COMBINATORICA 3 (2) (1983) 181—192

SUPERSATURATED GRAPHS AND HYPERGRAPHS

PAUL ERDŐS and MIKLÓS SIMONOVITS

Received 9 April 1982

We shall consider graphs (hypergraphs) without loops and multiple edges. Let $\mathcal{L}$ be a family of so called prohibited graphs and $\mathrm{ex}(n, \mathcal{L})$ denote the maximum number of edges (hyperedges) a graph (hypergraph) on n vertices can have without containing subgraphs from $\mathcal{L}$. A graph (hypergraph) will be called supersaturated if it has more edges than $\mathrm{ex}(n, \mathcal{L})$. If G has n vertices and $\mathrm{ex}(n, \mathcal{L})+k$ edges (hyperedges), then it always contains prohibited subgraphs. The basic question investigated here is: At least how many copies of $L\in\mathcal{L}$ must occur in a graph G^n on n vertices with $\mathrm{ex}(n, \mathcal{L})+k$ edges (hyperedges)?

Notation. In this paper we shall consider only graphs and hypergraphs without loops and multiple edges, and all hypergraphs will be uniform. If G is a graph or hypergraph, $e(G)$, $v(G)$ and $\chi(G)$ will denote the number of edges, vertices and the chromatic number of G, respectively. The first upper index (without brackets) will denote the number of vertices: G^n, S^n, $T^{n,p}$ are graphs of order n. $K_p^{(h)}(m_1, \ldots, m_p)$ denotes the h-uniform hypergraph with $m_1+\ldots+m_p$ vertices partitioned into classes $C_1, \ldots, C_p$, where $|C_i|=m_i$ $(i=1, \ldots, p)$ and the hyperedges of this graph are those h-tuples, which have at most one vertex in each C_i. For $h=2$ $K_p(m_1, \ldots, m_p)$ is the ordinary complete p-partite graph.

In some of our assertions we shall say e.g. that "changing $o(n^2)$ edges in G^n...". (Of course, $o(\cdot)$ cannot be applied to one graph.) As a matter of fact, in such cases we always consider a sequence of graphs G^n and $n\to\infty$.

Introduction

Let $T^{n,p}$ denote $K_p(n_1, \ldots, n_p)$ with the maximum number of edges with $n_1+\ldots+n_p=n$. In 1941, P. Turán [21] proved, that:

Among the graphs G^n with n vertices and containing no K_{p+1}, there exists exactly one having maximum number of edges, namely, $T^{n,p}$.

AMS subject classification (1980): 05 C 35,05 C 65

In particular, if $e(G^n) > \left[\frac{n^2}{4}\right]$, then G^n contains a K_3. Rademacher, (1941, unpublished) improved this result by showing that if $e(G^n) > \left[\frac{n^2}{4}\right]$, then G^n contains at least $\left[\frac{n}{2}\right]$ K_3's. Perhaps this was the first result in the area we call "the theory of supersaturated graphs". P. Erdős [5], [10], generalizing this result, showed that: *For every fixed p there exists a constant $c_p > 0$ such that if $0 < k < c_p n$ and $e(G^n) = e(T^{n,p}) + k$, then G^n contains at least as many K_{p+1}'s as the graph obtained from $T^{n,p}$ by putting k edges in one of its maximal classes (so that the new edges form no triangle).* (For further results see [1—3], [4], [18—19]).

In general, we shall always fix whether the considered objects are ordinary or (for some given $h \geqq 3$) h-uniform hypergraphs. We shall also fix some family $\mathscr{L}$ of prohibited graphs (hypergraphs). For a given $\mathscr{L}$, ex $(n, \mathscr{L})$ denotes the maximum number of edges a graph G^n can have without containing subgraphs from $\mathscr{L}$. Given E, we shall try to determine the *minimum* number of copies $L \in \mathscr{L}$ a graph G with n vertices and E edges must contain. This minimum will be denoted by $f(n, \mathscr{L}, E)$, and problems of this type will be called *extremal problems for supersaturated graphs.*

Theorems on supersaturated graphs are sometimes interesting for their own sake, in other cases for applications. Many of these applications have the following form: we would like to prove that ex $(n, \mathscr{L}) \leqq f(n)$ and

(a) for some L^* we know—by some theorem on supersaturated graphs—that if $e(G^n) > f(n)$, then G^n contains "many" cpoies of L^*.

(b) Further, we know, that if G^n contains "many" copies of L^*, then it con-. tains an $L \in \mathscr{L}$.

