Postgraduate Studies Programme
“Theoretical Mathematics”,
Department of Mathematics,
National and Kapodistrian University of Athens

Extremal Graph Theory: Basic Results

Nefeli Vidali
RN : 140401

Committee:
Thilikos Dimitrios M., Professor, Department of Mathematics, NKUA.
Kirousis Lefteris M., Professor, Department of Mathematics, NKUA.
Dracopoulos Michael C., Assistant Professor, Department of Mathematics, NKUA.

Advisor:
Thilikos Dimitrios M., Professor, Department of Mathematics, NKUA.

Μεταπτυχιακό Πρόγραμμα Ειδίκευσης
«Θεωρητικά Μαθηματικά»,
Τμήμα Μαθηματικών,
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Ακραία Γραφοθεωρία: Βασικά Αποτελέσματα

Νεφέλη Βιδάλη
ΑΜ : 140401

Τριμελής Επιτροπή:
Θηλυκός Δημήτριος Μ., Καθηγητής, Τμήμα Μαθηματικών, ΕΚΠΑ.
Κυρούσης Ελευθέριος Μ., Καθηγητής, Τμήμα Μαθηματικών, ΕΚΠΑ.
Δρακόπουλος Μιχάλης Χ., Επίκουρος Καθηγητής, Τμήμα Μαθηματικών, ΕΚΠΑ.

Επιβλέπων:
Θηλυκός Δημήτριος Μ., Καθηγητής, Τμήμα Μαθηματικών, ΕΚΠΑ.

Η παρούσα Διπλωματική Εργασία

εκπονήθηκε στα πλαίσια των σπουδών

για την απόκτηση του

Μεταπτυχιακού Διπλώματος Ειδίκευσης

«Θεωρητικά Μαθηματικά»

που απονέμει το

Τμήμα Μαθηματικών

του

Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών

Εγκρίθηκε την .... από Εξεταστική Επιτροπή

αποτελούμενη από τους:

Ονοματεπώνυμο	**Βαθμίδα**	**Υπογραφή**
1. Θηλυκός Δημήτριος	....	....
2. Κυρούσης Ελευθέριος	....	....
3. Δρακόπουλος Μιχάλης	....	....

Abstract

In this thesis, we take a general overview of extremal graph theory, investigating common techniques and how they apply to some of the more celebrated results in the field. The first chapter is an introduction to the subject and some preliminary definitions and results. The second chapter concerns substructures in dense graphs and focuses on important results such as Turán's theorem, Szemerédi's regularity lemma and the Erdős-Stone-Simonovits theorem. The third chapter concerns substructures in sparse graphs and investigates conditions which force a graph to contain a certain minor or topological minor. The fourth and final chapter is an introduction to the extremal theory of r-uniform hypergraphs and consists of a presentation of results concerning the conditions which force them to contain a complete r-graph and a Hamiltonian cycle.

Σύνοψη

Η παρούσα διπλωματική εργασία έχει σκοπό να παρουσιάσει μία σφαιρική εικόνα της θεωρίας των ακραίων γραφημάτων, διερευνώντας κοινές τεχνικές και τον τρόπο που εφαρμόζονται σε κάποια από τα πιο διάσημα αποτελέσματα του τομέα. Το πρώτο κεφάλαιο είναι μία εισαγωγή στο θέμα και κάποιοι προαπαιτούμενοι ορισμοί και αποτελέσματα. Το δεύτερο κεφάλαιο αφορά υποδομές πυκνών γραφημάτων και εστιάζει σε σημαντικά αποτελέσματα όπως είναι το θεώρημα του Turán, το λήμμα κανονικότητας του Szemerédi και το θεώρημα των Erdős-Stone-Simonovits. Το τρίτο κεφάλαιο αφορά υποδομές αραιών γραφημάτων και ερευνά συνθήκες που εξαναγκάζουν ένα γράφημα να περιέχει ένα δοθέν έλασσον ή τοπολογικό έλασσον. Το τέταρτο και τελευταίο κεφάλαιο είναι μία εισαγωγή στην θεωρία ακραίων r-ομοιόμορφων υπεργραφημάτων και περιέχει αποτελέσματα που αφορούν συνθήκες οι οποίες τα εξαναγκάζουν να περιέχουν πλήρη r-γραφήματα και Χαμιλτονιανούς κύκλους.

Acknowledgements

Foremost, I would like to express my heartfelt gratitude to my advisor, Professor Dimitrios M. Thilikos, for his valuable guidance, encouragement and helpful critique.

I would also like to give my sincere thanks to the other members of my thesis committee, Professor Lefteris M. Kirousis and Assistant Professor Michael C. Dracopoulos, for generously providing their time and support.

Lastly, I would like to express my gratefulness to my friends and family for the emotional support they offered me, not only during the writing of this work but throughout my entire experience in this postgraduate programme.

Contents

Chapter 1

Introduction

> Extremal graph theory, in its strictest sense, is a branch of graph theory developed and loved by Hungarians.
>
> ---
>
> B. Bollobás

1.1 Overview

Extremal graph theory is a branch of graph theory that seeks to explore the properties of graphs that are in some way extreme, where extremality can be taken with respect to different graph invariants, such as order, girth, chromatic number etc. More abstractly, it studies how global parameters of a graph influence its local substructures. For example, a simple extremal graph theory question is "which acyclic graphs on n vertices have the maximum number of edges?". The extremal graphs for this question are trees on n vertices, which have $n-1$ edges. More generally, a typical question is the following.

> Given a graph property P, an invariant u and a set of graphs $\mathcal{H}$, we wish to find the minimum value of m such that every graph in $\mathcal{H}$ which has u larger than m possesses property P.

A classical extremal graph theoretic result, and the one that started the study in the field, is Turán's theorem, proved in 1941, which reveals not only the edge-density but also the structure of those graphs that are extremal without a complete subgraph of some fixed size. Another crucial result for the subject appeared in 1975 when Szemerédi proved his regularity lemma, which became a vital tool in attacking extremal problems.

The material of this thesis falls into two main parts. The first one concerns substructures in *dense* graphs, that is, graphs whose number of edges is about quadratic in their number of vertices. The number $|E(G)|/\binom{|V(G)|}{2}$, the proportion of its potential edges that G actually has, is the *edge-density* of G. The question of exactly which edge-density is needed to force a given subgraph is a classical graph problem. We shall concentrate on a few important results, namely, Turán's theorem, the Erdős-Stone-Simonovits theorem, and Szemerédi's regularity lemma. Our main purpose in proving the regularity lemma will be to give a proof of the Erdős-Stone-Simonovits theorem, although it is worth noting that its use is widespread throughout extremal graph theory. We shall also consider extremal graphs that do not contain paths or cycles of a given length.

The second main part concerns substructures in *sparse* graphs, that is, graphs whose number of edges is about linear in their number of vertices. In this part we shall investigate which conditions can force a graph G to contain some given graph H as a minor or topological minor. We shall also consider Hadwiger's conjecture, which is one of the most famous open problems in graph theory.

The final chapter is an introduction to hypergraphs and a brief summary of results that concern extremal r-uniform hypergraphs.

1.2 Definitions

Before we begin our study of extremal graphs, we will provide the necessary definitions and notation. All graphs considered in this thesis are finite and simple, that is, undirected and loop-free with no multiple edges. Given a graph $G = (V, E)$, we denote by $V(G)$ and $E(G)$ its set of vertices and edges, respectively. We denote by xy an edge with endpoints $x, y \in V(G)$.

Two vertices connected by an edge are called *adjacent*, and two edges that share an endpoint are called *incident*.

A *clique* is a set of pairwise adjacent vertices, whereas an *independent set* is a set of pairwise non-adjacent vertices. A maximum independent set is an independent set of largest possible size. This size is called the independence number of a graph G, and denoted by $\alpha(G)$. An *independent set of edges*, often also called a *matching*, is a set of edges no two of which share an endpoint. A *perfect matching* is a matching in which every vertex of the graph is incident to exactly one edge of the matching. A *vertex cover* of a graph is a set of vertices such that each edge of the graph is incident to at least one vertex of the set.

If $G = (V, E)$ and $G' = (V', E')$ are two graphs, a map $\phi\colon V \to V'$ is a *homomorphism* from G to G' if it preserves the adjacency of vertices. If ϕ is bijective and its inverse ϕ^{-1} is also a homomorphism, then ϕ is an *isomorphism*. A class of graphs that is closed under isomorphism is called a *graph property*. A map taking graphs as arguments is called a *graph invariant* if it assigns equal values to isomorphic graphs.

If $G' = (V', E')$ is a graph such that $V' \subseteq V$ and $E' \subseteq E$, then G' is called a *subgraph* of G and we write $G' \subseteq G$. Given a graph H, a graph G which does not contain it as a subgraph is called H*-free*. If $G' \subseteq G$ and G' contains all the edges $xy \in E$ with $x, y \in V'$, then G' is called an *induced subgraph* of G; we say that V' *induces* or *spans* G' in G and write $G' = G[V']$. A graph on n vertices induced by a clique is called a *complete* graph and is denoted by K_n.

If $G = (V, E)$ and $U \subseteq V$, we define $G \backslash U = G[V \backslash U]$, that is, $G \backslash U$ is obtained from G by removing all the vertices in U and their incident edges. If $U = \{v\}$ is a singleton, we simply write $G \backslash v$. Similarly, if F is a subset of E, we define $G \backslash F = (V, E \backslash F)$, and write $G \backslash e$ if $F = \{e\}$ is a singleton.

The complement $\overline{G}$ of a graph $G = (V, E)$ is a graph on V with edge set $\binom{V}{2} \setminus E$. The line graph $L(G)$ of G is the graph on E in which $e, f \in E$ are adjacent as vertices if and only if they are adjacent as edges in G.

Given an $x \in V(G)$, we denote by $N_G(x)$ the set of all vertices adjacent to x. More generally, for $U \subseteq V(G)$, the set of neighbours in $V(G) \backslash U$ of vertices in U is denoted by $N_G(U)$. The size of $N_G(x)$ is called the *degree* of x and is denoted by $d_G(x)$. When the reference is clear, we may drop the index and simply write $N(x)$, $N(U)$ and $d(x)$, respectively. The number $\delta(G) = \min\{d(v) : v \in V(G)\}$ is called the *minimum degree* of G, the number $\Delta = \max\{d(v) : v \in V(G)\}$ is called its *maximum degree*, and the number $d(G) = \frac{1}{|V(G)|} \sum_{v \in V(G)} d(v)$ is called its *average degree*. The ratio $\frac{|E(G)|}{|V(G)|}$ is often denoted by $\epsilon(G)$.

A path is a graph $P = (V, E)$ of the form

$$V = \{x_0, x_1, \ldots, x_k\} \quad \text{and} \quad E = \{x_0x_1, x_1x_2, \ldots, x_{k-1}x_k\},$$

where $x_i \neq x_j$ for every $i \neq j$. The vertices x_0 and x_k are the *endpoints* of the path, and the vertices x_i for $1 \leqslant i \leqslant k-1$ are its *inner vertices*. The number of edges of a path is its *length* and a path of length k is denoted by P_k. We will refer to a path P by the natural sequence of its vertices, writing, say, $P = x_0x_1 \ldots x_k$ and calling P a path *from* x_0 *to* x_k, or simply an $x_0 - x_k$ *path*. The *distance* $d_G(x, y)$ (or, as before, simply $d(x, y)$) in G of two vertices x and y is the length of a shortest (x, y)-path in G. Given two sets of vertices A and B, we call $P = x_0 \ldots x_k$ an (A, B)*-path* if $V(P) \cap A = \{x_0\}$ and $V(P) \cap B = \{x_k\}$. When $A = \{a\}$ is a singleton, we may simply speak of an (a, B)-path. Two or more paths are called *independent* or *disjoint* if none of them contains an inner vertex of another. A set of (a, B)-paths is called an (a, B)-*fan* if any two of the paths have only a in common.

If $P = (V, E)$ is the path $x_0 \dots x_{k-1}$ and $k \geqslant 3$, then the graph $C = (V, E \cup \{x_{k-1}, x_0\})$ is called a *cycle*. As with paths, we will refer to a cycle by its (cyclic) sequence of vertices, writing $x_0 \dots x_{k-1}x_0$. The *length* of a cycle is the number of its edges (or vertices). The cycle of length k is called a k-*cycle* and is denoted by C^k.

The minimum length of a cycle contained in a graph G is the *girth* of G, denoted by $g(G)$. An edge which joins two vertices of a cycle but is not itself an edge, is a *chord* of the cycle. Thus, an *induced cycle* in G, that is, a cycle in G forming an induced subgraph, is one that has no chords.

A graph $G = (V, E)$ is *connected* if any two of its vertices can be joined by a path in G. A subgraph-maximal connected subgraph of G is a *component* of G. Clearly, the components are induced subgraphs and their vertex sets partition V.

If $A, B \subseteq V$ and $X \subseteq V \cup E$ are such that every (A, B)-path in G contains a vertex or an edge from X, we say that X *separates* the sets A and B in G. We say that X separates two vertices a, b, if it separates the sets $\{a\}$ and $\{b\}$ but $a, b \notin X$, and that X separates G if it separates some two vertices of G. A separating set of vertices is called a *separator*.

The unordered pair $\{A, B\}$ is a *separation* of G if $A \cup B = V$ and G has no edge between $A \backslash B$ and $B \backslash A$. if both $A \backslash B$ and $B \backslash A$ are non-empty, the separation is *proper*. The number $|A \cap B|$ is the *order* of the separation $\{A, B\}$, and the sets A and B are its *sides*.

Given a positive integer k, a graph $G = (V, E)$ is called k-*connected* if $|V| > k$ and $G \backslash X$ is connected for every set $X \subseteq V$ with $|X| < k$. The greatest integer k such that G is k-connected is the *connectivity* of G, denoted by $\kappa(G)$.

An *acyclic* graph, that is, a graph that contains no cycles, is called a *forest*. A connected forest is called a *tree*. The vertices of degree 1 in a tree are its *leaves*, and the others are its *inner vertices*.

Let $r \geqslant 2$ be an integer. A graph $G = (V, E)$ is called r-*partite* if V admits a partition into r classes such that every edge of G has its ends in different classes: vertices in the same class must not be adjacent. When $r = 2$, the graph H is called *bipartite*.

An r-partite graph in which every two vertices from different partition classes are adjacent is called *complete*. We denote by K_s^r the complete r-partite graph in which every partition class contains exactly s vertices, and by $K_{s,t}$ the complete bipartite graph in which one partition class contains s vertices and the other contains

t vertices.

A *subdivision* of a fixed graph X is any graph obtained from X by replacing some (or all) of its edges with new paths between their ends, so that none of these paths as an inner vertex in $V(X)$ or on another path. When G is a subdivision of X, we say that G is a TX. The original vertices of X are the *branch vertices* of the TX, while its new vertices are called *subdividing vertices*. If a graph Y contains a TX as a subgraph, then X is a *topological minor* of Y.

A graph G is an IX if its vertex set admits a partition $\{V_x : x \in V(X)\}$ into connected subsets V_x such that distinct vertices $x, y \in V(X)$ are adjacent in X if and only if G contains a $V_x - V_y$ edge. The sets V_x are the *branch sets* of the IX. If a graph Y contains an IX as a subgraph, then X is a *minor* of Y.

Thus, X is a minor of Y if and only if there is a map ϕ from a subset of $V(Y)$ onto $V(X)$ such that for every vertex $x \in V(X)$ its inverse image $\phi^{-1}(x)$ is connected in Y and for every edge xx' of X there is an edge in Y between the branch sets $\phi^{-1}(x)$ and $\phi^{-1}(x')$ of its ends. If the domain of ϕ is all of $V(Y)$ and $xx' \in E(X)$ whenever $x \neq x'$ and Y has an edge between $\phi^{-1}(x)$ and $\phi^{-1}(x')$ (so that Y is an IX), we call ϕ a *contraction* of Y onto X.

If $e = xy \in E(G)$, we denote by G/e the graph obtained from G by *contracting* the edge e into a new vertex v_e, which becomes adjacent to all the former neighbours of x and y.

An *embedding* of G in H is an injective map $\phi\colon V(G) \to V(H)$ that preserves the kind of structure we are interested in. Thus, ϕ embeds G in H "as a subgraph" if it preserves the adjacency of vertices, "as an induced subgraph" if it preserves both adjacency and non-adjacency etc.

A graph is called *planar* if it can be embedded in the plane, that is, it can be drawn on the plane in such a way that its edges intersect only at their endpoints. A planar graph s called maximal if it cannot be extended to a larger planar graph by adding an edge. A planar graph with $n \geqslant 3$ vertices is maximally planar if and only if it has $3n - 6$ edges.

A *face* of a plane embedding of a planar graph is a connected component of the complement of the graph. An *outer face* is a face with infinite area. Any other face is an *inner face*. A plane embedding of a finite planar graph can only have one outer face.

A *vertex colouring* of a graph $G = (V, H)$ is a map $c : V \to S$ such that $c(v) \neq c(w)$ whenever v and w are adjacent. The elements of the set S are called the available *colours*. The smallest integer k such that G has a *k-colouring*, that is, a vertex colouring $c : V \to \{1, \dots, k\}$, is called the *chromatic number* of G and is denoted by $\chi(G)$. If $\chi(G) = k$, then G is called *k-chromatic*, whereas if $\chi(G) \leqslant k$, then it is called *k-colourable*. A k-colouring is basically a vertex partition into k independent sets, which we call *colour classes*; the non-trivial 2-colourable graphs, for example, are precisely the bipartite graphs.

An *edge-colouring* of $G = (V, E)$ is a map $c : E \to S$ with $c(e) \neq c(f)$ for any incident edges e, f. The smallest integer k for which a *k-edge-colouring* exists, that is, an edge-colouring $c : E \to \{1, \dots, k\}$ is the *edge-chromatic number*, or *chromatic index* of G, and is denoted by $\chi'(G)$.

1.3 Preliminary results

We shall now provide a few basic results that will be useful throughout our discussion.

Proposition 1.3.1. *Every graph G with at least one edge has a subgraph H with $\delta(H) > \epsilon(H) \geqslant \epsilon(G)$.*

Proof. Let us repeatedly delete a vertex whenever it has degree less than ϵ. After each such deletion, the number of vertices decreases by 1 and the number of edges by at most ϵ, so the overall ratio ϵ of edges to vertices will not decrease.

Formally, we construct a sequence $G = G_0 \supseteq G_1 \supseteq \dots$ of induced subgraphs of G as follows. If G_i has a vertex v_i of degree $d(v_i) \leqslant \epsilon(G_i)$, we let $G_{i+1} = G_i \backslash v_i$. If not, we terminate our sequence and set $H = G_i$. By the choices of v_i, we have $\epsilon(G_{i+1}) \geqslant \epsilon(G_i)$ for all i and, therefore, $\epsilon(H) \geqslant \epsilon(G)$.

Since $\epsilon(K_1) = 0 < \epsilon(G)$, none of the graphs in our sequence is trivial, so in particular $H \neq \emptyset$. The fact that H has no vertex suitable for deletion thus implies $\delta(H) > \epsilon(H)$, as claimed. □

Proposition 1.3.2. *Every graph G contains a path of length $\delta(G)$ and a cycle of length at least $\delta(G) + 1$ (provided that $\delta(G) \geqslant 2$).*

Proof. Let $x_0 \dots x_k$ be a longest path in G. Then, all the neighbours of x_k lie on this path. Hence $k \geqslant d(x_k) \geqslant \delta(G)$. If $i < k$ is minimal with $x_i x_k \in E(G)$, then $x_i \dots x_k x_i$ is a cycle of length at least $\delta(G) + 1$. □

The following proposition states that a large minimum degree implies the existence of a highly connected subgraph.

