© DISCRETE MATHEMATICS 5 (1973) 323–334. North-Holland Publishing Company

ON A VALENCE PROBLEM IN EXTREMAL GRAPH THEORY

P. ERDŐS

Hungarian Academy of Sciences, Budapest V, Hungary

and

M. SIMONOVITS

Eötvös Loránd University, Budapest, Hungary

Received 20 September 1972

Abstract. Let $L \neq K_p$ be a p-chromatic graph and e be an edge of L such that $L - e$ is $(p-1)$-chromatic. If G^n is a graph of n vertices without containing L but containing K_p, then the minimum valence of G^n is

$$\leqslant n\left(1 - \frac{1}{p - 3/2}\right) + O(1).$$

0. Notation

We consider only graphs without loops and multiple edges. The number of edges, vertices and the chromatic number of a graph G will be denoted by

$$e(G), v(G), \chi(G),$$

respectively. The number of vertices will also be indicated sometimes by the upper indices, e.g. G^n will always denote a graph of n vertices. $N(x)$ will denote the neighbours of the vertex x in G, i.e., the set of vertices joined to x; $\sigma(x)$ denotes the valence of x (= cardinality of $N(x)$) and $\sigma(G)$ denotes the minimum valence in G. If E is any set, $|E|$ denotes its cardinality.

Let $G_1, \ldots, G_d$ be given graphs, no two of which have common vertices. Joining every vertex of G_i to every vertex of G_j if $i \neq j$, we obtain the product

$$\mathop{\mathsf{X}}_{i=1}^{d} G_i = G_1 \times G_2 \times \ldots \times G_d.$$

K_p will denote the complete p-graph. $K_p(n_1, \ldots, n_p)$ denotes the complete p-chromatic graph having n_i vertices in its i^{th} class.

1. Introduction

B. Andràsfai asked the following question in connection with the well-known theorem of P. Turán [8]:

Problem. Determine

$$\psi(n, K_p, t) = \max \{\sigma(G^n) : K_p \not\subset G^n, \chi(G^n) \geqslant t\}.$$

In other words, what is the minimum value of k such that if every vertex of G^n has valence $\geqslant k$ and G^n is at least t-chromatic, then G^n contains a complete p-graph (if n, p and $t \geqslant p$ are given).

For $t \leqslant p-1$, Turán's theorem gives $\psi(n, K_p, t) = [n(1 - 1/(p-1))]$.

B. Andràsfai, P. Erdös and V.T. Sós [1] proved that

$$\psi(n, K_p, p) = (1 - 1/(p - \tfrac{4}{3}))n + \mathrm{O}(1). \tag{1}$$

The extremal graph, i.e. the graph for given n which attains the maximum, is the following one:

$$T^n = P^{m_0} \times K_{p-3}(m_1, \ldots, m_{p-3}),$$

where $\Sigma_{i=0}^{p-3} m_i = n$. The vertices of P^{m_0} are divided into the non-empty classes $C_1, \ldots, C_5$ and each vertex of C_i is joined to each vertex of C_{i+1} (where $C_6 = C_1$). (See Fig. 1.) Then $m_0, \ldots, m_{p-3}$ are chosen so that the minimum valence should be as large as possible.

One can easily show that, in this case,

$$|C_i| = n/(3p-4) + \mathrm{O}(1), \quad m_i = 3n/(3p-4) + \mathrm{O}(1),$$

and from this (1) follows immediately.

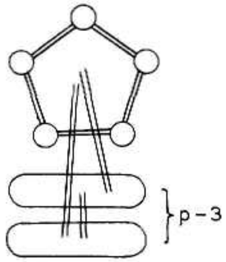

Fig. 1.

