EXTREMAL GRAPH THEORY WITH EMPHASIS ON PROBABILISTIC METHODS

Conference Board of the Mathematical Sciences
REGIONAL CONFERENCE SERIES IN MATHEMATICS

supported by the
National Science Foundation

Number 62

EXTREMAL GRAPH THEORY WITH EMPHASIS ON PROBABILISTIC METHODS

Béla Bollobás

Published for the
Conference Board of the Mathematical Sciences
by the
American Mathematical Society
Providence, Rhode Island

Expository Lectures
from the CBMS Regional Conference
held at Emory University
June 18–22, 1984

This conference was supported in part by National Science Foundation Grant DMS-8403620.

1980 *Mathematics Subject Classifications* (1985 *Revision*). Primary 05C35, 05C25, 05C10, 05C40, 05C65; Secondary 60C05.

Library of Congress Cataloging-in-Publication Data

Bollobás, Béla.

Extremal graph theory with emphasis on probabilistic methods.

(Regional conference series in mathematics, ISSN 0160-7642; no. 62)

"Based on the marathon lecture series given at the NSF-CBMS Regional Conference on Extremal Graph Theory held in June 1984 at Emory University, Atlanta" – Preface.

Bibliography: p.

1. Graph theory. 2. Extremal problems (Mathematics) I. Title. II. Series.

QA1.R33 no. 62 [QA166] 510s [511'.5] 85-30670

ISBN 0-8218-0712-9 (alk. paper)

The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.

Preface

This monograph is based on the marathon lecture series given at the NSF-CBMS Regional Conference on Extremal Graph Theory held in June 1984 at Emory University, Atlanta. The author is grateful to Dwight Duffus, Ron Gould, and Peter Winkler for their superb organization of the meeting; the additional lectures by Dick Duke, Ralph Faudree, Ron Graham, and Tom Trotter greatly enriched the conference.

Since the publication of the author's book, *Extremal Graph Theory* (Academic Press, London, New York, and San Francisco, 1978, to be referred to as EGT), a number of important results have been proved, and one of the aims of the lectures was to update EGT by presenting some of these developments.

Over the past few years a noticeable shift has been taking place in extremal graph theory towards probabilistic methods. The most obvious sign is that random graphs are used more and more, but that is not all. Even more significantly, a probabilistic frame of mind was needed to find many of the proofs, which on the surface have nothing to do with probabilistic ideas. In several beautiful and difficult proofs the underlying philosophy is that we do not have to care about single vertices, say, for it suffices to make use of the fact that there are many subsets of vertices of a given cardinality with the right properties. To give a simple example, one often makes use of the fact that if $X_1, X_2, \ldots, X_N$ are nonnegative integers bounded by A, $\sum_{i=1}^{N} X_i = Na$ and $0 < b < a$, then at least $(a-b)N/(A-b)$ of the X_i's are greater than b. Equivalently, if X_i is a random variable, $0 < X < A$ and $E(X) = a$, then

(1) $$P(X > b) \geqslant (a-b)/(A-b) \quad \text{for all } 0 < b < a.$$

Inequality (1) has the following reformulation in graph-theoretic terms. If B is a bipartite graph with bipartition (X, Y), $X = \{x_1, x_2, \ldots, x_m\}$, $Y = \{y_1, y_2, \ldots, y_n\}$, $d(y_j) \leqslant \Delta$ for all j, $1 \leqslant j \leqslant n$, then for $d' < d = \sum_{i=1}^{m} d(x_i)/n$ there are at least $(d-d')n/(\Delta - d')$ vertices y_j of degree at least d'.

Needless to emphasize, in the great majority of the cases the merit is in finding the need for probabilistic inequalities and applying them cleverly, and not in proving the inequalities. The main aim of the lecture was to show how fruitful a probabilistic frame of mind is in tackling main line extremal problems.

The notation in these notes is taken from EGT. In particular, $|G|$ is the order of a graph G, i.e. the number of vertices, and $e(G)$ is the size of G, i.e. the number of edges. The cardinality of a set U is denoted by $|U|$, and the collection of r-subsets of U is $U^{(r)}$. Though these notes are practically self-contained, familiarity with at least some parts of EGT will certainly help the reader. A conscious effort has been made to prevent the lectures from turning into a long catalogue of results; without this effort, the monograph could have ended up with several hundred results. However, there seems little doubt that it is much more useful to present just a few of the deeper results and thereby leave time to dwell on the proofs.

