COLLOQUIA MATHEMATICA SOCIETATIS JÁNOS BOLYAI
60. SETS, GRAPHS AND NUMBERS, BUDAPEST (HUNGARY), 1991

How Many Edges should be Deleted to Make a Triangle-Free Graph Bipartite?

P. ERDŐS, E. GYŐRI and M. SIMONOVITS[1]

ABSTRACT

In this paper we shall prove that if L is an arbitrary 3-chromatic graph and G_n is a simple graph on n vertices not containing L, and having at least

$$\frac{n^2}{5} - o(n^2)$$

edges, then it can be made bipartite by throwing away at most

$$\frac{n^2}{25} - o(n^2)$$

edges. This was known for $L = K_3$.

Let us call a graph pentagonlike if we can colour its 5 classes so that the vertices coloured by i are joined only to vertices coloured by $i \pm 1 \pmod 5$.

In addition to the above assertions, we shall prove that under the above conditions, there is a "pentagonlike graph" H_n with $e(H_n) = e(G_n)$ for which we have to delete more edges than in case of G_n to make it bipartite. We shall also prove a related stability theorem, according to which, if $D(G_n)$ denotes the minimum number of edges to be deleted to make G_n bipartite, then either $D(G_n) \le D(H_n) - cn^2$ (i.e. G_n is significantly better than H_n — though they both may be far from any bipartite graph, — or the structure of G_n is very near to that of a pentagonlike graph.

[1] This research was supported by Hungarian National Research Grant N1909.

Notation. Below, we shall regard simple graphs, i.e. graphs without loops and multiple edges. Given a graph G, $e(G)$, $v(G)$ and $\chi(G)$ will denote the number of edges, the number of vertices and the chromatic number of G respectively. Besides, in case if a graph is denoted by a capital letter with a subscript, (like G_n, H_n, S_k, ...) the subscript will always denote the number of vertices. K_p will denote the complete p-graph, C_p the p-cycle.

For a graph G, $N(x)$ and $d(x)$ will denote the neighbourhood of a vertex x and its degree. Sometimes we shall have two graphs at the same time: G and a $Z \subset G$. In such cases we will occasionally use subscripts to indicate, which graph are we speaking about: $N_Z(x)$ and $d_Z(x)$ will denote the neighbourhood of x and its degree in Z.

Given a graph L, with $v = v(L)$ vertices, $a_1, \dots, a_v$, and the integers $n_1, \dots, n_v \geq 0$, then the graph $L[n_1, \dots, n_v]$ is defined as follows: the ith vertex of L, a_i is replaced by n_i new, independent vertices, forming a set A_i , $(i = 1, \dots, v)$ and we join in $L[n_1, \dots, n_v]$ a vertex $x \in A_i$ to a vertex $y \in A_j$ iff $a_i a_j$ is an edge in L.

As a special case of this, take $H = C_5$, i.e., denote by $H[n_1, \dots, n_5]$ the pentagonlike graph with n_i vertices in its ith class. We shall write $H[n_1, \dots, n_5] \sim H[n_1', \dots, n_5']$ if $n_1 + \ldots + n_5 = n_1' + \ldots + n_5' = n$ and $n_i - n_i' = o(n)$. [2]

We shall use a "distance" to describe the structure of a graph G: $D(G)$ is the minimum number of edges representing all the odd cycles of G, or, in other words, the minimum number of edges to be deleted to turn G into a bipartite graph. Similarly $D_p(G)$ is the minimum number of edges to be deleted to turn G into a p-colorable graph.

$K(n_1, \dots, n_p)$ will denote the complete p-partite graph with n_i vertices in its ith class.

Given a graph L, $\text{ext}(n, L)$ denotes the maximum number of edges a graph G_n of order n can have without containing L as a (not necessarily induced) subgraph.

---

[2] To be precise, this makes sense only if $n \to \infty$, and mostly we shall regard large but fixed n's, where we should rather speak of ε and δ...

1. Introduction

To start with, below we occasionally disregard the "integer parts", or other similar "divisibility nuisances" in our formulas.

The results of this paper can basically be motivated in three different ways: by three different (though not too distant) problems.

(A) It is well known [1] that for every graph G

$$D(G) \leq \frac{1}{2}e(G). \tag{1}$$

In general this is sharp (apart from some error terms). Thus e.g. (1) is sharp for the complete graph, or (more generally), for random graphs with some fixed edge-probability $p > 0$. We are interested in the question, under which natural extra conditions can (1) essentially be improved.

(B) If a graph is bipartite, it cannot contain triangles. It is known, (see e.g. [3], [4], [10], [6]) that graphs having almost $\mathrm{ext}(n, K_3)$ edges and not containing K_3's can be changed into a bipartite graph by deleting relatively few edges. The question to be considered here is:

Given a graph G_n of n vertices and $E = e(G_n)$ edges, at least how many edges are needed to be deleted to make G_n bipartite?

There are many conjectures (mostly due to Erdős) asserting that in some similar cases the following graph is the "extremal" one: we put $n/5$ vertices into 5 classes $U_1, \dots, U_5$ and join each vertex of U_i to each vertex of U_{i+1} $(i = 1, \dots, 5,\ U_6 = U_1.)$ One of the most intriguing open conjectures in this field is

Conjecture 1. *(Erdős) Prove that if $K_3 \not\subseteq G_n$ then $D(G_n) \leq \frac{1}{25}n^2$, i.e. one can delete $\leq \frac{1}{25}n^2$ edges of G_n to end up with a bipartite graph.*

This problem is motivated by the fact that $H_n := C_5[n/5, \dots, n/5]$ contains no K_3 and

$$D(H_n) = \frac{1}{25}n^2.$$

Being interesting on its own, this problem would also have some applications related to the remainder terms in the Erdős-Simonovits Theorem [8], as well. (See also [10,2].)

This conjecture is proven for

$$e(G_n) \geq \frac{n^2}{5}$$

(see below) and the following (other) weakening is also known, [7].

Theorem. *If $K_3 \not\subseteq G_n$ then*

$$D(G_n) \leq \frac{n^2}{18+\delta}.$$

for some (calculable constant) $\delta > 0$.

In fact, Erdős, Faudree, Pach, and Spencer [7] proved that

Theorem. *For every triangle-free graph G with n vertices and m edges*

$$D(G_n) \leq \max\left\{\frac{1}{2}m - \frac{2m(2m^2-n^3)}{n^2(n^2-2m)}\ ,\ m - \frac{4m^2}{n^2}\right\} \tag{2}$$

Since the second term of (2) decreases in $\left[\frac{1}{8}n^2, \frac{1}{2}n^2\right]$, and its value is exactly $\frac{1}{25}n^2$ for $m = \frac{1}{5}n^2$, therefore (2) implies that if

$$e(G_n) > \frac{n^2}{5}, \tag{3}$$

and $K_3 \not\subseteq G_n$, then $D(G_n) \leq \frac{1}{25}n^2$. Again, by (1), it is trivial, that if

$$e(G_n) \leq \frac{2n^2}{25} \tag{4}$$

then $D(G_n) \leq \frac{1}{25}n^2$. However, the general conjecture is still open: the middle interval

$$\frac{2n^2}{25} < e(G_n) < \frac{n^2}{5}$$

is unsettled.