The case of $\psi(n, K_p, t)$ for $t > p$ seems to be much more difficult. E.g. even in the simplest case of $\psi(n, K_3, 4)$ we do not know whether

$$\psi(n, K_3, 4) \approx \tfrac{1}{3}n$$

or not. The authors of this paper and of [1] thought that there exists a sequence $\epsilon_t \to 0$ (when $t \to \infty$) such that

$$\psi(n, K_3, t) \leqslant \epsilon_t n.$$

An example, obtained in collaboration with A. Hajnal, will disprove this conjecture, showing that

$$\psi(n, K_3, t) \geqslant (\tfrac{1}{3} - o(1))n.$$

We conjecture that

$$\psi(n, K_3, t) \approx \tfrac{1}{3}n \qquad (t \geqslant 4),$$

but we can prove only

$$\overline{\lim_{n \to \infty}}\ n^{-1}\psi(n, K_3, 4) < \tfrac{2}{5}.$$

In this paper, we investigate

$$\psi(n, L, t) = \max\{\sigma(G^n) : G^n \not\supset L, \chi(G^n) \geqslant t\},$$

where L is a given (so called sample) graph. The valence-problems are interesting only in those cases, when they are not trivial consequences of the corresponding edge-problem. The edge-extremal problem of L is to determine

$$\max\{e(G^n): G^n \not\supset L\}.$$

The solution of such problems is fairly well described in [2,3,7] and if we suppose that $t \leqslant p-1$, then

$$(2) \qquad \psi(n, L, t) = n(1-p^{-1}) + o(n)$$

will follow immediately from the result on the corresponding edge extremal problem. Therefore we do not deal with this case. The behaviour of $\psi(n, L, t)$ is too complicated if $t > \chi(L)$; as we have mentioned, we cannot solve it even if $L = K_3$. Therefore we restrict our investigation to the case $t = \chi(L) = p$. But even in this case, (2) is almost always valid. The only exception is when

(∗) L contains an edge e such that $\chi(L-e) < \chi(L)$.

Such edges are called (colour-)critical and from now on we shall suppose that $\chi(L) = p$ and L satisfies (∗).

We shall prove that in this case the result obtained by Andràsfai, Erdös and Sòs remains valid.

Theorem 1. *Let $\chi(L) = p$ and L satisfy (∗). Then $\psi(n, L, p) \leqslant \psi(n, K_p, p)$ if n is large enough.*

Since

$$\psi(n, K_p, p)/n \approx 1 - 1/(p - \tfrac{4}{3}) < 1 - 1/(p - \tfrac{3}{2}).$$

Theorem 1 is an immediate consequence of

Theorem 2. *Let $\chi(L) = p$ and $L \neq K_p$. If L satisfies (∗), then*

$$(3) \qquad \widetilde{\psi}(n, L, p) = \max\{\sigma(G^n): L \not\subset G^n,\ K_p \subset G^n\}$$

$$\leqslant (1 - 1/(p - \tfrac{3}{2}))n + o(1).$$

Indeed, if G^n of Theorem 1 does not contain K_p, then $\sigma(G^n) \leqslant$

$\psi(n, K_p, p)$. If $G^n \supset K_p$, then Theorem 2 gives that

$$\sigma(G^n) \leqslant (1 - 1/(p - \tfrac{3}{2}))n + o(n) < \psi(n, K_p, p).$$

Hence Theorem 1 is really an easy consequence of Theorem 2.

Remark 3. One can prove, by much more complicated arguments, that

$$(3^*) \qquad \widetilde{\psi}(n, L, p) \leqslant (1 - 1/(p - \tfrac{3}{2}))n + O_L(1)$$

and this result cannot be improved since, (as we shall see) for every constant M, there exists a graph L such that

$$\widetilde{\psi}(n, L, p) \geqslant (1 - 1/(p - \tfrac{3}{2}))n + M.$$

2. Proof of Theorem 2

Let

$$q = 1 - 1/(p - \tfrac{3}{2}).$$

(A) First we give an example, showing that Theorem 2 cannot be improved. We fix an l and put $r = 2l + 1$. Let

$$T_r = K_2 \times K_{p-2}(r, \ldots, r).$$

This T_r will be the sample graph. Now we construct a graph

$$U^n = W^{6m+3l} \times K_{p-3}(4m + l, \ldots, 4m + l)$$

of

$$n = (p - 3)(4m + l) + (6m + 3l)$$

vertices containing K_p but not containing T_r. W^{6m+3l} is defined as follows (see Fig. 2).