The first two sections are closely related. They deal with subdivisions of graphs and subcontractions. Both areas owe a considerable amount to Mader, who proved that for every $p \in \mathbf{N}$ there are constants $s(p)$, $c(p)$ such that every graph of order n and size greater than $s(p)n$ contains a topological complete graph of order p, and every graph of order n and size greater than $c(p)n$ has a subcontraction to K^p. (Needless to say, $s(p)$ and $c(p)$ are taken to be the smallest values that will do in the statements above.) Consequently for every fixed graph H there is a constant $s(H)$ such that every graph of order n and size greater than $s(H)n$ contains a subdivision of H and there is an analogous constant $c(H)$. Bollobás started the study of subdivisions of graphs with some constraints on the subdivisions one allows. For example, we may wish to restrict the number of times we subdivide an edge, at least modulo some integer k. The main aim of §1 is to present recent result of Thomassen in this area, with a considerably better bound than the original one given by Thomassen.

The second section, on subcontractions, is devoted to a new result of Thomason and Kostochka, improving the upper bound on $c(p)$ proved by Mader. Together with a rather easy result of Bollobás, Catlin, and Erdős, this result implies that the order of $c(p)$ is $p(\log p)^{1/2}$, a fact not many of us would have expected.

The third and fourth sections concern different aspects of essentially the same problem. At least how many vertices must we have if the minimal degree is δ and the girth is at least g? At most how many vertices can we have if the maximal degree is at most Δ and the diameter is at most D? An "ideal" graph would answer both questions, but the trouble is that there are very few such ideal graphs. One is left with approximating the appropriate functions either by constructing suitable functions or by showing, usually by probabilistic methods, that suitable graphs do exist. As far as the bounds are concerned, the nonconstructive methods due to Erdős, Sachs, Bollobás, de la Vega, and others give better results, but the constructions have obvious advantages. In these sections the emphasis is on new constructive methods due to Bermond, Delorme, Farhi, Leland, Solomon, Jerrum, Skyum, and Margulis.

In §5 we concentrate on a substantial recent result of Gyárfás, Komlós, and Szemerédi concerning the distribution of cycle lengths in graphs with fairly many edges. Though the theorem is interesting, it is the proof, rich in ideas and techniques, that really justified spending two lectures on the result.

The sixth section contains a telegraphic review of the theory of random graphs. The highlights are the classical theorem of Erdős and Rényi on the evolution of random graphs and its recent extensions due to Bollobás.

In §7 we present a surprising and beautiful result of Beck on size Ramsey numbers. As a simple application of random graphs, Beck proved that there are graphs $G_1, G_2, \ldots$ such that G_i has at most ci edges, where c is a constant, and in any coloring of the edges of G_i with two colors we can find a monochromatic path of length s.

Saturated graphs were introduced over twenty years ago by Erdős, Hajnal, and Moon. Their result was extended considerably by Bollobás, who also introduced weakly saturated graphs. The main conjecture concerning weakly saturated graphs was proved recently by Alon, Frankl, and Kalai; the simple and elegant proof, based on exterior products (!), is presented in §8.

The last section, §9, concerns restricted colorings of graphs. We know from Vizing's theorem that a graph of maximal degree Δ is $(\Delta + 1)$-colorable. What happens if we prescribe a list for each edge from which the color of the edge has to be chosen? What is the maximal length of the lists that always let us color our graph? It has been conjectured that lists of length $\Delta + 1$ will do. This conjecture, if true, would clearly be best possible.

At the moment, the conjecture is far from being proved and many graph theorists suspect it to be false. The main aim of the section is to present a recent result of Bollobás and Harris, implying that for some constant $c < 2$ lists of size at most $c\Delta$ will do for every graph of maximal degree $\Delta \geqslant 3$. As a trivial consequence of this result, Bollobás and Harris made the first substantial progress towards a proof of a long-standing conjecture of Behzad concerning the total chromatic number.

It is a pleasure to thank Fan Chung, Dwight Duffus, Ron Graham, Hal Kierstead, and Andrew Thomason for their ideas and suggestions, many of which have been incorporated into the text. Finally, I would like to express my gratitude to all participants of the conference for their enthusiasm for the subject and the warm reception of the lectures.

Research was partially supported by National Science Foundation Grant DMS-8400643.

TRINITY COLLEGE, CAMBRIDGE, ENGLAND

LOUISIANA STATE UNIVERSITY

Contents

1. Subdivisions

A *subdivision* of a graph H is a graph obtained from H by subdividing the edges of H by some vertices (possibly none), i.e. by replacing each edge ab by an a-b path, such that any two of these paths have at most their endvertices in common. A subdivision of H is also said to be a *topological H-graph*; we write TH for such a graph. The *m-subdivision* of H, denoted by $T_m H$, is obtained from H be replacing each edge by a path with m internal vertices. Note that $T_m H$ is unique up to isomorphism, $T_0 H = H$, and if H has at least two edges then some TH is not an m-subdivision of H for any m.

