THEOREM 1 (Mantel's theorem)

If G_n has no triangles, then $e(G_n) \leq \lfloor \frac{n^2}{4} \rfloor$ with equality iff $G_n \simeq K_{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2}+1 \rfloor}$

proof :

Let $xy \in E(G)$

$N(x) \cap N(y) = \phi$ since G is triangle free

In particular, $d(x) + d(y) \leq n$

$$\therefore \sum_{xy \in E(G)} d(x) + d(y) \leq n \times e(G)$$

But $$\sum_{xy \in E(G)} d(x) + d(y) = \sum_{x \in V(G)} (d(x))^2$$

since each vertex contributes $d(x)$, $d(x)$ times (once for each of its $d(x)$ neighbours)

$$\therefore \sum_{x \in V(G)} (d(x))^2 \leq n \cdot e(G)$$

$$\Rightarrow \frac{\left(\sum_{x \in V(G)} d(x)\right)^2}{n} \leq \sum_{x \in V(G)} (d(x))^2 \leq n \cdot e(G)$$

$$\Rightarrow \frac{(2e(G))^2}{n} \leq n \cdot e(G)$$

$$\Rightarrow e(G) \leq \frac{n^2}{4}$$

$$\Rightarrow e(G) \leq \left\lfloor \frac{n^2}{4} \right\rfloor \quad (\because e(G) \in \mathbb{Z})$$

For n being even, $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil} = K_{m,m}$ $(n = 2m)$

$$e(K_{m,m}) = m^2 = \frac{(2m)^2}{4}$$

For n being odd, $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil} = K_{m,m+1}$ $(n = 2m+1)$

$$e(K_{m,m+1}) = m(m+1) = \left\lfloor \frac{(2m+1)^2}{4} \right\rfloor$$

Thus the inequality is tight. We now show that $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$ is the only graph that satisfies the equality of Mantel's theorem.

Let n be even with $n = 2m$

We want equality to hold in $d(x) + d(y) \leq n$ and also in the Cauchy Schwartz Sedrakyan lemma in which we claimed $\frac{(\Sigma\, d(x))^2}{n} \leq \frac{\Sigma\, (d(x))^2}{1}$

Suppose $d(x) + d(y) = n$ $\forall (x, y)$ and $d(x)$ is constant (since equality holds in Cauchy Schwartz if all are the same), then it is an m regular graph. Fix $x_0, y_0 \in G$ s.t. $x_0 y_0 \in E(G)$

Then $|N(x_0)| = |N(y_0)| = m$ and $d(x_0) = m$.

Further $N(x_0) \cap N(y_0) = \phi \Rightarrow$ ~~[illegible]~~

~~[illegible]~~ $N(x_0)$, $N(y_0)$ give rise to a bipartite graph $K_{m,m}$

Let $n = 2m+1$

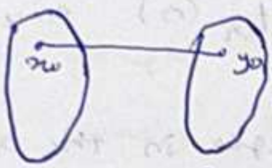
$d(u) + d(v) = n \quad \forall\, uv \in E(G) \implies$ graph is connected

Let $x_0 y_0 \in E(G)$

$d(x_0) + d(y_0) = n = 2m+1$

WLOG, let $d(x_0) = r$ be odd.

It is now easy to see that $G \cong K_{n-r,\,r}$



$n - r = |N(y_0)|$ $\qquad$ $r = |N(x_0)|$

we want maximum tightness in

$$\frac{(2r(n-r))^2}{n} \leq \sum_{x \in V(G)} (d(x))^2 = (n-r)r^2 + (n-r)^2 r$$

i.e. $\frac{4}{n} r(n-r) \leq n$

$\equiv \quad n^2 - 4nr + 4r^2 \geq 0$

i.e. $(n-2r)^2 \geq 0$

Since r cannot be $n/2$, maximum tightness occurs at $r = \lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$ (choosing whichever is odd since r was taken to be odd.

THEOREM 2 (Turan's theorem)

If G_n is K_{r+1} free, $e(G_n) \leq t_r(n)$ which is the Turan number i.e. no of edges in an r-partite graph with parts that are of sizes $\lceil \frac{n}{r} \rceil$ or $\lfloor \frac{n}{r} \rfloor$ maximally. Equality holds iff $G_n \cong T_r(n)$ the Turan graph.

proof :

We prove something much stronger. Suppose G_n is K_{r+1} free, then there exists an r-partite graph H that uses vertices of G_n such that $d_G(x) \le d_H(x)$

Suppose this is true, then,

$$\sum_{x \in V(G_n)} d_G(x) \le \sum_{x \in V(H)} d_H(x) \le t_r(n) \quad \text{since } H \text{ is}$$

r-partite and $T_r(n)$ has all edges in the r-partite graph. (Turan does best in r-partite is easy to prove !)

We prove our claim using induction. For $r = 1$ it is trivially true since, if G is K_2 free, then $d_G(x) = 0 \ \forall x \in G$ and we construct a bipartite graph H on G in any way we want.

Suppose the statement is true for graphs that are K_r free, we prove it for graphs which are K_{r+1} free. Let G be K_{r+1} free.

Let x_0 have maximum degree in G.

Consider $N(x_0)$ and $V \setminus \{x_0 \cup N(x_0)\}$.

Let U be the graph induced by G on $V \setminus \{x_0 \cup N(x_0)\}$

$G[N(x_0)]$ is K_r free and hence $\exists$ $(r-1)$ partite graph on vertices of the graph induced by G on $N(x_0)$.

call the $(r-1)$ partite graph as H'

For every $u \in U$, delete all current edges and add edges from u to every vertex in $G([N(x)])$ so that the new $N(u)$ will be $N(x_0)$

claim: $H' \cup \{x_0 \cup U\}$ is the required r-partite graph (call it H)

It is clearly r partite, we only need to check the inequality for vertices

For x_0, $d_G(x_0) = d_H(x_0)$

For $y \in U$, $d_G(y) \le d_G(x_0)$

$= d_H(x_0)$

$= d_H(y)$

For $y \in G([N(x_0)])$,

$$d_H(y) = d_{H'}(y) + \overbrace{n - d_G(x_0)}^{x_0 + \text{all vertices of } U}$$

$\geq d_{N(x_0)}(y) + n - d_G(x_0)$ ($\because$ induc hypo)

$\geq d_{N(x_0)}(y) + d_U(y) + 1$ ($\because$ y may not have been connected to all in U).

$= d_G(y)$

THEOREM 3

1) $e(G_n) = t_r(n) \Rightarrow \delta(G_n) \le \delta(T_r(n)) = n - \lceil \frac{n}{r} \rceil \le n - \lfloor \frac{n}{r} \rfloor$

$= \Delta(T_r(n)) \le \Delta(G_n)$

2) $t_r(n-1) = t_r(n) - (n - \lceil \frac{n}{r} \rceil)$

3) $\forall k \in \mathbb{N}$, $t_r(n) \geq k\cdot(n-k) + t_{r-1}(n-k)$ with

equality iff $k = \lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$

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

5) $t_{r-1}(n) \leq t_r(n)$

proof :

1) suppose FTSOC, $\delta(G_n) > \delta(T_r(n))$

Then $d(x) > \delta(T_r(n))$ $\forall x \in V(G)$

$\therefore 2e(G_n) > n\left(n - \lceil \frac{n}{r} \rceil\right)$; (summing over all x)

$\therefore e(G_n) = t_r(n) > \dfrac{n(n - \lceil \frac{n}{r} \rceil)}{2}$

$\therefore t_r(n) > e\big(K_{\underbrace{\lceil \frac{n}{r} \rceil, \lceil \frac{n}{r} \rceil, \ldots, \lceil \frac{n}{r} \rceil}_{r \text{ times}}}\big)$

This is ~~a~~ contradiction since $t_r(n)$ is the

no of edges in $K_{\lceil \frac{n}{r} \rceil, \lceil \frac{n}{r} \rceil, \ldots, \lceil \frac{n}{r} \rceil, \lfloor \frac{n}{r} \rfloor, \lfloor \frac{n}{r} \rfloor, \ldots, \lfloor \frac{n}{r} \rfloor}$

Similarly $\Delta(T_r(n)) \leq \Delta(G_n)$

2) consider $T_r(n-1)$. we wish to add a vertex and get $T_r(n)$. The vertex will have to be added to the part (with $\lfloor \frac{n-1}{r} \rfloor$ vertices to keep the balance of parts being 'roughly' same size maximally

Now we connect the vertex to every vertex outside its own bubble making $T_r(n)$

$\therefore$ no. of edges added $= n - \lceil \frac{n}{r} \rceil$

$\therefore \quad t_r(n-1) + \left(n - \lceil \frac{n}{r} \rceil\right) = t_r(n)$

if how the $n - \lceil \frac{n}{r} \rceil$ came from $n - \lfloor \frac{n-1}{r} \rfloor$ bothers you, just kick out a vertex from $T_r(n)$.

(To avoid this particular disparity, we added the new vertex to $\lfloor \frac{n-1}{r} \rfloor$ vertices part)

3) Start with $T_{r-1}(n-k)$. Add an extra bubble with k more vertices & connect each of these k vertices to all $n-k$ of $T_{r-1}(n-k)$. We added $k(n-k)$ new edges. We have a complete r-partite graph with $t_{r-1}(n-k) + k(n-k)$ edges but we know that it has to be less than $t_r(n)$ since the Turan graph maximizes the no. of edges

4) Starting with $t_r(n)$, select one point from each of the r bubbles. These points themselves induce K_r. Delete these points & edges incident on them. No of edges deleted $= \binom{r}{2} + (n-r)(r-1)$. The second term appears since each vertex that is NOT deleted is connected to $(r-1)$ vertices that get deleted. There are $n-r$ such vertices that will not get deleted.

5) View $T_{r-1}(n)$ as an r-partite graph with one part empty. For r-partite, $t_r(n)$ is maximum and hence $t_{r-1}(n) \le t_r(n)$. (Equivalent to putting $k=0$ in (3)

THEOREM 4

If G_n is P_k (path of length k) free, $e(G_n) \le \frac{n(k-1)}{2}$

proof:

WLOG assume G is connected else work on connected parts of G.

Let $v_1, v_2, \dots, v_\ell$ be the longest path in G (length $= \ell-1$)
(by definition, $v_1 \neq v_\ell$)

v_1 and v_2 have no 'off-path' neighbours (else longer path would be possible)

Further, v_1, v_ℓ cannot be connected else we get a longer path through an 'off-path' neighbour of cycle (suppose no such vertex is present, then $G = C_\ell$ (biggest cycle) in which case $e(G_n) \le \frac{\ell(k-1)}{2}$ holds true anyways since $G_n \not\supseteq P_k \Rightarrow \ell-1 \le k$)

Also note that v_1 connected to v_{i+1}, v_i connected to v_ℓ is not possible for the same reason of C_ℓ being formed.

$\therefore$ If $S = \{ i \mid x_{i+1} x_1 \in E(G) \}$

$T = \{ i \mid x_i x_l \in E(G) \}$

then $S \cap T = \emptyset$, $S \subseteq \{1, 2, \ldots, l-2\}$

and $T \subseteq \{2, 3, \ldots, l-1\}$

$d(x_1) = |S|$, $d(x_l) = |T|$

$\therefore d(x_1) + d(x_l) = |S| + |T| = |S \cup T| \le l-1$

claim: If for every non adjacent $u, v \in V(G)$, we have that $d(u) + d(v) \geq k$, then $G \supseteq P_k$ (further assuming that in longest path $v_1 \cdots v_m$, $v_1 v_m \notin E(G)$)

~~[illegible]~~

~~[illegible]~~

proof of claim: in particular $d(x_1) + d(x_l) \geq k$

$\therefore l - 1 \geq k$

But $l-1$ is length of $v_1 \cdots v_l$ which is the longest path in G

$\therefore P_k \subseteq G$.

Now if $n \le k$, there is nothing to prove since

$$e(G_n) \le \binom{n}{2} = \frac{n(n-1)}{2} < \frac{n(k-1)}{2}$$

if $n > k$, let x_0 be min. degree vertex.

we must have that $d(x_0) \le \frac{k-1}{2}$ else,

$d(y) + d(x) \geq 2d(x_0) > k-1 \quad \forall x, y \in V(G)$

$\therefore d(x) + d(y) \geq k \quad \forall x, y$ non edges in G

$\therefore P_k \subseteq G$ (contradiction)

$$\therefore e(G_n) = e(G\setminus\{x_0\}) + d(x_0)$$

$$\le \frac{(n-1)(k-1)}{2} + \frac{k-1}{2} \quad \text{(induction - hypothesis)}$$

$$= \frac{(k-1)n}{2}$$

Also note that inequality is tight since equality is achieved whenever $k|n$ and we have G having $\frac{n}{k}$ parts of K_k

THEOREM 5 (Zarankiewicks problem)

We define $z(m,n,s,t)$ to be the maximum no. of edges in a bipartite graph having parts of sizes m, n s.t. $\nexists$ $K_{s,t}$ in the graph i.e. $\nexists$ $S \subseteq X$, $T \subseteq Y$ s.t. $|S|=s$, $|T|=t$; $e(S,T)=st$

we have

1) $2\,ex(n, K_{s,t}) \le z(n,n,s,t) \le ex(2n, K_{s,t})$

2) $z(m,n,s,t) \le (m-s+1)(t-1)^{1/s} n^{1-1/s} + (s-1)n$

for $s \ge t$ (and for $t \ge s$, we swap t, s WLOG)

3) $z(n,n,2,2) \le (k^2-k+1)k$ for $k = \frac{1+\sqrt{4n-3}}{2}$

4) $z(n,n,t,t) = \Omega_t(n^{2-\frac{2}{t+1}})$

proof:

1) $z(n,n,s,t) \le ex(2n, K_{s,t})$ follows trivially since $z(n,n,s,t)$ applies to bipartite graphs but

$ex(2n, K_{s,t})$ can be any graph on $2n$ vertices.

In general, $z(m,n;s,t) \le ex(m+n, K_{s,t})$

Now let G_n be $K_{s,t}$ free. Create another copy G_n' and define H to be the bipartite graph $G_n \cup G_n'$ where if $uv \in E(G_n)$ then $uv', u'v$ are edges in H.

Aim: H has $K_{s,t} \Rightarrow G_n$ has $K_{s,t}$

H has $K_{s,t} \Rightarrow \{v_1, \dots, v_s\}$ & $\{w_1, \dots, w_t\}$ form $K_{s,t}$

$\therefore v_i w_j \in E(H) \quad \forall \; 1 \le i \le s, \; 1 \le j \le t$

$\therefore v_i m_j \in E(G_n) \quad \forall \; 1 \le i \le s, \; 1 \le j \le t$ where

m_j is pre image of w_j ie. m_j became w_j

$\therefore K_{s,t} \subseteq G_n$ (note that $v_i \ne w_j \; \forall \, i, j$)

Also $e(H) = 2e(G_n)$

$\therefore K_{s,t} \not\subseteq G_n \Rightarrow K_{s,t} \not\subseteq H_{\text{bipartite}}$

$\therefore 2\,ex(n, K_{s,t}) \le z(n,n,s,t)$

2) Let $S \subseteq X$ be chosen arbitrarily with $|S| = s$. Define $N_Y(S) =$ common neighbourhood of S in $Y = \{ y \in Y \mid xy \in E(G)$ for all $x \in S \}$

Clearly $|N_Y(S)| < t$ (else $K_{s,t} \subseteq G$)

$\therefore |N_Y(S)| \le t-1$

$$E(|N_Y(S)|) = \sum_{y \in Y} \mathbb{P}(y \in N_Y(S))$$

$$= \sum_{y \in Y} \frac{\binom{d(y)}{s}}{\binom{m}{s}}$$

$$\therefore \sum_{y \in Y} \binom{d(y)}{s} \leq \binom{m}{s}(t-1)$$

Now $\frac{1}{n} \sum_{y \in Y} \binom{d(y)}{s} \geq \binom{\frac{\sum d(y)}{n}}{s}$ (Jensens inequality)

$\left(f\left(\frac{\sum a_i x_i}{\sum a_i} \right) \leq \frac{\sum a_i f(x_i)}{\sum a_i} \text{ for convex } f \right)$

$$\therefore n \binom{\frac{\sum_{y \in Y} d(y)}{n}}{s} \leq \binom{m}{s}(t-1)$$

$$\therefore n \binom{e/n}{s} \leq \binom{m}{s}(t-1)$$

where e = no. of edges in $G = (X, Y) = \sum_{y \in Y} d(y)$.