For $i = 1, \ldots, 6$, $|A_i| = m$, for $i = 1, 2, 3$, $|B_i| = l$, and the 9 sets are pairwise disjoint. The indices are counted mod 6 and mod 3, respectively. Each vertex of A_i is joined to each vertex of A_{i+1}. Each vertex of B_i is

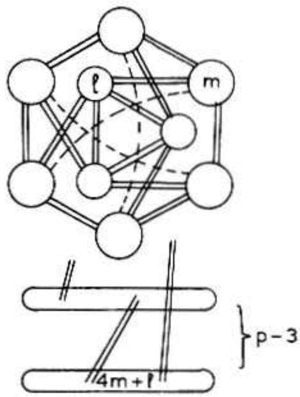

Fig. 2.

joined to each vertex of $A_i \cup A_{i+3}$ and to $B_{i-1} \cup B_{i+1}$. Finally, each vertex of A_i is joined to exactly l vertices of A_{i+3}. The minimum valence in U^n is

$$\sigma(U^n) = (p-3)(4m+l) + 2(l+m).$$

Therefore,

$$\sigma(U^n) = qn + 3l/(2p-3). \tag{4}$$

Trivially, $K_p \subset U^n$. On the other hand, T_r is a p-chromatic graph satisfying (*) and $T_r \not\subset U^n$.

$T_r = K_p(1, 1, r, ..., r)$ has p classes. At most $p-3$ classes can be contained by $K_{p-3}(4m+l, ..., 4m+l) \subset U^n$. Therefore, at least 3 classes of T_r are in W^{6m+3l}. Thus W^{6m+3l} has an edge with r triangles on it. But one can easily check that every edge of W^{6m+3l} is contained in at most $2l < r$ triangles. This proves (A) (see also Remark 3).

(B) We reduce the general case to the case of T_r showing that if L is the p-chromatic graph satisfying (*), then from $\sigma(G^n) \geq qn$ and $T_r \subset G^n$ follows $L \subset G^n$, if r and n are large enough. If we prove also

$$\widetilde{\psi}(n, T_r, p) \leq (q + o(1))n,$$

then, for every $\bar{q} > q$ and $n > n_0(\bar{q})$,

$$\sigma(G^n) \geq \bar{q}n$$

will imply that a G^n (containing K_p) must contain T_r and therefore L too. Thus it will be proved that

$$\widetilde{\psi}(n, L, p) \leqslant \overline{q}n$$

for every $\overline{q} > q$ and $n > n_0(\overline{q})$, i.e.,

$$\widetilde{\psi}(n, L, p) = (q + o(1))n.$$

(B_1) Let us suppose that $\sigma(G^n) > qn$ and

$$K_{p-2}(r, \ldots, r) \subset G^n.$$

The classes of $K_{p-2}(r, \ldots, r)$ will be denoted by $C_1, \ldots, C_{p-2}$. The method used here will be repeated later twice more and we shall refer to it as "estimation of the sum of valencies". This means that we consider those edges which join $K_{p-2}(r, \ldots, r)$ to $G^n - K_{p-2}(r, \ldots, r)$. Their number is at least

$$(p-2)r\,qn - O(1).$$

If x is the number of vertices joined to at least $(p-3+\delta)r$ vertices of $K_{p-2}(r, \ldots, r)$ (where $\delta > 0$ is a small constant, to be fixed later), then