Subdivisions were first studied because of Kuratowski's theorem. Almost twenty years ago Mader [74] proved the deep result that for $p \in \mathbb{N}$ every graph of order n and size at least $p2^{\binom{p-1}{2}}n$ contains a TK^p, a topological complete graph of order p. Consequently, for every graph H there is a constant $c = c(H)$ such that if the minimal degree $\delta(G)$ of a graph G is at least c then G contains a TH.

The bound $p2^{\binom{p-1}{2}}n$ was first improved by Halin [61] and then, substantially, by Mader [75], who proved the following result.

THEOREM 1. *For* $p \in \mathbb{N}$ *set* $t(p) = 3 \cdot 2^{p-3} - p$. *Then* $\mathrm{ex}(n; TK^p) \leqslant t(p)n - (t(p)+1)t(p)/2$. □

In Theorem 1 above the function ex has its customary meaning: given a family $\mathscr{F}$ of graphs, $\mathrm{ex}(n; \mathscr{F})$ denotes the maximal number of edges in a graph of order n which contains no member of $\mathscr{F}$ as a subgraph. An immediate consequence of Theorem 1 is that if $\delta(G) \geqslant 2t(p)$ then $G \supset TK^p$.

If H is not a forest and $m \geqslant 0$ then we cannot hope to find $T_m H$ in every graph of order n and size at least cn for some constant c, for there are graphs of arbitrarily large girth and minimal degree (see §3, Theorem 2 and Corollary 3). However, every forest can be found as a subgraph of a subdivision of a given graph in every graph of sufficiently large minimal degree. This was proved by Bollobás [17], and to state it precisely, we need the concept of a semitopological graph. Let F_0 be a subgraph of F. A *semitopological graph F with kernel* F_0 denoted by $\mathrm{ST}(F, F_0)$, is a subdivision of F in which no edge of F_0 has been subdivided.

THEOREM 2. *Given a graph F containing a forest F_0, there is a constant $c(F, F_0)$ such that every graph of minimal degree at least $c(F, F_0)$ contains an* $\mathrm{ST}(F, F_0)$. □

As an easy consequence of Theorem 2, one obtains the following result, conjectured by Burr and Erdős and first proved by Bollobás [16].

COROLLARY 3. *There is a function f: $\mathbf{N} \times \mathbf{N} \to \mathbf{N}$ such that if $\delta(G) \geqslant f(k, s)$ then the graph G contains a cycle of length $2s$ modulo k.*

PROOF. Given $k \in \mathbf{N}$ and $1 \leqslant s \leqslant k$, let F be the graph obtained from a path $P = x_0 x_1 \cdots x_k$ by adding to it a vertex x and independent $x - x_i$ paths of lengths s, $i = 0, 1, \ldots, k$. Let $F_0 \subset F$ be the tree formed by the $k + 1$ independent paths and set $f(k, s) = c(F, F_0)$.

By Theorem 2 all we have to check is that every $\mathrm{ST}(F, F_0)$ contains a cycle of length $2s \pmod k$. Let $\tilde{P}$ be the subdivision of P and pick two vertices x_i, x_j whose distance on $\tilde{P}$ is divisible by k. Then the paths $x - x_i$, $x - x_j$ and the $x_i - x_j$ segment of the path $\tilde{P}$ form a cycle of length $2s \pmod k$. □

Bipartite graphs show that in Corollary 3 we cannot replace $2s$ by s.

Recently Thomassen [84] proved an extension of Theorem 1 which also implies Corollary 3.

THEOREM 4. *There is a function f: $\mathbf{N} \times \mathbf{N} \times \mathbf{N} \to \mathbf{N}$ such that if $\delta(G) \geqslant f(p, k, s)$ then G contains a subdivision of K^p in which each edge has been subdivided into $2s$ edges modulo k.* □

In [17] another assertion was conjectured which also implies Corollary 3. This conjecture was also proved by Thomassen [85].

THEOREM 5. *There is a function g: $\mathbf{N} \times \mathbf{N} \to \mathbf{N}$ such that if $\delta(G) \geqslant g(k, m)$ then $G \supset T_m H$ for some graph H of minimal degree k.* □

Note once again that if H is any fixed graph with $\delta(H) = k \geqslant 2$, $m \geqslant 0$, and c is an arbitrarily large constant then there is a graph G of minimal degree at least c which does not contain an m-subdivision of H.

The function $g(k, m)$ which Thomassen showed has the property described in Theorem 5 grows rather fast; in fact, $g(k, 2^l - 1) > (8k)^{4^l}$. The main purpose of this section is to prove a slight improvement of Thomassen's theorem: we shall show that it suffices if our function grows considerably slower (Theorem 8). We shall need some simple results about domination in bipartite graphs.

