Sh:627
Journal of Combinatorial Theory, Series A **82**, 179–185 (1998)
Article No. TA972845

NOTE

Erdős and Rényi Conjecture

Saharon Shelah

Institute of Mathematics, The Hebrew University, Jerusalem, Israel; and Department of Mathematics, Rutgers University, New Brunswick, New Jersey

Communicated by the Managing Editors

Received October 10, 1996

Affirming a conjecture of Erdős and Rényi we prove that for any (real number) $c_1 > 0$ for some $c_2 > 0$, if a graph G has no c_1 $(\log n)$ nodes on which the graph is complete or edgeless (i.e., G exemplifies $|G| \nrightarrow (c_1 \log n)^2_2$), *then* G has at least $2^{c_2 n}$ non-isomorphic (induced) subgraphs. © 1998 Academic Press

0. INTRODUCTION

Erdős and Rényi [ER] conjectured (letting $I(G)$ denote the number of (induced) subgraphs of G up to isomorphism and $\mathrm{Rm}(G)$ be the maximal number of nodes on which G is complete or edgeless):

$(*)$ for every $c_1 > 0$ for some $c_2 > 0$ for n large enough for every graph G_n with n points,

$$\otimes\ \mathrm{Rm}(G_n) < c_1(\log n) \Rightarrow I(G_n) \geqslant 2^{c_2 n}.$$

They succeeded in proving a parallel theorem which replaces $\mathrm{Rm}(G)$ with the bipartite version:

$$\begin{aligned}\mathrm{Bipartite}(G) =: \mathrm{Max}\{k\colon & \text{ there are disjoint sets } A_1, A_2 \text{ of } k \text{ nodes of } G,\\ & \text{such that } (\forall x_1 \in A_1)(\forall x_2 \in A_2)(\{x_1, x_2\} \text{ an edge) or}\\ & (\forall x_1 \in A_1)(\forall x_2 \in A_2)(\{x_1, x_2\} \text{ is not an edge)}\}.\end{aligned}$$

It is well known that $\mathrm{Rm}(G_n) \geqslant \frac{1}{2} \log n$. On the other hand, Erdős [Er7] proved that for every n for some graph G_n, $\mathrm{Rm}(G_n) \leqslant 2 \log n$. In his construction G_n is quite a random graph; it seems reasonable that any

0097-3165/98 $25.00

graph G_n with small $\mathrm{Rm}(G_n)$ is of similar character and this is the rationale of the conjecture.

Alon and Bollobas [AlBl] and Erdős and Hajnal [EH9] affirm a conjecture of Hajnal:

$$(*) \quad \text{if } \mathrm{Rm}(G_n) < (1-\varepsilon)n \text{ then } I(G_n) > \Omega(\varepsilon n^2)$$

and Erdős and Hajnal [EH9] also prove

$$(*) \quad \text{for any fixed } k, \text{ if } \mathrm{Rm}(G_n) < \frac{n}{k} \text{ then } I(G_n) > n^{\Omega(\sqrt{k})}.$$

Alon and Hajnal [AH] noted that those results give poor bounds for $I(G_n)$ in the case $\mathrm{Rm}(G_n)$ is much smaller than a multiple of $\log n$, and prove an inequality weaker than the conjecture:

$$(*) \quad I(G_n) \geqslant 2^{n/2t^{20\log(2t)}} \qquad \text{when} \quad t = \mathrm{Rm}(G_m)$$

so in particular if $t \geqslant c \log n$ they obtained $I(G_n) \geqslant 2^{n/(\log n)^{c \log\log n}}$, that is, the constant c_2 in the conjecture is replaced by $(\log n)^{c \log\log n}$ for some c.

1

$\langle 0.1 \rangle$ *Notation.* $\log n = \log_2 n$. Let c denote a positive real, and G, H denote graphs, which are here finite, simple and undirected. V^G is the set of nodes of the graph G, and E^G is the set of edges of the graph G so $G = (V^G, E^G)$ and E^G is a symmetric, irreflexive relation on V^G, i.e. a set of unordered pairs. Thus $\{x, y\} \in E^G$, xEy, and $\{x, y\}$ an edge of G all have the same meaning. $H \subseteq G$ means that H is an induced subgraph of G, i.e., $H = G \upharpoonright V^H$. Let $|X|$ be the number of elements of the set X.

$\langle 1 \rangle$ Definition. $I(G)$ is the number of (induced) subgraphs of G up to isomorphisms.