Thus we obtain that ex $(n, \mathscr{L}) \leqq f(n)$. Arguments of this type were used, e.g., in the proof of [17]:

$$\text{ex}\,(n, K_2(p, q)) \leqq \frac{1}{2} \sqrt[p]{q-1}\, n^{2-(1/p)} + O(n),$$

where $L^* = K_2(1, p)$ was used. Similarly, let Q^8 be the cube graph. We proved in [14] that

$$\text{ex}\,(n, Q^8) = O(n^{8/5}).$$

In order to do this we counted the number of $C^4 \subseteqq G^n$ under the condition that $e(G^n) > cn^{8/5}$. Finally we mention that [23] implies a "supersaturated graph theorem" for the number of walks W^{k+1} (of k edges) and using this result we have proved several sharp results in [12], among others, that

$$\text{ex}\,(n, \{C^4, C^5\}) = \left(\frac{n}{2}\right)^{3/2} + o(n^{3/2}).$$

In the last paragraph we shall give another application of this kind: combining Theorem 2 and the Lovász—Simonovits theorem we deduce the Erdős-Simonovits sharpening of the Erdős—Stone theorem, (see below).

One of our most general results will be

Theorem 1. *Given a family $\mathcal{L}$ of h-uniform hypergraphs, let*

$$\mathcal{L}_t = \{L \in \mathcal{L}: v(L) \geqq t\}.$$

For every $c>0$ there exists a $c'>0$ such that if

$$e(G^n) > \operatorname{ex}(n, \mathcal{L}_t) + cn^h, \tag{1}$$

then G^n contains at least $c'n^t$ prohibited $L \in \mathcal{L}_t \subseteqq \mathcal{L}$.

Remark. One could say, that in the formulation of Theorem 1 $\mathcal{L}$ is not needed at all, we could restrict ourselves to $\mathcal{L}_t$. This is true. However, this form of Theorem 1 enables us to formulate a sharp result for the general case.

Corollary 1. *Restrict ourselves to h-uniform hypergraphs and assume also that a finite family $\mathcal{L}$ of prohibited graphs is given. Find the largest t for which*

$$\operatorname{ex}(n, \mathcal{L}) - \operatorname{ex}(n, \mathcal{L}_t) = o(n^h). \tag{2}$$

Then, for every $c>0$, there exist $c', c''>0$ such that

$$c'n^t \leqq f(n, \mathcal{L}, \operatorname{ex}(n, \mathcal{L}) + cn^h) \leqq c''n^t.$$

(Or, in simpler words, if an h-uniform hypergraph has by cn^h more hyperedges than what is allowed in the extremal graph problem, then it contains at least cn^t prohibited $L \in \mathcal{L}$, but nothing more can be guaranteed.)

Here the lower bound follows immediately from Theorem 1, while the upper bound can be obtained as follows. By the maximality of t, (2) does not hold for $t+1$. In [16] it is proved that

$$\lim_{n\to\infty} \operatorname{ex}(n, \mathcal{L}) \Big/ \binom{n}{h}$$

exists, (moreover, the ratio is monotone decreasing). Let

$$c_1 = \lim \frac{\operatorname{ex}(n, \mathcal{L}_{t+1}) - \operatorname{ex}(n, \mathcal{L})}{\binom{n}{h}} > 0.$$

By definition, if $\tilde{c} < c_1$ and n is sufficiently large, then we may choose G^n so that $e(G^n) = \operatorname{ex}(n, \mathcal{L}) + \left[\tilde{c}\frac{n^h}{h!}\right]$, but G^n contains no graphs from $\mathcal{L}_{t+1}$. In other words, the prohibited graphs in G^n have at most t vertices. This proves the upper bound for small values of $\tilde{c}$, and if we know it for small values, that implies the statement for large values as well.

Theorem 1 has many equivalent forms.

Theorem 1*. *We consider h-uniform hypergraphs. For every L with $t = v(L)$ for every $c>0$ there exists a $c'>0$ such that if*

$$e(G^n) > \operatorname{ex}(n, L) + cn^h,$$

then G^n contains at least $c'n^t$ copies of L.