THEOREM 6. *Let G be a bipartite graph with vertex classes U and W such that $d(x) \geqslant r$ for all $x \in W$. Let $a < u = |U|$, $b < w = |W|$, and $s < r$ be natural numbers such that*

$$\frac{b+1}{w}\binom{u}{a} > \sum_{l=0}^{s-1} \binom{r}{l}\binom{u-r}{a-l}.$$

Then there are sets $A \subset U$ and $B \subset W$ such that $|A| = a$, $|B| = b$, and each vertex of $W \setminus B$ is adjacent to at least s vertices of A.

PROOF. Let H be the bipartite graph with classes $U^{(a)}$ and W in which a vertex $A \in U^{(a)}$ (i.e., an a-subset A of U) is joined to a vertex $x \in W$ if $|\Gamma(x) \cap A| \leqslant s-1$. Then for $x \in W$ we have

$$d_H(x) \leqslant \sum_{l=0}^{s-1} \binom{r}{l}\binom{u-r}{a-l}.$$

Hence there is a vertex $A \in U^{(a)}$ such that

$$d_H(A) \leqslant \left\lfloor w \sum_{l=0}^{s-1} \binom{r}{l}\binom{u-r}{a-l} \Big/ \binom{u}{a} \right\rfloor \leqslant b. \quad \square$$

COROLLARY 7. *If* $a < u$, $b < w$, $3 \leqslant s < r$, $ar \geqslant 2us$, *and* $(a-s)(r-s) \geqslant \frac{9}{10}ar$ *(say* $a \geqslant 20s$ *and* $r \geqslant 20s$*) then one can find appropriate sets* A *and* B *provided* $w < (b+1)e^{s/10}\sqrt{2\pi s}$.

PROOF. Note that with $\lambda = ar/us \geqslant 2$ we have

$$\sum_{l=0}^{s-1} \binom{r}{l}\binom{u-r}{a-l} \Big/ \binom{u}{a} < \binom{r}{s}\binom{u-r}{a-s} \Big/ \binom{u}{a}$$

$$< \frac{1}{\sqrt{2\pi s}}\left(\frac{er}{s}\right)^s \frac{(a)_s(u-a)_{r-s}}{(u)_r} < \frac{1}{\sqrt{2\pi s}}\left(\frac{er}{s}\right)^s\left(\frac{a}{u}\right)^s\left(\frac{u-a}{u-s}\right)^{r-s}$$

$$< \frac{1}{\sqrt{2\pi s}}\left(\frac{era}{su}\right)^s \exp\left\{-\frac{(a-s)(r-s)}{u-s}\right\} \leqslant \frac{1}{\sqrt{2\pi s}}(e\lambda)^s \exp\left\{-\frac{9}{10}\lambda s\right\}$$

$$< \frac{1}{\sqrt{2\pi s}} e^{-s/10} < \frac{b+1}{w}.$$

Hence the conditions in Theorem 6 are satisfied. $\square$

In the proof of the main theorem we shall need some simple definitions. Given a rooted tree T with root x_0, the *height* of T is $h(T) = \max_{x \in T} d(x, x_0)$ and the jth *level* is $L_j(T) = \{x \in T\colon d(x, x_0) = j\}$. An *a-tree* of height b is a rooted tree of height b in which every vertex at level $h < b$ is joined to a vertex at level $h+1$. An *(a, b)-join* is a $(b-1)$- subdivision of a star $K(1, a)$, i.e. it is a union of a paths of lengths b, any two of them having precisely the same vertex in common. The common vertex is the center of the join.

THEOREM 8. *Let* k *and* m *be natural numbers and let* G *be a graph satisfying* $\delta(G) \geqslant \delta = 50(24k+1)^{m+1}$. *Then* G *contains an* m*-subdivision of a bipartite graph of minimal degree* k.

PROOF. As the result is trivial for $k = 1$, we may assume that $k \geqslant 2$. Set $k_1 = 24k$ and $r = 12(k_1+1)^{m+1}$. Denote by p the maximal number of vertex disjoint (k_1+1)-trees of height m in G; let $T_1, T_2, \ldots, T_p$ be vertex disjoint (k_1+1)-trees of height m in G. Then each T_i contains a k_1-set of vertices, say A_i, such that $A_i \subset L_m(T_i)$; the vertices of A_i belong to distinct branches of T_i and each vertex of A_i is joined to at most $2(k_1+1)^m - 1$ vertices outside the trees. Indeed, otherwise T_i would have two branches all of whose endvertices are joined to at least $2(k_1 + 1)^m$ vertices not belonging to the trees. However, then we could join each of these endvertices to $k_1 + 1$ vertices not belonging to the trees, without joining two endvertices to the same vertex. Hence we could obtain $p + 1$ vertex disjoint trees by replacing T_i by the two trees obtained from these two branches.