$\therefore n(\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-s+1)) \leq (t-1)(m(m-1)\cdots(m-s+1))$

($\because$ setting $\alpha = e/n$)

$$\therefore \frac{\alpha(\alpha-1)\cdots(\alpha-s+1)}{m(m-1)\cdots(m-s+1)} \leq \frac{t-1}{n}$$

$\therefore \left(\frac{\alpha-s+1}{m-s+1}\right)^s \leq \frac{t-1}{n}$ $\left(\because \frac{\alpha}{m} \geq \frac{\alpha-1}{m-1} \Leftrightarrow \alpha \leq m \Leftrightarrow e \leq mn \text{ which is true} \right)$

$$\therefore \alpha \leq (m-s+1)\frac{(t-1)^{1/s}}{n^{1/s}} + (s-1)$$

$$\therefore e \leq (m-s+1)(t-1)^{1/s} n^{1-1/s} + n(s-1)$$

3) Proceeding as in 2,

$$\frac{\alpha(\alpha-1)}{n(n-1)} \leq \frac{1}{n}$$

Instead of loosely writing $\left(\frac{\alpha-1}{n-1}\right)^2 \leq \frac{1}{n}$, we

solve explicitly for α to get

$$\alpha \leq \frac{1+\sqrt{4n-3}}{2}$$

$$\therefore \ e \leq n\left(\frac{1+\sqrt{4n-3}}{2}\right)$$

using a number theoretic fact, $\sqrt{4n-3}$ is a perfect

square iff $n = k^2-k+1$ & in this case $\sqrt{4n-3} = 2k-1$

$$e \leq (k^2-k+1)\,k$$

Note: If we want equality, we must ensure

equality in $\sum_{y\in Y} \binom{d(y)}{2} \leq \binom{n}{2}$ (=Jensen's)

and this only happens for $d(y) = k =$ const. (no proof here)

4) Notation: $f = \Omega(g)$ if $f > Cg$ for some C after a

certain point in domains of f, g

$N =$ no. of pairs (S, T) such that $|S| = |T| = t$, $e(S,T) = t^2$

then $E(N) = \sum_{\substack{S\subseteq X \\ T\subseteq Y}} \mathbb{P}(e(S,T) = t^2)$

$= \binom{n}{t}\binom{n}{t} p^{t^2}$ if p is $\mathbb{P}(xy$ is an edge$)$

(for all $x, y \in V(G)$)

Also, $E(e(G)) = p \cdot n^2$ (G is bipartite of size n, n.)

we know $z(n,n,t,t) \leq O(n^{2-1/t})$ with equality iff $t=2$

Demand $n^2 p - \binom{n}{t}^2 p^{t^2} \geq \frac{n^2 p}{2}$

$\therefore \binom{n}{t}^2 p^{t^2} \leq \frac{n^2 p}{2}$

$\therefore \exists$ G with at most $\frac{1}{2} e(G)$ copies of $K_{t,t}$ in it

($\because$ If we delete one edge from each copy, we obtain G' such that $e(G') \geq \frac{n^2 p}{2}$ and $K_{t,t} \not\subset G$)

$\therefore$ p is such that $\binom{n}{t}^2 p^{t^2} \leq \frac{n^2 p}{2}$

Now, $\binom{n}{t}^2 \leq \left(\frac{e \cdot n}{t}\right)^{2t}$ ($\because n! \geq \left(\frac{n}{e}\right)^n$, $e \approx 2.71828\ldots$)

$\therefore$ enough to demand $\left(\frac{e \cdot n}{t}\right)^{2t} p^{t^2-1} \leq \frac{n^2}{2}$

$\therefore$ $p = C_t \cdot \frac{1}{n^{2/t+1}}$

$\therefore z(n,n,t,t) \geq \Omega_t(n^{2-2/t+1})$

($\because E(N) = \binom{n}{t}^2 p^{t^2} \Rightarrow n^2, p^{t^2}$ factors)

THEOREM 6 (Erdős-Stone theorem) (weak)

for $r \geq 1$, $0 < \varepsilon < \frac{1}{2}$, $n >> 0$, if $\delta(G_n) \geq (1 - \frac{1}{r} + \varepsilon) n$

then $\exists\ K_{r+1}(V_1, V_2, \ldots, V_{r+1})$ s.t $V_i \subseteq V$ are disjoint

with $|V_i| = t \geq \left\lceil \frac{\varepsilon \log n}{2^{r-1}(r-1)!} \right\rceil$ $\forall\ i = 1, 2, \ldots, r+1$

proof :

we induct on the parameter r

Base case: $r = 1$

$\delta(G_n) \geq \varepsilon n$ is known.

suppose $K_{t,t} \not\subset G$ then consider $\beta = \{(x, T)\}$

where x varies in $V(G)$, $T \in N(x)$ s.t $|T| = t$

Clearly $|\beta| \leq t \cdot \binom{n}{t}$ (since $K_{t,t} \not\subset G$)

Also, $|\beta| = \sum_{x \in V} \binom{d(x)}{t} \geq n \cdot \binom{\varepsilon n}{t}$

$\therefore\ n \binom{\varepsilon n}{t} \leq t \binom{n}{t}$

$\Rightarrow 1 \leq \frac{t}{n \cdot \varepsilon^t} \left(1 - \frac{(t-1)}{\varepsilon n}\right)^{-t}$

observe that the inequality fails for $t > \lfloor \varepsilon \log n \rfloor$.

$\therefore$ if $t \geq \lceil \varepsilon \log n \rceil$, $K_{t,t} \subset G$ and

we are done since this is what we wanted

Assume that the theorem holds for $k = 1, 2, \ldots, r-1$

We shall prove it for $k = r$

$$\delta(G_n) \geq \left(1 - \frac{1}{r} + \varepsilon\right) n$$

$$\geq \left(1 - \frac{1}{r}\right) n$$

$$= \left(1 - \frac{1}{r-1} + \frac{1}{r(r-1)}\right) n$$

Letting $\varepsilon' = \frac{1}{r(r-1)}$,

$$K_r(t) \subseteq G \quad \text{for} \quad t' \geq \frac{\varepsilon' \log n}{2^{r-2}(r-2)!} = \frac{\lg n}{2^{r-2}\, r!}$$

Let $K_r(t') = K = V_1 \cup V_2 \cup \cdots \cup V_r$

consider $V \setminus K$

Let $U = \left\{ u \in V \setminus K \mid d(u, K) \geq \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|K| \right\}$

Now $e(K, V\setminus K) \geq |K| \left(\left(1 - \frac{1}{r} + \varepsilon\right)n - |K|\right)$.

since $\delta(G_n) \geq \left(1 - \frac{1}{r} + \varepsilon\right)$ and each vertex in K

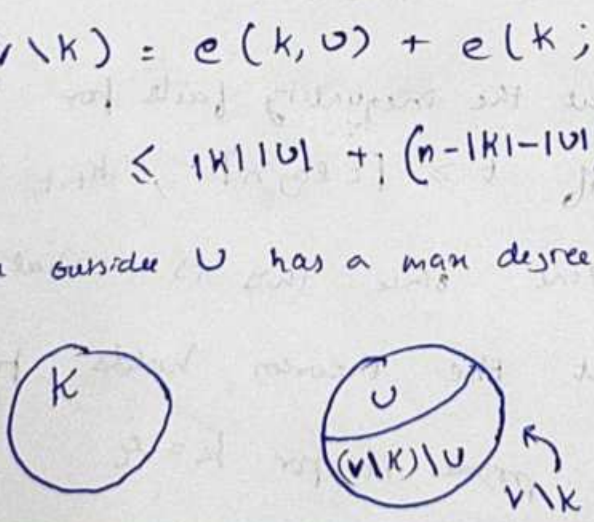
has degree at least $\left(1 - \frac{1}{r} + \varepsilon\right) n - |K|$.

Also, $e(K, V\setminus K) = e(K, U) + e(K; (V\setminus K)\setminus U)$

$$\leq |K||U| + (n - |K| - |U|)\left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|K|$$

since each vertex outside U has a max degree $\left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|K|$

Diagram:



$\therefore \left(1-\frac{1}{r}+\epsilon\right)n - |K| \le |U| + (n-|U|-|K|)\left(1-\frac{1}{r}+\frac{\epsilon}{2}\right)$

$\Rightarrow |U| \geq n\left(\frac{\epsilon}{2\left(\frac{1}{r}-\frac{\epsilon}{2}\right)}\right) - |K|$

$\Rightarrow |U| \geq \Omega_\epsilon(n)$ (ie U is large enough)

Now, each vertex in U has at least $\left(1-\frac{1}{r}+\frac{\epsilon}{2}\right)|K|$

neighbours in K

Now $\left(1-\frac{1}{r}+\frac{\epsilon}{2}\right)|K| = \left(1-\frac{1}{r}+\frac{\epsilon}{2}\right)rt'$

$= (r-1)t' + \frac{\epsilon}{2}rt' \geq \left(\frac{\epsilon r t'}{2}\right)r$

$\therefore$ each $u \in U$ has at least $\frac{\epsilon r t'}{2}$ neighbours in each V_i'

(If not, then neighbours of u in $K \leq \left(\frac{\epsilon}{2}rt'\right)r$)

THEOREM 7 (Erdős - Stone Theorem) (Strong)

Suppose $r \geq 1$, $0 < \varepsilon < \frac{1}{2}$, $n >> 0$ and if we have

$$e(G_n) \geq \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}, \text{ then } \exists\, K_{r+1}(t) \subseteq G$$

with $t = |V_i| \geq c_{r,\varepsilon} \log(n) = \dfrac{\varepsilon \log n}{2^{r+1}(r-1)!}$

proof :

Claim: $e(G_n) \geq (c+\varepsilon)\binom{n}{2} \Rightarrow G_n$ admits a subgraph G_m such that $m \geq \sqrt{\varepsilon}\, n$, $\delta(G_m) \geq cm$

proof of claim: suppose the subgraph G_n itself and it doesn't satisfy the conclusion. Then $\exists\, x_n \in V(G_n)$ s.t. $d(x_n) < cn$. Consider $G_{n-1} = G_n \setminus \{x_n\}$ and continue the argument. Suppose we reach G_m where all vertices have degree $\geq cm$, then $d(x_i) < c \times i$ $\forall\, i = m+1, m+2, \dots n$ (kicked out vertices).

$$\binom{m}{2} \geq e(G_m) \geq (c+\varepsilon)\binom{n}{2} - c\left[(m+1) + (m+2) + \dots + n\right]$$

$$= (c+\varepsilon)\binom{n}{2} - c\left(\binom{n}{2} - \binom{m}{2} + (n-m)\right)$$

$$= \varepsilon\binom{n}{2} + c\left(\binom{m}{2} - (n-m)\right)$$

$$\therefore \binom{m}{2} \geq \varepsilon\binom{n}{2} - cn + c\binom{m+1}{2}$$

At $m = \lfloor \varepsilon n \rfloor$ we get a contradiction

Thus we have proved our claim

for proving the strong Erdos Stone theorem, we use

$c = 1 - \frac{1}{r} + \frac{\varepsilon'}{2}$ and $\varepsilon = \frac{\varepsilon'}{2}$, we get that

$e(G_n) \geq \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2} \Rightarrow G_n \supset G_m$ with

$m \geq \sqrt{\varepsilon'}\, n$ and $\delta(G_m) \geq \left(1 - \frac{1}{r} + \varepsilon\right) m$. Here

onwards, the problem reduces to the weak Erdos Stone problem

THEOREM 8

$ex(n, C_k) =$ max no. of edges in C_k free graph $\geq c \cdot n^{1+\frac{1}{k-1}}$

(for some c (given k), $n \gg 0$)

Proof

Let $N(G)$ be the number of copies of C_k in G.

If C is the set of k-cycles,

$$|C| = \frac{n(n-1)\cdots(n-k+1)}{2k} = \frac{n!}{2k \cdot k!}$$

(writing $v_1 \dots v_k$ reverse or cycling it gives the same cycle and hence we divide by $2k$)

$$E(N(G_{n,p})) = \sum_{v_1, v_2, \dots v_k \in C} \mathbb{P}(v_1v_k, v_2v_3, \dots, v_{k-1}v_k, v_kv_1 \in E(G))$$

$$= \sum_{v_1 v_2 \dots v_k \in C} p^k$$

$$= \frac{n!\, p^k}{2k \cdot k!} \leq \frac{n!\, p^k}{2k} \leq \frac{(np)^k}{2k}$$

$E(e(G_{n,p})) = \binom{n}{2} \cdot p$

Thus given G, $\exists$ graph on same vertex set with $e(G) - N(G_{n,p})$ edges that is C_k free

$E(e(G_{n,p}) - N(G_{n,p})) \geq \binom{n}{2} p - \frac{(np)^k}{2k}$

Setting p = something appropriate $= \left(\frac{k}{2}\right)^{1/k-1} n^{-1+\frac{1}{k-1}}$

$\binom{n}{2} p - \frac{(np)^k}{2k} \geq C \cdot n^{1+\frac{1}{k-1}}$

$\left(\binom{n}{2} p - \frac{(np)^k}{2k} \geq \frac{n(n-1)}{4} p \right)$

THEOREM 9 (Stability theorem)

Given $\varepsilon > 0$, $\exists \delta > 0$ s.t. G_n is C_{2k+1} free and $e(G_n)$ is at least $\left(\frac{1}{4} - \delta\right) n^2 \Rightarrow G_n$ can be made bipartite by deleting at most εn^2 edges where $\varepsilon = o(\delta^2)$

Proof:

$e(G_n) \geq \left(\frac{1}{4} - \delta\right) n^2 \Rightarrow \delta(G_n) \geq \left(\frac{1}{2} - 2\sqrt{\delta}\right) n$

Also, $ex(n; C_{2k}) = \Omega\left(n^{1+\frac{1}{2k-1}}\right)$

Here, $e(G_n) = \Omega(n^2) \Rightarrow C_{2k} \subseteq G_n$

Let $K = \{x_1, x_2, \dots x_{2k}\}$ be the $2k$-cycle

Let $W = V \setminus K$

$N(x_i) \cap N(x_{i+1}) = \emptyset$ else $C_{2k+1} \subset G$

Let $A_i = N(x_i) \setminus K = N(x_i, W)$

$$|A_i| \geq \left(\frac{1}{2} - 2\sqrt{\delta}\right)n - 2k$$

$$\geq \left(\frac{1}{2} - \frac{5\sqrt{\delta}}{2}\right)n \quad \left(\because 2k \leq \frac{\sqrt{\delta}\, n}{2}\right)$$

Also notice that $G_1[A_i]$ cannot contain path of length $2k-1$ else we get a $2k+1$ cycle using x_i

$$\therefore e(A_i) \leq \frac{(2k+1-1)n}{2} = kn$$

Let $R = W \setminus (A_1 \cup A_2)$

$$\left\{ \therefore |R| \leq n - (1 - 5\sqrt{\delta})\, n \right\} = 5\sqrt{\delta}\, n$$

Now we ~~det~~ delete all edges inside A_1, A_2, all edges touching X and all edges touching R

Then (A_1, A_2) is our bipartite graph.