Proposition 1.3.3 (Mader's theorem)**.** *Let $k \in \mathbb{N}$. Every graph G with $d(G) \geqslant 4k$ has a $(k+1)$-connected subgraph H such that $\epsilon(H) < \epsilon(G) - k$.*

Proof. Consider the subgraphs $G' \subseteq G$ such that

$$|V(G')| \geqslant 2k \qquad \text{and} \qquad |E(G')| > \epsilon(G)(|V(G')| - k) \tag{1}$$

Such graphs G' exist since G is one; let H be one of smallest order.

No graph G' as in (1) can have order exactly $2k$, since this would imply that $|E(G')| > \epsilon(G)k \geqslant 2k^2 > \binom{|V(G')|}{2}$. The minimality of H therefore implies that $\delta(H) > \epsilon(G)$: otherwise we could delete a vertex of degree at most $\epsilon(G)$ and obtain a graph $G' \subsetneq H$ still satisfying (1). In particular, we have $|V(H)| \geqslant \epsilon(G)$. Dividing the inequality $|E(H)| > \epsilon(G)|V(H)| - \epsilon(G)k$ from (1) by $|V(H)|$ therefore yields $\epsilon(H) > \epsilon(G) - k$, as desired.

It remains to show that H is $(k+1)$-connected. If not, then H has a proper separation $\{U_1, U_2\}$ of order at most k; put $H[U_i] =: H_i$. Since any vertex $v \in U_1 \backslash U_2$ has all its $d(v) \geqslant \delta(H) > \epsilon(G)$ neighbours from H in H_1, we have $|V(H_1)| \geqslant \epsilon(G) \geqslant 2k$. Similarly, $|V(H_2)| \geqslant 2k$. As by the minimality of H neither H_1 nor H_2 satisfies (1), we further have

$$|E(H_1)| \leqslant \epsilon(G)(|V(H_i)|0 - k)$$

for $i = 1, 2$. But then

$$\begin{aligned}
|E(H)| &\leqslant |E(H_1)| + |E(H_2)| \\
&\leqslant \epsilon(G)(|V(H_1)| + |V(H_2)| - 2k) \\
&\leqslant \epsilon(G)(|V(H)| - k) \qquad (\text{as } |V(H_1) \cap V(H_2)| \leqslant k),
\end{aligned}$$

which contradicts (1) for H. □

The following theorem is one of the most fundamental results in graph theory.

Theorem 1.3.4 (Menger's theorem)**.** *Let $G = (V, E)$ be a graph and $A, B \subseteq V$. Then the minimum number of vertices separating A from B in G is equal to the maximum number of disjoint (A, B)-paths in G.*

Proof. We shall denote by k the minimum number of vertices separating A from B. Clearly, G cannot contain more than k disjoint (A, B)-paths, so our task is to show that such k paths exist. If $e \in E$, we denote by G/e the graph obtained by contracting the edge e, and by v_e the resulting new vertex.

We apply induction on $|E|$. If G has no edge, then $|A \cap B| = k$ and we have k trivial (A, B)-paths. So we assume that G has an edge $e = xy$. If G has no k disjoint (A, B)-paths, then neither does G/e; here, we count the contracted v_e as an element of A (respectively B) in G/e if in G at least one of x, y lies in A (respectively B). By the induction hypothesis, G/e contains an $A - B$ separator Y of fewer than k vertices. Among these must be the vertex v_e, since otherwise $Y \subseteq V$ would be an $A - B$ separator in G. Then $X = (Y \backslash \{v_e\}) \cup \{x, y\}$ is an $A - B$ separator in G of exactly k vertices.

We now consider the graph $G \backslash e = (V, E \backslash \{e\})$. Since $x, y \in X$, every $A - X$ separator in $G \backslash e$ is also an $A - B$ separator in G and hence contains at least k vertices. So by induction there are k disjoint (A, X)-paths in $G \backslash e$, and similarly there are k disjoint (X, B)-paths in $G \backslash e$. As X separates A from B, these two path systems do not meet outside X, and can thus be combined to k disjoint (A, B)-paths. $\square$

When mentioning Menger's theorem, what is often meant is its following corollary, which is sometimes referred to as the global version of Menger's theorem.

Theorem 1.3.5 (Global version of Menger's theorem)**.** *A graph is k-connected if and only if it contains k independent paths between any two vertices.*

Proof. If a graph G contains k independent paths between any two vertices, then $|V(G)| > k$ and G cannot be separated by fewer than k vertices; thus G is k-connected.

Conversely, suppose that G is k-connected (and, in particular, has more than k vertices) but contains vertices a, b not joined by k independent paths. Note that,by applying Menger's theorem to $G \backslash \{a, b\}$ with $A = N_G(a)$ and $B = N_G(b)$, we have that the minimum number of vertices separating a from b is equal to the maximum number of independent (a, b) paths in G. We can, therefore, conclude that a and b are adjacent. Indeed, let $G' = G \backslash ab$. Then G' contains at most $k - 2$ independent (a, b)-paths. Again, we can separate a and b in G' by a set X of at most $k - 2$ vertices. As $|V(G)| > k$, there is at least one further vertex $v \notin X \cup \{a, b\}$ in G. Now X separates v in G' from either a or b, say from a. But then $X \cup \{b\}$ is a set of at most $k - 1$ vertices separating v from a in G, contradicting the k-connectedness of G. $\square$

Proposition 1.3.6. *Every graph G with m edges contains a bipartite subgraph with $\lceil \frac{m}{2} \rceil$ edges.*

Proof. Let $A \cup B$ be a bipartition of $V(G)$ with the maximum number of edges between A and B. Any vertex $a \in A$ has at least as many neighbours in B as in A (otherwise, we could move v to B, thereby increasing the number of edges between A and B). The same holds for the vertices in B. Therefore, the number of edges between A and B is at least $\frac{m}{2}$. Deleting the extra edges between A and B yields the desired result. $\square$

The following result is known as Hall's theorem, often also called Hall's marriage theorem, and gives us a sufficient condition for a graph to contain a perfect matching[1].

Theorem 1.3.7 (Hall's theorem). *Let G be a bipartite graph with partition sets A and B of equal size. If, for every $U \subseteq A$, $|N_G(U)| \geqslant |U|$, then G contains a perfect matching.*

Proof. Let $|A| = |B| = n$. We will proceed by induction on n. Clearly, the result is true for $n = 1$. We therefore assume that it is true for $n - 1$ and prove it for n.

If $|N_G(U)| \geqslant |U| + 1$ for every non-empty proper subset U of A, pick an edge ab of G and consider the graph $G' = G \backslash ab$. Then, every non-empty subset U of $A \backslash \{a\}$ satisfies

$$|N_{G'}(U)| \geqslant |N_G(U)| - 1 \geqslant |U|.$$

Therefore, there is a perfect matching between $A \backslash \{a\}$ and $B \backslash \{b\}$. Adding the edge from a to b gives the full matching.

Suppose, on the other hand, that there is some non-empty proper subset U of A for which $|N_G(U)| = |U|$, and let $V = N_G(U)$. By induction, since Hall's condition holds for every subset of U, there is a matching between U and V. But Hall's condition also holds between $A \backslash U$ and $B \backslash V$. If this weren't the case, there would be some W in $A \backslash U$ with fewer than $|W|$ neighbours in $B \backslash V$. Then $W \cup U$ would be a subset of A with fewer than $|W \cup U|$ neighbours in B, contradicting our assumption. Therefore, there is a perfect matching between $A \backslash U$ and $B \backslash V$. Putting the two matchings together completes the proof. □

Proposition 1.3.8. *If H is a minor of G such that $\Delta(H) \leqslant 3$, then H is a topological minor of G.*

Proof. In order to show that H is a topological minor of G, we need to show that there is a subdivision of H that is isomorphic to a subgraph of G. This means that H can be obtained from G by a sequence of deletions of vertices and edges, and then contractions of edges of the form uv where one of the vertices u or v has degree exactly 2.

When forming H from G as a minor, we can do this so as to perform all deletions of vertices and edges first, which leaves us with a graph G'. Note that while we might want to later contract an edge uv where one of u or v have degree 1, we can convert this operation to just deleting the degree 1 vertex, so we do this here instead.

[1] The theorem has many other interesting "non-marital" applications as well. For example, take a standard deck of cards, and deal them out into 13 piles of 4 cards each. Then, using the marriage theorem, we can prove that it is always possible to select exactly 1 card from each pile, such that the 13 selected cards contain exactly one card of each rank (Ace, 2, 3, etc.).

In other words, we first construct a graph G' by deletions of vertices and edges that contains H as a minor, such that H is obtained from G' by contracting edges uv both of whose endpoints have degree at least 2.

If $G' = H$, we are done. Otherwise, we need to perform some edge contractions. Since $\Delta(H) \leqslant 3$, we can never contract an edge whose endpoints both have degree 3. This is because otherwise we would create a vertex of degree 4 at this step, and since we are only now contracting edges joining two vertices of degree at least 2, H would have a vertex of degree 4.

Therefore, in forming H from G', we are only contracting edges joining two vertices at least one of which has degree 2. Thus H is a topological minor of G. □

Moving on to planar graphs, it is worth noting that this theory contains a multitude of interesting and beautiful results but, as this thesis is not concerned with them, we shall only provide what is arguably the most important ones, that is, their characterization in terms of forbidden graphs.

The first such theorem was proved by Kuratowski [37], who characterized planar graphs in terms of forbidden topological minors.

Theorem 1.3.9 (Kuratowski 1930). *A graph is planar if and only if its topological minors include neither K_5 nor $K_{3,3}$*

Wagner [61], on the other hand, characterized planar graphs in terms of minors. Before we state Wagner's theorem, we will prove the following simple results.

Lemma 1.3.10. *If a planar graph G is 2-connected, then every face is bounded by a cycle.*

Proof. Assume there is a face F not bounded by a cycle. Choose a vertex v so that the boundary walk of F passes through v twice. Then we may draw a closed curve starting and ending at v with interior contained in F. This curve separates the plane into two components, each of which must contain a vertex of G, so we conclude that v is a cut vertex of G, that is, a vertex whose removal disconnects G. Thus, G is not 2-connected, a contradiction. □

Lemma 1.3.11. *Let C be a cycle and let $X, Y \subseteq V(C)$. Then one of the following holds.*

(i) $|X| \leqslant 1$ *or* $|Y| \leqslant 1$;

(ii) $X = Y$;

(iii) *there exists* $x_1, x_2 \in X$ *and* $y_1, y_2 \in Y$ *so that* x_1, y_1, x_2, y_2 *are distinct and occur in* C *in this order;*

(iv) *there exist* $u, v \in V(C)$ *so that if* P *and* Q *are the two paths of* C *between* u *and* v*, then* $X \subseteq V(P)$ *and* $Y \subseteq V(Q)$*.*

Proof. We assume that $|X|, |Y| \geqslant 2$ and that $X \neq Y$, as otherwise one of (i) or (ii) holds. By possibly switching X and Y, we may assume that $X \backslash Y \neq \emptyset$ and choose $x_1 \in X \backslash Y$. Let y_1 and y_2 be the first vertices in Y clockwise and counter-clockwise, respectively, from x_1. Since $|Y| \geqslant 2$, we have that $y_1 \neq y_2$. Let P and Q be the two paths of C between y_1 and y_2 and assume that x_1 lies on P. If $X \subseteq V(P)$, then (iv) holds, otherwise (iii) holds. □

Theorem 1.3.12 (Wagner 1937)**.** *A graph is planar if and only of its minors include neither* K_5 *nor* $K_{3,3}$*.*

Proof. It is easy to see that if a graph is planar then the removing of any edge or vertex or the contraction of any edge results in a graph that is again planar. It is also easy to check that neither K_5 nor $K_{3,3}$ are planar. Thus, if a graph is planar then it contains neither K_5 nor $K_{3,3}$ as a minor.

For the converse, let G be a graph with no K_5 or $K_{3,3}$ minor. We will proceed by induction on $|V(G)|+|E(G)|$. If G is not connected then, by applying the induction hypothesis to each component, we obtain a plane embedding of each component and, by combining these, we get a plane embedding of G. We may, therefore, assume that G is connected.

If G is not 2-connected, then we may choose a separation $\{H_1, H_2\}$ of G (unless $|V(G)| \leqslant 2$, in which case the result is trivial). Since H_1 and H_2 have no K_5 or $K_{3,3}$ minor, by applying the induction hypothesis we may embed them in the plane. Combining these we obtain a plane embedding of G. We may, therefore, further assume that G is 2-connected.

If G is not 3-connected, then we may choose a separation $\{H_1, H_2\}$ of G where $V(H_1) \cap V(H_2) = \{u, v\}$ (unless $|V(G)| \leqslant 3$, in which case the result is trivial). Add a new edge uv to H_1 and H_2 to form the graphs H_1' and H_2'. Choose vertices $z_1 \in V(H_1) \backslash V(H_2)$ and $z_2 \in V(H_2) \backslash V(H_1)$ and apply Menger's theorem to choose two internally disjoint paths from z_1 to z_2. It follows that H_1' is a minor of G; to see this, delete all vertices and edges of H_2 not in the two paths chosen above and then contract all but one of the edges in H_2 that remain. Similarly, H_2' is a minor of G. This means that H_1' and H_2' have no K_5 or $K_{3,3}$ minor so, by induction, we may choose plane embeddings of them. By combining these on the edge uv and then removing it, we obtain a plane embedding of G. We may, therefore, further assume that G is 3-connected.

Lastly, let us assume that G has an edge uv so that $G\backslash\{u,v\}$ is not 2-connected. Choose a separation $\{H_1, H_2\}$ of $G\backslash uv$ where $V(H_1) \cap V(H_2) = \{u, v, w\}$. For $i = 1, 2$, let H_i' be the graph obtained from H_i by adding a new vertex adjacent to u, v and w. Then choose vertices $z_1 \in V(H_1)\backslash V(H_2)$ and $z_2 \in V(H_2)\backslash V(H_1)$ and apply Menger's theorem to find three internally disjoint paths from z_1 to z_2. It follows that H_1' is a minor of G; to see this, delete all vertices and edges of H_2 not on the three paths chosen above, and then contract all edges in H_2 except for one on each of these paths. Similarly, H_2' is a minor of G. It follows that H_1' and H_2' have no K_5 or $K_{3,3}$ minor so, by induction, we may choose plane embeddings of them. By combining these, we obtain a plane embedding of G. Thus, we may assume that G has no such edge uv.

We now have sufficient connectivity in order to proceed. Choose an edge $e = xy$ of G and let $G' = G/e$. Let x be the vertex we obtain by contracting xy and let $G'' = G\backslash\{x, y\} = G\backslash z$. Now, G' has no K_5 or $K_{3,3}$ minor so, by induction, we may choose a plane embedding of it. Furthermore, it follows from our assumptions that G'' is 2-connected so, by Lemma 1.3.10, the face of G'' that contains z is bounded by a cycle C. Thus, all neighbours of x and y in G lie on C. We will now embed G'' on the plane and try to extend this to a plane embedding of G. Note that $N_G(x)$ and $N_G(y)$ are subsets of $V(C)$.

We now apply Lemma 1.3.11 to C for $X = N_G(x)$ and $Y = N_G(y)$. If either (i) or (ii) holds and $|X| = |Y| = 2$, then we may extend our plane embedding of G'' to one of G. If (ii) holds with $|X| = |Y| \geqslant 3$, then G contains a K_5 minor, contradicting our assumption. If (iii) holds, then G contains a $K_{3,3}$ minor, again contradicting our assumption. Lastly, if (iv) holds, then we may again extend our plane embedding of G'' to one of G, thus proving the desired result. □

The following theorem is one of the most well-known results in graph theory.

Theorem 1.3.13 (Four Colour Theorem). *Every planar graph is 4-colourable.*

Lastly, we shall prove a couple of results concerning vertex and edge colourings.

Proposition 1.3.14. *Every k-chromatic graph has a k-chromatic subgraph of minimum degree at least $k - 1$.*

Proof. Given G with $\chi(G) = k$, let $H \subseteq G$ be minimal with $\chi(H) = k$. If H had a vertex v of degree $d_H(v) \leqslant k - 2$, we could extend a $(k-1)$-colouring of $H\backslash v$ to one of H, contradicting the choice of H. □

Theorem 1.3.15 (Vizing). *For any finite, simple graph G,*

$$\Delta(G) \leqslant \chi'(G) \leqslant \Delta(G) + 1.$$

Proof. The lower bound is trivial since, if G has a vertex u of degree d, at least d edges have u as an endpoint and cannot be coloured with fewer than d colours.

For the upper bound, we proceed by induction on $|E(G)|$. The result clearly holds when $|E(G)| = 0$, so suppose that $|E(G)| > 0$ and that a proper $(\Delta+1)$-edge-colouring exists for all $G\backslash xy$, where $xy \in E(G)$ and $\Delta = \Delta(G)$.

We say that a colour $\alpha \in \{1, \dots, \Delta+1\}$ is *absent* at a vertex x with respect to a proper $(\Delta+1)$-edge-colouring c, if $c(xy) \neq \alpha$ for all $y \in N(X)$. Let *α/β-path from* x denote the unique maximal path which starts from x with an α-coloured edge and alternates colours of edges (the second edge has colour β, the third has colour α, and so on). Note that if c is a proper $(\Delta+1)$-edge-colouring of G, then every vertex has an absent colour with respect to c.

Suppose, towards a contradiction, that no proper $(\Delta+1)$-edge-colouring of G exists. This is equivalent to the following statement.

> Let $xy \in E(G)$ and c be an arbitrary proper $(\Delta+1)$-edge-colouring of $G\backslash xy$ and suppose that α is absent in x and β is absent in y with respect to c. (1)
> Then the α/β-path from y ends in x.

This is equivalent because if (1) doesn't hold, then we can interchange the colours α and β on the α/β-path and set the colour of xy to be α, thus creating a proper $(\Delta+1)$-edge-colouring of G from c. Conversely, if a proper $(\Delta+1)$-edge-colouring exists, then we can delete an edge, restrict the colouring and (1) won't hold either.

Now, let $xy_0 \in E(G)$ and let c_0 be a proper $(\Delta+1)$-edge-colouring of $G\backslash xy_0$ and suppose that α is absent in x with respect to c_0. We define $y_0, \dots, y_k$ to be a maximal sequence of neighbours of x such that $c_0(xy_i)$ is absent in y_{i-1} with respect to c_0 for all $0 < i \leqslant k$. We also define a sequence of colours $c_1, \dots, c_k$ such that $c_i(xy_j) = c_0(xy_{j+1})$ for all $0 \leqslant j < i$, $c_i(xy_i)$ is not defined, and $c_i(e) = c_0(e)$ otherwise.

Then c_i is a proper $(\Delta+1)$-edge-colouring of $G\backslash xy_i$ due to the definition of $y_0, \dots, y_k$. Note that the absent colours in x are the same with respect to c_i, for all $0 \leqslant i \leqslant k$.

Let β be the absent colour in y_k with respect to c_0. Then β is also absent in $Y-k$ as well, with respect to C_i, for all $0 \leqslant i \leqslant k$. Note that β cannot be absent in x, otherwise we could easily extend c_k; therefore, an edge with colour β is incident to x for all c_j. From the maximality of k, there exists $1 \leqslant i < k$ such that $c_0(xy_i) = \beta$. From the definition of $c_1, \dots, c_k$, we have that $c_0(xy_i) = c_{i-1}(xy_i) = c_k(xy_{i-1}) = \beta$.

