Let $G(n)$ be a graph of $n$ vertices, $G(n; k)$ a graph of $n$ vertices and $k$ edges. $f(n; G(m))$ is the smallest integer so that every $G(n; k)$ contains a subgraph isomorphic to $G(m)$. More generally let $G_1, \ldots$ be a finite or infinite family of finite graphs. $f(n; G_1, \ldots)$ is the smallest integer so that every $G(n; f(n; G_1, \ldots))$ contains one of the $G_k$ as a subgraph. Many papers have been published in the last few years on the determination or estimation of these functions. In one of my recent papers I give a far from complete list of papers dealing with extremal problems in graph theory. Bollobás is about to publish a comprehensive book on this subject which will also contain a very extensive list of references.

In this paper I first of all state a few of my favorite unsolved extremal problems. Then I prove the following theorems:

**Theorem 1.** Assume that $G(n)$ does not contain a $C_{2k+1}$ for $3 \leq k \leq r$. Then the independence number of $G(n)$ is greater than $c_k^n^{1-1/r}$.

$C_k$ is a circuit of $k$ edges and the independence number of $G(n)$ is the cardinal number of the largest set of vertices no two of which are joined by an edge.

$K(m)$ is the complete graph of $m$ vertices. Denote by $K_{\text{top}}(m)$ an arbitrary subdivision of $K(m)$ (i.e. a topological complete $K(m)$. $K_{\text{top}}(3)$ is simply a circuit).

**Theorem 2.** There is a function $f(c) > 0$ so that every $G(n; \lceil cn^2 \rceil)$ contains a $K_{\text{top}}(\ell)$ with $\ell \geq f(c)n^2.$

Before proving the theorems we will state several related conjectures.

P. Erdős, Some recent progress on extremal problems in graph theory, Proc. sixth Southeastern conference on combinatorics graph theory and computing 1975, Utilitas Math. press 3-14, Congress Num XIV. We refer to this paper as I.

For further problems see my paper: Problems and results in graph theory and combinatorial analysis, Proc. fifth British comb conference 1975, 169-192, Utilitas Math. Cong. Num XV. I refer to this paper as II. For some historical remarks see P. Erdős, Problems in number theory and combinatorics, Proc. sixth Manitoba conference on numerical math, Congress Num. XVIII, 35-58. For some further extremal and other problems see my paper, Some recent problems and results in graph theory combinatorics and number theory. Proc. seventh Southeastern conference. Ut. Math. press 3-14,

1. Simonovits and I conjectured that if $G$ is bipartite (unless stated otherwise $G$ is always bipartite) then there is a rational number $a, 1 \leq a < 2$ so that

$$\lim_{n \to \infty} \frac{f(n; G)}{n^a} = c_a, \quad 0 < c_a < \infty.$$  

We are very far from being able to prove (1). As a first step one should prove that to every bipartite $G$ there is an $a$ so that for every $\varepsilon > 0$ and $n > n_0(\varepsilon)$

$$n^{a-\varepsilon} < f(n; G) < n^{a+\varepsilon}.$$  

(1.2) perhaps will not be very hard to prove.

We further conjecture that to every rational $a, 1 \leq a < 2$, there is a $G$ for which (1.1) is satisfied.


Nothing like (1) holds for hypergraphs. This follows from a result of Szemerédi and Rursa see II p. 179.

For non-bipartite graphs the results of Simonovits, Stone and myself cleared up the situation to some extent, though many problems remain.


2. Define $V(G)$ as the minimum valency (or degree) of all the vertices of $G$. Put $V_1(G) = \max V(G')$ where the maximum is taken over all the subgraphs of $G$. Simonovits and I asked: Is it true that

$$f(n; G) < c n^{3/2} \text{ if } V_1(G) = 2?$$  

We now expect that (2.1) is false, but can prove nothing.

Assume $V_1(G) = r$. A result of Rényi and myself implies $f(n; G) > c n^{2(1-1/r)}$. Define $a_1(r)$ and $a_2(r)$ as follows: For $V_1(G) = r$ and every $\varepsilon > 0$ if $n > n_0(\varepsilon)$,
\[ n^{a_1(r)-\epsilon} < f(n; G) < n^{a_2(r)+\epsilon}. \]

Our result with Rényi implies \( a_1(r) \geq 2(1 - 1/r) \). Is this best possible? Is it true that for every \( r \), \( a_2(r) < 2 \)? Unfortunately we do not know this even for \( r = 2 \).


