Записки научных
семинаров ПОМИ
Том 488, 2019 г.

D. D. Cherkashin

ON THE ERDŐS–HAJNAL PROBLEM IN THE CASE OF 3-GRAPHS

Abstract. Let $m(n, r)$ denote the minimal number of edges in an n-uniform hypergraph which is not r-colorable. For the broad history of the problem see [10]. It is known [4] that for a fixed n the sequence

$$\frac{m(n,r)}{r^n}$$

has a limit.

The only trivial case is $n = 2$ in which $m(2, r) = \binom{r+1}{2}$. In this note we focus on the case $n = 3$. First, we compare the existing methods in this case and then improve the lower bound.

§1. Introduction

A hypergraph $H = (V, E)$ consists of a finite set of *vertices* V and a family E of the subsets of V, which are called *edges*. A hypergraph is called *n-uniform* if every edge has size n. A *vertex r-coloring* of a hypergraph $H = (V, E)$ is a map from V to $\{1, \dots, r\}$. A coloring is *proper* if there are no monochromatic edges, i.e., any edge $e \in E$ contains two vertices of different color. The *chromatic number* of a hypergraph H is the smallest number $\chi(H)$ such that there exists a proper $\chi(H)$-coloring of H. Let $m(n, r)$ be the minimal number of edges in an n-uniform hypergraph with chromatic number more than r.

Erdős and Hajnal [7] introduced problems on determining $m(n, r)$ and related quantities. We are interested in the case when n is much smaller than r (see [10] for general case and related problems).

1.1. Upper bounds. Erdős conjectured [6] that

$$m(n,r) = \binom{(n-1)r+1}{n},$$

for $r > r_0(n)$, that is achieved on the complete hypergraph.

Key words and phrases: extremal combinatorics, hypergraph colorings.
The work was supported by the Russian Scientific Foundation grant 16-11-10014.

However Alon [3] disproved the conjecture by using the estimate

$$m(n,r) < \min_{a \geqslant 0} T(r(n+a-1)+1, n+a, n),$$

where the Turán number $T(v,k,n)$ is the smallest number of edges in an n-uniform hypergraph on v vertices such that every induced subgraph on k vertices contains an edge. Different bounds on Turán numbers refine the complete hypergraph construction when $n > 3$ (see [11] for a survey). So the case $n = 3$ is in some sense the most interesting.

Also note, that using the same inequality with better bounds on Turán numbers Akolzin and Shabanov [2] showed that

$$m(n,r) < Cn^3 \ln n \cdot r^n.$$

Alon conjectured that the sequence $m(n,r)/r^n$ has a limit which was proved by Cherkashin and Petrov [4]. Denote the corresponding limit by L_n. In this paper we are interested in estimates on L_3. The best known upper bound follows from the complete hypergraph:

$$L_3 \leqslant \frac{4}{3}.$$

1.2. Lower bounds. There are several ways to show an inequality of type $m(n,r) \geqslant c(n)r^n$ (i.e. $L_n \geqslant c(n)$). Note that Erdős conjecture implies in particular that

$$L_n = \frac{(n-1)^n}{n!}.$$

Alon [3] suggested to color vertices of an n-uniform hypergraph in $a < r$ colors uniformly and independently, and then recolor a vertex in every monochromatic edge in unused color. The expected number of monochromatic edges is

$$|E| \cdot a^{1-n}.$$

Note that we have $r - a$ remaining colors, and we can color $n - 1$ vertices in each unused color such that no new monochromatic edge appears. Summing up, if

$$|E| < a^{n-1}(r-a)(n-1)$$

then a hypergraph $H = (V,E)$ has a proper r-coloring. Substituting $a = \left\lfloor \frac{n-1}{n} r \right\rfloor$, we get

$$m(n,r) \geqslant (n-1) \left\lceil \frac{r}{n} \right\rceil \left\lfloor \frac{n-1}{n} r \right\rfloor^{n-1}.$$

This method gives $L_3 \geqslant 8/27 > 0.296$.

Another way is due to Pluhár [9]. He introduced the following useful notion. A sequence of edges $a_1, \dots, a_r$ is an *r-chain* if $|a_i \cap a_j| = 1$ if $|i - j| = 1$ and $a_i \cap a_j = \varnothing$ otherwise; it is an *ordered r-chain* if $i < j$ implies that every vertex of a_i is not bigger than any vertex of a_j (with respect to a certain fixed linear ordering on V).