For $x \in A_i$ write $\tilde{d}(x)$ for the number of neighbors of x which belong to $\bigcup_{j=1;\, j\neq i}^{p} V(T_j)$. What we have just shown implies that

$$\begin{aligned}\tilde{d}(x) &\geqslant \delta - \{2(k_1+1)^m - 1\} - (t(k_1) - 1)\\ &> \delta - 3(k_1+1)^m\\ &\geqslant 48(k_1+1)^{m+1} = 4r.\end{aligned}$$

Here $t(k_1) = |V(T_i)| = 1 + (k_1 + 1)(k_1^m - 1)/(k_1 - 1)$. From this we see that $4r \leqslant (p-1)t(k_1)$, and so $p > 40k_1 = 960k$.

Every tree T_i contains a (k_1, m)-fan whose center is the root of T_i and whose endvertices are the vertices in A_i. Our aim is to show that parts of some t of these fans can be continued to $(2k, m+1)$-fans such that the endvertices of these fans form a set of cardinality at most t. Then the union of these fans is an m-subdivision of a graph which has a subgraph of minimum degree k.

Consider partitions $P = \{1, 2, \ldots, p\} = P_1 \cup P_2$, $|P_1| = \lceil p/2 \rceil$, and $|P_2| = \lfloor p/2 \rfloor$. Set $W(P_1) = \bigcup_{i\in P_1} A_i$ and $U(P_2) = \bigcup_{i\in P_2} V(T_i)$. Note that there are $\pi = \binom{p}{\lfloor p/2 \rfloor}$ such partitions. Furthermore, for any fixed vertex $x \in \bigcup_{i\in P} A_i$ precisely $\Pi_0 = \binom{p-1}{\lfloor p/2 \rfloor}$ of these partitions satisfy $x \in W(P_1)$. Write π_x for the number of partitions among these π_0 partitions such that x is joined to at least r vertices in $\bigcup(P_2)$. We wish to show that π_x is fairly large. To this end consider at random one of the π_0 partitions with $x \in W(P_1)$. Since $\lfloor p/2 \rfloor \geqslant (p-1)/2$, the expected number of neighbors of x in $U(P_2)$ is at least $\tilde{d}(x)/2$. Hence

$$(\pi_0 - \pi_x)(r-1) + \pi_x \tilde{d}(x) \geqslant \pi_0 \tilde{\tilde{d}}(x)/2$$

from which it follows that

$$\pi_x \geqslant \frac{\tilde{d}(x)/2 - (r-1)}{\tilde{d}(x) - (r-1)}\pi_0 \geqslant \frac{r+1}{3r+1}\pi_0 > \frac{1}{3}\pi_0 \geqslant \frac{\pi}{6}. \tag{1}$$

The alert reader must have noticed that our inequality is, in fact, just inequality (1) of the Preface, with $N = \pi_0$, $A = \tilde{d}(x)$, $a \geqslant \tilde{d}(x)/2$, and $b = r - 1$.

Inequality (1) implies that we can find a partition $P_1 \cup P_2$ and a set $W \subset W(P_1)$ such that $|W| = pk_1/6 = 4pk$ and each $x \in W$ is joined to at least r vertices in $U = U(P_2)$. Indeed, let H be the bipartite graph whose classes are $\bigcup_1^p A_i$ and the set of all π partitions. A vertex x of the first class is adjacent with a partition $P_1 \cup P_2$ if $x \in W(P_1)$ and, in our original graph G, x is joined to at least r vertices in $U(P_2)$. Then (1) implies that in H every vertex in the first class is joined to more than one sixth of the vertices of the second class. Hence some vertex in the second class of H is joined to more than one sixth of the vertices in the first class.

Consider the bipartite graph formed by the U-W edges of G. The parameters $w = |W| = 4pk$, $u = |U| = \lfloor p/2 \rfloor t(k_1)$, $a = q = \lceil p/12 \rceil$, $b = pk$, and $s = 2k$ satisfy the conditions of Corollary 7, so there are sets $W_0 = W - B$, $|W_0| = 3pk$, and $U_0 = A \subset U$, $|U_0| = q$, such that each vertex of W_0 is joined to at least $2k$ vertices in U_0.

The A_i's partition W_0: $W_0 = \bigcup_{i \in P_1}(A_i \cap W_0) = \bigcup_{i \in P_1} \tilde{A}_i$. At least $q = \lceil p/12 \rceil$ of the $\tilde{A}_i$'s have at least $2k$ elements, for otherwise we would have

$$(\lceil p/12 \rceil - 1)k_1 + (\lceil p/2 \rceil - \lceil p/12 \rceil)(2k - 1) \geqslant 3pk$$

and this is false.