Let p be the α/β-path from Y_k with respect to c_k. From (1), P has to end in x. But α is absent in x, so it has to end with an edge of colour β. Therefore, the last edge of P is $y_{i-1}x$. Now let P' be the α/β-path from Y_{i-1} with respect to c_{i-1}. Since P' is uniquely determined and the inner edges of P are not changed in $c_0, \dots, c_k$, the pathP' uses the same edges as P in reverse order and visits Y_k. The edge leading to y_k clearly has colour α. But β is missing in y_k, so P' ends in y_k, which contradicts (1), thus completing the proof. $\square$

Chapter 2

Subgraphs

2.1 Turán graphs

We now start systematically investigating the local structure of graphs. Local structure refers to the intrinsic relations that hold between the answers to the questions "which subgraphs appear in a graph G?" and "how many of them are there?". The first serious result of this kind was proved by Mantel [42] in 1907, in which he studies the maximum number of edges that a graph on n vertices can have without containing a triangle as a subgraph.

Theorem 2.1.1 (Mantel)**.** *If a graph G on n vertices contains no triangles, then it has at most $\frac{n^2}{4}$ edges.*

Proof. Let A be the largest independent set in G. Since G contains no triangles, the neighbourhood of every vertex x is an independent set. Therefore, for every $x \in V(G)$, we have that $d(x) \leqslant |A|$. Let B be the complement of A. Every edge in G must meet a vertex of B. Therefore, the number of edges in G satisfies

$$|E(G)| \leqslant \sum_{x \in B} d(x) \leqslant |A||B| \leqslant \left(\frac{|A| + |B|}{2}\right)^2 = \frac{n^2}{4}.$$

Suppose that n is even. Then, equality holds if and only if $|A| = |B| = \frac{n}{2}$, $d(x) = |A|$ for every $x \in B$ and B has no internal edges. This easily implies that the unique structure with $\frac{n^2}{4}$ edges is a bipartite graph with equal partition sets. For n odd, the number of edges is maximised when $|A| = \left\lceil \frac{n}{2} \right\rceil$ and $|B| = \left\lfloor \frac{n}{2} \right\rfloor$. Again, this yields a unique bipartite structure. $\square$

This proof tells us that not only is $\left\lfloor \frac{n^2}{4} \right\rfloor$ the maximum number of edges in a triangle-free graph, but also that any triangle-free graph with this number of edges is bipartite with partition sets of almost equal size.

The natural generalisation of Mantel's theorem to complete graphs of size r is the following, proved by Turán [59] in 1941.

Theorem 2.1.2 (Turán)**.** *If a graph G on n vertices contains no copy of K_{r+1}, then it contains at most $\left(1-\frac{1}{r}\right)\frac{n^2}{2}$ edges.*

First proof. We will prove the claim by induction on n. The theorem is trivially true for $n = 1, \ldots, r$. We will therefore assume that it is true for all values less than n and prove it for n. Let G be a graph on n vertices which contains no K_{r+1} and has the maximum possible number of edges. Then G contains copies of K_r, otherwise we could add edges to G, contradicting maximality.

Let A be a clique of size r and let B be its complement. Since B has size $n-r$ and contains no K_{r+1}, there are at most $\left(1-\frac{1}{r}\right)\frac{(n-r)^2}{2}$ edges in B. Moreover, since every vertex in B can have at most $r-1$ neighbours in A, the number of edges between A and B is at most $(r-1)(n-r)$. Summing, we see that

$$\begin{aligned}|E(G)| &= |E(A)| + |E(A,B)| + |E(B)| \\ &\leqslant \binom{r}{2} + (r-1)(n-r) + \left(1-\frac{1}{r}\right)\frac{(n-r)^2}{2} \\ &= \left(1-\frac{1}{r}\right)\frac{n^2}{2},\end{aligned}$$

thus completing the proof. □

Second proof. We again assume that G is edge-maximal without containing K_{r+1} as a subgraph. We will begin by proving that if $xy \notin E(G)$ and $yz \notin E(G)$, then $xz \notin E(G)$.

Suppose, towards a contradiction, that $xy \notin E(G)$ and $yz \notin E(G)$, but $xz \in E(G)$. If $d(y) < d(x)$, then we may construct a new K_{r+1}-free graph G' by deleting y and creating a new copy of the vertex x, say x', which we join to exactly the neighbours of x (but not to x itself). Since any clique in G' can contain at most one of x, x', we see that G' is K_{r+1}-free. Moreover,

$$|E(G')| = |E(G)| - d(y) + d(x) > |E(G)|,$$

contradicting the maximality of G. A similar conclusion holds if $d(y) < d(z)$. We may therefore assume that $d(y) \geqslant d(x)$ and $d(y) \geqslant d(z)$. We create a new graph G'' by deleting x and z and creating two extra copies of the vertex y which, again, we join to exactly the neighbours of y. As before, G'' contains no K_{r+1} and

$$|E(G'')| = |E(G)| - (d(x) + d(z) - 1) + 2d(y) > |E(G)|,$$

so, again, we arrive at a contradiction.

Hence, the vertices of G can be partitioned into equivalence classes where vertices in the same class are non-adjacent and vertices in different classes are adjacent. Therefore, the graph is a complete multipartite graph and, clearly, it can have at most r parts. We will show that the number of edges is maximised when all of these parts have sizes which differ by at most 1. Indeed, if there are two parts A and B with $|A| > |B| + 1$, we could increase the number of edges in G by moving one vertex from A to B. We would lose $|B|$ edges by doing so, but gain $|A| - 1$. Overall, we would gain $|A| - 1 - |B| \geqslant 1$ edges, thus proving our claim. □

The second proof, as did the one of Mantel's theorem, determines the structure of the extremal graph, that is, it must be r-partite with all parts having size as close as possible. In particular, if $n = pr + q$, then G has q partition sets of size $p + 1$ and $r - q$ partition sets of size p. The unique complete r-partite graphs on $n \geqslant r$ vertices whose partition sets differ in size by at most 1 are called *Turán graphs* and are denoted by $T^r(n)$. The number of edges of the Turán graph $T^r(n)$ is often denoted by $t_r(n)$.

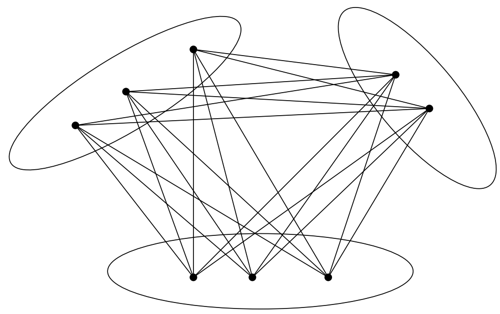

Figure 2.1: The Turán graph $T^3(8)$

By the pigeonhole principle, every set of $r + 1$ vertices in $T^r(n)$ includes two vertices in the same partition set. Therefore, $T^r(n)$ does not contain a K_{r+1}. Turán's theorem states that $T^r(n)$ is edge-maximal among all K_{r+1}-free graphs on n vertices.

2.2 Szemerédi's regularity lemma

Our next aim is to prove Szemerédi's regularity lemma, which was developed by Szemerédi [56] in his work of what is now known as Szemerédi's theorem. This theorem states that for any $\delta > 0$ and $k \geqslant 3$ there exists a $n_0 \in \mathbb{N}$ such that, for $n \geqslant n_0$, any subset of $\{1, \ldots, n\}$ with at least δn elements must contain an arithmetic progression of length k. The particular case where $k = 3$ had been proven earlier by Roth [48] and is accordingly known as Roth's theorem.

Roughly speaking, Szemerédi's regularity lemma says that any graph may be partitioned into a finite number of sets such that most of the bipartite graphs between different sets are random-like. In order to be precise, we shall need some notation and a few definitions.

Let G be a graph and let A and B be subsets of $V(G)$. The *density* of edges between A and B is given by

$$d(A,B) = \frac{|E(A,B)|}{|A||B|}.$$

Given some $\epsilon > 0$, the pair (A,B) ia said to be *ϵ-regular* if, for every $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geqslant \epsilon|A|$ and $|B'| \geqslant \epsilon|B|$, we have that

$$|d(A',B') - d(A,B)| \leqslant \epsilon.$$

We say that a partition $\{X_1,\ldots,X_k\}$ of $V(G)$ is *ϵ-regular* if

$$\sum \frac{|X_i||X_j|}{n^2} \leqslant \epsilon,$$

where the sum is taken over all pairs (X_i,X_j) which are not ϵ-regular.

That is, a bipartite graph is ϵ-regular if all small induced subgraphs have approximately the same density as the full graph, and a partition of the vertex set of a graph G into smaller sets is ϵ-regular if almost every pair forms a bipartite graph which is ϵ-regular. We are now ready to state the regularity lemma.

Theorem 2.2.1 (Szemerédi's regularity lemma). *For every $\epsilon > 0$ there exists an M such that, for every graph G, there is an ϵ-regular partition of the vertex set of G with at most M pieces.*

Given a partition $\{X_1,\ldots,X_k\}$ of $V(G)$, the *mean square density* of this partition is given by

$$\sum_{1\leqslant i,j\leqslant k} \frac{|X_i||X_j|}{n^2} d(X_i,X_j)^2.$$

Lemma 2.2.2. *For every partition of the vertex set of a graph G, the mean square density lies between 0 and 1.*

Proof. Since $\sum_{1\leqslant i,j\leqslant k} \frac{|X_i||X_j|}{n^2} = 1$ and $0 \leqslant d(X_i,X_j) \leqslant 1$, the mean square density also lies between 0 and 1. □

Another important property of the mean square density is that it cannot decrease under refinement of a partition. That is, we have the following.

Lemma 2.2.3. *Let G be a graph. If $\{X_1, \ldots, X_k\}$ is a partition of $V(G)$ and $\{Y_1, \ldots, Y_l\}$ is a refinement of $\{X_1, \ldots, X_k\}$, then the mean square density of $\{Y_1, \ldots, Y_l\}$ is at least the mean square density of $\{X_1, \ldots, X_k\}$.*

Proof. Since $\{Y_1, \ldots, Y_l\}$ is a refinement of $\{X_1, \ldots, X_k\}$, every X_i may be rewritten as a disjoint union $X_{i1} \cup \cdots \cup X_{ia_i}$, where each $X_{ia_i} = Y_j$ for some j. Now, by the Cauchy-Schwarz inequality, we have

$$\begin{aligned} d(X_i, X_j)^2 &= \Big(\sum_{s,t} \frac{|X_{is}||X_{jt}|}{|X_i||X_j|} d(X_{is}, X_{jt}) \Big)^2 \\ &\leqslant \Big(\sum_{s,t} \frac{|X_{is}||X_{jt}|}{|X_i||X_j|} \Big) \Big(\sum_{s,t} \frac{|X_{is}||X_{jt}|}{|X_i||X_j|} d(X_{is}, X_{jt})^2 \Big) \\ &= \sum_{s,t} \frac{|X_{is}||X_{jt}|}{|X_i||X_j|} d(X_{is}, X_{jt})^2. \end{aligned}$$

Therefore,

$$\frac{|X_i||X_j|}{n^2} d(X_i, X_j)^2 \leqslant \sum_{s,t} \frac{|X_{is}||X_{jt}|}{n^2} d(X_{is}, X_{jt})^2.$$

Adding over all values of i and j implies the lemma. □

An analogous result also holds for bipartite graphs. That is, if G is a bipartite graph with partition sets X and Y, $\bigcup_i X_i$ and $\bigcup_i Y_i$ are partitions of X and Y, and $\bigcup_i Z_i$ and $\bigcup_i W_i$ refine these partitions, then

$$\sum_{i,j} \frac{|X_i||Y_j|}{n^2} d(X_i, Y_j)^2 \leqslant \sum_{i,j} \frac{|Z_i||W_j|}{n^2} d(Z_i, W_j)^2.$$

We will now show that if X and Y are two sets of vertices and the graph between them is not ϵ-regular then there is a partition of each of X and Y for which the mean square density increases.

Lemma 2.2.4. *Let G be a graph and suppose X and Y are subsets of $V(G)$. If $d(X, Y) = a$ and the graph between X and Y is not ϵ-regular, then there are partitions $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$ such that*

$$\sum_{1 \leqslant i,j \leqslant 2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 \geqslant a^2 + \epsilon^4.$$

Proof. Since the graph between X and Y is not ϵ-regular, there must be two subsets X_1 and Y_1 of X and Y respectively, with $|X_1| \geqslant \epsilon|X|$, $|Y_1| \geqslant \epsilon|Y|$ and $|d(X_1, Y_1) -$

$a| > \epsilon$. Let $X_2 = X\backslash X_1$, $Y_2 = Y\backslash Y_1$ and $u(X_i, Y_j) = d(X_i, Y_j) - a$. Then

$$\begin{aligned}
\epsilon^4 &\leqslant \sum_{1\leqslant i,j\leqslant 2} \frac{|X_i||Y_j|}{|X||Y|} u(X_i, Y_j)^2 \\
&= \sum_{1\leqslant i,j\leqslant 2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 - 2a \sum_{1\leqslant i,j\leqslant 2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j) + a^2 \sum_{1\leqslant i,j\leqslant 2} \frac{|X_i||Y_j|}{|X||Y|} \\
&= \sum_{1\leqslant i,j\leqslant 2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 - a^2.
\end{aligned}$$

Note that the second equality holds since

$$\sum_{1\leqslant i,j\leqslant 2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j) = d(X, Y) = a.$$

The result follows. □

In order to complete the proof of the regularity lemma, we need to prove that if a partition is not ϵ-regular there is a refinement of this partition which has a higher mean square density. This is taken care of in the following lemma.

Lemma 2.2.5. *Let G be a graph and let $\{X_1, \ldots, X_k\}$ be a partition of its vertex set which is not ϵ-regular. Then there is a refinement of $\{X_1, \ldots, X_k\}$, say $\{X_{11}, \ldots, X_{1a_1}, \ldots, X_{k1}, \ldots, X_{ka_k}\}$, such that $a_i \leqslant 2^{2k}$, for every $1 \leqslant i \leqslant k$, and the mean square density of the refinement is at least ϵ^5 larger than the one of the original partition.*

Proof. Let $I = \{(i,j) : (X_i, X_j) \text{ is not } \epsilon\text{-regular}\}$. Let a^2 be the mean square density of $\{X_1, \ldots, X_k\}$.

For each $(i,j) \in I$, the previous lemma gives us partitions $X_i = A_1^{ij} \cup A_2^{ij}$ and $X_j = B_1^{ij} \cup B_2^{ij}$ for which

$$\sum_{1\leqslant p,q\leqslant 2} \frac{|A_p^{ij}||B_q^{ij}|}{|X_i||X_j|} d(A_p^{ij}, B_q^{ij})^2 \geqslant d(X_i, X_j)^2 + \epsilon^4.$$

For each i, let $\{X_{i1}, \ldots, X_{ia_i}\}$ be the partition of X_i which refines all partitions that arise from splitting X_i or X_j into A_i's or B_i's. Note that this partition has at most 2^{2k} pieces, that is, $a_i \leqslant 2^{2k}$. Moreover, since refining bipartite partitions does not decrease the mean square density, we have

$$\sum_{p=1}^{a_i} \sum_{q=1}^{a_j} \frac{|X_{ip}||X_{jq}|}{|X_i||X_j|} d(X_{ip}, X_{jq})^2 \geqslant d(X_i, X_j)^2 + \epsilon^4,$$

for all $(i,j) \in I$. Multiplying both sides of the inequality by $\frac{|X_i||X_j|}{n^2}$ and summing over all (i,j), we have

$$\sum_{1 \leqslant i,j \leqslant k} \sum_{p=1}^{a_i} \sum_{q=1}^{a_j} \frac{|X_{ip}||X_{jq}|}{n^2} d(X_{ip}, X_{jq})^2 \geqslant$$
$$\sum_{1 \leqslant i,j \leqslant k} \frac{|X_i||X_j|}{n^2} d(X_i, X_j)^2 + \epsilon^4 \sum_{(i,j)\in I} \frac{|X_i||X_j|}{n^2} \geqslant$$
$$a^2 + \epsilon^5.$$

The result follows. □

We now have all the ingredients necessary to prove the regularity lemma.

Proof of Szemerédi's regularity lemma. Let us start with a trivial partition into one set. If it is ϵ-regular we are done. Otherwise, there is a partition into at most 4 sets where the mean square density increases by ϵ^5.

If, at stage i, we have a partition into k pieces and this partition is not ϵ-regular, there is a partition into at most $k2^{2k} \leqslant 2^{2^k}$ pieces whose mean square density is at least ϵ^5 greater. Since the mean square density is bounded above by 1, this process must end after at most ϵ^{-5} steps. The number of pieces in the final partition is at most a tower of 2's of height $2\epsilon^{-5}$. □

The tower function $t_i(x)$ is defined recursively by $t_0(x) = x$ and, for $i \geqslant 0$, $t_{i+1}(x) = 2^{t_i(x)}$. The bound in the proof of the regularity lemma is $t_{2\epsilon^5}(2)$, which is clearly enormous. Surprisingly, Gowers [26] proved that there are graphs where, in order to get an ϵ-regular partition, one need roughly that many pieces in the partition, thus putting the hopes of finding a better bound to rest.

2.3 Consequences

One of the interesting consequences of the regularity lemma is the *triangle removal lemma*, from which we shall deduce Roth's theorem. The triangle removal lemma states that, if a graph contains very few triangles, then one may remove all of them by removing very few edges. We begin with what is known as a *counting lemma*, which states that if a tri-partite graph is pairwise "dense" and regular, then it contains a positive portion of triangles .

Lemma 2.3.1. *Let G be a graph and let X, Y, Z be subsets of $V(G)$. Suppose that $(X,Y), (Y,Z)$ and (Z,X) are ϵ-regular and that $d(X,Y) = \alpha$, $d(Y,Z) = \beta$ and*

$d(Z,X) = \gamma$. Then, provided $\alpha, \beta, \gamma \geqslant 2\epsilon$, the number of triangles xyz with $x \in X$, $y \in Y$ and $z \in Z$ is at least

$$(1-2\epsilon)(\alpha-\epsilon)(\beta-\epsilon)(\gamma-\epsilon)|X||Y||Z|.$$

Proof. For every x, let $d_Y(x)$ and $d_Z(x)$ denote the number of neighbours of x in Y and Z, respectively. Then the number of $x \in X$ with $d_Y(x) < (\alpha-\epsilon)|Y|$ is at most $\epsilon|X|$; otherwise, there would be a subset X' of X of size at least $\epsilon|X|$ such that the density of edges between X' and Y is less than $\alpha - \epsilon$, contradicting regularity. We may similarly show that there are at most $\epsilon|X|$ values of x for which $d_Z(x) < (\gamma-\epsilon)|Z|$. If $d_Y(x) > (\alpha-\epsilon)|Y|$ and $d_Z(x) > (\gamma-\epsilon)|Z|$, the number of edges between $N(x) \cap Y$ and $N(x) \cap Z$ and, consequently, the number of triangles containing x, is at least

$$(\alpha-\epsilon)(\beta-\epsilon)(\gamma-\epsilon)|Y||Z|.$$

Summing over all $x \in X$, we get the desired result. □

Theorem 2.3.2 (Triangle removal lemma)**.** *For every $\epsilon > 0$ there exists $\delta > 0$ such that any graph G on n vertices with at most δn^3 triangles may be made triangle-free by removing at most ϵn^2 edges.*

Proof. Let $\{X_1, \dots X_M\}$ be an $\frac{\epsilon}{4}$-regular partition of $V(G)$. We remove an edge xy from G if

1. $(x,y) \in X_i \times X_j$, where (X_i, X_j) is not an $\frac{\epsilon}{4}$-regular pair;
2. $(x,y) \in X_i \times X_j$, where $d(X_i, X_j) < \frac{\epsilon}{2}$;
3. $x \in X_i$, where $|X_i| \leqslant \frac{\epsilon}{4M} n$.

