Discrete Applied Mathematics 25 (1989) 37-52
North-Holland

RAMSEY-TYPE THEOREMS*

P. ERDÖS and A. HAJNAL

Mathematical Institute of the Hungarian Academy of Sciences, Budapest, V. Reáltanoda, u. 13-15, P.O.B. 127, H-1364, Hungary

Received 9 January 1989

In this paper we will consider Ramsey-type problems for finite graphs, r-partitions and hypergraphs. All these problems ask for the existence of large homogeneous (monochromatic) configurations of a certain kind under the condition that the size of the underlying set is large. As it is quite common in Ramsey theory, most of our results are not sharp and almost all of them lead to new problems which seem to be difficult. The problems we treat are only loosely connected. So we will state and explain them section by section.

Introduction, notation

Our notation will be standard (we hope) hence we state only a few conventions in advance.

For $n,k\in\mathbb{N}$, $[n]$ denotes the set $\{1,2,\dots,n\}$, $[A]^k=\{B\subset A\colon |B|=k\}$, but we write $[n]^k$ for $[[n]]^k$. $G=\langle V_G,E_G\rangle=\langle V,E\rangle$ is a *simple graph*, i.e., $E\subset[V]^2$. $\mathcal{G}^n$ is the *set of graphs having n vertices*, $\bar{G}$ is the complement of G, i.e., $\bar{G}=\langle V,[V]^2\setminus E\rangle$, K_n is (the isomorphism class) of the *complete graph on n vertices*. For $r,k\in\mathbb{N}$ and V a set, an r-partition of V with k colors is a map $f:[V]^r\to[k]$, 2-partitions of V with 2-colors can be canonically identified to graphs with vertex set V, by the formula

$$f(e)=1 \Leftrightarrow e\in E \quad \text{for } e\in[V]^r.$$

A subset $A\subset V$ is homogeneous (monochromatic) for f if f is constant on $[A]^r$.

For a graph $G=\langle V,E\rangle$, $G[A]$ is the *subgraph* of G *induced* by A, i.e., $G[A]=\langle A,E\cap[V]^2\rangle$. Hence a subset $A\subset V$ is homogeneous for G if and only if $G[A]$ is either complete or independent. We write $H\Subset G$ if H is isomorphic to an induced subgraph of G.

For $A\subset V$ we denote by $G(A)$ the neighbourhood of A in G, i.e., $G(A)=\{y\in V\colon \exists x\in A\ \{x,y\}\in E\}$. For $x\in V$ we write $G(\{x\})=G(x)$.

* Research supported by OTKA grant 1805.

0166-218X/89/$3.50

1. On subgraphs of graphs not containing large homogeneous sets

As it is well known, Ramsey's theorem implies that large graphs contain large homogeneous subsets. The following estimates are the starting point of Ramsey theory.

All $G \in \mathcal{G}^n$ contain a homogeneous set of size at least $(\log n)/(2 \log 2)$ [12]. (1.1)

For all sufficiently large n there is a $G \in \mathcal{G}^n$ not containing homogeneous sets of size $(2 \log n)/(\log 2)$ [1]. (1.2)

The graphs establishing (1.2) are nowadays called Ramsey graphs or Ramsey-like graphs and are proved to exist with a probabilistic argument. Recently Frankl and Wilson [14] constructed graphs $G \in \mathcal{G}^n$ which, for sufficiently large n and for $c>0$, do not contain homogeneous subsets of size

$$e^{c\sqrt{\log n}}. \tag{1.3}$$

We have been interested in the structure of Ramsey-like graphs for a long time. We investigated such problems for infinite graphs in [6]. There are many ways to express that such graphs are complicated and similar to random graphs. One of them uses the following definition.

A graph G is *l-universal* if for all graphs H with l vertices $H \Subset G$ holds, i.e., H is isomorphic to an induced subgraph of G. We can expect that for sufficiently large n and fixed l all Ramsey-like graphs are l-universal. In [7] we proved this for graphs not containing homogeneous sets of size $c \log n$. In an addendum we claimed the following stronger result:

Theorem 1.1. *Assume $l \in \mathbb{N}$, $0<c<1/l$. Then there is an $n_0 = n_0(l, c)$ such that for all $n > n_0$, $G \in \mathcal{G}^n$ and $k < e^{c\sqrt{\log n}/2}$ either G contains a homogeneous subset of size k or G is l-universal.*

The first aim of this section is to give a proof of this result. Before doing this we want to make some remarks. We do not know if this result is the best possible. The appearance of the expression

$$e^{c\sqrt{\log n}}$$

in both (1.3) and Theorem 1.1 seems to be a coincidence, since the graphs constructed by Frankl and Wilson [14] are l-universal.

The following could be true.

For all graphs H there is an $\varepsilon>0$ such that for all sufficiently large n and for all $G \in \mathcal{G}^n$ either $H \Subset G$ or G contains a homogeneous set of size n^{ε}. (1.4)

First let us remark that well-known results show that for $H=K_r$ (1.4) can only be true with $\varepsilon \le (r-2)/(\binom{r}{2}-1)$ (see e.g. [16]). Hence the ε of (1.4) may tend to 0 if the size of H tends to infinity. We knew for a long time that (1.4) is true for some special graphs H. For example $\varepsilon=\frac{1}{3}$ will do for C_4. Note that we do not know if this is the best possible, we only know that $\varepsilon<\frac{1}{2}$, for C_4.

Our next result says that (1.4) is true for a class of graphs.

> We define by induction a class $\mathscr{S}$ of graphs, which we will call very simple. $K_1 \in \mathscr{S}$. Assume $G_1, G_2 \in \mathscr{S}$, $V_1 \cap V_2 = \emptyset$. Then $G=\langle V,E\rangle \in \mathscr{S}$ if $V=V_1 \cup V_2$ and E is either $E_1 \cup E_2$ or $E_1 \cup E_2 \cup [V_1, V_2]$, where $[V_1,V_2]=\{\{u_1,u_2\}: u_1 \in V_1 \wedge u_2 \in V_2\}$. (1.5)

Theorem 1.2. $\forall H \in \mathscr{S}\ \exists \varepsilon>0\ \exists n_0\ \forall n>n_0\ \forall G \in \mathscr{G}^n$ *either* $H \subseteq G$ *or* G *contains a homogeneous set of size at least* n^{ε}, *i.e.*, (1.4) *is true for the very simple graphs* H.

