

# COMPACTNESS RESULTS IN EXTREMAL GRAPH THEORY

P. ERDŐS and M. SIMONOVITS

Dedicated to Tibor Gallai on his seventieth birthday

Received 15 April 1982

Let  $\mathbf{L}$  be a given family of so called prohibited graphs. Let  $\text{ex}(n, \mathbf{L})$  denote the maximum number of edges a simple graph of order  $n$  can have without containing subgraphs from  $\mathbf{L}$ . A typical extremal graph problem is to determine  $\text{ex}(n, \mathbf{L})$ , or at least, find good bounds on it. Results asserting that for a given  $\mathbf{L}$  there exists a much smaller  $\mathbf{L}^* \subseteq \mathbf{L}$  for which

$$\text{ex}(n, \mathbf{L}) \approx \text{ex}(n, \mathbf{L}^*)$$

will be called *compactness* results. The main purpose of this paper is to prove some compactness results for the case when  $\mathbf{L}$  consists of cycles. One of our main tools will be finding lower bounds on the number of paths  $P^{k+1}$  in a graph of  $n$  vertices and  $E$  edges., witch is, in fact, a "super-saturated" version of a wellknown theorem of Erdős and Gallai.

## Notations

Below we shall consider only graphs without loops and multiple edges. For a graph  $G$ , let  $e(G)$ ,  $v(G)$ , and  $\chi(G)$  denote the number of edges, vertices and the chromatic number, respectively. The first upper index always denotes the number of vertices, e.g.,  $T^{n,d}$ ,  $G^n$ ,  $S^n$  are graphs of order  $n$ .  $C^t$ ,  $P^t$ , and  $K_t$  denote the cycle, path and complete graph of  $t$  vertices. The degree of  $x$  in  $G^n$  is  $d_G(x)$ .

Given a family of so called prohibited graphs,  $\mathbf{L}$ , then  $\text{ex}(n, \mathbf{L})$  denotes the maximum number of edges a graph  $G^n$  can have without containing subgraphs from  $\mathbf{L}$ . (If  $\mathbf{L}$  consists of the sole graph  $L$ , we shall use the simpler notation  $\text{ex}(n, L)$ .)

## Introduction

The well known theorem of P. Turán [16] asserts that among the graphs  $G^n$  not containing  $K_{p+1}$  there exists one having more edges than all the others and this graph  $T^{n,p}$  can be defined as follows.

We partition  $n$  vertices into  $p$  classes  $C_1, \dots, C_p$  as equally as possible:  $|C_i| = \left\lfloor \frac{n}{p} \right\rfloor$  or  $\left\lceil \frac{n}{p} \right\rceil$ . Two vertices of  $T^{n,p}$  are joined iff they belong to different classes.

Generally one can replace  $K_{p+1}$  by an arbitrary finite or infinite class  $\mathbf{L}$  and ask:

What is the maximum number of edges,  $\text{ex}(n, \mathbf{L})$ , a graph  $G^n$  can have without containing subgraphs from  $\mathbf{L}$ ? It is somewhat surprising, [9], that the maximum depends very loosely on  $\mathbf{L}$ . If

$$(1) \quad p = p(\mathbf{L}) = \min_{L \in \mathbf{L}} \chi(L) - 1,$$

then

$$(2) \quad \text{ex}(n, \mathbf{L}) = \left(1 - \frac{1}{p} + o(1)\right) \binom{n}{2}.$$

This shows that there is a sharp difference between the cases when  $\mathbf{L}$  contains bipartite graphs (and therefore  $p=1$ ,  $\text{ex}(n, \mathbf{L})=o(n^2)$ ), and the other cases. The case, when  $p=1$ , will be called *degenerate*. Another interesting corollary of (2) is, that if  $L^* \in \mathbf{L}$  is of the minimum chromatic number ( $=p+1$ ), then

$$(3) \quad \text{ex}(n, \mathbf{L}) = \text{ex}(n, L^*) + o(n^2).$$

If the extremal problem is non-degenerate, then (3) implies that

$$(4) \quad \frac{\text{ex}(n, L^*)}{\text{ex}(n, \mathbf{L})} \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

Unfortunately, (3) does not imply (4) for degenerate extremal graph problems. We shall call results asserting that for some  $\mathbf{L}$  there is a "much smaller"  $\mathbf{L}^* \subseteq \mathbf{L}$  for which

$$\frac{\text{ex}(n, \mathbf{L}^*)}{\text{ex}(n, \mathbf{L})} \rightarrow 1 \quad \text{as } n \rightarrow \infty$$

*compactness theorems*.

Compactness theorems can be proved not only for graphs but digraphs as well, [4]. Here we restrict ourselves to ordinary graphs. As we have seen, for  $p \geq 2$  the compactness problem has a trivial answer. This is why we are interested primarily in compactness results connected with *degenerate* problems. The field of degenerate extremal problems is very wide and interesting with plenty of unsolved problems. Just to indicate the line of these problems we formulate two conjectures.

**Conjecture 1.** For every finite  $\mathbf{L}$  (containing bipartite graphs as well) there exists an  $L^* \in \mathbf{L}$  for which

$$(5) \quad \text{ex}(n, \mathbf{L}) = O(\text{ex}(n, L^*)).$$

**Conjecture 2.** For every finite  $\mathbf{L}$  containing bipartite graphs there exists a (probably rational) constant  $c = c_{\mathbf{L}} \geq 1$  for which

$$(6) \quad \frac{\text{ex}(n, \mathbf{L})}{n^c} \text{ converges to a positive limit as } n \rightarrow \infty.$$

**Remark.** For infinite families the above conjectures do not hold. If e.g.,  $\mathbf{L}$  consists of all the cycles, then  $\text{ex}(n, \mathbf{L}) = n - 1$ , but by a (random graph) result of Erdős [7], for

every finite  $L^* \subseteq L$ ,  $\text{ex}(n, L^*) > \bar{c}(L^*) \cdot n^{1+c(L^*)}$  for some  $c(L^*) > 0$ . This disproves Conjecture 1 for infinite families; a slightly more complicated example disproves Conjecture 2 for infinite  $L$ , as well.