Let $I = \{(i,j) \colon (X_i, X_j) \text{ is not } \frac{\epsilon}{4}\text{-regular}\}$. The number of edges removed by condition (1) is at most

$$\sum_{(i,j)\in I} |X_i||X_j| \leqslant \frac{\epsilon}{4} n^2.$$

The number removed by condition (2) is clearly at most $\frac{\epsilon}{2} n^2$. Finally, the number removed by condition (3) is at most

$$Mn \frac{\epsilon}{4M} n = \frac{\epsilon}{4} n^2.$$

Overall, we have removed at most ϵn^2 edges.

Suppose that some triangle remains in the graph, say xyz, where $x \in X_i$, $y \in X_j$ and $z \in X_k$. Then the pairs (X_i, X_j), (X_j, X_k) and (X_k, X_i) are all $\frac{\epsilon}{4}$-regular with density at least $\frac{\epsilon}{2}$. Therefore, since $|X_i|, |X_j|, |X_k| \geqslant \frac{\epsilon}{4M}n$, the counting lemma implies that the number of triangles is at least

$$\left(1 - \frac{\epsilon}{2}\right)\left(\frac{\epsilon}{4}\right)^3\left(\frac{\epsilon}{4M}\right)^3 n^3.$$

Taking $\delta = \frac{\epsilon^6}{2^{20}M^3}$ yields a contradiction. □

We will now use the triangle removal lemma in order to prove Roth's theorem. We are actually going to prove the following stronger result.

Theorem 2.3.3. *For any $\delta > 0$ there exists $n_0 \in \mathbb{N}$ such that, for $n \geqslant n_0$, any $A \subseteq [n]^2$ with at least δn^2 elements must contain a triple of the form $(x, y), (x + d, y), (x, y + d)$ with $d > 0$.*

Proof. The set $A + A = \{x + y : x, y \in A\}$ is contained in $[2n]^2$. There must, therefore, be some z which is represented as $x + y$ in at least

$$\frac{(\delta n^2)^2}{(2n)^2} = \frac{\delta^2 n^2}{4}$$

different ways. Pick such a z and let $A' = A \cap (z - A)$ and $\delta' = \frac{\delta^2}{4}$. Then $|A'| \geqslant \delta' n^2$ and if A' contains a triple of the form $(x, y), (x + d, y), (x, y + d)$ for $d < 0$, then so does $z - A$. Therefore, A will contain such a triple with $d > 0$. That being so, we may forget about the constraint that $d > 0$ and try to find some non-trivial triple with $d \neq 0$.

Consider the tripartite graph on vertex sets X, Y and Z, where $X = Y = [n]$ and $Z = [2n]$. The set X will correspond to vertical lines through A, the set Y to horizontal lines and the set Z to diagonal lines with constant values of $x + y$. We form a graph G by joining $x \in X$ to $y \in Y$ if and only if $(x, y) \in A$. We also join $x \in X$ and $z \in Z$ if and only if $(x, z - x) \in A$ and $y \in Y$ and $z \in Z$ if and only if $(z - y, y) \in A$.

If there is a triangle xyz in G, then (x, y), $(x, y+(z-x-y))$ and $(x+(z-x-y), y)$ will all be in A and thus we will have the required triple unless $z = x+y$. This means that there are at most $n^2 = \frac{1}{64n}(4n)^3$ triangles in G. By the triangle removal lemma, for n sufficiently large, one may remove $\frac{\delta}{2}n^2$ edges and make the graph triangle-free. But every point in A determines a degenerate triangle. Hence, there are at least δn^2 degenerate triangles, all of which are edge disjoint. We cannot, therefore, remove them all by removing $\frac{\delta}{2}n^2$ edges. This contradiction implies the required result. □

This implies Roth's theorem as follows.

Theorem 2.3.4 (Roth)**.** *For all $\delta > 0$ there exists $n_0 \in \mathbb{N}$ such that, for $n \geqslant n_0$, any $A \subseteq [n]$ with at least δn elements contains an arithmetic progression of length 3.*

Proof. Let $B \subseteq [2n]^2$ be the set $\{(x, y) : x - y \in A\}$. Then $|B| \geqslant \delta n^2 = \frac{\delta}{4}(2n)^2$, so we have a triple (x, y), $(x + d, y)$, $(x, y + d)$ in B. This translates back to tell us that $x - y - d$, $x - y$ and $x - y + d$ are in A, as required. □

2.4 Erdős-Stone-Simonovits theorem

For general graphs H, we are interested in the function $\mathsf{ex}(n, H)$, defined as

$$\mathsf{ex}(n, H) = \max\{|E(G)| : |V(G)| = n,\ H \nsubseteq G\}.$$

Turán's theorem itself tells us that

$$\mathsf{ex}(n, K_{r+1}) \leqslant \left(1 - \frac{1}{r}\right)\frac{n^2}{2}.$$

We are now going to deal with the general case. We will show that the behaviour of the extremal function $\mathsf{ex}(n, H)$ is tied intimately to the chromatic number of the graph H.

The fundamental result which we shall prove is known as the Erdős-Stone-Simonovits theorem. Although the theorem is named after the three mathematicians, Erdős, Stone and Simonovits never actually wrote a paper together. Erdős and Stone [21] proved the result when H is a complete multipartite graph, and Erdős and Simonovits [22] then proved it for general H.

Theorem 2.4.1 (Erdős-Stone-Simonovits)**.** *For every H, r and $\epsilon > 0$, such that $\chi(H) = r$, there exists an $n_0 = n_0(r, H, \epsilon)$ so that, for any $n \geqslant n_0$,*

$$\left(1 - \frac{1}{r-1} - \epsilon\right)\frac{n^2}{2} \leqslant \mathsf{ex}(n, H) \leqslant \left(1 - \frac{1}{r-1} + \epsilon\right)\frac{n^2}{2}.$$

For the complete graph K_r, the chromatic number is r, so in this case the Erdős-Stone-Simonovits theorem reduces to an approximate version of Turán's theorem. For bipartite H, it gives $\mathsf{ex}(n, H) \leqslant \epsilon n^2$, for all $\epsilon > 0$.

The lower bound is easy; the Turán graph $T^{r-1}(n)$ is $(r-1)$-colourable, hence H-free, and it is easy to check that $t_{r-1}(n) \geqslant \left(1 - \frac{1}{r-1} - o(1)\right)\frac{n^2}{2}$.

For the upper bound, it is sufficient to prove that the bound holds for the t-blowup[1] of K_r, since if $\chi(H) = r$, then H is a subgraph of K_r^t for $t = |V(H)|$.

[1] The t-blowup of a graph K, denoted by K^t, is the graph obtained by replacing every vertex x of K by an independent set I_x of size t, and replacing every edge xy by a complete bipartite graph between I_x and I_y.

This, however, is not the proof we shall present. Instead, our proof will rely on the regularity lemma in an attempt to illustrate how powerful a tool it is. To begin, we shall need another counting lemma which generalises the one given earlier for triangles.

Lemma 2.4.2. *Let $\epsilon > 0$ and let G be a graph. Suppose that $V_1, \ldots, V_r$ are subsets of $V(G)$ such that $|V_i| \geqslant 2\epsilon^{-\Delta} t$ for each $1 \leqslant i \leqslant r$ and the graph between V_i and V_j has density $d(V_i, V_j) \geqslant 2\epsilon$ and is $\frac{1}{2}\epsilon^{\Delta}\Delta^{-1}$-regular for all $1 \leqslant i < j \leqslant r$. Then G contains a copy of any graph H on t vertices with chromatic number at most r and maximum degree Δ.*

Proof. Since the chromatic number of H is at most r, we may split $V(H)$ into r independent sets $U_1, \ldots, U_r$. We will give an embedding f of H into G so that $f(U_i) \subseteq V_i$ for all $1 \leqslant i \leqslant r$.

Let the vertices of H be $u_1, \ldots, u_t$. For each $1 \leqslant h \leqslant t$, let $L_h = \{u_1, \ldots, u_h\}$. For each $y \in U_j \backslash L_h$, let T_y^h be the set of vertices in V_j which are adjacent to all already embedded neighbours of y. That is, letting $N_h(y) = N(y) \cap L_h$, T_y^h is the set of vertices in V_j adjacent to every element of $f(N_h(y))$. We will find, by induction, an embedding of L_h such that, for each $y \in V(H)\backslash L_h$, $|T_y^h| \geqslant \epsilon^{|N_h(y)|}|V_j|$.

For $h = 0$ there is nothing to prove. We may, therefore, assume that L_h has been embedded consistent with the induction hypothesis and attempt to embed $u = u_{h+1} \in U_k$ into an appropriate $v \in T_u^h$. Let Y be the set of neighbours of u which are not yet embedded. We wish to find an element $v \in T_u^h \backslash f(L_h)$ such that,for all $y \in Y$, $|N(v) \cap T_y^h| \geqslant \epsilon |T_y^h|$. If such a vertex v exists, taking $f(u) = v$ and $T_y^{h+1} = N(v) \cap T_y^h$ will complete the proof.

Let B_y be the set of vertices in T_u^h such that $|N(v) \cap T_y^h| < \epsilon |T_y^h|$. Note that, by induction, if $y \in U_l$, then $|T_y^h| \geqslant \epsilon^{\Delta}|V_l|$. Therefore, we must have $|B_y| < \frac{1}{2}\epsilon^{\Delta}\Delta^{-1}|V_k|$, for otherwise the density between B_y and T_y^h would be less than ϵ, contradicting the regularity assumption on G. Hence, since $|V_k| \geqslant 2\epsilon^{\Delta} t$, we have

$$\left| T_u^h \backslash \bigcup_{y \in Y} B_y \right| > \epsilon^{\Delta}|V_k| - \Delta \frac{1}{2}\epsilon^{\Delta}\Delta^{-1}|V_k| \geqslant t.$$

Since at most $t - 1$ vertices have already been embedded, an appropriate choice for $f(u)$ exists. □

In fact, there are at least $\frac{1}{2}\epsilon^{\Delta}|V_k| - t$ choices for each vertex u. Therefore, if H had d_i vertices in U_i, the lemma tells us that, for $|V_i| \gg 2\epsilon^{-\Delta} t$, we have at least

$$c_H(\epsilon) \prod_{i=1}^{r} |V_i|^{d_r}$$

copies of H, where $c_H(\epsilon)$ is the appropriate constant. Like the triangle counting lemma, we could make the constant $c_H(\epsilon)$ reflect the densities between the various V_i, but we simply wanted to note that the graph G contained a positive proportion of the total number of possible copies of H.

We are now ready to give the proof of the Erdős-Stone-Simonovits theorem.

Proof of Erdős-Stone-Simonovits. As we mentioned after the statement of the theorem, the lower bound is easy to prove. We shall therefore only consider the upper bound.

Let H be a graph with t vertices, chromatic number r and maximum degree Δ. Suppose that G is a graph on n vertices with at least $(1 - \frac{1}{r-1} + \epsilon)\frac{n^2}{2}$ edges. We will show how to embed H in G.

Let $\{X_1, \dots, X_M\}$ be a $\frac{1}{2}(\frac{\epsilon}{8})^\Delta \Delta^{-1}$-regular partition of $V(G)$. We remove edges as in the triangle removal lemma, removing an edge xy if

1. $(x, y) \in X_i \times X_j$, where (X_i, X_j) is not $\frac{1}{2}(\frac{\epsilon}{8})^\Delta \Delta^{-1}$-regular;
2. $(x, y) \in X_i \times X_j$, where $d(X_i, X_j) < \frac{\epsilon}{4}$;
3. $x \in X_i$, where $|X_i| < \frac{\epsilon}{16M} n$.

The total number of edges removed by condition (1) is at most $\frac{\epsilon}{16} n^2$, since if I is the set of (i, j) corresponding to non-regular pairs (X_i, X_j), we have

$$\sum_{(i,j)\in I} |X_i||X_j| \leqslant \frac{1}{2}\left(\frac{\epsilon}{8}\right)^\Delta \Delta^{-1} n^2 \leqslant \frac{\epsilon}{16} n^2.$$

The total number of edges removed by condition (2) is clearly at most $\frac{\epsilon}{4} n^2$ and the total number removed by condition (3) is at most $\frac{\epsilon}{16} n^2$.

Overall, we have removed at most $\frac{3\epsilon}{8} n^2$ edges. Hence, the graph G' that remains after all these edges have been removed has density at least $1 - \frac{1}{r-1} + \frac{\epsilon}{8}$. It must, therefore, contain a copy of K_r. We may suppose that this lies between sets $V_1, \dots, V_r$ (some of which may be equal). Because of our removal process, $|V_j| \geqslant \frac{\epsilon}{16M} n$, and the graph between V_i and V_j has density at least $\frac{\epsilon}{4}$ and is $\frac{1}{2}(\frac{\epsilon}{8})^\Delta \Delta^{-1}$-regular. Therefore, if

$$\frac{\epsilon}{16M} n \geqslant 2\left(\frac{\epsilon}{8}\right)^{-\Delta} t,$$

an application of the previous lemma with $\frac{\epsilon}{8}$ implies that G contains a copy of H. $\square$

Because of the observation made after the previous lemma we know that, for n sufficiently large, G not only contains one copy of any given r-chromatic graph H, it must contain $cn^{|V(H)|}$ copies. This phenomenon, that once one passes the extremal density one gets a very large number of copies rather than a single one, is known as *supersaturation.*

2.5 Bipartite graphs

The *Zarankiewicz problem*, an unsolved problem in extremal graph theory, asks for the largest possible number of edges in a bipartite graph that has a given number of vertices but has no complete bipartite subgraphs of a given size. It is named after Kazimierz Zarankiewicz, who proposed several special cases of the problem in 1951 [62]. We have already seen that if H is a bipartite graph, then $\mathsf{ex}(n, H) \leqslant \epsilon n^2$ for any $\epsilon > 0$. We will now prove a much stronger result, which was first shown by Kővári, Sós and Turán [33] in 1954.

Theorem 2.5.1 (Kővári-Sós-Turán)**.** *For any $s, t \in \mathbb{N}$ with $s \leqslant t$, there exists a constant c such that*

$$\mathsf{ex}(n, K_{s,t}) \leqslant cn^{2-\frac{1}{s}}.$$

Proof. Suppose that G is a graph on n vertices with at least $cn^{2-\frac{1}{s}}$ edges. We will count the pairs (v, S) consisting of a vertex v and a set $S \subseteq N(v)$ of size s. The number of such pairs is

$$\sum_v \binom{d(v)}{s} \geqslant n\binom{\frac{1}{n}\sum_v d(v)}{s} \geqslant n\binom{2cn^{1-\frac{1}{s}}}{s} \geqslant n\frac{c^s n^{s-1}}{s!} = c^s\frac{n^s}{s!},$$

for n sufficiently large. If $c^s > t - 1$, then there exists a set S that is counted at least t times in the equation above. This gives a copy of $K_{s,t}$ in G. □

An interesting example is $H = K_{2,2}$. The following $K_{2,2}$-free construction, due to Erdős, Rényi and Sós [20], allows us to show that $\mathsf{ex}(n, K_{2,2}) \approx \frac{1}{2}n^{\frac{3}{2}}$.

Construction of a $K_{2,2}$-free graph. Let p be a prime number and consider the graph on $n = p^2 - 1$ vertices whose vertex set is $\mathbb{Z}_p \times \mathbb{Z}_p \backslash \{(0, 0)\}$ and where (x, y) is joined to (a, b) if and only if $ax + by = 1$.

For a fixed (x, y), there are exactly p solutions (a, b) to $ax + by = 1$. To see this, we must split into some subcases. If $x = 0$, then there is a unique non-zero solution for b and anything works for a. Similarly, if $y = 0$, then a is uniquely determined and b may be anything. If both x and y are non-zero, it is elementary to see that any choice of b gives rise to a unique choice of a, that is, $a = x^{-1}(1 - by)$.

Therefore, (x, y) has degree at least $p-1$ (one of the solutions could be $(a, b) = (x, y)$, which we ignore) and the graph has at least $\frac{1}{2}n(p-1) \approx \frac{1}{2}n^{\frac{3}{2}}$ edges. Moreover, the graph does not contain a $K_{2,2}$; suppose otherwise and that (a, b), (x, y), (a', b'), (x', y') is a $K_{2,2}$. Then the set of simultaneous equations $ux+vy = 1$ and $ux'+vy' = 1$ would have two solutions, $(u, v) = (a, b)$ and (a', b'), which is clearly impossible, since any two distinct lines meet at at most one point. □

There is also a construction, due to Brown [8], which gives a lower bound $\mathsf{ex}(n, K_{3,3}) \geqslant c'n^{\frac{5}{3}}$. Roughly speaking, take a prime $p \equiv (3 \mod 4)$ and consider the graph on p^3 vertices whose vertex set is $\mathbb{Z}_p^3$, where (x, y, z) is joined to (a, b, c) if and only if $(a-x)^2 + (b-y)^2 + (c-z)^2 = 1$. For any given (x, y, z), there will be on the order of p^2 elements (a, b, c) to which it is connected. There are, therefore, approximately $c'n^{\frac{5}{3}}$ edges in the graph. Moreover, the unit spheres around the three distinct points (x, y, z), (x', y', z') and (x'', y'', z'') cannot meet in more than two points, so the graph does not contain a $K_{3,3}$. The result for all n follows from an argument similar to the above.

Almost the best known lower bound is a result of a random construction, but before we can present it we need to dip our toes in the theory of random graphs.

Let V be a fixed set of n elements, say $V = \{0, \ldots, n-1\}$. We want to turn the set $\mathcal{G}$ of all graphs on V into a probability space. Intuitively, given a set $G \in \mathcal{G}$, we should be able to generate it randomly as follows. For each $e \in [V]^2$ we decide by some random experiment whether or not e shall be an edge of G; these experiments are performed independently, and for each the probability of success, that is, of accepting e as an edge for G, is equal to some fixed number $p \in [0, 1]$. Then, if G_0 is some fixed graph on V, with m edges say, the elementary event $\{G_0\}$ has a probability of $p^m q^{\binom{n}{2}-m}$, where $q = 1-p$. With this probability, our randomly generated graph G is this particular graph G_0. (The probability that G is *isomorphic* to G_0 will usually be greater.) But if the probabilities of all the elementary events are thus determined, then so is the entire probability measure of our desired space $\mathcal{G}$. Hence, all that remains to be checked is that such a probability measure on $\mathcal{G}$, one for which all individual edges occur independently with probability p, does indeed exists.

In order to construct such a measure on $\mathcal{G}$ formally, we start by defining for every potential edge $e \in [V]^2$ its own little probability space $\Omega_e := \{0_e, 1_e\}$, choosing $\mathbb{P}(\{1_e\}) := p$ and $\mathbb{P}(\{0_e\}) := q$ as the probabilities of its two elementary events. As our desired probability space $\mathcal{G} = \mathcal{G}(n, p)$, we then take the product space

$$\Omega := \prod_{e \in [V]^2} \Omega_e.$$

Thus, formally, an element of Ω is a map ω assigning to every $e \in [V]^2$ either 0_e or 1_e, and the probability measure $\mathbb{P}$ on Ω is the product measure of all the measures $\mathbb{P}_e$. In practice, of course, we identify ω with the graph G on V whose edge set is

$$E(G) = \{e : \omega(e) = 1_e\}$$

and call G a *random graph* on V with edge probability p.