Pluhár's theorem states that existence of an order on V without ordered r-chains is equivalent to r-colorability of $H = (V, E)$. Let us prove a lower bound on $m(n, r)$ via this theorem. Consider a random order on the vertex set. Note that the probability of an r-chain to be ordered is

$$\frac{[(n-1)!]^2[(n-2)!]^{r-2}}{((n-1)r+1)!}.$$

From the other hand, the number of r-chains is at most $2|E|^r/r!$ since every set of r edges generates at most 2 chains. So if

$$2\frac{|E|^r}{r!}\frac{[(n-1)!]^2[(n-2)!]^{r-2}}{((n-1)r+1)!} < 1,$$

then we have a proper r-coloring of H. After taking r-root and some calculations we have

$$m(n, r) > c\sqrt{n}r^n,$$

and in particular $L_3 \geqslant 4/e^3 > 0.199$.

Combining two previous arguments with Cherkashin–Kozik approach [5] Akolzin and Shabanov [2] proved that

$$m(n, r) \geqslant c\frac{n}{\ln n}r^n,$$

without explicit bounds on c. We show that this method gives the bound $L_3 > 0.205$ in Section 3.

Cherkashin and Petrov [4] suggested an approach, based on the evaluation of the inverse function, to show that the sequence $m(n, r)/r^n$ has a limit. Denote by $f(N)$ the maximal possible chromatic number of an n-uniform hypergraph with N edges. Also $f(0) = 1$ by agreement. The function f non-strictly increases and satisfies

$$m(n, r) = \min\{N : f(N) > r\}.$$

Therefore $m(n, r) \sim Cr^n$ if and only if $f(N) \sim (N/C)^{1/n}$. The following lemmas were proved in [4].

Lemma 1. *For any $N > 0$ and any positive integer p we have*

$$f(N) \leqslant \max_{a_1+a_2+\cdots+a_p \leqslant N/p^{n-1}} f(a_1) + f(a_2) + \cdots + f(a_p).$$

Lemma 2. *Denote $c_n = \lceil (1 - 2^{1/n-1})^{-n} \rceil$. For any $M > 0$ the inequality*

$$f(N) \leqslant N^{1/n} \cdot \max_{M \leqslant a < c_n M} f(a) \cdot a^{-1/n}$$

holds for all $N \geqslant M$.

It is known that $f(0) = 1$, $f(1) = \ldots = f(6) = 2$, $f(7) = \ldots = f(26) = 3$ (see [1]). Lemmas 1, 2 and computer calculations were used to get

$$L_3 > 0.324.$$

The contribution of the paper is the following theorem, which is proved by refining Pluhár approach via inducibility arguments.

Theorem 1.

$$L_3 \geqslant \frac{4}{e^2} > 0.54.$$

Structure of the paper. In Section 2 we show how to apply inducibility to the chain argument and prove Theorem 1. In Section 3 we find the constant in Akolzin–Shabanov theorem for $n = 3$ and show that even if we apply Theorem 2 to the corresponding part of the proof, the constant will be still worse than in Theorem 1.

§2. Inducibility tool

Theorem 2. *Suppose $H = (V, E)$ is a hypergraph. Then it has at most*

$$|E| \left(\frac{|E| - 1}{r - 1} \right)^{r-1}$$

r-chains.

We need a notion of inducibility. Denote by $I(G, H)$ the number of induced subgraphs of G, isomorphic to H. Let P_r be a graph with r vertices and $r - 1$ edges which form a simple path. The following basic bound was proved by Pippenger and Golumbic.

Lemma 3 (Pippenger–Golumbic [8])**.** *Let G be a graph on N vertices. Then*

$$I(G, P_r) \leqslant \frac{N}{2} \left(\frac{N - 1}{r - 1} \right)^{r-1}.$$

It turns out that the bound is close to optimal. The following example is about $e^2/2$ times worse than the bound in Lemma 3 (we assume that r is fixed and n tends to infinity).

Example 1. *We construct the sequence of graphs G_k inductively. Let G_1 be a copy of C_{r+1}. Define an auxiliary graph $F_k = (V_k, E_k)$ (which is the $(r+1)^{k-1}$-blow-up of C_{r+1}):*

$$V_k := W_k^1 \sqcup W_k^2 \sqcup \cdots \sqcup W_k^{r+1}$$

with $|W_k^i| = (r+1)^{k-1}$ for all i; edges connect all the pairs of vertices from parts with adjacent indices (i is adjacent to $i+1$ modulo $r+1$, in particular $r+1$ is adjacent to 1). Then, G_k is obtained from F_k by drawing the graph G_{k-1} on each vertex set W_k^i.