3. Denote by \( D_n \) the graph of the \( n \) dimensional cube, it has \( 2^n \) vertices and \( n 2^{n-1} \) edges, \( D_2 = C_4 \).

Simonovits and I proved \( f(n; D_3) < c n^{8/5} \). Probably the exponent \( 8/5 \) is best possible, but we have not even been able to prove \( f(n; D_3)/n^{3/2} \) is.

Brown, V.T. Sós, Rényi and I proved

\[ f(n; C_4) = (3 + o(1))n^{3/2}. \]

Let \( \theta \) be a power of a prime. We also proved

\[ f(\theta^2 + \theta + 1; C_4) \geq \frac{3}{2}(p^2 + p) + p^2 + 1, \]

perhaps there is equality in (3.2). I proved in I that

\[ f(n; C_4) \leq \frac{1}{2} n^{3/2} + \frac{n}{4} - \left(\frac{3}{16} + o(1)\right)n^{1/4}. \]

I conjectured

\[ f(n; C_4) = \frac{1}{2} n^{3/2} + \frac{n}{4} + o(n). \]

It is not impossible that in (3.4) the error term is \( o(n^{3/4}) \).

\( K(u, v) \) is the complete bipartite graph of \( u \) white and \( v \) black vertices. Kövári, V.T. Sós, P. Turán and I proved

\[ f(n; K(u, v)) < c n^{2-1/r}. \]

Very likely the exponent in (3.5) is best possible. For \( r = 2 \) this is implied by (3.3) and Brown proved it for \( r = 3 \), but for \( r > 3 \) nothing is known.

Denote by \( G - e \) the subgraph of \( G \) from which the edge \( e \) has been omitted.

Simonovits and I proved

\[ f(n; D_3 - e) < c n^{3/2} \]

and I proved

\[ f(n; K(r, r) - e) < c n^{1-1/(r-1)}. \]

Simonovits and I tried to characterize the graphs \( G \) with the property that for every proper subgraph \( G' \)
We were of course unsuccessful, but in view of (3.6) and (3.7) it seemed to us that highly symmetric graphs are likely to satisfy (3.8).

Our paper with Simonovits is quoted in 1.


4. We have

\[ \lim_{n \to \infty} \frac{1}{n^{3/2}} \cdot f(n; c_3, c_4) = 0. \]

The lower bound is a result of Reiman and E. Klein (Mrs. Szekeres). The upper bound is (3.1). Determine the value of the limit in (4.1). I never managed to get anywhere with this question and cannot decide whether it is really difficult or whether I overlook a simple argument. I was never able to improve (4.1).

More generally, let \( G_1, \ldots, G_k \) be a family of graphs some of which are bipartite. I hope and expect that

\[ \lim_{n \to \infty} \frac{f(n; G_1, \ldots, G_k)}{n^a} = c \]

Assume that the conjecture (1.1) holds and let \( a_1 \) be the rational number for which

\[ \lim_{n \to \infty} \frac{f(n; G_1)}{n^{a_1}} = c_1, \quad 0 < c_1 < \infty. \]

Perhaps \( a = \min a_i \). I have of course no real evidence for this. I am sure that the situation changes completely for infinite families of graphs \( \{G_k\}, 1 \leq k < \infty \).

At the moment I do not know an example of an infinite family of graphs \( \{G_k\}, 1 \leq k < \infty \) so that there is an \( a \) with \( f(n; G_k)/n^a \to \infty \) for every \( k \), but for some \( \beta < a \)

\[ \lim_{n \to \infty} \frac{f(n; G_1, \ldots)}{n^\beta} = 0. \]

Probably the family of graphs \( G \) with \( V_1(G) \geq 3 \) satisfies (4.3) for every \( \beta > 1 \). This is an old conjecture of Sauer and myself. If true then, since for these graphs \( f(n; G) > c n^{4/3} \) by our result with Rényi stated in 2, this family would have the
above property.

Our problem with Gauer is discussed in I. p. 10 and II p. 178.


5. As far as I know G. Dirac was the first to investigate $f(n; K_{top}(m))$. Trivially
$f(n; K_{top}(3)) = n$ and G. Dirac proved $f(n; K_{top}(4)) = 2n - 2$. He conjectured
$f(n; K_{top}(5)) = 3n - 5$. It is surprising that this attractive conjecture is still open. Mader proved that
$f(n; K_{top}(m)) \leq 2^{m-2} n$, he conjectured

\[(5.1)\]

$$f(n; K_{top}(m)) < c m^n.$$ 

(5.1) is probably rather deep. It is easy to see (as was of course known to Mader) that the conjecture if true is best possible - apart from the value of $c$.