Before turning to the proofs we make some more comments. The only $H \notin \mathscr{S}$ for which we know that (1.4) holds is P_4, the path of length 4 (having four vertices), since known results imply that graphs not containing an induced copy of P_4 are perfect (see e.g. [15]), hence $\varepsilon=\frac{1}{2}$ works for them. The simplest graph for which we do not know (1.4) is C_5. However there is an even simpler problem pointed out to us by Lovász. We cannot even prove that if neither G nor $\bar{G}$ contains an odd circuit as an induced subgraph, then G contains a homogeneous set of size n^{ε}. On the other hand if the strong perfect graph conjecture holds, $\varepsilon=\frac{1}{2}$ must work for such graphs.

Finally we state a generalization of Theorem 1.1 without proof.

Theorem 1.3. *Assume* $l,s \in \mathbb{N}$, $s \ge 2$ *and* $h:[l]^2 \to s$. *There are* n_0 *and* $c>0$ *such that for all* $n \ge n_0$ *and for all 2-partitions* g *of length* s *of* $[n]$, $g:[n]^2 \to s$, *either there is a* $B \subset [n]$, $|B|=l$ *with* $h \cong g \mid B$ *or there is a subset* $H \subset V$ *of size* $e^{c\sqrt{\log n}}$ *which is not totally inhomogeneous for* g, *i.e., for some* $i \in [s]$, $g(e) \ne i$ *for all* $e \in [H]^2$.

We mention that there are estimates analogous to (1.1) and (1.2) for the size of the largest not totally inhomogeneous set for a $g:[n]^2 \to [s]$. The size of these sets is $c_s \log n$ where c_s tends to infinity with s [13].

We need some lemmas. The first one is about very simple graphs, which will be used in the proof of Theorem 1.1 as well.

Lemma 1.4. *Assume* $H \in \mathscr{S}$, $|H|=k$. *Then* H *contains a homogeneous set of size at least* $k^{1/2}$.

Proof. It is a well-known fact that very simple graphs are perfect. The simplest way to see this is the following. Denote, as usual, by $\alpha(H)$ the size of the *largest independent* set and by $\omega(H)$ the size of the *largest complete graph* and prove by induction that for $H \in \mathscr{S}$,

$$\alpha(H)\omega(H) \geq |H|. \qquad \square$$

The following is the crucial lemma needed for the proof of Theorem 1.1.

Lemma 1.5. *Assume* $G \in \mathscr{G}^n$, $H \in \mathscr{G}^l$, $1 \leq l \leq n$, $0 < c < 1$ *and* H *is not isomorphic to an induced subgraph of* G. *Then there are* $i \in [l]$ *and two disjoint subsets* $A, B \subset V = V_G$ *such that*

$$|A| \geq c^{i-1} \left\lfloor \frac{n}{l^2} \right\rfloor, \qquad |B| \geq c^{i-1} \left\lfloor \frac{n}{l} \right\rfloor$$

and either for all $x \in A$, $|G(x) \cap B| \leq c^i \lfloor n/l \rfloor$ *or for all* $x \in A$, $|\bar{G}(x) \cap B| \leq c^i \lfloor n/l \rfloor$.

Proof. Let $V = A_1 \cup \cdots \cup A_l$ be a disjoint partition of V with $|A_i| \geq \lfloor n/l \rfloor$ for $i \in [l]$ and let $y_1, \ldots, y_l$ be an enumeration of V_H. We want to choose a sequence $x_i \in A_i$ for $i \in [l]$ in such a way that

$$H \cong G[\{x_1, \ldots, x_l\}].$$

We use induction on l.

Assume first that there is an $x \in A_1$ such that for $j = 2, \ldots, l$,

$$|G(x) \cap A_j| \geq c|A_j| \text{ if } \{y_1, y_j\} \in E_H \text{ and } |\bar{G}(x) \cap A_j| \geq c|A_j| \text{ if } \{y_1, y_j\} \notin E_H. \tag{1.6}$$

We choose $x_1 \in A_1$ to be such a point and write

$$A_j^1 = \begin{cases} G(x) \cap A_j & \text{if } \{y_1, y_j\} \in E_H, \\ \bar{G}(x) \cap A_j & \text{if } \{y_1, y_j\} \notin E_H \end{cases}$$

for $j = 2, \ldots, l$. Applying the induction hypothesis for the sets $\{A_j^1 : j = 2, \ldots, l\}$ we get that either there are $x_j \in A_j$ such that $H[\{y_2, \ldots, y_l\}] \cong G[\{x_2, \ldots, x_l\}]$ and, as a corollary of this,

$$H \cong G[\{x_1, \ldots, x_l\}]$$

or there are sets A, B satisfying the requirement of the lemma for an i, $2 \leq i \leq l$.

If there are no $x \in A_1$ satisfying (1.6), then for all $x \in A_1$ there is a $j(x) = j$, $2 \leq j \leq l$ such that $|G(x) \cap A_j| < c|A_j|$ in case $\{y_1, y_j\} \in E_H$ and $|\bar{G}(x) \cap A_j| < c|A_j|$ in case $\{y_1, y_j\} \notin E_H$. There is a $j \in [l]$ such that for $A = \{x \in A_1 : j(x) = j\}$ and $B = A_j$ we have $|A| \geq |A_1|/l$ and A, B satisfy the requirements of the lemma with $i = 1$. $\square$

Proof of Theorem 1.1. We may assume $l \geq 3$. Let $H \in \mathscr{G}^l$ and let

$$f_H(n) = \min\{\max\{|A|: A \subset V_G \wedge G[A] \in \mathscr{S}\}: G \in \mathscr{G}^n \wedge H \not\subset\cdot\, \mathscr{G}\}.$$

By Lemma 1.4, it is sufficient to prove that

$$f_H(n) \geq e^{c\sqrt{\log n}} \quad \text{for sufficiently large } n.$$

Set $K = (l \log 2 + 2 \log l)/l$. We prove by induction on n that

$$f_H(n) \geq e^{c\sqrt{\log n} - K} \text{ holds for every } n. \tag{1.7}$$

The claim is trivial in case $e^{c\sqrt{\log n} - K} \leq 2$, as all two-element graphs are very simple. Hence we may assume $c\sqrt{\log n} > K$ and (1.7) holds for all $n' < n$. Write $g(n) = e^{c\sqrt{\log n} - K}$. Let $G \in \mathscr{G}^n$ and $H \not\subset\cdot\, G$.