$$(p-2)r\,qn - O(1) \leqslant (p-3+\delta)r(n-x) + (p-2)rx + O(1)$$

$$= (p-2)rn - (1-\delta)(n-x)r + O(1).$$

Hence

$$(1-\delta)n - (p-2)(1-q)n - O(1) \leqslant (1-\delta)x.$$

If δ is sufficiently small, then $x \geqslant c_0 n$ (where $c_0 > 0$ is a constant). But even the much weaker condition $x > r$ would imply (as we shall prove in B_2)) that there exist λ vertices outside of $K_{p-2}(r, \ldots, r)$ and λ vertices in each class of $K_{p-2}(r, \ldots, r)$ forming a $K_{p-1}(\lambda, \ldots, \lambda) \subset G^n$, where $\lambda \to \infty$, if $r \to \infty$.

Let the original $K_{p-2}(r, \ldots, r)$ be just the $K_{p-2}(r, \ldots, r)$ of $T_r \subset G^n$, then replacing 2 vertices of the λ new ones by the two vertices of T_r joined to each (other) vertex of T_r we obtain a $K_{p-1}(\lambda, \ldots, \lambda)$ and with an additional edge.

This graph will be denoted by $T((p-1)\lambda, (p-1), 1)$. One can easily prove that L satisfies (*) if and only if $L \subset T((p-1)\lambda, (p-1), 1)$ for $\lambda = v(L)$. Therefore, if r is large enough, $T_r \subset G^n$ and $G(G^n) \geqslant qn$ imply

$$L \subset T((p-1)\lambda, (p-1), 1) \subset G^n.$$

This proves the possibility of reduction to the case $L = T_r$.

(B_2) We have to prove that, if $x > r$, then λ vertices in each class of $K_{p-2}(r, ..., r)$ and λ vertices outside can be determined so that the graph spanned by them should contain $K_{p-1}(\lambda, ..., \lambda)$. One short but not too elementary proof of this fact is the following one: Let $\eta > 0$ be a small constant, depending on δ and fixed only later. We select ηr vertices from those joined to $K_{p-2}(r, ..., r)$ by at least $(p-3+\delta)r$ edges. Let G^* be a graph, the vertices of which are the considered $(p-2+\eta)r$ vertices and the edges of which join either two different classes of $K_{p-2}(r, ..., r)$ or a class of it to a vertex outside. An easy computation gives that if η is a fixed sufficiently small constant, then

$$\varliminf_{r \to \infty} e(G^*)/v(G^*)^2 > \tfrac{1}{2}(1 - 1/(p-2)).$$

Now we apply a theorem of Erdös and Stone [4] according to which, if

$$\lim_{v(\theta^*) \to \infty} e(G^*)/v(G^*)^2 > \tfrac{1}{2}(1 - 1/(\tau - 1)),$$

then, for every λ and $v(G^*) > n(\lambda)$, G^* contains $K_\tau(\lambda, ..., \lambda)$. In our case, $G^* \supset K_{p-1}(\lambda, ..., \lambda)$ and, since we did not consider the edges of G^n joining two vertices of the same class of $K_{p-2}(r, ..., r)$ or two vertices outside, there must be λ vertices outside and λ vertices in each class, forming a $K_{p-1}(\lambda, ..., \lambda)$.

(C) Now we prove Theorem 2 for $L = T_r$ by induction on p. The case $p = 3$ is trivial and is a special case of the proof below. Let us suppose that Theorem 2 is known already for $p-1$, and that $\bar{q} > q = 1 - 1/(p - \frac{3}{2})$,

$$\sigma(G^n) \geqslant \bar{q} n, \quad K_p \subset G^n.$$

We have to prove that $T_r \subset G^n$. Let a be a vertex of $K_p \subset G^n$ and let G^{qn} be a subgraph of G^n spanned by qn vertices of $N(a)$. We suppose also that $K_p - a = K_{p-1} \subset G^{qn}$. ([ ] is usually omitted!)