Following standard probabilistic terminology, we may now call any set of graphs on V an *event* in $\mathcal{G}(n,p)$. In particular, for every $e \in [V]^2$, the set

$$A_e := \{\omega : \omega(e) = 1_e\}$$

of all graphs G on V with $e \in E(G)$ is an event: the event that e is an edge of G. The events A_e are independent and occur with probability p.

In the context of random graphs, each of the familiar graph invariants (like average degree, girth, chromatic number etc.) may be interpreted as a non-negative *random variable* on $\mathcal{G}(n,p)$, a function

$$X : \mathcal{G}(n,p) \to [0,\infty).$$

The *mean* or *expected* value of X is the number

$$\mathbb{E}(X) := \sum_{G \in \mathcal{G}(n,p)} \mathbb{P}(\{G\}) \cdot X(G).$$

If X takes integers as values, we can compute $\mathbb{E}(X)$ alternatively by summing over these values k:

$$\mathbb{E}(X) = \sum_{k \geqslant 1} \mathbb{P}[X \geqslant k] = \sum_{k \geqslant 1} k \cdot \mathbb{P}[X = k].$$

Note that the operator $\mathbb{E}$, the *expectation*, is linear.

Lastly, we note that, if $p = p(n)$ is a fixed function (possibly constant), and $\mathcal{P}$ is a graph property, we may ask how the probability $\mathbb{P}[G \in \mathcal{P}]$ behaves for $G \in \mathcal{G}(n,p)$ as $n \to \infty$. If this probability tends to 1, we say that $G \in \mathcal{P}$ for *almost all* (or *almost every*) $G \in \mathcal{G}(n,p)$, or that $G \in \mathcal{P}$ *almost surely.*

We are now able to present the random construction that provides us with almost the best lower bound for $\mathsf{ex}(n, K_{s,t})$.

Theorem 2.5.2. *For any $s,t \geqslant 2$, there exists a constant c' such that*

$$\mathsf{ex}(n, K_{s,t}) \geqslant c' n^{2-\frac{s+t-2}{st-1}}.$$

Proof. Choose each edge in the graph randomly with probability $p = \frac{1}{2}n^{-\frac{s+t-2}{st-1}}$. The expected number of copies of $K_{s,t}$ is

$$p^{st}\binom{n}{s}\binom{n}{t} \leqslant p^{st}n^{s+t}.$$

Phrased differently, if J is the random variable counting copies of $K_{s,t}$, then $\mathbb{E}(J) \leqslant p^{st}n^{s+t}$. On the other hand, the expected number of edges in the graph is $p\binom{n}{2} \geqslant \frac{1}{4}pn^2$. Again, if I is the random variable counting the number of edges in the graph, then $\mathbb{E}(I) \geqslant \frac{1}{4}pn^2$. By linearity of expectation, we have that

$$\mathbb{E}(I - J) = \mathbb{E}(I) - \mathbb{E}(J) \geqslant \frac{1}{4}pn^2 - p^{st}n^{s+1} \geqslant \frac{1}{8}pn^2 = \frac{1}{16}n^{2-\frac{s+t-2}{st-1}}.$$

The final inequality follows from the fact that $p^{st}n^{s+t} \leqslant \frac{1}{8}pn^2$. This, in turn, follows from

$$p^{st-1}n^{s+t-2} \leqslant \left(\frac{1}{2}\right)^{st-1} \leqslant \frac{1}{8}.$$

Therefore, there exists some graph G on n vertices for which $I - J \geqslant \frac{1}{16}n^{2-\frac{s+t-2}{st-1}}$. We may therefore remove one edge from each of the $K_{s,t}$, removing all copies of $K_{s,t}$ and still be left with a graph containing $\frac{1}{16}n^{2-\frac{s+t-2}{st-1}}$ edges. □

Our next aim is to prove that if a bipartite graph H has one side whose maximum degree is Δ, then $\mathsf{ex}(n, H) \leqslant c(H)n^{2-\frac{1}{\Delta}}$. This clearly generalises our previous result that $\mathsf{ex}(n, K_{s,t}) \leqslant n^{2-\frac{1}{s}}$ when $s \leqslant t$. Before we prove it we shall need the following lemmas, the first of which is a typical result proved by dependent random choice and demonstrates that every dense graph contains a large vertex subset U such that all small subsets of U have large common neighbourhood.

Lemma 2.5.3. *Let α, m and r be positive integers and let G be a graph on n vertices and average degree d. If there is a positive integer t such that*

$$\frac{d^t}{n^{t-1}} - \binom{n}{r}\left(\frac{m}{n}\right)^t \geqslant \alpha,$$

then G contains a subset U of at least α vertices such that every r vertices in U have at least m common neighbours.

Proof. Pick a set T of t vertices of $V(G)$ uniformly at random with repetition. Let I be the random variable counting the size of $N(T)$. By linearity of expectation, we have that

$$\begin{aligned}\mathbb{E}(I) = \sum_{v \in V(G)} \left(\frac{|N(v)|}{n}\right)^t &= n^{-t} \sum_{v \in V(G)} |N(v)|^t \\ \geqslant n^{1-t}\left(\frac{\sum_{v\in V(G)}|N(v)|}{n}\right)^t &= \frac{d^t}{n^{t-1}},\end{aligned}$$

where the last inequality is by convexity of the function x^t.

Let J be the random variable counting the number of subsets $S \subseteq N(T)$ of size r with fewer than m common neighbours. For a given such S, the probability that it is a subset of $N(T)$ equals $\left(\frac{|N(S)|}{n}\right)^t$. Since there are at most $\binom{n}{r}$ subsets S of size r for which $|N(S)| < m$, it follows that

$$\mathbb{E}(J) < \binom{n}{r}\left(\frac{m}{n}\right)^t \geqslant \alpha.$$

Hence, there exists a choice of T for which $N(T)$ satisfies $I - J \geqslant \alpha$. Delete one vertex from each subset S of $N(T)$ of size r with fewer than m neighbours and let U be the remaining subset of $N(T)$. The set U has at least $I - J \geqslant \alpha$ vertices and all of its subsets of size r have at least m common neighbours. □

Lemma 2.5.4. *Let H be a bipartite graph with partition sets A and B, where the vertices of B have degree at most Δ. If G is a graph with a vertex subset U such that all subsets of U of size Δ have at least $|A| + |B|$ common neighbours, then H is a subgraph of G.*

Proof. We want to find an embedding of H in G given by an injection $A \cup B \to V(G)$. Let us start by defining an injection $f : A \to V(G)$ arbitrarily. Suppose that the current vertex to embed is $v_i \in B$. Let $N_i \subseteq A$ be those vertices in H adjacent to v_i, so that $|N_i| \leqslant \Delta$. Since $f(N_i)$ is a subset of U of cardinality at most Δ, there are at least $|A| + |B|$ vertices adjacent to all vertices in $f(N_i)$. As the total number of vertices already embedded is less than $|A| + |B|$, there is a vertex $w \in V(G)$ which is not yet used in the embedding and is adjacent to all vertices in $f(N_i)$. Set $f(v_i) = w$.

It is immediate from the above description that f provides an embedding of H in G. □

Theorem 2.5.5. *If H be a bipartite graph with partition sets A and B in which the vertices in B have degree at most Δ, then there exists a constant c such that $\mathsf{ex}(n, H) \leqslant cn^{2-\frac{1}{\Delta}}$.*

Proof. Suppose G is a graph on n vertices with at least $cn^{2-\frac{1}{\Delta}}$ edges. Then the average degree d of G satisfies $d \geqslant cn^{1-\frac{1}{\Delta}}$. Let $c = \max\left\{|A|^{1/\Delta}, \frac{3(|A|+|B|)}{\Delta}\right\}$. Then, using the fact that $r! \geqslant (\frac{r}{e})^r$, it is easy to check that

$$\begin{aligned}\frac{d^\Delta}{n^{\Delta-1}} - \binom{n}{\Delta}\left(\frac{|A|+|B|}{n}\right)^\Delta &\geqslant (2c)^\Delta - \frac{n^\Delta}{\Delta!}\left(\frac{|A|+|B|}{n}\right)^\Delta \\ &\geqslant (2c)^\Delta - \left(\frac{e(|A|+|B|)}{\Delta}\right)^\Delta \\ &\geqslant c^\Delta \geqslant |A|.\end{aligned}$$

Therefore, we can use Lemma 2.5.3 (with $t = \Delta = r$, $m = |A| + |B|$ and $\alpha = |A|$) to find a vertex subset U of G with $|U| = |A|$ such that al subsets of U of size Δ have at least $|A| + |B|$ common neighbours. The previous embedding lemma completes the proof. □

A graph is said to be d-degenerate if every subgraph contains a vertex of degree at most d. Equivalently, H is d-degenerate if there is an ordering $\{v_1, \dots, v_t\}$ of the vertices, such that any v_j has at most d neighbours v_i with $i < j$. An old conjecture of Burr and Erdős [9] states that if a bipartite graph H is d-degenerate, then $\mathsf{ex}(n, H) \leqslant cn^{2-\frac{1}{d}}$. This would be strictly stronger than the result we proved above. The best result currently known, due to Alon, Krivelevich and Sudakov [2], is that $\mathsf{ex}(n, H) \leqslant cn^{2-\frac{1}{4d}}$.

2.6 Paths and cycles

The problem of determining the maximum number of edges in a graph on n vertices if it contains no path with $k+1$ vertices was first considered by Erdős and Gallai [19]. In the following, given two graphs G and H, we denote by $G \cup H$ their disjoint union, and by $G + H$ their union together with al the edges between them.

Theorem 2.6.1 (Erdős-Gallai). *For every $k \geqslant 1$, $\mathsf{ex}(n, P_k) \leqslant \frac{k-1}{2}n$, with equality if and only of $n = kt$, in which case the extremal graph is $\bigcup_{i=1}^{t} K_k$.*

This bound is not difficult to come by. Indeed, consider a graph G with n vertices and more than $\frac{k-1}{2}n$ edges. By successively removing any vertex v with degree $d(v) < \frac{k-1}{2}$, we are left with a non-empty graph G' with n' vertices and more than $\frac{k-1}{2}n'$ edges, where every vertex u has degree $d(u) \geqslant \frac{k-1}{2}$.

We will find a path of length k in G'. Without loss of generality, assume that G' is connected (otherwise we can focus on any connected component of G' with the largest ratio of edges to vertices).

Consider a maximum length path $v_0 \dots v_t$ in G'. If $t \geqslant k$, then G', and hence G, contains a P_k. Otherwise, $t \leqslant k-1$. By maximality of the path, all the neighbours of v_0 and all the neighbours of v_t lie on the path.

If $v_0 v_t \in E(G')$, then $v_0 \dots v_t v_0$ is a cycle, and hence $v_i v_{i+1} \dots v_t v_0 \dots v_{i-1}$ is also a path of maximum length, and therefore we deduce that all the neighbours of v_i are also on the path. Thus, $v_0, \dots, v_t$ are all the vertices of G', so $|E(G')| = \binom{t+1}{2} \leqslant \frac{k-1}{2}n'$, a contradiction.

Thus, assume that v_0 is not adjacent to v_t. Since v_0 and v_t both have degree at least $\frac{k-1}{2}$ and all their neighbours are amongst $v_1, \dots, v_{t-1}$, there must exist some

v_i, v_{i+1} such that v_i is adjacent to v_t and v_{i+1} is adjacent to v_0. This gives us a cycle $v_0v_1\dots v_iv_tv_{t-1}\dots v_{i+1}v_0$. We thus conclude, as before, that $v_0,\dots,v_t$ are all the vertices of G', and get a contradiction. Therefore $t \geqslant k$, completing the proof.

In 1975, this result was improved by Faudree and Schelp [23], determining $\mathsf{ex}(n, P_k)$ for all $n > k > 0$ as well as the corresponding extremal graphs.

Theorem 2.6.2 (Faudree and Schelp)**.** *If G is a graph with $|V(G)| = kt + r$, $0 \leqslant r < k$, containing no P_k, then $|E(G)| \leqslant t\binom{k}{2} + \binom{r}{2}$, with equality if and only if G is either $\bigcup_{i=1}^{t} K_k$ or $\left(\bigcup_{i=1}^{t-l-1} K_k\right) \cup \left(K_{\frac{k-1}{2}} + \overline{K}_{\frac{k+1}{2}+lk+r}\right)$, for some l such that $0 \leqslant l < t$, when k is odd, $t > 0$, and $r = \frac{k\pm1}{2}$.*

Finally, Balister, Győri, Lehel and Schelp [5] considered the problem over all connected graphs and determined the extremal numbers as well as the corresponding extremal graphs. These extremal graphs are particular examples of graphs of the form $G_{n,k,s}$ which, for $n \geqslant k > 2s > 0$ are defined as $G_{n,k,s} = K_{k-2s} \cup \overline{K}_{n-k+s}$. Note that $|E(G_{n,k,s})| = \binom{k-s}{2} + s(n-k+s)$ and, since $k > 2s$, $G_{n,k,s}$ contains no P_k.

In particular, they proved the following.

Theorem 2.6.3. *Let G be a connected graph on n vertices that contains no P_k, for $n > k \geqslant 3$. Then*

$$|E(G)| \leqslant \max\left\{\binom{k-1}{2} + n - k + 1,\ \binom{\lceil\frac{k+1}{2}\rceil}{2} + \left\lfloor\frac{k-1}{2}\right\rfloor\left(n - \left\lceil\frac{k+1}{2}\right\rceil\right)\right\}.$$

If equality occurs, then G is either $G_{n,k,1}$ or $G_{n,k,\lfloor\frac{k-1}{2}\rfloor}$.

Our next objective is to consider the extremal problem for cycles of even lengths. Our main result is that $\mathsf{ex}(n, C_{2k}) \leqslant cn^{1+\frac{1}{k}}$, proved by Bondy and Simonovits [7] in 1974. For k =2, 3 and 5, that is, for C_4, C_6 and C_{10}, this is known to be sharp. A quick probabilistic argument, similar to that used earlier for complete bipartite graphs, gives the general lower bound.

Theorem 2.6.4. *There exists a constant c such that*

$$\mathsf{ex}(n, C_{2k}) \geqslant cn^{1+\frac{1}{2k-1}}.$$

There is also an explicit construction, due to Lazebnik, Ustimenko and Woldar [38], which does better giving $\mathsf{ex}(n, C_{2k}) \geqslant cn^{1+\frac{2}{3k-3}}$.

In order to prove the Bondy-Simonovits theorem, we will need two preliminary lemmas, both of which concern cycles with an extra chord.

Lemma 2.6.5. *Let H be a cycle with an extra chord. Let (A, B) be a non-trivial partition of $V(H)$. Then, unless H is bipartite with partition sets A and B, it contains paths of every length $l < |V(H)|$ which begin in A and end in B.*

Proof. Label the vertices of H as $0, 1, \ldots, t-1$, where $t = |V(H)|$. Suppose that H does not contain paths which start in A and end in B for every possible length $l < t$. We will focus on a particular class of path, saying that a path is *good* if it begins in A and ends in B and does not use the chord of H. Let s be the smallest integer such that there is no good path of length s. Then $s > 1$, since there is at least one edge between A and B (as we have assumed that (A, B) is a non-trivial partition). If this edge is a chord, it will automatically imply that there is some other edge across this partition. We also have that $s \leqslant \frac{t}{2}$. This is because, by symmetry, the existence of a good path of length j implies the existence of a good path of length $t - j$.

Let χ be the characteristic function of A. Then, for any j, we have that $\chi(j+s) = \chi(j)$, where addition is taken modulo t. Let $d = hcf(s, t)$. Then there are p and q such that $ps + qt = d$ and, therefore, $\chi(j) = \chi(j + d)$, for all j. But then there is no good path of length d. Therefore, since s was the smallest number with this property, $d = s$ and s divides t. This also implies that for every i which is not a multiple of s, there will be good paths of length i.

We will now find paths of all remaining lengths is, where $1 \leqslant i \leqslant \frac{t}{s} - 1$, by using the chord. Suppose first that the chord joins two vertices at distance r, where $1 < r \leqslant s$, say 0 and r. We know from above that there are good paths of length $s + r - 1$. In particular, there is some j such that $\chi(j) \neq \chi(j + s + r - 1)$. By shifting, we may assume that $-s < j \leqslant 0$. Therefore, since $j + s + r - 1 \geqslant r$ and $\chi(j) \neq \chi(j + is + r - 1)$, the path $j, j+1, \ldots, 0, r, r+1, \ldots, j + is + r - 1$ is a path of length is beginning in A and ending in B. We need to verify that $j + is + r - 1 < t + j$, that is, that it doesn't loop all the way around, but this follows easily for $i \leqslant \frac{t}{s} - 1$.

We therefore assume that the chord is $0r$, where $s < r < t - s$. Let $-s < j < 0$ and consider the paths $j, j+1, \ldots, 0, r, r-1, \ldots, r - j - s + 1$ and $s + j, s + j - 1, \ldots, 0, r, r+1, \ldots, r - j - 1$, each of length s. If either of them is a path starting in A and ending in B, we may extend it to produce a well-behaved path of length is until the number of unused vertices in the two arcs defined by the chord is less than s in both arcs. At this point, $is + 1 \geqslant t - 2(s-1)$ and, since s divides t, $is = t - s$, so we already have everything. Similarly, if either of the paths $0, r, r-1, \ldots, r - s + 1$ or $0, r, r+1, \ldots, r + s - 1$ begin in A and end in B, then H contains well-behaved paths of all lengths less than t.

We may, therefore, assume that, for $-s < j < 0$,

$$\chi(r - j + 1) = \chi(r - j - s + 1) = \chi(j) = \chi(s + j) = \chi(r - j - 1).$$

The first and third equalities are by shifting. The second and fourth follow from the fact that the paths $j, j+1, \ldots, 0, r, r-1, \ldots, r-j-s+1$ and $s+j, s+j-1, \ldots, 0, r, r+1, \ldots, r-j-1$ must each have both endpoints in one of A or B. Similarly, we may assume that $\chi(r+s+1) = \chi(r+s-1)$. This implies that $\chi(i) = \chi(i+2)$ for every vertex i. Therefore, $s = 2$.

Thus, we may conclude that t is even and that the vertices of the cycle alternate between A and B. It is easy now to see that if the chord is contained in one of A and B, then the graph contains paths of all length less than t which start in A and end in B. Hence, the chord goes between A and B and H is bipartite, as required. □

The second lemma we need is a condition for a graph to contain a cycle with an extra chord.

Lemma 2.6.6. *Any bipartite graph G with minimum degree $d \geqslant 3$ contains a cycle of length at least $2d$ with an extra chord.*

Proof. Let P be a longest path in G, visiting vertices $x_1, \ldots, x_p$ in that order. The vertex x_1 has $d \geqslant 3$ neighbours in G and, by maximality of P, they all lie in P. Suppose that they are $x_{i_1}, \ldots, x_{i_d}$ with $i_1 < \cdots < i_d$. Every two neighbours of x_1 must have distance at least 2, since G is bipartite. Therefore, since $i_1 \geqslant 2$, we must have that $i_d \geqslant 2d$. The required cycle with chord is formed by taking the path from x_1 to x_{i_d} and adding the edges $x_1 x_{i_1}$ and $x_1 x_{i_d}$. □

We are now ready to prove the Bondy-Simonovits theorem.