Total edges deleted $\leq 2kn + (2k)n + (5\sqrt{\delta} n)n$

$\therefore$ no. of edges deleted $= o(n^2)$

Definitions

1) Density: For $U, W \subseteq V(G_n)$. $d(U, W) := \dfrac{e(U, W)}{|U|\,|W|}$

2) ε-reg pair: A pair $(U, W) \subseteq V(G_n)^2$ is called ε-regular if we have $\forall\ |A| \geq \varepsilon |U|$, $|B| \geq \varepsilon |W|$,

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

3) Energy: $\underset{\subseteq (V(G))^2}{q(U, W)} := \dfrac{|U|\,|W|}{n^2}\, d^2(U, W) = \dfrac{e^2(U, W)}{n^2\,|U|\,|W|}$

4) Partition energy: $q(\mathcal{U}, \mathcal{W}) = \displaystyle\sum_{\substack{1 \leq i \leq m \\ 1 \leq j \leq \ell}} q(U_i, W_j)$

THEOREM 10 (Triangle counting lemma)

Let X, Y, Z be disjoint subsets of $V(G)$ with X, Y; Y, Z; Z, X being ε-regular. Suppose,

$d(X, Y) = d_{XY}$ is the notation for density. Then no. of triangles (x, y, z) with $x \in X$, $y \in Y$, $z \in Z$ is at least $(1-2\varepsilon)(d_{XY}-\varepsilon)(d_{YZ}-\varepsilon)(d_{ZX}-\varepsilon)|X||Y||Z|$

proof:

Let $X_{bad, Y} = \{ x \in X \mid d_Y(x) < (d_{XY}-\varepsilon)|Y| \}$

Then $|X_{bad, Y}| < \varepsilon |X|$. If not, then,

$|X_{bad, Y}| \geq \varepsilon |X|$

$\therefore e(X_{bad, Y}, Y) \geq \varepsilon |X| (d_{XY}-\varepsilon)|Y|$

$\therefore d(X_{bad, Y}, Y) \geq d_{XY} - \varepsilon$

This is a contradiction since X, Y is ε-reg pair

Similarly $|X_{bad, Z}| < \varepsilon |X|$

Let $X' = X \setminus (X_{bad, Y} \cup X_{bad, Z})$

Then $|X'| \geq |X|(1-2\varepsilon)$

Let, for $x \in X'$, $N_Y(x) = Y(x) = N(x) \cap Y$

Then $\left.\begin{array}{l} |Y(x)| \geq (d_{XY}-\varepsilon)|Y| \\ |Z(x)| \geq (d_{XZ}-\varepsilon)|Z| \end{array}\right\}$

Since for most vertices in X', $d_Y(x) \approx d_{XY}|Y|$

and $|d_{XY} - d_{X',Y}| < \varepsilon$

$$\Rightarrow \quad d_{XY} - \varepsilon < d(X', Y)$$

$$\Rightarrow \quad (d_{XY} - \varepsilon)\,|Y| < \frac{e(X', Y)}{|X'|}$$

But, $$|Y(x)| \geq \frac{e(X', Y)}{|X'|}$$

$$\therefore \; |Y(x)| \geq (d_{XY} - \varepsilon)\,|Y|$$

Also, we have,

$$e(Y(x), Z(x)) \geq (d_{YZ} - \varepsilon)\,|Y(x)|\,|Z(x)|$$

$\therefore$ no of triangles in G is at least

$$\underbrace{(1 - 2\varepsilon)|X|}_{|X'| \geq \cdots} \cdot \underbrace{(d_{XY} - \varepsilon)|Y||Z|(d_{YZ} - \varepsilon)\cdot(d_{ZX} - \varepsilon)}_{\text{no of } \Delta\text{'s corresponding to } x \in X'}$$

THEOREM 11 (Szemerdi's regularity lemma)

Given $\varepsilon > 0$, $\exists$ $M = M(\varepsilon) > 0$ such that for any graph G_n we have a partition $\{V_0, \ldots, V_k\}$ of the vertex set $V(G_n)$ such that

(i) $|V_0| \leq \varepsilon n$ is called the exceptional set

(ii) $|V_1| = |V_2| = \cdots = |V_k| = c$

(iii) almost all (V_i, V_j) are ε-regular ($1 \leq i, j \leq k$, at most εk^2 of them are not ε-regular)

(iv) $k \leq M(\varepsilon)$

Note: without (iv), we could just take $k = n$ and be done!

proof :

lemma 1 : $q(U, W) \leq q(\bar{U}, \bar{W})$ where $\bar{U}$ and $\bar{W}$ are partitions of U, W which are subsets of $V(G_n)$

proof : Define the r.v $Z(u, \omega) = \begin{cases} d(U', W') & * \\ 0 & \text{otherwise} \end{cases}$

where $*$ is $u \in U' \in \bar{U}$, $\omega \in W' \in \bar{W}$

Then $E(Z) = \sum_{\substack{U' \in \bar{U} \\ W' \in \bar{W}}} d(U', W') \dfrac{|U'||W'|}{|U||W|}$

$\uparrow \mathbb{P}(x \in U')$ $\uparrow \mathbb{P}(y \in W')$

$$= \frac{1}{|U||W|} \sum_{\substack{U' \in \bar{U} \\ W' \in \bar{W}}} e(U', W')$$

$$= \frac{e(U, W)}{|U|\,|W|} = d(U, W)$$

$$E(Z^2) = \sum_{\substack{U' \in \bar{U} \\ W' \in \bar{W}}} d^2(U', W') \frac{|U'||W'|}{|U||W|}$$

$$= \frac{n^2}{|U||W|} q(\bar{U}, \bar{W})$$

$$(E(Z))^2 = (d(U, W))^2 = \frac{n^2}{|U||W|} q(U, W)$$

$\Rightarrow q(\bar{U}, \bar{W}) \geq q(U, W)$ (variance ≥ 0)

lemma 2 : if (U, W) is ε-irregular, $\exists\ \bar{U} = \{U_1, U_2\}$ and $\bar{W} = \{W_1, W_2\}$ so that

$$q(\bar{U}, \bar{W}) \geq q(U, W) + \varepsilon^4 \frac{|U||W|}{n^2}$$

proof : $\exists\ U_1 \subseteq U$, $W_1 \subseteq W$ s.t. $|U_1| \geq \varepsilon |U|$ and $|W_1| \geq \varepsilon |W|$ but

$$|d(U_1, W_1) - d(U, W)| > \varepsilon$$

using the random variable Z defined previously,

$$\text{var}(Z) = E(Z^2) - (E(Z))^2$$

$$= \frac{n^2}{|U||W|} \cdot (q(\bar{U}, \bar{W}) - q(U, W))$$

we need to show : $\text{var}(Z) \geq \varepsilon^4$

we consider $\bar{U} = \{U_1, U \setminus U_1\}$, $\bar{W} = \{W_1, W \setminus W_1\}$

Var(Z) will consider U_1 $U \setminus U_1$, W_1 $W \setminus W_1$ but we can consider only U_1, W_1 and say $\text{var}(Z) \geq E((Z - d(U,W))^2) \times \frac{|U_1|}{|U|} \times \frac{|W_1|}{|W|}$

$$\geq \varepsilon^2 \times \varepsilon \times \varepsilon$$

$$= \varepsilon^4$$

lemma 3 : For $0 < \varepsilon < 1/4$, $P = (V_0, \dots, V_k)$, a partition of V, if P is such that $P \setminus \{V_0\}$ has at least εk^2 irregular pairs, $P \setminus \{V_0\}$ is equitable (blocks of same size t), and $|V_0| \leq \varepsilon n$, then we have a refinement $Q = (V_0', V_1', \dots, V_l')$ of P s.t. $l \leq k \cdot 4^k$, $Q \setminus \{V_0'\}$ is equitable,

$$|V_0'| \leq |V_0| + \frac{n}{2^k}, \quad q(Q) \geq q(P) + \frac{\varepsilon^5}{2}$$

proof: Let $|V_i| = t \quad \forall\ i = 1, 2, \dots k$

Let (V_i, V_j) be ε-irregular and let $(V_{i,1}, V_{i,2})$ and $(V_{j,1}, V_{j,2})$ be partitions of V_i, V_j as described in lemma 2. If $Q_1 = (P \cup \{V_{i,1}, V_{i,2}, V_{j,1}, V_{j,2}\}) \setminus \{V_i, V_j\}$ is this corresponding refinement of P, then,

$$q(Q_1) \geq q(P) + \varepsilon^4 \frac{|V_i||V_j|}{n^2}$$

(note: $q(P) := q(P, P)$ for a partition P)

Do the same process for all irregular pairs till we reach a ~~[illegible]~~ Q' = common refinement of all Q_i's

Then $q(Q') \geq q(P) + \varepsilon^4 \frac{t^2}{n^2} (\varepsilon k^2)$

$$= q(P) + \varepsilon^5 \frac{(tk)^2}{n^2}$$

$$\geq q(P) + \frac{\varepsilon^5}{2}$$

(since $tk = n - |V_0| \geq (1-\varepsilon) n \geq \frac{3n}{4}$ and $\frac{9}{16} \geq \frac{1}{2}$)

Now we create Q from Q'

Let V_i get broken down to $V_{i1}, V_{i2}, \dots V_{ip}$. Cut down every part of Q' into parts of size $\lfloor \frac{t}{4^k} \rfloor$. whatever remains and can't be cut, put it in V_0 as V_0'.

Size of any block of Q (other than V_0) is $b = \lfloor \frac{t}{4^k} \rfloor$.

Thus, no. of blocks $\leq \frac{n}{b} \leq \frac{n}{t} \times 4^k = k \cdot 4^k$

Also, energy of $Q = q(Q) > q(Q') > q(P) + \frac{\varepsilon^5}{2}$

Now we only need to verify $|V_0'|$

Q' has at most $k \cdot 2^{k-1}$ parts (breaking down process → at most $(k-1)$ 2 part partitions of each block of P & P had k blocks excluding the special V_0)

$\therefore |V_0'| \leq |V_0| + |Q'| b$ (size of all blocks of Q other than V_0' is b)

$$\leq |V_0| + k \cdot 2^{k-1} \cdot \frac{t}{4^k}$$

$$\leq |V_0| + \frac{n}{2 \cdot 2^k}$$

$$\leq |V_0| + \frac{n}{2^k}$$

$\boxed{}$

Now, we begin with a partition P_0 having k_0 parts with $2^{k_0} \geq 2/\varepsilon$. Given P_k, let P_{k+1} be the partition given by lemma 3. Since energy ≤ 1 always, the process will terminate after $\lceil 2/\varepsilon^5 \rceil$ steps. Invoking lemma 3 as above, we are done $\boxed{}$

Cleaning process

Given a graph G_n, do the following

(i) given $\varepsilon > 0$, find the Szemeredi partition

(ii) Delete all edges between irregular pairs

(iii) Delete all edges between 'sparse' pairs i.e. pairs with density lower than ε.

(iv) Delete all edges inside V_0

At most, we delete $(\varepsilon k^2) t^2 + (\varepsilon t^2) \frac{k^2}{2} + \frac{\varepsilon^2}{2} n^2$ number of edges

And hence no. of deleted edges $\leq 2 \varepsilon n^2$.

The resulting graph has all pairs ε-regular

THEOREM 12 (triangle removal lemma)

$\forall \varepsilon > 0, \exists \delta > 0$ s.t. for $n \gg 0$, any graph G_n with at most δn^3 triangles can be made triangle free by deleting at most εn^2 edges

proof:

Start with $\varepsilon/4$ and obtain the Szemeredi partition. Apply the cleaning process. Further, delete all edges within every V_i. We lost additional $k \cdot \binom{t}{2}$ edges at max. $= \frac{k t^2}{2} = \frac{k^2 t^2}{2k} \leq \frac{\varepsilon}{4} n^2$

Thus we lost $\frac{\varepsilon}{4}n^2 + \frac{\varepsilon}{4}n^2$ edges.

Loosely, at most εn^2 edges.

Thus, if there is a triangle, it must be generated with (V_i, V_j, V_k) where all 3 pairs are $\frac{\varepsilon}{4}$-reg and density is at least $\frac{\varepsilon}{2}$

Using the triangle counting lemma, we have at least $t^3 \left(\frac{\varepsilon}{4}\right)^3 \left(1-\frac{\varepsilon}{2}\right) > \frac{n^3 \left(\frac{\varepsilon}{8}\right)^3}{(M(\varepsilon))^3}$ triangles

Letting $\frac{\delta}{2} = \frac{1}{6} \frac{\left(\frac{\varepsilon}{8}\right)^3}{M(\varepsilon)^3}$, we are done

in case ijk not distinct

since $\geq 2\delta n^3$ Δ's will contradict the hypothesis

THEOREM 13 (Graph counting lemma)

Let G_n be a graph and H be a graph on $[k]$. Let $V_1, V_2, \ldots, V_k$ be a partition of $V(G)$ with (V_i, V_j) being ε-regular whenever $ij \in E(H)$. ~~Then, if~~ Let

$$S = \{ (v_1, \ldots, v_k) \mid v_i \in V_i, \ v_1 \ldots v_k \text{ form a copy of } H \text{ in } G \}$$

Then $$\left| |S| - \left(\prod_{i=1}^{k} |V_i|\right)\left(\prod_{ij \in E(H)} d(V_i, V_j)\right) \right| \leq \varepsilon\, e(H) \prod_{i=1}^{k} |V_i|$$

proof:

We induct on $e(H)$. If $e(H)=0$, nothing to show

We wish to show,

$$\left| \frac{|S|}{\prod_{i \in [k]} |V_i|} - \prod_{ij \in E(H)} d(V_i, V_j) \right| \leq \varepsilon\, e(H)$$

i.e. If $v_i \in V_i$ is picked independently & randomly,

$$\left| \mathbb{P}(v_i v_j \in E(G)\ \forall\ ij \in E(H)) - \prod_{ij \in E(H)} d(V_i, V_j) \right| \leq \varepsilon\, e(H)$$

WLOG let $12 \in E(H)$

It will suffice to show that

$$\left| \mathbb{P}(v_i v_j \in E(G)\ \forall\ ij \in E(H)) - d(V_1, V_2)\, \mathbb{P}(v_i v_j \in E(G)\ \forall\ ij \in E(H) \setminus \{(1,2)\}) \right| \leq \varepsilon \quad \text{— ①}$$

(Why? will be used in the induction step later on. Towards the end we add and subtract $\mathbb{P}_{H'}$. See last few lines of next page)

Let $A_1 = \{ v_1 \in V_1 \mid v_1 v_i \in E(G)\ \forall\ i \in N_H(1) \setminus \{2\} \}$

$A_2 = \{ v_2 \in V_2 \mid v_2 v_i \in E(G)\ \forall\ i \in N_H(2) \setminus \{1\} \}$

If $|A_1| \leq \varepsilon |V_1|$ or $|A_2| \leq \varepsilon |V_2|$, then

$$\frac{e(A_1, A_2)}{|V_1||V_2|} \leq \frac{|A_1||A_2|}{|V_1||V_2|} \leq \varepsilon$$

$$\Rightarrow d(V_1, V_2) \frac{|A_1||A_2|}{|V_1||V_2|} \leq \varepsilon \qquad (\because d \leq 1)$$

$$\therefore \left| \frac{e(A_1, A_2)}{|V_1||V_2|} - d(V_1, V_2)\frac{|A_1||A_2|}{|V_1||V_2|} \right| \leq \varepsilon$$

If $|A_1| > \varepsilon |V_1|$, $|A_2| > \varepsilon |V_2|$,

$$\left| \frac{e(A_1, A_2)}{|V_1||V_2|} - d(V_1, V_2)\frac{|A_1||A_2|}{|V_1||V_2|} \right|$$

$$= \left| \frac{e(A_1, A_2)}{|A_1||A_2|} - d(V_1, V_2) \right| \cdot \frac{|A_1||A_2|}{|V_1||V_2|}$$

$$\leq \varepsilon \cdot 1$$

$$\therefore \left| \frac{e(A_1, A_2)}{|V_1||V_2|} - d(V_1, V_2)\frac{|A_1||A_2|}{|V_1||V_2|} \right| < \varepsilon$$

This is equivalent to 1

[Why? We fixed the other vertices in other components $V_3, V_4, \ldots$ etc and we are now probabilistically looking at v_1, v_2 such that they satisfy the conditions]

Now remove 12 from H and call it H'.