(C) Our third starting point was the following problem. In Conjecture 1 we try to estimate $D(G_n)$ in terms of n, the number of vertices. It is much more natural to try to find estimates which use functions of $e(G)$ or functions of $e(G)$ and $v(G)$, to bound $D(G)$. (Actually, one of the main goals of [7] is to give estimates on $D(G)$ in terms of $e(G)$ or in terms of both $v(G)$ and $e(G)$.) Erdős first thought that

Conjecture 2. *If $K_3 \not\subseteq G_n$, then $D(G_n) \le \frac{1}{5}e(G_n)$.*

This would be sharp e.g. for the pentagonlike graph .... However, later Erdős has disproved this, by proving in [2] that there exists an infinite sequence of graphs H_k such that $K_3 \not\subseteq H_k$ and

$$D(H_k) \ge \left(\frac{1}{2} - o(1)\right) e(H_k) \qquad \text{as} \qquad k \to \infty.$$

One natural way to try to save Conjecture 2 would be to add the extra condition that G_n has many edges: $e(G_n) \ge cn^2$ for some constant $c > 0$ while $n \to \infty$.

This natural conjecture is unfortunately also false. We shall disprove this weaker conjecture using the "random construction" of Erdős, mentioned above. We shall prove

Theorem 1. *For every $\varepsilon > 0$ there exists a constant $c = c_\varepsilon > 0$ such that for infinitely many n, there exists a G_n with $e = e(G_n)$ and $K_3 \not\subseteq G_n$, with $e(G_n) > c_\varepsilon n^2$, for which*

$$D(G_n) > \left(\frac{1}{2} - \varepsilon\right) e.$$

This is sharp in the obvious sense (guaranteed by (1)).

Our main tool to prove Theorem 1 will be to regard graphs of form $L_k[n_1 \dots, n_k]$ and show that to make such a graph bipartite by deleting the minimum number of edges we may always find an edge-deletion where the fact if we delete an edge or not depends only on the classes of its endpoints.

Definition 1. *Canonical edge deletion.* Given a graph H_k and the integers $n_1, \dots, n_k$, put $G_n = H_k[n_1, \dots, n_k]$. We shall call an edge-deletion *canonical* if for each pair of vertex-groups of G_n either we delete all the connecting edges or none of them. (In other words, the resulting G' has the form $G' = H'[n_1, \dots, n_k]$, for some subgraph $H' \subseteq H$. This implies, among others, that – to determine $D_p(G_n)$ we have to check only a bounded number of cases, independently of n.)

The main goal of this paper is to show that for any fixed 3-chromatic L, for every graph G_n not containing this L and having sufficiently many edges, it is easier to make G_n bipartite than to make bipartite a corresponding pentagonlike graph with at least the same number of edges[3] .

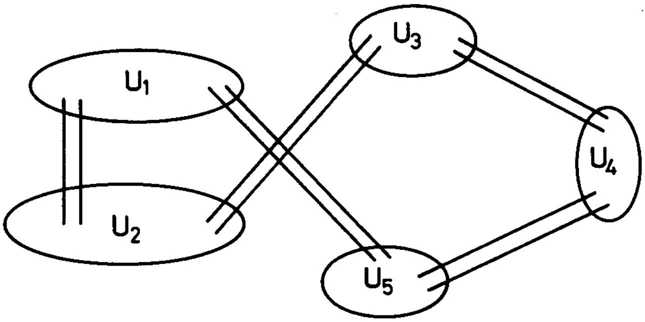

Figure 1.

For given B we may regard the pentagonlike graph $H[n_1, \ldots, n_5]$ with $n_3 = n_4 = n_5 = \sqrt{B}$. If all the other classes are bigger, then (by Theorem 7) we need to delete at least B edges from this graph to make it bipartite.

Theorem 2. *If $K_3 \not\subseteq G_n$ and*

$$e(G_n) \geq \frac{n^2}{5},$$

then there is a pentagonlike graph H_n^* with at least the same number of edges: $e(G_n) \leq e(H_n^*)$, for which

$$D(G_n) \leq D(H_n^*). \tag{5}$$

Greedy Algorithm. *(a) Pick an edge (x, y) with maximum $d(x) + d(y)$. Let $A_1 = N(x)$, $A_2 = N(y)$, $C = V - A_1 - A_2$. Put the vertices $v_1 \ldots$ of C successively into C_1 and C_2, so that the number of monochromatic edges (i.e. edges incident to v_i from C_i to $A_i \cup C_i$)) be the minimum possible. The number of monochromatic edges — obtained in this procedure — will be denoted by $D'(G_n)$. (This may depend on the choices of xy and the order of the elements of C.)*

By definition, $D'(G) \geq D(G)$. Instead of proving (5) we shall prove the stronger

$$D'(G_n) \leq D(H_n^*). \tag{6}$$

[3] Below Theorem 2 is an "exact statement", i.e. holds without error-terms, but the other results are "asymptotic structure"-type results.

Theorem 3. *If $K_3 \not\subseteq G_n$ and*

$$e(G_n) > \frac{n^2}{5},$$

and applying the above Greedy Algorithm to G we get a class $C = C_1 \cup C_2$ with $c = |C|$ vertices, representing cd monochromatic edges of G_n:

$$D'(G_n) := cd,$$

then there exists an $H_n \sim H[\sqrt{cd}, \sqrt{cd}, \sqrt{cd}, u, u]$ s.t.

$$e(G_n) \le e(H_n).$$

Theorem 3 implies Theorem 2. Indeed, a standard but somewhat tedious argument shows that for given n, as $e = e(H_n)$ increases, $\min\limits_{e(H_n) \ge e} D(H_n)$ decreases.

The next theorem is a stability theorem connected to Theorems 2–3. It asserts that either H_n is significantly more difficult to make bipartite than G_n or they are similar to each other in the sense that even G_n can be turned into a pentagonlike graph quite easily.

Theorem 4. *For every $\varepsilon > 0$ there exists a $\delta > 0$ such that if — in Theorem 3 —*

$$e(H_n) - \delta n^2 \le e(G_n),$$

then G_n can be turned into a pentagonlike graph by deleting at most εn^2 edges.

Theorem 5. *For every $\varepsilon > 0$ there exists a $\delta > 0$ such that if G_n contains at most δn^3 K_3's and*

$$e(G_n) > \frac{n^2}{5} - \delta n^2,$$

then for some $H_n \sim H[m, m, \mu, \mu, \mu]$, $(\mu \le m)$

$$e(G_n) \le e(H_n).$$

If further,

$$e(H_n) - \delta n^2 \le e(G_n),$$

then G_n can be turned into a pentagonlike graph by deleting at most εn^2 edges.

