height  $x$  (so  $Tower(4) = 2^{2^{2^2}}$ ). As we will see later on the function  $T(\varepsilon)$  from the statement of the Regularity Lemma is given by  $Tower(1/\varepsilon^5)$ . It follows that the function  $\delta(\varepsilon)$  from the triangle-removal lemma is also a tower-type function. This in turn, implies that the above bound for Roth’s Theorem gives that if  $S \subseteq [n]$  does not contain a 3-term arithmetic progression then  $|S| = O(n/\log^* n)$ , which is just “barely”  $o(n)$ . The best known upper bound for Roth’s Theorem is  $O(n(\log \log n)^5/\log n)$  and the best lower bound is  $n/2^{c\sqrt{\log n}}$  (we will prove the lower bound later on).

A surprising result of Gowers showed that  $T(\varepsilon) \geq \text{Tower}(\varepsilon^{-1/16})$ , that is, that there are graphs with the property that every  $\varepsilon$ -regular graphs must be of size at least  $\geq \text{Tower}(\varepsilon^{-1/16})$ . The best known lower bound for the Triangle Removal Lemma is roughly  $\delta(\varepsilon) > 1/\text{Tower}(\log 1/\varepsilon)$ . Although this is still a Tower-type bound it is better than then  $1/\text{Tower}(1/\varepsilon)$  bound that follows from the regularity lemma (as we have done above). The best upper bound for the triangle removal lemma is  $(1/\varepsilon)^{\log(1/\varepsilon)}$  which is just barely super-polynomial. We will prove this bound in a later lecture.  $\square$

### 6.3 Proof of the Regularity Lemma

**Definition 6.8.** Assume  $P = \{V_1, V_2, \dots, V_k\}$  is a partition of  $G$  to sets of size  $a_1n, \dots, a_kn$ , where  $0 \leq a_i \leq 1$  and  $\sum a_i = 1$ . Define

$$q(V_i, V_j) = a_i a_j d^2(V_i, V_j)$$

$$q(P) = \sum_{i,j} q(V_i, V_j)$$

Note that since  $0 \leq d(V_i, V_j) \leq 1$  we always have  $0 \leq q(P) \leq 1$ . The key part in proving the Regularity Lemma is the following lemma.

**Lemma 6.9 (Key Lemma).** *If  $P = \{V_1, V_2, \dots, V_k\}$  is a equipartition which is not  $\varepsilon$ -regular and  $k \geq \frac{1}{\varepsilon^6}$  then it is possible to refine  $P$  to form a new equipartition,  $P'''$ , such that  $q(P''') \geq q(P) + \frac{\varepsilon^5}{2}$  and if  $k'$  is the size of the partitions of  $P'''$  then  $k' \leq k^2 4^k$ .*

Let us quickly derive the regularity lemma from the above lemma.

**Proof (Regularity Lemma):** We start with an arbitrary equipartition of order  $\frac{1}{\varepsilon^6}$  and then repeatedly apply Lemma 6.9 until we get an  $\varepsilon$ -regular partition. As we observed above, we always have  $0 \leq q(P) \leq 1$ , and since each application of Lemma 6.9 increases  $q(P)$  by at least  $\varepsilon^5/2$  this process must stop after at most  $\frac{2}{\varepsilon^5}$  iterations with a partition of order at most  $\text{Tower}(4/\varepsilon^5)$ .  $\square$

We turn to prove Lemma 6.9. We start with the following simple calculation.

**Claim 6.10.** *Assume  $A, B$  are vertex sets of sizes  $an, bn$  in some graph  $G$ , and assume  $A$  was partitioned into  $l$  subsets  $A_1, \dots, A_l$  of sizes  $x_1an, \dots, x_lan$  and  $B$  was partitioned into  $l$  subsets  $B_1, \dots, B_l$  of sizes  $y_1bn, \dots, y_lbn$  (where  $0 \leq x_i, y_i \leq 1$  and  $\sum x_i = \sum y_i = 1$ ). Suppose for every  $i, j$  we have  $d(A_i, B_j) = d(A, B) + \varepsilon_{ij}$  ( $\varepsilon_{ij}$  can be either positive or negative). Then we have*

$$\sum_{i,j} q(A_i, B_j) = q(A, B) + ab \sum_{i,j} x_i y_j \varepsilon_{ij}^2$$

**Proof:** Set  $d = d(A, B)$  and just note that

$$\begin{aligned} \sum_{i,j} q(A_i, B_j) &= \sum_{i,j} x_i a y_j b d^2(A_i, B_j) = ab \sum_{i,j} x_i y_j (d + \varepsilon_{ij})^2 \\ &= ab \sum_{i,j} x_i y_j (d^2 + 2d\varepsilon_{ij} + \varepsilon_{ij}^2) = q(A, B) + 0 + ab \sum_{i,j} x_i y_j \varepsilon_{ij}^2 \end{aligned}$$

where in the last equality we use the facts that  $\sum_{i,j} x_i y_j = 1$ , and that  $\sum_{i,j} x_i y_j \varepsilon_{ij} = 0$ . This last equality follows from the fact that

$$d = \frac{e(A, B)}{abn^2} = \frac{\sum_{i,j} e(A_i, B_j)}{abn^2} = \sum_{i,j} x_i y_j d(A_i, B_j) = \sum_{i,j} x_i y_j (d + \varepsilon_{ij}) = d + \sum_{i,j} x_i y_j \varepsilon_{ij} .$$

□

The claim above gives us the following immediate corollaries:

**Corollary 6.11.** *If we take two vertex sets,  $A, B$ , participating in some partition  $P$ , and refine them into finer sets  $A_1, \dots, A_l$ , and  $B_1, \dots, B_l$  then we have*

$$\sum_{i,j} q(A_i, B_j) \geq q(A, B) \quad (6.1)$$

**Corollary 6.12.** *Assume  $A, B$  are two vertex sets, participating in some equipartition  $P$ , and  $(A, B)$  is not  $\varepsilon$ -regular, namely there exists  $A_1 \subseteq A$  and  $B_1 \subseteq B$  such as  $|A_1| \geq \varepsilon|A|$  and  $|B_1| \geq \varepsilon|B|$  such that  $|d(A_1, B_1) - d(A, B)| > \varepsilon$ . Set  $A_2 = A \setminus A_1$  and  $B_2 = B \setminus B_1$ . Then if  $P$  has order  $k$  (so  $|A| = |B| = \frac{n}{k}$ ) then we have*

$$\sum_{1 \leq i, j \leq 2} q(A_i, B_j) \geq q(A, B) + \frac{\varepsilon^4}{k^2}$$

Now we can tackle the proof of Lemma 6.9.

**Proof:** Assume  $P = \{V_1, \dots, V_k\}$  is not  $\varepsilon$ -regular and denote by  $I$  the set of pairs  $(i, j)$  such that  $(V_i, V_j)$  is not  $\varepsilon$ -regular. So  $|I| \geq \varepsilon k^2$ . We construct a partition (not equipartition)  $P'$  satisfying  $q(P') \geq q(P) + \varepsilon^5$  as follows. Assume  $(V_i, V_j)$  is not  $\varepsilon$ -regular. Then there are  $V_1^{i,j} \subseteq V_i$  and  $V_1^{j,i} \subseteq V_j$  which witness this fact, that is, they satisfy  $|V_1^{i,j}| \geq \varepsilon|V_i|$ ,  $|V_1^{j,i}| \geq \varepsilon|V_j|$  and  $|d(V_i, V_j) - d(V_1^{i,j}, V_1^{j,i})| > \varepsilon$ . We then define  $P'$  to be the “atoms” of the Venn diagram of the sets  $\{V_1, \dots, V_k\}$  and  $\{V_1^{i,j}, V_1^{j,i} : (i, j) \in I\}$ . As for each set  $V_i$  we have at most  $k-1$  sets  $V_1^{i,j}$  we get that each  $V_i$  is refined into at most  $2^{k-1}$  sets and so the total number of sets is  $k2^{k-1}$ . We now compare  $q(P')$  to  $q(P)$ . To this end, suppose each set  $V_i$  of  $P$  is refined in  $P'$  into sets  $V_i^1, \dots, V_i^l$ . Also for every pair  $(i, j) \in I$  set  $V_2^{i,j} = V_i \setminus V_1^{i,j}$  and  $V_2^{j,i} = V_j \setminus V_1^{j,i}$  and note that for any  $j$ , the sets  $V_i^1, \dots, V_i^l$  refine the pair  $V_1^{i,j}, V_2^{i,j}$ . Corollary 6.11 tells us that for any  $i, j$  we have  $\sum_{i',j'} q(V_i^{i'}, V_j^{j'}) \geq q(V_i, V_j)$ . Furthermore, if  $(i, j) \in I$  then applying Corollary 6.12 and then Corollary 6.11 we get

$$\sum_{i',j'} q(V_i^{i'}, V_j^{j'}) \geq q(V_1^{i,j}, V_1^{j,i}) + q(V_1^{i,j}, V_2^{j,i}) + q(V_2^{i,j}, V_1^{j,i}) + q(V_2^{i,j}, V_2^{j,i}) \geq q(V_i, V_j) + \varepsilon^4/k^2 .$$

In other words, we recover all the contributions  $q(V_i, V_j)$  and when  $(i, j)$  is not  $\varepsilon$ -regular we in fact gain at least  $\varepsilon^4/k^2$ . Since we assume that there are at least  $\varepsilon k^2$  pairs that are not  $\varepsilon$ -regular we conclude that  $q(P') \geq q(P) + \varepsilon^5$ .

We are now just left with turning  $P'$  into an equipartition while not decreasing  $q$  by much. We start with partitioning cluster of  $P'$  into clusters of size  $n/k^2$ . After pulling for each of the clusters of  $P'$  subsets of size sets might also leave a remainder of size at most  $\frac{n}{k^2}$ . As  $P''$  is a refinement of  $P'$  we get from Corollary 6.11 that we still have  $q(P'') \geq q(P') \geq q(P) + \varepsilon^5$ . Assume  $P'' = \{V_1, \dots, V_{k''}, U_1, \dots, U_{k''}\}$  where  $V_1, \dots, V_{k''}$  are the subsets of size exactly  $\frac{n}{k^2}$  and  $U_1, \dots, U_{k''}$  are the remainders (of size  $< \frac{n}{k^2}$ ). Note that we also have  $\sum_i |U_i| \leq k'' \cdot \frac{n}{k^2} = \frac{n}{k'} \leq \varepsilon^6 n$ . To get a true equipartition we now simply “distribute” the vertices of  $U_1, \dots, U_{k''}$  equally into the sets  $V_1, \dots, V_{k''}$  thus obtaining an equipartition  $P'''$  of order  $k''$ . It should be clear that since  $\sum_i |U_i| \leq \varepsilon^6 n$  this cannot decrease  $q(P'')$  by much. In fact, it follows from Exercise 6.13 below that  $q(P''') \geq q(P'') - 8\varepsilon^6$ . Since we can always assume that  $\varepsilon$  is small enough (make sure you see why), we get

$$q(P''') \geq q(P'') - 8\varepsilon^6 \geq q(P) + \varepsilon^5 - 8\varepsilon^6 \geq q(P) + \varepsilon^5/2 ,$$

as needed. Finally, note that  $P'''$  has order  $k^2$  and since  $k' \leq k2^{k-1}$ , the proof is complete. □

**Exercise 6.13.** Suppose  $P = \{V_1, \dots, V_k, U_1, \dots, U_t\}$  is a partition of  $V(G)$  with  $|V_1| = \dots = |V_k|$  and  $\sum_{i=1}^t |U_i| \leq \varepsilon n$ . Show how to turn  $P$  into an equipartition  $P'$  of order  $k$  satisfying  $q(P') \geq q(P) - 8\varepsilon$ .

## 7 Seventh Lecture

### 7.1 The Counting Lemma

The following lemma, which is usually referred to as the counting lemma, states that if a collection of vertex sets are regular/dense enough then we can find in them any small subgraphs we would expect to find, if the graphs were genuinely random. It generalizes a similar lemma we proved last time for triangles.

**Theorem 7.1** (Embedding Lemma). *For every  $d > 0$  and  $\Delta$  there exists  $\varepsilon = \varepsilon(d, \Delta)$  and  $c = c(d, \Delta)$  so that the following holds: Assume  $V_1, \dots, V_r$  are vertex-sets of size at least  $ch$  such that for every  $i, j$  the pair  $(V_i, V_j)$  is  $\varepsilon$ -regular and satisfies  $d(V_i, V_j) \geq d$ . Then, if  $H$  is a graph on at most  $h$  vertices with  $\chi(H) \leq r$  and  $\Delta(H) \leq \Delta$  then  $V_1, \dots, V_r$  span a copy of  $H$ .*

We will prove that there are in fact  $c(\varepsilon, h)|V|^h$  copies of  $H$  (assuming all sets are of the same size  $|V|$ ). We will first need the following simple observation regarding  $\varepsilon$ -regular pairs.

**Claim 7.2.** *Assume  $(A, B)$  is a  $\varepsilon$ -regular pair with density  $d$ . Then, if  $B' \subseteq B$  is of size at least  $\varepsilon|B|$ , then there are at most  $\varepsilon|A|$  vertices in  $A$  that have less than  $(d - \varepsilon)|B'|$  neighbors in  $B'$ .*

**Proof:** Otherwise, taking  $A'$  to be the subset of vertices with less than  $(d - \varepsilon)|B'|$  neighbors in  $B'$  we see that  $A', B'$  contradict the  $\varepsilon$ -regularity of  $(A, B)$ .  $\square$

**Proof (Embedding Lemma):** Assume an  $r$ -coloring of  $H$ 's vertices induces the partition into  $r$  vertex-disjoint sets  $U_1, \dots, U_r$ . We shall find an embedding of  $H$  in  $V_1, \dots, V_h$ , where the vertices of  $U_i$  are chosen out of  $V_i$ . Namely, every vertex  $v_i \in H$  has a set  $V_{\sigma(i)}$  in which it shall be embedded. The idea of the proof is to simply embed the vertices  $v_1, \dots, v_h$  in a ‘‘greedy’’ way. We will just have to show that we never get ‘‘stuck’’.

Suppose that for some  $0 \leq j \leq h - 1$  we have already found vertices  $x_1, \dots, x_j$  in  $V_{\sigma(1)}, \dots, V_{\sigma(j)}$  so that  $x_i, x_j$  are connected if  $v_i, v_j$  are connected in  $H$ . For every  $i > j$  we define a subset  $Y_i^j \subseteq V_{\sigma(i)}$  of all the possible vertices which can be chosen for  $v_i$  given the choice of  $x_1, \dots, x_j$ , namely all the vertices  $x$  that satisfy the following condition: if  $j' \leq j$  and  $v_i$  is connected to  $v_{j'}$  in  $H$  then  $x$  is connected to  $x_{j'}$ . Observe that for  $j = 0$  we can set  $Y_i^0 = V_{\sigma(i)}$  for every  $1 \leq i \leq h$ .