$\langle 2 \rangle$ Theorem. *For any $c_1 > 0$ for some $c_2 > 0$ we have (for n large enough) that if G is a graph with n edges and G has neither a complete subgraph with $\geqslant c_1 \log n$ nodes nor a subgraph with no edges with $\geqslant c_1 \log n$ nodes, then $I(G) \geqslant 2^{c_2 n}$.*

$\langle 3 \rangle$ *Remarks.* (1) Suppose $n \nrightarrow (r_1, r_2)$ and m are given. Choose a graph H on $\{0, ..., n-1\}$ exemplifying $n \nrightarrow (r_1, r_2)^2$ (i.e. with no complete subgraphs with r_1 nodes and no independent set with r_2 nodes). Define the graph G with set of nodes $V^G = \{0, ..., mn-1\}$ and set of edges $E^G = \{\{mi_1 + \ell_1, mi_2 + \ell_2\} : \{i_1, i_2\} \in E^H \text{ and } \ell_1, \ell_2 < m\}$. Clearly G has nm nodes and it exemplifies $mn \nrightarrow (r_1, mr_2)$. So $I(G) \leqslant (m+1)^n \leqslant 2^{n \log_2(m+1)}$ (as the isomorphism type of $G' \subseteq G$ is determined by $\langle |G' \cap [mi, mi+m)| : i < n \rangle$). We conjecture that this is the worst case.

(2) Similarly, if $n \nrightarrow ([{r_1 \atop r_2}])^2_2$; i.e., there is a graph with n nodes and no disjoint $A_1, A_2 \subseteq V^G$, $|A_1| = r_1$, $|A_2| = r_2$ such that $A_1 \times A_2 \subseteq E^G$ or $(A_1 \times A_2) \cap E^G = \emptyset$, then there is G exemplifying $mn \to ([{n_1 m \atop r_2 m}])^2_2$ such that $I(G) \leqslant 2^{n \log(m+1)}$.

Proof. Let c_1, a real > 0, be given.

Let m_1^* be[1] such that for every n (large enough) $(n/(\log n)^2 \log\log n) \to (c_1 \log n, (c_1/m_1^*) \log n)$.

[Why does it exist? By Erdős and Szekeres [ErSz] $\binom{n_1+n_2-2}{n-1} \to (n_1, n_2)^2$ and hence for any k letting $n_1 = km$, $n_2 = m$ we have $\binom{km+m-2}{m-1} \to (km, m)^2$, now $\binom{m+m-2}{m-1} \leqslant 2^{2(m-1)}$ and

$$\binom{(k+1)\,m+m-2}{m-1} \Big/ \binom{km+m-2}{m-1} = \prod_{i=0}^{m-2} \left(1 + \frac{m}{km+i}\right)$$

$$\leqslant \prod_{i=0}^{m-2} \left(1 + \frac{m}{km}\right) = \left(1 + \frac{1}{k}\right)^{m-1},$$

hence $\binom{km+m-2}{m-1} \leqslant (4 \cdot \prod_{\ell=0}^{k-2} (1 + (1/\ell+1)))^{m-1}$, and we choose k large enough (see below). For (large enough) n we let $m = (c_1 \log n)/k$; more exactly, the first integer is not below this number so

$$\log \binom{km+m-2}{m-1} \leqslant \log \left(4 \cdot \prod_{\ell=0}^{k-2} \left(1 + \frac{1}{\ell+1}\right)\right)^{m-1}$$

$$\leqslant (\log n) \cdot \frac{c_1}{k} \cdot \log \left(4 \cdot \prod_{\ell=0}^{k-2} \left(1 + \frac{1}{\ell+1}\right)\right) \leqslant \frac{1}{2} (\log n)$$

(the last inequality holds as k is large enough). Finally, let m_1^* be such a k. Alternatively, just repeat the proof of Ramsey's theorem.]

Let m_2^* be minimal such that $m_2^* \to (m_1^*)^2_2$. Let $c_2 < 1/m_2^*$ (be a positive real), let $c_3 \in (0, 1)_{\mathbb{R}}$ be such that $0 < c_3 < (1/m_2^*) - c_2$, let $c_4 \in \mathbb{R}^+$ be $4/c_3$ (even $(2+\varepsilon)/c_3$ suffices), and let $c_5 = (1 - c_2 - c_3)$ (it is > 0). Finally, let $\varepsilon \in (0, 1)_{\mathbb{R}}$ be small enough.

Now suppose

$(*)_0$ n is large enough, G is a graph with n nodes, and $I(G) < 2^{c_2 n}$.

[1] The $\log\log n$ can be replaced by a constant computed from m_1^*, m_2^*, c_ℓ later.