Let $\mathbb{P}_H = \mathbb{P}(\{v_i, v_j\} \in E(G) \ \forall \ ij \in E(H))$

Then,

$$\left| \mathbb{P}_H - \prod_{ij \in E(H)} d(V_i, V_j) \right| \leq d(V_1, V_2) \left| \mathbb{P}_{H'} - \prod_{ij \in E(H')} d(V_i, V_j) \right|$$

$$+ \left| \mathbb{P}_H - d(V_1, V_2)\, \mathbb{P}_{H'} \right| \cdot \leq \overbrace{d(V_1, V_2)\, e(H') \cdot \varepsilon}^{\text{induc hypo}} + \varepsilon$$

$$\leq (e(H') + 1)\,\varepsilon = e(H)\,\varepsilon$$

THEOREM 14 (Graph Removal lemma)

Given $\varepsilon > 0$ and any graph H, $\exists\ \delta = \delta(\varepsilon)$ s.t for $n \gg 0$, any graph G_n with at most $\delta n^{v(H)}$ copies of H, can be made ~~___~~ H-free by deleting at most εn^2 edges

Proof: Let $v(H) = h$. Choose a $\frac{\varepsilon^h}{4h}$ - regular partition of $v(G)$ using Szemeredi lemma. Choose the δ as $\delta = (2h)^{-2h} \varepsilon^m K^{-h}$ where $m = e(H)$ and K is the upper bound on no of parts in the partition as given by Szemeredi lemma.

If $n < \delta^{-1/h}$, no. of copies of H in G is at most $\delta n^h < 1$ and G is already H free.

So, if $n \geq \delta^{-1/h}$.

Apply a cleaning as follows -

$\rightarrow$ delete edges between $\frac{\varepsilon^h}{4h}$ - irregular pairs

at most $\left(\gamma \frac{k^2}{2}\right)\left(\frac{2n}{k}\right)^2$ edges lost $= 2\gamma n^2$

(denoting $\frac{\varepsilon^h}{4h} = \gamma$)

$\rightarrow$ delete edges between pairs with density $< \varepsilon$

at most $\frac{\varepsilon n^2}{2}$ edges lost

$\therefore$ Total no of edges lost $= 2\gamma n^2 + \frac{\varepsilon n^2}{2} < \varepsilon n^2$

$(\because \gamma \le \frac{\varepsilon}{4})$

claim: G is now H free.

Suppose not. Then a copy of H in the cleaned graph G' will have edges b/w parts which are pairwise γ-regular and having density at least ε.

By the graph counting lemma,

$$\text{no. of copies of H in G} \ge 2^{-h}\varepsilon^{m}\left(\frac{n}{2k}\right)^{h}$$

$$> h!\,\delta n^{h}$$

This contradicts that copies of H in G is at most δn^h

$\boxtimes$

THEOREM 15

If every edge of G_n is in exactly 1 triangle, then $e(G_n) = o(n^2)$ (Better description: given $\varepsilon > 0$, $\exists\, n$ so that G_n is a graph for which every edge is in 1 triangle & $|e(G_n)| \le \varepsilon n^2$

proof:

number of $e(G_n) = 3 \times$ no. of triangles $\le \frac{n^2}{2}$

$\therefore$ no. of triangles $= o(n^2) \le o(n^3)$

$\therefore$ By triangle removal lemma, we need to remove $o(n^2)$ edges to make it triangle free

$$\therefore\ o(n^2) = \frac{e(G_n)}{3}$$

$$\Rightarrow e(G_n) = o(n^2)$$

THEOREM 16 (No corner theorem)

Suppose $A \subseteq [N]^2$ has no corners, then $|A| = o(N^2)$.

(where $A \subseteq [N]^2$ has a corner if (x,y), $(x+d,y)$, $(x, y+d) \in A$ for some $d > 0$, $x, y \in N$)

proof:

$A + A = \{ a+b : a, b \in A \} \subseteq [2N]^2$

$x + A = \{ x+a : a \in A \}$

We want to get rid of condition on d being positive

$\exists\ z \in [2N]^2$ s.t. $z = a+b$ in at least $\frac{|A|^2}{(2N)^2}$ ways (pigeonhole)

Let $A' = A \cap (z - A)$

Then $|A'| \geq \frac{|A|^2}{(2N)^2}$

(if $m \in A'$, then $m = a = z - b \Rightarrow z = a+b$)

Suffices to show $|A'| = o(N^2)$

. A' is symmetric about $z/2$ since $A' = z - A'$

Thus, if λ' has a $\llcorner$ corner, it also has $\urcorner$ corner & hence we relaxed $d < 0$

Want to embed N^2 lattice into graph so that triangles correspond with corners

Build a tripartite graph (X, Y, Z) where $X = [n]$, $Y = [N]$, $Z = [2N]$.

X enumerates all vertical lines, Y is horizontal, Z is diagonal lines. (slope -1)

join vertices if their corresponding lines meet inside A

i.e. $(x, z-x) \in A \iff$ edge between $x \in X$, $z \in Z$

$(x, y) \in A \iff$ edge between $x \in X$, $y \in Y$

$(z-y, y) \in A \iff$ edge between $z \in Z$, $y \in Y$

no of edges in this graph $= 3|A|$

(every element of A gives 3 edges)

triangles correspond to corners

But A is triangle free & hence there must be trivial corners only.

Also notice that each edge is exactly in one unique triangle (given x, y we can find z ~~or~~ $y, z \to x$ or $x, z \to y$)

$\therefore$ By theorem 15, $3|A| = o(N^2) \Rightarrow |A| = o(N^2)$ $\blacksquare$

THEOREM 17 (Roth's theorem)

Given $\varepsilon > 0$, $\exists$ $N_0 \in \mathbb{N}$ s.t. $\forall n > N_0$, $A \subseteq [n]$,

$|A| \geq \varepsilon n$ $\Rightarrow$ A contains a non trivial 3 term arithmetic progression

proof:

Let $A \subseteq [N]$ have no 3-AP. We will show $|A| = o(N)$

Let A be embedded into add cyclic group $\mathbb{Z}/M\mathbb{Z}$ with $M = 2N+1$. ie. $A \subseteq \mathbb{Z}/M\mathbb{Z}$

Construct tripartite $G = (X, Y, Z)$ where X, Y, Z are all defined on $\mathbb{Z}/M\mathbb{Z}$ and the edges are given by (for notation, $x \in X$, $y \in Y$, $z \in Z$)

edge between x and y iff $y - x \in A$

edge between y and z iff $z - y \in A$

edge between z and x iff $\frac{z-x}{2} \in A$

(note: 2 is invertible in $\mathbb{Z}/M\mathbb{Z}$ since M is odd)

If $x\,y\,z$ form a triangle, then

$$y - x,\ \frac{z-x}{2},\ z - y \in A$$

$\therefore$ we have 3-AP in A

Then this has to be trivial

$\therefore$ $y - x = \frac{z-x}{2} = z - y$

Solving, we get $2y = x + z$

$\therefore x, y, z$ form AP in $\mathbb{Z}/M\mathbb{Z}$

$\therefore$ given x, y, z can be found and given yz, x can be found & so on.

$\therefore$ Every edge of G lies in exactly one triangle.

$\therefore$ By theorem 15, we have $o(M^2)$ edges

(technically $o((3M)^2) = o(M^2)$)

But by construction,

no. of edges in $G = 3M|A|$

$$\therefore 3M\,|A| = o(M^2)$$

$$\Rightarrow |A| = o(M)$$

$$= \underline{\underline{o(N)}}$$

$\blacksquare$

(since every element of A generates M edges between XY, M between YZ, M between XZ)

THEOREM 18 (Erdős Stone Simonovits theorem)

Fix $t \in \mathbb{N}$, $\epsilon > 0$. Assume $n \gg 0$.

If G_n is $K_{r+1}(t)$ free, then,

$$e(G_n) \leq \left(1 - \frac{1}{r}\right)\frac{n^2}{2} + \epsilon n^2$$

Proof:

Suppose $e(G_n) > (1-\frac{1}{r})\frac{n^2}{2} + \varepsilon n^2$.

Let $\delta > 0$ (TBD later)

Get an δ-regular Szemeredi partition with

$|V_0| \leq \delta n$, $|V_i| = l \;\; \forall\, i = 1, 2, \dots, k$

and at most δk^2 pairs are δ-irregular

Clean the graph as follows

→ delete edges within all V_i (at most $l^2 \frac{k}{2}$, edges lost)

→ delete edges related to V_0 (at most $(\delta n)\, n$ edges lost)

→ delete edges in δ-irregular pairs (at most $\delta k^2 l^2$)

→ delete edges b/w pairs of density lower than δ (at most $l\delta(\frac{k^2}{2})$ edges lost)

$\therefore$ total edges lost $\leq 4\delta n^2$ ($\because kl = n$)

$\therefore$ cleaned graph has at least $(1-\frac{1}{r})\frac{n^2}{2} + (\varepsilon - 4\delta)n^2$ edges left

Let H be a graph with vertex set $\{V_i : 1 \leq i \leq k\}$ and $V_iV_j \in E(H)$ iff (V_i, V_j) is $\frac{\varepsilon}{8}$-regular with density at least $\varepsilon/4$

By pigeonhole principle and $e(V_i, V_j) \le \ell^2$

$$\Rightarrow e(H) \geq \left(1 - \frac{1}{r} + \varepsilon\right) \frac{n^2}{2\ell^2} = \left(1 - \frac{1}{r} + \varepsilon\right) \frac{k^2}{2}$$

(why?

By Turan's theorem, $H \supseteq K_{r+1}$. Say, $(V_1 \cdots V_{r+1})$ form this K_{r+1}. All (V_i, V_j) are $\frac{\varepsilon}{8}$-reg with density at least $\varepsilon/4$

from graph counting lemma, for $F = K_{r+1}(t)$, no. of copies of F in $(V_1 \cdots V_{r+1})$ is at least $\prod_{i=1}^{r+1} |V_i|^t \left(\left(\frac{\varepsilon}{4}\right)^{t^2\binom{r+1}{2}} - \delta t^2 \binom{r+1}{2} \right)$

So if $\delta \le \dfrac{\left(\frac{\varepsilon}{4}\right)^{t^2\binom{r+1}{2}}}{2t^2\binom{r+1}{2}}$, then there are many copies of $K_{r+1}(t)$ in G.

Thus, G is not $K_{r+1}(t)$ free & we get a contradiction. $\square$

THEOREM 19 (Frankl Pach Theorem)

$$ex(n, K_{r+1}, K_{s,\lceil cn\rceil}) \leq \left(c^{1/s}\left(1-\frac{1}{r}\right)^{1/s} + o(1)\right)n^2$$

$(s, r \in \mathbb{N},\ 0<c<1,\ n \gg 0)$

proof

Let $\varepsilon > 0$. Use Szemeredi with ε and apply cleaning to get graph G' s.t.

- $V(G') = V_1 \cup \dots \cup V_k$; $k = O_\delta(1)$
- $|V_i| = \ell\ \forall i$, $V(G') \geq (1-\delta)n$
- all pairs (V_i, V_j) are ε regular
- if $e(V_i, V_j) > 0$ then $d(V_i, V_j) \geq \eta$
- $e(G) - e(G') \leq \varepsilon n^2$;

Look at the telescopic graph i.e. construct graph H on $[k]$ with $ij \in E(H)$ iff $d(V_i, V_j) \geq \eta$.

By graph counting, $K_{r+1} \subseteq H \Rightarrow K_{r+1} \subset G$

and hence $K_{r+1} \subseteq G$. Thus H must be K_{r+1} free and by Turan's theorem, $e(H) \leq \left(1-\frac{1}{r}\right)\frac{k^2}{2}$

If $K_{s,\lceil cn\rceil}$ appears in G, one of the V_i's must have at least s vertices since $k = O_\delta(1)$, $n \gg 0$ (pigeonhole)

We count the number of pairs (x, S) where $|S| = s$, $S \subseteq V_i$ for some i and x is adjacent to every $y \in S$. We call such a structure an (i,s)-class.

$$\text{no. of such } (x, S) = \sum_{x \in V(G')} \sum_{i=1}^{k} \binom{|N(x) \cap V_i|}{s}$$

But G' (and G) is $K_{s, \lceil cn \rceil}$-free

$$\therefore \text{ no. of } (x, S) \leq (\lceil cn \rceil - 1) \sum_{i=1}^{k} \binom{|V_i|}{s}$$

$$\leq cnk \binom{l}{s}$$

But
$$\sum_{(x,i) \in P} \binom{|N(x) \cap V_i|}{s} \geq |P| \binom{\frac{1}{|P|} \sum_{(x,i) \in P} |N(x) \cap V_i|}{s}$$

(Jensens inequality)

$$= |P| \binom{2e(G') \times \frac{1}{|P|}}{s}$$

$$= |P| \binom{u}{s}$$

where $u = \frac{2e(G')}{|P|}$

and $P = \{ (x, i) \mid |N(x) \cap V_i| > 0 \}$

$\therefore \quad cnk \frac{l^s}{s!} \geq cnk \binom{l}{s} \geq |P| \binom{u}{s}$

$$\geq |P| \frac{(u-s+1)^s}{s!}$$

$\therefore \quad e(G') = \frac{u|P|}{2} \leq \frac{1}{2}\left((cnk)^{1/s} l \, |P|^{1-1/s} + |P|(s-1) \right)$

Now

$$|P| \leq e(H) \cdot 2l$$

since each edge in H gives at most $2l$ choices for x (edge ~~is~~ ij then $x \in V_i$ or V_j ?)

$$\therefore |P| \leq \left(1 - \frac{1}{r}\right) \frac{k^2}{\cancel{2}} \cdot \cancel{2} l$$

Finally,

$$e(G') \leq \left(c^{1/s} \left(1 - \frac{1}{r}\right)^{1/s} + o(1) \right) \frac{n^2}{2}$$

since $\quad e(G) - e(G') \leq \varepsilon n^2$,

our result follows.

THEOREM 20 (Pseudorandomness - Chung, Graham, Wilson theorem)

The following six notions of a pseudo random graph are equivalent.

Consider the Erdős Renyi graph $G_n(p)$ where each edge occurs independently with probability $p \in (0,1)$.