(Below, if $\mathscr{L}=\{L\}$, we shall often write $\operatorname{ex}(n, L)$ instead of $\operatorname{ex}(n, \{L\})$.) Clearly, Theorem 1* is a special case of Theorem 1. The following assertion is a very important subcase of Theorem 1*. Let, for $p \geqq h$, $K_p^{(h)}(m_1, \ldots, m_p)$ denote the following generalization of the *complete p-partite* graphs. Fix p disjoint sets of vertices, $C_1, \ldots, C_p$, where C_i has m_i elements, and take *all* those h-tuples which contain *at most* one vertex from each C_i. The corresponding generalization of the *complete bipartite* graph is $K_h^{(h)}(m_1, \ldots, m_h)$. A theorem of Erdős [7] asserts, that

(3) $$\operatorname{ex}(n, K_h^{(h)}(m, \ldots, m)) = O(n^{h-1/m^{(h-1)}}) = o(n^h).$$

(In case $h=2$ we get the Kővári—T. Sós—Turán theorem [17]). Now, applying Theorem 1* to this $K_h^{(h)}(m, \ldots, m)$ we obtain

Corollary 2. *Given $c>0$, there exists a $c'>0$ such that if an h-uniform hypergraph G^n has at least cn^h hyperedges, then it contains at least $c'n^{hm}$ copies of $L=K_h^{(h)}(m, \ldots, m)$.*

One can easily see that Corollary 2 is not only a special case of Theorem 1, but is also equivalent to it. We shall prove the more general

Theorem 1.** *There exist $c=c_{m,h}>0$ and $c^*=c^*_{m,h}>0$ such that every h-uniform G^n with*

(4) $$E = e(G^n) > c_{m,h} n^{h-(1/m^{h-1})}$$

edges contains at least

$$c^*_{m,h} n^{mh} \cdot \left(\frac{E}{n^h}\right)^{m^h}$$

copies of $K_h^{(h)}(m, \ldots, m)$. This is sharp: almost all hypergraphs with E hyperedges contain

$$O\left(n^{mh} \cdot \left(\frac{E}{n^h}\right)^{m^h}\right)$$

*copies of $K_h^{(h)}(m, \ldots, m)$.**

The behaviour of $\operatorname{ex}(n, \mathscr{L})$ is fairly complicated and our knowledge on it is rather poor. Some of the "most elementary" problems defy all our efforts to solve them. The situation is much nicer for ordinary graphs. The authors have described the function $\operatorname{ex}(n, \mathscr{L})$ and the structure of extremal graphs for $\mathscr{L}$ sufficiently well in [13], [8], [9], [20]:

Erdős—Stone—Simonovits theorem [13]. *Let $p=p(\mathscr{L})=\min_{L \in \mathscr{L}} \chi(L)-1$. Then*

$$\operatorname{ex}(n, \mathscr{L}) = \left(1-\frac{1}{p}\right)\binom{n}{2} + o(n^2).$$

Erdős—Simonovits theorem [8], [9], [20]. *Let $p=p(\mathscr{L})=\min \chi(L)-1$. Then every extremal graph S^n for $\mathscr{L}$ can be obtained from the Turán graph $T^{n,p}$ by deleting and adding $o(n^2)$ edges. Further, the minimum degree $\underline{d}(S^n)=\left(1-\frac{1}{p}+o(1)\right)n$.*

* The proof of the sharpness will be left to the reader.

These results are interesting for us in two different ways. On the one hand we shall deduce them from our general result, (for the simpler Erdős—Stone—Simonovits theorem this have already been done in [11]), on the other hand, we can formulate Theorem 1 for ordinary graphs in a more explicit form.

Theorem 2. *Let us consider ordinary graphs and let* $p=\min\chi(L)-1$ *for a given family* $\mathcal{L}$ *of prohibited graphs. If there exists a* $(p+1)$*-chromatic* $L\in\mathcal{L}$ *with* t *vertices, then for every* $c>0$ *there is a* $c'>0$ *such that if*

$$e(G^n) > \left(1-\frac{1}{p}+c\right)\binom{n}{2},$$

then G^n *contains at least* $c'n^t$ *prohibited subgraphs.*

The above theorems gave lower bounds on the number of prohibited subgraphs in a supersaturated graph. The next theorem is a stability theorem in the following sense: it asserts that either a supersaturated (ordinary) graph G^n contains many prohibited subgraphs or it has almost the same structure as the extremal graphs for $\mathcal{L}$.