Theorem 2.6.7 (Bondy-Simonovits)**.** *For any natural number $k \geqslant 2$, there exists a constant c such that*

$$\mathsf{ex}(n, C_{2k}) \leqslant cn^{1+\frac{1}{k}}.$$

Proof. Suppose that G is a C_{2k}-free graph on n vertices with at least $cn^{1+\frac{1}{k}}$ edges. Then the average degree of G is at least $2cn^{\frac{1}{k}}$ so, reminding ourselves that every graph contains a subgraph whose minimum degree is at least half the average degree of the graph, there exists some subgraph H for which the minimum degree is at least $cn^{\frac{1}{k}}$.

Fix an arbitrary vertex x of H. Let $i \geqslant 0$ and let V_i be the set of vertices that are at distance i from x with respect to the graph H. In particular, $V_0 = \{x\}$ and $V_1 = N(x)$. Let $v_i = |V_i|$ and let H_i be the bipartite subgraph $H[V_{i-1}, V_i]$ induced by the disjoin sets $V_{i=1}$ and V_i.

We claim that, for $1 \leqslant i \leqslant k-1$, none of the graphs $H[V_i]$ or H_{i+1} contain a bipartite cycle of length at least $2k$ with a chord; suppose otherwise and let F be such a cycle, contained in $H[V_i]$. Let $Y \cup Z$ be the bipartition of $V(F)$.

Let $T \subseteq H$ be a breadth-first search tree beginning at x. That is, we begin at the root vertex x. The first layer will consist of the neighbourhood of x, labelling them as we uncover them. At the j-th step, we look at layer $j-1$. For the first vertex in the ordering, we look at its neighbours that have not yet occurred and label them as they occur. Then we do the same in order for every vertex in the $(j-1)$-st level. This will give us the j-th level with all vertices labelled.

Let y be the vertex which is farthest from x in the tree T and which still dominates the set Y, that is, every vertex in Y is a descendant of y.

Clearly, the paths leading from y to Y must branch at y. Pick one such branch (leading to a non-trivial subset of Y), defined by some child z of y and let A be the set of descendants of z which lie in Y. Let $B = (Y \cup Z)\backslash A$. Since $Y\backslash A \neq \emptyset$, B is not an independent set of F.

Let l be the distance between x and y. Then $l < i$ and $2k-2i+2l < 2k \leqslant |V(F)|$. By Lemma 2.6.5, since F is not bipartite with respect to the partition into A and B, we can find a path P in F of length $2k-2i+2l$ which starts in $a \in A$ and ends in $b \in B$. Since the path has even length and the partition into Y and Z is bipartite, b must be in Y. Let P_a and P_b be the unique paths in T that connect y to a and b, respectively. These intersect only at y, since a is a descendant of z and b is not. Also, they each have length $i-l$. Therefore, the union of the paths P, P_a and P_b forms a C_{2k} in H, which contradicts our assumption. The proof follows similarly for H_{i+1} if we take $Y = V(F) \cap V_i$, thus proving our claim that, for $1 \leqslant i \leqslant k-1$, none of the graphs $H[V_i]$ or H_{i+1} contain a bipartite cycle of length at least $2k$ with a chord.

We also know, by Lemma 2.6.6, that if a bipartite graph has minimum degree $d \geqslant 3$, then it contains a cycle of length at least $2d$ with an extra chord. We may, therefore, assume that, for $1 \leqslant i \leqslant k-1$, the average degrees $d(H[V_i])$ and $d(H_{i+1})$ of $H[V_i]$ and H_{i+1}, respectively, satisfy

$$d(H[V_i]) \leqslant 4k-4 \quad \text{and} \quad d(H_{i+1}) \leqslant 2k-2.$$

For example, if $H[V_i]$ had average degree greater than $4k-4$, it would contain a bipartite subgraph with average degree greater than $2k-2$ and, therefore, a bipartite subgraph with minimum degree greater than $k-1$. This would then imply that the graph contained a bipartite cycle of length at least $2k$ with a chord, which would contradict the claim. The bound for $d(H_{i+1})$ follows similarly.

We will now show inductively that, provided n is sufficiently large,

$$\frac{|H_{i+1}|}{v_{i+1}} \leqslant 2k$$

for every $0 \leqslant i \leqslant k-1$. For $i=0$, this is true since every edge in V_1 is connected to x by only one edge. Suppose that we want to prove it for some $i>0$. Then, by the induction hypothesis and the bound on $d(H[V_i])$,

$$\begin{aligned}|E(H_{i+1})| = &\sum_{y\in V_i} d_{V_{i+1}}(y) \geqslant \left(\delta(H) - \frac{4k-4}{2} - 2k\right)v_i \\ \geqslant &(cn^{\frac{1}{k}} - 4k + 2)v_i \geqslant \frac{c}{2}n^{\frac{1}{k}}v_i \geqslant 2kv_i.\end{aligned}$$

In particular, $V_{i+1} \neq \emptyset$ and the average degree of vertices of V_i with respect to H_{i+1} is at least $2k$. But since $d(H_{i+1}) \leqslant 2k-2$, we must have that the average degree of V_{i+1} with respect to H_{i+1} is at most $2k-2$, that is, $|E(H_{i+1})| \leqslant (2k-2)v_{i+1}$, implying the required bound.

Note now that we have

$$\frac{c}{2}n^{\frac{1}{k}}v_i \leqslant |E(H_{i+1})| \leqslant 2kv_{i+1}.$$

Therefore,

$$\frac{v_{i+1}}{v_i} \geqslant \frac{c}{4k}n^{\frac{1}{k}}.$$

This, in turn, implies that

$$v_k \geqslant \left(\frac{c}{4k}\right)^k n.$$

This is a contradiction if $c \geqslant 4k$, completing the proof. □

Our next aim is to consider the extremal number for odd cycles. We already know, by the Erdős-Stone-Simonovits theorem, that $\mathsf{ex}(n, C_{2k+1}) \approx \frac{n^2}{4}$. Here, we will use the so-called *stability approach* to prove that, for n sufficiently large, $\mathsf{ex}(n, C_{2k+1}) = \left\lfloor \frac{n^2}{4} \right\rfloor$.

The idea behind the stability approach is to show that a C_{2k+1}-free graph with roughly the maximal number of edges is approximately bipartite. Then one uses this approximate structural information to prove an exact result.

Lemma 2.6.8. *For every natural number $k \geqslant 2$ and $\epsilon > 0$, there exists $\delta > 0$ and a natural number n_0 such that, if G is a C_{2k+1}-free graph on $n \geqslant n_0$ vertices with at least $(\frac{1}{4} - \delta)n^2$ edges, then G may be made bipartite by removing at most ϵn^2 edges.*

Proof. We will prove the result for $\delta = \frac{\epsilon^2}{100}$ and n sufficiently large. We begin by finding a subgraph G' of G with large minimum degree. We do this by deleting vertices one at a time, forming graphs $G = G_0, G_1, \dots, G_l$, at each stage removing a vertex with degree less than $\frac{1}{2}(1 - 4\delta^{\frac{1}{2}})|V(G_l)|$, should it exist. By doing so, we delete at most $4\delta^{\frac{1}{2}}n$ vertices. Otherwise, we would have a C_{2k+1}-free graph G' on $n' = (1 - a\delta^{\frac{1}{2}})n$ vertices with at least

$$\begin{aligned}
|E(G')| &> |E(G)| - \sum_{i=n'+1}^{n} \frac{1}{2}(1 - 4\delta^{\frac{1}{2}})i \\
&\geqslant \left(\frac{1}{4} - \delta\right)n^2 - \frac{1}{2}(1 - 4\delta^{\frac{1}{2}})\left(\binom{n+1}{2} - \binom{n'+1}{2}\right) \\
&\geqslant \frac{n'^2}{4} + 2\delta^{\frac{1}{2}}n^2 - 4\delta n^2 - \delta n^2 - \frac{1}{2}(1 - 4\delta^{\frac{1}{2}})(n - n')n \\
&= \frac{n'^2}{4} + 2\delta^{\frac{1}{2}}n^2 - 5\delta n^2 - 2\delta^{\frac{1}{2}}n^2 + 8\delta n^2 \\
&\geqslant \frac{n'^2}{4}(1 + \delta)
\end{aligned}$$

edges. But, by the Erdős-Stone-Simonovits theorem, for n sufficiently large, G' contains a copy of C_{2k+1}, so we have reached a contradiction. We therefore have a subgraph G' with $n' \geqslant (1 - 4\delta^{\frac{1}{2}})n$ vertices and minimum degree at least $\frac{1}{2}(1 - 4\delta^{\frac{1}{2}})n'$.

Since $\mathsf{ex}(n, C_{2k+1}) = o(n^2)$, we know that for n (and therefore n') sufficiently large, the graph G' will contain a cycle of length $2k$. Let $a_1 a_2 \dots a_{2k}$ be such a cycle. Note that $N(a_1)$ and $N(a_2)$ cannot intersect, for otherwise there would be a cycle of length $2k + 1$. Moreover, each of the two neighbourhoods must contain a small number of edges. Indeed, if $N(a_1)$ contained more than $4kn'$ edges, then Lemma would imply that there was a path of length $2k$ in $N(a_1)$. But then the endpoints could be joined to a_1 to give a cycle of length $2k + 1$. Therefore, we have two large disjoint sets $N(a_1)$ and $N(a_2)$, each of size at least $\frac{1}{2}(1 - 4\delta^{\frac{1}{2}})n' \geqslant \frac{1}{2}(1 - 8\delta^{\frac{1}{2}})n$, such that each contains at most $4kn'$ edges. We can make the graph bipartite by deleting all the edges between $N(a_1)$ and $N(a_2)$ and all of the edges which have one end in the complement of these two sets. In total, this is at most

$$8kn' + 8\delta^{\frac{1}{2}}n^2$$

edges. Hence, for n sufficiently large and $\delta = \frac{\epsilon^2}{100}$, we wil have deleted at most ϵn^2 edges, which gives the required result. □

We showed that a C_{2k+1}-free graph with roughly $\frac{n^2}{4}$ edges must be approximately bipartite. We will now refine this structure to prove that the graph must be exactly bipartite for C_{2k+1}-free graphs of maximum size.

Theorem 2.6.9. *For n sufficiently large,*

$$\mathsf{ex}(n, C_{2k+1}) = \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Proof. Let G be a C_{2k+1}-free graph on n vertices with the maximum number of edges. Then G has at least $\lfloor \frac{n^2}{4} \rfloor$ edges. Note that it suffices to prove the result in the case where G has minimum degree at least $\frac{1}{2}(1 - 4\epsilon^{\frac{1}{2}})n$. For suppose that we knew the result under this assumption for all $n \geqslant n_0$. As in the previous lemma, we form a sequence of graphs $G = G_0, G_1, \ldots, G_l$. If there is a vertex in G_l of degree at least $\frac{1}{2}(1 - 4\epsilon^{\frac{1}{2}})|V(G_l)|$, we remove it, forming G_{l+1}. This process must stop before we reach a graph G' with $n' = (1 - 4\epsilon^{\frac{1}{2}})n$ vertices. Otherwise, we would have a graph with n' vertices and more than $(1 - \epsilon)\frac{n'^2}{4}$ edges. It would, therefore, for n sufficiently large, contain a copy of C_{2k+1}, which would be a contradiction. When we reach the required graph, we will have a graph with $n' > (1 - 4\epsilon^{\frac{1}{2}})n$ vertices, minimum degree at least $\frac{1}{2}(1 - 4\epsilon^{\frac{1}{2}})n'$ and more than $\lfloor \frac{n'^2}{4} \rfloor$ edges, so we will have a contradiction if the removal process begins at all. Hence, we may assume that the minimum degree of G is at least $\frac{1}{2}(1 - 4\epsilon^{\frac{1}{2}})n$.

By the previous lemma, we know that G is approximately bipartite between two sets of size roughly $\frac{n}{2}$. Consider a bipartition $V(G) = A \cup B$ such that $|E(A)| + |E(B)|$ is minimised. Then $|E(A)| + |E(B)| < \epsilon n^2$, where ϵ may be taken to be arbitrarily small provided that n is sufficiently large. We may assume that A and B have size $(\frac{1}{2} \pm \epsilon^{\frac{1}{2}})n$. Otherwise, $|E(G)| < |A||B| + \epsilon n^2 < \frac{n^2}{4}$, contradicting the choice of G as having maximum size. Let $d_A(x) = |A \cap N(x)|$ and $d_B(x) = |B \cap N(x)|$ for any vertex x. Note that for any $a \in A$, $d_A(a) \leqslant d_B(a)$. Otherwise, we could improve the partition by moving vertex a to B. Similarly, $d_B(b) \leqslant d_A(b)$ for any $b \in B$.

Let $c = 2\epsilon^{\frac{1}{2}}$. We claim that there are no vertices $a \in A$ with $d_A(a) \geqslant cn$.If $d_A(a) \geqslant cn$, then $d_B(a) \geqslant cn$ as well. Moreover, $A \cap N(a)$ and $B \cap N(a)$ span a bipartite graph with no path of length $2k - 1$ and, therefore, there are at most $4kn$ edges between them. For n sufficiently large, this gives $(cn)^2 - 4kn > |E(A)| + |E(B)|$ missing edges between A and B. Hence, $|E(G)| < |A||B| \leqslant \frac{n^2}{4}$,a contradiction. Similarly, there no vertices $b \in B$ with $d_B(b) \geqslant cn$.

Now suppose that there is an edge in A, say aa'. Then

$$|N_B(a) \cap N_B(a')| > d(a) - cn + d(a') - cn - |B| > \left(\frac{1}{2} - 9\epsilon^{\frac{1}{2}}\right)n.$$

Let $A' = A \backslash \{a, a'\}$ and $B' = N_B(a) \cap N_B(a')$. There is no path of length $2k - 1$ of the form $b_1 a_1 b_2 a_2 \ldots b_{k-1} a_{k-1} b_k$ between A' and B'. But this implies that there is no path of any type of length $2k$ (remember that, since the graph is bipartite, a path must alternate sides). This implies that the number of edges between A' and

B' is at most $4kn$. This, in turn, implies that the number of edges in the graph is at most

$$|E(A', B')| + |E(A \backslash A', V(G))| + |E(V(G), B \backslash B')| \leqslant 4kn + 2n + 10\epsilon^{\frac{1}{2}}n^2,$$

a contradiction for n large. □

More generally, there is a result of Simonovits [54] which shows that if H is a graph with $\chi(H) = t$ and $\chi(H \backslash e) < t$, for some edge e, then $\mathsf{ex}(n, H) = \mathsf{ex}(n, K_t)$ for n sufficiently large. We say that such graphs are colour-critical and it is easy to verify that odd cycles are colour-critical.

Let G be a graph of order n. A *Hamiltonian path* is a path that visits each vertex of G exactly once. Similarly, a *Hamiltonian cycle* is a cycle that visits each vertex of G exactly once. The graph G is said to be *Hamiltonian* if it contains a Hamiltonian cycle. As opposed to the previous results concerning cycles, the question "what is the maximum number of edges in a graph so that it does not contain a Hamiltonian cycle?" is not particularly interesting, as the answer is close to the maximum possible number of edges. Consider the following graph G. Take a K_{n-1} together with an isolated vertex v, and add one edge between K_{n-1} and v. Then G contains Hamiltonian cycle, since $d(v) = 1$ and no cycle can go through v. On the other hand

$$|E(G)| = \binom{n-1}{2} + 1 = \binom{n}{2} - (n-2)$$

Thus, the edge density of G is

$$\frac{|E(G)|}{\binom{n}{2}} = 1 - \frac{2(n-2)}{n(n-1)} \to 1,$$

as $n \to \infty$. This means that even if the graph contains almost all the possible edges, it could still be non-Hamiltonian.

Note that the degrees of the example we constructed above are distributed fairly unevenly. There are $n-1$ vertices with degree at least $n-2$ and one vertex with degree 1. Thus, a more interesting question is "how can we guarantee the existence of Hamiltonian cycles by lower bounding the minimum degree of the graph?". The answer is given in the following theorem, which was proved by Dirac [17] in 1952.

Theorem 2.6.10 (Dirac)**.** *Let G is a graph of order $n \geqslant 3$ and $\delta(G) \geqslant \frac{n}{2}$, then G is Hamiltonian.*

Proof. We claim that G is connected. Suppose otherwise and pick one of the smallest components of G. This must contain at most $\frac{n}{2}$ vertices. Hence, any vertex in this component has degree at most $\frac{n}{2} - 1$, a contradiction.

Now suppose that G is not Hamiltonian. Consider a maximal path P with length $l \leqslant n-1$. That is, $P = x_0x_1 \dots x_l$, where $x_ix_{i+1} \in E(G)$ for all $0 \leqslant i \leqslant l-1$. Since P os maximal, the neighbours of x_0 and the neighbours of X_l must all lie on P. Let $A = N(x_1)$ and $B = \{x_{i+1} \colon x_i \in N(x_l)\}$. Since $\delta(G) \geqslant \frac{n}{2}$, we have that $|A|, |B| \geqslant \frac{n}{2}$. On the other hand, it is easy to see that $x_0 \notin A$ and $x_0 \notin B$. Hence, $A \cup B \subseteq \{x_1, \dots, x_l\}$. Thus, $|A \cup B| \leqslant l \leqslant n-1$. This implies that $A \cap B \neq \emptyset$, as otherwise we would have that $|A \cup B| \geqslant \frac{n}{2} + \frac{n}{2} = n$.

Suppose $x_t \in A \cup B$ for some t. Then consider the cycle

$$C = x_1x_tx_{t+1} \dots x_{l-1}x_lx_{t-1}x_{t-2} \dots x_2x_1.$$

If $l = n$, then C is a Hamiltonian cycle, contradicting our assumption.

If $l < n$, there exists at least one vertex $v \notin C$. However, G is connected, hence there exists a path from v to some vertex x_r in C. We construct a path P' as follows. We start from v, to x_r, and then traverse C to x_{r-1}. The length of P' is at least $l+1$, contradicting the maximality of P.

Thus, G is Hamiltonian. □

The following generalisation was obtained by Ore [46] in 1960.

Theorem 2.6.11 (Ore)**.** *Let G be a graph on $n \geqslant 3$ vertices. If $d(x) + d(y) \geqslant n$ for each pair of non-adjacent vertices x and y, then G is Hamiltonian.*

Proof. Suppose, towards a contradiction, that G is not Hamiltonian. Choose any two non-adjacent vertices in G and add an edge between them. We keep doing so until we obtain a graph H which is Hamiltonian.

Let G' be the graph obtained immediately before H in the process of adding edges, and suppose that xy is the edge added to G' in order to obtain H. Let $z_1 \dots z_nz_1$ be the Hamiltonian cycle in H. This must use the edge xy at some point, otherwise G' would be Hamiltonian and the process of adding edges would stop before we reached H. If $z_nz_1 = xy$, then $z_1 \dots z_n$ is a Hamiltonian path in G'. Otherwise, there is some r such that $1 \leqslant r < n$, $z_r = x$ and $z_{r+1} = y$. Now, $z_{r+1} \dots z_nz_1 \dots z_r$ is a Hamiltonian path in G'. Note that, either way, all the edges used in this path appear in G'; it is only xy that appears in H but not in G'. Relabel the vertices so that this path is $x_1 \dots x_n$.

Suppose we could find a vertex x_i such that x is adjacent to x_i, and y is adjacent to x_{i-1}. Then, $xx_i \dots x_{n-1}yx_{i-1} \dots x$ would be a Hamiltonian cycle on G', which would contradict our assumption.

It remains to show that there must be such a vertex x_i. Note that since G' is obtain from G by adding edges, it still satisfies the hypotheses of the theorem. As x and y are not adjacent in G', we have that $d(x) + d(y) \geqslant n$. Hence, $N_{G'}(x)$ and $N_{G'}(y)$ are subsets of $\{2, \ldots, n\}$ containing at least n elements between them. It follows that they must intersect non-trivially, thus providing us with the desired vertex x_i which can be any vertex in $N_{G'}(x) \cap N_{G'}(y)$. $\square$

The following result, which associates the existence of a Hamiltonian cycle with the connectivity and independence number of G, was proved by Chvátal and Erdős [12] in 1972.