Let G_n have $p \cdot \binom{n}{2} + o(1)\binom{n}{2}$ edges

1) Disc : $|e(X,Y) - p|X||Y|| = o(n^2)$

$\forall X, Y \subseteq V(G)$

2) Disc' : $|e(X) - p\binom{|X|}{2}| = o(n^2)$

$\forall X \subseteq V(G)$

3) Count : The no. of labelled copies of H in a graph G is $(1+o(1))\, p^{e(H)} n^{|V(H)|}$

4) C_4-count : The no. of C_4 graphs (labelled) in G is $(p^4 + o(1))\, n^4$

5) Codegree : If codeg (u,v) denotes no. of common neighbours of u, v,

$$\sum_{x,y \in V} |\text{codeg}(x,y) - p^2 n| = o(n^3).$$

6) Eigen : If $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are evals of G_n, then $\lambda_1 = pn(1+o(1))$, $\max_{i=2,3,\dots,n} |\lambda_i| = o(n)$

proof:

we establish equivalence in the following manner

$$\begin{tikzcd} 2 \arrow[r, shift left] & 1 \arrow[l, shift left] \arrow[r] & 3 \arrow[r] & 4 \arrow[r] \arrow[d, shift left] & 5 \arrow[lll, bend right] \\ & & & 6 \arrow[u, shift left] & \end{tikzcd}$$

$(1 \Rightarrow 2)$: using $X = Y$,

$$|e(X,X) - p|X|^2| = o(n^2)$$

$$\therefore \left| \frac{e(X,X)}{2} - \frac{p|X|^2}{2} \right| = o(n^2)$$

$$\therefore \left| e(X) - p\binom{|X|}{2} \right| = o(n^2)$$

(since in $e(X,X)$ we count edges twice)

$(2 \Rightarrow 1)$: $e(X,Y) = e(X \cup Y) + e(X \cap Y) - e(X \setminus Y) - e(Y \setminus X)$

$$\therefore e(X,Y) = p\left[\binom{|X \cup Y|}{2} + \binom{|X \cap Y|}{2} - \binom{|X \setminus Y|}{2} - \binom{|Y \setminus X|}{2}\right] + o(n^2)$$

$$= p|X||Y| + o(n^2)$$

$$\therefore |e(X,Y) - p|X||Y|| = o(n^2)$$

$(1 \Rightarrow 3)$: nothing but the graph counting lemma

$(3 \Rightarrow 4)$: Put $e(H) = v(H) = 4$ in count to get C_4-count.

$(4 \Rightarrow 5)$:

Assume C_4

$$\sum_{u,v \in V} \text{codeg}(u,v) = \sum_{x \in V} (\deg(x))^2$$

$$\geq \frac{1}{n} \left(\sum_{x \in V} (\deg x) \right)^2$$

$$= \frac{4\, e(G))^2}{n} = (1 + o(1))\, p^2 n^3$$

$$\left(\because e(G_n) = p\binom{n}{2} + o(1)\binom{n}{2} = p\binom{n}{2} + o(n^2) \right)$$

$$\sum_{u,v \in V} (\text{codeg}(u,v))^2 = \text{no. of labelled copies of } C_4 \text{ in } G + o(n^4)$$

$$\leq (p^4 + o(1))\, n^4$$

since

$$\sum (\text{codeg})^2 \geq \frac{1}{n^2} (\sum \text{codeg})^2$$

$$= \frac{1}{n^2} (\sum (\deg)^2)^2$$

$$\geq \frac{1}{n^2} \left(\frac{1}{n} (\sum \deg)^2 \right)^2$$

$$= \frac{1}{n^4} ((pn^2)^2)^2$$

$$= p^4 n^4$$

(given G has $e(G) \geq p\frac{n^2}{2}$)

$\therefore \sum_{u,v} |\text{codeg}(u,v) - p^2 n|$

$$\leq n \left(\sum_{u,v} (\text{codeg}(u,v) - p^2 n)^2 \right)^{1/2}$$

$$= n \left(\sum_{u,v} \text{codeg}^2 - 2p^2 n \sum \text{codeg} + p^4 n^2 \times n^2 \right)^{1/2}$$

$$\leq n \left(p^4 n^4 - 2p^2 n \cdot p^2 n^3 + p^4 n^4 + o(n^4) \right)^{1/2}$$

$$= o(n^3)$$

$(5 \Rightarrow 1)$: we know that

$$\sum_{u \in V} |\deg(u) - pn| \leq n^{1/2} \left(\sum (\deg(u) - pn)^2 \right)^{1/2}$$

$$= n^{1/2} \left(\sum \deg^2 - 2pn \sum \deg + p^2 n^3 \right)^{1/2}$$

$$= n^{1/2} \left(\underbrace{\sum \text{codeg}}_{\text{using } 5} - 4pn\, e(G) + p^2 n^3 \right)^{1/2}$$

$$= n^{1/2} \left(p^2 n^3 - 2p^2 n^3 + p^2 n^3 + o(n^3) \right)^{1/2}$$

$$= o(n^2)$$

$$\therefore |e(X,Y) - p|X||Y|| = \left| \sum_{x \in X} \deg(x,Y) - p|Y| \right|$$

$$\leq n^{1/2} \left(\sum_{x \in X} (\deg(x,Y) - p|Y|)^2 \right)^{1/2}$$

$$\cancel{\leq n^{1/2} \left(\sum_{x \in V} \deg^2(x,Y) - 2p|Y| \sum_{x \in} \right.}$$

$\le \sum n^{1/2} \left(\sum_{x \in V} (\deg(x, Y)^2) - 2p|Y| \sum_{x \in V} \deg(x, Y) + p^2 n |Y|^2 \right)^{1/2}$

$= n^{1/2} \left(\sum_{y, y' \in Y} \mathrm{codeg}(y, y') - 2p|Y| \sum_{y \in Y} \deg y + p^2 n |Y|^2 \right)^{1/2}$

using 5

$= n^{1/2} \left(|Y|^2 p^2 n - 2p|Y| \cdot |Y| p n + p^2 n |Y|^2 + o(n^3) \right)^{1/2}$

$= o(n^2)$

$(6 \Rightarrow 4)$:

no. of labelled copies in C_4 = no. of closed walks of length 4.

$= \mathrm{tr}(A_4^4)$

$= \sum_{i=1}^{n} \lambda_i^4$

$= \lambda_1^4 + \sum_{i=2}^{n} \lambda_i^4$

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

Now $\sum_{i \geq 2} \lambda_i^4 \leq \max_i |\lambda_i|^2 \sum_{i \geq 2} \lambda_i^2$

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

closed walks of length 2 are just degrees

$= o(n^2) \cdot 2e(G)$

$= o(n^4)$

$\therefore$ no of labelled copies of $C_4 = p^4n^4 + o(n^3) + o(n^4)$

$$= p^4 n^4 + o(n^4)$$

$(4 \Rightarrow 6)$:

$$\lambda_1 = \sup_{\|x\|=1} x^T A x$$

$$\geq \frac{2e(G)}{n} \qquad \left(\because \text{using } \vec{x} = \frac{1}{\sqrt{2}}(1,1,1,\dots,1)\right)$$

$$= (p + o(1))\, n$$

from using C_4 count,

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

$$\therefore \lambda_1 \leq pn + o(n)$$

But $\lambda_1 \geq pn + o(1)\, n$

$$\therefore \lambda_1 = pn + o(1)$$

Also,

$$\max_{i \neq 1} |\lambda_i|^4 \leq tr(A^4) - \lambda_1^4$$

$$\leq p^4n^4 - p^4n^4 + o(n^4)$$

$$= o(n^4)$$

$$\therefore \max_{i \neq 1} |\lambda_i| = o(n)$$

THEOREM 21 (Expander mixing lemma)

Suppose G is a d-regular graph and $\lambda = \max\{|\lambda_2|, |\lambda_n|\}$ where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the e. vals of G_n. Then,

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

proof: $\lambda = \max\{|\lambda_2|, |\lambda_n|\} = \max\limits_{i \neq 1} |\lambda_i|$

Notice that $\lambda_1 = d$ ($\because$ d-regular graph)

Let J be the all 1's $n \times n$ matrix.

Notice that $e(X,Y) = 1_X^T A \, 1_Y$

and $|X||Y| = 1_X^T J \, 1_Y$

where 1_X is the vector having 1's in vertex positions described by vertex set X

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

$$= \left| 1_X^T \left(A_G - \frac{d}{n} J \right) 1_Y^T \right|$$

$$\leq \underbrace{\left\| A_G - \frac{d}{n} J \right\|}_{\text{matrix norm}} \underbrace{\|1_X\| \, \|1_Y\|}_{\text{vector norms}}$$

$$= \left\| A_G - \frac{d}{n} J \right\| \sqrt{|X||Y|}$$

$A_G - \frac{d}{n} J$ has same e.vecs as A_G ~~(...)~~

~~$A_G - \frac{d}{n} J$ ...~~ This is because

$\mathbb{1}$ is an e.vec of A_G with eval d and is also an e.vec of $A_G - \frac{d}{n} J$ with eval 0.

If $v \neq \mathbb{1}$ is an evec of A_G, then $v \perp \mathbb{1}$

i.e. $v \cdot \mathbb{1} = \sum_{i=1}^{n} v_i = 0 \Rightarrow Jv = 0 \Rightarrow v$ is also evec of $A_G - \frac{d}{n} J$ with same eval

$\therefore$ evals of $A_G - \frac{d}{n} J$ are $0, \lambda_2, \lambda_3, \ldots, \lambda_n$

$\therefore$ largest |e.val| of $A_G - \frac{d}{n} J = \lambda$

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

$$\leq \lambda \sqrt{|X||Y|}$$

GRAPH LIMITS

Definitions:

- A graphon is a symmetric measurable function $W : [0,1]^2 \to [0,1]$
- Given a graph G, the associated graphon (considering labelled vertices of $G - 1, 2, \ldots, n$) is $W_G : [0,1]^2 \to [0,1]$ obtained by partitioning $[0,1]$ as $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\}$ such that if $(x,y) \in \left[\frac{i-1}{n}, \frac{i}{n}\right] \times \left[\frac{j-1}{n}, \frac{j}{n}\right]$, then $W(x,y) = 1$ provided i, j are connected in G and zero otherwise

- Hom (H, G) : set of all homomorphisms from H to G

- we define homomorphism density

$$t(H, G) = \frac{|\text{Hom}(H, G)|}{|V(G)|^{|V(H)|}}$$

- we go on to define $t(H, W) = \int_{[0,1]^h} \prod_{ij \in E(H)} W(x_i, x_j) \, dx_1 \cdots dx_h$

where $h = |V(H)|$

- A sequence of graphs $\{G_n\}$ converges if $t(H, G_n)$ converges as n goes to ∞, for every graph H.

- we define a cut norm of $W : [0,1]^2 \to [0,1]$ as

$$\|W\|_{\square} = \sup_{\substack{S, T \subseteq [0,1] \\ S, T \text{ measurable}}} \left| \int_{S \times T} W \right|$$

- Some more norms:

L^p norm : $\|W\|_p = \left(\int |W|^p \right)^{1/p}$

L^∞ norm : $\inf_{m \in \mathbb{R}} \left\{ m \mid \{(x, y) \mid W(x, y) > m\} \text{ has measure zero} \right\}$

- We also define an isomorphism density for graphons as

$$d(H, W) = \frac{\overbrace{|V(H)|}^{=h}!}{|\text{Aut } H|} \int_{[0,1]^h} \prod_{ij \in E(H)} W(x_i, x_j) \times \prod_{ij \notin E(H)} (1 - W(x_i, x_j)) \, dx_1 dx_2 \cdots dx_h$$

THEOREM 22

From a norm, we can get a distance and vice versa. For our cut norm, the cut distance is

$$d_\square(U,W) = \sup_{\substack{S,T \subseteq [0,1] \\ S,T \text{ measurable}}} \left| \int_{S\times T} W - U \right|$$

The theorem states that

$$d_\square(U,W) = \sup_{\substack{a,b:[0,1]\to[0,1] \\ a,b \text{ Borel measurable}}} \left| \int_{[0,1]^2} a(x)\, b(y) \left(W(x,y) - U(x,y)\right) dx\,dy \right|$$

or equivalently,

$$\|W\|_\square = \sup_{\substack{a,b:[0,1]\to[0,1] \\ a,b \text{ measurable}}} \left| \int_{[0,1]^2} a(x)\,b(y)\, W \right|$$

proof:

We show (quite easily) that the two norms are equal

$$\|W\|_\square := \sup_{\substack{S,T \subseteq [0,1] \\ S,T \text{ measurable}}} \left| \int_{S\times T} W \right|$$

$$= \int_{[0,1]^2} W \quad \text{since } W \text{ is non negative}$$

$$\sup_{\substack{a,b:[0,1]\to[0,1] \\ a,b \text{ measurable}}} \int_{[0,1]^2} a(x)\, b(y)\, W = \int_{[0,1]^2} W = \|W\|_\square$$

(since $|a(x)\,b(y)| \le 1 \quad \forall x,y \in [0,1]^2$) ∎

THEOREM 23

For graphons W, U and graph H, we have,

$$\left| t(H, W) - t(H, U) \right| \le |E(H)| \cdot d_\square(W, U)$$

proof:

let $V(H) = \{1, 2, \dots, n\}$

· WLOG, let $\{1,2\} \in E(H)$

let $x_3, \dots, x_n \in [0,1]$

Define
$$a(x) = \prod_{\substack{i \geq 3 \\ 1,i \in E(H)}} W(x, x_i)$$

$$b(x) = \prod_{\substack{i \geq 3 \\ 2,i \in E(H)}} W(x, x_i)$$

$$c = \prod_{\substack{i,j \geq 3 \\ ij \in E(H)}} W(x_i, x_j)$$

Then,
$$\int_{[0,1]^2} \left(W(x_1, x_2) - U(x_1, x_2) \right) \prod_{\substack{ij \in E(H) \\ ij \neq 12 \text{ or } 21}} W(x_i, x_j) \, dx_1 dx_2$$

$$= \int_{[0,1]^2} \left(W(x_1, x_2) - U(x_1, x_2) \right) a(x_1) \, b(x_2) \, c \, dx_1 dx_2$$

$$= c \int_{[0,1]^2} a(x_1) \, b(x_2) \left(W(x_1, x_2) - U(x_1, x_2) \right) dx_1 dx_2$$

$$\therefore \left| \int_{[0,1]^2} \left(W(x_1, x_2) - U(x_1, x_2) \right) \prod_{\substack{ij \in H \\ ij \neq 12 \text{ or } 21}} W(x_i, x_j) \, dx_1 dx_2 \right| \le c \, \delta_\square(W, U)$$
$$\le \delta_\square(W, U)$$

Doing this for all edges of H and adding in an abstract sense we get our result.

$$|t(H,W)-t(H,U)|$$

$$= \left| \int \left(\prod_{u_iv_i \in E} W(u_i v_i) - \prod_{u_iv_i \in E} U(u_i,v_i) \right) \cdot \prod_{v \in V} dv \right|$$

$$\leq \sum_{i=1}^{|e(H)|} \left| \int \left(\prod_{j=1}^{i-1} U(u_j,v_j) \left(W(u_i,v_i) - U(u_i v_i) \right) \prod_{k=i+1}^{e(H)} W(u_k,v_k) \right) \prod_{v \in V} dv \right|$$

$$\leq \sum_{i=1}^{e(H)} \delta_\square (W,U)$$

$$= e(H) \cdot d_\square (W,U)$$

As an example to make it clear,

$$t(K_3,U) - t(K_3,U) = \int W(x,y) W(y,z) W(x,z) - U(x,y) U(y,z) U(x,z) \, dx\, dy\, dz.$$

$$= \int (W(x,y) - U(x,y)) W(x,z) W(y,z) \, dx\, dy\, dz$$

$$+ \int U(x,y) (W(x,z) - U(x,z)) W(y,z) \, dx\, dy\, dz$$

$$+ \int U(x,y) U(x,z) (W(y,z) - U(y,z)) \, dx\, dy\, dz$$

(MPB)

Definition : A measure preserving bijection $\varphi : [0,1] \to [0,1]$ is a function that is measurable and for any measurable A, $m(A) = m(\varphi^{-1}(A))$ where m is measure

Definition : we define the cut metric arising from the cut norm (or equivalently cut distance) as follows.

$$\delta_\square (U, W) = \inf_{\varphi \text{ is MPB}} \| U^\varphi - W \|_\square$$

where U^φ stands for $U(\varphi(x), \varphi(y))$

φ analogically stands for any relabelling

extra notes :

- $\delta_\square$ is not a distance yet because there are cases where $\delta_\square = 0$ but $U \neq W$
- $U = W$ iff $m \{ (x,y) : U(x,y) \neq W(x,y) \} = 0$
- Let $U \sim W$ if $\delta_\square (U, W) = 0$

  Then $\sim$ is an equivalence relation (check)

Definition : Given a partition $P = \{ S_1, \dots, S_k \}$ of $[0,1]$, $W : [0,1]^2 \to \mathbb{R}$ (a graphon), we define the stepping operator $W_P : [0,1]^2 \to \mathbb{R}$ which is constant on each $S_i \times S_j$ with value $\frac{1}{\lambda(S_i)\lambda(S_j)} \int_{S_i \times S_j} W$

note : dont bother if measure turns out zero (denominator problem)

THEOREM 24

Let W be a graphon & P be a partition of $[0,1]$ such that $\|W - W_P\|_\square > \varepsilon$. Then there exists P' a refinement of P into no more than 4 parts per part of P such that

$$(\|W_{P'}\|_2)^2 > (\|W_P\|_2)^2 + \varepsilon^2$$

proof:

$\|W - W_P\|_\square > \varepsilon$

$\Rightarrow \exists\, S, T \subset [0,1]$ s.t. $\int_{S\times T} W - W_P > \varepsilon$

Let P' be the refinement of P wrt S, T i.e. we divide each part of P depending on whether it is in $S \setminus T$, $T \setminus S$, $S \cap T$ or $S^c \cap T^c$ which gives at most 4 parts per part of P

Now, $\|\cdot\|_2 = \sqrt{\langle \cdot, \cdot \rangle}$ by definition.