Theorem 2 can be considered as proving (5.1) for large values of $m$, but it is very doubtful if it will help in proving (5.1). Before we prove our theorems we give a preliminary discussion and state some conjectures, some of which are in my opinion more interesting than the theorems. First of all it would be of interest to determine the largest $f(c)$ for which Theorem 2 holds. I am sure that it will be a continuous strictly increasing function of $c$. It is not hard to prove that $f(c) + 0$ as $c + 0$ and $f(c) = o$ as $c + \frac{1}{2}$. It would of course be interesting to determine $f(c)$ explicitly.

I am sure that the following strengthening of Theorem 2 holds.

Conjecture 1: Every $G(n; [c_1 n^2])$ contains $[c_2 n^2]$ vertices $x_1, \ldots, x_r$, $r = [c_2 n^2]$ so that $x_i$ and $x_j$, $1 \leq i < j \leq r$ are joined by vertex disjoint paths of length 2

This conjecture is clearly connected with the following problem of perhaps greater independent interest.

Let $|S| = n, A_k \subset S, |A_k| > c n, 1 \leq k \leq m$. Determine the largest $f(n, m, \varepsilon, c)$ so that there always are sets $A_{k_1}, 1 \leq i \leq f(n, m, c, \varepsilon)$ for which for every

$1 \leq i_1 < i_2 \leq f(n, m, c, \varepsilon)$

$$|A_{k_{i_1}} \cap A_{k_{i_2}}| > cn.$$ 

$\varepsilon > 0$ can be chosen as small as we wish but must be independent of $n$ and $m$. Observe that if $c > \frac{1}{2}$ then for sufficiently small $\varepsilon = \varepsilon(c)$, $f(n, m, c, \varepsilon) = m$. Thus the problem is of interest only for $c \leq \frac{1}{2}$.

The connection between this problem and the conjecture is easy to establish. First of all it is well known and easy to see that every $G(n; c n^2)$ contains a
subgraph $G(N)$, $N > c_1 n$ each vertex of which has valency greater than $(2c + o(1))N$.
(To prove the lemma omit successively the vertices of smallest valency).

Let the vertices of $G(N)$ be $x_1, \ldots, x_N$. The sets $A_k$ are the vertices joined to $x_k$.

It is immediate that Conjecture 1 is a consequence of

Conjecture 2. For $n = m$ and $\varepsilon = \varepsilon(c)$ sufficiently small

$$(5.2) \quad f(n, m, c, \varepsilon) > n^{\frac{2}{m}}$$

for some $n = n(c, \varepsilon) > 0$.

I can not even disprove

$$(5.3) \quad f(n, m, c, \varepsilon) \geq n m$$

for $m < n$ and $n = n(c, \varepsilon)$. On the other hand I can not prove (5.3) for $c = 1$ even

for $c = \varepsilon$.

Perhaps for every $m \leq 2^n$ and $c = \varepsilon(n)$

$$(5.4) \quad f(n, m, c, \varepsilon) > m^{1-\gamma}.$$ 

(5.4), if true, is best possible. To see this, let the $A$'s be all subsets of $S$ having

at least on elements, and let $m = 2^n - \sum_{i=1}^{n} (\frac{m}{i})$. It is easy to see that in this case

(5.4) can not be improved.

These conjectures have many connections with other interesting questions in

graph theory. First of all an old conjecture of Kneser states as follows: Let

$|S| = 2n + k$. The vertices of $G$ are the $\binom{2n+k}{n}$ subsets of size $n$ of $S$. Join two

vertices if the corresponding n-sets are disjoint. Prove that the chromatic number

$\chi(G)$ of $G$ is $k + 2$. This conjecture has recently been proved by Lovász and Bárány in

a surprisingly simple way. Their proofs will appear soon.

Define now a graph $G(n, m, c)$ as follows. Its vertices are the $m$ sets $A_k \subset S$.

Two A's are joined if $|A_k \cap A_{k'}| < cn$. Determine or estimate $\chi(G(n, m, c))$. (5.2) would

follow from $\chi(G(n, m, c)) < c n^{\frac{2}{m}}$. Perhaps very much more is true e.g. $\chi(G(n, m, c)) < c_1 (\log m)^{2}$.