We choose $A \subseteq V^G$ in the following random way: for each $x \in V^G$ we flip a coin with probability $c_3/\log n$, and let A be the set of $x \in V^G$ for which we succeed. For any $A \subseteq V^G$ let $\approx_A$ be the following relation on V^G, $x \approx_A y$ *iff* $x, y \in V^G$ and $(\forall z \in A)[zE^G x \leftrightarrow zE^G y]$. Clearly, $\approx_A$ is an equivalence relation, and $\approx'_A = \approx_A \restriction (V^G \backslash A)$.

For distinct $x, y \in V^G$ what is the probability that $x \approx_A y$? Let

$$\mathrm{Dif}(x, y) =: \{z: z \in V^G \text{ and } zE^G x \leftrightarrow \neg zE^G y\},$$

and $\mathrm{dif}(x, y) = |\mathrm{Dif}(x, y)|$, so the probability of $x \approx_A y$ is

$$\left(1 - \frac{c_3}{\log n}\right)^{\mathrm{dif}(x, y)} \sim e^{-c_3\, \mathrm{dif}(x, y)/\log n}.$$

Hence the probability that for some $x \neq y$ in V^G satisfying $\mathrm{dif}(x, y) \geqslant c_4(\log n)^2$ we have $x \approx_A y$ is at most

$$\binom{n}{2} e^{-c_3(c_4(\log n)^2)/\log n} \leqslant \binom{n}{2} e^{-4 \log n} \leqslant 1/n^{0.5}$$

(remember $c_3 c_4 = 4$ and $(4/\log e) \geqslant 2.5$). Hence for some set A of nodes of G we have

$$(*)_1 \quad A \subseteq V^G \text{ and } A \text{ has } \leqslant \frac{c_3}{\log n} \cdot n \text{ elements and } A \text{ is non-empty}$$

and

$$(*)_2 \quad \text{if } x \approx_A y \text{ then } \mathrm{dif}(x, y) \leqslant c_4(\log n)^2.$$

Next,

$$(*)_3 \quad \ell =: |(V^G \backslash A)/\approx_A| \text{ (i.e. the number of equivalence classes of } \approx'_A = \approx_A \restriction (V^G \backslash A)) \text{ is } < (c_2 + c_3) \cdot n.$$

[Why? Let $C_1, ..., C_\ell$ be the $\approx'_A$-equivalence classes. For each $u \subseteq \{1, ..., \ell\}$ let $G_u = G \restriction (A \cup \bigcup_{i \in u} C_i)$. So G_u is an induced subgraph of G and $(G_u, c)_{c \in A}$ for $u \subseteq \{1, ..., \ell\}$ are pairwise nonisomorphic structures, so

$$2^\ell = |\{u: u \subseteq \{1, ..., \ell\}\}| \leqslant |\{f: f \text{ a function from } A \text{ into } V^G\}| \times I(G)$$
$$\leqslant n^{|A|} \times I(G),$$

hence (first inequality by the hypothesis toward contradiction)

$$2^{c_2 n} > I(G) \geqslant 2^{\ell} \times n^{-|A|} \geqslant 2^{\ell} \cdot n^{-c_3 n/\log n} = 2^{\ell} \times 2^{-c_3 n}$$

and hence

$$c_2 n > \ell - c_3 n \text{ so } \ell < (c_2 + c_3) n \text{ and we have gotten } (*)_3.]$$

Let $\{B_i : i < i^*\}$ be a maximal family such that:

(a) each B_i is a subset of some $\approx'_A$-equivalence class

(b) the B_i's are pairwise disjoint

(c) $|B_i| = m_1^*$

(d) $G \restriction B_i$ is a complete graph or a graph with no edges.

Now if $x \in V^G \setminus A$ then $(x/\approx'_A) \setminus \bigcup_{i<i^*} B_i$ has $< m_2^*$ elements (as $m_2^* \to (m_1^*)_2^2$ by the choice of m_2^* and "$\langle B_i : i < i^* \rangle$ is maximal"). Hence

$$\begin{aligned} n = |V^G| &= |A| + \left| \bigcup_{i<i^*} B_i \right| + \left| V^G \setminus A \setminus \bigcup_{i<i^*} B_i \right| \\ &\leqslant c_3 \frac{n}{\log n} + m_1^* \times i^* + |(V^G \setminus A)/\approx'_A| \times m_2^* \\ &\leqslant c_3 \frac{n}{\log n} + m_1^* \times i^* + m_2^*(c_2 + c_3) n \\ &= c_3 \frac{n}{\log n} + m_1^* \times i^* + (1 - m_2^* c_5) \cdot n \end{aligned}$$

hence

$$(*)_4 \quad i^* \geqslant \frac{n}{m_1^*} \left(m_2^* c_5 - \frac{c_3}{\log n} \right).$$