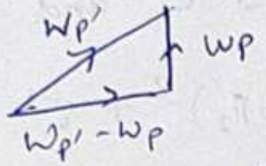
Define $\langle W, U \rangle$ to be $\int_{[0,1]^2} WU$

claim: $\langle W_{P'}, W_P \rangle = \langle W_P, W_P \rangle$

This is because while computing $\langle W, W_P \rangle$, on $S_i' \times S_j'$, we can replace W_P by its average without affecting anything.

$\therefore \langle W_{P'} - W_P, W_P \rangle = 0$

By pythagorean theorem,



$$\therefore (\|W_{P'}\|_2)^2 = (\|W_{P'} - W_P\|_2)^2 + (\|W_P\|_2)^2$$

$$\geq \varepsilon^2 + (\|W\|_2)^2$$

since $\|W_{P'} - W_P\|_2 \geq |\langle W_{P'} - W_P, 1_{S\times T}\rangle|$

(by cauchy schwartz)

(where $1_{S\times T}$ is the constant 1 function on $S \times T$ and
S is the union of parts of P'
T is the same as S)

since $\underbrace{\|1_{S\times T}\|_2}_{1} \|W_{P'} - W_P\|_2 \geq |\langle W_{P'} - W_P, 1_{S\times T}\rangle|$

$$= |\langle W - W_P, 1_{S\times T}\rangle|$$

$$> \varepsilon \quad \text{(given)}$$

THEOREM 25 (weak regularity for graphons)

$\forall\, \varepsilon > 0$, W (graphon) and P_0 (partition of $[0,1]$), there is a partition P refining parts of P_0 into no more than $4^{1/\varepsilon^2}$ parts such that $\|W - W_P\|_\square \leq \varepsilon$

Proof:

We iterate on the energy argument finitely.

Keep applying prev theorem to obtain $P_0, P_1, \dots$

At each step, either $\|W - W_P\|_\square \le \varepsilon$ and hence we stop else $(\|W_{P'}\|_2)^2 \ge (\|W_P\|_2)^2 + \varepsilon^2$ and we continue.

Since $(\|W_{P_i}\|_2)^2 \le 1$, we are guaranteed to stop after fewer than ε^{-2} steps. Also, since each subpart is divided into at most 4 parts at each stage, after ε^{-2} stages, we get each part divided into $4^{\varepsilon^{-2}}$ parts and hence we are done.

__Remark__ we state the weak regularity lemma for graphs as follows:

$\forall \varepsilon > 0$ and graph G, $\exists$ partition $P = \{V_1, \dots, V_k\}$ such that $k \le 4^{\varepsilon^{-2}}$ and $\forall A, B \subseteq V(G)$,

$$\left| e(A,B) - \sum_{i,j} d(V_i, V_j)\, |A \cap V_i|\, |B \cap V_j| \right| \le \varepsilon\, |V(G)|^2$$

THEOREM 26 (compactness of graphons)

The metric space $(\widetilde{W_0}, \delta_\square)$ is compact where $\widetilde{W_0} = W_0 / \sim$ (W_0 is space of graphons and $\sim$ was defined in definitions section)

proof:

we use tools from analysis & probability and state them here beforehand

lemma 1: Hilbert regularity lemma

suppose H is a real hilbert space and suppose $K_1, K_2, \ldots$ are all non empty subsets of H;

Given $\varepsilon > 0$, $f \in H$, $\exists\ m \le \frac{1}{\varepsilon^2}$ and $\gamma_i \in \mathbb{R}$ and $f_i \in K_i$ such that,

$$\forall\ g \in K_{m+1} \quad , \quad \left| \left< g , f - \sum_{i=1}^{m} \gamma_i f_i \right> \right| \le \varepsilon \|f\| \|g\|$$

lemma 2: Doob-Martingale convergence theorem

Every bounded martingale converges almost surely

lemma 3:

If $\delta_\square (U, W) = 0$, then $\exists\ \varphi, \phi$ which are measure preserving bijections such that $U^{\varphi} = W^{\phi}$

Now we can begin the proof of the theorem.

Since $\tilde{W}_0$ is a metric space, we will be done if we show sequential compactness i.e. given $W_1, W_2, \ldots$ graphons, we want to show there is a subsequence which converges, with respect to $\delta_\square$, to some W

Let $\{W_n\}$ be the sequence of graphons.

For each n, apply weak regularity lemma repeatedly to get a sequence of partitions $P_{n,1}, P_{n,2}, P_{n,3}, \dots$ such that

(i) $P_{n,k+1}$ refines $P_{n,k}$ $\forall\, n, k$

(ii) $|P_{n,k}| = m_k$ (a function of only k)

(iii) $\| W_n - W_{n,k} \|_\square \leq 1/k$ where $W_{n,k}$ is $(W_n)_{P_{n,k}}$

WLOG, we can assume parts of each $P_{n,k}$ are intervals. This is because swapping intervals or parts of $P_{n,k}$ is a measure preserving transformation.

Considering only the first parts of $P_{n,1}$ $\forall n$, there is a subsequence (Bolzano-Weierstrass) s.t. the lengths of the intervals along this seq converges. Fix that subseq and apply to second intervals ie first parts of $P_{n,2}$. Since there are only m_k of these, the process is finite & we have all lengths converging.

By passing to a subsequence assume that the endpoints of $P_{n,1}$ converge as $n \to \infty$

Also assume that $W_{n,1}$ converges to some graphon U_1 pointwise (through a subseq)

Repeat for $W_{n,k}$ for every k, to get U_k → understand W_i as $W_{n_i,k}$

Pass to a subsequence $W_1, W_2, W_3, \dots$

for $k=1$ this produces $W_{1,1}, W_{2,1}, W_{3,1} \dots$ which converge to U_1 pointwise

... so on for all K.

But why does $W_{n,1} \longrightarrow U_1$ (through subseq ?

Initially it may not but since there are only m_1 number of parts in $P_{n,1}$, we have an $m_1 \times m_1$ (matrix of real numbers. So everything is finite and bounded & hence bolzano weierstrass is applied on it

Now U's are step graphons. Since $P_{n,k+1}$ refines $P_{0,k}$, if we look at $W_{n,k+1}$ and step it by $P_{0,k}$, it technically 'steps down' ie. $(W_n)_{P_{n,k}} = ((W_n)_{P_{n,k+1}})_{P_{n,k}}$

ie. $W_{n,k} = (W_{n,k+1})_{P_{n,k}}$

so, in the limit $n \to \infty$,

$U_k = (U_{k+1})_{P_k}$ where P_k is the limit of $P_{n,k}$ (end points of $P_{n,k}$ converge)

These equalities imply that $\{U_t\}_{t=1}^{\infty}$ is a martingale

Since range of each U_t is $[0,1]$, it is a bounded martingale and by lemma 1, it converges to some graphon U (pointwise, almost everywhere)

We claim $W_1, W_2, \ldots$ converges to U (we passed into its subsequences & noticed it does converge to U) under $\delta_\square$

We wish to show $\delta_\square(W_n, U) \to 0$ as $n \to \infty$.

Let $\varepsilon > 0$. $\exists\ k > \frac{3}{\varepsilon}$ s.t. $\|U - U_k\|_1 < \frac{\varepsilon}{3}$

(follows by pointwise converge & DCT)

Since $W_{n,k} \to U_k$ pointwise, almost everywhere,

$\exists\ n_0 \in \mathbb{N}$ s.t. $\|U_k - W_{n,k}\|_1 < \frac{\varepsilon}{3}$ $\forall\ n > n_0$

(again by DCT). Since $k > 3/\varepsilon$,

$$\delta_\square(W_n, W_{n,k}) < \frac{\varepsilon}{3} \quad \forall\ n$$

$$\therefore \delta_\square(U, W_n) \leq \delta_\square(U, U_k) + \delta_\square(U_k, W_{n,k}) + \delta_\square(W_{n,k}, W_n)$$

$$\leq \|U - U_k\|_1 + \|U_k - W_{n,k}\|_1 + \delta_\square(W_{n,k}, W_n)$$

$$\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$ ☒

Note:

$\delta_\square(W_1, W_2)$
$\leq \|W_1 - W_2\|_\square$
$\leq \|W_1 - W_2\|_1$

THEOREM 27

- $\forall\ \varepsilon > 0$, graphon W, $\exists\ G$ (graph) s.t. $\delta_\square(G, W) < \varepsilon$
- $\forall\ \varepsilon > 0$, $\exists\ N \in \mathbb{N}$ s.t. for any graphon W, $\exists\ G_N$ so that $\delta_\square(G_N, W) < \varepsilon$

proof:

From measure theory, $\exists$ step function U so that $\|W - U\|_1 < \varepsilon/2$

Also, for any constant graphon p $(0 \le p \le 1)$ there is a graph G such that $\|G - p\|_\square < \varepsilon/2$

$\therefore$ we can find a graph G s.t. $\|G - U\|_\square < \varepsilon/2$

(patching together the G from $\|G - p\|_\square < \varepsilon/2$)

$$\therefore \delta_\square(G, W) = \|G - W\|_\square \le \|W - U\|_1 + \|U - G\|_\square$$
$$\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

To prove the second part,

consider $B_\varepsilon(G) = \{W : \delta_\square(G, W) < \varepsilon\}$

claim: $B_\varepsilon(G)$ form an open cover of $\tilde{W}_0$

This follows from the first part or you could also use that every W is the limit of some sequence of graphs

$\therefore$ By compactness, there is a finite subcover

Let the subcover arise from $G_1, \dots, G_k$

Let $N = \operatorname{lcm}(|V(G_1)|, \dots, |V(G_k)|)$

Then for each of $G_1, \dots, G_k$, there are some N vertex graphs $H_1, H_2, H_3, \dots$ such that $\delta_\square(H_i, G_i) = 0$

(The graph H_i is obtained by buffing up each G_i as follows. Replace every vertex of G_i with $N/|V(G_i)|$ number of vertices & get a complete $|V(G_i)|$ partite graph)

(Note: The graphon representations of H_i, G_i are same and hence $\delta_\square = 0$)

Since $G_1 \dots G_k$ span a finite subcover i.e. for every graphon W, $\delta_\square(W, G_i) < \varepsilon$ for some i and hence $\delta_\square(W, H_i) < \varepsilon$ and H_i is our required graph on N vertices

THEOREM 28 (Strong regularity lemma)

Let $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots) \in \mathbb{R}^\infty$. Then $\exists M (= M(\varepsilon))$ such that every graphon W can be written as

$$W = W_{str} + W_{psr} + W_{sml} \quad \text{where}$$

W_{str} is a step function with $k\ (\leq M)$ parts, and

$$\| W_{psr} \|_\square \leq \varepsilon_k, \quad \| W_{sml} \|_1 \leq \varepsilon_1$$

proof :

Fun fact : $\varepsilon_k = \varepsilon$ gives weak regularity lemma

$\varepsilon_k = \frac{\varepsilon}{k^2}$ gives Szemeredis lemma (approx.)

Now we prove the theorem

$\forall W$, $\exists$ step function U s.t. $\|W - U\|_1 \leq \varepsilon_1$

(from a measure theoretic result)

Define $k(W)$ to be the minimum k such that some k-step graphon U satisfies $\|W - U\|_1 \leq \varepsilon_1$.

Since $\{B_{\varepsilon_{k(W)}}(W)\}_{W \in \tilde{W}_0}$ is an open cover of $\tilde{W}_0$

and $\tilde{W}_0$ is compact, we can get a finite subcover. Call it $\mathcal{S}$ ($\subseteq W_0$)

Let $M = \max_{W \in \mathcal{S}} k(W)$. Then for every graphon W,

$\exists W' \in \mathcal{S}$ s.t. $\delta_\square(W, W') \leq \varepsilon_{k(W')}$ and

there is a graphon U with $k = k(W') \leq M$ s.t.

$\|W' - U\|_1 \leq \varepsilon_1$

Hence $W = \underbrace{U}_{str} + \underbrace{(W - W')}_{psr} + \underbrace{(W' - U)}_{sml}$

U has $\leq M$ steps

$\delta_\square(W, W') \leq \varepsilon_k$

$\|W' - U\|_1 \leq \varepsilon_1$

__Remark__ We state the analytic & graph forms of this strong regularity lemma.

__Analytic form :__

Given $\varepsilon > 0$, $\exists\ M = M(\varepsilon) \in \mathbb{N}$ s.t. if $W \in \mathcal{W}_0$ is a graphon, $\exists\ P = \{S_1, \dots, S_m\}$, a partition of $[0,1]$ into sets of equal measure with $m \leq M$ such that $\forall\ R \subseteq [0,1]$ which is the union of at most m^2 rectangles,

$$\left| \int_R W - W_P \right| < \varepsilon$$

__graph form__

Given $\tilde{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots)$ with $\varepsilon_1 \geq \varepsilon_2 \geq \dots$, $\exists\ M = M(\tilde{\varepsilon}) \in \mathbb{N}$ s.t. any graph G admits partition P, Q of $V(G)$ satisfying

(i) Q refines P

(ii) $|Q| \leq M$

(iii) Q is $\varepsilon_{|P|}$ - regular

(iv) $q(Q) \leq q(P) + \varepsilon_1$

where as usual $q(Q) = \sum_{ij} q(V_i, V_j)$ with Q being made up from $V_1 \dots V_k$ and $q(V_i, V_j)$ is given by $d^2(V_i, V_j) \cdot \frac{|V_i||V_j|}{n}$

THEOREM 29 (Moments lemma)

Let U, W be s.t. $t(H, W) = t(H, U)$ $\forall$ graphs H which are finite. Then $\delta_\square(U, W) = 0$

proof (sketch):

Define a W-random graph as a graph on vertex set $[n]$ with vertices i, j connected with probability $W(x_i, x_j)$ where $x_1, \dots, x_n$ are picked uniformly from $[0,1]$ (also independently)

Construct a W-random graph $G(k, W)$. Then for any k-vertex graph H, we claim,

$$P\left(G(k,W) \simeq H_{(\text{labelled})}\right) = \sum_{G \supseteq H} (-1)^{e(G)-e(H)} \, t(G, W)$$

$\therefore G(k,W)$, $G(k,U)$ have same distributions.