We apply Lemma 1.5 choosing the c of Lemma 1.5 as $1/(2g(n))$. Let A, B be the two sets given by Lemma 1.5. We may assume

$$n > |A| \geq \frac{n}{\lfloor g(n)^l 2^l l^2 \rfloor} \geq 1, \qquad |B| \geq \frac{n}{\lfloor g(n)^l 2^l l \rfloor}$$

and w.l.o.g.

$$|G(x) \cap B| \leq \frac{1}{2g(n)} |B| \quad \text{for all } x \in A.$$

By the induction hypothesis, there is $A_1 \subset A$,

$$|A_1| \geq g\left(\frac{n}{\lfloor g(n)^l 2^l l^2 \rfloor}\right) \quad \text{with } G[A_1] \in \mathscr{S}.$$

We may assume $|A_1| < g(n)$, otherwise we are home. It follows that for $B' = B \setminus \bigcup\{G(x): x \in A_1\}$ we have

$$|B'| \geq |B| - \tfrac{1}{2}|B| \geq \tfrac{1}{2}|B| \geq \frac{n}{\lfloor g(n)^l 2^l l^2 \rfloor}.$$

Using the induction hypothesis again, we find a $B_1 \subset B'$, $G[B_1] \in \mathscr{S}$ with

$$|B_1| \geq g\left(\frac{n}{\lfloor g(n)^l 2^l l^2 \rfloor}\right).$$

As, by our construction, $G[A_1 \cup B_1]$ is very simple too, and $G \in \mathscr{G}^n \wedge H \not\subset\cdot\, G$ is arbitrary we obtain that

$$f_H(n) \geq 2g\left(\frac{n}{\lfloor g(n)^l 2^l l^2 \rfloor}\right).$$

Now we compute this last number (neglecting the integer part sign) and show that this is at least $g(n)$.

$$2g\left(\frac{n}{g(n)^l 2^l l^2}\right)$$
$$= \exp\{\log 2 + c(\log n - l\log 2 - 2\log l - c\sqrt{\log n} + lK)^{1/2} - K\}$$
$$\geq \exp\{\log 2 + c(\log n - lc\sqrt{\log n})^{1/2} - K\}$$
$$\geq \exp\{c\sqrt{\log n} - K\}$$

provided $c(\log n^{1/2} - (\log n - lc\log n)^{1/2}) \leq \log 2$.

But $\log n^{1/2} - (\log n - lc\sqrt{\log n})^{1/2} \leq c$ holds trivially, hence the left-hand side of the last inequality is at most c, since $cl < 1$, and $c < \log 2$, holds by $c < \frac{1}{3}$. □

Note that, as the proof shows, Theorem 1.1 remains true if l tends to infinity slowly enough, say if $l = o(\sqrt{\log n}/\log\log n)$.

Proof of Theorem 1.2. It is clearly sufficient to prove the following statement.

Assume $H \in \mathscr{G}^{2k}$, $k \in \mathbb{N}$, $V_H = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$, $|V_1| = |V_2| = k$, $[V_1, V_2] \subset E_H$. Assume further that there exists $0 < \varepsilon_1 < 1$ such that for sufficiently large n and for all $G \in \mathscr{G}^n$ not containing homogeneous subsets of size at least n^{ε_1}, $H_1, H_2 \Subset G$ for $H_i = H[V_i]$, $i = 1,2$.

Let $\varepsilon_2 = \varrho\varepsilon_1$, $\varrho < 1/(2k+1)$. Then for all $G \in \mathscr{G}^{n^\varrho}$ not containing homogeneous subsets of size n^{ε_2}, $H_1 \Subset G$ holds for all sufficiently large n. (1.8)

To prove (1.8) assume $G \in \mathscr{G}^n$ and G does not contain homogeneous sets of size n^{ε_2}. Then, for all $A \in [V]^{n^\varrho}$, there is a $B \in [A]^k$ with $G[B] \cong H_1$ provided n is large enough. We may assume in the proof that $|G(x)| \geq n^{l-\varepsilon_2}$ holds for all $x \in V$ otherwise we get a large independent set. Let $\mathscr{F} = \{(B,x)\colon B \in [V]^k \wedge x \in V \wedge B \subset G(x) \wedge H_1 \cong G[B]\}$. By averaging, we get

$$|\mathscr{F}| \geq n\binom{n^{1-\varepsilon_2}}{n^\varrho}\binom{n^{1-\varepsilon_2}-k}{n^\varrho - k}^{-1} \geq c_1 n^{1+(1-\varepsilon_2-\varrho)k}$$

for some $c_1 > 0$.

Hence there is a $B_1 \in [V]^k$ such that for $A = \{x \in V \setminus B_1\colon B_1 \subset G(x)\}$ we have

$$|A| \geq c_2 n^{1-(\varepsilon_2+\varrho)k} \quad \text{for some } c_2 > 0.$$

Now, if $\varepsilon_1(1-(\varepsilon_2+\varrho)k) \geq \varepsilon_2$, then there is a $B_2 \subset A$, $|B_2| = k$ such that $G[B_2] \cong H_2$, and then $G[B_1 \cup B_2] \cong H$. But the inequality required is equivalent to the inequality $1 \geq \varrho(1+\varepsilon_1)k + \varrho$, which follows from the assumption $\varrho < 1/(2k+1)$. □

Note that the proof yields that for $H \in \mathscr{G}^{2^t}$, $\varepsilon = [(2t+1)(2t-1)\cdots 1]^{-1}$ works. This bound is clearly not best possible, but not very bad either.