One version of the extremal graph problems usually considered is, when we restrict our considerations to bipartite graphs. In other words, if  $C^*$  denotes the family of all odd cycles, we wish to determine  $\text{ex}(n, L \cup C^*)$  instead of  $\text{ex}(n, L)$ . Clearly, if  $p(L) > 2$ , then

$$\text{ex}(n, L \cup C^*) = \text{ex}(n, C^*) = \left\lfloor \frac{n^2}{4} \right\rfloor.$$

On the other hand, if  $L$  contains bipartite graphs, then

$$(7) \quad \frac{1}{2} \text{ex}(n, L) \leq \text{ex}(n, L \cup C^*) \leq \text{ex}(n, L).$$

Indeed, the right side of (7) is trivial. The left side follows from the fact that every  $G^n$  contains a bipartite  $H^n$  with maximum number of edges and for this  $H^n$   $d_H(x) \cong \frac{1}{2} d_G(x)$  for every vertex. Hence  $e(H^n) \cong \frac{1}{2} e(G^n)$ , proving (7).

**Conjecture 3.** For every finite  $L$  there exists a  $t$  such that

$$\frac{\text{ex}(n, L \cup C^*)}{\text{ex}(n, L \cup \{C^3, C^5, \dots, C^{2t+1}\})} \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

This conjecture asserts that the exclusion of sufficiently many odd cycles has the same effect as excluding them all.

### Cycles in graphs

There are many interesting results on the extremal graph problems of cycles.

If  $L$  contains no bipartite graphs but it contains an (odd) cycle, then  $\text{ex}(n, L) = \left\lfloor \frac{n^2}{4} \right\rfloor$  for  $n \geq n_0(L)$ . This case will not be considered below. The question of the even cycles is a much more intriguing one. An old number theoretic result of Erdős and E. Klein [6], asserts in a graph theoretic language that if  $C^*$  denotes the family of odd cycles, then

$$(8) \quad \text{ex}(n, C^* \cup \{C^4\}) = \left(\frac{n}{2}\right)^{3/2} + o(n^{3/2}).$$

Reiman [14] proved that the construction of [6] is sharp for infinitely many values of  $n$ . Later Kővári, T. Sós and Turán [13], Erdős, Rényi and V. T. Sós [8], W. G. Brown [3] proved that

$$(9) \quad \text{ex}(n, C^4) = \frac{n}{2}^{3/2} + o(n^{3/2}).$$

These two results show that the exclusion of the odd cycles diminishes the maximum

number of edges by a factor of  $\sqrt{2}$ . One can conjecture that the exclusion of any odd cycle does the same:

**Conjecture 4.** For any  $k$  and  $t \geq 2$

$$(10) \quad \text{ex}(n, \{C^{2k}, C^{2t-1}\}) = \left(\frac{n}{2}\right)^{1+\frac{1}{k}} + o\left(n^{1+\frac{1}{k}}\right).$$

Further,

**Conjecture 5.**

$$\text{ex}(n, C^{2k}) = \frac{n^{1+\frac{1}{k}}}{2} + o\left(n^{1+\frac{1}{k}}\right).$$

The background of the above conjectures is the following. Many people investigated the so called cage problem: for given  $r$  and  $g$  find the minimum number  $n(r, g)$  of vertices for which there exists an  $r$ -regular  $G^n$  with this girth  $g$  and order  $n = n(r, g)$ . Singleton [15] and Benson [1] constructed infinite families of  $r$ -regular graphs  $G^n$  of girth  $2k+2$  for  $r \approx \sqrt[k]{\frac{n}{2}}$  for  $k=2, 3, 5$ . These are bipartite graphs with  $\frac{n}{2}$  vertices in each colour-class and for any vertex  $x$  the number of vertices  $y$  of distance  $s$  from  $x$  is  $\left(\frac{n}{2}\right)^{s/k}$  ( $s=1, 2, \dots, k$ ). Conjecture 4 asserts on the one hand that if we drop

the condition of regularity, the average degree remains still  $\sqrt[k]{\frac{n}{2}}$ , and even more, if we exclude only the longest even cycle ( $=C^{2k}$ ) and any odd cycle, we still get the same upper bound. Further, Conjecture 4 asserts that this bound is sharp for every  $k$ . We cannot prove but the much weaker results stated below.

**Theorem 1.** Let  $k$  be a natural number. Then

$$(11) \quad \text{ex}(n, \{C^3, \dots, C^{2k}, C^{2k+1}\}) \leq \left(\frac{n}{2}\right)^{1+\frac{1}{k}} + 2^k \cdot \left(\frac{n}{2}\right)^{1-\frac{1}{k}}.$$

**Theorem 2.**

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

Conjecture 5 asserts that if we exclude only  $C^{2k}$  (and probably the same holds if we exclude all the shorter even cycles as well), then  $G^n$  may have significantly more edges: the number of vertices of distance  $s$  from  $x$  can increase as  $n^{s/k}$  (instead of  $\left(\frac{n}{2}\right)^{s/k}$ ). This is known for  $k=2$ , [3], [8], but we cannot prove it even for  $k=3$ .

**Theorem 3.** If  $\text{ex}(n, C^{2n}) \geq cn^{1+\frac{1}{k}}$ , then there is a  $t$  such that

$$\lim_{n \rightarrow \infty} \frac{\text{ex}(n, \{C^{2k}, C^3, C^5, C^7, \dots\})}{\text{ex}(n, \{C^{2k}, C^3, \dots, C^{2t-1}\})} = 1.$$

There is a wider known class of graphs to which Theorem 3 generalizes. To formulate a more general theorem we need a definition.