Let $Y(k,W)$ be a weighted W-random graph with weight of edge ij being $W(x_i, x_j)$

Then we claim

- $\delta_\square(Y(k,W), G(k,W)) \to 0$ as $k \to \infty$ with probability 1
- $\delta_\square(Y(k,W), G(k,U)) \to 0$ as $k \to \infty$ with probability 1
- $\delta_1(Y(k,W), W) \to 0$ as $k \to \infty$

with probability 1

- $\delta_1 (Y(k,U), U) \to 0$ as $k \to \infty$ with probability 1

$\therefore$ from triangle inequality & the fact that $G(k,W)$ and $G(k,U)$ have same distributions,

$\delta_\square (W, U) = 0$

THEOREM 30 (Inverse counting lemma)

$W_n \longrightarrow W$ as $n \to \infty$ iff $\delta_\square (W_n, W) \to 0$ as $n \to \infty$

proof:

we state a ~~lemma~~ corollary without proof (follows from theorem 29 and 26)

lemma : Given $\varepsilon > 0$, $\exists$ $k \in \mathbb{N}$ and $\eta > 0$ s.t. if U, W satisfy $| t(F,U) - t(F,W) | < \eta$ for every F with at most k vertices, then

$\delta_\square (U, W) \le \varepsilon$

Now we prove the theorem.

The counting lemma (theorem 23) tells,

$|t(H,U) - t(H,W)| \leq e(H) \, \|U - W\|_{\square}$

$\therefore$ If $\delta_{\square}(W_n, W) \to 0$,

then $t(H, W_n) \to t(H, W) \;\; \forall \; H$ (finite)

and hence we say $W_n \to W$ (by definition)

Let $\phi : \tilde{W}_0 \to [0,1]^{\mathcal{F}}$ as follows

$\phi(W) = \{t(H; W)\}_{H \in \mathcal{F}}$ (some sequence)

where $\mathcal{F}$ is family of all finite graphs

From the moments lemma, ϕ is injective.

ϕ is also continuous and $[0,1]^{\mathcal{F}}$ is compact

Thus $\varphi : \tilde{W}_0 \to \mathrm{Im}(\phi)$ is a bijection and

an open map. Hence the inverse is

continuous

THEOREM · 31

If $0 < p < 1$ and $t(K_2, W) = p$, $t(C_4, W) = p^4$, then $W \equiv p$ almost everywhere

proof:

Let $w(z) = \int_{[0,1]} W(x, z)\, dx$

Then, $\int_{[0,1]} w(z)\, dz = \int_{[0,1]^2} W(x,y)\, dx\, dy = p.$

$$\left(\int_{[0,1]} 1 \cdot w(z) \cdot dz\right)^2 = p^2 \leq 1 \cdot \int_{[0,1]} w^2(z)\, dz$$

(Cauchy Schwartz)

$$\Rightarrow \int_{[0,1]} w^2(z)\, dz \geq p^2$$

Now, $$\int_{[0,1]^2} \left(\int_{[0,1]} (W(x,z)W(y,z) - p^2)\, dz\right)^2 dx\, dy$$

$$= \int_{[0,1]^2} \left(\int_{[0,1]^2} (W(x,z)\, W(y,z) - p^2)\big)\big(W(x,z')\, W(y,z') \ldots\, dz\, dz'\right) dx\, dy$$

$$= \int_{[0,1]^2} \left[\int_{[0,1]^2} W(x,z) W(y,z)\, W(x,z') W(y,z')\, dz\, dz' - 2p^2 \int_{[0,1]} W(x,z)\, W(y,z)\, dz + p^4\right] dx\, dy$$

$$= 2p^2 \left(p^2 - \int_{[0,1]} w^2(z)\, dz \right)$$

$$\therefore\ 0 \leq 2p^2 \left(p^2 - \int_{[0,1]} w^2(z)\, dz \right)$$

$$\Rightarrow p^2 \leq \int_{[0,1]} w^2(z)\, dz \leq p^2$$

$$\Rightarrow \quad w(z) = p \quad \text{almost everywhere}$$

Also, cauchy schwartz must have had equality

$$\Rightarrow \int_{[0,1]} w(x,z)\, w(y,z)\, dz = p^2 \quad (a.e.)$$

$$\cancel{\Rightarrow \int_{[0,1]} w(x,z)\, dz = p \quad (a.e.)}$$

$$p^2 = \left(\int_{[0,1]} w(z)\, dz \right)^2 \leq \int_{[0,1]} w^2(x,z)\, dz = p^2$$

$$\Rightarrow w(x,z) = p \quad (a.e.)$$

We used the fact that if $f : [0,1]^2 \to [0,1]$

such that $\displaystyle\int_{[0,1]} f(x,z)\, f(y,z)\, dz = \lambda \quad (a.e.)$

then $\displaystyle\int_{[0,1]} f(x,z)^2\, dz = \lambda \quad (a.e.)$.

GRAPH COLOURINGS

THEOREM 32

1) $H \subseteq G \Rightarrow \chi(G) \geq \chi(H)$

- In particular $\chi(G) \geq \omega(G)$ = clique number

2) $\chi(G) \geq \dfrac{n}{\alpha(G)}$

3) $\chi(G) \leq \Delta(G) + 1$

4) $\chi(G) \leq d(G) + 1$

where $d(G)$ = degeneracy number $:= \max_{H \subseteq_{ind} G} \delta(H)$

proof:

1) Trivially follows by restriction of the colouring function φ to the subgraph

2) each set $V_i := \{ v \in V(G) \mid v \text{ is coloured colour } i \}$ for $i = 1, 2, \dots, k$ is an independent set (where $k = \chi(G)$)

$\therefore |V_i| \leq \alpha(G) \quad \forall i$

$\therefore n = \left| \bigcup_{i=1}^{k} V_i \right| = \sum_{i=1}^{k} |V_i| \leq \sum_{i=1}^{k} \alpha(G) = k\alpha(G)$

$\therefore n \leq \chi(G)\,\alpha(G)$

3) order the vertices $\{v_1, v_2, \ldots, v_n\}$ in any order randomly.

Consider the set of colours $\{1, 2, \ldots, \Delta(G)+1\}$

Assign a colour to vertex v_i which is different from all colours of its neighbours to its left in the list.

(Note: The different colour should be picked ideally i.e. If v_1 is coloured 1, v_2 is coloured 2 & v_3 is not a neighbour of v_1 but is a neighbour of v_2; then we shouldn't colour it 3 but colour it 1)

Since every vertex has maximum degree $\Delta(G)$, it can have at most $\Delta(G)$ neighbours to its left in which case we colour it with the $\Delta(G)+1$ th colour in the worst case when all neighbours have all the colours $1, 2, \ldots, \Delta(G)$.

4) We proceed by induction on n

The statement is trivially true for $n = 1$ since $d(G) = 0$

Pick a vertex x of minimum degree

Then $d(x) \leq d(G)$

Let $H = G \setminus \{x\}$

By induction, H can be $d(H)+1$ coloured

Clearly $d(H) \leq d(G)$ and hence H can be $d(G)+1$ coloured

x can have at most $d(G)$ neighbours in H

and hence we can always assign it the $d(G)+1^{th}$ colour

Thus, every colouring of H is extendable to G and we are done

THEOREM 33 (Brook's theorem)

If G is connected, $G \neq K_n, C_{2n+1}$, then

$\chi(G) \leq \Delta(G)$

ie. $\chi(G) = \Delta(G)+1$ iff $G = K_n, C_{2n+1}$

proof :-

Clearly $G = K_n, C_{2n+1} \Rightarrow \chi(G) = \Delta(G)+1$

Suppose $G \neq K_n, C_{2n+1}$.

Let $\Delta(G) = k$. We wish to prove $\chi(G) = k$.

Suppose G is not k-regular.

Let v_0 be a vertex of G of degree less than k.

Consider a spanning tree of G. We now give an algorithm to create a vertex order wrt v_0.

Consider the longest paths starting from v_0 and ending at some collection of vertices $\{v_i\}_{i=1}^{n_1}$ ie all paths in the spanning tree are chosen, that start from

v_0 and end at v_1 or v_2 or ... or v_{n_1} are the longest paths.

Put $v_1, \ldots, v_{n_1}$ as the first n_1 vertices in our list in any order

Then the next n_2 vertices, $v_{n_1+1}, v_{n_1+2}, \ldots, v_{n_2}$ are precisely those that form the second longest paths with v_1

continue and put v_0 to be the last vertex.

In this ordering, for every vertex different from v_0, there is at least one ~~....~~ neighbour to the right of it. (since to reach that vertex ~~.........~~ from v_0, we must pass a neighbour of it which will come later in the list because its path length will be 1 lesser)

Thus, maximum no. of colours assigned to neighbours of ~~.~~ v_i $(\neq v_0)$ that appear before v_i is $k-1$ and hence v_i can be given a colour from the list $\{1, 2, \ldots, k\}$. v_0 has degree $< k$ and is also colourable from the same list $\{1, 2, .. k\}$

Hence $\chi(G) = k$.

Now let G be k-regular.

Suppose G is k-regular and G has a cut vertex i.e. there is a vertex v so that $G \setminus \{v\}$ is disconnected and is a union of connected graphs $G_1 \cup G_2 \cdots \cup G_m$

Consider G_i and include v in it with all edges of v relevant to G_i i.e. consider the induced subgraph $G([V(G_i) \cup \{v\}])$. Then v has degree $< k$ in this graph else all other G_i will become empty (since v has degree k in G, all of its neighbours cannot be in G_i else v will not be a cut vertex)

$\therefore$ Going back to prev case, $H_i = G([V(G_i) \cup \{v\}])$ has a k-colouring and let this colouring assign colour 1 to v. Similarly we have k-colourings in other H_i that assign colour 1 to v & put together, G has a k-colouring.

Now let G have no cut vertices. Suppose some vertex v has neighbours x_1, x_2 such that x_1, x_2 are not adjacent and $G \setminus \{x_1, x_2\}$ is still connected, then we consider $H = G([V(G) \setminus \{x_1, x_2\}])$ and consider its spanning tree. Applying the idea of vertex ordering with respect to v, we get that, each vertex has at most $k-1$ neighbours before it and can be coloured from $\{1, 2, \dots, k\}$. Coming to the last vertex v, we have at most $k-2$ neighbours before it since x_1, x_2 are already gone. Hence it is colourable from $1, 2, \dots, k-1$. Now, x_1, x_2 can be given the same colour k and we are done.

Now we show that such a triple of (x_1, x_2, v) can always be found in 2-connected k-regular graphs (2-connected $\equiv$ no cut vertex)

Choose any vertex x. If $\kappa(G \setminus \{x\}) \geq 2$, (where κ denotes connectivity i.e. min no. of vertices to remove to get a disconnected graph, choose x_1 to be x and x_2 to be some vertex at dist. 2 from x_1. (it exists since G is regular but not a complete graph) and we are done.

If $\kappa(G \setminus \{x\}) = 1$ we let v be x. G has no cut vertex $\Rightarrow$ x has neighbour in ~~at least 2 different 'blocks' else~~ every leaf block of $G \setminus \{x\}$ (leaf block is block having exactly 1 cut vertex), choose two neighbours of x x_1, x_2 & we are done $\blacksquare$

THEOREM 34

The list chromatic number χ_ℓ satisfies :

(i) $\chi_\ell(G) \le \Delta(G) + 1$

(ii) $\chi(G) \le \chi_\ell(G)$

(iii) $\chi_\ell(G) \le d(G) + 1$ where $d(G)$ = degeneracy number

proof:

(i) We use a similar argument as we did before. For any given vertex ordering, every vertex has at most k neighbours to its left where $k = \Delta(G)$

Thus, given any list $\{x_1, x_2, \dots, x_k, x_{k+1}\}$ on a vertex x, there will always be some x_i which was not used on one of x's previous neighbour & colouring greedily thus, we are done

(ii) Assign the same list $\{1, 2, \dots, m\}$ to each vertex where $m = \chi_\ell(G)$ since G is m-list colourable, on this m sized list also it is L-choosable.

Hence $\chi(G) \le m$

(iii) we prove by induction (in a similar way we proved $\chi(G) \le d(G) + 1$)

$$d(G) = \max_{H \,(\subseteq \text{ind})\, G} \delta(H)$$

Consider vertex v of min degree in G

Clearly $d(v) \leq d(G) = \lambda$ (say)

(oh! By the way the statement is trivially true for a single vertex graph)

Let $F = G \setminus \{v\}$

F is a graph on $n-1$ vertices & by induction hypothesis, it can be $d(H)+1$ list coloured and hence $d(G)+1$ list coloured (since $d(H) \leq d(G)$)

Thus, if v is given a list $\{x_1, \ldots, x_\lambda, x_{\lambda+1}\}$, then since v only has atmost λ neighbours in F, a colour can be assigned to v with no issues.

THEOREM 35 (Brook for list)

$\chi_\ell(G) \leq \Delta(G)$ if G is not a clique or odd cycle (and of course, connected)

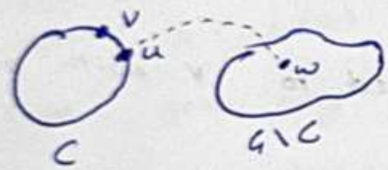
Proof:

If the graph is not Δ-regular, we go exactly along lines of theorem 33 by constructing the vertex ordering with respect to a given vertex v_0 whose degree is less than Δ

Suppose G is Δ-regular. We show that G is list-Δ-colourable. Fix some list assignment L to each vertex with each list having Δ colours, for this purpose.

Suppose G has a k-cycle $(k < n)$ and a vertex v on the cycle C such that all neighbours of v are only within C (hence $k > \Delta$)

Since G is connected, v has a neighbour u such that u has a neighbour w in $G \setminus C$



By induction hypothesis, $G \setminus C$ is list colourable with list sizes of Δ. Obtain a colouring. We now extend this colouring to C and hence G as follows.

Case 1: If colour of w is not even mentioned in list of u, colour u arbitrarily

Case 2: If colour of w is in list of u

Case 2.1: colour of w is in list of v

Case 2.2: colour of w is not in list of V

Case 2.1 : Choose colour of v to be that of w

Case 2.2 : Choose color of v to be ~~anything~~ anything from .

List of v \ list of u

We note that in all the three cases,

$$|L(u) \cap \{f(v), f(w)\}| \leq 1$$

Now pick colours for $v_2, v_3, \ldots, v_k$ in that order where $v_1 = v$, $v_k = u$. For every $2 \leq i \leq k-1$, there is a right neighbour and hence an available colour for v_i can be chosen (since v_i will have $\Delta - 1$ (atmost) previous neighbours. For $v_k = u$, at most $\Delta - 1$ colours have been used for neighbours of u (since v and w together took at most 1) and hence u can also be coloured

Now we come to the general case. If G is connected and not a clique, we can find $v_1 v_2 v_3$ s.t. v_1v_2, v_2v_3 are edges & v_1v_3 isn't one. (To do this just pick x, y s.t. $xy \notin E(G)$, consider the shortest path from x to y and $v_1 v_2 v_3$ are the last 3 ~~neighbours~~ vertices of the path i.e. $v_3 = y$.