Ramsey's theorem can be used to obtain weaker inequalities than (5.2). Let

$1/r > c > 1/r+1$. A simple argument shows that for sufficiently small $\varepsilon$ the largest

independent set of $G(n, m, c)$ is at most $r$. ($G$ is the complementary graph of $G$, i.e.

two vertices are joined in $G$ if and only if they are not joined in $G$. A set of vertices

is independent if no two of them are joined by an edge). Thus by a well known theorem

of Szekeres and myself it contains a complete graph of size $> c n^{1/r}$. (for $r = 1$ the

whole graph of course is complete). In other words (5.2) holds with $1/r$ instead of $\frac{1}{2}$.

Now we show that it is possible to obtain considerably stronger results. First
of all assume $1/3 < c < \frac{1}{3}$. Clearly in our graph the largest independent set has size 2, but we can easily get some further information. Let $A_1, \ldots, A_5, |A_i| > cn$, $c > 1/3$ be any five of our sets. Then a simple argument shows that there is a set $S_1 = S$, $|S_1| > cn$ which is contained in three of the $A$'s. In other words every five points of our graph spans a triangle. Thus the complementary graph of our graph contains no triangle and no pentagon. But then by Theorem 1 it contains an independent set of size $> c n^{2/3}$ or our graph contains a complete graph of size greater than $cn^{2/3}$.

More generally assume $|A_k| > n/r + 1 (1 + n)$ for some $n > 0$ ($A_2 \subseteq S$, $|S| = n$, $1 \leq k \leq m$). Join two sets $A_k_1$ and $A_k_2$ if $|A_k_1 \cap A_k_2| > cn$, $c = c(n)$ is sufficiently small. Then these graphs belonging to the set system the graphs depend on $c$ has the following property: For every fixed $t = t(\varepsilon)$ and $1 \leq t$ every set of $t(r + 1)$ vertices contains a $k(1 + 1)$. I hope that for sufficiently large $t = t(r, \delta)$ this condition implies that our graph contains a complete graph of size $> m^{1-\delta}$. (For Conjecture 1 it suffices to prove this for $\delta = \frac{1}{r}$).

for a very simple proof see:

6. Now we prove Theorem 1. Let the vertices of our graph $G(n)$ be $x_1, \ldots, x_m$.
Denote by $S_i$ the set of those $x_j$'s which can be joined to $x_i$ by a path of length $i$ but not by a shorter path ($S_0$ is defined to be $x_1$). Observe that the set $S_i$ is independent of the set $\bigcup_{j=1}^{i+2} S_j$ (i.e. no vertex of $S_i$ is joined (by an edge) to a vertex of $\bigcup_{j=1}^{i+2} S_j$).
Observe further that for $1 \leq i \leq r$, $S_i$ is an independent set. For if two vertices of $S_i$ are joined then our $G(n)$ contains an odd circuit of size $\leq 2i + 1$, which contradicts our assumptions. Observe next that for some $i$, $0 \leq i \leq r - 1$,

$$\frac{|S_i|}{|S_{i-1}|} < n^{1/r}. \quad (6.1)$$

(6.1) follows immediately from the fact that $S_{i+1} \cap S_{i+2} = \emptyset$ and that $\bigcup_{j=1}^{i+2} S_j \subseteq S_i$.
(In fact we can assume $|\bigcup_{j=1}^{i+2} S_j| < n$ for if not then $\max_{1 \leq i \leq r} |S_i| \geq (n-1)/r$ which implies...
Let now $i \geq 0$ be the smallest index satisfying (6.1). We construct our large independent subset of $G(n)$ as follows: The vertices of $S_i$ will be in our large independent set. $G_i$ is the subgraph of $G$ spanned by those vertices of $G$ which are not in $\bigcup_{j=0}^{i+1} S_j$. Clearly by (1) and the minimum property of $i$

\[(6.2) \quad |\bigcup_{j=0}^{i+1} S_j| < (n^{1/r} + 1) |S_i| \]

or $G_i$ has at least $n - (n^{1/r} + 1) |S_i|$ vertices and no vertex of $S_i$ is joined to any vertex of $G_i$. Repeat the same construction for $G_i$ and continue until all vertices are exhausted. The union of the $S_i$ belonging to the $G_i$ will be our large independent set of size $> (1 - n^{1-1/r}) n$ for every $n > 0$ if $n > n_o(n)$. This last statement easily follows from (6.1) and (6.2).

Probably the exponent $1 - 1/r$ cannot be improved this is known only for $r = 1$. I expect that $cn^{1-1/r}$ can be improved by a logarithmic factor but this also is known only for $r = 1$.