2. Large weakly homogeneous r-sequences

Let $G=\langle V,E\rangle$ be a graph, $r\in\mathbb{N}$. A sequence $\{A_1,\dots,A_t\}\subset[V]^r$ of pairwise disjoint r-element subsets of V will be called a *weakly complete* r-sequence of length t for G, if for every pair $\{i,j\}\in[t]^2$, $[A_i,A_j]$ contains an edge of G.

$\{A_1,\dots,A_t\}$ will be called *weakly homogeneous* for G if it is weakly complete, either for G or $\bar{G}$. It is clear that weakly homogeneous 1-sequences correspond canonically to homogeneous sets, and the larger the r, the easier it is to get weakly homogeneous r-sequences.

Large graphs not containing long weakly homogeneous r-sequences, $r\in\mathbb{N}$, play an important role in infinite combinatorics and especially in applications to topology. Though it is not clear (even to us) that the finite version of this problem is very useful, in this section we give estimates on the length of the weakly homogeneous r-sequences necessarily contained in graphs on n vertices. Our main motivation is that this investigation leads to a problem which seems to be of intrinsic interest.

Theorem 2.1. (A) *Assume $c>2^{r^2+1}r$. There is an n_0 such that for all $n>n_0$ there is a $G\in\mathscr{G}^n$ without weakly homogeneous r-sequences of length $c\log n$.*

(B) *Assume $c<(\frac{3}{2})^r$. There is an n_0 such that for all $n>n_0$ every $G\in\mathscr{G}^n$ contains a weakly homogeneous r-sequence of length at least $\frac{1}{2}c\log n$.*

Proof. The proof of (A) is a straightforward probabilistic computation. Taking each edge with probability $\frac{1}{2}$,

$$\Pr(\text{There is a weakly complete } r\text{-sequence of length } t \text{ for } G)$$
$$\le n^{tr}\left(1-\frac{1}{2^{r^2}}\right)^{\binom{t}{2}}<1 \quad \text{if } t>2^{r^2+1}r\log n.$$

For the proof of (B) we need the following:

Lemma 2.2. $\forall r\in\mathbb{N}\ \forall\varepsilon>0\ \exists n_0\ \forall n>n_0\ \forall G\in\mathscr{G}^n\ \exists A\in[V]^r$ *such that either*

$$|G(A)|\ge(1-(\tfrac{2}{3})^r(1+\varepsilon))n \quad \textit{or} \quad |\bar{G}(A)|\ge(1-(\tfrac{2}{3})^r(1+\varepsilon))n.$$

Before giving the proof we remark that the example of the random graph shows that $\frac{2}{3}$ cannot be replaced by anything smaller than $\frac{1}{2}$ and we are left with the annoying problem, if Lemma 2.2 holds with $\frac{1}{2}$ instead of $\frac{2}{3}$. Note that we can prove this for $r=1,2$ only.

Proof of Lemma 2.2. For $r=1$ the claim is trivial. We use induction on r. Let $\Delta_G(A)=\{x\in V\colon \forall y\in A,\ \{x,y\}\in E\}$, and let d_x denote the degree of the vertex in G, i.e., $d_x=|G(x)|$ for $x\in V$. We set $\bar{d}_x$ for $|\bar{G}(X)|$.

Assume first that there are $x,y\in V$ with $d_x>\frac{2}{3}n$ and $\bar{d}_y>\frac{2}{3}n$. Then $|G(x)\cap$

$\bar{G}(y)| > \frac{1}{3}n$. Let $V_1 \subset G(x) \cap \bar{G}(y)$, $|V_1| = \lceil \frac{1}{3}n \rceil$ and $V_2 = V \setminus V_1$. Applying the induction hypothesis for $G[V_2]$ we get an $A' \in [V_2]^{r-1}$ such that either $G(A') \cap V_2$ or $\bar{G}(A') \cap V_2$ has cardinality at least $\frac{2}{3}n(1-(\frac{2}{3})^r(1+\varepsilon))$ for large enough n. Clearly either $A = \{x\} \cup A'$ or $A = \{y\} \cup A'$ satisfies the requirements of Lemma 2.2.

Hence w.l.o.g. we may assume that $d_x \leq \frac{2}{3}n$ for $x \in V$. Let $\mathscr{F} = \{(A,x): A \in [V]^r \wedge x \in V \wedge A \subset G(x)\}$. Counting the number of pairs in $\mathscr{F}$ in two ways, we get that

$$\sum_{x \in V} \binom{d_x}{r} = \sum_{A \in [V]^r} |\Delta_G(A)|.$$

This implies that $\min\{|\Delta_G(A)|: A \in [V]\} \leq n(1+\varepsilon)(\frac{2}{3})^r$ if n is large enough. □

Returning to the proof of Theorem 2.1(B), it follows from Lemma 2.2, by a standard argument, that for $c < (\frac{2}{3})^r$ we can pick a sequence $A_1', \dots, A_{t'}'$, $t' = \lceil c \log n \rceil$ of pairwise disjoint r-element sets in such a way that for $1 \leq i < t'$ either

$$\bigcup_{j=i+1}^{t} A_j' \subset G(A_i') \quad \text{or} \quad \bigcup_{j=i+1}^{t} A_j' \subset \bar{G}(A_i').$$

We then can choose $t \geq \frac{1}{2}t'$ sets $A_1, \dots, A_t$ from the sets $A_1', \dots, A_t'$ such that either $A_j \subset G(A_i)$ for $1 \leq i < j \leq t$ or $A_j \subset \bar{G}(A_i)$ for $1 \leq i < j < t$. The sequence $A_1, \dots, A_t$ is then weakly homogeneous for G. □

Note that the proof gives a sequence having a stronger property than required by Lemma 2.2, and it can be seen with a slight change of the probabilistic argument used for the proof of Theorem 2.1(A) that such sequences of length $2^r r \log n$ do not necessarily exist for a $G \in \mathscr{G}^n$. We omit the details.

3. *r*-almost homogeneous subsets of a graph

Let $G = \langle V, E \rangle$ be a graph, $r \in \mathbb{N}$, $A \subset V$. A is said to be *r-almost complete* for G if $\bar{G}[A]$ does not contain a K_{r+1}. It is clear that 1-almost complete sets are complete.