For every  $0 \leq j < i \leq h$  let  $b_{j,i}$  denote the number of vertices among  $v_1, \dots, v_j$  that are adjacent to  $v_i$  in  $H$ . We claim that if we choose  $\varepsilon$  appropriately then we can pick the vertices  $x_1, \dots, x_h$  so that for every  $0 \leq j < i \leq h$  we have  $|Y_i^j| \geq (d - \varepsilon)^{b_{j,i}} |V_{\sigma(i)}|$ . This clearly holds for  $j = 0$  (since  $b_{0,i} = 0$ ) so assuming it holds for  $j - 1$  we prove it for  $j$ . In this case we need to pick a vertex  $x_j \in Y_j^{j-1} \subseteq V_{\sigma(j)}$  that will play the role of  $v_j$ . The definition of  $Y_j^{j-1}$  guarantees that any vertex in this set is adjacent to the appropriate vertices among  $x_1, \dots, x_{j-1}$ . Once we pick a vertex  $x_j$  we will need to set  $Y_i^j$  (with  $i > j$ ) to be the subset of vertices of  $Y_i^{j-1}$  that are adjacent to  $x_j$ . Recall that we know from induction that  $|Y_i^{j-1}| \geq (d - \varepsilon)^{b_{j-1,i}} |V_{\sigma(i)}|$  for every  $i \geq j$ . This means that if  $i > j$  and  $v_i$  is not adjacent to  $v_j$  then after picking  $x_j$  we can take  $Y_i^j = Y_i^{j-1}$  so the condition still holds for such vertices. As to  $i$  for which  $v_i$  is adjacent to  $v_j$ , note that in this case we just need to make sure that  $|Y_i^j| \geq (d - \varepsilon)^{b_{j,i}} |V_{\sigma(i)}|$ . Observe that if  $\varepsilon \leq (d/2)^\Delta$  then since  $b_{i,j} \leq \Delta$  we get from induction that

$$|Y_i^{j-1}| \geq (d - \varepsilon)^{b_{j-1,i}} |V_{\sigma(i)}| \geq (d - \varepsilon)^\Delta |V_{\sigma(i)}| \geq \varepsilon |V_{\sigma(i)}|.$$

We thus get from Claim 7.2, with  $A, B$  being the sets  $V_{\sigma(j)}, V_{\sigma(i)}$  and  $B' = Y_i^{j-1}$ , that  $V_{\sigma(j)}$  contains at most  $\varepsilon|V_{\sigma(j)}|$  vertices that have less than  $(d - \varepsilon)|Y_i^{j-1}|$  neighbors in  $Y_i^{j-1}$ . Since  $H$  has maximum degree  $\Delta$  we get that all but  $\Delta\varepsilon|V_{\sigma(j)}|$  of the vertices in  $V_{\sigma(j)}$  have at least  $(d - \varepsilon)|Y_i^{j-1}|$  neighbors in each of the sets  $Y_i^{j-1}$  for which  $v_i$  is a neighbor of  $v_j$  in  $H$ . Note that choosing any of these vertices will make sure that  $|Y_i^j| \geq (d - \varepsilon)|Y_i^{j-1}|$  as we wanted. We thus get that we have at least

$$|Y_j^{j-1}| - \Delta\varepsilon|V_{\sigma(j)}| \geq (d - \varepsilon)^{b_{j-1,j}}|V_{\sigma(j)}| - \Delta\varepsilon|V_{\sigma(j)}| \geq (d - \varepsilon)^\Delta|V_{\sigma(j)}| - \Delta\varepsilon|V_{\sigma(j)}| \geq \varepsilon|V_{\sigma(j)}|,$$

choices for a vertex  $x_j \in Y_j^{j-1}$  satisfying the required condition. The only thing we need to make sure is that  $\varepsilon$  is small enough to that the last inequality holds. It is easy to check that taking  $\varepsilon = (d/2)^\Delta / 2\Delta$  satisfies this inequality as well as the previous one we wanted  $\varepsilon$  to satisfy. Finally, note that up to  $h - 1$  of the vertices of  $Y_{j-1}^j$  might have already been chosen as vertices  $x_{j'}$  for some  $j' < j$  so we need to make sure that  $\varepsilon|V_{\sigma(j)}| - h \geq 1$ . Taking  $c = 1/\varepsilon$  makes sure this condition holds. This completes the proof of the induction.

We finally note that if each set is of size at least  $(2/\varepsilon)h$  then  $\varepsilon|V_{\sigma(j)}| - h \geq (\varepsilon/2)|V_{\sigma(j)}|$  so at each iteration we actually have at least  $(\varepsilon/2)|V_{\sigma(j)}|$  choices for the vertex  $x_j$ .  $\square$

## 7.2 Another Proof of the Erdős-Stone-Simonovits Theorem

The embedding lemma gives us a pretty simple proof of the **Erdős-Stone-Simonovits**:

**Theorem.** *For every  $b, r$  and  $\delta > 0$  there exists  $n_0(\delta, r, b)$  so that if  $G$  is a graph on at least  $n_0$  vertices with  $(1 - \frac{1}{r} + \delta) \frac{n^2}{2}$  edges then  $G$  contains a copy of  $K_{r+1}^b$  (a  $b$ -blowup of  $K_{r+1}$ ).*

**Proof:** Define  $\varepsilon = \varepsilon(\frac{\delta}{8}, rb)$  and  $c = c(\frac{\delta}{8}, rb)$  and apply the **Regularity Lemma** with  $\min(\varepsilon, \frac{\delta}{8})$ . We get an  $\varepsilon$ -regular partition into  $k$  sets  $V_1, \dots, V_k$  where  $\frac{1}{\varepsilon} \leq k \leq T(\varepsilon)$  and therefore  $|V_i| \geq \frac{n}{T(\varepsilon)}$ . Remove the following edges from the graph:

1. Edges within  $V_i$ . There are at most  $\frac{\delta n^2}{8}$  such edges.
2. Edges connecting  $V_i$  to  $V_j$  where  $d(V_i, V_j) < \frac{\delta}{8}$ . There are at most  $\frac{\delta n^2}{4}$  such edges.
3. Edges connecting  $V_i$  to  $V_j$  where  $(V_i, V_j)$  is not  $\varepsilon$ -regular. There are at most  $\frac{\delta n^2}{8}$  such edges.

All in all, we removed less than  $\frac{\delta n^2}{2}$  edges. Therefore  $G$  still has at least  $(1 - \frac{1}{r}) \frac{n^2}{2}$  edges, so by **Turán's Theorem (weak version)**  $G$  contains a copy of  $K_{r+1}$ . This  $K_{r+1}$  must have a single vertex in each one of  $r+1$  sets  $V_1, \dots, V_{r+1}$  so that all pairs  $(V_i, V_j)$  are  $\varepsilon$ -regular and satisfy  $d(V_i, V_j) \geq \frac{\delta}{8}$ . We have  $r+1$  sets of size at least  $\frac{n}{T(\varepsilon)}$  such as the density between them is at least  $\frac{\delta}{8}$  and every one of them is  $\varepsilon(\frac{\delta}{8}, rb)$ -regular. The graph  $K_{r+1}^b$  is an  $(r+1)$ -colorable graph with maximal degree  $rb$ . Hence, the Embedding Lemma would imply that  $V_1, \dots, V_{r+1}$  have a copy of  $K_{r+1}^b$  if

$$|V_i| \geq \frac{n}{T(\varepsilon)} \geq c \left( \frac{\delta}{8}, rb \right) \cdot (r+1)b$$

So to conclude the proof, we just set

$$n_0 = c \left( \frac{\delta}{8}, rb \right) \cdot b(r+1) \cdot T \left( \min \left\{ \frac{\delta}{8}, \varepsilon \left( \frac{\delta}{8}, rb \right) \right\} \right).$$

Since  $n_0$  is just a function of  $\delta, r$  and  $b$  the proof is complete.  $\square$

### 7.3 Ramsey Numbers of Bounded Degree Graphs

We can state Ramsey's Theorem as follows.

**Theorem** (Ramsey). *Every 2-coloring of  $K_{4^n}$  has a monochromatic copy of  $K_n$ .*

One can thus consider the more general function  $R(K, H)$  which is the smallest integer  $N$  so that every 2-coloring of  $K_N$  contains either a monochromatic  $K$  or a monochromatic  $H$ . Of course, if both  $K$  and  $H$  are of size  $n$  then we have the trivial bound  $R(K, H) \leq R(K_n, K_n) \leq 4^n$  but it is reasonable to suspect that for "simple" graphs one should be able to come up with better bounds. As it turns out, if the graphs have bounded degree, then it is enough to consider a 2-coloring of the complete graph on  $O(n)$  vertices (as opposed to the  $4^n$  bound for  $K_n$ ).

**Theorem 7.3.** *For every  $\Delta$  there is  $c' = c'(\Delta)$  so that if  $K, H$  are graphs on  $n$  vertices with maximal degree at most  $\Delta$  then  $R(H, K) \leq c'n$ .*

**Proof:** Set  $\varepsilon = \varepsilon(\frac{1}{2}, \Delta)$ ,  $c = c(\frac{1}{2}, \Delta)$ , and  $m = 4^{\Delta+1}$ . Given 2-coloring of  $K_{c'n}$  we apply the regularity lemma with  $\varepsilon' = \min\{\varepsilon, \frac{1}{4m}\}$  on the black edges. We get a regular partition of size at least  $T(\varepsilon') \geq 2m$ . There are  $k \geq 2m$  sets and at most  $\frac{k^2}{4m}$  of the pairs are not regular. Therefore at least  $\frac{k^2}{2}(1 - \frac{1}{2m})$  pairs are regular. Define a graph on  $k$  vertices and connect vertex  $i$  to  $j$  iff  $(V_i, V_j)$  is an  $\varepsilon'$ -regular pair. This graph has at least  $\binom{k}{2} - \frac{k^2}{4m}$  edges. Since  $k \geq 2m$  we have  $\binom{k}{2} - \frac{k^2}{4m} \geq \frac{k^2}{2}(1 - \frac{1}{m})$ , so by Turán's Theorem this graph contains  $K_m$ , meaning that there exists  $m$  subsets  $V_1, \dots, V_m$  for which every pair is  $\varepsilon'$ -regular.

Consider the following 2-coloring of a graph on  $m = 4^{\Delta+1}$  vertices;  $(i, j)$  is colored green if  $d(V_i, V_j) \geq \frac{1}{2}$  and  $(i, j)$  is blue if  $d(V_i, V_j) < \frac{1}{2}$ . By Ramsey's Theorem, this coloring contains a monochromatic  $K_{\Delta+1}$ . This gives us  $\Delta + 1$  sets  $V_1, \dots, V_{\Delta+1}$  so that every pair is  $\varepsilon'$ -regular and one of the following holds:

1. All the densities of black edges are at least  $\frac{1}{2}$ .
2. All the densities of red edges are at least  $\frac{1}{2}$ .

Assuming case 1 above holds, we now apply the Embedding Lemma on  $V_1, \dots, V_{\Delta+1}$  and find a black copy of  $H$  (case 2 is identical, in which case we find a red copy of  $K$ ). Since  $H$  has maximal degree  $\Delta$  it is  $\Delta + 1$  colorable. The choice of  $\varepsilon$  also guarantees that the pairs are regular and dense enough for an application of the Embedding Lemma. We only need to make sure the sets  $V_i$  are large enough, that is that

$$|V_i| \geq \frac{|G|}{T(\varepsilon')} \geq c \cdot n.$$

Hence, we can set  $c'(\Delta) = c \cdot T(\varepsilon')$ . □

## 8 Eighth Lecture

We continue with more applications of the [Regularity Lemma](#).

### 8.1 The Induced Ramsey Theorem

Recall the [Regularity Lemma](#) and the [Embedding Lemma](#):

**Theorem.** *For every  $\varepsilon > 0$  exists  $T = T(\varepsilon)$  such as every graph has an  $\varepsilon$ -regular equipartition of size  $k$  where  $\frac{1}{\varepsilon} \leq k \leq T(\varepsilon)$ .*



**Theorem.** For every  $d > 0$  and  $\Delta$  exists  $\varepsilon = \varepsilon(d, \Delta)$  and  $c = c(d, \Delta)$  such that if  $V_1, \dots, V_r$  are vertex-sets of size  $\geq ch$  such as  $d(V_i, V_j) \geq d$  for every  $i, j$  and also every pair  $(V_i, V_j)$  is  $\varepsilon$ -regular, then for every graph  $H$ , where  $|H| \leq h$  vertices with  $\chi(H) \leq r$  and  $\Delta(H) \leq \Delta$ , we can find an embedding of  $H$  in  $V_1, \dots, V_r$ .

We now state an induced version of the **Embedding Lemma**:

**Theorem 8.1** (Induced Embedding Lemma). For every  $h$  there exists constants  $\varepsilon'(h)$  and  $c'(h)$  such that if  $V_1, \dots, V_h$  are vertex sets of size  $\geq c'$  and every pair  $(V_i, V_j)$  is  $\varepsilon$ -regular with density  $\in [\frac{1}{8}, \frac{7}{8}]$  then it is possible to embed in  $V_1, \dots, V_h$  an induced copy of  $H$  for every graph  $|H| \leq h$ .

**Exercise 8.2.** Prove the above Induced Embedding Lemma. The case of  $H = K_h$  follows as a special case of the Embedding Lemma we have previously proved. The case of general  $H$  can be reduced to the case of  $K_h$ .

We recall the statement of Ramsey's Theorem.

**Theorem** (Ramsey). For every  $t$ , there exists  $N = N(t)$  so that every 2-coloring of  $K_N$  has a monochromatic copy of  $K_t$ .

It is natural to look for an induced version of Ramsey's Theorem. That is, given a fixed graph  $H$  we want to find a graph  $R = R(H)$  so that in any 2-coloring of  $E(H)$  one can find a monochromatic induced copy of  $H$ . Note that Ramsey's Theorem states that we can take  $K_{4^n}$  as  $R(K_n)$ . Observe that for any other  $H$ , the graph  $R(H)$  cannot be a complete graph. It was an open problem for some time whether  $R(H)$  indeed exists for every  $H$ . The following theorem states that they indeed exist.

**Theorem 8.3** (Induced Ramsey Theorem). For every graph  $H$  there exists a graph  $R = R(H)$  such that in every 2-coloring of  $E(R)$  it is possible to find a monochromatic induced copy of  $H$ .