Theorem 2.6.12 (Chvátal-Erdős)**.** *If G is a graph on $n \geqslant 3$ vertices such that $\alpha(G) \leqslant \kappa(G)$, then G is Hamiltonian.*

Proof. Let $\kappa(G) = k$ and let $C = v_1 \ldots v_t v_1$ be a longest cycle in G. Let us assume, towards a contradiction, that C is not a Hamiltonian cycle. Pick a vertex $v \in V(G)\backslash C$ and a $v - C$ fan $\mathcal{F} = \{P_i \colon i \in I\}$ in G, where $I \subseteq [n]$ and each P_i ends in v_i. Let $\mathcal{F}$ be chosen with maximum cardinality; then $vv_j \notin E(G)$ for any $j \notin I$, and by Menger's theorem

$$|\mathcal{F}| \geqslant \min\{k, |C|\}. \tag{1}$$

For every $i \in I$, we have that $i + 1 \notin I$, otherwise $(C \cup P_i \cup P_{i+1}) - v_i v_{i+1}$ would be a longer cycle than C. Thus, $|\mathcal{F}| < |C|$ and hence $|I| = |CF| \geqslant k$ by (1).

Furthermore, $v_{i+1}v_{j+1} \notin E(G)$ for all $i, j \in I$, as otherwise $(C \cup P_i \cup P_j) + v_{i+1}v_{j+1} - v_i v_{i+1} - v_j v_{j+1}$ would be a cycle longer than C. Hence, $\{v_{i+1} \colon i \in I\} \cup \{v\}$ is a set of $k + 1$ or more independent vertices in G, contradicting $\alpha(G) \leqslant k$. $\square$

Chapter 3

Minors

3.1 Minors

In this section, we consider the analogue of Turán's theorem for graph minors, that is, how many edges on n vertices can force a graph to contain a K_p minor. The answer is that, unlike for K_p subgraphs, a number of edges linear in n is enough; it suffices to assume that the graph has large enough average degree (depending on p).

Proposition 3.1.1. *Every graph of average degree at least 2^{p-2} has a K_p minor, for all integers $p \geqslant 2$.*

Proof. We apply induction on p. For $p \leqslant 2$ the assertion is trivial. For the induction step let $p \geq 3$, and let G be any graph of average degree at least 2^{p-2}. Then $\epsilon(G) \geq 2^{p-3}$. Let H be a minimal minor of G with $\epsilon(H) \geq 2^{p-3}$ and pick a vertex $x \in V(H)$. By the minimality of H, x is not isolated and each of its neighbours y has at least 2^{p-3} common neighbours with x; otherwise, by contracting the edge xy, we would lose one vertex and at most 2^{p-3} edges, yielding a smaller minor H' with $\epsilon(H') \geq 2^{p-3}$. The subgraph induced in H by the neighbours of x, therefore, has minimum degree at least 2^{p-3} and hence has a K_{p-1} minor by the induction hypothesis. Together with x, this yields the desired K_p minor of G. □

Mader [40] was the first one to prove the existence of a constant $c_p \leqslant \lceil 2c_{p-1} \rceil$ such that $\mathsf{ex}(n, IK_p) \leqslant c_p n$, for each $p \geqslant 2$. Later, the same author showed that $c_p \leqslant 8p\log(p)$. De la Vega [16] and Thomason [57], by using random graphs, advanced in the study of c_p proving that $c_p \geqslant \frac{1}{4}\sqrt{\log(p)}$. The study was improved by Kostochka [32], who showed that this is the correct order of magnitude for c_p, and Thomason [57], who proved that $c_p \leqslant 2{,}68p\sqrt{\log(p)}$. Finally, Thomason [58] completed the problem by showing that $c_p = (\alpha + o(1))t\sqrt{\log t}$, where $\alpha = 0{,}319...$ is an explicit constant.

All of the above provide us with interesting upper bounds for the extremal number $\mathsf{ex}(n, IK_p)$ in an asymptotical way, that is, when p is a fixed integer and n is much larger than p. Regarding the lower bounds, by considering the graph $G = K_{p-1} + \overline{K_{n-p+2}}$, we obtain

$$\mathsf{ex}(n, IK_p) \geqslant (p-2)n - \binom{p-1}{2}. \tag{1}$$

However, inequality (1) is an equality only for small values of p. Dirac [18] proved that equality holds for $p \leqslant 5$, and Mader [41] proved it for $p \leqslant 7$. Jørgensen [29] obtained the exact value

$$\mathsf{ex}(n, IK_p) = \begin{cases} 6n - 20, & \text{if 5 divides } n \\ 6n - 21, & \text{otherwise.} \end{cases}$$

for $p = 8$.

Cera *et al.* [10] found a lower bound that improves (1) and is best possible for infinitely many values of n and p. This lower bound also allows us to prove that every Turán graph $T^r(n)$ contains K_p as a minor for all $n \geqslant 2p - 2$. Their proof goes as follows.

Let us first relate the minimum cardinality of a vertex cover of a certain graph with the minimum number of edge contractions that are necessary in a graph in order to get a complete graph.

Lemma 3.1.2. *Let G be a graph and let $U = \{v_1, \ldots, v_p\}$ be a subset of $V(G)$. Let us denote by $s \leqslant p$ the minimum cardinality of a vertex cover of $H = \overline{G}[U]$. If G is contractible to a K_p with vertex set U, then the minimum number of edge contractions that are necessary to obtain K_p is at least s.*

Proof. If H has no edges, then the result holds trivially. We therefore assume that $E(H) \neq \emptyset$. Let $W = \{w_1, \ldots, w_s\} \subseteq U$ be a vertex cover of H with minimum cardinality. Suppose that there exists a set of edges $E' = \{e_1, \ldots, e_r\} \subseteq E(G)$, with $r < s$, such that the graph obtained from G by contracting these edges contains a complete K_p with vertex set U. Then, there exists at least one vertex w_j such that none of the edges of E' is contracted in w_j. But since G is contractible to a K_p with vertex set U, we have that

$$N_H(w_j) \subseteq \bigcup_{i=1, i\neq j}^{s} N_H(w_i).$$

Hence, $W \backslash \{w_j\}$ is a vertex cover of H with cardinality $s - 1$, contradicting the minimality of W. $\square$

Lemma 3.1.2 allows us to deduce a sufficient condition in order to show that a graph G is contractible to a complete graph K_p, as is stated in the following result.

Proposition 3.1.3. *Let n, p be positive integers with $n \geqslant p$, and let G be a graph on n vertices. If, for each subset U of $V(G)$, the minimum cardinality of a vertex cover of $\overline{G}[U]$ is $s > n - p$, then G does not contain K_p as a minor.*

Proposition 3.1.3 permits us to split the region of pairs (n, p), with $n \geqslant p$, into two parts separated by the curve $n = 2p - 3$. On the one hand, there exist Turán graphs on n vertices not containing K_p as a minor, if $p \leqslant n \leqslant 2p - 3$. On the other hand, every Turán graph on n vertices contains K_p as a minor, if $n \geqslant 2p - 2$.

Theorem 3.1.4. *Let n, p, r be positive integers such that $p \leqslant n \leqslant 2p - 3$ and $2 \leqslant r \leqslant 2p - n - 1$. Then, the Turán graph $T^r(n)$ is not contractible to K_p.*

Proof. Let us denote by C_i, with $i \in \{1, \ldots, r\}$, the partition sets of the vertices of $T^r(n)$. Let $U = \{v_1, \ldots, v_p\}$ be any vertex set of $T^r(n)$ and $H = \overline{T^r(n)}[U]$. Finally, let $r_i = |U \cap C_i|$, for $i = 1, \ldots, r$, and consider the set $I = \{j \in \{1, \ldots, r\} \colon r_j \geqslant 1\}$.

Observe that $H = \bigcup_{i \in I} K_{r_i}$. Then, if we denote by s the minimum degree cardinality of a vertex cover of H, it is clear that

$$s = \sum_{i \in I}(r_i - 1) = \sum_{i \in I} r_i - |I| \geqslant p - r \geqslant p - (2p - n - 1) > n - p.$$

Thus, by applying Proposition 3.1.3, $T^r(n)$ is not contractible to K_p. □

As an immediate consequence of Theorem 3.1.4, we have the following.

Corollary 3.1.5. *Let p, n be positive integers such that $p \leqslant n \leqslant 2p - 3$. Then,*

$$\mathsf{ex}(n, IK_p) \geqslant t_{2p-n-1}(n).$$

Finally, we shall prove that the pairs of values of n and p described in Theorem 3.1.4 are the only ones for which there are Turán graphs not containing K_p as a minor.

Theorem 3.1.6. *Let n, p, r be positive integers such that $n \geqslant 2p - 2$ and $r \geqslant 2$. Then, every Turán graph $T^r(n)$ contains K_p as a minor.*

Proof. If $r \geqslant p$, the result clearly holds. We can therefore assume that $2 \leqslant r \leqslant p-1$. Let us denote by C_i, with $i \in \{1, \ldots, r\}$, the partition sets of vertices of $T^r(n)$. Let U be a vertex set of $T^r(n)$ chosen in such a way that $|U \cap C_i| = \left\lceil \frac{p}{r} \right\rceil$, for $i \in \{1, \ldots, p - r\lfloor \frac{p}{r} \rfloor\}$, and $|U \cap C_i| = \left\lfloor \frac{p}{r} \right\rfloor$, for $i \in \{p - r\lfloor \frac{p}{r} \rfloor + 1, \ldots, r\}$. Clearly, $|U| = p$ and $T^r(n)[U]$ is a Turán graph $T^r(p)$. Therefore, if $H = \overline{T^r(n)}[U]$, we deduce that H is a graph on p vertices formed by r disjoint copies of complete graphs. Some of them are copies of $K_{\lceil \frac{p}{r} \rceil}$ and the rest of them are copies of $K_{\lfloor \frac{p}{r} \rfloor}$. Thus, the minimum cardinality of a vertex cover of H is

$$s = p - r \leqslant p - 2 = (2p - 2) - p \leqslant n - p.$$

Notice that the edges of H are the necessary ones in $T^r(n)[U]$ to be a complete graph K_p, so we need to contract some edges in $T^r(n)$ in order to obtain a K_p with vertex set U. Let $W = \{w + 1, \ldots, w_s\}$ be a vertex cover of H with minimum cardinality and consider the bipartite graph B whose vertex classes are W and $Z = V(T^r(n))\backslash U$ defined in such a way that a vertex w_i is adjacent to v_j in B if $w_i v_j \in E(T^r(n))$.

If B has a perfect matching M, then it suffices to contract in $T^r(n)$ the edges of M to obtain a new graph $T^r(n)/M$ containing a copy of K_p with vertex set U. Let us see that a perfect matching in B does, indeed, exist.

Let A be a subset of W. If there exist $w, w' \in A$ such that $w \in U \cap C_j$ and $w' \in U \cap C_k$, with $j \neq k$, then $N(A) = Z$, because $N(\{w\}) = Z\backslash(Z \cap C_j)$ and $N(\{w'\}) = Z\backslash(Z \cap C_k)$. Then, $|N(A)| = |Z| = n - p \geqslant s = |W| \geqslant |A|$.

For each fixed $j \in \{1, \ldots, r\}$, we have

$$\begin{aligned}
\sum_{i=1, i\neq j}^{r} |Z \cap C_i| &= n - p - |Z \cap C_j| \\
&= (n - p - |C_j| + 1) + (|C_j| - 1 - |Z \cap C_j|) \\
&= (n - p - |C_j| + 1) + |W \cap C_j| \\
&\geqslant \left(n - p - \left\lceil \frac{n}{r} \right\rceil + 1\right) + |W \cap C_j| \\
&\geqslant \left(n - p - \left\lceil \frac{n}{2} \right\rceil + 1\right) + |W \cap C_j| \\
&= \left\lfloor \frac{n - 2p + 2}{2} \right\rfloor + |W \cap C_j| \\
&\geqslant |W \cap C_j|,
\end{aligned}$$

because $n \geqslant 2p - 2$. Thus, by Hall's theorem, there exists a perfect matching in the graph B and, therefore, $T^r(n)$ contains K_p as a minor. $\square$

So far, we have seen that we can force a graph to contain K_p as a minor by raising its average degree, that is, we have the following.

Theorem 3.1.7. *There exists a constant $c \in \mathbb{R}$ such that, for every $n \in \mathbb{N}$, every graph G of average degree $d(G) \geqslant cp\sqrt{\log p}$ contains K_p as a minor. Up to the value of c, this bound is best possible as a function of p.*

Another interesting result is that we can force a K_p minor in a graph simply by raising its girth. This may seem questionable as a result, but it seems more believable if, rather than trying to force a K_p minor directly, we instead try to force a minor just of large minimum or average degree -which suffices, by Theorem 3.1.7. For if the girth g of a graph is large then the ball $\left\{v : d(x,v) < \left\lfloor \frac{g}{2} \right\rfloor \right\}$ around a vertex x induces a tree with many leaves, each of which sends all but one of its incident edges away from the tree. Contracting enough disjoint such trees, we can thus hope to obtain a minor of large average degree, which in turn will have a large complete minor.

The following lemma realizes this idea.

Lemma 3.1.8. *Let $d, k \in \mathbb{N}$ with $d \geqslant 3$, and let G be a graph of minimum degree $\delta(G) \geqslant d$ and girth $g(G) \geqslant 8k+3$. Then G has a minor H of minimum degree $\delta(H) \geqslant d(d-1)^k$.*

Proof. Let $X \subseteq V(G)$ be maximal with $d(x,y) > 2k$ for all distinct $x, y \in X$. Foe each $x \in X$ put $T_x^0 := \{x\}$. Given $i < 2k$, assume that we have defined disjoint trees $T_x^i \subseteq G$ (one for each $x \in X$) whose vertices together are precisely the vertices at distance at most i from X in G. Joining each vertex at distance $i+1$ from X to a neighbour at distance i, we obtain a similar set of disjoint trees T_x^{i+1}. As every vertex of G has distance at most $2k$ from X (by the maximality of X), the trees $T_x := T_x^{2k}$ obtained this way partition the entire vertex set of G. Let H be the minor of G obtained by contracting every T_x.

To prove that $d(H) \geqslant d(d-1)^k$, note first that the T_x are induced subgraphs of G, because $diam(T_x) \leqslant 4k$ and $g(G) > 4k+1$. Similarly, there is most one edge in G between any two trees T_x and T_y: two such edges, together with the paths joining their ends in T_x and T_y, would form a cycle of length at most $8k+2 < g(G)$. So all the edges leaving T_x are preserved in the contraction.

It remains to count how many such edges there are. Note that, for every vertex $u \in T_x^{k-1}$, all its $d_G(u) \geqslant d$ neighbours v also lie in T_x: since $d(v,x) \leqslant k$ and $d(x,y) > 2k$ for every other $y \in X$, we have $d(v,y) > k \geqslant d(v,x)$, so v was added to T_x rather than to T_y when those trees were defined. Therefore T_x^k, and hence also

T_x has at least $d(d-1)^{k-1}$ leaves. But every leaf of T_x sends at least $d-1$ edges away from T_x, so T_x sends at least $d(d-1)^k$ edges to (distinct) other threes T_y. □

Lemma 3.1.8 provides Theorem 3.1.7 with the following corollary.

Theorem 3.1.9. *There exists a function $f : \mathbb{N} \to \mathbb{N}$ such that every graph of minimum degree at least 3 and girth at least $f(r)$ has a K_p minor, for all $p \in \mathbb{N}$.*

Proof. We prove the theorem with $f(p) := 8\log p + 4\log\log p + c$, for some constant $c \in \mathbb{R}$. Let $k = k(p) \in \mathbb{N}$ be minimal with $3 \cdot 2^k \geqslant c' p\sqrt{\log p}$, where $c' \in \mathbb{R}$ is the constant from Theorem 3.1.7. Then, for a suitable constant $c \in \mathbb{R}$, we have $8k+3 \leqslant 8\log p + 4\log p \log p + c$, and the result follows by Lemma 3.1.8 and Theorem 3.1.7. □

3.2 Topological minors

In the previous section, we saw that we need an average degree of 2^{p-2} in order to force a K_p minor. Forcing a topological K_p minor is slightly harder. We shall fix its branch vertices in advance and then construct its subdivided edges inductively, which requires an average degree of $2^{\binom{p}{2}}$ to start with.

Proposition 3.2.1. *Every graph of average degree at least $2^{\binom{p}{2}}$ has a topological K_p minor, for every integer $p \geq 2$.*

Proof. The assertion is clear for $p = 2$, so let us assume that $p \geq 3$. We shall show by induction on $m = p, \ldots, \binom{p}{2}$ that every graph G of average degree $d(G) \geq 2^m$ has a topological minor X with p vertices and m edges.

If $m = p$ then, by Propositions 1.3.1 and 1.3.2, G contains a cycle of length at least $\epsilon(G) + 1 \geq 2^{p-1} + 1 \geq p - 1$ and the assertion follows with $X = C^r$.

Now let $r < m \leqslant \binom{p}{2}$ and assume the assertion holds for smaller m. Let G with $d(G) \geq 2^m$ be given; thus, $\epsilon(G) \geq 2^{m-1}$. Since G has a component C with $\epsilon(C) \geq \epsilon(G)$, we may assume that G is connected. Consider a maximal set $U \subseteq V(G)$ such that U is connected in G and $\epsilon(G \backslash U) \geq 2^{m-1}$. Such a set U exists because G itself has the form $G \backslash U$ with $|U| = 1$. Since G is connected, we have $N(U) \neq \emptyset$.

Let $H := G[N(U)]$. If H has a vertex v of degree $d_H(v) < 2^{m-1}$, we may add it to U and obtain a contradiction to the maximality of U; when we contract the edge v_U in $G \backslash U$, we lose one vertex and $d_H(v) + 1 \leqslant 2^{m-1}$ edges, so ϵ will still be at least 2^{m-1}. Therefore $d(H) \geq \delta(H) \geq 2^{m-1}$. By the induction hypothesis, H contains a topological minor Y with $|V(Y)| = r$ and $|E(Y)| = m - 1$. Let x, y be two branch vertices of this Y that are non-adjacent in Y. Since x and y lie in $N(U)$ and U is connected in G, G contains an (x, y)-path whose inner vertices lie in U. Adding this path to Y, we obtain de desired X. $\square$

In order to prove our next result we shall need a definition and a theorem concerning graph connectivity, whose proof we shall omit for the sake of brevity.

Let G be a graph and let $X \subseteq V(G)$ be a set of vertices. We call X *linked* in G if for any distinct vertices $s_1, \ldots, s_l, t_1, \ldots, t_l$ in X there exist disjoint paths $P_1, \ldots, P_l$ in G such that each P_i links s_i to t_i and has no inner vertices in X. Thus, unlike Menger's theorem, we are not merely asking for disjoint paths between two sets of vertices, but insist that each of these paths link a specific pair of endpoints.

If $|V(G)| \geq 2k$ and every set of at most $2k$ vertices is linked in G, then G is *k-linked.* As is easily checked, this is equivalent to requesting that disjoint paths $P_i = [s_i, \ldots, t_i]$ exist for every choice of exactly $2k$ vertices $s_1, \ldots, s_k, t_1, \ldots, t_k$.