A is said to be *r-almost homogeneous* for G if it is r-almost complete either for G or for $\bar{G}$. Clearly 1-almost homogeneous means the same as homogeneous.

One can define a generalization of the Ramsey function $R(k,l)$ say $R(k|r,l|s) = \min\{n: \forall G \in \mathscr{G}^n$ either there is an $A \subset V$, $|A| = k$, A is r-almost complete for G or there is a $B \subset V$, $|B| = l$, B is s-almost complete for $\bar{G}\}$. There is no existence problem, since clearly

$$R(k|r,l|s) \leq R(k,l) \quad \text{for } r,s \in \mathbb{N}.$$

It is also clear that this function can be generalized for partitions to more that 2-classes, and for hypergraphs as well. Indeed we raised the problem in [11] in full generality, but we treated there the infinitary problems only. In this case we are more convinced that the problems arising are relevant. Unfortunately on this problem we have less information than on the problem treated in Section 2.

First we give a rather weak result for the symmetric case, $R(k\,|\,r)=R(k\,|\,r,k\,|\,r)$ and then we state some miscellaneous remarks and problems about the non-symmetric case.

Theorem 3.1. (A) $\exists c>0\ \forall r,n\in\mathbb{N}\ \exists G\in\mathcal{G}^n$ *having no r-almost homogeneous subsets of size*

$$c2^{\binom{r+1}{2}}(r+1)^2\log n.$$

(B) $\forall r\ \exists n_0\ \forall n>n_0\ \forall G\in\mathcal{G}^n$ *there is an* $(r+1)$*-almost homogeneous subset of size at least*

$$\frac{r\log n}{2\log 2}.$$

It is left to the reader to compute the estimates for $R(k\,|\,r)$ from these results.

Proof. (A) We work in the probability space where the edges are chosen independently with probability $\frac{1}{2}$. We need the following fact.

$$\text{There is a constant } c_0>0 \text{ such that for all } k,r\in\mathbb{N} \text{ and for all } |A|=k \text{ there is a system } \mathcal{F}\subset[A]^{r+1} \text{ such that } |\mathcal{F}|\geq c_0k^2/(r+1)^2 \text{ and } |X\cap Y|\leq 1 \text{ holds for } X\neq Y\in\mathcal{F}. \tag{3.1}$$

This follows from a theorem of Wilson [17]. Let $\mathcal{A}$ be the event that $A\in[V]^k$ is r-almost complete, and B_X the event that X is not complete for $\bar{G}$, for $X\in\mathcal{F}$. Then

$$\Pr(B_X)\leq 1-\frac{1}{2^{\binom{r+1}{2}}},\qquad \Pr(\mathcal{A})\leq\Pr\Big(\bigcap_{X\in\mathcal{F}}B_x\Big)$$

and the B_X being pairwise independent,

$$\Pr(\mathcal{A})\leq\left(1-\frac{1}{2^{\binom{r+1}{2}}}\right)^{c_0k^2/(r+1)^2}$$

It follows that

$$\Pr(\text{There is an } r\text{-almost homogeneous subset of size } k)$$

$$\leq 2\binom{n}{k}\left(1-\frac{1}{2^{\binom{r+1}{2}}}\right)^{c_0k^2/(r+1)^2}<1 \text{ provided } k>c2^{\binom{r+1}{2}}(r+1)^2\log n$$

for some c.

(B) If n is large enough and $G\in\mathcal{G}^n$, then by the Erdös–Szekeres theorem (1.1), there are $2r-1$ pairwise disjoint homogeneous subsets of size at least $\frac{1}{2}(\log n)/$

(log 2) for G. Either r of them is complete or r of them is independent. The union of these r sets is r-almost homogeneous. □

Clearly Theorem 3.1(B) is very weak, but we know no nontrivial improvement. Here is a very simple problem. Let $r(n)$ be the inverse of the Ramsey function, i.e.,

$$r(n)=\min\{\max\{|A|: A\subset G\wedge A \text{ is homogeneous for } G\}: G\in\mathscr{G}^n\}.$$

Is it true that for all (or for many) n for all $G\in\mathscr{G}^n$ there is an $A\subset V$ that is 2-almost homogeneous for G of size $2r(n)+1$?

Turning to the nonsymmetric case we remind the reader the Erdös-Szekeres theorem

$$R(k,l)\le\binom{k+l-2}{k-1}.$$

We cannot even formulate a conjecture for the upper bound on $R(k\,|\,r, l\,|\,s)$ in the general case. We are going to make some comments on the relation

$$n=R(k\,|\,1, l\,|\,s)=R(k, l\,|\,s). \tag{3.2}$$

The following is just a restatement of some old results.

Theorem 3.2. *Assume* $s\ge 2$. *Then in* (3.2)

(A) $n=k+s-1$ *for* $l\le 2s$;

(B) $k^{1+\eta}<n<ck^{1+1/s}$ *for* $l=2s+1$ *for some* $\eta, c>0$;

(C) $R(k,5\,|\,2)\le ck^{4/3}$ *for some* $c>0$.

Proof (Sketch). (A) can be reformulated as follows: Assume $r\in\mathbb{N}$ and for all $A\subset V$, $|A|\le 2r+2$, $\tau(G[A])\le r$ holds. Then $\tau(G)\le r$. Here $\tau(G)$ is the minimal number of points representing all edges. This is a theorem of Erdös and Gallai [5].

The upper bound in (B) follows from [4]. If every $(2s+1)$-element set contains a K_{s+1} of $G\in\mathscr{G}^n$, then $C_{2s+1}\not\subset\bar{G}$. Then, by the theorem mentioned, there is a $K_k\subset G$ provided $n>ck^{(1+1/s)}$. To prove the lower bound we use the following theorem of [2]. For all $s\in\mathbb{N}$ there are $\varepsilon>0$ and n_0 such that for all $n>n_0$ there is an $H\in\mathscr{G}^n$ with girth greater than $2s+1$ and not containing an independent set of size $n^{1-\varepsilon}$. $G=\bar{H}$ shows that (B) holds. Note that $\bar{H}[A]$ is bipartite on every set $|A|\le 2s+1$, hence $K_{s+1}\subset H[A]$. (C) is just a better lower bound then the one obtained in (B) for the special case $s=2$. The proof of (C) is implicitly contained in [16]. □

As we have already mentioned the Ramsey function for almost homogeneous sets can be generalized for triple systems (3-uniform hypergraphs). We define a special case $R^3(k,l\,|\,s)$ is the minimal n such that for all triple systems $H=\langle V,E\rangle$, $E\subset[V]^3$, with $|V|=n$ either $K_k^3\subset H$, i.e., $\exists A\subset V\ |A|=k\wedge[A]^3\subset E$ or $\exists B\subset V$ $|B|=l$ such that $K_{s+1}^3\not\subset H[B]$.