**Proof:** The idea is very simple; we will show that a large enough random graph can be taken as the graph  $R$ . Of course, we need to define some deterministic property. So set  $\varepsilon = \varepsilon'(h)$ ,  $c = c'(\varepsilon)$ ,  $m = 4^h$ , and  $T = T(\min\{\frac{\varepsilon}{2}, \frac{1}{4m}\})$ . We will show that if  $G$  is a graph on  $n$  vertices such that:

1.  $n/T \geq c$
2. For every pair  $(A, B)$  of disjoint sets of size  $\geq \frac{\varepsilon n}{T}$  it holds that  $|d(A, B) - \frac{1}{2}| \leq \varepsilon/2$

then  $G$  can be taken as the graph  $R(H)$ . Note that we claim that a single graph can be used for all  $H$  on  $h$  vertices.

Let us first explain why there are graphs satisfying the above two conditions. The first condition just asks the graph to be large enough. For the second condition we note that by a simple averaging argument, if the condition  $|d(A, B) - \frac{1}{2}| \leq \varepsilon/2$  holds for subsets  $A, B$  of size  $k$  then it also holds for bigger subsets as well. Now, a simple applications of Chernoff's inequality gives that in  $G(n, 1/2)$ , with high probability every pair of sets  $A, B$  of size  $100 \log n / \varepsilon^2$  satisfy  $|d(A, B) - \frac{1}{2}| \leq \varepsilon/2$ . So we just need to take  $n$  to be large enough so that the first condition will hold, and also large enough so that  $\varepsilon n / T \geq 100 \log n / \varepsilon^2$ , thus making sure that the second condition also holds. So we see that indeed almost all (large enough) graphs satisfy the above two conditions.

We now turn to show that if  $R$  is a graph satisfying the above two conditions then it satisfies the assertions of the theorem. So consider a red/black coloring of  $E(R)$ , and apply the **Regularity Lemma** on the black edges with error parameter  $\min\{\frac{\varepsilon}{2}, \frac{1}{4m}\}$ . The lemma will return an equipartition of the vertices to  $k$  sets where  $2m \leq k \leq T$  so that at most  $\frac{k^2}{4m}$  of the pairs are not  $\frac{\varepsilon}{2}$ -regular.

By **Turán's Theorem (weak version)** we get  $m$  sets  $V_1, \dots, V_m$  such that every pair  $(V_i, V_j)$  is  $\frac{\varepsilon}{2}$ -regular. Now observe that our second assumption about  $G$  implies that all pairs satisfy  $d(V_i, V_j) \in [0.49, 0.51]$

(assuming  $\varepsilon \leq 1/50$  which we can clearly impose). **Ramsey's** Theorem then implies that those  $4^h = m$  sets contain  $h$  sets  $V_1, \dots, V_h$  where every pair  $(V_i, V_j)$  is  $\varepsilon/2$ -regular as above but also one of the following holds:

1. In all the pairs, the density of the black edges is  $\in [1/4, 0.51]$ .
2. In all the pairs, the density of the black edges is  $\in [0, 1/4]$ .

Assume the first case holds, and consider the vertex sets  $V_1, \dots, V_h$ . We now pick some of the edges between these clusters as follows. For every  $(i, j)$ :

1. If  $(i, j) \in H$  then take the black edges between  $V_j$  and  $V_i$ .
2. If  $(i, j) \notin H$  then take the black and red edges between  $V_j$  and  $V_i$ .

We claim that all the requirements of the **Induced Embedding Lemma** hold. This is obvious for  $(i, j) \in H$  since we assume that in these cases  $d(V_i, V_j) \in [0.25, 0.51]$  and that  $(V_i, V_j)$  is  $\varepsilon$ -regular. For  $(i, j) \notin H$  it follows from condition 2 on the graph  $G$  that when taking all (i.e. black and red) edges between  $(V_i, V_j)$  that the resulting graph is  $\varepsilon$ -regular, since all the sets in the partition are of size at least  $n/T$ . As to the density, we know from condition 2 that  $d(V_i, V_j) \in [0.49, 0.51]$ . Finally, note that the first condition on  $G$  gives that  $|V_i| \geq \frac{n}{T} \geq c$ , so we see that we can indeed apply the **Induced Embedding Lemma** to get an induced copy of  $H$ . It is now easy to see that this will be an induced black copy of  $H$  in  $G$ .

Assume now that the second condition holds. As before consider the vertex sets  $V_1, \dots, V_h$ , and pick some of the edges between these clusters as follows. For every  $(i, j)$ :

1. If  $(i, j) \in H$  then take the red edges between  $V_j$  and  $V_i$ .
2. If  $(i, j) \notin H$  then take the black and red edges between  $V_j$  and  $V_i$ .

□

We claim that all the requirements of the **Induced Embedding Lemma** hold. This is obvious for  $(i, j) \in H$  since we assume that in these cases  $d(V_i, V_j) \in [0.25, 0.51]$  and that  $(V_i, V_j)$  is  $\varepsilon$ -regular. For  $(i, j) \notin H$  it follows from condition 2 on the graph  $G$  that when taking all (i.e. black and red) edges between  $(V_i, V_j)$  that the resulting graph is  $\varepsilon$ -regular, since all the sets in the partition are of size at least  $n/T$ . As to the density, we know from condition 2 that  $d(V_i, V_j) \in [0.49, 0.51]$ . Finally, note that the first condition on  $G$  gives that  $|V_i| \geq \frac{n}{T} \geq c$ , so we see that we can indeed apply **Induced Embedding Lemma** to get an induced copy of  $H$ . It is now easy to see that this will be an induced black copy of  $H$  in  $G$ .

We recall the **Triangle Removal Lemma**:

**Theorem.** *For every  $\varepsilon > 0$  there exists  $\delta = \delta(\varepsilon) > 0$  such as if  $G$  is a graph which at least  $\varepsilon n^2$  edges must be removed before it becomes triangle free, then  $G$  has at least  $\delta n^3$  triangles.*

We wish to generalize it to arbitrary graphs.

**Theorem** (Graph Removal Lemma). *For every  $\varepsilon > 0$  and graph  $H$ , there exists  $\delta = \delta_H(\varepsilon)$  such that if  $G$  is  $\varepsilon$ -far from being  $H$ -free<sup>10</sup> then  $G$  has at least  $\delta n^h$  copies of  $H$  ( $|H| = h$ ).*

**Proof:** Exercise. Similar to the case of triangles. □

---

<sup>10</sup>A graph is  $\varepsilon$ -far from being  $H$ -free if one has to remove  $\varepsilon n^2$  edges before it becomes  $H$ -free.

## 8.2 The (6, 3)-Problem and the Induced Matchings Problem

**Definition 8.4** ((6, 3) Problem). Let  $f(n)$  denote the largest number of edges a 3-uniform hypergraph on  $n$  vertices can have, if it does not contain 6 vertices that span at least 3 edges.

**Definition 8.5** (Induced Matching Problem). Let  $g(n)$  be the maximal number of edges in a graph on  $n$  vertices whose edges can be partitioned into at most  $n$  induced matchings.

**Definition 8.6** (One Edge One Triangle Problem). Denote by  $h(n)$  the maximal number of edges in a graph on  $n$  vertices where each edge belongs to exactly one triangle.

**Theorem 8.7.** *We have the following equivalence between the above three problems  $h = \Theta(g) = \Theta(f)$ .*

We will apply the following simple but very useful claim.

**Claim.** *Every graph has a bipartite subgraph with  $\frac{n}{2}$  edges.*

**Proof:** Partition  $V(G)$  randomly to two sets  $A, B$ . For every edge  $(xy)$ , the probability  $x, y$  are in different partitions is  $\frac{1}{2}$ . Hence  $\mathbb{E}[m] = \frac{n}{2}$ , therefore a partition with at least this many edges exists.  $\square$

**Exercise.** Find an explicit way (i.e. an efficient deterministic algorithm) to produce a partition of  $V(G)$  so that at least half the edges cross the cut.

We will prove the equivalence between  $h, g, f$  in the following three claims.

**Claim.**  $g(n) \leq 2f(2n)$ .

**Proof:** Assume  $G$  is a graph on  $n$  vertices and  $M_1, \dots, M_n$  is a partition of its vertices into induced matchings. We take a bipartite subgraph of  $G$  with at least half of the edges.

We construct an hypergraph  $H$  on  $2n$  vertices whose number of edges will be the same as the number of edges in  $G$  and will fulfill the (6, 3)-Problem. Such construction will prove that  $\frac{g(n)}{2} \leq f(2n)$ .

Let  $A = \{v_1, \dots, v_n\}$  be the extra  $n$  vertices we add to  $H$ . For every  $i$ , let  $(x, y)$  be an arbitrary matching in  $M_i$ , in  $H$  we add the edge  $(v_i, x, y)$ .

If we take a set of 6 vertices in  $H$  and assume that they span 3 edges, then the only the following cases exist:

1. If only one vertex is in  $A$ , then at least 6 of the vertices are from  $G$ , which is a contradiction.
2. If two of the vertices are in  $A$ , then 4 are from  $G$  and it is easy to verify that this contradicts the conditions of induced matching.
3. 3 vertices in  $A$ , 3 in  $G$ . This case would've been possible unless we considered a bipartite graph.
4. The rest of the cases (4-6 vertices in  $A$ ) are trivial to reduce into contradictions.

Thus, no set of 6 vertices in  $H$  spans 3 edges, giving  $\frac{g(n)}{2} \leq f(2n)$ .  $\square$

**Claim.**  $h(n) \leq g(n)$

**Proof:** Assume  $G$  fulfills the condition that every edge belongs to one triangle. Define  $n$  matchings  $M_1, \dots, M_n$  in the following way: For every  $i$   $M_i$  would contain the edges spanned by  $N(v_i)$ . Note that

$$\left| \bigcup M_i \right| = |E(G)|$$

as every edge belongs to one triangle, and thus is part of the neighborhood of the opposed vertex. It is easy to verify that every  $M_i$  is indeed an induced matching. Therefore,  $h(n) \leq g(n)$ .  $\square$

**Claim.**  $f(n) \leq h(n) + n$

**Proof:** Assume  $H$  is a 3-uniform hypergraph which fulfills the  $(6, 3)$ -condition. As long as there is a vertex whose degree is 1, remove the vertex and the relevant edge. Assume we performed  $k$  such iterations. The new graph,  $H'$ , fulfills the  $(6, 3)$  condition as a subgraph of  $H$ .

Define a graph  $G$  on  $n$  vertices in the following way: For every edge in  $H'$  we put in  $G$  a triangle on the same vertices. We would like to show that  $G$  fulfills the One-Edge-One-Triangle condition. It is obvious that every edge belongs to  $\geq$  one triangle. Assume that an edge  $e$  belongs to two triangles, then  $H$  as two vertices  $a, b$  who form two edges with  $d, c$ . But,  $d$  must be part of another edge, as we removed all vertices which participate in a single edge, which gives us 3 edges on 6 vertices, which is a contradiction.

Ergo,  $G$  contains  $n - k$  vertices and it has 3 times the number of edges as  $H'$ .  $G$  also fulfills the OEOT condition, implying

$$\begin{aligned} 3m' &\leq h(n - k) \\ m' &\leq \frac{h(n - k)}{2} \\ m &\leq \frac{h(n - k)}{3} + k \leq h(n) + n \end{aligned}$$

□

**Claim.** *The Triangle Removal Lemma implies that  $h(n) = o(n^2)$ .*

**Proof:** Assume  $G$  fulfills the OEOT condition and has  $\varepsilon n^2$  edges. Then  $G$  is the disjoint union of  $\frac{\varepsilon n^2}{3}$  vertex-disjoint triangles, implying that  $G$  is  $\frac{\varepsilon}{3}$ -far from being triangle-free. Therefore, by the Triangle Removal Lemma  $G$  has at least  $\delta(\varepsilon)n^3$  triangles. For sufficiently large  $n$ , we have  $\delta(\varepsilon)n^3 \geq \frac{\varepsilon n^2}{3}$ . Hence, there exists an “extra” triangle. This triangle must have edges from at least two different triangles (in fact, three), making those edges belong to two different triangles, which is a contradiction. □

The result of  $f(n) = o(n^2)$  was first proved by [Ruzsa and Szemerédi, 1978]. Recall that by the equivalences we proved above, we also get that  $g(n) = o(n^2)$  and  $f(n) = o(n^2)$ .

## 9 Ninth Lecture

### 9.1 Behrend’s Construction

In the last lecture we considered three extremal problems; The  $(6, 3)$  Problem, The Induced Matching Problem and The One Edge One Triangle Problem. Denoting by  $f(n), g(n), h(n)$  their respective extremal functions, we proved that  $f = \Theta(g) = \Theta(h) = o(n^2)$ . What the actually gave, is that  $h(n) = O\left(n^2 / (\log^* n)^{1/5}\right)$ , which is just barely sub-quadratic. We will first prove a lower bound for these three extremal-problems. We first prove a very surprising relation between them and the question of bounding  $r_3(n)$ .

**Theorem 9.1.** *Let  $r_3(n)$  denote the size of the largest subset of  $\{1, \dots, n\}$  that contains no 3-term arithmetic progression. Then  $h(n) \geq n \cdot r_3(n)$ .*

We will then prove the following lower bound on  $r_3(n)$ , due to Behrend.

**Theorem 9.2** ([Behrend, 1946]).  $r_3(n) \geq n/2^{c\sqrt{\log n}} = (n^{1-o(1)})$ .

As an immediate corollary of the above two theorems, we will get the following lower bounds for  $f, g, h$ .