**Definition.** (cf [12]). If  $T$  is a connected bipartite graph, coloured by "blue" and "red", then  $L_k = L_k(T, c)$  denotes the graph obtained from  $T$  by joining a vertex  $x$  outside of  $T$  by *disjoint* paths of length  $k-1$  to each blue vertex of  $T$ , where

- (a)  $c$  is the fixed colouring,  
 (b) "disjoint" means that any two of these paths have only  $x$  in common, and only their other endvertices belong to  $T$ . Some illustrations are given on figure 1.

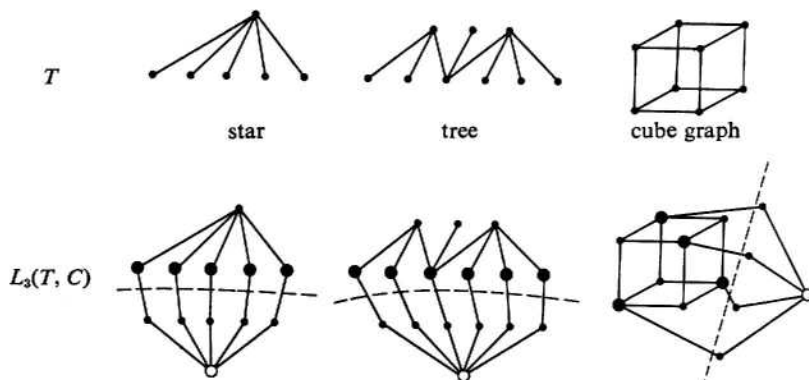


Fig. 1

In [12] Faudree and Simonovits proved that if  $T$  is a tree and  $L = L_k(T, c)$ , then  $\text{ex}(n, L) = O(n^{1+\frac{1}{k}})$ . Obviously, this generalizes the corresponding theorem on  $C^{2k}$ .

**Theorem 3\*.** If  $\text{ex}(n, L) \cong cn^{1+\frac{1}{k}}$  for some tree  $T$  and colouring  $c$  and  $L := L_k(T, c)$ , then there is a  $t$  for which

$$\lim_{n \rightarrow \infty} \frac{\text{ex}(n, \{L, C^3, C^5, \dots\})}{\text{ex}(n, \{L, C^3, C^5, \dots, C^{2t-1}\})} = 1.$$

### Walks in graphs

In many degenerate (and nondegenerate) extremal graph problems we prove that  $G^n$  contains some prohibited  $L \in \mathbf{L}$  by choosing another family  $\mathbf{P}$  and proving that  $G^n$  contains very many  $P \in \mathbf{P}$  under the conditions considered. Assume we have a family  $\mathbf{P}$  of prohibited graphs and  $e(G^n) = \text{ex}(n, \mathbf{P}) + k$  for some  $k > 0$ . Clearly, in such cases  $G^n$  will have prohibited subgraphs  $P \in \mathbf{P}$ . What is somewhat surprising,  $G^n$  will have extremely many prohibited subgraphs. Theorems of this type will be called *supersaturated extremal graph problems* and we shall devote a whole paper to their study, [10]. Many such result have already been proved by various authors, for

$P = \{K_p\}$ . Here we are interested in the number of paths  $P^k$  in a graph  $G^n$ , since lower bounds on it will be the basic tools in some of our proofs.

Let us consider a regular graph  $G^n$  of degree  $d$ . If we call a sequence of vertices  $(x_0, x_1, \dots, x_k)$  a walk of length  $k$  assumed that  $x_{i-1}$  and  $x_i$  are joined by an edge, and denote such a walk by  $W^{k+1}$ , then it is much easier to count the number of walks  $W^{k+1}$ , since we do not have to care with the repetitions.

**Notation.**  $w_k(G^n)$  denotes the number of  $W^{k+1}$ 's in  $G^n$ , divided by  $n$ , where  $(x_0, x_1, \dots, x_k)$  and  $(x_k, x_{k-1}, \dots, x_1, x_0)$  are generally counted different.

Obviously, in a  $d$ -regular graph  $G^n$

$$(12) \quad w_k(G^n) = d^k.$$

**Remark.** The word "generally" above means that e.g. the triangle  $(x_0, x_1, x_2, x_0)$  yields two walks, however, the  $(x_0, x_1, x_2, x_1, x_0)$  is only one walk.

The following conjecture asserts that in some sense the regular graphs are the minimal.

**Conjecture 6.** If  $d$  is the average degree in  $G^n$ , i.e.,  $d = \frac{2e(G^n)}{n}$  then

$$(13) \quad w_k(G^n) \cong d^k,$$

further, if  $k \cong t$ , and both  $t$  and  $k$  are odd, then

$$(14) \quad w_k(G^n)^t \cong w_t(G^n)^k.$$

Observe that for  $t=1$  (14) reduces to (13),  $w_1 = d$ .

**Theorem 4.** If  $k$  is even, then (14) holds.<sup>1</sup>

**Remarks.** In the conjecture and theorem above the case "t ≅ k, t is even, k is odd" is left out. In this case (14) does not necessarily hold: a trivial counterexample is  $K_2(a, b)$  if  $a+b=n$  and  $a$  is much smaller than  $b$ . If  $A$  denotes the adjacency matrix of  $G^n$ , then  $w_k(G^n)$  obviously equals the sum of entries of  $A^k$ . Thus the above results and conjectures have purely matrix-theoretical formulations as well. We have recently discovered that (13) was proved for every  $k \cong 1$  in [11], [17] and [18].

There is another case when (13) is proved, [11]:

**Faudree—McKay theorem.** If  $k=3^p$ , then  $w_k(G^n) \cong d^k$ . ■

We shall also prove the following generalization of (13):

**Theorem 4\*.** Let  $f(x) = x^k + a_1x^{k-1} + a_2x^{k-2} + \dots + a_k$ ,  $k$  be even and

$$F(G^n) = w_k(G^n) + a_1w_{k-1}(G^n) + \dots + a_kw_0(G^n).$$

If  $f$  is a convex increasing function for  $x \cong x^*$  and  $m$  denotes its minimum, in  $(-\infty, \infty)$  then for  $d = d(G^n) \cong x^*$

$$F(G^n) \cong f(d(G^n)) + m - f(x^*).$$

<sup>1</sup> The proof given here is due to C. D. Godsil. (Cf. "Added in proof".)