We now discuss an analogue of the previous theorem, concerning the relation

$$n = R(k, l \mid 3). \tag{3.3}$$

Theorem 3.3. *In* (3.3)
(A) $n = k$ *for* $l = 4$;
(B) $\exists \alpha > 1$ $k^\alpha < n < k^3$ *for* $l = 5$;
(C) $\exists c > 0$ $2^{ck} < n$ *for* $l = 8$.

We do not know the right order of magnitude for $l = 6, 7$.

Proof. (A) is trivial. For the lower bound of (B) we use a theorem of Erdös and Hajnal [8], a generalization of the Erdös theorem used in the previous proof. A triple system K is t-circuitless if for every $K' \subset K$, $1 \le |K'| \le t$, $|\bigcup K'| > 2|K'|$. We proved that for every t there are $\varepsilon > 0$ and n_0 so that for every $n > n_0$ there is a t-circuitless triple system K on n vertices having no independent set of size $n^{1-\varepsilon}$. Applying this with $t = 3$ and choosing $H = \bar{K}$, we see that every 5-element set contains a K_4^3 of H and H has no complete set of size $n^{1-\varepsilon}$. Set $\alpha = (1-\varepsilon)^{-1}$.

The upper bound is due to Füredi. Assume H is a triple system such that every 5-element set contains a K_4^3 of H. For each $e \in [V]^2$ let $\bar{H}(e) = \{x \in V\colon e \cup \{x\} \in \bar{H}\}$. As, by the assumption, $\bar{H}(e)$ is a complete set of H for every $e \in [V]^2$, we may assume $|H(e)| \le n^{1/3}$. If now A is a maximal complete set of H, then

$$V \setminus A = \bigcup \{\bar{H}(e)\colon e \in [A]^2\},$$

consequently $n - |A| \le \binom{|A|}{2} n^{1/3}$ from which we get $|A| \ge n^{1/3}$.

Finally, to see (C) consider a random tournament on n vertices and let H be the collection of nontransitively oriented triples.

As every tournament on 8 vertices contains a transitive subtournament of four vertices, H satisfies our requirements. For triple systems constructed in such a way see e.g. [9]. $\square$

4. Remarks on the Ramsey function for 3-partitions

We will denote by $\mathcal{H}^{3,n}$ the class of 3-uniform hypergraphs. For $H \in \mathcal{H}^{3,n}$, $H = \langle V_H, E_H \rangle = \langle V, E \rangle$ where $E \subset [V]^3$. $\bar{E}$ denotes the complement of E. If $f\colon [V]^3 \to [2]$ the formula $f(e) = 1 \Leftrightarrow e \in E_H$ defines a corresponding hypergraph. The Ramsey function $R^3(k_1, \ldots, k_s)$ is defined as usual:

$$R^3(k_1, \ldots, k_s) = \min\{n : \forall f\colon [n]^3 \to [s]\ \exists i \in [s]\ \exists A \in [n]^{k_i}\ f \text{ is constant on } [A]^3\},$$

$$R^3(k_1, \ldots, k_s) = R_s^3(k) \quad \text{for } k_1 = \cdots = k_s = k.$$

We remind the reader of the following facts:

Let $n=R_2^3(k)$. Then for some $c>0$,

$$c\log\log n\le k\le\left(\frac{6\log n}{\log 2}\right)^{1/2}. \tag{4.1}$$

But:

For $n=R_4^3(k)$, for some $c>0$, $k\le c\log\log n$. (4.2)

Fact (4.1) is due to Erdös and Rado, (4.2) is due to Hajnal. For both statements see [10]. Note that the upper estimate in (4.1) is obtained by probabilistic methods. It is a concensus among experts that the probabilistic method as it is known today cannot help to close the huge gap in (4.1), while the method used for the proof of (4.2) does not seem to work for fewer then 4 colors.

We are going to prove two theorems:

Theorem 4.1. *For $n=R_3^3(k)$,*

$$k<\left(\frac{\log n\log\log\log n}{\log\log n}\right)^{1/2}$$

for n large enough.

Here we did not bother to get the best possible upper bound. We only wanted to get an essentially better upper bound then the one given by the probabilistic argument.

To state the next result we need one more definition. For $H=\langle V,E\rangle\in\mathscr{H}^{3,n}$, $A\subset V$ we define the density of H on A, $D_H(A)$, as follows

$$D_H(A)=|E\cap[A]^3|\Big/\binom{|A|}{3}.$$

The next result is kind of a discrepancy result contrasting the situation known for graphs.

Theorem 4.2.

$$\forall c<1/(2\log 2)\ \exists n_0\ \forall n>n_0\ \forall H\in\mathscr{H}^{3,n}\ \exists A\subset V$$
$$\tfrac{1}{5}c\sqrt{\log n}\le|A|\le c\sqrt{\log n}\wedge|D_H(A)-\tfrac{1}{2}|\ge\tfrac{1}{24}\sqrt{3}.$$

Before giving the proofs we state a few questions raised by this result. Can one improve this result by specifying the cardinality of A? Can one exhibit a "fixed configuration" of density $>\frac{1}{2}$ to occur monochromatically? Finally, can one replace $\frac{1}{24}\sqrt{3}$ by $\frac{1}{2}-\varepsilon$ for $\varepsilon>0$ for sufficiently small c?