Theorem 3. *Let us consider ordinary graphs. Let* $\mathcal{L}$ *be a finite family of prohibited subgraphs and* $p(\mathcal{L})=p=\min\limits_{L\in\mathcal{L}}\chi(L)-1$. *Let* t *be the maximum number of vertices in the* $(p+1)$*-chromatic graphs of* $\mathcal{L}$. *If* $e(G^n)>\operatorname{ex}(n,\mathcal{L})$ *and* G^n *contains only* $o(n^t)$ *prohibited subgraphs, then it can be obtained from* $T^{n,p}$ *by changing* $o(n^2)$ *edges.*

This will easily imply the Erdős—Simonovits theorem and also

Theorem 4. *Let us consider ordinary graphs. Let* $\mathcal{L}$ *be a family of prohibited subgraphs,* $p=\min\limits_{L\in\mathcal{L}}\chi(L)-1$, *and assume that, for some* $k=k_n=o(n^2)$, we have $e(S^n)=$ $=\operatorname{ex}(n,\mathcal{L})+k=E$. *Suppose further that* S^n *is extremal for the supersaturated problem, that is, it contains the minimum of* $L\in\mathcal{L}$ *among the graphs* G^n *with* E *edges. Then* S^n *can be obtained from* $T^{n,p}$ *by changing* $o(n^2)$ *edges.*

Some recursions on ex $(n, \mathcal{L})$ and $f(n, \mathcal{L}, E)$

The following graph theoretical operation will be used in our proofs.

Definition. If L is an $(h-1)$-uniform hypergraph, $L^{(r)}$ will denote the h-uniform hypergraph obtained by fixing r new vertices $x_1, \dots, x_r$ and taking the hypergraph whose vertices are the vertices of L and the new vertices; the hyperedges are those h-tuples which consist of an edge of L and a new vertex.

Example. If $L=K_p^{(h)}(m, \dots, m)$, then $L^{(r)}=K_{p+1}^{(h+1)}(r, m, \dots, m)$.

The next theorems are recursions on the number of prohibited subgraphs and on ex $(n, \mathcal{L})$. They may seem rather technical at the first sight, however, they are often very useful.

Theorem 5. *Let* L *be an* $(h-1)$*-uniform hypergraph,* $v(L)=v$. *Assume that a sequence* $E_n>\operatorname{ex}(n, L)$ *is fixed and* $f(n, E)$ *is a function for which*

(i) $f(n, E)=0$ *for* $E<E_n$;

$f(n, E)$ is monotone increasing and convex for every fixed n while $E_n \leqq E \leqq \binom{n}{2}$.

(ii) $f(n, E+E_n) \geqq f(n, E)+f(n, E_n)$.

(iii) *If H^n is an $(h-1)$-uniform hypergraph with E edges, then it contains at least $f(n, E)$ copies of L.*

Then every h-uniform hypergraph S^n with F hyperedges contains at least

$$c\frac{f\left(n, \frac{hF}{4n}\right)^r}{n^{vr-v-r}} \tag{5}$$

copies of $L^{(r)}$, for some constant $c>0$.

Remarks. Conditions (i) and (ii) are merely technical assumptions. We are interested primarily in functions of form

$$f(n, E) = \begin{cases} t(n)E^a & \text{if } E \geqq E_n, \\ 0 & \text{if } E < E_n, \end{cases} \tag{6}$$

for some $a>1$. For these functions (i) and (ii) automatically hold and (5) reduces to

$$f(n, L^{(r)}, F) \geqq g(n, F) = \begin{cases} \frac{c^*}{4^{ar}} \cdot \frac{t(n)^r F^{ar}}{n^{vr-ar-v-r}} & \text{if } F \geqq F_n = \frac{4nE_n}{h}; \\ 0 & \text{if } F < F_n. \end{cases} \tag{7}$$

One further comment on (i) and (ii) is that if $f(n, L, E)$ denotes the minimum number of L's an $(h-1)$-uniform hypergraph G^n with E edges must contain, then f satisfies (ii), but often is not convex. This is, why we must choose a smaller function f which is already convex and still satisfies (ii).

The following theorem is very similar to the previous one.

Theorem 6. *Under the conditions of Theorem 5 every h-uniform hypergraph S^n with F hyperedges contains at least*

$$n^r \cdot f\left(n, \frac{cF^r}{n^{h(r-1)+1}}\right) \tag{8}$$

copies of $L^{(r)}$ for some constant $c>0$.

These recursive theorems on oversaturated graphs contain regular extremal graph theorems as a particular case. Thus, e.g., we shall use Theorem 6 to prove

Theorem 7. *There exist a constant* $c>0$ *and an integer* $n_0(L, c, h)$ *such that*

$$\mathrm{ex}\,(n, L^{(r)}) \leqq c \cdot n^{h\left(1-\frac{1}{r}\right)+\frac{1}{r}} \cdot \mathrm{ex}\,(n, L)^{\frac{1}{r}} \quad \textit{if} \quad n>n_0(L, c, h).$$

This recursion theorem will yield the old result of Erdős, [7]:

Erdős-theorem. ex $\left(n, K_h^{(h)}(m, \ldots, m)\right)=O\left(n^{h-(1/m^{h-1})}\right)$.