One could ask: how do our theorems change if we replace K_3 by an arbitrary 3-chromatic excluded subgraph L? The answer is, that the theorems remain valid.

Theorem 6. *Let L be an arbitrary fixed 3-chromatic graph. For every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $L \not\subseteq G_n$ and*

$$e(G_n) > \frac{n^2}{5} - \delta n^2,$$

and if applying the above greedy algorithm to G we get a class C with c vertices, representing cd monochromatic edges of G_n:

$$D'(G_n) := cd,$$

then for some $H_n \sim H[\sqrt{cd}, \sqrt{cd}, \sqrt{cd}, u, u]$,

$$e(G_n) \le e(H_n).$$

If, additionally,

$$e(G_n) \le e(H_n),$$

and

$$e(H_n) - \delta n^2 \le e(G_n),$$

then G_n can be turned into a pentagonlike graph by deleting at most εn^2 edges.

One could ask if one can — under the conditions of Theorem 6 — guarantee that G_n is contained by a pentagonlike graph, i.e. we do not have to delete edges to get a subgraph of some $H_n(n_1, \dots, n_5)$. However this is not the case: if Q is the Petersen graph, then an appropriate $Q[n_1, \dots, n_{10}]$ will provide a counterexample to this "conjecture", (see Figure 2).

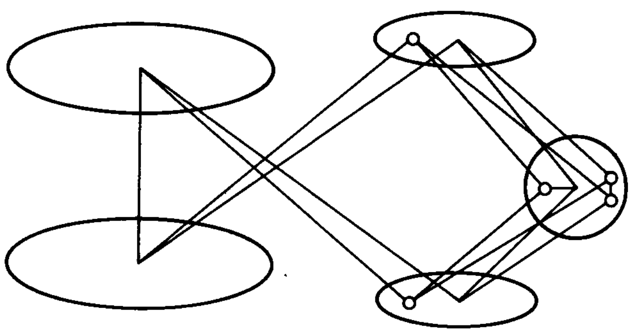

Figure 2.

Theorem 6*. *Let L be an arbitrary fixed 3-chromatic graph, $v := v(L)$. For every $\varepsilon > 0$ there exists a $\delta > 0$ such that if G_n contains at most δn^v copies of L and*

$$e(G_n) > \frac{n^2}{5} - \delta n^2,$$

and if applying the above Greedy Algorithm to G we get a class C with c vertices, representing cd edges of G_n:

$$D'(G_n) := cd,$$

then for some appropriate $H_n \sim H[m, m, \mu, \mu, \mu]$

$$e(G_n) \leq e(H_n).$$

If, additionally,

$$e(H_n) - \delta n^2 \leq e(G_n),$$

then G_n can be turned into a pentagonlike graph by deleting at most εn^2 edges.

2. Lemmas on Pentagonlike graphs

Here our first aim is to prove that if for a given p we wish to make a graph $G_n = L_k[n_1, \dots, n_k]$ into a p-chromatic graph by deleting the minimum number of edges needed, then we can always delete these edges canonically.

The meaning of this is the following.

Assume that $k \mid n$. It is trivial that if $G_n = L_k[n/k, \dots, n/k]$ then

$$\frac{D_p(G_n)}{n^2} \leq \frac{D_p(L_k)}{k^2}. \tag{7}$$

Indeed, we can delete $D_p(L_k)$ edges from L_k to get a p-chromatic graph, and therefore we can delete (canonically)

$$D_p(L_k)\left(\frac{n}{k}\right)^2$$

edges to get a p-chromatic graph from G_n. This proves (7). The meaning of the next theorem is that this is roughly the best one can do for these graphs.

Theorem 7. *Let*

$$G_n = H_k[n_1, \ldots, n_k].$$

Then there is a canonical "edge-deletion" achieving the minimum $D_p(G_n)$.

Corollary 1. *Assume that for some p, H_k is a graph for which $D_p(H_k) > \alpha k^2$. Let (G_n) be a graph sequence defined by $G_n = H_k[n_1, \ldots, n_k]$ for some $n_i = \frac{n}{k} + o(n)$. Then $D_p(G_n) > \alpha n^2 - o(n^2)$ as $n \to \infty$. Moreover for the case when $k \mid n$, we have*

$$\frac{D_p(G_n)}{n^2} \geq \frac{D_p(H_k)}{k^2}. \tag{8}$$

Proof of Theorem 7. Assume we have an H_k and construct a

$$G_n = H_k[n_1, \ldots, n_k]$$

from it. Then someone deletes $m = D_p(G_n)$ edges from G_n, producing a p-chromatic graph Z_n. Transforming Z_n step by step we will show that there is an equally optimal Z_n^* which is already canonical. This will imply Theorem 7.

Let $U_1, \ldots, U_k$ be the classes of G_n. We shall call two vertices symmetric if they have the same neighbourhood.

Let $x, y \in U_j$. We apply symmetrization to this graph Z iteratively: if $d_Z(x) \geq d_Z(y)$, then we delete all the edges of Z_n adjacent to y and join y to all the Z-neighbours of x. During this the number of edges does not decrease. Further, the chromatic number cannot increase either. Observe that if x and y are symmetric and we symmetrize u to v then this uv-symmetrization does not ruin the symmetry of x and y. So let us iterate this step. We can do this so that the symmetrization will increase the number of pairs of symmetric vertices in each step. This ensures that the procedure will sooner or later terminate. Finally we end up with a graph in which all the pairs x, y belonging to the same U_i have the same neighbourhood. ∎

Corollary 2. *If $H[n_1, \ldots, n_5]$ is a pentagonlike graph, then*

$$D(H[n_1, \ldots, n_5]) = \min_i n_i n_{i+1}.$$

In other words, the most efficient way to turn a pentagonlike graph into a bipartite one is to choose two connected classes and delete all the edges between them.

Remarks. 1. It can easily happen that we may turn a graph $G_n = L_k[n_1, \dots, n_k]$ into a bipartite graph so that we have deleted between any pairs of groups only half of the edges. E.g. in a pentagonlike graph we can split each group U_i into two almost equal groups U_i^1 and U_i^2 and delete half of the edges to get the 10-cycle-like graph with classes

$$U_1^1, \quad U_2^1, \quad \dots \quad ,U_5^1, \quad U_1^2, \quad U_2^2, \quad \dots \quad ,U_5^2, \quad U_1^1.$$

This is not an optimal edge-deletion.

2. One could ask if the maximum can be attained only by canonical subgraphs. The answer is that there are many non-canonical maximum-size bipartite graphs of a graph, e. g. for $H = C_5$ $H[a,b,c,d,b]$ can be turned into a bipartite graph by splitting U_1 into two classes U' and U" and deleting all the edges between U' and U_2 and all the edges between U" and U_5. If $a, b \le c, d$ then all these splittings are optimal.