Proof of Theorem 4.1. Let $n=r^s$, $r,s\ge 2$. We are going to define a 3-coloring of the set $V=\{x\colon x=(x_1,\dots,x_s)\wedge x_i\in[r]$ for $i\in[s]\}$. Note that $|V|=r^s=n$. For $x\neq y\in V$ let $\delta(x,y)=\min\{i\in[r]\colon x_i\neq y_i\}$ be the so called first discrepancy for x and y.

First we split the triples of V into two parts. For $\{x,y,z\}\in[V]^3$ let $\Delta(x,y,z)=\{\delta(x,y),\delta(x,z),\delta(y,z)\}$. Let $\{x,y,z\}\in\mathscr{A}$ if $|\Delta(x,y,z)|\geq 2$ and $\{x,y,z\}\in\mathscr{B}$ if $|\Delta(x,y,z)|=1$. For $\{x,y,z\}\in\mathscr{B}$ there is a $\delta(x,y,z)=i$ such that

$$x_j=y_j=z_j \quad \text{for } j<i\wedge j\in[s], \qquad \{x_i,y_i,z_i\}\in[r]^3.$$

Put briefly $c_0=(\frac{1}{6}\log 2)^{1/2}$. By (4.1), there is a mapping $g\colon[r]^3\to[2]$ without a homogeneous set of size $c_0\sqrt{\log r}$.

Define a mapping $f\colon[V]^3\to[3]$ as follows. For $\{x,y,z\}\in[V]^3$ let $f(\{x,y,z\})=3$ if $\{x,y,z\}\in\mathscr{A}$ and $f(\{x,y,z\})=g(\{x_i,y_i,z_i\})$ if $\{x,y,z\}\in\mathscr{B}$ and $\delta(x,y,z)=i$.

Assume now that $A\subset[V]^3$ is homogeneous for f in the class $\nu\in[3]$, $|A|\geq 3$.

If $\nu=1$ or $\nu=2$ it is obvious that there is an $i\in[s]$ such that $\delta(x,y,z)=i$ for all $\{x,y,z\}\in[A]^3$ and $\{x_i\colon x\in A\}$ is homogeneous for the partition g. Hence $|A|\leq c_0\sqrt{\log r}$.

Assume that A is homogeneous for the third colorclass of f. Then $|\Delta(x,y,z)|\geq 2$ for $\{x,y,z\}\in[A]^3$. Clearly $|\{x_1\colon x\in A\}|\leq 2$ and also for $i\in[s-1]$ and $(y_1,\dots,y_i)$ with $y_j\in[r]$ for $j\in[i]$,

$$|\{x_{i+1}\colon x\in A\wedge\forall j\in[i]\ x_j=y_j\}|\leq 2.$$

It follows easily that $|A|\leq 2^s$ for such an A.

With a suitable choice of r and s this partition gives a result even better we claimed.

We just indicate the computation. Choose

$$\log r=\frac{1}{c_0^2}\,\frac{\log n\,\log_{(3)}(n)}{\log_{(2)}(n)}.$$

Then

$$r=n^{(1/c_0^2)(\log_{(3)}(n)/\log_{(2)}(n))}, \qquad x=c_0^2\,\frac{\log_{(2)}(n)}{\log_{(3)}(n)}.$$

The first two colors do not contain a homogeneous set of the claimed size, while

$$2^x=\log n^{c_0^2(1/\log_{(3)}(n))}=o\left(\left(\frac{\log n\,\log_{(3)}(n)}{\log_{(2)}(n)}\right)^{1/2}\right). \qquad \square$$

For the proof of Theorem 4.2 we need the following:

Lemma 4.3 (Erdös [3]). *There are an n_0 and a $c>0$ such that for all $n>n_0$ and for all $H\in\mathscr{H}^{3\ n}$ there are two disjoint sets A, B with*

$$|A|\geq\left\lfloor\frac{\sqrt{\log n}}{2\log 2}\right\rfloor, \qquad |B|=cn^{1-1/(2\log 2)}$$

such that either $[A,B]^{2,1}\subset E_H$ or $[A,B]^{2,1}\cap E_H=\emptyset$ where

$$[A,B]^{2,1} = \{X \subset A \cup B\colon\ |X \cap A| = 2 \wedge |Y \cap B| = 1\}.$$

For the convenience of the reader we outline the proof. Let $A_1 \subset V$, $|A_1| = \lceil e^{\sqrt{\log n}} \rceil$. For $x \in V \setminus A_1$ define a graph $G_x = \langle A_1, E_x \rangle$ in such a way that for $e \in [A_1]^2$,

$$e \in E_x \Leftrightarrow e \cup \{x\} \in E = E_H.$$

By (1.1), for each $x \in V \setminus A_1$ there is an $A_x \subset A_1$ homogeneous for G_x, $|A_x| = \lfloor \sqrt{\log n}/2 \log 2 \rfloor$. As

$$|[A_1]^{\lfloor \sqrt{\log n}/2 \log 2 \rfloor}| \le n^{1/(2 \log 2)},$$

there is a set

$$B_1 \subset V \setminus A_1, \quad |B_1| \ge |V \setminus A_1| n^{-1/(2 \log 2)}$$

such that $A_x = A$ for $x \in B_1$ and there is $|B \subset B_1|$, $|B| \ge \frac{1}{2}|B_1|$ satisfying the requirements of the lemma. □

Proof of Theorem 4.2. Assume $c < 1/(2 \log 2)$, $H = \langle V, E \rangle \in \mathscr{H}^{3,n}$. Put $m = c \log n$. Let $\alpha = \frac{1}{2} + \frac{1}{6}\sqrt{3}$, $\beta = \frac{1}{2} - \frac{1}{6}\sqrt{3}$. If n is sufficiently large then, by Lemma 4.3, there are disjoint sets $A, B \subset V$, $|A| = \alpha m$, $|B| = \beta m$ such that either $[A,B]^{2,1} \subset E$ or $[A,B]^{2,1} \cap E = \emptyset$. W.l.o.g. we may assume $[A,B]^{2,1} \subset E$.

Note that, by the choice of α and β, we have $\frac{1}{5}m \le |A|$, $|B| \le m$. So we may assume $|D_H(A) - \frac{1}{2}| \le \frac{1}{24}\sqrt{3}$ and $|D_H(B) - \frac{1}{2}| \le \frac{1}{24}\sqrt{3}$ otherwise we are home.