Assume now that $G(n)$ has girth greater than $2r + r$. (i.e. $G(n)$ has no circuit of length $\leq 2r + 2$). I cannot prove more than Theorem 1, i.e. I can only show that $G(n)$ has an independent set of size greater than $cn^{1-1/r}$. I wonder if the exponent $1 - 1/r$ is best possible. The case $r = 1$ is perhaps most interesting, i.e. $G(n)$ has no triangle and rectangle. Is there an independent set of size $> n^{3+\epsilon}$? I do not know.


For a penetrating and deep study of extremal problems on cycles in graphs see;


To finish our paper we now prove Theorem 2. First of all observe that Theorem 2 clearly holds for $c > \lambda$. To see this observe that, by the lemma stated in 5, our $G(n; (\lambda + \delta)n^2)$ contains a subgraph $G'$ of $N > c_n$ vertices each vertex of which has valency greater than $N(1+\delta)/2$. But then to every two vertices of $G'$ there exist $\delta N > c_n$ vertices which are joined to both of them. But then it is immediate that every set $y_1, \ldots, y_t, t = [c_n n]$ of vertices is a $K_{\top}(t)$ i.e. any two are joined by vertex disjoint paths of length two. Thus Theorem 2 is proved for $c > \lambda$.

Assume now that Theorem 2 is false. Let $C$ be the upper bound of the numbers for which Theorem 2 fails. In other words, for every $\epsilon > 0$ there is an infinite sequence $n_1 < n_2 < \ldots$ and graphs $G(n_i; (C - \epsilon)n^2)$ which do not contain a $K_{\top}(t)$ for
i > n \frac{2}{n}, for any fixed n if \( n_i > n(n, e) \), but no such sequence of graphs \( G(n; (C + e)n^2) \) exist. We now easily show that this assumption leads to a contradiction.

First of all our assumption means that there is an infinite sequence of integers \( n_1 < \ldots \) so that there is a graph \( G(n_1; (C - o(1)n^2) \) the largest \( K_{\text{top}}(l) \) of which satisfies \( l/n_1^2 = 0 \) and that \( C \) is the largest number with this property. Further by the trivial lemma stated in 5, we can assume that every vertex of our \( G \) has valency not less than \( (2C - o(1))n_1^2 \). Our assumption implies that there is a sequence \( n_i \to 0 \) and \( l_i \to \infty \) so that our \( G(n_i; (C - o(1)n_1^2) \) has the property that we can omit \( n_i \) of its vertices, so that in the remaining graph \( G'(n_i - n_i) = G'_i \) there are two vertices which cannot be joined by a path of length less than \( k_i \). To see this, observe that if our statement would be false then for sufficiently small \( n \) every set of \( n_1 \) sets of vertices of our \( G(n_i) \) would be a \( K_{\text{top}}(l) \).

To arrive at the contradiction let \( y_1 \) and \( y_2 \) be two vertices of our \( G'_i \) which cannot be joined by a path of length less than \( k_i \). Observe that every vertex of our \( G'_i \) has valency not less than \( (2C - o(1))n_1^2 \). Denote by \( S_1(j) \), respectively \( S_2(j) \), the set of vertices which can be joined to \( y_1 \), respectively \( y_2 \), with \( i \) but not with fewer edges. Clearly for every \( t \leq \frac{k_i - 1}{2} \) the two sets \( S_1(j) \) and \( S_2(j) \) are disjoint. \( (S_1^{(0)} = y_1, S_2^{(0)} = y_2) \) (otherwise there would be a path of length less than \( k_i \) joining \( y_1 \) and \( y_2 \)). Without loss of generality we can thus assume
\[
|S_1^{(t)}| < \frac{n_i}{2}, \quad |S_1^{(1)}| > (2C - o(1)n_i).
\]
From (7.1) we obtain that there is an \( 2 \leq r < t \) for which
\[
|S_i^{(r)}| < \frac{n_i}{2(2t-1)}.
\]
Let now \( G_i^{(r)} \) be the subgraph of \( G'_i \) spanned by the vertices of \( C_0 \). The valency of every one of its vertices is at least \( (2C - o(1)) \). \( n_i \) which are joined to a vertex of \( G'_i \) are all in \( S_1^{(r)} \) which implies our statement by (7.2).

The sequences of graphs \( G_i^{(r)} \) establish our contradiction. The \( i \)-th graph has by (7.1) and (7.2) more than \( (2C - o(1))n_i \) and fewer than \( \frac{n_i}{2} \) vertices each of which has valency not less than \( (2C - o(1))n_i \) and the largest \( K_{\text{top}}(l) \) of it is \( o(n_i) \). This contradicts the maximality property of \( C \) and hence Theorem 2 is proved.