Now consider the graph G_k on $N = (r+1)^k$ vertices. Note that

$$\begin{aligned} I(G_k, P_r) &= I(F_k, P_r) + (r+1)I(G_{k-1}, P_r) \\ &= (r+1)\left(\frac{N}{r+1}\right)^r + (r+1)I(G_{k-1}, P_r). \end{aligned}$$

Proof of Lemma 3. Let $X(q, l)$ denote the largest possible number of ways of sequentially choosing q objects $w_0, w_1, \dots, w_{q-1}$ from among l objects, subject to rules whereby the set of objects that are eligible to be chosen as w_i depends only on the previous choices $w_0, w_1, \dots, w_{i-1}$, and whereby no object that is eligible to be chosen as w_i will be eligible to be chosen as w_j for any $i+1 \leqslant j \leqslant q-1$. Also, define $X(0, l) = 1$. If $q > 0$, let m denote the number of objects eligible to be chosen as w_0. For any choice of w_0, the remaining $q-1$ objects can be chosen in at most $X(q-1, l-m)$ ways. Thus

$$X(q, l) \leqslant \max_{1 \leqslant m \leqslant l} mX(q-1, l-m).$$

From these relations, we obtain

$$X(q, l) \leqslant \left(\frac{l}{q}\right)^q \tag{1}$$

by induction on q: the base $q = 1$ is obvious. To prove the step it is enough to maximize the right-hand side of

$$X(q, l) \leqslant \max_{1 \leqslant m \leqslant l} m \left(\frac{l-m}{q-1}\right)^{q-1}.$$

Taking the derivative with respect to m, we get the maximum at $m = l/q$, and we are done.

Now we are ready to prove the initial statement. Fix the first vertex v_0. The number of ways to continue an induced r-path is at most $X(r-1, N-1)$. There are N ways to choose the first vertex and every copy of induced P_r is counted twice. Substitution of (1) finishes the proof. □

Proof of Theorem 2. Consider an auxiliary graph $G=(E,F)$ with vertex set being equal to the edge set of H and edges connecting pairs of vertices which intersect (as hyperedges) on exactly one vertex.

Note that every r-chain forms induced P_r in G (note that the reverse consequence is wrong, because a non-edge in G can correspond to the pair of hyperedges with large intersection, which is impossible in r-chain). Every copy of P_r is formed by at most two different r-chains, so the number of r-chains is at most $2I(G, P_r)$. Hence, Lemma 3 finishes the proof. □

Proof of Theorem 1. Let us try to color H via Pluhár's greedy algorithm. Recall that the probability of an r-chain to be ordered is

$$\frac{[(n-1)!]^2[(n-2)!]^{r-2}}{((n-1)r+1)!} = \frac{4}{(2r+1)!}.$$

Using Theorem 2 we get that if

$$\frac{|E|(|E|-1)^{r-1}}{(r-1)^{r-1}} \frac{4}{(2r+1)!} < 1,$$

then hypergraph is r-colorable. Summing up,

$$L_3 \geqslant \lim_{r\to\infty} \sqrt[r]{\frac{(2r+1)!(r-1)^{r-1}}{4}} \frac{1}{r^3} = \frac{4}{e^2}. \qquad \square$$

§3. Analysis of the Akolzin–Shabanov proof

We rewrite the proof from [2] with optimization in the case $n=3$.

First, for every vertex v introduce the weight $w(v)$ as randomly (accordingly to the uniform distribution and independently) chosen number from $[0,1]$. Fix parameters $p \in [0,1]$, $a < r$. An edge e is called *bad* if

$$\max_{v\in e} w(v) - \min_{v\in e} w(v) \leqslant \frac{1-p}{a};$$

otherwise it is called *good*.

The coloring algorithm is the following. First we color a (random) subhypergraph, consisting of all good edges, in a colors via Pluhár approach; then we color (or recolor) some vertices from bad edges in unused $r-a$ colors. If Pluhár approach succeeds (i.e. there are no ordered

a-chains) and we have at most $(n-1)(r-a)$ bad edges, then the algorithm return a proper r-coloring. Let us evaluate the probability of success.

Lemma 4 (Akolzin–Shabanov [2]). *Let e be an edge, then*

$$P\Big[e \text{ is bad}\Big]=\Big(\frac{1-p}{a}\Big)^{n-1}\Big(\frac{1-p}{a}+n\Big(1-\frac{1-p}{a}\Big)\Big)\leqslant n\Big(\frac{1-p}{a}\Big)^{n-1}=3\Big(\frac{1-p}{a}\Big)^{2}.$$

Let $C(A_1,\dots,A_a)$ denote the event that all the edges A_j are good and $(A_1,\dots,A_a)$ is an ordered a-chain.