Let $C = A \cup B$. Then

$$D_H(C) \ge \frac{D_H(A)\alpha^3 \frac{1}{6}m^3 + D_H(B)\beta^3 \frac{1}{6}m^3 + \alpha^2 \beta \frac{1}{2}m}{\frac{1}{6}m^2}$$
$$\ge \tfrac{1}{2}(\alpha^3 + \beta^3) + 3\alpha^2\beta - \tfrac{1}{24}(\alpha^3 + \beta^3)\sqrt{3}.$$

Now the reader should compute that for the choice of α and β we made, $\frac{1}{2}(\alpha^3 + \beta^3) + 3\alpha^2\beta = \frac{1}{2} + \frac{1}{12}\sqrt{3}$ holds. Indeed if $\alpha = \frac{1}{2} + \frac{1}{6}\sqrt{3}$, $\beta = \frac{1}{2} - \frac{1}{6}\sqrt{3}$, then from $1 = \alpha^3 + 3\alpha^2\beta + 3\alpha\beta^2 + \beta^3$ we get

$$\tfrac{1}{2}(\alpha^3 + \beta^3) = \tfrac{1}{2} - \tfrac{3}{2}(\alpha^2\beta + \alpha\beta^2) = \tfrac{1}{2} - \tfrac{3}{2}\alpha\beta,$$

$$\tfrac{1}{2}(\alpha^2 + \beta^3) + 3\alpha^2\beta = \tfrac{1}{2} + 3\alpha^2\beta - \tfrac{3}{2}\alpha\beta = \tfrac{1}{2} + 3\alpha(1-\alpha)(\alpha - \tfrac{1}{2})$$
$$= \tfrac{1}{2} + 3(\tfrac{1}{4} - \tfrac{3}{36})\tfrac{1}{6}\sqrt{3} = \tfrac{1}{2} + \tfrac{1}{12}\sqrt{3}.$$

Since $\alpha^3 + \beta^3 \le 1$ if follows that $D_H(C) \ge \frac{1}{2} + \frac{1}{24}\sqrt{3}$.

Though we wrote down the proof above with some abuse of the equality sign, we consider the proof finished. □

Note added in proof

We were wrong in guessing that Theorem 3.1(B) is weak. It is Theorem 3.1(A) that can be improved considerably, in fact it is almost exact.

Using the same random graph we can prove $\forall r \in \mathbb{N}\ \forall \varepsilon > 0\ \exists n_0\ \forall n > n_0\ \exists G \in G^n$ having no $(r+1)$-almost homogeneous subsets of size $((2r+\varepsilon)/\log 2)\log n$. This follows from the result stated below.

If $G_r(k)$ is the number of graphs with k vertices not containing K_r as a subgraph, then

$$\frac{\log(G_r(k))}{\log 2} = \frac{k^2}{2}\left(1 - \frac{1}{r-1}\right) + o(k^2).$$

This in turn was proved in [18].

The estimate given in our original proof gives some orientation in case r tends to infinity, in a range where the estimate for $G_r(k)$ is no longer valid.

An affirmative answer to the problem stated below Lemma 2.2 is given in the paper "Domination in colored complete graphs", by P. Erdös, R. Faudree, A. Gyárfás and R.H. Schelp, to appear in Journal of Graph Theory.

Acknowledgment

We are indebted to M. Simonovits for valuable conversations on the topics treated in this paper.

References

[1] P. Erdös, Some remarks on the theory of graphs, Bull. Amer. Math. Soc. 53 (1947) 292-294.
[2] P. Erdös, Graph theory and probability, Canad. J. Math. 11 (1959) 34-38.
[3] P. Erdös, On some extremal problems on r-graphs, Discrete Math. 1 (1971) 1-6.
[4] P. Erdös, R.J. Faudree, C.C. Rousseau and R.H. Schelp, On cycle-complete graph Ramsey numbers, J. Graph Theory 2(1) (1978) 53-64.
[5] P. Erdös and T. Gallai, On the maximal number of vertices representing the edges of a graph, MTA MKI Közl. 6 (1961) 181-204.
[6] P. Erdös and A. Hajnal, Embedding theorems for graphs establishing negative partition relations, Period. Math. Hungar. 9 (1978) 205-229.
[7] P. Erdös and A. Hajnal, On spanned subgraphs of graphs, Graphentheorie und Ihre Anwendungen (Oberhof, 1977).
[8] P. Erdös and A. Hajnal, On chromatic number of graphs and set-systems, Acta Math. Acad. Sci. Hungar. 17 (1966) 61-99.
[9] P. Erdös and A. Hajnal, On Ramsey-like theorems, problems and results, in: Proceedings of the Oxford Conference (1972).
[10] P. Erdös, A. Hajnal, Máté and R. Rado, Combinatorial set theory, Partition relations for cardinals, in: Studies in Logic and the Foundations of Mathematics (North-Holland, Amsterdam, 1984).

[11] P. Erdös, A. Hajnal and R. Rado, Partition relations for cardinals, Acta Math. Acad. Sci. Hungar. 16 (1965) 93–196.

[12] P. Erdös and G. Szekeres, A combinatorial problem in geometry, Compositio Math. 2 (1931) 463-470.

[13] P. Erdös and E. Szemerédi, On a Ramsey type theorem, Period. Math. Hungar. 2 (1972) 295–299.

[14] P. Frankl and R.M. Wilson, Intersection theorems with geometric consequences, Combinatorica 1 (1981) 357–368.

[15] H. Meynier, On the perfect graph conjecture, Discrete Math. 16 (1976) 339–342.

[16] J. Spencer, Asymptotic lower bounds for Ramsey functions, Discrete Math. 20 (1977) 69–76.

[17] R.M. Wilson, An existence theory for pairwise balanced designs I, II, Combin. Theory Ser. A 13 (1972) 220–245, 246–273.

[18] P. Erdös, D.J. Kleitman and B.L. Rothschild, Asymptotic enumeration of K_n-free graphs, Call. Intern sulle Theorie Combinatorie, Rome (1973).