One important case of this theorem is, when  $f$  has  $k$  real roots and  $x^*$  is the maximum of them. Then  $f(x^*)=0$ , of course. In the applications of this theorem the coefficients of  $f$  depend on  $n$ .

Our next proposition shows that (13) is valid apart from the exact value of the multiplicative constant, even if  $k$  is odd.

**Proposition 1.** For every  $k$ ,  $w_k(G^n) \cong 4^{-k} \cdot d(G^n)^k$ .

### Paths in graphs

The next two assertions show that in some sense the difference of the number of walks and paths is negligible.

**Proposition 2.** Let  $q < 1 + \frac{1}{k-1}$  and assume that the maximum degree  $\bar{d}(G^n) \cong \leq d(G^n)^q$ . If  $w_k^*(G^n)$  denotes the number of walks  $W^{k+1}$  in  $G^n$  which are not paths, then

$$(15) \quad w_k^*(G^n) = o(n \cdot w_k(G^n)), \quad \text{if } d(G^n) \rightarrow \infty.$$

Observe that the total number of walks is  $n \cdot w_k(G^n)$ : Proposition 2 asserts that almost all walks are paths if the maximum degree is not too large. An easy example shows that without having bounds on the maximum degree we cannot guarantee (15):

Let e.g.  $G^n$  be the graph obtained from an  $r$ -regular graph  $H^n$  by partitioning the vertices into the classes  $A_1, \dots, A_m$ , choosing an  $a_i$  from each  $A_i$  and joining this  $a_i$  to all other vertices of *this*  $A_i$ . If e.g.  $r \approx \log n$  and  $|A_i| \approx n/\log n$ , then we get the desired counterexample.

**Theorem 5.** If  $f(n, d)$  is the minimum number of walks  $W^{k+1}$  a graph  $G^n$  can have with average degree  $d$ , then every graph of order  $n$  and average degree  $d$  contains at least  $(1/2) \cdot f(n, d) - o(f(n, d))$  paths  $P^{k+1}$ , as  $d \rightarrow \infty$ .

Here the factor  $1/2$  compensates that each path yields two walks.

### Proofs

**Proof of Theorem 4.** It is known, [5], that if  $A$  is the adjacency matrix of the graph  $G^n$  and  $\lambda_1, \dots, \lambda_n$  are the eigenvalues, where a  $\lambda$  of multiplicity  $\mu$  is listed  $\mu$  times, then

$$(16) \quad w_m(G^n) = \sum_i c_i \lambda_i^m$$

for some constants  $c_i \geq 0$  depending only on  $G^n$ . Putting  $m=0$  we get, by  $w_0=1$ , that

$$(17) \quad \sum_i c_i = 1.$$

Since  $w_1=d$  is just the average degree, we get that

$$(18) \quad w_k(G^n) = \sum_i c_i \lambda_i^k \cong \left(\sum_i c_i \lambda_i\right)^k = w_1(G^n)^k = d^k,$$

by Jensen's inequality, since  $k$  is even. More generally, if  $t < k$  and  $k$  is even, put  $r:=k/t$ . By the convexity of  $x^r$ ,

$$w_k(G^n) = \sum_i c_i \lambda_i^k = \sum_i c_i |\lambda_i|^r \cong \left(\sum_i c_i |\lambda_i|\right)^r \cong w_t(G^n)^r. \quad \blacksquare$$

**Proof of Theorem 4\*.** Clearly, by (16), putting  $a_0=1$ , we have

$$F(G^n) = \sum_j \sum_i a_j c_i \lambda_i^{k-j} = \sum_i c_i f(\lambda_i).$$

Let

$$g(x) = \begin{cases} f(x) & \text{if } x \cong x^* \\ f(x^*) & \text{if } x < x^*. \end{cases}$$

Then  $g$  is convex in  $(-\infty, \infty)$  and

$$f(x) = g(x) + (f(x) - g(x)) \cong g(x) + m - f(x^*).$$

(Here  $m$  is mostly (?) negative.) Thus, by (17),

$$F(G^n) = \sum_i c_i f(\lambda_i) \cong \sum_i c_i g(\lambda_i) + (m - f(x^*)) \cdot \sum_i c_i \cong g\left(\sum_i c_i \lambda_i\right) + m - f(x^*).$$

Since  $\sum_i c_i \lambda_i = d$ , we are home.  $\blacksquare$

**Proof of Proposition 1.** We use induction on  $n$ . For  $n=1, 2, 3$  there is nothing to prove. Assume that we know the assertion for  $n-1$  and have a graph  $G^n$  with  $E$  edges. If  $G^n$  has a vertex  $y$  of degree  $< E/2n$ , let  $G^{n-1} := G^n - y$ . Clearly,  $e(G^{n-1}) > E - E/2n = E \cdot (n - \frac{1}{2})/n$ . Thus  $G^{n-1}$  contains at least

$$\begin{aligned} \frac{E^k \left(\frac{n-(1/2)}{n}\right)^k}{2^k (n-1)^{k-1}} &\cong \frac{E^k}{2^k n^{k-1}} \cdot \frac{n-1}{n} \cdot \left(\frac{n-(1/2)}{n-1}\right)^k \cong \\ &\frac{E^k}{2^k n^{k-1}} \left(\frac{n-(1/2)}{n-1}\right)^2 \cdot \frac{n-1}{n} \cong \frac{E^k}{2^k n^{k-1}} \cdot \frac{n^2 - n + (1/4)}{n^2 - n} \end{aligned}$$

walks  $W^{k+1}$ , proving the assertion.