Now if $|\tilde{A}_i| \geqslant 2k$ then a $(2k, m)$-fan in T_i ending in vertices of $\tilde{A}_i$ can be extended to a $(2k, m+1)$-fan all of whose endvertices are in U_0. Hence our graph G contains q internally disjoint $(2k, m+1)$-fans whose endvertices belong to a set with q elements and this set contains no internal vertex of any fan. The union of these fans is an m-subdivision of a q by q bipartite graph with $2kq$ edges. A graph of order $2q$ and size $2kq$ contains a subgraph of minimum degree k, so the theorem is proved. □

Consider the bipartite graph formed by the [illegible] edges of G. [illegible] $d_1 = |T|$ [illegible] and d_2 [illegible] $\ge pk$, and $\alpha > 2k$ satisfy the conditions of Corollary [illegible] so there are sets W_0 [illegible] and U_0 [illegible] $\ge q$, such that each vertex of W_0 is joined to at least $2k$ vertices in U.

[illegible] contains [illegible] $\ge$ [illegible] at least (p, q) [illegible] of the A's have at least $2k$ elements, for otherwise we would have

$$\{(p-2k+1)k\} + \{(q-2k+1)\}(2k-1) \ge 3pk$$

[illegible] is false.

Now [illegible] $\ge 2k$ then a $(2k, m)$-fan in T_1 ending in vertices of A_i can be extended to a $(2k, m+1)$-fan all of whose end-vertices are in U_0. Hence our graph G contains [illegible] disjoint $(2k, m+1)$-fans whose end-vertices belong to a set with q elements and this set contains no internal vertex of any fan. The union of these fans is an even subdivision of a p by q bipartite graph with $2kq$ edges. A graph of order $2q$ and size $2kq$ contains a subgraph of minimum degree k, so the theorem is proved. □

[illegible]

2. Contractions

An *elementary contraction* of a graph G is obtained from G by identifying two adjacent vertices; the result of a sequence of elementary contractions is a *contraction* of G. A graph H is a *subcontraction* or *minor* of G if H is a contraction of a subgraph of G; in notation $G > H$. Subcontractions to complete graphs are of special interest because of one of the deepest conjectures in graph theory, Hadwiger's conjecture [59]: $\chi(G) = p$ implies $G > K^p$. Define the *contraction clique number* of a graph G as $\mathrm{ccl}(G) = \max\{ p: G > K^p\}$.

Let us recall some of the basic results concerning subcontractions (see EGT, Chapter VII, §1). Hadwiger's conjecture has been proved for $\chi(G) \leqslant 5$. The conjecture is trivial for $\chi(G) = 3$. It was proved by Dirac [34] for $\chi(G) = 4$, and for $\chi(G) = 5$ it follows from the four color theorem due to Appel and Haken [4]. For $\chi(G) \geqslant 6$ the conjecture is still open.

Mader [74] was the first to prove that a graph with small contraction clique number must have small minimal degree.

THEOREM 1. *Let G be a graph of order n and size $2^{p-3}n$ where $p \geqslant 3$. Then* $\mathrm{ccl}(G) \geqslant p$.

PROOF. Let us apply induction on p. The assertion is clear for $p = 3$ so assume that $p \geqslant 4$ and the result holds for smaller values of p. Consider the set of finite graphs

$$\mathscr{F} = \{ G: e(G) \geqslant 2^{p-3}|G|\},$$

partially ordered by " < ". Let G_0 be a minimal element of this partially ordered set. For an edge $xy \in E(G_0)$ the elementary contraction G_0/y does not belong to $\mathscr{F}$ so

$$e(G_0/xy) = e(G_0) - 1 - |\Gamma(x) \cap \Gamma(y)| < 2^{p-3}(|G_0| - 1),$$

implying

$$|\Gamma(x) \cap \Gamma(y)| \geqslant 2^{p-3}. \tag{1}$$

Let now $x \in G$ and set $G_1 = G_0[\Gamma(x)]$. Then (1) implies that $\delta(G_1) \geqslant 2^{p-3}$ and so

$$e(G_1) \geqslant \frac{1}{2} 2^{p-3}|G_1| = 2^{p-4}|G_1|.$$

Hence $G_1 > K^{p-1}$ and so $G_0 > K^p$. □

Because of Theorem 1 it makes sense to study edge densities implying large contraction clique numbers. To be precise, for $p \in \mathbf{N}$, $p \geqslant 2$, set

$$c(p) = \inf\{c\colon e(G) \geqslant c|G| \text{ implies } G > K^p\}.$$

Theorem 1 shows that $c(p) \leqslant 2^{p-3}$ for all $p \geqslant 2$. This bound was improved considerably by Mader [75].