Since each vertex of G^{qn} is joined to at least

$$\bar{q}n-(1-q)n=(q+\bar{q}-1)n$$

vertices of G^{qn} and since

$$n(q+\bar{q}-1)>n(2q-1)=(1-1/((p-1)-\tfrac{3}{2}))\cdot qn,$$

we may apply the hypothesis to G^{qn} with $p-1$ and ν, obtaining a $K_2\times K_{p-3}(\nu,\ldots,\nu)\subset G^{qn}$. Hence

$$V_\nu=K_3\times K_{p-3}(\nu,\ldots,\nu)\subset G^n.$$

Here K_3 will be called the triangle of V_ν.

(D) We apply the method of "estimation of the sum of valencies" to K_3 of V_ν. Let X be the set and x be the number of vertices, joined to at least 2 vertices of the K_3 of V_ν.

(5) $$3\bar{q}n-3\leqslant(n-x+O(1))+3x=n+2x+O(1).$$

Thus

(6) $$x\geqslant\tfrac{1}{2}(3\bar{q}-1)n+O(1).$$

The method of (B_2) now gives that either X contains at most 3ν vertices joined to $\geqslant(p-4+\delta)\nu$ vertices of $K_{p-3}(\nu,\ldots,\nu)$ of V_ν or there exist r vertices in X joined to the same pair of vertices of the triangle of V_ν and r vertices in each class of $K_{p-3}(\nu,\ldots,\nu)$, determining together a $K_{p-2}(r,\ldots,r)$. If we add the edge of the triangle of V_ν, to which each considered vertex outside is joined by 2 edges, then we obtain a

$$K_2\times K_{p-2}(r,\ldots,r)=T_r\subset G^n.$$

In this case our proof is finished. In the other case, when at most 3ν vertices of X are joined to $K_{p-3}(\nu,\ldots,\nu)$ by $\geqslant(p-4+\delta)\nu$ edges, we shall obtain a contradiction by applying again the method of "estimation of the sum of valencies". Now we apply it to $K_{p-3}(\nu,\ldots,\nu)$:

$$\nu(p-3)\bar{q}n \leqslant (p-3)(n-x+O(1))\nu + (p-4+\delta)\nu(x-O(1))$$

$$= (p-3)n\nu - (1-\delta)x\nu + O(1).$$

This means that

$$(7) \qquad (1-\delta)x \leqslant (p-3)n(1-\bar{q}) + O(1).$$

(6) is a lower, (7) an upper bound for x. Comparing them we get

$$(8) \qquad \tau(\bar{q}) = 2(p-3)(1-\bar{q})/(3\bar{q}-1) > 1-\delta.$$

Here first $\bar{q} > q$, then δ (and then ν which does not occur in (8)) are fixed. But a trivial computation shows that $\tau(q) = 1$. Further, it is also trivial that $\tau(\bar{q})$ is a monotone decreasing function of $\bar{q}$, hence $\tau(\bar{q}) < 1$. Therefore, if δ is small enough (what can be assumed), then (8) gives the contradiction.

3. The lower estimation of $\psi(n, K_3, t)$

In this section, we give an example of a graph G^n which does not contain K_3, is p-chromatic and $G(G^n) = \frac{1}{3}n + o(n)$.

Kneser conjectured [6] that the following graph is $l+2$-chromatic:

For a given m, we consider the $\binom{2m+l}{m}$ m-tuples of a given set of $2m+l$ elements. These are the vertices of our graph. Two m-tuples are joined if and only if their intersection is empty.

Szemerèdi obtained some lower bounds for the chromatic number of this graph. We shall need the simplest case of Szemerèdi's (unpublished) results.