The other case is, when all the vertices of  $G^n$  have degree  $\cong E/2n$ . Then it is trivial, that for any fixed  $y$  the number of walks  $W^{k+1}$  starting from  $y$  is at least  $(E/2n)^k$ .  $\blacksquare$

**Proof of Proposition 2.** We know from Proposition 1 that  $w_k(G^n) \cong 4^{-k} \cdot d^k$ . Let us count the number of walks  $W^{k+1}$  in  $G^n$  which are not paths. For each of them we may fix a  $j$  such that if  $(x_0, \dots, x_k)$  is the walk considered, then  $x_j = x_t$  for some  $t < j$ . For given  $x_0$  and  $j$  the walks  $(x_0, \dots, x_j)$  can be chosen in at most  $j \bar{d}^{(j-1)}$  ways. Thus the walks  $(x_0, \dots, x_k)$  can be chosen in at most  $j \cdot \bar{d}^{(j-1)} \cdot \bar{d}^{(k-j)} = j \cdot \bar{d}^{(k-1)}$  ways. By the assumption  $\bar{d} < d^q$ , (15) immediately follows.  $\blacksquare$



**Proof of Theorem 5.** Consider a graph  $U^n$  with average degree  $d(U^n)=d$  and maximum degree  $\bar{d}(U^n)\leq d+1$ . Clearly, the number of walks  $W^{k+1}\subseteq U^n$  is at most  $n \cdot (d+1)^k$ . Hence

$$(19) \quad f(n, d) \leq n \cdot (d+1)^k.$$

Now, let  $G^n$  be an arbitrary graph with  $d(G^n)=d$ . We wish to prove Theorem 5 by reducing it to the case when there are no high degrees. Fix two functions  $a(d)$  and  $b(d)$  tending very slowly to  $\infty$  as  $d \rightarrow \infty$  and satisfying  $b(d) \ll a(d)$ . E.g.  $a(d)=\log d$  and  $b(d)=\log \log d$  will do.

Now we delete all vertices of low and high degrees as follows.  $A_1, \dots, A_k$  are sets of vertices defined recursively.  $A_1$  is the set of vertices of  $G^n$  of degree  $\leq d/b(d)$ .  $A_j$  is the set of vertices of  $G^n - A_1 - A_2 - \dots - A_{j-1} = G_j$  having degree  $\leq d/b(d)$  in  $G_j$ , ( $j=2, 3, \dots, k$ ). Let  $Q$  be the set of vertices of degree  $\geq d \cdot a(d)$  in  $G^n - \cup A_j = H$ .

If  $Q=(x_1, \dots, x_m)$  and  $d_H(x_j)=d_j$ , then we have at least  $2 \sum_j \binom{d_j}{2}$  walks of form  $ux_jv$  and each can be continued in at least  $\left(\frac{d}{b(d)} - k\right)^{k-2}$  ways into a path  $(u, x_j, v, v_1, \dots, v_{k-2})$ : indeed, by definition, each  $v_i \in G^n - A_1 - A_2 - \dots - A_{k-i}$  is joined to at least  $d/b(d)$  vertices of  $G^n - A_1 - \dots - A_{k-i-1}$  and less than  $k$  of them are excluded because they have already been chosen in  $(u, x_j, v, \dots, v_i)$ , (see Fig. 2). If  $D=(\sum d_j)/m$ , then, using Jensen's inequality, we get at least

$$(20) \quad \left(\frac{d}{b(d)} - k\right)^{k-2} \sum_{j=1}^m (d_j^2 - d_j) \geq \left(\frac{d}{b(d)} - k\right)^{k-2} \cdot m(D^2 - D)$$

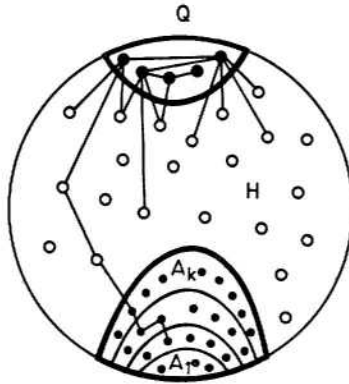


Fig. 2

paths  $P^{k+1}$  in  $G^n$ . (If  $k=2$ , then every term in (20) should be divided by 2, but that will cause no problems.) There are two possibilities: either (20) yields more paths than  $f(n, d)$  (estimated from above by (19)), and we are home, or else by  $m \leq n/a(d)$

$$(21) \quad \sum d_j \leq c_k \cdot E \cdot \sqrt{\frac{b(d)^{k-2}}{a(d)}} = o(E).$$

Delete all edges incident to  $\cup A_i \cup Q$ . Thus we get a new graph  $G^*$  with  $E^* = E - o(E)$  edges. Here we used (21) and that  $e(G_{j-1}) - e(G_j) \cong |A_j| \cdot d/b(d)$ . Applying Proposition 2 to this  $G^*$  we get that for some  $d^* = d - o(d)$

$$(22) \quad p_k(G^n) \cong p_k(G^*) = f(n, d^*) - o(f(n, d^*)).$$

This is slightly weaker than what we stated.

To get the desired result we shall try to put back the edges so that in the resulting  $H^n$

- (a) each vertex is incident to at most  $O(d/\sqrt[3]{a(d)})$  new edges.
- (b) the number of new  $P^{k+1}$ 's is  $o(f(n, d))$ .

If we succeed, then, applying Proposition 2 to  $H^n$  we get that it contains at least  $(1/2)f(n, d) - o(f(n, d))$   $P^{k+1}$ 's. Hence, by (b),  $G^*$  (and therefore  $G^n$  as well) contains at least  $(1/2)f(n, d) - o(f(n, d))$   $P^{k+1}$ 's. This will complete the proof.