THEOREM 2. *If $p \geqslant 4$ then $c(p) \leqslant 8(p-2)\lfloor \log_2(p-2) \rfloor$.* □

For small values of p the function $c(p)$ looks very pleasant. It is trivial that $c(2) = 0$ and $c(3) = 1$. A little work enables us to show that $c(4) = 2$ and $c(5) = 3$, and Mader [74] proved that $c(6) = 4$ and $c(7) = 5$. In view of this it is not surprising that I wrote (EGT, Chapter VIII, p. 378): "It is rather hard not to conjecture that $c(p)$ is always $p - 2$." However, I did stop short of conjecturing $c(p) = p - 2$ for all $p \geqslant 2$. This was just as well, since a little later Catlin, Erdős, and I [25] proved the following result, which implies that $c(p)$ is considerably larger than $p - 2$, provided p is large.

THEOREM 3. *Almost every random graph G of order n with probability $1/2$ of an edge is such that*

$$n\left((\log_2 n)^{1/2} + 4\right)^{-1} \leqslant \operatorname{ccl}(G) \leqslant n\left((\log_2 n)^{1/2} - 1\right)^{-1}. \quad □$$

Since the probability that a random graph $G \in \mathscr{G}(n, 1/2)$ has at least $n^2/4$ edges tends to $1/2$, we find that if $\varepsilon > 0$ and p is sufficiently large then there is a graph of order $n = \left\lceil (1-\varepsilon)p(\log_2 p)^{1/2} \right\rceil$ and size $\lceil n^2/4 \rceil$ whose contraction clique number is at most $p - 1$. Hence $c(p) \geqslant \frac{1}{4}(1-\varepsilon)p(\log_2 p)^{1/2}$ if p is large enough.

A version of Theorem 3 for a different edge probability gives an even better lower bound for $c(p)$. As stated in [25], if $0 < P < 1$ is fixed then almost every random graph G of order n with probability P of an edge is such that $\operatorname{ccl}(G) \sim n(\log_{1/Q} n)^{-1/2}$, where $Q = 1 - P$. As remarked by Thomason [83], on taking $P = 0.716$ the last relation has the following consequence.

COROLLARY 4. *If p is sufficiently large then $c(p) > 0.265p(\log_2 p)^{1/2}$.* □

Having thus disproved the possibility of $c(p) = p - 2$, what is the correct order of magnitude of $c(p)$? This question was answered independently by Kostochka [71] and Thomason [83]. Since the proof given by Thomason is simpler and gives a better constant (2.68 as opposed to 324), that is the one we shall present.

THEOREM 5. *If p is sufficiently large then $c(p) \leqslant 2.68p(\log_2 p)^{1/2}$.*

The proof of this theorem is based on four lemmas. The first of these is in the vein of the proof of Theorem 1. The peculiar choice of the constant α is not essential but does result in the best upper bound for $\overline{\lim}_{p\to\infty} c(p)p^{-1}(\log_2 p)^{-1/2}$ which can be obtained by the latter part of the proof.

Define $\alpha = 2.678\cdots$ by the relation $\alpha = 1 + \log 2\alpha$. For $r \in \mathbf{R}$ and $\alpha r \in \mathbf{N}$, $\alpha r \geqslant 3$, set

$$\mathscr{F}_r = \{G\colon |G| \leqslant \alpha r, 2\delta(G) - |G| \geqslant \lfloor r \rfloor - 3\}.$$

LEMMA 6. *If $G \in \mathscr{F}_r$ implies $G > K^{p-2}$ then $c(p) \leqslant \alpha r$.* □

This lemma shows that Theorem 5 follows if we prove that for $\varepsilon > 0$ and sufficiently large values of p, $G \in \mathscr{F}_r$ implies $G > K^{p-2}$, where $r = (1+\varepsilon)p(\log_2 p)^{1/2}$. The set $\mathscr{F}_r$ consists of rather dense graphs, i.e., graphs whose minimal degree is not much smaller than $(1+\alpha^{-1})/2$ times their order. How can we show that a large complete graph K^{p-2} is a minor of such a graph G? We have to find $p-2$ vertex disjoint connected subgraphs of G such that any two of these subgraphs is joined by an edge.

There are two requirements: the subgraphs have to be connected and they have to be joined to each other. Lemma 8 will take care of the first condition. It implies that by the addition of rather few vertices we can make our subgraphs connected. The real difficulty is finding many small disjoint subsets of vertices such that any two of them are joined by an edge. If we choose our subsets one by one (as we shall), it is advantageous to pick subsets which dominate as many vertices as possible. Furthermore, what we need is that many small sets of vertices dominate many (all but rather few) vertices of our graph, for then we can select one which not only dominates many vertices but is also joined to all previously selected sets. This will follow from Lemma 9.

Lemmas 8 and 9 depend on a simple lemma on domination in graphs.