Indeed, for $h=1$ the theorem is trivial. (For $h=2$ it follows from the Kővári—T. Sós—Turán theorem, but we do not need this.) We may use induction on h: if we know Erdős theorem for $K_{h-1}^{(h-1)}(m, \ldots, m)=L$, then, applying Theorem 7 to this L with $r=m$, we obtain the theorem for $L^{(m)}=K_h^{(h)}(m, \ldots, m)$.

Proofs

As we have mentioned, most of our results will be derived from Theorem 5.

Proof of Theorem 5. Given a vertex x, S_x will denote the $(h-1)$-uniform hypergraph whose vertices are the same as the vertices of S^n and whose edges are those $(h-1)$-tuples, which together with x form a hyperedge of S^n. We shall count the pairs $\{(L, x): L \subseteqq S_x\}$, *first fixing x, then fixing L.* These pairs will be called "incidences".

(A) If x is fixed, we have at least $f(n, e(S_x))$ copies of L in S_x (by the definition of f). Thus *the number of incidences is at least* $\sum_x f(n, e(S_x))$. Let

$$g(E) = g(n, E) := \begin{cases} f(n, E) & \text{if } E > E_n \\ f(n, E_n) & \text{if } E \leqq E_n. \end{cases} \tag{9}$$

Clearly, $g(E)$ is convex, further, $\sum_x e(S_x) = hF$. Thus

$$\sum_x f(n, e(S_x)) \geqq \sum_x \big(g(e(S_x)) - f(n, E_n)\big) \geqq n\left[g\left(\frac{hF}{n}\right) - f(n, E_n)\right]. \tag{10}$$

(B) Let q_L denote the number of x's such that $L \subseteqq S_x$. Then the *number of $L^{(r)} \subseteqq S^n$ is exactly* $\frac{1}{T} \sum \binom{q_L}{r}$, where T denotes the number of $L \subseteqq L^{(r)}$ generating $L^{(r)}$ (in the way described in the definition of $L^{(r)}$). Extending $\binom{x}{r}$ to the reals by

$$\binom{x}{r} := \begin{cases} \dfrac{x(x-1)(x-2)\ldots(x-r+1)}{r!} & \text{if } x > r-1 \\ 0 & \text{if } x \leqq r-1, \end{cases}$$

we get a convex function. We know that $\sum_L q_L$ is also the number of incidences. Thus by (10)

$$\sum q_L \geqq n\left[g\left(\frac{hF}{n}\right) - f(n, E_n)\right]. \tag{11}$$

By the convexity of $\binom{x}{r}$ we get that the number of $L^{(r)} \subseteqq S^n$ is at least

$$\frac{1}{T} \sum \binom{q_L}{r} \geqq \frac{1}{T} cn^v \cdot \binom{\dfrac{f\left(n, \frac{hF}{n}\right) - f(n, E_n)}{cn^{v-1}}}{r} \tag{12}$$

(since the number of summands is the number of L's on n vertices, which is asymptotically cn^v).

(C) As a matter of fact, (12) is just the sort of formula we needed, and it does not use (ii). However, it is not a nice looking formula. This is, why we assume (ii) and derive a simpler form of (12). (Only relatively simple formulas can effectively be used.) We *try* to prove

$$f\left(n, \frac{hF}{n}\right) - f(n, E_n) > f\left(n, \frac{hF}{4n}\right). \tag{13}$$

If we can prove this, then (12) immediately will give Theorem 5. However, for $\frac{hF}{4n} < E_n$, $f=0$, therefore the theorem guarantees no $L^{(r)}$ at all. Thus it holds. For $\frac{hF}{4n} \geqq E_n$,

$$\frac{hF}{n} - E_n > \frac{hF}{4n},$$

therefore (ii) yields (13), completing the proof. ∎

Proof of Theorem 1**. We use induction on h, applying Theorem 5 to $L = K_{h-1}^{(h-1)}(m, \dots, m)$ and

$$f(n, E) = \begin{cases} 0 & \text{if} \quad E < cn^{h-1-\frac{1}{m^{h-2}}} = E_n \\ n^{m(h-1)} \left(\dfrac{E}{n^{h-1}}\right)^{m^{h-1}} & \text{if} \quad E \geqq E_n. \end{cases} \tag{14}$$