Let  $S$  be the set of vertices  $v$  of  $G^*$  such that, for every  $p \leq k$ ,  $v$  is the endpoint of at most  $d^p \cdot b(d)^{k+1}$  paths  $P^{p+1}$  in  $G^*$ . ( $S$  includes  $\cup A_i$  and  $Q$ .) We shall distinguish whether

- (i)  $|S| < n - \frac{n}{b(d)}$  or
- (ii)  $|S| \cong n - \frac{n}{b(d)}$ .

In case (ii) we put back the edges so that each new edge joins vertices of  $S$ . One can easily achieve that (a) also be satisfied. Indeed, by Proposition 2,  $d(G^n) = O(n^{1-c})$  can be assumed; then we use a greedy algorithm to put back the edges.

Now, a new path  $P^{k+1}$  consists of old paths and new edges, and choosing a new edge at a given vertex  $v$  we have at most  $O(d/\sqrt[3]{a(d)})$  possibilities, choosing an old path  $P^{p+1}$  we have at most  $d^p b(d)^{k+1}$  possibilities. Therefore the number of new paths  $P^{k+1}$  is at most  $O(nd^k/\sqrt[4]{a(d)}) = o(f(n, d))$ , by Proposition 1.

In case (i) for at least one  $p \leq k$  we have at least  $n/p \cdot b(d)$  vertices in  $G^*$  emanating at least  $d^p b(d)^{k+1}$   $P^{p+1}$ 's. Repeating the argument above we may extend these paths into at least  $(1/2)n \cdot d^p \cdot b(d)^{k+1} \cdot (d/b(d) - k)^{k-p} \cong nd^k \cdot b(d)/4$  paths  $P^{k+1}$ . ■

**Proof of Theorem 1.** Let us call a walk  $W^{k+2}$  "good" if it is not contained in any path  $P^k$ . Observe that if  $x_0$  and  $y_0$  are joined by two good walks in  $G$ , then  $G$  contains a cycle of length  $\cong 2(k+1)$ . Assume that  $c = 2^k$  and

$$(23) \quad e(G^n) \cong \left(\frac{n}{2}\right)^{1+\frac{1}{k}} + c \left(\frac{n}{2}\right)^{1-\frac{1}{k}}.$$

The number of good walks  $W^{k+2} \subseteq G^n$  is, neglecting  $-m$  in Theorem 4\*, at least

$$(24) \quad n(w_{k+1}(G^n) - cw_{k-1}(G^n)) \geq n(d^{k+1} - cd^{k-1}) \geq n \cdot \frac{2e(G^n)}{n} (d^k - cd^{k-2}) \\ \geq 2e(G^n) \left\{ \frac{n}{2} + c(k-1) \left(\frac{n}{2}\right)^{1-2/k} \right\},$$

since the number of walks  $W^{k+2} \subseteq P^k$  and having the same first vertex  $x_0$  is at most  $2^k$ , and for  $x \geq 1$  and  $k \geq 2$

$$\left(x + \frac{c}{x}\right)^k - c\left(x + \frac{c}{x}\right)^{k-2} \geq x^k + c(k-1)x^{k-2}.$$

By (23) and (24) there is an edge  $(x_0, y_0)$  contained in at least  $n + 2c(k-1)(n/2)^{1-2/k}$  good walks  $(x_0, y_0, \dots, u)$  or  $(y_0, x_0, \dots, u)$ . Thus there must be at least two such walks with the same other endvertex  $u = u_k$ . Denote them by  $W$  and  $\tilde{W}$ .

There are two subcases:

Case (a).  $W = (x_0, y_0, u_3, \dots, u_k)$  and  $\tilde{W} = (x_0, y_0, v_3, \dots, u_k)$  or  
 Case (b).  $W = (x_0, y_0, u_3, \dots, u_k)$  and  $\tilde{W} = (y_0, x_0, v_3, \dots, u_k)$ . In the first case  $G^n$  contains two good  $W^{k+1}$  joining  $y_0$  to  $u_k$ . Thus  $G^n \supseteq C^j$  for some  $j \geq 2k$ . In the second case the union of  $W$  and  $\tilde{W}$  contains a  $C^{2j+1} \subseteq G^n$  for some  $j \leq k$ . ■

**Proof of Theorem 2.** We use the argument above, however, since  $C^3 \subseteq G^n$  is allowed, we have to be more careful in Case (b). Now we have two walks  $(x_0, y_0, a, u_4)$  and  $(y_0, x_0, b, u_4)$ .

Case (b<sub>1</sub>). If  $a \neq b$ , then  $(x_0, y_0, a, u_4, b, x_0) = C^5 \subseteq G^n$ .

Case (b<sub>2</sub>). The case  $a = b$  yields a  $C^3$  which is not excluded. However, if there are two such  $a$ 's, then  $(x_0 a_1 y_0 a_2 x_0) = C^4 \subseteq G^n$  and we are home again. Thus we may assume that for each edge there is at most one  $a$  such that, for some  $u$ ,  $(x_0, y_0, a, u)$  and  $(y_0, x_0, a, u)$  are in  $G^n$ .

First we assume not only (23) (with  $k=2$ ), but also

$$(25) \quad \underline{d}(G^n) \leq 3\sqrt{\frac{n}{2}}.$$

Then we find  $(x_0, y_0)$  as in the proof of Theorem 1 and observe that it is contained in at least  $n + 4\sqrt{n/2}$  "good" walks.\* Cases (a) and (b<sub>1</sub>) are already settled, in case (b<sub>2</sub>)  $a$  is uniquely determined, therefore  $u$  can be chosen in at most  $3\sqrt{n/2}$  ways. Forget these walks, we have still a pair of walks with the same endvertex  $u^*$ , for which we have Case (b<sub>1</sub>):  $C^5 \subseteq G^n$ .