3. The method of symmetrization was first used by Zykov [14] to prove Turán's theorem. Since that the method was successfully used in various cases, e. g. Simonovits used it in a very similar setting.

4. One final remark should be made. Given a $C_5[n_1, \dots, n_5]$ with a given number of vertices, $n_1 + \dots + n_5 = n$, one is curious, how to choose the n_i's to get the maximum number of edges under the condition that $n_i n_{i+1} \ge B$. The extremal distribution in such problems is sometimes asymptotically unique, sometimes not. Here e. g. one optimal graph is often obtained by taking $C_5[a,a,c,c,a]$ with $3a + 2c = n$, $a \le c$. However, in this case we have infinitely many equally good graphs. Generally, moving from $C_5[n_1, n_2, n_3, n_4, n_5]$ to $C_5[n_1, n_2+1, n_3, n_4-1, n_5]$ the number of edges changes only by $(n_1 - n_5)$. Therefore it remains constant for $n_1 = n_5$.

Theorem 7 implies Theorem 1 but what is the dependence of the constants in Theorem 1?

Claim 1. *For every (small) $c > 0$, there exists a $c' > 0$ such that if $K_3 \not\subseteq G_n$ and $e(G_n) \ge cn^2$ then*

$$D(G_n) < \left(\frac{1}{2} - c'\right) e(G_n).$$

Proof. We may assume that $e(G_n) = cn^2$. Then we may also assume that each vertex has degree $\ge cn$. Pick a vertex x and consider its neighbourhood $A = N(x)$. It is a set of independent vertices and so $e(N(x), V - N(x)) \ge$

c^2n^2. On the average a set $B \subseteq V - N(x)$ of size $\frac{1}{2}cn$ contains $\frac{1}{4}c^3n^2 + O(c^4n^2)$ edges and on the average it is joined to $N(x)$ by $> c^3n^2/2$ edges. Hence, by averaging, there is a set $B \subseteq V - N(x)$ of $\frac{1}{2}cn$ vertices such that the magnitude of

$$e(B, N(x)) - e(G[B])$$

is $> c^3n^2/5$. So we have found two sets A, B so that the number of edges joining different classes is greater by at least $c^3n^2/5$ than the number of edges joining vertices of the same class. Let us call this quantity the "surplus". Now, putting the vertices outside of $A \cup B$ one by one either into A or into B, we can always put a new vertex x into the class to which it is joined by the fewer (or equal) number of edges. Hence the surplus will not decrease and we shall get a partition $A \cup B = V(G_n)$ of the whole vertex set. The above argument shows that $c' > c^3/5$. ■

3. Proof of Theorem 3

To make the proof easier to follow we fix some notations in advance.

We have a graph G_n and pick an edge x^*y^* for which

$$d(x^*) + d(y^*) \tag{MAX}$$

is the maximum possible. By Lemma 2,

$$d(x^*) + d(y^*) > \frac{4}{5}n. \tag{9}$$

Assume that $d(x^*) = a_2$, $d(y^*) = a_1$, $A_2 = N(x^*)$, $A_1 = N(y^*)$. (Thus $a_i = |A_i|$.) Define $C = V - A_1 - A_2$. Apply the greedy algorithm to C: put each vertex $u \in C$ either into C_1 or into C_2 depending on if

$$|N(u) \cap A_2| \geq |N(u) \cap A_1|$$

or not. Then $V(G_n) = A_1 \cup A_2 \cup C_1 \cup C_2$.

Put $c_i = |C_i|$ and $n_i = |A_i| + |C_i|$ and put

$$D_i = \max_{u \in C_i} |N(u) \cap A_i|$$

and $D = \max\{D_1, D_2\}$.

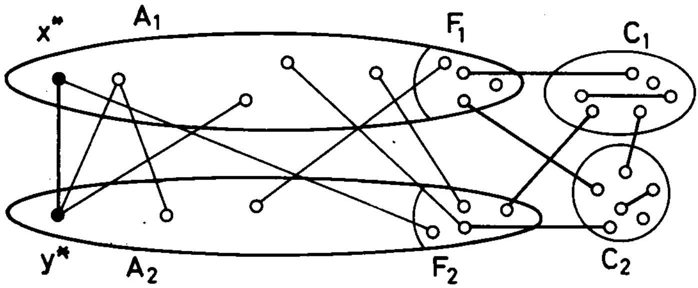

Figure 3.

Definition of d. We shall call the edges joining two vertices of $A_1 \cup C_1$ or two vertices of $A_2 \cup C_2$ *monochromatic* edges. (Sometimes, referring to most of the figures here, these edges will be called *horizontal*, the edges joining $A_1 \cup C_1$ to $A_2 \cup C_2$ will be called *vertical*.

Now, d is defined by that cd is the number of monochromatic edges in G, These edges will determine $t \geq d$ independent horizontal edges by Lemma 1 below, some of which will join C_1 to A_1. Their number will be denoted by γ_1, their endvertices in A_1 form a subset F_1. γ_2 and F_2 are similarly defined, further, φ_i is the number of horizontal edges in C_i.

One crucial point in the proof is that for any edge $e = xy$ — since $K_3 \not\subseteq G_n$, — each vertex can be joined to at most one of x and y. Therefore, if we have a big set of vertices not joined to this edge, then we have a good upper bound on $d(x) + d(y)$. We shall try to cover the "critical" part of the graph by independent edges so that we could apply this bound to them. Then we shall be able to get an upper bound on the degree-sum of G_n.

The next lemma seems to be so harmless, still, this is one important point where we use that $K_3 \not\subseteq G_n$.

Lemma 1. *Let G be a graph not containing any K_3. Let $Z \subseteq V(G)$ and $z := |Z|$. Assume that $d \leq z$ and $d(x) \leq D$ for $x \in V(G)$. If*

$$d(Z) = e(G) - e(G - Z) > \max\{D, z\} \cdot (d - 1),$$

then there exist d independent edges in G incident to Z.

(They may join C to $C - Z$ but also they may join pairs of vertices in Z.)

Proof. Induction on d.

If $d(x) \le d-1$ for $x \in Z$ then $(d-1)\max\{D, z\} < d(Z) \le z(d-1)$, a contradiction. So, we may assume that there is a vertex $x \in Z$ with $d(x) \ge d$. Then

$$d_{G-x}(Z-x) > \max\{D, z-1\}(d-2)$$

since $d(x) \le \max\{D, z-1\}$, and $G-x$ contains a matching M of $d-1$ edges incident to Z by the induction hypothesis. At most $d-1$ edges lead from x to M by triangle-freeness and so M can be extended with some edge xy to a matching of d edges incident to Z. ■

Later we shall also need a slightly modified, asymptotic version of this lemma.