Let P = longest path in G starting with $v_1 v_2 v_3$

$P = (v_1\ v_2\ v_3 \dots v_l)$

All neighbours of v_l are on P

Let v_i be the neighbour of v_l with lowest index i

i.e. farthest from v_l.

$\therefore C = (v_i, v_{i+1}, \dots, v_l)$ has all neighbours of v_l on it.

If C is not hamiltonian we are done by the previous case

$\therefore l = n$, $i = 1$ is the remaining case

Let v_j be a neighbour of v_2 other than v_1, v_3 (since G is not 2-regular v_j exists)

Now consider the ordering of vertices:

$(v_1, v_3, v_4, v_5, \dots, v_{j-1}, v_n, v_{n-1}, v_{n-2}, \dots, v_j, v_2)$

Same way as before we can assign colours $f(v_1)$, $f(v_3)$ to v_1, v_3 so that

$$|L(v_2) \cap \{f(v_1), f(v_3)\}| \leq 1$$

Now we choose rest of the colours in order. Every v_i other than v_2 has at least one neighbour after it

This is clear for $i = 1, 3, 4, 5, \ldots, j-2, n, n-1, \ldots, j+1$.

v_{j-1} has neighbour v_j

and v_j has neighbour v_2

Now, for v_2, colour it something from its list distinct from colours assigned to its neighbours.

Doing this is possible because we saw that $\{f(v_1), \ldots f(v_j)\}$ 'eat up' at most 1 colour together from $L(v_2)$

THEOREM 36 (Alons theorem)

$$\chi_\ell(G) \geq \Omega\left(\frac{\log d}{\log \log d}\right)$$

In fact, $\chi_\ell(G) \geq \left(\frac{1}{2} - o_d(1)\right) \log_2 d$

where $d = \delta(G)$

and $o_d(1) \xrightarrow{d \to \infty} 0$.

proof:

we prove a slightly stronger statement i.e.

if $\delta > \dfrac{4(s^2+1)^2\, 2^{2s}}{(\log_2 e)^2}$, then $\chi_\ell(G) > s$

(δ = min degree)

Our aim is to show that G is not s-list choosable.

Let B be a subset of V with each vertex of V getting chosen to be in B with probability $1/\sqrt{d}$.

Let $S = \{1, 2, \dots, s^2\}$ be the set of colours.

For every $b \in B$, let $S(b)$ denote the s sized list of vertex b. There are $\binom{s^2}{s}$ possible options.

Call a vertex $v \in V$ good if $v \notin B$ and for every

~~v has at least 1 neighbour b in B s.t.~~

~~S(b)~~ $T \subset S$ with cardinality $|T| = \lceil \frac{s^2}{2} \rceil$

v has a neighbour $b \in B$ s.t. $S(b) \subset T$

$$\mathbb{P}(v \text{ is not good}) = \mathbb{P}(v \in B) + \mathbb{P}(v \notin B \text{ and } v \text{ is not good})$$

$$= \mathbb{P}(v \in B) + \mathbb{P}(v \notin B) \cdot \mathbb{P}(E)$$

Where E is the event that there is some subset $T \subset S$ s.t. $\forall$ neighbours w of v either $w \notin B$ or $S(w) \not\subseteq T$

$$\therefore \mathbb{P}(v \text{ is not good}) \leq p + (1-p)\binom{s^2}{\lceil \frac{s^2}{2} \rceil}\left(1 - \mathbb{P}\left(\begin{array}{c}\text{every neighbour of } u \text{ is in } B \text{ and} \\ S(u) \subseteq T\end{array}\right)\right)^{d}$$

where $p = \frac{1}{\sqrt{d}}$ since $d(v) \geq \delta$

$$= \cancel{\frac{1}{\sqrt{d}} + \left(1 - \frac{1}{\sqrt{d}}\right)\left(\left\lceil \frac{s^2}{2} \right\rceil\right)\left(1 - \frac{1}{\sqrt{d}} \left\lceil \frac{s}{2} \right\rceil\right)}$$

$$= \frac{1}{\sqrt{d}} + \left(1 - \frac{1}{\sqrt{d}}\right)\binom{s^2}{\lambda}\left(1 - \frac{1}{\sqrt{d}}\left(\frac{\lambda(\lambda-1)\cdots(\lambda-s+1)}{s^2(s^2-1)\cdots(s^2-s+1)}\right)\right)^d$$

where $\lambda = \lceil \frac{s^2}{2} \rceil$

Now $$\frac{\lambda(\lambda-1)\cdots(\lambda-s+1)}{s^2(s^2-1)\cdots(s^2-s+1)} \geq \frac{1}{2^s}\prod_{i=0}^{s-1}\frac{s^2-2i}{s^2-i}$$

$$= \frac{1}{2^s}\prod_{i=0}^{s-1}\left(1 - \frac{i}{s^2-i}\right)$$

$$\geq \frac{1}{2^s}\left(1 - \frac{\sum_{i=0}^{s-1} i}{s^2-s}\right)$$

$$= \frac{1}{2^{s+1}}$$

$$\therefore P(v \text{ is not good}) \leq \frac{1}{\sqrt{d}} + \frac{2^{s^2}}{4}\left(1 - \frac{1}{\sqrt{d}\cdot 2^{s+1}}\right)^d$$

$$\leq \frac{1}{\sqrt{d}} + \frac{2^{s^2}}{4} e^{-\frac{\sqrt{d}}{2^{s+1}}} < \frac{1}{4}$$

$$\left(\because d > \frac{4(s^2+1)^2}{(\log_2 e)^2} 2^{2s}.\right)$$

$\therefore \mathbb{E}$(vertices not good) $< \frac{n}{4}$.

$\therefore$ By markov's inequality,

$\mathbb{P}$(at least $n/2$ good vertices) $> \frac{1}{2}$

Also $\mathbb{E}(|B|) = \frac{n}{\sqrt{d}}$

$\therefore \mathbb{P}\left(|B| > \frac{2n}{\sqrt{d}}\right) < \frac{1}{2}$

$\therefore$ With positive probability, $|B| < \frac{2n}{\sqrt{d}}$; there are at least $\frac{n}{2}$ good vertices

Now fix a choice of B, $S(b)$ $\forall b \in B$ s.t. $|B| < \frac{2n}{\sqrt{d}}$ and there is a set A of good vertices of size at least $n/2$.

For each $a \in A$ choose a set $S(a) \subset S$ randomly and independently

We will show that with +ve prob, $\nexists$ no proper colouring $c : V \to S$ of G assigning to each $v \in A \cup B$ some colour from its list

There are at most $s^{|B|}$ choices for restricting the colouring c to b s.t. $c(b) \in S(b)$. Fix a restriction. We want to extend it to a proper colouring of $G[(A \cup B)]$

Since each $a \in A$ is good, the set T_a consisting of all colours assigned to neighbours of a in B by $c|_B$ intersects every subset of size $\lceil \frac{s^2}{2} \rceil$ of S and hence its cardinality is at least $\lceil \frac{s^2}{2} \rceil$. If $S(a) \subseteq T_a$ we don't have a proper colouring

$$\therefore \mathbb{P}(a \text{ can be coloured}) \leq 1 - \frac{\lambda(\lambda-1)\cdots(\lambda-s+1)}{s^2(s^2-1)\cdots(s^2-s+1)}$$

$$\leq 1 - \frac{1}{2^{s+1}}$$

$$\therefore \left(1 - \frac{1}{2^{s+1}}\right)^{|A|} \leq \left(1 - \frac{1}{2^{s+1}}\right)^{n/2} \leq e^{-n/2^{s+2}}$$

$$\therefore \mathbb{P}(c|_B \text{ extends to } A) \leq e^{-n/2^{s+2}}$$

$\therefore \mathbb{P}(\text{some } c|_B \text{ extends to } A)$

$$\leq s^{|B|} e^{-\frac{n}{2^{s+2}}}$$

using hypothesis and also that $s \geq 3$

$$\leq e^{\frac{2n}{\sqrt{d}} \ln s - \frac{n}{2^{s+2}}} \leq 1$$

$\therefore$ no colouring of desired type with the probability ☒

THEOREM 37

Given $k, g \in \mathbb{N}$. $\exists$ G_n s.t. $girth(G) > g$

and $\chi(G) > k$

where girth denotes size of smallest cycle other than ~~the~~ 2 cycle (edge) & 1 cycle (vertex)

Proof:

Pick the Erdos Renyi graph $G_{n,p}$ where p will be decided based on n

Set $p = \frac{f(n)}{n}$ for some f to be declared

Let $N =$ no of cycles in g of size at most g

$$= \sum_{\substack{v_1 v_2 \cdots v_t \\ 3 \le t \le g}} p^t$$

$$= \sum_{t=3}^{g} \frac{n(n-1)\cdots(n-t+1)}{2t} p^t$$

$$\le \sum_{t=3}^{g} \frac{(np)^t}{2t}$$

$$\le \frac{g}{6}(np)^g = \frac{g}{6}(f(n))^g$$

$$\mathbb{P}(\alpha(G) \geq l) = \binom{n}{l}(1-p)^{\binom{l}{2}}$$

$$\leq \frac{n^l e^{-p\binom{l}{2}}}{l!}$$

$$\leq \frac{1}{l!} e^{-p\frac{l^3}{3} + l \log n}$$

$$\leq \frac{1}{l!} e^{-p\frac{l^2}{3} + l \log n}$$

We ensure $p\frac{l}{3} \geq 2 \log n$

i.e. $p \geq \frac{6 \log n}{l}$

choose $l = \frac{n}{2k}$ and $f(n) = 12 k \log n$

so that $\frac{g}{6} (\log n)^g \, (12 k)^g$ is small

and $\mathbb{P}(\alpha(G) \geq l) \leq \frac{1}{l!} e^{l\left(-\frac{pl}{3} + \log n\right)}$

is also small (so that $\mathbb{P}(\alpha(G) < l)$ is big)

Now $E(N) < O_{g,k}((\log n)^g)$

$\Rightarrow \mathbb{P}(N > C_{g,k} (\log n)^g) < \frac{1}{2}$

by Markov inequality

$\therefore \alpha(G) \le \frac{n}{2k}$ and $N < C_{g,k} (\log n)^g$ with positive probability.

Throw away a single vertex from every cycle of size $\le g$ and get the new graph G^*

$\alpha(G^*) \le \alpha(G) \le \frac{n}{2k}$

and $n - C_{g,k} (\log n)^g \geq 0.9n$

vertices are left

$\Rightarrow \chi(G^*) \geq \frac{0.9n}{(n/2k)} = 1.8k$

THEOREM 38 (Vizings theorem)

$\Delta(G) \le \chi'(G) \le \Delta(G) + 1$

where $\chi'(G)$ = edge colouring number

= colouring number of line graph

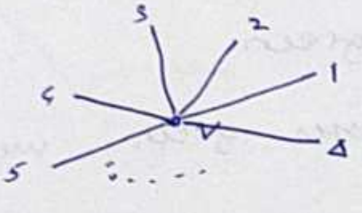
Proof:

we first prove the obvious

$\chi'(G) \geq \Delta(G)$ bound.

In the line graph, two vertices, that represent two edges of G, are adjacent iff they have a common vertex

Consider max degree vertex v



we are forced to use at least Δ colours else two adjacent edges get the same colour by pigeonhole.

we now show that a $\Delta+1$ colouring is possible through induction.

Pick any edge xy_0 in G and delete it.

By induction we have an edge colouring of $G - xy_0$

Now we want to colour xy_0 using the already present $\Delta+1$ colours.

Notice that each vertex misses at least one colour since degree $\leq \Delta$

Consider all neighbours of x which are $y_0, y_1, \ldots$

Suppose y_0 misses red colour

Then if x_0 also misses red we are directly done by colouring xy_0 = red

If not, x will have a red edge xy_1

suppose y_1 misses green

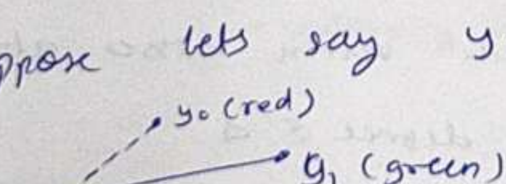
Then if x also misses green, we colour xy_1 green and red gets freed up from both x and y_1 and we can colour xy_0 as red.

Suppose not, then x has a green edge xy_2

Keep proceeding till we reach xy_k and y_k misses a colour already seen

ie in the above example of xy_2,

suppose lets say y_2 misses red

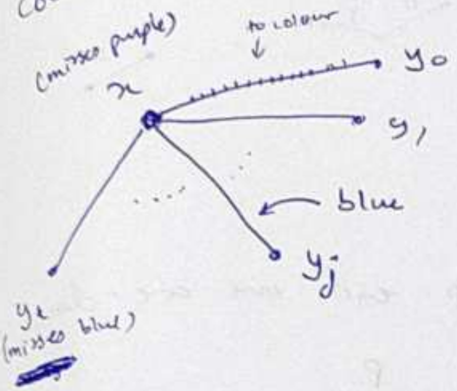


Say we stopped at y_k

colour missing at y_i = colour of xy_{i+1}

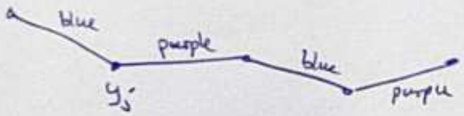
$\forall\, i = 0, 1, \dots, k-1$

colour missing at y_k = colour of some xy_j



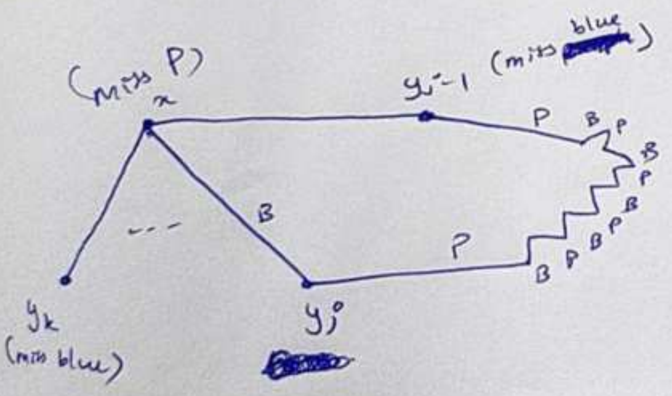
Then y_j <u>must</u> have some purple edge

(else we could not have gone till y_k)



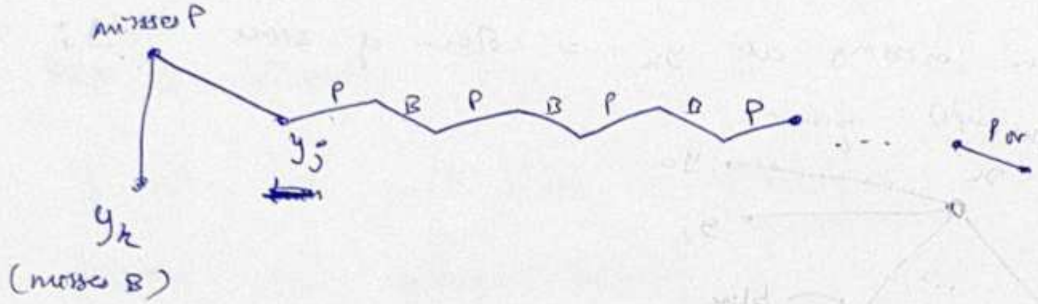
a longest path is constructed (alternating PBPB ...)

<u>Case I</u> The path ends at v_{j-1}



swap all B's and P's. and then x & y_k both will miss blue & we can be done

Case 2 : Suppose the path ends somewhere else



Swap all B, P again and we are done

note that for w, either P or B must have been free else path wasnt longest