Hence our only problem is to ensure (25) and then the proof is complete. First we show that (25) can be replaced by

$$(26) \quad \underline{d}(G^n) \leq \frac{1}{3}\sqrt{\frac{n}{2}} + 1.$$

Indeed, assume (26). Let  $x$  be an arbitrary vertex and  $N(x) = \{y_1, \dots, y_d\}$  be the set of its neighbours. Let  $G_x$  be the graph spanned by  $N(x)$ . Since  $C^4 \subseteq G^n$ ,  $\underline{d}(G_x) \leq 1$ . Thus each  $y_i$  has at least  $d(y_i) - 2$  neighbours in  $V(G^n) - N(x) - x$ . Thus (fig. 4)

$$\sum (d(y_i) - 2) \leq n - d(x) - 1,$$

which means that  $d(x) \cdot (\underline{d}(G^n) - 1) \leq n - 1$ . This and (26) yield (25).

\* Theorem 3\* is not applicable, since  $k$  is odd. However, the estimate on  $w_3(G^n) - cw_1(G^n) = w_3(G^n) - ce(G^n)$  is trivial by [11].

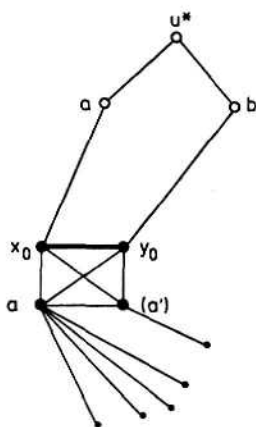


Fig. 3

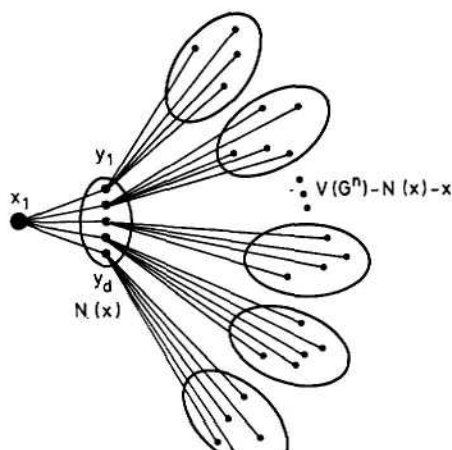


Fig. 4

To ensure (26) we use a standard "regularization" argument. Clearly, the theorem is proved for graphs satisfying (26) (beside (23)). Further, it holds for  $n=1, 2, 3$ , since its condition (23) is never satisfied.

Let now  $G^n$  be an arbitrary graph satisfying (23) with  $k=2$ . We show by induction on  $n$  that  $G^n$  contains either a  $C^4$  or a  $C^5$ . If in addition, (26) also holds, then we are home. Therefore we may assume the existence of a vertex  $x$  violating (26):  $d(x) < (1/3)\sqrt{n/2} + 1$ . Deleting this  $x$  we get a  $G^{n-1}$  with

$$e(G^{n-1}) = e(G^n) - d(x) \cong \left(\frac{n}{2}\right)^{3/2} + 4\sqrt{\frac{n}{2}} - \frac{1}{3}\sqrt{\frac{n}{2}} - 1 \cong \left(\frac{n-1}{2}\right)^{3/2} + 4\sqrt{\frac{n-1}{2}}.$$

(This inequality is an easy consequence of the Lagrange mean value theorem for  $n \geq 12$ . However, we need it for  $4 \leq n \leq 11$  as well. One can check it for each  $n$  separately.)

Thus, by induction,  $G^{n-1}$  contains a  $C^4$  or a  $C^5$ . This completes our proof. ■

**Proof of Theorem 3\*.** In [12] Faudree and Simonovits proved that  $ex(n, L) = O(n^{1+1/k})$ . In their proof the main tool was that if  $\tilde{G}^n$  is a bipartite graph with

$$(27) \quad d(\tilde{G}^n) \cong \frac{1}{2} a \sqrt[k]{n} \quad \text{and} \quad L \not\subseteq \tilde{G}^n,$$

$x$  is an arbitrary fixed vertex and  $S_p$  is the set of vertices at distance  $p$  from  $x$ , then  $|S_p| \cong c_p (a \cdot \sqrt[k]{n})^p$  for some fixed constant  $c_p > 0$  ( $p=1, 2, \dots, k$ ). Hence

$$(28) \quad |S_k| \cong c_k a^k n.$$

This fact will be needed below.

Let

$$a = \limsup_{n \rightarrow \infty} \frac{ex(n, \{L, C^3, C^5, C^7, \dots\})}{n^{1+1/k}} > 0.$$

We shall assume that

$$(29) \quad d(G^n) > a \sqrt[k]{n} \text{ and } L \not\subseteq G^n.$$

Let  $\tilde{G}^n \subseteq G^n$  be a bipartite graph with the maximum number of edges (in  $G^n$ ). The maximality trivially implies (27). Put  $t = 2(k+1)(\lceil (c_k a^k)^{-1} \rceil + 1)$ . First we prove that if  $G^n \not\supset L$  and  $G^n \not\supset C^{2q+1}$  for  $q \leq t$ , then (under the additional condition (29))  $G^n$  is bipartite. Assume indirectly, that  $G^n$  contains an odd cycle  $C^{2q+1}$  for some  $q > t$ . (For  $q \leq t$  this is excluded.) Let  $C^{2q+1}$  be the shortest odd cycle in  $G^n$ , with the vertices  $x_0, \dots, x_{2q}$ . Take the vertices  $x_{2j(k+1)}$  for  $j=0, 1, 2, \dots, \lceil (c_k a^k)^{-1} \rceil$ . Let  $S_k(j) = S(j)$  be the set of vertices of  $G^n$  having distance  $k$  from  $x_{2j(k+1)}$ . By (28),  $|S(j)| \cong c_k a^k n$ . Thus there exists a  $u_0$  belonging to two sets  $S(j)$ , say  $u_0 \in S(j) \cap S(h)$ . The path joining  $u_0$  to  $x_{2j(k+1)}$  and  $x_{2h(k+1)}$  and the odd arc of our  $C^{2q+1}$  yield a closed odd walk containing an odd cycle  $C^{2r+1}$  with  $r < q$ . Since  $C^{2q+1}$  was the shortest walk, we arrived at a contradiction.