(Theorem 1** is trivial for $h=1$; the function $f(n, E)$ trivially satisfies (i) and (ii), and—by the induction hypothesis—(iii) as well, with some constant $c = c_{h-1}$.) By Theorem 5 any h-uniform S^n with F edges contains at least

$$c' \cdot \frac{\left\{ n^{m(h-1)} \cdot \left(\dfrac{F}{n^{h-1}n}\right)^{m^{h-1}} \right\}^m}{n^{(h-1)m^2 - (h-1)m - m}} \tag{15}$$

copies of $L^{(m)} = K_h^{(h)}(m, \dots, m)$. Indeed, here $m = r$, $v = (h-1)m$, and we may use the second case of definition (14), since

$$F \geqq An^{h - \frac{1}{m^{h-1}}}$$

implies

$$\frac{hF}{4n} \geqq cn^{h-1-\frac{1}{m^{h-1}}}.$$

(15) yields that S^n contains at least

$$c^* \cdot \left(\frac{F}{n^h}\right)^{m^h} \cdot n^{hm} \tag{16}$$

copies of $L^{(m)}$, completing the proof. ∎

Proof of Corollary 2. If $F = an^h$ for some constant $a > 0$, then (16) yields $c'n^{hm}$ for some $c' > 0$. ▮

Proof of Theorem 1. In [11] Erdős proved the following simple but very useful lemma:

Lemma. *For a fixed integer m and fixed constants $q \geqq 0$, $c > 0$ there exists a constants $\eta > 0$ such that if*

$$e(G^n) \geqq (q+c)\binom{n}{h}, \tag{17}$$

then for at least $\eta\binom{n}{m}$ induced subhypergraphs $G^m \subseteqq G^n$

$$e(G^m) \geqq \left(q + \frac{c}{2}\right)\binom{m}{h}. \tag{18}$$ ▮

Let now

$$q = \lim_n \frac{\operatorname{ex}(n, \mathscr{L}_t)}{\binom{n}{h}}.$$

We may fix an m such that $\operatorname{ex}(m, \mathscr{L}_t) < \left(q + \frac{c}{2}\right)\binom{m}{h}$. Apply the above lemma to this m, c, and the G^n in Theorem 1. We obtain that G^n must contain at least $\eta\binom{n}{m}$ subhypergraphs G^m satisfying

$$e(G^m) > \operatorname{ex}(m, \mathscr{L}_t). \tag{19}$$

Thus each of these G^m's contains a prohibited $L \in \mathscr{L}_t$. Thus we obtain at least $\eta\binom{n}{m}$ prohibited subgraphs $L \in \mathscr{L}_t$ in G^n; however, many of these are *counted many times*. This does not really matter, since each L is contained in at most $\binom{n-t}{m-t}$ induced $G^m \subseteqq G^n$. Thus we obtain at least

$$\eta\binom{n}{m}\bigg/\binom{n-t}{m-t} \geqq c'n^t$$

different copies of prohibited subgraphs from $\mathscr{L}_t$. ▮

Observe, that Theorem 2 immediately follows from the Erdős—Stone—Simonovits theorem and Theorem 1. Next we shall prove Theorem 3. For this we shall use the following result of Lovász and Simonovits [19] (which is a particular case of a more general theorem).

Lovász—Simonovits theorem. *Let C be an arbitrarily large constant, p be a fixed integer. There exist a $\delta > 0$ and a $C' > 0$ for which, if*

$$e(G^n) = e(T^{n,p}) + k, \quad (0 < k < \delta n^2),$$

and G^n contains only Ckn^{p-1} copies of K_{p+1}, then G^n can be obtained from a $T^{n,p}$ by changing at most $C'k$ edges.

Remark. This is a "stability theorem": $T^{n,p}$ contains no K_{p+1}, and adding k edges to it we get a graph U^n with $\approx \left(\frac{n}{p}\right)^{p-1}$ K_{p+1}'s on each of the new edges. The above theorem asserts that either G^n has "much more" K_{p+1}'s, or it has almost the same structure as U^n.

Proof of Theorem 3. Let us fix an arbitrary $c>0$. Define a $(p+1)$-uniform hypergraph H^n on the vertices of G^n as follows. The hyperedges of H^n are the $(p+1)$-tuples forming a $K_{p+1} \subseteq G^n$. Select an $L \in \mathscr{L}$ with $\chi(L)=p+1$ and $v(L)=t$. Now,

(a) either $e(H^n) \leqq cn^{p+1}$, or
(b) take a minimal $K_{p+1}^{(p+1)}(m_1, \dots, m_{p+1})$ for which $K_{p+1}^{(2)}(m_1, \dots, m_{p+1}) \supseteq L$.