Lemma 1*. *Let (G_n) be a graph sequence in which each edge is contained in at most $o(n)$ triangles. Let $Z \subseteq V(G_n)$ and $z := |Z|$. Assume that $d \le z$ and $d(x) \le D$ for $x \in V(G)$. If*

$$d(Z) = e(G_n) - e(G_n - Z) > d\max\{D, z\}$$

and $d \to \infty$, then there exist $d - o(d)$ independent edges in G_n adjacent to Z.

The proof is almost the same as that of the "exact" version. It is left to the reader.

Lemma 1.** *If in Lemma 1* we have almost equality: if G_n contains at most $d+o(d)$ independent edges adjacent to Z, then changing $o(zd)$ edges of G_n we can achieve that if $z \ge D$ then for an appropriately chosen B of size $d+o(d)$, all the edges incident to Z join Z to B (and consequently), all these edges belong to the (modified) G_n and if $D > z$ then for an appropriately chosen B of size $D+o(D)$ and $Z_0 \subset Z$ of size $d+o(d)$, all edges incident to Z join Z_0 to B and consequently all these edges belong to the modified G_n.*

This lemma is somewhat more complicated to prove, still it is a standard argument left to the reader.

Lemma 2. *(Folklore). Given a graph G with average degree $\delta = \dfrac{2e(G)}{v(G)}$, it contains an edge xy for which $d(x) + d(y) \ge 2\delta$.*

Proof. For any graph

$$\sum_{xy\in E(G)} (d(x)+d(y)) = \sum_{x\in V(G)} d(x)^2 \geq \frac{1}{n}\left(\sum_{x\in V(G)} d(x)\right)^2 = n\cdot\left(\frac{2e(G)}{n}\right)^2.$$

So the average of $d(x)+d(y)$ for the edges xy of G is at least $4e(G)/n = 2\delta$. ∎

Proof of Theorem 3.

(A) Given the graph G_n, we shall distinguish two cases:

— the maximum horizontal degree $D \leq c$

— $D > c$.

In some sense the first case is the more involved. So let us assume first that $D \leq c$. In this case we shall build a pentagonlike graph as follows:

Later we shall find some set of independent edges in G_n and some of them will be represented in A_1, their number will be t_1, some others in A_2: their number will be t_2. With these parameters we build a pentagonlike H_n by taking a complete bipartite graph with $a_i+\frac{1}{2}(c_i-t_i)$ vertices in its classes ($i=1,2$), which will be split later into U_1 and U_3, (U_2 and U_5). Next put $\mu=\frac{1}{2}(c_1+c_2+t_1+t_2)$ vertices into the residual class. This will be U_4. Fix μ from the $a_1+\frac{1}{2}(c_1-t_1)$ vertices and further μ from the $a_2+\frac{1}{2}(c_2-t_2)$ vertices. These will be U_3 and U_5. Delete the edges between U_3 and U_5 and join U_4 to both U_3 and U_5 completely. The resulting H_n has

$$a_1+\frac{1}{2}(c_1-t_1)+a_2+\frac{1}{2}(c_2-t_2)+\frac{1}{2}(c_1+c_2+t_1+t_2) = a_1+a_2+c = n$$

vertices.

We shall estimate the degree-sums both in G_n and H_n. Let us start with H_n.

Clearly,

$$\begin{aligned} 2e(H_n) =& 2(a_1+\frac{1}{2}(c_1-t_1))(a_2+\frac{1}{2}(c_2-t_2))+2\mu^2 \\ &\geq 2a_1a_2+(c_2-t_2)a_1+(c_1-t_1)a_2+\frac{1}{2}(c+t)^2. \end{aligned} \tag{10}$$

where $c=c_1+c_2,\ t=t_1+t_2$.

(B) To estimate the degree-sum in G_n we shall cover as many vertices of C as possible by a set of independent monochromatic edges. By Lemma 1, we shall have at least $t \geq d$ such edges and we shall estimate for these edges the degree-sum $d(x) + d(y)$ by $a_i + t + c$, (if x or both x and y are in C_i). For the remaining vertices we shall use the estimate a_1 or a_2, depending on the position of the vertex.

We choose maximum number $t \geq d$ of independent edges from the cd monochromatic ones so that the number of chosen edges from C to $A_1 \cup A_2$ be the maximum one. (In other words, the number of chosen edges inside C be the minimum possible.) Then we have some edges joining vertices of the same C_i. These form a matching M.

(B_1) For technical reasons (to be able to estimate the degrees of C not covered by M) we add some further edges to M joining C_i either to C_{3-i} if possible, or to A_{3-i} if we have already run out of the previous possibility. The extended matching will be denoted by M^*. It can happen that some vertices of C are not covered by this (extended) M^*.

These conditions on the choice of M^* together will be called the #-condition.

Now we have to check, how can we estimate $d(x) + d(y)$ for the various types of edges in M^*. We have 7 types.

$(*)_1$	$C_1 \longleftrightarrow C_1$	φ_1	$(*)_2$	$C_2 \longleftrightarrow C_2$	φ_2
$(**)_1$	$C_1 \longleftrightarrow A_1$	γ_1	$(**)_2$	$C_2 \longleftrightarrow A_2$	γ_2
$(***)_1$	$C_1 \longleftrightarrow A_2$	S_1	$(***)_2$	$C_2 \longleftrightarrow A_1$	S_2
		$(****)$	$C_1 \longleftrightarrow C_2$	R	

The set of endvertices of the edges of type $(**)_1$ and $(**)_2$ in A_1, C_1, A_2, C_2 will be denoted by F_1, F_1^*, F_2, F_2^*, respectively.

Let us start with type $(*)_1$ type $C_1 \longleftrightarrow C_1$.

Clearly, $d(x) + d(y) \leq a_2 + c + \gamma_1$ for this type, since no edge goes to such an edge from $A_1 - F_1$ (otherwise we could "improve" the selected edge-system with respect to #). Thus from $A_1 \cup A_2$ we have at most $(a_1 + a_2) - (a_1 - \gamma_1) = a_2 + \gamma_1$ edges. Further, at most c edges can join vertices of C to such an edge (because of the K_3-freeness). Thus, if we have φ_1 such edges, we get altogether degree-sum

$$(a_2 + c + \gamma_1)\varphi_1 \tag{11}$$

for these edges in M and similarly $(a_1 + c + \gamma_2)\varphi_2$ for the edges of type $(*)_2$.

For edges xy of type $(^{**})_1$ $(x \in F_1^* \subseteq C_1,\ y \in F_1 \subseteq A_1)$, we prove that $d(x) + d(y) \le a_2 + c + \gamma_1 + c_2$. Notice that either $N(x) \cap A_1 \subseteq F_1$ or $N(y) \cap C_1 \subseteq F_1^*$. Suppose not and, say, $x_1 \in N(y) \cap (C_1 - F_1^*)$ and $y_1 \in N(x) \cap (A_1 - F_1)$. Then taking $x_1 y$ and $x y_1$ instead of xy we get $\gamma_1 + 1$ independent edges of type $(**)_1$ and we have to delete at most one edge of type $(*)_1$ (if x_1 was covered by some edge of this type), a contradiction to the choice of M.