Lemma 5 (Akolzin–Shabanov [2]).

$$P\left[C(A_1,\dots,A_a)\right]\leqslant a^{-a(n-2)}\frac{p^{a-1}}{(a-1)!}=a^{-a}\frac{p^{a-1}}{(a-1)!}.$$

By Theorem 2 we have at most $(|E|/(a-1))^{a-1}$ a-chains. Define $c=|E|/r^3$; we need

$$\Big(\frac{|E|}{a-1}\Big)^{a-1}a^{-a}\frac{p^{a-1}}{(a-1)!}=\Big((1+o(1))\frac{|E|pe}{a^3}\Big)^{a}=\Big((c+o(1))\frac{r^3pe}{a^3}\Big)^{a}<1.$$

Also we need at most $(n-1)(r-a)=2(r-a)$ bad edges:

$$P[X>2(r-a)]\leqslant\frac{1}{2(r-a)}\frac{3(1-p)^2|E|}{a^2}<1.$$

Define $x=r/a$. Then we need $cx^3pe<1$ and

$$\frac{3c(1-p)x^3}{2(x-1)}<1.$$

Computer simulations give that for $p=0.741$ and $x=1.05$ the algorithm with $c=0.42$ returns a proper coloring with positive probability, which implies $L_3>0.42$.

If we simply follow the initial proof, the required inequalities are

$$cx^3pe^2<1 \quad\text{and}\quad \frac{3c(1-p)x^3}{2(x-1)}<1.$$

So pure Akolzin–Shabanov approach gives $L_3>0.205$. Both constants are worse than in Theorem 1.

§4. Open problems

• First, recall that the Erdős conjecture is still open in the case $n = 3$.
• Also it is natural to ask if $m(n, r)$ is regular on the first variable, i.e.

$$\lim_{n\to\infty} \frac{m(n+1, r)}{m(n, r)} = r?$$

• In the proof of Theorem 2 we consider an auxiliary graph G. The problem is to describe the set of graphs, which may be achieved from an r-chromatic n-uniform hypergraph. Also it may be reasonable to evaluate the minimal number of vertices $N(r)$ in a graph G, which has an ordered induced r-path in every linear order of $V(G)$.

Acknowledgements. I am grateful to Alexander Sidorenko for an introduction in inducibility theory and to Fedor Petrov, who noted some deficiencies in the paper. Also Alexander Sidorenko pointed to an inattention, relating to the use of Turán numbers. Finally, I am grateful to the anonymous referee, who made a lot of helpful remarks.

References

1. I. A. Akolzin, *On* 3*-homogeneous hypergraphs colorings in* 3 *colors.* — Itogi Nauki i Tekhniki. Seriya Sovrem. Matem. i ee Prilozhen. Tenaticheskie Obzory, **150** (2018), 26–39.
2. I. Akolzin, D. Shabanov, *Colorings of hypergraphs with large number of colors.* — Discrete Mathematics **339**, No. 12 (2016), 3020–3031.
3. N. Alon, *Hypergraphs with high chromatic number.* — Graphs and Combinatorics **1**, No. 1 (1985), 387–389.
4. D. Cherkashin, F. Petrov, *Regular behavior of the maximal hypergraph chromatic number.* `arXiv preprint arXiv:1808.01482`, 2018.
5. D. Cherkashin, J. Kozik, *A note on random greedy coloring of uniform hypergraphs.* — Random Structures & Algorithms **47**, No. 3 (2015), 407–413.
6. P. Erdős, *Some old and new problems in various branches of combinatories* in: Proceedings of the Tenth Southeastern Conference on Combinatorics, Graph Theory and Computing. Congressus Numerantium XXIII, pp. 19–37. Winnipeg: Utilitas Mathematica, 1979.
7. P. Erdős, A. Hajnal, *On a property of families of sets.* — Acta Mathematica Hungarica **12**, No. 1–2 (1961), 87–123.
8. N. Pippenger, M. Ch. Golumbic, *The inducibility of graphs.* — J. Combinatorial Theory, Series B **19**, No. 3 (1975), 189–203.
9. A. Pluhár, *Greedy colorings of uniform hypergraphs.* — Random Structures & Algorithms **35**, No. 2 (2009), 216–221.

10. A. M. Raigorodskii, D. A. Shabanov, *The Erdős–Hajnal problem of hypergraph colouring, its generalizations, and related problems.* — Russian Mathematical Surveys **66**, No. 5 (2011), 933–1002.
11. A. Sidorenko, *What we know and what we do not know about Turán numbers.* — Graphs and Combinatorics **11**, No. 2 (1995), 179–199.

Chebyshev Laboratory,
St.Petersburg State University;
Moscow Institute of Physics and Technology,
Moscow region, 141700, Russia;
National Research University Higher School of Economics
St. Petersburg, Russia
E-mail: matelk@mail.ru

Поступило 14 ноября 2019 г.