The minimality means that first we embed L into $K_{p+1}(t, \dots, t)$, and then delete all vertices of $K_{p+1}(t, \dots, t)$ not belonging to L. Thus $m_1+\dots+m_{p+1}=v(L)=t$. By Theorem 1*, H^n contains at least $c'n^t$ copies of $K_{p+1}^{(p+1)}(m_1, \dots, m_{p+1})$. Thus G^n contains at least $c'n^t$ copies of $K_{p+1}^{(2)}(m_1, \dots, m_{p+1}) \supseteq L$, which yields, that "$G^n$ contains at least cn^t prohibited subgraphs". In Theorem 3 we assumed that G^n has only $o(n^t)$ prohibited L's. Thus, for any $c>0$ and $n>n_0(c)$, G^n contains at most cn^{p+1} copies of K_{p+1}. In other words, G^n contains at most $o(n^{p+1})$ K_{p+1}'s. By the Lovász—Simonovits theorem, G^n can be obtained by changing $o(n^2)$ edges in $T^{n,p}$. ∎

Proof of Theorem 4. Again, add k edges to $T^{n,p}$ and regard the resulting graph U^n. If the new edges form a bipartite graph (what can be assumed), then U^n contains no $(p+2)$-chromatic $L \in \mathscr{L}$. Thus it contains at most $ck\left(\frac{n}{p}\right)^{t-2}$ prohibited L: indeed, if $L \subseteq U^n$, then $v(L) \leqq t$ and L contains one of the new edges. This shows that if S^n is extremal for the supersaturated problem in the sense that it has minimum number of prohibited $L \in \mathscr{L}$, then it contains at most $ck\left(\frac{n}{p}\right)^{t-2}$ prohibited subgraphs. Applying Theorem 3 to S^n we complete the proof. ∎

Proof of Theorem 6. Let us consider an h-uniform S^n with F hyperedges. If U is an $(h-1)$-tuple of vertices and R is an r-tuple such that each $x_i \in R$ forms a hyperedge of S with U, then (U, R) will be called "flower". The proof below will be very similar to that of Theorem 5, with the difference that we shall *count these "flowers"* instead of the "incidences".

Let s_U and a_R be the number of "flowers" (U, R) for fixed U and R, respectively. The number of flowers is

$$s = \sum_U s_U = \sum_R a_R. \tag{20}$$

Clearly, if T denotes the number of ways an $L^{(r)}$ is obtained from an $L \subseteq L^{(r)}$, then the number of $L^{(r)} \subseteq S^n$ is at least

$$\frac{1}{T}\sum_R f(n, a_R) \geqq \frac{1}{T}\sum_R (g(a_R)-g(E_n)) \geqq \frac{1}{T}\binom{n}{R}\left(\sum_R g\left(\frac{\sum a_R}{\binom{n}{R}}\right)-g(E_n)\right), \tag{21}$$

by the convexity of g. Let d_U denote the number of hyperedges containing the $(h-1)$-tuple U. By (20),

$$s=\sum_R a_R=\sum_U s_U=\sum_U\binom{d_U}{r}\geqq\binom{n}{h-1}\binom{\frac{hF}{\binom{n}{h-1}}}{r}\geqq\frac{cF^r}{n^{h(r-1)+1}} \tag{22}$$

since $\sum_U d_U=hF$. A short calculation, using (ii), yields the desired result. ∎

Proof of Theorem 7. Using the notations of the above proof, observe that if S^n contains no $L^{(r)}$, then, for every R, $a_R\leqq\operatorname{ex}(n, L)$, otherwise, there were $\operatorname{ex}(n, L)+1$ $(h-1)$-tuples U such that each $x\in R$ forms an edge of S with each U. Some of them would form an L which generates with R an $L^{(r)}$ in S^n.

Thus (22) yields that

$$\binom{n}{r}\operatorname{ex}(n, L)\geqq\binom{n}{h-1}\binom{\frac{hF}{\binom{n}{h-1}}}{r}\geqq\left(\frac{1}{r!}+o(1)\right)\frac{(hF)^r}{\binom{n}{h-1}^{r-1}}. \tag{23}$$

Rearranging (23) we get the desired assertion. ∎

Remark. Formally we may deduce Theorem 7 from Theorem 6: Set

$$f(n, E)=\begin{cases}\dfrac{E}{\operatorname{ex}(n, L)+1} & \text{if}\quad E\geqq E_n=\operatorname{ex}(n, L)+1\\ 0 & \text{otherwise.}\end{cases}$$

It is easy to see that Theorem 6 can be applied with this $f(n, E)$ and yields Theorem 7.