For $i < i^*$ let

$$B_i = \{x_{i,0}, x_{i,2}, ..., x_{m_1^* - 1}\},$$

and let

$$u_i =: \{ j < i^* : j \neq i \text{ and for some } \ell_1 \in \{1, ..., m_1^* - 1\} \text{ and } \ell_2 \in \{0, ..., m_1^* - 1\} \text{ we have } x_{j, \ell_2} \in \mathrm{Dif}(x_{i,0}, x_{i, \ell_1}) \}$$

Clearly

$$(*)_5 \quad |u_i| \leqslant m_1^*(m_1^* - 1)\, c_4 (\log n)^2.$$

Next we can find W such that

$$(*)_6 \quad \text{(i)} \quad W \subseteq \{0, ..., i^* - 1\}$$

$$\text{(ii)} \quad |W| \geqslant i^*/(m_1^*(m_1^* - 1)\, c_4(\log n)^2)$$

$$\text{(iii)} \quad \text{if } i \neq j \text{ are members of } W \text{ then } j \notin u_i.$$

[Why? See de Bruijn and Erdős [BrEr]; however, we shall give a proof when we weaken the bound. First we weaken the demand to

(iii)′ $i \in W$ and $j \in W$ and $i < j \Rightarrow j \notin u_i$.

We get this as follows: First we choose the ith member by induction. Next we find $W' \subseteq W$ that W' satisfies (iii); this is chosen similarly but the members are chosen from the top down (inside W) so the requirement on i is that $i \in W$ and $(\forall j,\ i < j \in W' \rightarrow i \notin u_j)$ so our situation is similar. So we have proved the existence, except that we get a somewhat weaker bound, which is immaterial here.]

Now for some $W' \subseteq W$,

$(*)$ $W' \subseteq W$, $|W'| \geqslant \frac{1}{2}|W|$, and all the $G \upharpoonright B_i$ for $i \in W'$ are complete graphs or all are independent sets.

By symmetry we may assume the former.

Let us sum up the relevant points:

(A) $W' \subseteq \{0, ..., i^* - 1\}$, $|W'| \geqslant ((m_2^* c_5 - (c_3/\log n)) \cdot n)/(2(m_1^*)^2 (m_1^* - 1)\, c_4(\log n)^2)$

(B) $G \upharpoonright B_i$ is a complete graph for $i \in W'$

(C) $B_i = \{x_{i,\ell} : \ell < m_1^*\}$ without repetition and $i_1, i_2 < i^*$, $\ell_1, \ell_2 < m_1^* \Rightarrow x_{i_1,\ell_1} E^G x_{i_2,\ell_2} \equiv x_{i_1,0} E^G x_{i_2,0}$.

But by the choice of m_1^* (and as n is large enough, $|W'|$ is large enough) we know $|W'| \rightarrow ((c_1/m_1^*) \log n, (c_1/1) \log n)^2$.

We apply this to the graph $\{x_{i,0} : i \in W'\}$. So one of the following occurs:

(α) there is $W'' \subseteq W'$ such that $|W''| \geqslant (c_1/m_1^*) \log n$ and $\{x_{i,0} : i \in W''\}$ is a complete graph

or

(β) there is $W'' \subseteq W'$ such that $|W'| \geqslant c_1(\log n)$ and $\{x_{i,0} : i \in W''\}$ is a graph with no edges.

Now if possibility (β) holds, then $\{x_{i,0}: i \in W''\}$ is as required, and if possibility (α) holds then $\{x_{i,t}: i \in W'', t < m_1^*\}$ is as required (see (C) above).

ACKNOWLEDGMENTS

I thank Alice Leonhardt for the beautiful typing. I thank Andras Hajnal for telling me about the problem and Mariusz Rabus and Andres Villaveces for some corrections. This research was partially supported by the NSF, Publication Number 627.

REFERENCES

[AlBl] N. Alon and B. Bollobas, Graphs with a small number of distinct induced subgraphs, *Discrete Math.* **75** (1989), 23–30.

[AH] N. Alon and A. Hajnal, Ramsey graphs contains many distinct induced subgraphs, *Graphs and Combin.* **7** (1991), 1–6.

[BrEr] G. de Bruijn and P. Erdős, A colour problem for infinite graphs and a problem in the theory of relation, *Akad. Amsterdam* **13** (1951), 371–373.

[Er7] P. Erdős, Some remarks on the theory of graphs, *Bull. Am. Math. Soc.* **53** (1947), 292–294.

[EH9] P. Erdős and A. Hajnal, On the number of distinct induced subgraphs of a graph, *Discrete Math.* **75** (1989), 145–154.

[ErSz] P. Erdős and G. Szekeres, A combinatorial problem in geometry, **2** (1935), 463–470.