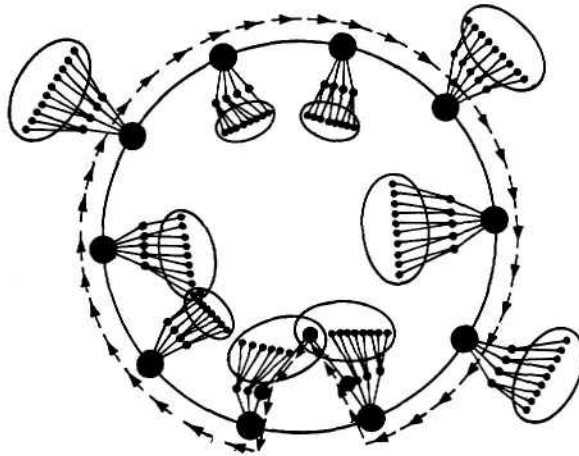


Fig. 5

In the general case we have to drop the condition (29). We shall use here a variant of the “regularization argument” of the proof of Theorem 2.

The graphs  $G^{n-j}$  are defined recursively:  $G^{n-j}$  is obtained from  $G^{n-j+1}$  by deleting a vertex  $x_j$  of degree  $< a \cdot \sqrt[k]{n-j+1}$ , if such a vertex exists. If all the vertices of  $G^{n-j+1}$  have degree  $\cong a \cdot \sqrt[k]{n-j+1}$ , then we see by the argument above that  $G^{n-j+1}$  is a bipartite graph. Thus

$$(30) \quad e(G^n) \cong \text{ex}(n-j+1, \{L, C^3, \dots, C^{2r+1}, \dots\}) + \sum_{m=n-j+1}^n a \cdot \sqrt[k]{m}.$$

We agree to stop the recursion if either  $G^{n-j+1}$  is bipartite or if  $j = \lfloor n/2 \rfloor$ . Then, by (30),

$$e(G^n) \cong (a + e_{n-j})(n-j)^{1+\frac{1}{k}} + a \cdot \sum \sqrt[k]{m}$$

where  $\varepsilon_{n-j} \rightarrow 0$ . Since

$$(n-j)^{1+\frac{1}{k}} + \sum^k \sqrt[k]{m} \leq n^{1+\frac{1}{k}},$$

we obtain that

$$e(G^n) \leq (a + o(1))n^{1+\frac{1}{k}}. \quad \blacksquare$$

**Added in proof.** We would like to thank very much C. D. Godsil who was originally a co-author of the present note for his valuable contribution. Recently he has discovered that most of his results in this paper are covered by [11], [17], [18], and he withdrew his name. May his decision be right, we still owe him a great debt of gratitude for, among others, the proof of Theorem 4 which is due to him. We also note that at an earlier stage we heavily relied on some results of Faudree and McKay.

### References

- [1] C. BENSON, Minimal regular graphs of girth eight and twelve, *Canadian J. Math.* **18** (1966) 1091—1094.
- [2] J. A. BONDY and M. SIMONOVITS, Cycles of even length in graphs, *Journal of Combinatorial Theory*, **16** (1974) 97—105.
- [3] W. G. BROWN, On graphs that do not contain a Thomsen graph, *Canadian Math. Bull.* **9** (1966) 281—285.
- [4] W. G. BROWN, P. ERDŐS and M. SIMONOVITS, Algorithmic solution of extremal digraph problems, *to be published*.
- [5] D. CVETKOVIĆ, M. DOOB and H. SACHS, *Spectra of graphs — Theory and Application* — VEB Deutscher Verlag der Wissenschaften, Berlin, 1980.
- [6] P. ERDŐS, On sequences of integers no one of which divides the product of two others, and some related problems, *Mitt. Forschungsinstitut Math. u. Mech. Tomsk* **2** (1938) 74—82.
- [7] P. ERDŐS, Graph theory and probability, *Can J. Math.* **11** (1959) 34—38.
- [8] P. ERDŐS, A. RÉNYI and V. T. SÓS, On a problem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 215—235.
- [9] P. ERDŐS and M. SIMONOVITS, A limit theorem in graph theory *Studia Sci. Math. Hungar.* **1** (1966) 51—57.
- [10] P. ERDŐS and M. SIMONOVITS, Supersaturated graphs, *submitted to Combinatorica*.
- [11] G. R. BLAKLEY and P. ROY, Hölder type inequality for symmetric matrices with nonnegative entries. *Proc. AMS* **16** (1965) 1244—1245.
- [12] R. J. FAUDREE and M. SIMONOVITS, On a class of degenerate extremal problems, *submitted to Combinatorica*.
- [13] T. KÖVÁRI, V. T. SÓS and P. TURÁN, On a problem of Zarankiewicz, *Coll. Math.* **3** (1954) (50—57).
- [14] I. REIMAN, Über ein problem von Zarankiewicz, *Acta Acad. Sci. Hungar.* **9** (1958) 269—279.
- [15] R. SINGLETON, On minimal graphs of maximum even girth, *J. Combin. Theory* **1** (1966) 306—332.
- [16] P. TURÁN, On an extremal problem in graph theory, (in Hungarian), *Mat. Fiz. Lapok* **48** (1941) 436—452.
- [17] D. LONDON, Inequalities in quadratic forms, *Duke Math. J.* **83** (1966) 511—522.
- [18] H. P. MULHOLLAND and C. A. B. SMITH, An inequality arising in genetical theory, *Amer. Math. Monthly* **66** (1969), 673—683.

P. Erdős

*Mathematical Institute of the  
Hungarian Academy of Sciences  
Reáltanoda u. 13—15  
Budapest, Hungary H—1053*

M. Simonovits

*Dept. Analysis I.  
Eötvös University  
Budapest, Hungary H—1088*