Appendix

We will deduce the Erdős—Simonovits theorem mentioned in the introduction from Theorem 3. Let S^n be an extremal graph for $\mathscr{L}$. Fix an $L_0\in\mathscr{L}$ with $\chi(L_0)=p+1$, and with the maximum number of vertices, t. Since S^n contains no L_0 at all, we may apply Theorem 3: S can be obtained from $T^{n,p}$ by changing $o(n^2)$ edges.

(The proof above is that easy, because we used the fairly deep Lovász—Simonovits theorem.)

References

[1] B. Bollobás, *Extremal graph theory*, London Math. Soc. Monographs, No. **11** Academic Press, 1978.

[2] B. Bollobás, Relations between sets of complete subgraphs, *Proc. Fifth British Comb. Conference*, Aberdeen, 1975, 79—84.

[3] B. Bollobás, On complete subgraphs of different orders, *Math. Proc. Cambridge Phil. Soc.* **79** (1976), 19—24.

[4] B. BOLLOBÁS, P. ERDŐS and M. SIMONOVITS. On the structure of edge graphs II, *J. London Math. Soc.* **12** (2) (1976) 219—224.
[5] P. ERDŐS, On a theorem of Rademacher—Turán, *Illinois J. Math.* **6** (1962) 122—127.
[6] P. ERDŐS, On the number of complete subgraphs contained in certain graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* 7 (1962) 459—464.
[7] P. ERDŐS, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1965) 183—190.
[8] P. ERDŐS, Some recent results on extremal problems in graph theory, *Theory of Graphs, Intern. Symp. Rome,* 1966, 118—123.
[9] P. ERDŐS. On some new inequalities concerning extremal properties of graphs, *Theory of Graphs, Proc. Coll. Tihany,* 1966, 77—81.
[10] P. ERDŐS, On the number of complete subgraphs and circuits contained in graphs, *Casopis Pest. Mat.* **94** (1969) 290—296.
[11] P. ERDŐS, On some extremal problems on r-graphs, *Discrete Math.* **1** (1971) 1—6.
[12] P. ERDŐS and M. SIMONOVITS, Compactness results in extremal graph theory, *Combinatorica* **2** (3) (1982), 275—288.
[13] P. ERDŐS and M. SIMONOVITS, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966) 51—57.
[14] P. ERDÖS and M. SIMONOVITS, Some extremal problems in graph theory, *Coll. Math. Soc. János Bolyai* **4** (1969) 377—390.
[15] P. ERDŐS and A. H. STONE, On the structure of linear graphs, Bull. Amer. Math. Soc. **52** (1946) 1089—1091.
[16] GY. KATONA, T. NEMETZ and M. SIMONOVITS, A new proof on a theorem of Turán and some remarks on a generalization of it, (in Hungarian) *Matematikai Lapok* **15** (1964) 228—238.
[17] T. KŐVÁRI, V. T. SÓS and P. TURÁN, On a problem of Zarankievicz, *Coll. Math.* 3 (1954) 50—57.
[18] L. LOVÁSZ and M. SIMONOVITS, On the number of complete subgraphs of a graph, *Proc. Fifth British Combinatorial Coll.,* Aberdeen, (1975) 431—442.
[19] L. LOVÁSZ and M. SIMONOVITS, On the number of complete subgraphs of a graph. II., *Studies in Pure Mathematics,* (1983), 459—495.
[20] M. SIMONOVITS, A method for solving extremal problems in graph theory, *in: Theory of Graphs, Proc. Coll. Tihany,* Hungary, 1966, 279—319.
[21] P. TURÁN, On an extremal problem in graph theory, Mat. Fiz. Lapok **48** (1941) 436—452.
[22] G. R. BLAKLEY and P. ROY, *Proc. Amer. Math. Soc.* **16** (1965) 1244—1245; see also H. P. MULHOLLAND and C. A. B. SMITH, *American Mathematical Monthly* **66**) (1959) 673—683, and D. LONDON, *Duke Math. Journal* **33** (1966) 511—522.

Paul Erdős, Miklós Simonovits

Mathematical Institute of the
Hungarian Academy of Sciences
Budapest, Hungary