If $N(x) \cap A_1 \subseteq F_1$ then $N(x) \cup N(y) \subseteq A_2 \cup C \cup F_1$ and since no vertex is joined to both x and y, so

$$d(x) + d(y) \le a_2 + c + \gamma_1 .$$

If $N(y) \cap C_1 \subseteq F_1^*$ then $|N(x) \cap (A_1 \cup C_1)| \le D \le c$ (by definition of D), $N(y) \cap (A_1 \cup C_1) \subseteq F_1^*$ and since no vertex in $A_2 \cup C_2$ is joined to both x and y, we have

$$d(x) + d(y) \le c + \gamma_1 + a_2 + c_2 .$$

Similarly, for the edges xy of type $(**)_2$, we have

$$d(x) + d(y) \le a_1 + c + \gamma_2 + c_1 , \tag{12}$$

and it yields the upper estimate

$$\gamma_1(a_2 + c + \gamma_1 + c_2) + \gamma_2(a_1 + c + \gamma_2 + c_1)$$

for the degree sum in $F_1 \cup F_1^* \cup F_2 \cup F_2^*$.

The next types are $(***)_1$ type $C_1 \longleftrightarrow A_2$, $(***)_2$ type $C_2 \longleftrightarrow A_1$ or $(****)$ type $C_1 \longleftrightarrow C_2$. Now we can use the estimate

$$d(x) + d(y) \le a_1 + a_2$$

by (MAX).

Finally, we have to estimate the degrees for the vertices of C not covered by M^*. For these "free" vertices x we know — by the greediness — that no y outside of M^* can be joined to x, otherwise we could extend M^*. Further, only one endvertex of each of the edges of M^* can be joined to such an x:

$$d(x) \le |M^*| \le c. \tag{13}$$

We will prove that $c \le a_1$, $c \le a_2$, implying $d(x) \le a_1$ and $d(x) \le a_2$. Suppose that, say, $a_2 < c \le n/5$. Notice that $d(v) \le a_2$ for $v \in A_1$, by (MAX) and $|E(G[A_2 \cup C])| \le (a_2+c)^2/4$ by Turán's theorem. Thus

$$|E(G)| \le a_1 a_2 + \frac{(a_2+c)^2}{4} < (n-a_2)a_2 + \frac{n^2}{25} \le \frac{4n^2}{25} + \frac{n^2}{25} = \frac{n^2}{5}$$

a contradiction.

(C) Now, the degree-sum $\sum_{x \in V} d(x)$ can be estimated as follows. Let us count the contribution of the vertices of various kind to $\sum d(x_i)$. The vertices in $A_1 - F_1$ contribute a_2, the vertices of $A_2 - F_2$ contribute a_1, by (MAX). If there are R_1 edges in the matching M joining C_1 to C_2, their contribution can be estimated by $a_1 + a_2$, by (MAX).

Let S_1, S_2, R be the number of matching edges of type $C_1 \longleftrightarrow A_2$, $C_2 \longleftrightarrow A_1$, $C_1 \longleftrightarrow C_2$. Applying (11), (12) and $d(x) \le a_1$ for $x \in C_2 - V(M^*)$, $d(x) \le a_2$ for $x \in C_1 - V(M^*)$, we get

$$\begin{aligned}\sum_{x \in V} d(x) \le & (a_1 - \gamma_1 - S_2)a_2 + (a_2 - \gamma_2 - S_1)a_1 \\ & + S_1(a_1+a_2) + S_2(a_1+a_2) + R(a_1+a_2) \\ & + \varphi_1(a_2 + \gamma_1 + c) + \varphi_2(a_1 + \gamma_2 + c) \\ & + \gamma_1(a_2 + c + \gamma_1 + c_2) + \gamma_2(a_1 + c + \gamma_2 + c_1) \\ & + (c_1 - 2\varphi_1 - \gamma_1 - R - S_1)a_2 \\ & + (c_2 - 2\varphi_2 - \gamma_2 - R - S_2)a_1.\end{aligned} \tag{14}$$

We assumed that $D \le c$. Put $t_1 = \gamma_1 + \varphi_1$ and $t_2 = \gamma_2 + \varphi_2$, $t = t_1 + t_2$. Thus

$$\begin{aligned}\sum_{x \in V} d(x) \le & 2a_1a_2 + \varphi_1(\gamma_1 + c) + \varphi_2(\gamma_2 + c) + \gamma_1(c + c_2 + \gamma_1) \\ & + \gamma_2(c + c_1 + \gamma_2) \\ & + (c_1 - \varphi_1 - \gamma_1)a_2 + (c_2 - \varphi_2 - \gamma_2)a_1 \\ = & 2a_1a_2 + \varphi_1 c + \varphi_2 c + \gamma_1(t_1 + c_2 + c) + \gamma_2(t_2 + c_1 + c) \\ & + (c_1 - t_1)a_2 + (c_2 - t_2)a_1 \\ = & 2a_1a_2 + tc + \gamma_1(t_1 + c_2) + \gamma_2(t_2 + c_1) + (c_1 - t_1)a_2 \\ & + (c_2 - t_2)a_1.\end{aligned} \tag{15}$$

Here $t_1 + c_2 \le c_1 + c_2 = c$, and $t_2 + c_1 \le c$, $\gamma_1 + \gamma_2 \le t$. Thus

$$\gamma_1(t_1 + c_2) + \gamma_2(t_2 + c_1) \le (\gamma_1 + \gamma_2)c \le ct.$$

By (15),

$$\sum_{x\in V} d(x) \leq 2a_1a_2 + 2tc + (c_1 - t_1)a_2 + (c_2 - t_2)a_1.$$

We have to prove that this is smaller than $2e(H_n)$, i.e. that

$$2a_1a_2 + 2tc + (c_1 - t_1)a_2 + (c_2 - t_2)a_1 \leq 2a_1a_2 + (c_2 - t_2)a_1 + (c_1 - t_1)a_2 + \frac{(c+t)^2}{2}. \tag{16}$$

■

(C_2) The second case is when the maximum horizontal degree $D \geq c$. Assume that x_0 has horizontal degree D. Having used the Greedy algorithm in defining the above vertex partitioning, we know that x_0 has to be joined to at least D vertices of the "other class". Hence we may fix on both sides a set B_i of size D joined to x_0. Let the number of horizontal edges in $A_1 \cup C_1$ be k_1D, while in $A_2 \cup C_2$ let it be k_2D. (These are the definitions of k_1, k_2.)