**Corollary 9.3.**  $f(n), g(n), h(n) \geq n^2/2^c \sqrt{\log n}$  ( $= n^{2-o(1)}$ ).

**Corollary.** *Combining the results we proved we get the following diagram of implications*

$$\begin{array}{c} TRL \rightarrow OEOT \rightarrow r_3(n) = o(n) \\ \downarrow \searrow \\ IM \leftarrow (6, 3) - P \end{array}$$

**Proof** (Theorem 9.1): We prove that  $h(6n) \geq 3nr_3(n)$ , using the same idea we used in the proof of **Roth's Theorem**. Assume  $S \subseteq \{1, \dots, n\}$  is of size  $r_3(n)$  and does not contain 3-term arithmetic progressions. We construct a graph on  $6n$  vertices as follows: For every  $1 \leq x \leq n$  and every  $s \in S$  we add a triangle on the vertices  $x \in A, x+s \in B, s+2s \in C$  (see Figure 6.1) and label its edges by  $s$ . As in the proof of **Roth's Theorem**, it is easy to see that these triangles are edge-disjoint, and that if the edges of some triangle are labeled by  $s_1, s_2, s_3$  then  $s_1 + s_2 = 2s_3$ . As  $S$  does not contain 3-term arithmetic progressions, it follows that necessarily  $s_1 = s_2 = s_3$  and thus any triangle is one of the original triangles we put in the graph. These two observations imply that each edge belongs to exactly one triangle.  $\square$

**Corollary 9.4.**  $(6, 3)$  Problem  $\rightarrow$  **Roth's Theorem** ( $r_3(n) = o(n)$ ).

**Proof** (Theorem 9.2): Denote  $n = d^k$ . We shall think about  $x \in \{1, \dots, n\}$  in base  $d$ , hence it has  $k$  digits. Define

$$\begin{aligned} \bar{S}_{d,k,r} &= \left\{ (x_1, \dots, x_k) \mid \left( 0 \leq x_i < \frac{d}{2} \right) \wedge \sum_{i=1}^k x_i^2 = r^2 \right\} \\ S_{d,k,r} &= \left\{ \sum_{i=1}^k x_i d^{k-i} \mid (x_1, \dots, x_k) \in \bar{S}_{d,k,r} \right\} \end{aligned}$$

Note that the total size of the sets  $\bar{S}_{d,k,r}$  is  $(d/2)^k = n/2^k$  since the union of these sets has all vectors  $x$  with digits  $0 \leq x_i < \frac{d}{2}$ . Since  $|\bar{S}_{d,k,r}| = |S_{d,k,r}|$  we get that the total size of the sets  $S_{d,k,r}$  is also  $n/2^k$ . We will shortly show that each set  $S_{d,k,r}$  does not contain a 3-term arithmetic progression, but let us first show how to finish the proof given this fact. To this end we just need to pick  $d, k, r$ . For every  $d, k$  the value of  $r^2 = \sum_{i=1}^k x_i^2$  is an integer in  $[0, kd^2]$ . Since the total size of the sets  $S_{d,k,r}$  is  $n/2^k$  so for some  $r$  we have  $|S_{d,k,r}| \geq n/k2^k d^2$ . If we pick  $k = \sqrt{\log n}$  and  $d = 2\sqrt{\log n}$  we get that at least one of the sets  $S_{d,k,r}$  has no 3-term arithmetic progression and is of the required size.

We now prove the required claim. Assume  $x, y, z \in S_{d,k,r}$  and  $x + y = 2z$ . By rewriting the numbers in base  $d$  we get that  $\bar{x} + \bar{y} = 2\bar{z}$  that is, that  $x_i + y_i = 2z_i$  for every  $i$  (this follows from the fact that all the digits belong to  $[0, \frac{d}{2}]$  hence there is no carry).  $\bar{S}_{d,k,r}$  is actually a sphere, hence no three distinct points can be on the same line, hence  $\bar{x} = \bar{y} = \bar{z}$  and from here we get  $x = y = z$ . More detailed version:  $x_i + y_i = 2z_i$ , hence by Jensen's inequality  $x_i^2 + y_i^2 \geq 2z_i^2 \implies \sum x_i^2 + \sum y_i^2 \geq 2 \sum z_i^2$  but all of these are  $r^2$ , hence all the inequalities are actually equalities.  $\square$

**Fact 9.5.** *The best known bounds for  $r_3(n)$  are  $n/2^c \sqrt{\log n} \leq r_3(n) \leq O(n(\log \log n)^5 / \log n)$ .*

## 9.2 Lower Bound for the Triangle Removal Lemma

The lower-bound for  $r_3(n)$  gives us a lower-bound for the (6,3)-Problem. Now we shall seek a lower bound for the **Triangle Removal Lemma**.

**Theorem 9.6.** *For every  $\varepsilon > 0$  exists  $\delta = \delta(\varepsilon) > 0$  such as if  $G$  is a graph which at least  $\varepsilon n^2$  edges must be removed before it becomes triangle free, then  $G$  has at least  $\delta n^3$  triangles.*

Unfortunately, as we mentioned in an earlier lecture, the best known lower bound are of the form  $\delta(\varepsilon) \geq 1/\text{Tower}(1/\varepsilon)$ . It is thus natural to ask if a more civilized bound can be obtained. For example, is it true that  $\delta(\varepsilon) \geq \varepsilon^C$  for some absolute constant  $C$ ? The following theorem gives a negative answer.

**Theorem 9.7.** *For all small enough  $\varepsilon > 0$  we have  $\delta(\varepsilon) \leq \varepsilon^{c \log(\frac{1}{\varepsilon})}$ . Namely, there exist graphs for which  $\varepsilon n^2$  edges need to be removed to make them triangle-free but they contain “only”  $\varepsilon^{c \log(\frac{1}{\varepsilon})} n^3$  triangles.*

**Proof:** We take a graph  $H$  on  $m$  vertices from the lower bound of OEOT (see theorem 9.1), so it has  $\frac{m^2}{2^{\sqrt{\log m}}}$  triangles, where each edge is part of exactly one triangle. Now we take  $\frac{n}{m}$ -blowup of  $H$  to the graph  $G$ . We note:

1.  $G$  has  $n$  vertices.
2.  $G$  has at least  $\left(\frac{n}{m}\right)^3 \frac{m^2}{2^{\sqrt{\log m}}} = \frac{n^3}{m 2^{\sqrt{\log m}}}$  triangles, and from the construction of  $H$  we get that we have exactly this number of triangles.
3. In order to remove all the triangles in a blowup of an original triangle we need to remove at least  $\frac{n^2}{m^2}$  edges (all the edges between two sets of size  $\frac{n}{m}$ ), hence over all we need to remove at least  $\frac{n^2}{m^2} \frac{m^2}{2^{\sqrt{\log m}}} = \frac{n^2}{2^{\sqrt{\log m}}}$  edges (as the original triangles are edge-disjoint so we need to remove the blowups of all the original triangles).

So we need to remove  $\frac{n^2}{2^{\sqrt{\log m}}}$  edges in order to make  $G$  triangle-free and  $G$  has  $\leq \frac{n^3}{m}$  triangles. We shall require of

$$2^{\sqrt{\log m}} = \frac{1}{\varepsilon} \implies m = 2^{\log^2(\frac{1}{\varepsilon})} = \left(\frac{1}{\varepsilon}\right)^{\log(\frac{1}{\varepsilon})}$$

and we shall get exactly the bound we sought. □

## 9.3 Deducing Szemerédi’s Theorem from the Hypergraph Removal Lemma

We have seen that the Regularity Lemma implies the Triangle Removal Lemma, which in turn implies Roth’s Theorem. A famous theorem of Szemerédi extends Roth’s theorem by showing that for any fixed  $k$ , if  $S \subseteq [n]$  does not contain a  $k$ -term arithmetic progression then  $|S| = o(n)$ . It is thus natural to ask if one can come up with a graph theoretic proof of Szemerédi’s Theorem using the Regularity Lemma. As it turns out, this seems to be impossible. Therefore we shall attempt to do so using hypergraphs. To this end one first needs to prove a removal lemma for hypergraphs. Let us state such a removal lemma for  $k$ -uniform hypergraphs and for the graph  $K_{k+1}^k$  that is the complete  $k$ -uniform hypergraph on  $k+1$  vertices.

**Theorem 9.8 (Hypergraph Removal Lemma).** *If one should removal at least  $\varepsilon n^k$  edges from a  $k$ -uniform hypergraph  $H$  in order to make it  $K_{k+1}^k$ -free then  $H$  contains at least  $\delta(\varepsilon, k) \cdot n^{k+1}$  copies of  $K_{k+1}^k$ .*

Our goal now is to prove the following:

**Theorem 9.9.** *The Hypergraph removal lemma implies Szemerédi’s Theorem.*

Before we prove the theorem, let us comment on the proof of the Hypergraph Removal Lemma. Recall that our plan for proving the Triangle Removal Lemma consisted of the following steps:

1. We defined the concept of “graph regularity”.
2. We proved the Regularity Lemma, which states that every graph has an equipartition satisfying the notion of “graph regularity” we came up with.
3. We proved a counting lemma, that is, that 3 vertex sets that satisfy our notion of “graph regularity” contain many copies of  $K_3$ .

When one tries to extend this plan and prove the Hypergraph Removal Lemma, one first needs to come up with a concept of hypergraph regularity. The first choice that comes to mind is the following:

**Definition 9.10** (Regularity condition for 3-uniform hypergraph).  $A, B, C$  are  $\varepsilon$ -regular if for every  $A' \subseteq A, B' \subseteq B, C' \subseteq C$  such that  $|A'| \geq \varepsilon|A|, |B'| \geq \varepsilon|B|, |C'| \geq \varepsilon|C|$  we have

$$|d(A', B', C') - d(A, B, C)| \leq \varepsilon$$

where  $d(A, B, C) = \frac{e(A, B, C)}{|A||B||C|}$ .

It is easy to adapt the proof of the graph regularity lemma and prove the following.

**Theorem 9.11.** *Every 3-uniform hypergraph has an  $\varepsilon$ -regular partition of size at most  $T\left(\frac{1}{\varepsilon^5}\right)$ .*

The problem with this regularity lemma is that the concept of regularity it uses is not strong enough to satisfy the third condition of our plan. That is, it is not true that if 4 vertex sets  $A, B, C, D$  are such that each 3 of them are dense and regular then they contain many copies of  $K_4^3$ . In fact, they might not contain any copy of  $K_4^3$ ! Here is how to construct such an example. Take 4 vertex sets of size  $n$  and for every pair of vertices (in two different vertex sets) pick a direction randomly, uniformly and independently. We take  $x, y, z$  to be an edge iff they form a directed cycle. Hence  $x, y, z$  is an edge with probability  $\frac{1}{4}$  and it is easy to verify that each of the 4 triples of vertex sets are  $o(1)$ -regular (as the density in every triplet of vertex-sets of size  $\Omega(n)$  will be  $\frac{1}{4} \pm o(1)$  with high probability). But it is also easy to see that this 3-uniform hypergraph has no copy of  $K_4^3$ -free.

One can also come up with very strong notions of regularity that satisfy condition 3, but do not satisfy condition 2, that is, that it is not the case that all hypergraphs have an equipartition satisfying these conditions. We finally note that it is possible to define a “right” notion of regularity and thus prove the removal lemma, but it is rather complicated and hence we shall omit it.

We now return to our original task of deducing Szemerédi’s Theorem from the hypergraph removal lemma. We will do so for 4-term arithmetic progressions. The proof for general  $k$  is identical and is thus left as an exercise.

**Proof** (Theorem 9.9): Given  $S \subset \{1, \dots, n\}$  we define a 3-uniform hypergraph  $H$  over 4 vertex-sets  $X_1, X_2, X_3, X_4$ , each of size  $O(n)$ . For every  $x_1, x_2 \in [n]$  and  $s \in S$  we place a copy of  $K_4^3$  whose vertices are  $x_1 \in X_1, x_2 \in X_2, a_1x_1 + a_2x_2 + s \in X_3$ , and  $b_1x_1 + b_2x_2 + s \in X_4$ , where  $a_1, a_2, b_1, b_2$  are absolute constants (i.e. independent of  $n$  and  $S$ ) that will be determined later. For ease of reference, we label the 4 edges of this copy with the integer  $s$ . Note that we have placed in  $H$  exactly  $n^2|S|$  copies of  $K_4^3$  (one for every  $x_1, x_2 \in [n]$  and  $s \in S$ ). We will shortly prove that:

**Claim 9.12.** *If  $a_1 \neq b_1$  and  $a_2 \neq b_2$  then these  $|S|n^2$  copies of  $K_4^3$  are edge disjoint.*

Since  $|S| \geq \varepsilon n$  this implies that one should remove from  $H$  at least  $\varepsilon n^3$  edges in order to make it  $K_4^3$ -free, implying that  $H$  is  $\varepsilon/c$ -far from being  $K_4^3$ -free (the  $c$  comes from the fact that  $H$  has  $O(n)$  vertices). Hence by the [Hypergraph Removal Lemma](#) we get that  $H$  contains  $\delta(\varepsilon)n^4$  copies of  $K_4^3$ .

The key part of the proof will be the following:

**Claim 9.13.** *It is possible to choose  $a_1 \neq b_1$  and  $a_2 \neq b_2$  so that if we take the labels  $s_1, s_2, s_3, s_4$  written on the 4 edges of some copy of  $K_4^3$  in  $H$ , then  $s_1, s_2, s_3, s_4$  form a 4-term arithmetic progression.*

Therefore, as soon as  $n$  is large enough so that  $\delta(\varepsilon)n^4 \geq \varepsilon n^3 \geq n^2|S|$  we get a copy of  $K_4^3$  whose edges are not marked by the same  $s$ , which gives us a non-trivial 4-term arithmetic progression. So given the above claims the proof is complete.  $\square$

We now finish the proof of Theorem 9.9 by proving the two claims we stated above.