LEMMA 7. *Let G be a graph of order n and let $\varnothing \neq U \subset V = V(G)$. Suppose $d(u) \geqslant d$ for all $u \in U$.*

(i) *G has a vertex which dominates at least $(d+1)/n$ of the vertices of U.*

(ii) *If $d \geqslant (n+m)/2$ and $m \leqslant n-3$ then there are at least $m+3$ vertices dominating at least half of the vertices of U.*

PROOF. For $v \in V$ define $E(v) = \{(v,u)\colon u \in U \text{ and } u = v \text{ or } uv \in E(G)\}$ and set

$$E^* = \bigcup_{v\in V} E(v) \subset V \times U. \tag{2}$$

By our assumption,

$$|E^*| \geqslant (d+1)|U|. \tag{3}$$

(i) Relations (2) and (3) imply that there is a vertex $v \in V$ satisfying

$$|E(v)| \geqslant (d+1)|U|/n.$$

(ii) Suppose $d \geqslant (n+m)/2$, $k < m+3 \leqslant n$, and W is a set of k vertices each of which dominates at least half of the vertices of U. All we have to show is that there is a vertex $v \in V \setminus W$ which also dominates at least half of the vertices of U. To see this, note that by (3),

$$\left| E^* - \bigcup_{w \in W} E(w) \right| \geqslant (d+1)|U| - k|U|,$$

so there is a vertex $v \in V \setminus W$ satisfying

$$|E(v)| \geqslant \frac{(d+1)-k}{n-k}|U| \geqslant \frac{(n+m)/2 - m - 1}{n-m-2} = \frac{1}{2}|U|. \quad \square$$

LEMMA 8. *Let G be a graph of order n and minimal degree at least $(n-1)/2$. Let $\varnothing \neq U \subset V(G)$ and $|U| \leqslant 2^k$ for some $k \in N$. Then there is a set W of at most k vertices such that $G[U \cup W]$ is connected.*

PROOF. We apply induction on k. For $k = 0$ there is nothing to prove so suppose that $k \geqslant 0$ and the assertion holds for smaller values of k.

By Lemma 7(i) there is a vertex x_1 which dominates all but a set U_0 of at most $2^{k-1} - 1$ vertices of U. Set $U_1 = \{x_1\} \cup U_0$. Since $|U_1| \leqslant 2^{k-1}$, by the induction hypothesis there is a set W_1 of at most $k-1$ vertices such that $G[U_1 \cup W_1]$ is connected. Set $W = W_1 \cup \{x_1\}$. Then $|W| \leqslant k$ and the graph $G[U \cup W] = G[(U \setminus U_0) \cup \{x_0\} \cup U_0 \cup W_1]$ is connected since $G[(U \setminus U_0) \cup \{x_0\}$ and $G[\{x_0\} \cup U_0 \cup W_1]$ are connected. $\square$

LEMMA 9. *Let G be a graph of order n and minimum degree at least $(n+m)/2$, where $1 \leqslant m \leqslant n-3$. Then there are at least $\binom{m+3}{k}$ k-subsets of $V(G)$ each of which dominates at least $n - \lfloor 2^{-k} n \rfloor$ vertices of G.*

PROOF. As in the proof of Lemma 8, even the greedy algorithm produces sufficiently many suitable k-sets. To get a k-set, choose its vertices one by one, always dominating as many vertices as possible. To fill in the details, suppose there are at least $\binom{m+3}{k-1}$ $(k-1)$-subsets of $V(G)$ that dominate at least $n - \lfloor 2^{-k+1} n \rfloor$ vertices of G. By Lemma 7(ii) each such $(k-1)$-set is contained in at least $(m+3) - (k-1) = m + 4 - k$ k-sets dominating at least

$$n - \lfloor 2^{-k+1} n \rfloor + \lceil \lfloor 2^{-k+1} n \rfloor / 2 \rceil \geqslant n - \lfloor 2^{-k} n \rfloor$$

vertices. As each such k-set is obtained from at most k subsets, we obtain at least

$$\binom{m+3}{k-1} \frac{m+4-k}{k} = \binom{m+3}{k}$$

appropriate k-sets. $\square$

Thomason [83] deduced Theorem 5 from the following rather technical result.

THEOREM 10. *Let $p \in \mathbb{N}$ and $r \in \mathbb{R}$ be such that $p \geqslant 4$, $\alpha r \in \mathbb{N}$ and there is a $k \in \mathbb{N}$ satisfying $5 \leqslant k \leqslant r - (p-3)l$ and*

$$(p-3)\binom{\lfloor 2^{-k} \alpha r \rfloor}{k} < \binom{\lfloor r \rfloor - (p-3)l}{k},$$

where $l = k + \lfloor \log_2 k \rfloor + 1$. Then $c(p) \leqslant \alpha r$.