Let us consider the following graph: H_n has 5 classes U_i, where U_4 is partitioned into U' and U". The sizes are as follows: $|U'| = k_1$, $|U"| = k_2$, $U_1 = (a_1 + c_1) - D - k_1$, $U_2 = (a_2 + c_2) - D - k_2$, and $U_3 = D$, $U_5 = D$. All the vertices of the classes U_1 are joined completely to all the vertices of U_2 and U_5; all the vertices of the classes U_2 are joined completely to all the vertices of U_1 and U_3; U_4 is completely joined to U_3 and U_5.

We assert that H_n is "better" than G_n in the following sense:

$$e(G_n) \leq e(H_n).$$

Further,

$$D'(H_n) = D(H_n) \geq D'(G_n),$$

where $D'(H_n)$ is now defined as the number of edges incident to C to be deleted to make the graph bipartite.

Since the last graph is of form $H(n_1, \ldots, n_5)$, we know that the best way to make it bipartite is to delete all the edges between two groups. So $(k_1 + k_2)D$ is really the minimum number of edges to be deleted to make this graph into bipartite, assumed that *this* product is the smallest. If this product is not the smallest then $|U_1| < k_1 + k_2$, $|U_2| < k_1 + k_2$ or $|U_1|, |U_2| < D$. In these cases we will see that $n^2/5 > e(H_n) \geq e(G_n)$. We have $e(H_n) = n_1n_2 + n_2n_3 + n_3n_4 + n_4n_5 + n_5n_1$ with $n_1 = a_1 + c_1 - D - k$,

$n_2 = a_2 + c_2 - D - k_2$, $n_3 = n_5 = D$, $n_4 = k = k_1 + k_2 < D, n/5$. The complete 5-partite graph $K[n_1, n_2, n_3, n_4, n_5]$ has at most $2n^2/5$ edges by Turán's theorem, thus it is sufficient to prove that

$$n_1n_2 + n_2n_3 + n_3n_4 + n_4n_5 + n_5n_1 < n_1n_3 + n_3n_5 + n_5n_2 + n_2n_4 + n_4n_1$$

i.e.

$$n_1n_2 + n_1D + n_2D + 2Dk < n_1D + D^2 + n_2D + n_2k + n_1k$$

which is equivalent to

$$(D-k)^2 - (n_1-k)(n_2-k) > 0. \tag{17}$$

(i) If $n_1, n_2 \geq k$ then we have to prove (17) only if $n_1, n_2 < D$ and in this case (17) holds by $D - k > n_i - k \geq 0$.

(ii) if, say, $n_1 \geq k$, $n_2 < k$ then (17) holds since the left hand side of it has a positive and a non-negative term.

(iii) if $n_1, n_2 < k$ then suppose that

$$(D-k)^2 \leq (n_1-k)(n_2-k) \leq \left(k - \frac{n_1+n_2}{2}\right)^2$$

i.e.

$$D - k \leq k - \frac{n_1+n_2}{2}.$$

It yields

$$4k \geq n_1 + n_2 + 2D = n - k$$

which contradicts $k < n/5$.

Now, let us estimate and calculate the number of edges in G_n and H_n, respectively . There is no reason to calculate the horizontal edges, we just have taken care of them.

What about the vertical edges, vertical degrees?

In case of H_n we use the following calculation: Above, for $n_i = a_i + c_i$, $n_1 - D - k_1$ vertices are joined to $n_2 - k_2$ vertices; D vertices are joined to $n_2 - D$ vertices from the other class, and k_1 vertices are joined to D vertices from the other class. So

$$\begin{aligned} e(H_n) &= (n_1 - D - k_1)(n_2 - k_2) + D(n_2 - D) + k_1D \\ &= (n_1 - D - k_1)(n_2 - k_2) + (D - k_1)(n_2 - D) + k_1n_2. \end{aligned}$$

In G_n we can estimate the number of edges as follows:

The number of horizontal edges is estimated by the same amount: $(k_1 + k_2)D$. The number of vertical edges incident to the k_1 independent horizontal edges is estimated by k_1n_2, since each vertex in $A_2 \cup C_2$ is joined to at most one end of such an edge. For the remaining at least $D - k_1$ vertices in B_1 we can use the stronger estimate $n_2 - D$ and for all the other vertices above we can use the weaker bound $n_2 - k_2$, since these vertices are joined to at most one end of each edge of the k_2 horizontal selected edges.

We obtained the desired estimate and this completes the proof. ∎

4. Proof of Theorem 4

It is enough to prove that if (G_n) is a sequence of graphs not containing K_3 and having at least

$$\frac{n^2}{25} + o(n^2)$$

edges and — for the corresponding pentagonlike graphs H_n — G_n satisfies

$$e(G_n) \geq e(H_n) - o(n^2)$$

then G_n can be made pentagonlike by deleting $o(n^2)$ edges. We have to deal with the two cases $D \leq c$ and $D \geq c$ separately, like in the proof of Theorem 3. In general, we follow the estimations in the proof of Theorem 3 and use its notation.

(A) Assume that $D \leq c$.

Unless $t = c - o(n)$, we gain some $\alpha_1 n^2$ edges for infinitely many of our graphs in (16), which implies that

$$e(G_n) \leq e(H_n) - \alpha_1 n^2,$$

contradicting our assumptions. Therefore we may assume that $t = c - o(n)$. Further, since $c \geq \gamma_1 + \gamma_2 + 2\varphi_1 + 2\varphi_2 = t + \varphi_1 + \varphi_2$, therefore $\varphi_1 + \varphi_2 = o(n)$.

A similar argument shows that $|N(x) \cap A_i| = c - o(n)$ for almost all the vertices $x \in C_i$. Indeed, if for $\alpha_2 n$ of them we had $|N(x) \cap A_i| \leq c - \alpha_3 n$, then we could gain $\alpha_3 n$ in (12) and it would improve our final inequality (14) by $\alpha_2\alpha_3 n^2$, contradicting our assumption. It is also clear that for almost each $x \in C$ the pair

$$(N(x) \cap A_1 \ ; \quad N(x) \cap A_2)$$

is almost the same, say $(U_3; U_5)$ with $|U_3|, |U_5| \geq c - o(n)$ where $|U_3| = c - o(n)$ if $c_1 \neq o(n)$ and $|U_5| = c - o(n)$ if $c_2 \neq o(n)$. Finally, it is not too difficult to show that $a_1 - a_2 = o(n)$, in this case, otherwise we could move some vertices from the larger A_i to the smaller one, increasing the total number of edges by some αn^2.