**Proof (Claim 9.12):** It is sufficient to show that given an edge it is possible to determine  $x_1, x_2, s$ . If the edge contains a vertex from  $X_1$  and  $X_2$  this is obvious. If the edge contains vertices  $x_1 \in X_1, x_3 \in X_3$  and  $x_4 \in X_4$  we can recover  $x_2$  and  $s$  by solving the linear equations (observe that we “know”  $x_1, x_3$  and  $x_4$ )

$$\begin{cases} a_1x_1 + a_2x_2 + s = x_3 \\ b_1x_1 + b_2x_2 + s = x_4 \end{cases}$$

Therefore if  $a_2 \neq b_2$  it is possible to uniquely determine  $x_1, x_2$  and  $s$ . If the edge has vertices in  $X_2, X_3$  and  $X_4$  then we need the condition  $a_1 \neq b_1$ . So the copies of  $K_4^3$  are edge-disjoint if  $a_1 \neq b_1$  and  $a_2 \neq b_2$ .  $\square$

**Proof (Claim 9.13):** Note that  $z_1 \leq z_2 \leq z_3 \leq z_4$  form a 4-term arithmetic progression iff they satisfy the linear equations  $z_1 + z_3 = 2z_2$  and  $z_2 + z_4 = 2z_3$ . By performing elementary row operations we get that  $z_1 \leq z_2 \leq z_3 \leq z_4$  form a 4-term arithmetic progression iff they satisfy the linear equations

$$z_1 - 2z_2 = -z_3 \quad \text{and} \quad 2z_1 - 3z_2 = -z_4$$

Assume  $x_1 \in X_1, x_2 \in X_2, x_3 \in X_3, x_4 \in X_4$  span a copy of  $K_4^3$ . Suppose the edge  $(x_1, x_2, x_3)$  is marked by  $s_1$  and the edge  $(x_1, x_2, x_4)$  is marked by  $s_2$ . Then we know that

$$a_1x_1 + a_2x_2 + s_1 = x_3 \quad \text{and} \quad b_1x_1 + b_2x_2 + s_2 = x_4$$

Suppose the edge  $(x_1, x_3, x_4)$  is marked by  $s_3$ . This edge was originally defined using  $x_1, s_3$  and some integer  $x'_2$ . Hence,

$$a_1x_1 + a_2x'_2 + s_3 = x_3 \quad \text{and} \quad b_1x_1 + b_2x'_2 + s_3 = x_4$$

Combining these two ways of expressing  $x_3$  and  $x_4$  we deduce that

$$a_2(x_2 - x'_2) = s_3 - s_1 \quad \text{and} \quad b_2(x_2 - x'_2) = s_3 - s_2$$

Multiplying the first equation by  $-b_2$ , the second by  $a_2$  and then summing them we see that  $s_1, s_2, s_3$  satisfy the linear equation  $b_2s_1 - a_2s_2 = (b_2 - a_2)s_3$ . Recall that our goal is to show that  $s_1, s_2, s_3$  satisfy the linear equation  $z_1 - 2z_2 = -z_3$  so setting  $b_2 = 1$  and  $a_2 = 2$  we get the required equation. Note that this choice satisfies the requirement of the previous claim that  $b_2 \neq a_2$ . By an identical argument, this time using the edge  $(x_2, x_3, x_4)$  instead of the edge  $(x_1, x_3, x_4)$ , we get that  $b_1s_1 - a_1s_2 = (b_1 - a_1)s_4$ . Our goal is to show that  $s_1, s_2, s_4$  satisfy the linear equation  $2z_1 - 3z_2 = -z_3$  so setting  $b_1 = 2$  and  $a_1 = 3$  we get the required equation (note that again  $a_1 \neq b_1$ ). Since  $s_1, s_2, s_3, s_4$  satisfy the two linear equations that define a 4-term arithmetic progression, the proof is complete.  $\square$



## 10 Tenth Lecture

### 10.1 The Ramsey-Turán Problem

We know that a graph can be triangle-free and have  $n^2/4$  edges, but such a graph contains large independent sets. We thus ask the following: can a graph  $G$  be triangle-free, have only independent sets of size  $o(n)$  and still have  $cn^2$  edges, or does  $G$  necessarily have  $o(n^2)$  edges? It is easy to see that the answer is the latter. Indeed, if  $\alpha(G) \leq \delta n$  then  $\Delta(G) \leq \delta(n)$  (as an edge in the neighborhood of a vertex implies a triangle hence all the neighborhood of each vertex must be independent set). So we get that  $2m \leq n\delta n$ , namely if  $G$  is  $K_3$ -free and  $\alpha(G) = o(n)$  then  $m = o(n^2)$ .

But what if instead of  $K_3$ -free graphs we asked about  $K_5$ -free graphs? The answer is given by the next two theorems:

**Theorem 10.1.** *There exists a  $K_5$ -free graph, that contains only independent sets of size  $o(n)$  and contains more than  $\frac{n^2}{4}$  edges.*

**Theorem 10.2.** *Every graph with  $\geq (\frac{1}{4} + \varepsilon)n^2$  edges contains either a  $K_5$  or independent set of size  $\delta(\varepsilon) \cdot n$ .*

Note that Turán's Theorem guarantees a copy of  $K_3$  when  $m > \frac{1}{4}n^2$ . The above theorem states that if we forbid independent sets of size  $\Omega(n)$  then  $m > (\frac{1}{4} + \varepsilon)n^2$  edges give us a  $K_5$ .

**Proof (Theorem 10.1):** We start with  $K_{n/2, n/2}$  with bipartition into two sets  $A, B$ . This already gives us  $\frac{n^2}{4}$  edges. We now wish to find a graph  $H$  on  $n/2$  vertices that will have the following two properties

1. All independent sets in  $H$  will be of size  $o(n)$ .
2.  $H$  will be triangle-free.

It is clear that if we put one such  $H$  in  $A$  and one in  $B$ , the final graph will satisfy all the required properties.

It is not hard to show that such an  $H$  exists. In fact it is known that there exists a  $K_3$ -free graph where all independent sets are of size  $O(\sqrt{n \log n})$ . Since we do not need such a strong result (which is rather hard to prove), we will be content with showing that there is a  $K_3$ -free graph where all independent sets are of size  $n^{2/3}$ . To this end we take  $G(n, p)$ . Since we want the graph to be triangle-free we would like to have

$$\mathbb{E}[\#K_3] = p^3 n^3 < \frac{n}{4},$$

so we set  $p = \frac{1}{2n^{2/3}}$ . The probability to have an independent set of size  $n^c$  is at most

$$\binom{n}{n^c} \left(1 - \frac{1}{2n^{2/3}}\right)^{n^{2c}} \leq 2^{n^c \log n} e^{-\frac{n^{2c}}{2n^{2/3}}},$$

so we would like that  $n^{2/3}n^c \leq n^{2c}$ . Hence picking any  $c > 2/3$ , gives that with high probability the graph has no independent set of size  $n^c$  and that with probability at least  $1/2$  will have at most  $n/2$  copies of  $K_3$  (we use Markov's Inequality here). So with positive probability we get a graph with no independent set of size  $n^c$  and with at most  $n/2$  copies of  $K_3$ . Removing one vertex from each of these  $K_3$ 's gives the required graph.  $\square$

**Proof (Theorem 10.2):** Given  $G$  with  $\geq (\frac{1}{4} + \varepsilon)n^2$  edges we need to find in it either a  $K_5$  or an independent set of size  $\delta(\varepsilon)n$ . We now apply Claim 2.3, which tells us that  $G$  contains a subgraph  $G'$  on  $n' \geq \varepsilon n$  vertices with  $\delta(G') \geq (\frac{1}{2} + \frac{\varepsilon}{2})n'$ . So it is enough to find an independent set of size  $\delta(\varepsilon)n'$  in  $G'$

(or a  $K_5$ ). For the sake of simplicity we assume that  $G$  itself satisfies the condition  $\delta(G) \geq (\frac{1}{2} + \varepsilon)n$ . For every edge  $(x, y)$ , the vertices  $x, y$  have at least  $\varepsilon n$  common neighbors. If those vertices do not contain any edge then we get an independent set of size  $\varepsilon n$ , and we are done. So suppose they contain an edge  $z, w$ , which gives us a copy of  $K_4$  on the vertices  $x, y, z, w$ . As  $\delta(G) \geq (\frac{1}{2} + \varepsilon)n$  the number of edges between  $x, y, z, w$  and the rest of the vertices is at least  $4(\frac{n}{2} + n\varepsilon) = 2n + 4\varepsilon n$ . If there is a vertex connected to all the 4 vertices  $x, y, z, w$ , it is a copy of  $K_5$  and we are done. Otherwise, denote by  $U$  those vertices connected to 3 of the vertices  $x, y, z, w$ . The number of edges  $e$  between  $\{x, y, z, w\}$  to the rest of the vertices thus satisfies

$$2n + 4\varepsilon n \leq e \leq 3|U| + 2(n - |U|) = |U| + 2n$$

implying that  $|U| \geq 4\varepsilon n$ . At least  $1/4$  of the vertices in  $U$  must be connected to the same triplet of vertices from  $x, y, z, w$ . This set of vertices, call it  $U'$ , has size at least  $\varepsilon n$ . Finally, if  $U'$  contains an edge we get a  $K_5$ , and if it has no edge we get an independent set of size  $\varepsilon n$ .  $\square$

Having addressed the problem for  $K_3$  and  $K_5$  we now address the case of  $K_4$ . It turns out that answering such questions for even  $K_r$ 's is more difficult than for odd  $K_r$ 's. For  $K_4$  the answer turned out to be around  $n^2/8$  edges.

**Theorem 10.3.** *If  $G$  contains  $(\frac{1}{8} + \varepsilon)n^2$  edges then  $G$  either contains  $K_4$  or an independent set of size  $\delta(\varepsilon)n$  (this time  $\delta(\varepsilon)$  is something like  $1/\text{Tower}(1/\varepsilon)$ ).*

**Theorem 10.4.** *For every  $\varepsilon > 0$ , and large enough  $n > n_0(\varepsilon)$ , there exists a  $K_4$ -free graph with  $(\frac{1}{8} - \varepsilon)n^2$  edges that contains only independent set of size  $o(n)$ .*

We will not prove the above lower bound, and instead address Theorem 10.3. We will need a few preliminary claims.

**Claim 10.5.** *Assume  $H$  is a graph on two vertex-sets  $A, B$  of size  $m$  each, and  $d(A, B) \geq \frac{1}{2} + \varepsilon$ . Then  $H$  contains either  $K_4$  or an independent set of size  $\geq \varepsilon m$ .*

**Proof:** It is easy to see that  $A$  has at least  $\frac{\varepsilon m}{2}$  vertices whose degree is at least  $(\frac{1}{2} + \frac{\varepsilon}{2})m$  as

$$\frac{\varepsilon m}{2} \cdot m + \left(1 - \frac{\varepsilon}{2}\right)m \left(\frac{1}{2} + \frac{\varepsilon}{2}m\right)m < m^2 \left(\frac{1}{2} + \varepsilon\right)$$

If those vertices do not span an edge then we're done (we found an independent set of size  $\frac{\varepsilon m}{2}$ ). Otherwise, assume  $x, y$  are connected then they have  $\frac{\varepsilon m}{2}$  common neighbors in  $B$ . Then those vertices are either an independent set of size  $\frac{\varepsilon m}{2}$  or a copy of  $K_4$ .  $\square$

**Claim 10.6.** *If  $A, B, C$  are vertex-sets of size  $m$  such as  $(A, B)$ ,  $(A, C)$  and  $(B, C)$  are  $\varepsilon$ -regular with densities  $\geq 3\varepsilon$ , then they contain either a  $K_4$  or an independent set of size  $\varepsilon^2 m$ .*

**Proof:** Given  $(A, B)$  is  $\varepsilon$ -regular with density  $3\varepsilon$  all the vertices of  $A$ , apart of  $\varepsilon m$  of them, have a minimal degree  $\geq 2\varepsilon m$  in  $B$ . The same goes for  $C$ . Especially there exists a vertex (actually almost all of them) whose degree in  $B$  and  $C$  is at least  $2\varepsilon m$ . All of  $B$ 's vertices, apart for possibly  $\varepsilon m$  of them, have a minimal degree  $2\varepsilon|C'|$  in  $C'$ . Particularly,  $B$  has a vertex  $b$  whose degree in  $C'$  is at least  $4\varepsilon^2 m = 2\varepsilon|C'|$ . We call those neighbors  $C''$ . We know that  $a$  is connected to  $b$ , and both of them are connected to very vertex in  $C''$ . so we either have a  $K_4$  or an independent set of size  $\varepsilon^2 m$ .  $\square$

**Claim 10.7.** *If  $G$  has  $(\frac{1}{8} + \varepsilon)n^2$  edges then it either contains  $K_4$  or independent set of size  $\delta n$  where  $\delta = \frac{\varepsilon^2}{T(\varepsilon/6)}$ .*

**Proof:** Given  $G$  we apply the Regularity Lemma with  $\varepsilon/6$ . We remove from  $G$  the following edges:

1. Edges between non- $\varepsilon/6$ -regular pairs ( $\leq \varepsilon n^2/6$  edges).
2. Edges within the sets of the partition ( $\leq \varepsilon n^2/12$  edges).
3. Edges between pairs with density  $\leq \frac{\varepsilon}{2}$  ( $\leq \varepsilon n^2/4$  edges).

All in all, we removed at most  $\frac{\varepsilon n^2}{2}$  edges. The remaining graph has at least  $(\frac{1}{8} + \frac{\varepsilon}{2})n^2$  edges. All the pairs in the equipartition must have densities  $\leq (\frac{1}{2} + \frac{\varepsilon}{2})$  else by Claim 10.5 the remaining graph (and  $G$  as well) contains  $K_4$  or an independent set of size  $\geq \frac{\varepsilon}{2} \frac{n}{T}$ .  $\square$

How many pairs  $(V_i, V_j)$  with density  $> 0$  exist? If the partition has  $k$  sets then no more than  $\frac{k^2}{4}$ , otherwise Mantel's Theorem guarantees 3 sets such that the density between any two of them is  $\geq \varepsilon/2$  and they are  $\varepsilon/6$ -regular. Then Claim 10.6 proves the existence of either  $K_4$  or an independent set of size  $\frac{\varepsilon^2}{9} \frac{n}{T}$ . So the number of edges in  $G$  is

$$\frac{\varepsilon n^2}{2} + \frac{k^2}{4} \left[ \frac{n^2}{k^2} \left( \frac{1}{2} + \frac{\varepsilon}{2} \right) \right] = \frac{\varepsilon n^2}{2} + \frac{n^2}{8} + \frac{\varepsilon n^2}{8} < \left( \frac{1}{8} + \varepsilon \right) n^2$$

which is a contradiction. Hence, somewhere along the way we had to find either a  $K_4$  or an independent set of size  $\delta(\varepsilon) \cdot n$ .

## 10.2 Quasi-Random Graphs

Consider the following proposed definition:

**Definition.** We shall say a graph is  $(\varepsilon, p)$ -quasi-random if every set  $S$  contains  $\binom{|S|}{2}(p \pm \varepsilon)$  edges (where  $x \pm \varepsilon$  is a number between  $x - \varepsilon$  and  $x + \varepsilon$ ).

The above definition is flawed, as by Ramsey's Theorem it will not hold for sets of size  $o(\log n)$ . So we consider the following three alternative definitions:

**Definition 10.8.** We say  $G$  fulfills  $P_1$  if every set  $S$  contains  $p \binom{|S|}{2} \pm \varepsilon n^2 = p \frac{|S|^2}{2} \pm \varepsilon n^2$  edges.

**Definition 10.9.** We say  $G$  fulfills  $P_2$  if every set  $S$  of size  $\geq \varepsilon n$  contains  $\frac{|S|^2}{2}(p \pm \varepsilon)$  edges.

**Definition 10.10.** We say  $G$  fulfills  $P_3$  if it fulfills  $P_2$  for sets of size  $\frac{n}{2}$ .

We say that  $P_1 \implies P_2$  if for every  $\varepsilon > 0$  there exists  $\delta > 0$  such that if  $G$  fulfills  $P_1$  with an error of  $\delta$  then it fulfills  $P_2$  with an error of  $\varepsilon$ . We now wish to prove the following equivalence.

**Theorem 10.11.**  $P_1 \iff P_2 \iff P_3$

We break the proof into several claims.