Lemma 4. *Let $c > 0$ be a given small constant. For $l = cm$ and $m \to \infty$ the chromatic number of the Kneser-graph tends to infinity.*

Proof (Szemerèdi). Let us suppose that the n-tuples of $2m + l = N$ elements can be divided into t classes so that all sets belonging to the same class always have common elements. (This is equivalent to the assertion that the Kneser-graph is $\leqslant t$-chromatic.) We add a subset of the

N elements to the i^{th} class if this subset contains an m-tuple in the i^{th} class. According to a result of Kleitman [5], the number of these subsets is at most $2^N - 2^{N-t}$. Thus at least 2^{N-t} subsets of the N elements do not belong to any class. We know that exactly

$$\binom{N}{m-1} + \binom{N}{m-2} + \ldots + \binom{N}{1}$$

subsets do not belong to any class. Therefore

$$\sum_{k<m} \binom{N}{k} \geq 2^{N-t}. \tag{9}$$

It is a well-known fact that

$$\sum_{k<N/(2+c)} \binom{N}{k} = o(2^N). \tag{10}$$

Therefore $t \to \infty$. (To prove (10) we can apply the Tschebitshev inequality.)

Let us now consider the following graph. First we fix P and then $c > 0$. If m is large enough and $l \approx cm$, then the Kneser-graph of $\gamma = \binom{2m+l}{m}$ vertices will be $\geq p$-chromatic. Let the set of $2m + l$ elements be just $\{1, 2, \ldots, 2m + l\}$ and the subsets be $S_1, S_2, \ldots, S_\gamma$. Let $x_1, \ldots, x_h$ and $y_{i,j}$, $i = 1, 2, \ldots, 2m + l$, $j = 1, 2, \ldots, h/m$ be new vertices. (For the sake of simplicity we suppose that h is a multiple of m.) Let us join the set S_k (which is a vertex of our graph) to $y_{i,j}$ if $i \in S_k$. Clearly, each S_k is joined to h vertices, i.e., has the valence h. Each x_t and $y_{i,j}$ are joined, therefore $\sigma(x_t) > 2h$, $\sigma(y_{i,j}) > h$. If now n is the number of vertices in this graph G^n, then $\sigma(G^n) \approx n/(3 + c)$. Further, $\chi(G^n) \geq p$. It is not too hard to show that $K_3 \not\subset G^n$. Thus

$$\psi(n, K_3, t) \geq n/(3 + c).$$

Since c was an arbitrary positive constant,

$$\psi(n, K_3, t) \geq (\tfrac{1}{3} + o(1))n.$$

The construction can be modified to obtain this lower bound for every large n.

4. Open problems

We have already mentioned that we could not prove or disprove that $\psi(n, K_3, t) \approx \frac{1}{3}n$ if $t \geqslant 4$. Another problem, which we could not solve, is: whether there exists a sequence $\epsilon_t \to 0$ (if $t \to \infty$) such that

$$\max\{\sigma(G^n) : C^5 \not\subset G^n, \chi(G^n) \geqslant t\} \leqslant \epsilon_t n,$$

where C^5 is the pentagon.

References

[1] B. Andràsfai, P. Erdös and V.T. Sòs, On the connection between chromatic number, maximal clique and minimal degree of a graph, Discrete Mathematics, to appear.
[2] P. Erdös, Some recent results on extremal problems in graph theory, in: Theory of graphs; International Symp. Rome (1966) 118–123.
[3] P. Erdös, On some new inequalities concerning extremal properties of graphs, in: Theory of graphs, Proc. Coll. Tihany (1966) 77–81.
[4] P. Erdös and A.M. Stone, On the structure of linear graphs, Bull. Am. Math. Soc. 52 (1946) 1089–1091.
[5] D. Kleitman, Families of non-disjoint subsets, J. Combin. Theory 1 (1966) 153–155.
[6] A. Knéser, Ausgabe 360, Jber. Deutsch. Math. Verein. 58 (2) (1955).
[7] M. Simonovits, A method for solving extremal problems in graph theory, in: Theory of graphs, Proc. Coll. Tihany (1966) 279–319.
[8] P. Turán, On an extremal problem in graph theory, Mat. Fiz. Lapok 48 (1941) 436–452 (in Hungarian).