(B) We have applied Lemma 1 to get k_1, k_2 independent edges. If we can get more than $k_1 + k_2 + \alpha_4 n$ independent edges, then we gain at least $\alpha_5 n^2$ edges in the estimates on $e(G_n)$. This being excluded, we may apply Lemma 1**: There are k_1 vertices (forming a class W_1) in C_1 almost completely joined to D vertices (forming a class Q_1) of A_1 and k_2 vertices (forming a class W_2) in C_2 almost completely joined some other D vertices of A_2, (forming a class Q_2). Further, the number of other edges in $A_1 \cup C_1$ (and $A_2 \cup C_2$ respectively) is only $o(n^2)$: these edges can be disregarded. If $k_2 \neq o(n)$ then almost all of the vertices of $C_1 - W_1$ are joined to each edge of the k_2-matching by exactly one edge, and either they are joined to the vertices in Q_2 or to ones in W_2, and since $|Q_2| > |W_2| + o(n)$ and we cannot gain $\alpha_6 n^2$ in the estimate so almost all of the vertices of $C_1 - W_1$ are joined to almost all vertices of Q_2. Thus $C_1 - W_1$ can be added to $A_1 - Q_1$, and similarly $C_2 - W_2$ to $A_2 - Q_2$. It is also clear that $N(x) \approx Q_1 \cup Q_2$ for almost all $x \in W_1 \cup W_2$ and we get a pentagonlike graph with classes $A_1 \cup C_1 - W_1$, $A_2 \cup C_2 - W_2, Q_1, W_1 \cup W_2, Q_2$. ■

5. Proof of Theorem 5.

It is enough to prove Theorem 5 when we have a sequence of graphs (G_n) with

$$e(G_n) \geq \frac{n^2}{5} - o(n^2)$$

and when (G_n) contains $o(n^3)$ K_3's. We may delete $o(n^2)$ edges of G_n so that in the remaining graph G_n^* each edge is contained in at most $o(n)$ triangles. Now all the estimates (9)–(16) on the degrees and sums of degrees remain valid if we add an extra term $+o(n)$ to the right hand side (or $+o(n^2)$, when we estimate numbers of edges between groups of vertices). ■

6. Proof of Theorems 6,6*

Here and below the theory of supersaturated graphs will be used. Given a family $\mathbb{L}$ of forbidden graph, a graph G_n with $e(G_n) > \mathrm{ext}(n, \mathbb{L})$ will be called supersaturated. There are many results stating that supersaturated graphs contain many forbidden subgraphs. Below we shall use the following results.

Theorem. [7] *If G_n is a p-uniform hypergraph with cn^p p-edges, then it contains p classes $X_1, \dots, X_p$ of size t each so that all the t^p p-tuples*

$$(x_1, \dots, x_p) : x_i \in X_i \qquad \text{for} \qquad i = 1, \dots, p$$

are hyperedges of G_n.

Corollary 3. *If G_n is an ordinary graph containing at least cn^p copies of K_p for some fixed $c > 0$ and if n is sufficiently large, ($n > n_0(c,p,t)$) then G_n contains a $K_p(t, t, \dots, t)$.*

This result is extended by the following theorem of Erdős and Simonovits:

Theorem. [9] *Assume that L is a forbidden graph and $c > 0$. There exists a $c_L > 0$ such that if $e(G_n) > \mathrm{ext}(n, L) + cn^2$ then G_n contains at least $[c_L n^{v(L)}]$ copies of L.*

Both Theorem 6 and Theorem 6* easily follow from Theorem 5. Indeed, let $v = v(L)$. Now $L \subset K(v, v, v)$.

Let us consider a sequence of graphs, (G_n) satisfying the conditions of Theorems 6 or 6*, with $o(n^2)$ instead of δn^2.

If for some constant $\eta > 0$, G_n contains at least ηn^3 triangles, then — by the Erdős Theorem [7] applied to the triples of $V(G_n)$ formed by the triangles of G_n, $G_n \supseteq K(v, v, v)$, proving that $L \subseteq G_n$. This is excluded by the condition of Theorem 6. By a generalization of the above theorem of Erdős, [see 9], if G_n contains ηn^3 triangles, then it contains at least

γn^{3v} copies of $K(v,v,v)$, and therefore at least γn^v copies of L. (Each L can be extended in less than n^{2v} ways into a $K(v,v,v) \subseteq G_n$.) This case is excluded by the conditions of Theorem 6*. Thus in both cases we may assume that G_n contains $o(n^3)$ copies of K_3. Hence the assertion of Theorems 6,6* immediately follow from Theorem 5. ■

References

[1] P. Erdős, Oral communication, common knowledge

[2] P. Erdős, On bipartite subgraphs of graphs (in Hungarian), *Matematikai Lapok* **18**(1967), 283–288.

[3] P. Erdős, Some recent results on extremal problems in graph theory (Results), in: *Theory of Graphs, International symposium,* Rome, (1966), Gordon and Breach, New York and Dunod, Paris, 1967, 118–123.

[4] P. Erdős, On some new inequalities concerning extremal properties of graphs, in: *Theory of Graphs, Proc. Coll. Tihany, Hungary* (eds.: P. Erdős and G. Katona), Acad. Press. N. Y. 1968, 77–81.

[5] P. Erdős, On some extremal problems on r-graphs, *Discrete Math.* **1**(1971) 1–6.

[6] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. of Mathematics* **2/3**(1964), 183–190. (Reprinted in Art of Counting, MIT PRESS, 1973)

[7] P. Erdős, R. J. Faudree J. Pach and J. Spencer, How to make a graph bipartite, *J. Combinatorial Th. B* **45**(1988) 86–98.

[8] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1**(1966) 51–57, (reprinted in the Art of Counting, MIT PRESS, 1973, 194–200).

[9] P. Erdős and M. Simonovits, Supersaturated graphs and hypergraphs, *Combinatorica* **3**(1983), 181–192.

[10] M. Simonovits, A method for solving extremal problems in graph theory, in: *Theory of Graphs, Proc. Coll. Tihany, Hungary* (eds.: P. Erdős and G. Katona), Acad. Press. N. Y. 1968, 279–319.

[11] M. Simonovits, Extremal Graph Theory, in: *Selected Topics in Graph Theory* (eds.: Beineke and Wilson), Academic Press, London, New York, San Francisco, 1983, 161–200.

[12] M. Simonovits, Extremal graph problems with conditions, in: *Combinatorial Theory and its Applications,* (eds.: P. Erdős, A. Rényi and V. T. Sós), Colloq. Math. Soc. János Bolyai, **4**, North-Holland, Amsterdam, 1969, 999–1011.

[13] E. Szemerédi, On regular partitions of graphs, in: *Problemes Combinatoires et Théorie des Graphes,* (eds.: J. Bermond et al.), Orsay, 1976, CNRS Paris, 1978, 399–401.

[14] A. A. Zykov, On some properties of linear complexes, *Mat. Sbornik* **24** (1949), 163–188, *Amer. Math. Soc. Translations No 79.* 1952.

Paul Erdős

Mathematical Institute of the
Hungarian Academy of Sciences
H-1364 Hungary
Budapest, P.O.B. 127

Ervin Győri

Mathematical Institute of the
Hungarian Academy of Sciences
H-1364 Hungary
Budapest, P.O.B. 127

Miklós Simonovits

Mathematical Institute of the
Hungarian Academy of Sciences
H-1364 Hungary
Budapest, P.O.B. 127