**Claim 10.12.**  $P_1 \implies P_2$

**Proof:** We note that  $P_2$  is fulfilled iff it holds for sets of size  $\varepsilon n$  (by averaging argument it will hold for any bigger set). In this case we require that

$$e(S) = p \frac{|S|^2}{2} \pm \varepsilon |S|^2 = p \frac{|S|^2}{2} \pm \varepsilon^3 n^2$$

so it is enough that  $P_1$  will hold with  $\delta = \varepsilon^3$ .  $\square$

**Claim 10.13.**  $P_2 \implies P_1$

**Proof:** If  $G$  fulfills  $P_2$  with  $\delta$  then  $P_1$  holds for every  $S$  of size  $\geq \delta n$  since  $\frac{\delta|S|^2}{2} \leq \delta n^2$ . So if  $\delta \leq \varepsilon$  then  $P_1$  holds for all subsets of size at least  $\delta n$ . For sets of size at most  $\delta n$  we would like that  $P_1$  will be fulfilled trivially. Namely, that

$$\frac{p|S|^2}{2} - \varepsilon n^2 \leq 0 \leq e(S) \leq \frac{|S|^2}{2} \leq \frac{p|S|^2}{2} + \varepsilon n^2$$

So to get the left hand side we need to have  $\frac{p|S|^2}{2} - \varepsilon n^2 \leq 0$  so we also ask than  $|S| \leq \delta n \leq \sqrt{\frac{2\varepsilon}{p}}n$ , that is, that  $\delta \leq \sqrt{\frac{2\varepsilon}{p}}$ . For the other side we would need to have  $\frac{p|S|^2}{2} + \varepsilon n^2 \geq \frac{|S|^2}{2}$  so we also ask that  $|S| \leq \delta n \leq \sqrt{\frac{2\varepsilon}{1-p}}n$ , that is, that  $\delta \leq \sqrt{\frac{2\varepsilon}{1-p}}$ . So overall it is enough to take  $\delta = \min\left(\sqrt{\frac{2\varepsilon}{p}}, \sqrt{\frac{2\varepsilon}{1-p}}, \varepsilon\right)$ .  $\square$

**Claim 10.14.**  $P_2 \iff P_3$

**Proof:** It is obvious that  $P_2 \implies P_3$ . We will prove the other direction by showing that if  $G$  does not fulfill  $P_2$  with error  $\varepsilon$  then it does not fulfill  $P_3$  with some other small error  $\delta(\varepsilon)$ . Recall that if  $G$  fails to satisfy  $P_2$  with error  $\varepsilon$  then there exists a set  $A$  of size  $\varepsilon n$  whose density is not  $p \pm \varepsilon$ . We assume that it is  $p + \varepsilon$  (the case  $p - \varepsilon$  is similar). We will show that in this case  $G$  has a subset of size  $n/2$  with density  $p + \Omega(\varepsilon^4)$ . This means that if  $G$  satisfies  $P_3$  with error  $\delta(\varepsilon) = \varepsilon^4$  then it must satisfy  $P_2$  with error  $\varepsilon$ .

Denote  $B = V \setminus A$ . If the density of  $B$  is not  $(p \pm \delta)$  then we are done as by averaging, it must contain a subset of size  $\frac{n}{2}$  whose density is also not in  $(p \pm \delta)$ . Since  $G$  has  $pn^2/2$  edges we get that the density of the bipartite graph connecting  $A$  and  $B$  is given by

$$d(A, B) = \frac{\frac{pn^2}{2} - e(A) - e(B)}{\varepsilon n(1 - \varepsilon)n} = \frac{\frac{pn^2}{2} - \frac{\varepsilon^2 n^2}{2}(p + \varepsilon) - \frac{(1 - \varepsilon)^2 n^2}{2}(p \pm \delta)}{\varepsilon n(1 - \varepsilon)n}$$

We pick a random subset  $S \subseteq B$  of size  $(1/2 - \varepsilon)n$  and we examine the expected number of edges contained in  $A \cup S$ . This is a set of size  $\frac{n}{2}$  hence it is sufficient to show that the expected value is  $\geq (p + \Omega(\varepsilon^3))\frac{n^2}{8} > (p + \delta)\frac{n^2}{8}$  which will imply that  $P_3$  is not fulfilled with error  $\delta = \varepsilon^4$ . The expected number of edges in  $A$  is just the number of edges in  $A$  which is  $\frac{\varepsilon^2 n^2 p}{2} + \frac{\varepsilon^3 n^2}{2}$ . The expected number of edges in  $S$  is

$$\mathbb{E}[e(S)] = \frac{(1/2 - \varepsilon)^2 n^2}{2} (p \pm \delta),$$

and the expected number of edges between  $A$  and  $S$  is

$$\mathbb{E}[e(A, S)] = \varepsilon n \left(\frac{1}{2} - \varepsilon\right) n \frac{\left[\frac{pn^2}{2} - \frac{\varepsilon^2 n^2}{2}(p + \varepsilon) - \frac{(1 - \varepsilon)^2 n^2}{2}(p \pm \delta)\right]}{\varepsilon n(1 - \varepsilon)n}.$$

So the total expected number of edges in  $A \cup S$  is

$$\frac{\varepsilon^2 n^2 p}{2} + \frac{\varepsilon^3 n^2}{2} + \frac{(1/2 - \varepsilon)^2 n^2}{2} (p \pm \delta) + \varepsilon n \left(\frac{1}{2} - \varepsilon\right) n \frac{\left[\frac{pn^2}{2} - \frac{\varepsilon^2 n^2}{2}(p + \varepsilon) - \frac{(1 - \varepsilon)^2 n^2}{2}(p \pm \delta)\right]}{\varepsilon n(1 - \varepsilon)n}.$$

We want to show that this expression is at least  $(p + \Omega(\varepsilon^3))\frac{n^2}{8}$ , so dividing both expression by  $n^2/2$  we need to show that

$$\varepsilon^2 p + \varepsilon^3 + (1/2 - \varepsilon)^2 (p \pm \delta) + \frac{(\frac{1}{2} - \varepsilon)}{(1 - \varepsilon)} [p - \varepsilon^2 p - \varepsilon^3 - (1 - \varepsilon)^2 (p \pm \delta)] \geq \frac{p}{4} + \Omega(\varepsilon^3).$$

As we assumed  $\delta = \varepsilon^4$  all the terms involving  $\delta$  are  $O(\varepsilon^4)$  so we can disregard them as all the rest are of order at least  $\varepsilon^3$ . So we are down to proving that

$$\varepsilon^2 p + \varepsilon^3 + (1/2 - \varepsilon)^2 p + \frac{(\frac{1}{2} - \varepsilon)}{(1 - \varepsilon)} [p - \varepsilon^2 p - \varepsilon^3 - (1 - \varepsilon)^2 p] \geq \frac{p}{4} + \Omega(\varepsilon^3).$$

It is now easy to check that in the left hand side all the terms involving  $\varepsilon$  and  $\varepsilon^2$  cancel and we are left with  $p/4 + \varepsilon^3 - \varepsilon^3 \frac{(\frac{1}{2} - \varepsilon)}{(1 - \varepsilon)} > p/4 + \varepsilon^3/2$ , and the proof is complete.  $\square$

## 11 Eleventh Lecture

Throughout this lecture we use  $d(x, y)$  to denote the co-degree of  $x, y$ , that is, the number of vertices that are connected to both  $x$  and  $y$ .

### 11.1 The Chung-Graham-Wilson Theorem

**Definition 11.1.** A graph  $G$  is said to satisfy property  $P_4$  if the number of copies of  $C_4$  (i.e. the 4-cycle) in  $G$  is at most  $(p^4 + \varepsilon)n^4$ . Here, the number of  $C_4$  is the number of 4-tuples  $(x, y, z, w)$  so that  $(x, z), (x, w), (y, z), (y, w)$  are all edges in the graph.

The advantage of this notion of quasi-randomness, compared to the previous properties we considered last lectures, is that it can easily verified in polynomial time, compared to properties  $P_1, P_2, P_3$  which require checking exponentially many conditions.

**Theorem 11.2** (Chung-Graham-Wilson Theorem).  $P_3 \iff P_4$  [*Chung et al., 1989*]

The implication  $P_3 \implies P_4$  is the easy part of this theorem.

**Exercise 11.3.** Show that  $P_3 \implies P_4$ .

We shall thus focus on proving that  $P_4 \implies P_3$ . As a warm-up, we start with the following claims.

**Claim 11.4.** Every graph with  $pn^2/2$  edges contains at least  $p^2n^3$  copies of  $K_{1,2}$ , where a copy of  $K_{1,2}$  is a 3-tuple  $(x, y, z)$  such that both  $(x, y)$  and  $(x, z)$  are edges in the graph.

**Proof:**  $\#K_{1,2} = \sum d^2(x) \geq n \left[ \frac{\sum d(x)}{n} \right]^2 = p^2n^3.$   $\square$

**Claim 11.5.** Every  $n \times n$  bipartite graph on vertex sets  $A, B$ , that has  $pn^2$  edges contains at least  $p^2n^3$  copies of  $K_{1,2}$ , with one vertex in  $A$  and 2 in  $B$ .

**Proof:** Identical to the previous proof.  $\square$

**Claim 11.6.** If  $G$  contains  $K$  copies of  $K_{1,2}$  then it contains at least  $(n/K)^2$  copies of  $C_4$ .

**Proof:**  $\#C_4 = \sum_{x,y} d^2(x, y) \geq n^2 \left[ \frac{\sum d(x, y)}{n^2} \right]^2 = n^2 \left( \frac{K}{n^2} \right)^2 = \left( \frac{K}{n} \right)^2.$   $\square$

**Claim 11.7.** *Every graph with  $pn^2/2$  edges contains at least  $(pn)^4$  copies of  $C_4$ .*

**Proof:** Immediately from Claims 11.4 and 11.6.  $\square$

As a warm up towards proving Theorem 11.2, let us prove that  $P_4$  forces the graph to be nearly regular.

**Claim 11.8.** *If  $G$  satisfies  $P_4$  then all but  $\delta n$  vertices have degree  $(p \pm \delta)n$ .*

**Proof:** Assume the contrary and recall that if  $\frac{1}{n} \sum_{i=1}^n x_i = d$ , then setting  $x_i = d + \varepsilon_i$  we have

$$\sum_{i=1}^n x_i^2 = d^2 n + \sum_{i=1}^n \varepsilon_i^2. \quad (11.1)$$

Hence, if we denote  $d(x) = pn + \varepsilon_x n$  then we get

$$\#K_{1,2} = \sum_x d^2(x) = \sum_x (p + \varepsilon_x)^2 n^2 = p^2 n^3 + \sum_x \varepsilon_x^2 n^2 \geq p^2 n^3 + \delta^3 n^3.$$

We can now infer from Claim 11.6 that

$$\#C_4 \geq \left( \frac{p^2 n^3 + \delta^3 n^3}{n} \right)^2 \geq (p^4 + 2p^2 \delta^3) n^4,$$

which is a contradiction to the  $P_4$  property of  $G$ , upon taking  $\varepsilon = 2p^2 \delta^3$ .  $\square$

Our plan for proving Theorem 11.2 is similar to the one we employed in the last proof, that is, we show that if  $G$  fails to satisfy  $P_3$  then it has more than  $(pn)^4$  copies of  $C_4$ . Let us first observe that the above claim does not prove that  $P_4 \implies P_3$ . Indeed, while one expects a quasi-random graph to be nearly regular, it is easy to see that being regular is not equivalent to properties  $P_1, P_2, P_3$ . For example  $K_{n/2, n/2}$  is  $n/2$ -regular but fails to satisfy these three properties. In other words, a graph can fail to satisfy  $P_3$  but still have the correct number of  $K_{1,2}$ . In a more technical level, this means that we will have to rework the proof of Claim 11.6 rather than rework the proof of Claim 11.4 as we have done in the above claim.

**Proof (Chung-Graham-Wilson Theorem):** We show that if  $G$  fails to satisfy  $P_3$  then it fails to satisfy  $P_4$  as well. Assume  $G$  has  $pn^2/2$  edges and contains a vertex set  $A$  of size  $n/2$  whose density is not in  $p \pm \varepsilon$ . We need to show that  $G$  contains at least  $(p^4 + \delta)n^4$  copies of  $C_4$  for some  $\delta = \delta(\varepsilon) > 0$ . Set  $B = V \setminus A$  and suppose  $e(A) = an^2/8$ ,  $e(B) = bn^2/8$  and  $e(A, B) = cn^2/4$ . Then we have  $a + 2c + b = 4p$ . Since we assume that  $a \notin p \pm \varepsilon$  we get from (11.1) that

$$a^2 + 2c^2 + b^2 \geq 4p^2 + \varepsilon^2. \quad (11.2)$$

Claim 11.6 tells us that  $G$  has at least  $p^2 n^3$  copies of  $K_{1,2}$ . Our goal is thus to improve the argument of Claim 11.6 by finding  $\sim n^2$  pairs of vertices whose co-degree deviates from its expectation (which is  $p^2 n$ ) by  $\sim n$ .

To this end we will focus on the pairs of vertices  $x, y$  so that either  $x, y \in A$  or  $x, y \in B$ . We first calculate  $\sum_{x, y \in A} d(x, y)$ , which equals the number of copies of  $K_{1,2}$  (namely, triples  $(x, y, z)$  with both  $(x, z)$  and  $(y, z)$  edges of the graph), with  $x, y \in A$ . By Claim 11.6, applied to the graph induced by  $A$ , we get that the number of copies of  $K_{1,2}$  with all three vertices in  $A$  is at least  $a^2 n^3/8$ . By Claim 11.5, applied to the bipartite graph connecting  $A$  to  $B$ , we get that the number of copies of  $K_{1,2}$  with two vertices in  $A$  and one in  $B$  is at least  $c^2 n^3/8$ . So we get that  $\sum_{x, y \in A} d(x, y) \geq (a^2 + c^2) n^3/8$ . An identical argument implies that  $\sum_{x, y \in A} d(x, y) \geq (b^2 + c^2) n^3/8$ . These two facts, together with (11.2), imply that

$$\sum_{(x,y) \in A^2 \cup B^2} d(x,y) \geq (a^2 + b^2 + 2c^2)n^3/8 \geq p^2n^3/2 + \varepsilon^2n^3/8. \quad (11.3)$$

Denote  $d(x,y) = p^2n + \delta_{xy}$ . Then (11.3) is equivalent to saying that

$$\sum_{(x,y) \in A^2 \cup B^2} \delta(x,y) \geq (n^2/2) \cdot \varepsilon^2n/4. \quad (11.4)$$

Hence, we get

$$\begin{aligned} \#C_4 &= \sum_{x,y} d^2(x,y) = \sum_{x,y} (p^2n + \delta_{xy})^2 = p^4n^4 + \sum_{x,y} \delta_{x,y}^2 \geq p^4n^4 + \sum_{(x,y) \in A^2 \cup B^2} \delta_{xy}^2 \\ &\stackrel{(11.1)}{\geq} p^4n^4 + \frac{n^2}{2} \left( \frac{\sum \delta_{xy}}{n^2/2} \right)^2 \stackrel{(11.4)}{\geq} p^4n^4 + \frac{n^2}{2} (\varepsilon^2n/4)^2 = p^4n^4 + \varepsilon^4n^4/32. \end{aligned}$$

So we can finally conclude the proof by setting  $\delta = \varepsilon^4/32$ .  $\square$

## 11.2 An Algorithmic Version of the Regularity Lemma

Inspecting the proof of the regularity lemma, one sees that the proof actually supplies an algorithm that given  $\varepsilon$  a graph  $G$  constructs an  $\varepsilon$ -regular partition of  $G$ . However, there is one part which is not “explicit” in that it is not clear how to implement it in polynomial time. The part we are referring to is where given a bipartite graph on vertex sets  $A, B$  that is not  $\varepsilon$ -regular, we need to find two vertex sets  $A' \subseteq A$  and  $B' \subseteq B$  of sizes  $|A'| \geq \varepsilon|A|$  and  $|B'| \geq \varepsilon|B|$  satisfying  $d(A', B') \notin d(A, B) \pm \varepsilon$ . It is easy to see that if we could produce these sets in polynomial time, then the running time of the entire process would also be polynomial<sup>11</sup>. As it turns out solving this algorithmic problem is hard. Instead we will show the following:

**Theorem 11.9.** *There is an  $O(n^3)$  time algorithm that given a bipartite graph  $G$  on vertex sets  $A, B$  does the following; if  $G$  is not  $\varepsilon$ -regular, it finds two vertex sets  $A' \subseteq A$  and  $B' \subseteq B$  of sizes  $|A'| \geq \varepsilon^5|A|$  and  $|B'| \geq \varepsilon^5|B|$  satisfying  $d(A', B') \notin d(A, B) \pm \varepsilon^5$ . If  $G$  is bipartite the algorithm returns nothing.*

Note that the algorithm does not produce a pair of sets  $A', B'$  satisfying the properties we used in the proof of the regularity lemma, since the  $\varepsilon$  is replaced by  $\varepsilon^5$ . But still

**Exercise 11.10.** Show that plugging the algorithm of Theorem 11.9 in the proof of the regularity lemma still produces an  $\varepsilon$ -regular partition in time  $O(n^3)$ .

One way to solve the above exercise is by checking the details of the proof. A more “abstract” way to see this is to observe that we are thus “proving” the regularity lemma for  $\varepsilon^5$  instead of  $\varepsilon$  (so the partition will be of size  $Tower(1/\varepsilon^{25})$  instead of  $Tower(1/\varepsilon^5)$ ).

To prove of Theorem 11.9 we will apply an interesting equivalence between two quasi-random properties of bipartite graphs. We first make the following observation. Denote by  $P_5$  the graph property that all pairs of vertices  $(x, y)$ , possibly except  $\varepsilon n^2$  of them, satisfy  $d(x, y) = (p^2 + \varepsilon)n$ . We have the following:

**Exercise 11.11.** Show that  $P_2 \iff P_5$ . You can prove this either directly or by showing that  $P_5$  is equivalent to  $P_4$  and then using the equivalences we have already proved (try to do both).

<sup>11</sup>Here and later in the proof, when we say polynomial, we mean polynomial in  $n$ . The dependence on  $\varepsilon$  would still be of Tower-type since the size of the partition can be of this size. So one can think of  $\varepsilon$  as fixed and the input size being  $n$ .

The properties  $P_1, \dots, P_5$  of quasi-random graphs have each a natural variant that deals with quasi-random bipartite graphs. For example, suppose a bipartite graph  $G$ , on vertex sets  $A, B$ , with  $|A| = |B| = n$ , has  $pn^2$  edges. We say that  $G$  satisfies property  $\bar{P}_2$  if for every pair of vertex sets  $A' \subseteq A, B' \subseteq B$  where  $|A'|, |B'| \geq \varepsilon n$  we have  $d(A', B') = p \pm \varepsilon$ . This can be considered the “bipartite” version of property  $P_2$  from the previous lecture (see Definition 10.9). More importantly, note that this property is simply the property of a bipartite graph being  $\varepsilon$ -regular (in the sense of the regularity lemma).

We can also define a bipartite analogue of property  $P_5$ . Suppose  $G$  has vertex sets  $A, B$ , with  $|A| = |B| = n$ , and  $pn^2$  edges. We say that  $G$  satisfies property  $\bar{P}_5$  if all but at most  $\varepsilon n^2$  pairs of vertices  $\{x, y\} \subseteq A$ , satisfy  $d(x, y) = (p^2 \pm \varepsilon)n$ . Then we have the following bipartite version of the equivalence between  $P_2$  and  $P_5$ .

**Exercise 11.12.** Show that  $\bar{P}_2 \iff \bar{P}_5$ . In particular, if  $(A, B)$  is not  $\varepsilon$ -regular, then  $A$  contains  $\varepsilon^4 n^2$  pairs of vertices  $x, y$  for which  $d(x, y) \geq (d^2 + \varepsilon^4)n$ , where  $d = d(A, B)$ .

The fact that  $\bar{P}_2 \implies \bar{P}_5$ . The more interesting part is the other direction (which is the only direction we will need in the proof). To prove this direction one can check that all the arguments one needs to prove that  $P_5 \implies P_2$  carry almost word for word to the bipartite setting. Another shorter way is to reduce the fact that  $\bar{P}_5 \implies \bar{P}_2$  to the fact that  $P_5 \implies P_2$ .

**Proof** (Theorem 11.9): If  $G$  is not  $\varepsilon$ -regular, then by Exercise 11.12 the set  $A$  contains at least  $\varepsilon^4 n^2$  pairs of vertices  $x, y$  satisfying  $d(x, y) \geq (d^2 + \varepsilon^4)n$ . It is obvious that such pairs can be found in  $O(n^3)$  time. Assume we found a vertex  $x \in A$  satisfying  $d(x) = (d \pm \varepsilon^5)n$  so that  $d(x, x') \geq (d^2 + \varepsilon^4)n$  for at least  $\varepsilon^5 n$  other vertices  $x' \in A$ . We claim that in this case the algorithm can return  $N(x)$  as the set  $B'$  and the vertices  $x'$  satisfying  $d(x, x') \geq (d^2 + \varepsilon^4)n$  as the set  $A'$ . Indeed, note that<sup>12</sup>  $|B'| \geq (d - \varepsilon^5)n \geq \varepsilon^5 n$ , and  $|A'| \geq \varepsilon^5 n$  so both sets are of the required size. As to  $d(A', B')$ , we have<sup>13</sup>

$$d(A', B') = \frac{e(A', B')}{|A'||B'|} \geq \frac{|A'|((d^2 + \varepsilon^4)n)}{|A'|(d + \varepsilon^5)n} = \frac{d^2 + \varepsilon^4}{d + \varepsilon^5} \geq d + \varepsilon^5.$$

Let us now complete the proof. Assume first that  $A$  contains  $2\varepsilon^5 n$  vertices whose degree is not  $(d \pm \varepsilon^4)n$ , and assume w.l.o.g. that  $\varepsilon^5 n$  of them have degree at least  $(d + \varepsilon^4)n$ . Then the algorithm can return these vertices as the  $A'$  and as  $B'$  it can return  $B$  itself. So assume that there are at most  $2\varepsilon^5 n$  vertices whose degree is not  $(d \pm \varepsilon^8)n$ . As there are at least  $\varepsilon^4 n^2$  pairs  $x, y$  satisfying  $d(x, y) \geq (d^2 + \varepsilon^4)n$ , then there must be at least  $\varepsilon^4 n/2$  vertices  $x$  satisfying this inequality with at least  $\varepsilon^4 n/2$  other vertices  $x'$ . If we throw out those vertices whose degree is not  $(d \pm \varepsilon^8)n$  (at most  $2\varepsilon^5 n$  such vertices) then we still<sup>14</sup> have at least  $\varepsilon^4 n/4$  vertices which fulfill the required condition, so one of them can be taken as the vertex  $x$  from the previous paragraph, thus completing the proof.  $\square$

### 11.3 Approximating MAX-CUT using the Regularity Lemma

We now give an application of the algorithmic version of the regularity lemma we obtain in the previous subsection. Our focus will be the MAX-CUT problem. In this problem the input is a graph  $G$  and the goal is (as the name suggests) to find a partition of  $V(G)$  into two sets  $A, B$  so as to maximize the number of edges connecting  $A$  to  $B$ . This problem is known to be NP-hard so we try to find a good approximation. Observe that Claim 8.2 we previously proved shows that we can always find a partition

<sup>12</sup>Assuming that  $d \geq 2\varepsilon^5$ , otherwise the graph is sparse and will trivially be  $\varepsilon$ -regular anyway. In fact, already when  $d \leq \varepsilon^3$  the graph is trivially  $\varepsilon$ -regular.

<sup>13</sup>Here we are also assuming that  $\varepsilon < \varepsilon_0$ . We can justify this assumption by simply replacing  $\varepsilon$  with  $\varepsilon_0$  if this is not the case. The only affect this might have is that instead of having  $\varepsilon^5$  in the statement of the Theorem we will get  $c\varepsilon^5$  for some absolute constant  $c = \varepsilon_0^5$ .

<sup>14</sup>Here we are again assuming that  $\varepsilon$  is small enough.



which achieves at least half of the optimum (from now on let us use  $OPT = OPT(G)$  for the optimal solution for  $G$ ).

Using the Regularity Lemma we can obtain the following:

**Theorem 11.13.** *There is an  $O(n^3)$  algorithm that given a graph  $G$  finds a partition  $A, B$  satisfying*

$$e(A, B) \geq OPT - \varepsilon n^2 .$$

Note that the above result is useless when  $G$  has  $o(n^2)$  edges. On the other hand, when  $G$  has  $cn^2$  edges, we know from Claim 8.2 that  $OPT \geq \frac{1}{2}cn^2$  so applying the above theorem with  $\varepsilon/c$  we can find a partition  $A, B$  satisfying  $e(A, B) \geq (1 - \varepsilon)OPT$ .

**Proof** (Theorem 11.13): We already know (via Theorem 11.9) how to find an  $\varepsilon$ -regular equipartition  $V_1, \dots, V_k$  of  $G$  in time  $O(n^3)$ . Therefore, we only need to show how to find the partition  $A, B$  in time  $O(n^3)$  given  $V_1, \dots, V_k$ . We will actually show that given the densities between all pairs of clusters  $V_i, V_j$  we can find the partition in constant time! So suppose we are given the densities  $d_{i,j} = d(V_i, V_j)$ . For every choice of  $a_1 \in \{0, \dots, 1/\varepsilon\}, \dots, a_k \in \{0, \dots, 1/\varepsilon\}$  let us set  $x_i = \varepsilon a_i$  and define

$$e(x_1, \dots, x_k) = \sum_{i < j} d_{i,j} (x_i(1 - x_j) + (1 - x_i)x_j) |V_i| |V_j| .$$

Observe that we can interpret  $e(x_1, \dots, x_k)$  as  $e(A, B)$  when  $A$  is obtained by picking  $x_i|V_i|$  vertices from each of the sets  $V_i$ . This is of course just an approximation of  $e(A, B)$ , but it is easy to see that since  $V_1, \dots, V_k$  is  $\varepsilon$ -regular, that no matter how we pick the  $x_i|V_i|$  vertices from each cluster  $V_i$ , we always have

$$|e(A, B) - e(x_1, \dots, x_k)| \leq \varepsilon n^2 . \quad (11.5)$$

We thus claim that after computing  $e(x_1, \dots, x_k)$  for all possible  $(1/\varepsilon)^k$  assignments, we can pick the assignment  $x_1, \dots, x_k$  that achieves the maximum, define a partition  $A, B$  using  $x_1, \dots, x_k$  as above, and return it as the output.

To show that we indeed have  $e(A, B) \geq OPT - \varepsilon n^2$ , let  $X, Y$  be the partition of  $V(G)$  satisfying  $e(X, Y) = OPT$ . Now define a new partition  $X', Y'$  so that for every  $1 \leq i \leq k$  the set  $|X' \cap V_i| = \varepsilon a'_i |V_i|$ , where  $a'_i$  is an integer. This partition can clearly be obtained from  $X, Y$  by “moving around” at most  $\varepsilon n$  vertices. Since this can change the number of edges connecting the two sets by at most  $\varepsilon n^2$ , we have

$$e(X', Y') \geq e(X, Y) - \varepsilon n^2 = OPT - \varepsilon n^2 .$$

Setting  $x'_i = \varepsilon a'_i$  we now get from (11.5) that

$$e(x'_1, \dots, x'_k) \geq e(X', Y') - \varepsilon n^2 \geq OPT - 2\varepsilon n^2 .$$

Since we defined  $(A, B)$  using an assignment that maximizes  $e(x_1, \dots, x_k)$  we have

$$e(x_1, \dots, x_k) \geq e(x'_1, \dots, x'_k) \geq OPT - 2\varepsilon n^2 .$$

Finally we get from another application of (11.5) that

$$e(A, B) \geq e(x_1, \dots, x_k) - \varepsilon n^2 \geq OPT - 3\varepsilon n^2 ,$$

as needed. □

## 12 Twelfth lecture

### 12.1 Hypergraph Ramsey Numbers

Recall that  $N = R(s, t)$  is the smallest number such that every 2 coloring of  $K_N$  contains either a red  $K_s$  or black  $K_t$ . We showed that  $R(s, t) \leq \binom{s+t-2}{s-1}$  and in particular  $2^{t/2} \leq R(t, t) \leq 4^t$ . This result can be generalized to 3-graphs.

**Theorem 12.1.** *For every  $s, t$  there exists  $N = R_3(s, t)$  such that every 2 coloring of  $K_N^3$  contains a red  $K_s^3$  or a black  $K_t^3$ .*

**Proof:** We show that

$$R_3(s, t) \leq R_2(R_3(s-1, t), R_3(s, t-1)) + 1. \quad (12.1)$$

We begin by picking an arbitrary vertex  $v$  and let  $X$  denote the rest of the vertices. Define a 2-coloring of the complete graph on the vertex set  $X$  as follows; we color edge  $(i, j)$  using the color given to  $(v, i, j)$ . By Ramsey's Theorem for graphs, this 2-coloring has either a vertex set  $X_1$  that spans a red  $K_{R_3(s-1, t)}$  or a vertex set  $X_2$  that spans a black  $K_{R_3(s, t-1)}$ . Assume the former holds. By induction, the hypergraph induced by  $X_1$  contains either black  $K_t^3$  (and we are done) or a red  $K_{s-1}^3$ . It is not easy to see that the way we defined the coloring of  $X$  guarantees that  $X_1$  together with  $v$  gives a red  $K_s^3$ , thus completing the induction step for proving (12.1). The base cases can easily be checked, thus proving the theorem.  $\square$

Let us try to get a feeling of the type of bound we get from the above proof. Recalling that we have previously proved that  $R(t, t) \geq 2^{t/2}$  we see that the bound we get from (12.1) satisfies

$$R_3(t, t) \geq R_2(R_3(t-1, t-1), R_3(t-1, t-1)) \geq 2^{R_3(t-1, t-1)/2}$$

implying that this bound grows like a tower of  $\sqrt{2}$  of height  $t$ , which is a pretty weak. We now give a better bound based on an argument of Erdős and Rado.

**Theorem 12.2.**  $R_3(t, t) \leq 2^{\binom{R_2(t, t)}{2}}$ . In particular  $R_3(t, t) \leq 2^{2^{4t}}$ .

**Proof:** We try to imitate the first proof of Ramsey's Theorem for graphs we previously gave. We iteratively construct a 3-tuple  $(S_i, \phi_i, X_i)$  where  $S_i$  and  $X_i$  are two disjoint vertex sets,  $S_i = \{s_1, \dots, s_i\}$ , and  $\phi_i$  is a 2-coloring of the pairs of vertices of  $S_i$  satisfying the following two key conditions: (1) For every  $1 \leq a < b < c \leq i$  the edge  $(S_a, S_b, S_c)$  is colored  $\phi_i(S_a, S_b)$ , and (2) For every  $1 \leq a < b \leq i$  and  $x \in X_i$  the edge  $(S_a, S_b, x)$  is also colored  $\phi_i(S_a, S_b)$ .

Suppose first that we have managed to construct such a 3-tuple with  $i = R_2(t, t)$ . Then by Ramsey's Theorem, the coloring  $\phi_i$  of  $S_i$  (when viewed as a 2-coloring of the complete graph on  $R_2(t, t)$  vertices) contains a monochromatic  $K_t$ . It is now easy to see that the first key property of  $\phi_i$  guarantees that the vertices of this  $K_t$  also form a monochromatic  $K_t^3$  in the hypergraph.

To construct the 3-tuples  $(S_i, \phi_i, X_i)$  we start with  $S_1 = \{v\}$ , with  $v$  an arbitrary vertex, and  $X_1 = V \setminus v$ . Assuming we defined  $(S_i, \phi_i, X_i)$  we now define  $(S_{i+1}, \phi_{i+1}, X_{i+1})$ . Pick an arbitrary vertex from  $X_i$ , remove it from  $X_i$  and add it to  $S_i$ . This vertex is going to be  $s_{i+1}$ . We now need to define the new set  $X_{i+1}$  and the new coloring  $\phi_{i+1}$ . The new coloring  $\phi_{i+1}$  will agree with  $\phi_i$  on all pairs  $(S_a, S_b)$  with  $1 \leq a < b \leq i$ . So we just need to color the pairs  $(S_a, S_{i+1})$  with  $1 \leq a \leq i$ . The set  $X_i$  clearly contains a subset  $X_i^1$  of size at least  $|X_i|/2$  so that all the triples  $(s_1, s_{j+1}, x)$ , with  $x \in X_i^1$ , are colored with the same color. If this color is red we color  $(s_1, s_{j+1})$  red, otherwise we color it black. In a similar manner, the set  $X_i^1$  clearly contains a subset  $X_i^2$  of size at least  $|X_i^1|/2$  so that all the triples  $(s_2, s_{j+1}, x)$ , with  $x \in X_i^2$ , are colored with the same color. If this color is red we color  $(s_2, s_{j+1})$  red, otherwise we color it black. We continue in this manner and define  $X_{i+1} = X_i^i$ . It is easy to see that  $(S_{i+1}, \phi_{i+1}, X_{i+1})$  has

the required properties. Finally, note that when constructing  $(S_i, \phi_i, X_i)$ , we were forced to reduce the size of  $X_1$  by a factor of at most 2, for  $\binom{i}{2}$  times. Therefore, if the initial hypergraph is of size at least  $2^{\binom{R_2(t,t)}{2}}$  we are guaranteed that the process will be able to go on for  $i = R_2(t, t)$  iterations.  $\square$

We note that the best known upper bound for  $R_3(t, t)$  is given by  $R_3(t, t) \leq 2^{2^{2t}}$ . Somewhat annoyingly, the best lower bound follows from the obvious random construction as the following exercise suggests.

**Exercise 12.3.** Show that  $R_3(t, t) \geq 2^{t^2/6}$ .

We now turn to show that one can obtain a rather tight bound for  $R_3(t, t, t, t)$ , that is, the Ramsey number for 4-colors. The upper bound follows by adapting the above proof for 2-colors and is thus left as an exercise.

**Exercise 12.4.** Show that  $R_3(t, t, t, t) \leq 2^{2^{ct}}$ .

The lower bound is a bit harder.

**Theorem 12.5.** *Suppose there is a 2-coloring of  $K_m$  with no monochromatic  $K_{t-1}$ . Then  $R_3(t, t, t) \geq 2^m$ . In particular, we have  $R_3(t, t, t) \geq 2^{2^{t/2}}$ .*

**Proof:** We define a 4-coloring of  $K_{2^m}^3$  without a monochromatic  $K_t^3$ . We think of every vertex as a binary string of length  $m$ . For such strings  $x, y$  we define  $\delta(x, y)$  to be the largest index where the strings differ. We also assume that the strings/vertices are ordered lexicographically. We will rely on the following facts that are easy to verify.

1. If  $x < y < z$  then  $\delta(x, y) \neq \delta(y, z)$ .
2. If  $x < y < z$  then  $\delta(x, z) = \max\{\delta(x, y), \delta(y, z)\}$ .

Note that the second fact above implies that more generally, if  $x_1 < \dots < x_l$  then

$$\delta(x_1, x_l) = \max_{1 \leq i \leq l-1} \delta(x_i, x_{i+1}) \tag{12.2}$$

We define the coloring as follows: Given  $x < y < z$ , set  $\delta_1 = \delta(x, y)$ ,  $\delta_2 = \delta(y, z)$ , and color the edge  $(x, y, z)$  as follows:

1. If  $\delta_1 < \delta_2$  and  $(\delta_1, \delta_2)$  is red (in the assumed  $K_m$ ) then color  $(x, y, z)$  with color  $A$ .
2. If  $\delta_1 < \delta_2$  and  $(\delta_1, \delta_2)$  is black then color  $(x, y, z)$  with color  $B$ .
3. If  $\delta_1 > \delta_2$  and  $(\delta_1, \delta_2)$  is red then color  $(x, y, z)$  with color  $C$ .
4. If  $\delta_1 > \delta_2$  and  $(\delta_1, \delta_2)$  is black then color  $(x, y, z)$  with color  $D$ .

Assume that  $x_1 < x_2 < \dots < x_t$  form a monochromatic  $K_t^3$  whose color is  $A$  (the other 3 cases are identical). Define  $\delta_i = \delta(x_i, x_{i+1})$  for every  $1 \leq i \leq t-1$ . We will now show that  $\{\delta_i\}_{1 \leq i \leq t-1}$  forms a red clique of size  $t-1$  in the original graph  $K_m$ , which will be a contradiction. First note that since all triples  $x_i, x_{i+1}, x_{i+2}$  are colored  $A$ , we get

$$\delta_1 < \delta_2 < \delta_3 < \dots < \delta_{t-1} \tag{12.3}$$

so these vertices are indeed distinct. We also get from the definition of the coloring, that for every  $1 \leq i \leq t-2$  the edge  $(\delta_i, \delta_{i+1})$  is colored red. For every  $i < j$  consider the edge  $(x_i, x_{i+1}, x_{j+1})$ . We

now get from (12.2) and (12.3) that  $\delta(x_{i+1}, x_{j+1}) = \delta(x_j, x_{j+1}) = \delta_j$ . Since  $(x_i, x_{i+1}, x_{j+1})$  is colored  $A$ , we get that  $(\delta_i, \delta_j)$  must be colored red as well. We thus get that  $\delta_1, \dots, \delta_{t-1}$  form a red.  $\square$

We now return to the problem of bounding the 2-color Ramsey numbers  $R_k(t, t)$ . As the following exercise suggests, one can easily adapt the proof of our previous upper  $R_3(t, t) \leq 2^{\binom{R_2(t, t)}{2}}$  and obtain the following recursive relation.

**Exercise 12.6.** Show that for every  $k \geq 3$  we have  $R_k(t, t) \leq 2^{\binom{R_{k-1}(t, t)}{k-1}}$ .

Let us define  $twr_1(x) = x$  and  $twr_k(x) = 2^{twr_{k-1}(x)}$ . So, for example,  $twr_3(x)$  is the function  $2^{2^x}$ . We observed earlier that the bound  $R_2(t, t) \leq 2^{2t} = twr_2(2t)$ , together with the above recursive relation implies that  $R_3(t, t) \leq 2^{2^{4t}} = twr_3(4t)$ . More generally, it is easy to see this this relation implies that  $R_k(t, t) \leq twr_k((4 + o(1))k)$ . It is widely believed this is bound is tight. However, as we have mention earlier, this is not known even for  $k = 3$ . However, as the following Theorem shows, giving a tight bound for  $k = 3$ , that is showing that  $R_3(t, t) \geq 2^{ct}$ , will give a tight bound for general  $k$ , that is, imply that  $R_k(t, t) \geq twr_k(ct)$ .

**Theorem 12.7** (Step-up Lemma). *The following holds for every  $k \geq 3$ ; if there is a 2-coloring of  $K_m^k$  with no monochromatic  $K_t$  then there is a 2-coloring of  $K_{2m}^k$  with no monochromatic  $K_{2t+k}$ , that is,  $R_{k+1}(2t+k, 2t+k) \geq 2^m$ .*

**Proof:** We prove the result for  $k = 3$  and leave the general case as an exercise. Assuming we have a 2-coloring of  $K_m^3$  with no monochromatic  $K_{t-1}^3$  we define a 2-coloring of  $K_{2m}^4$  with no monochromatic  $K_{2t-1}^4$ . As in the last proof, we think of the vertices as binary strings of length  $m$ , and define  $\delta$  as before. Given  $x_1 < x_2 < x_3 < x_4$ , set  $\delta_1 = \delta(x_1, x_2)$ ,  $\delta_2 = \delta(x_2, x_3)$ ,  $\delta_3 = \delta(x_3, x_4)$ . We use the following rules:

1. If  $\delta_1 < \delta_2 < \delta_3$  we color  $(x_1, x_2, x_3, x_4)$  using the color given to  $(\delta_1, \delta_2, \delta_3)$ .
2. If  $\delta_1 > \delta_2 > \delta_3$  we color  $(x_1, x_2, x_3, x_4)$  using the color given to  $(\delta_1, \delta_2, \delta_3)$ .
3. If  $\delta_1 > \delta_2 < \delta_3$  we color  $(x_1, x_2, x_3, x_4)$  black.
4. If  $\delta_1 < \delta_2 > \delta_3$  we color  $(x_1, x_2, x_3, x_4)$  red.

Assume  $x_1 < \dots < x_{2t-1}$  is a monochromatic  $K_{2t-1}^4$ , and suppose it is red (the black case is identical). Consider  $\delta_1, \delta_2, \dots, \delta_{2t-2}$ . By coloring rules 3 and 4, there is at most one consecutive triplet which is not monotone. As the sequence is of length  $2t - 2$ , it must contain a consecutive monotone sub-sequence of length  $t - 1$ . Assume, without loss of generality that it is of the form  $\delta_1 < \delta_2 < \dots < \delta_{t-1}$  and consider  $x_1, \dots, x_t$ . We get that for every  $1 \leq i \leq t - 3$ , the edge  $(\delta_i, \delta_{i+1}, \delta_{i+2})$  is colored red. Consider now a general triple  $(\delta_i, \delta_j, \delta_k)$  and take the 4-edge  $(x_i, x_{i+1}, x_{j+1}, x_{k+1})$ . We get from the monotonicity of the  $\delta$ 's and (12.3) that  $\delta(x_{i+1}, x_{j+1}) = \delta_j$  and that  $\delta(x_{j+1}, x_{k+1}) = \delta_k$ . Since  $(x_i, x_{i+1}, x_{j+1}, x_{k+1})$  was colored red and  $\delta_i < \delta_j < \delta_k$ , this means that  $(\delta_i, \delta_j, \delta_k)$  was also colored red, implying that  $\delta_1, \delta_2, \dots, \delta_{t-1}$  form a red  $K_{t-1}^3$ , which is a contradiction.  $\square$

Let use finally mention that the above recursive relations we proved for  $R_k(t, t)$  can be easily adapted to the multicolor case. Since we already proved that for 4-colors we have  $R_3(t, t, t, t) = twr_3(\Theta(t))$  these recursive relations imply that  $R_k(t, t, t, t) = twr_k(\Theta(t))$ .

## 13 Notations

$d(x)$       The degree of vertex  $x$

$\delta(G)$	$\min_{x \in V(G)} d(x)$
$e(G)$	$ E(G) $ – the number of vertices in $G$
$ex(n, H)$	The maximal number of edges for a $H$ -free graph on $n$ vertices
$K_n$	The complete graph on $n$ vertices
$K_{n_1, \dots, n_r}$	The complete $r$ -partite graph with partitions of size $n_1, n_2, \dots, n_r$
$K_r^s$	$r$ -graph with a clique of size $s$ . Sometimes written with super- and subscript reversed.

## References

- F.A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci. USA*, 32:331–332, 1946.
- B. Bollobas and A. Thomason. Highly linked graphs. *Combinatorica*, 16(3):313–320, 1996.
- JA Bondy. Pancyclic graphs I. *Journal of Combinatorial Theory, Series B*, 11(1):80–84, 1971.
- F.R.K. Chung, R.L. Graham, and R.M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.
- K.F. Roth. On certain sets of integers. *Journal of the London Mathematical Society*, 1(1):104, 1953.
- I.Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. *Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bolyai*, 18:939–945, 1978.
- E. Szemerédi. On sets of integers containing no four elements in arithmetic progression. *Acta Mathematica Hungarica*, 20(1):89–104, 1969.
- E. Szemerédi. On sets of integers containing no  $k$  elements in arithmetic progression. *Acta Arith*, 27(585):199–245, 1975.