Theorem 3.2.2. *Let G be a graph and let $k \in \mathbb{N}$. If G is $2k$-connected and $\epsilon(G) \geq 8k$, then G is k-linked.*

We are now going to use Theorem 3.2.2 in order to reduce the bound in Proposition 3.2.1 from exponential to quadratic, which is best possible up to a multiplicative constant.

Theorem 3.2.3. *There is a constant $c \in \mathbb{R}$ such that, for every $p \in \mathbb{N}$, every graph G of average degree $d(G) \geqslant cp^2$ contains K_p as a topological minor.*

Proof. We prove the theorem with $c = 10$. Let G with $d(G) \geqslant 10p^2$ be given. By Mader's theorem, for $k = p^2$, G has a subgraph H with $\kappa(H) \geqslant p^2$ and $\epsilon(H) > \epsilon(G) - p^2 \geqslant 4p^2$. For a topological K_p minor in H, pick a set X of r vertices in H as branch vertices, and a set Y of $p(p-1)$ neighbours of X in H, $p-1$ for each vertex in X, as initial subdividing vertices. These are p^2 vertices altogether and they can be chosen distinct since $\delta(H) \geqslant \kappa(H) \geqslant p^2$.

It remains to link up the vertices of Y in pairs, by disjoint paths in $H' := H - X$ corresponding to the edges of K_p. This can be done if Y is linked in H'. We show more generally that H' is $\frac{1}{2}p(p-1)$-linked, by checking that H' satisfies the premise of Theorem 7.2 for $k = \frac{1}{2}p(p-1)$. We have $\kappa(H') \geqslant \kappa(H) - p \geqslant p(p-1) = 2k$. And as H' was obtained from H by deleting at most $r|V(H)|$ edges (as well as some vertices), we also have $\epsilon(H') \geqslant \epsilon(H) - p \geqslant 4p(p-1) = 8k$. $\square$

As with minors, large girths can also be used to force a topological K_p minor. We now need some vertices of degree at least $p-1$ to serve as branch vertices, but if we assume a minimum degree of $p-1$ to secure these, we can even get by with a girth bound that is independent of p, as was shown by Kühn and Osthus [34]:

Theorem 3.2.4. *There exists a constant g such that G contains K_p as a topological minor for every graph G satisfying $\delta(G) \geqslant p-1$ and $g(G) \geqslant g$.*

3.3 Hadwiger's conjecture

As we saw in the previous two sections, an average degree of $cp\sqrt{\log p}$ suffices to force an arbitrary graph to have a K_p minor, and an average degree of cp^2 forces it to contain a topological K_p minor. If we replace "average degree" with "chromatic number" then, with almost the same constants c, the two assertions remain true: this is because every graph with chromatic number k has a subgraph of minimum -and thus average- degree at least $k-1$.

Although both functions above, $cp\sqrt{\log p}$ and cp^2, are best possible (up to the value of c) for the said implications with respect to the average degree, the question arises of whether they are best possible with respect to the chromatic number, or whether some slower growing function would do in that case. The underlying significance of this question has to do with the nature of the invariant χ and the structural effect it may have on a graph.

This question is, unsurprisingly, somewhat related to the four colour theorem, whose proof has been deemed unsatisfactory, requiring as it does the extensive use of a computer. Since this is the case, it may be said that we do not know the *real* reason the four colour theorem is true, that is, the exact reason why planarity implies that four colours suffice. In order to answer this question, several attempts have been made to reduce the hypotheses of the theorem to a minimum core, in the hope of achieving a better understanding of the situation. Although this question has not yet been answered, it has given rise to some interesting problems. Of these, the most famous is Hadwiger's conjecture.

As we have seen before, planar graphs are precisely the graphs that do not contain K_5 or $K_{3,3}$ as a minor, so the four colour theorem states that every graph with no K_5 or $K_{3,3}$ minor is 4-colourable. Searching for the exact reason behind the four colour theorem, it is natural to exclude K_5, since it is not 4-colourable, but the same cannot be said about $K_{3,3}$. The question arises of whether all graphs without a K_5 minor are 4-colourable, and a natural generalisation is an analogous statement about graphs with no K_p minor, for any integer p. Hadwiger conjectured that this question has a positive answer, although the problem remains open.

Conjecture 3.3.1 (Hadwiger 1943)**.** *For every integer $p > 0$, every graph with no K_p minor is $(p-1)$-colourable.*

When Hadwiger introduced his conjecture, he proved it for $p \leqslant 4$. Wagner [61] had already shown in 1937 that the case $p = 5$ is equivalent to the four colour theorem, and so this case was finally proved in 1976 by Appel and Haken [3, 4]. In 1993, Robertson, Thomas and Seymour [47] proved the conjecture for $p = 6$. For $p \geqslant 7$, the conjecture remains open, but it is true for line graphs, a result easily obtained as a consequence of Vizing's theorem. Bollobás, Catlin and Erdős [6] proved that, rephrased as "G contains a $K^{\chi(G)}$", Hadwiger's conjecture is true for almost all graphs.

Hadwiger's conjecture raises the question of what the graphs without a K_p minor look like: any sufficiently detailed structural description of those graphs should enable us to decide whether or not they can be $(p-1)$-colourable.

The cases $p = 1$ and $p = 2$ are trivial. For $p = 3$, the graphs without a K_3 minor are precisely the forests, and those are indeed 2-colourable. For $p = 4$ there is also a simple structural characterization of the graphs without a K_p minor. Before we state it, we shall need a definition and a lemma.

If G is a graph with induced subgraphs G_1, G_2 and S, such that $G = G_1 \cup G_2$ and $S = G_1 \cap G_2$, we say that G arises from G_1 and G_2 by *pasting these graphs together along S.*

Lemma 3.3.2. *Let $\mathcal{X}$ be a set of 3-connected graphs. Let G be a graph with a proper separation $\{V_1, V_2\}$ of order $\kappa(G) \leqslant 2$. If G is edge-maximal without a topological minor in $\mathcal{X}$, then so are $G_1 := G[V_1]$ and $G_2 := G[V_2]$, and $G_1 \cap G_2 = K_2$.*

Proof. Note first that every vertex $v \in S := V_1 \cap V_2$ has a neighbour in every component of $G_i = S$, $i = 1, 2$; otherwise $S \backslash \{v\}$ would separate G, contradicting $|S| = \kappa(G)$. By the maximality of G, every edge e added to G lies in a $TX \subseteq G + e$ with $X \in \mathcal{X}$, For all the choices of e considered below, the 3-connectedness of X will imply that the branch vertices of this TX all lie in the same V_i, say in V_1. (The position of e will always be symmetrical with respect to V_1 and V_2, so this assumption entails no loss of generality.) Then the TX meets V_2 at most in a path P corresponding to an edge of X.

If $S = \emptyset$, we obtain an immediate contradiction by choosing e with one end in V_1 and the other in V_2. If $S = \{v\}$ is a singleton, let e join a neighbour v_1 of v in $V_1 \backslash S$ to a neighbour v_2 of v in $V_2 \backslash S$, as shown in the figure below. Then P contains both v and the edge $e = \{v_1, v_2\}$; replacing its segment vPv_2v_1 with the edge $\{v, v_1\}$ we obtain a TX in $G_1 \subseteq G$, a contradiction.

So $|S| = 2$, say $S = \{x, y\}$. If $\{x, y\} \notin E(G)$, let $e := \{x, y\}$, and in the arising TX replace e by an (x, y)-path through G_2. This yields a TX in G, a contradiction. Hence $\{x, y\} \in E(G)$, and $G[S] = K_2$ as claimed.

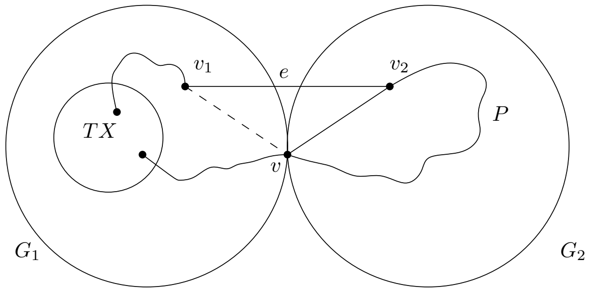

Figure 3.1: If $G + e$ contains a TX, then so does G_1 or G_2

It remains to show that G_1 and G_2 are edge-maximal without a topological minor in $\mathcal{X}$. So let e' be an additional edge for G_1, say. Replacing xPy with the edge $\{x, y\}$ if necessary, we obtain a TX either in $G_1 + e'$ (which shows the edge-maximality of G_1, as desired), or in G_2 (which contradicts $G_2 \subseteq G$). □

Proposition 3.3.3. *Let G be an edge-maximal graph without a K_4 minor. If $|V(G)| \geqslant 3$, then G can be constructed recursively from triangles by pasting along K_2's.*

Proof. Recall first that if K_4 is a minor of a graph G, then it is also a topological minor of G, because $\Delta(K_4) = 3$; the graphs without a K_4 minor, thus, coincide with those without a topological K_4 minor.

We will use induction on $|V(G)|$. Let G be given, edge-maximal without a K_4 minor. If $|V(G)| = 3$, then G itself is a triangle, so let $|V(G)| \geqslant 4$ for the induction step. Then G is not complete; let $S \subseteq V(G)$ be a separator of size $\kappa(G)$, and let C_1, C_2 be distinct components of $G \backslash S$. Since S is a maximal separator, every vertex in S has a neighbour in C_1 and another in C_2. If $|S| \geqslant 3$, this implies that G contains three independent paths P_1, P_2, P_3 between a vertex $v_1 \in C_1$ and a vertex $v_2 \in C_2$. Since $\kappa(G) = |S| \geqslant 3$, the graph $G \backslash \{v_1, v_2\}$ is connected and contains a (shortest) path P between two different P_i. Then $P \cup P_1 \cup P_2 \cup P_3$ is a subdivision of K_4, a contradiction.

Hence $\kappa(G) \leqslant 2$, and the assertion follows from Lemma 3.3.2 and the induction hypothesis. □

One of the interesting consequences of Proposition 3.3.3 is that all the edge-maximal graphs without a K_4 minor have the same number of edges, and thus are all extremal.

Corollary 3.3.4. *Every edge-maximal graph G without a K_4 minor has $2|V(G)| - 3$ edges.*

Corollary 3.3.5. *Hadwiger's conjecture holds for $p = 4$.*

Proof. If G arises from G_1 and G_2 by pasting along a complete graph, then $\chi(G) = \max\{\chi(G_1), \chi(G_2)\}$. Hence, Proposition 3.3.3 implies by induction on $|V(G)|$ that all edge-maximal (and hence all) graphs without a K_4 minor can be 3-coloured. □

It is also possible to prove Corollary 3.3.5 by a direct argument which, though simple, does not give us the structural insight that the proof above does. The proof is by induction on the order of G; If G has fewer than 4 vertices, we have at least one colour per vertex. If G has no cycles then it is a forest and thus 2-colourable; otherwise, let C be a minimum cycle in G. Since a chord would create a smaller cycle, C has no chords. Therefore, C is an induced cycle, so it is possible to 3-colour $G[V(C)]$ by travelling around the cycle alternating between two colours and colouring the last vertex with the third colour, if C has odd length. Fix such a colouring of C. If $G = C$ we are done. Otherwise, $G \backslash C$ is non empty.

Let $S_1, \ldots, S_k$ be the components of $G \backslash C$. Each has order strictly smaller than G, and does not contain K_4 as a minor. For any $1 \leqslant i \leqslant k$, the component S_i has at most two neighbours in C; otherwise, contracting S_i to one vertex and contracting C to remove all vertices of C except three neighbours, forms a K_4.

If S_i has no neighbours in C we get a colouring of S_i from the induction hypothesis and need not modify it. If S_i has only one neighbour in C, or two neighbours that both receive the same colour in the colouring of C, take $G[V(C) \cup V(S_i)]$ and contract C to a single vertex v_C to get a graph that does not contain K_4 as a minor and has smaller order, so is 3-colourable by the induction hypothesis. Take a 3-colouring of this graph and exchange colours so that v_C receives the same colour as the neighbour(s) of S_i did in the colouring of C. The result is a 3-colouring of S_i that is proper on S_i but also proper with respect to the colouring of C that was fixed.

Otherwise, S_i has neighbours x and y in C that received different colours. In this case, take $G[V(C) \cup V(S_i)]$ and contract C until only x and y remain from C. The result is again a graph that does not contain K_4 as a minor and is of smaller order, so is 3-colourable. Take a 3-colouring of this graph and permute colours so that x and y get the same colour they had in the fixed colouring of C. This is possible because x and y were adjacent in the original colouring, so they received different colours, just as they do in the colouring of C. The result is again a 3-colouring of S_i that is proper with respect to the colouring of C.

Be performing the above for each S_i, we obtain a 3-colouring of the entire graph G that is proper, thus completing the proof.

Hadwiger's conjecture for $p = 5$ follows from another structural theorem for graphs without a K_5 minor, just as it follows from Proposition 3.3.3 for $p = 4$.

The proof of the following theorem is similar to that of Proposition 3.3.3, although considerably longer. We therefore state the theorem without proof.

Theorem 3.3.6 (Wagner). *Let G be an edge-maximal graph without a K_5 minor. If $|V(G)| \geqslant 4$, then G can be constructed recursively by pasting along triangles and K_2's from plane triangulations and copies of the Wagner graph W.*

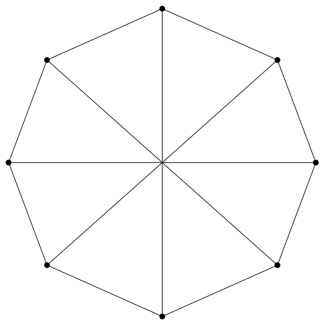

Figure 3.2: The *Wagner graph W*.

Using the fact that every maximally planar graph with n vertices has $3n-6$ edges, one can easily compute which of the graphs constructed as in Theorem 3.3.6 have the most edges. It turns out that these extremal graphs without a K_5 minor have no more edges than the maximal planar graphs:

Corollary 3.3.7. *A graph with n vertices and no K_5 minor has at most $3n-6$ edges.*

Since $\chi(W) = 3$, Theorem 3.3.6 and the four colour theorem imply Hadwiger's conjecture for $p = 5$:

Corollary 3.3.8. *Hadwiger's conjecture holds for $p = 5$.*

Hadwiger's conjecture for $p = 6$ is substantially more difficult than the case $p = 5$ and, again, it relies on the four colour theorem. Suppose that the conjecture does not hold for $p = 6$ and consider a smallest counterexample G. Robertson, Thomas and Seymour [47] showed, without using a computer and without assuming the four colour theorem, that G must be an apex graph, that is, there exists a vertex whose deletion makes G planar. Therefore, since the four colour theorem implies that the planar part of G is 4-colourable, we still have a colour left for the vertex we deleted, so G is not a counterexample after all.

The proof that G is apex is, very roughly, as follows. One can show that G is 6-connected and, in particular, all vertices have degree at least 6; and vertices of degree 6 belong to K_4 subgraphs, and it follows that there are not many of them(in fact at most two), or else we could piece together all these K_4's to make a K_6 minor. On the other hand, a theorem of Mader states that the average degree of G is less than 8, and we cannot make the average degree bigger than 8 even if we cleverly contract edges. This implies that there are edges that are in several triangles or squares. If, say, there is an edge uv in four triangles, then there is a K_4 minor of $G\backslash\{u, v\}$ on the four surviving vertices of the triangles (since G has no K_6 minor), and graphs with this property are well-understood; basically they have to be planar with the four special vertices on the infinite region. So $G\backslash\{u, v\}$ is planar and a little more thought shows that one of $G\backslash u$, $G\backslash v$ is planar. Hence G is apex.

Proving that graphs with no K_7 minor are 6-colourable is thus the first case of Hadwiger's conjecture that is still open. Albar and Gonçalves [1] proved the following result.

Theorem 3.3.9. *Every graph with no K_7 minor is 8-colourable, and every graph with no K_8 minor is 10-colourable.*

As mentioned earlier, the challenge posed by Hadwiger's conjecture is to devise a proof technique that makes better use of the assumption $\chi \geqslant r$ than just using its consequence $\delta \geqslant r - 1$ in a suitable subgraph, which we know cannot force a K_p minor (Theorem 3.1.7). So far, no such technique is known.

If we resign ourselves to using just $\delta \geqslant r - 1$, we can still ask what additional assumptions might help in making this force a K_p minor. Theorem 3.2.4 says that an assumption of large girth has this effect. In fact, a much weaker assumption suffices: for any fixed $s \in \mathbb{N}$ and all large enough d depending only on s, the graphs $G \not\supseteq K_{s,s}$ of average degree at least d can be shown to have K_r minors for r considerably larger than d. For Hadwiger's conjecture, this implies the following.

Theorem 3.3.10 (Kühn and Osthus [35])**.** *For every integer s there exists an integer p_s such that Hadwiger's conjecture holds for all graphs $G \not\supseteq K_{s,s}$ and $p \geqslant p_s$.*

The strengthening of Hadwiger's conjecture that graphs of chromatic number at least p contain K_p as a *topological* minor has become known as *Hajos' conjecture.* It is false in general, but Theorem 3.2.4 implies it for graphs of large girth:

Theorem 3.3.11. *There is a constant g such that for any graph G of girth at least g, if $\chi(G) \geqslant p$, then G contains K_p as a topological minor, for all p.*

Proof. Let g be the constant from Theorem 3.2.4. If $\chi(G) \geqslant p$ then G has a subgraph H of minimum degree $\delta(H) \geqslant p-1$. As $g(H) \geqslant g(G) \geqslant g$, Theorem 3.2.4 implies that H, and thus G, contains a topological K_p minor. $\square$

The following weakening of the assertion of Hadwiger's conjecture is true.

Theorem 3.3.12 (Halin [28])**.** *Every graph of chromatic number $\alpha \geqslant \aleph_0$ contains every K_β with $\beta < \alpha$ as a minor, as well as a topological minor.*

Chapter 4

Hypergraphs

We will now turn our attention to hypergraphs. A *hypergraph* is a generalisation of a graph in which an edge can join any number of vertices. Formally, a hypergraph $\mathcal{G}$ is a pair (V, E) where X is a set of elements and E is a set of non-empty subsets of X, that is, $E \subseteq \mathcal{P}(V)\backslash\{\emptyset\}$. As before, the elements of the set V are called *vertices* of $\mathcal{G}$, whereas the elements of the set E are called *hyperedges* (or, when there is no fear of confusion, simply *edges*) of $\mathcal{G}$.

An *r-uniform hypergraph*, which we sometimes call an *r-graph*, is a hypergraph such that all its hyperedges have cardinality r. The *complete* r-uniform hypergraph $K_n^{(r)}$ is a hypergraph on n vertices where every r-element subset of the vertex set is an edge.

For $X \subseteq V$, the *induced subhypergraph* $\mathcal{G}[X]$ has vertex set X and edge set all edges of $\mathcal{G}$ that are contained in X. We often abbreviate "subhypergraph" to "subgraph". By a k-set, we mean a set with cardinality k.

The extremal number $\mathsf{ex}(n, \mathcal{F})$ is, as before, the maximum number of edges in an $\mathcal{F}$ free r-graph on n vertices.

4.1 Complete r-graphs

Developing some understanding of extremal numbers for general r-graphs $\mathcal{F}$ is a long-standing open problem in extremal graph theory. By contrast with the graph case, there is comparatively little understanding of the hypergraph case. Having solved the problem for $\mathcal{F} = K_n$ for graphs, Turán posed the question of determining $\mathsf{ex}(n, \mathcal{F})$ when $\mathcal{F} = K_n^{(r)}$ is a complete r-graph on n vertices. To date, no case with $t > r > 2$ of this question has been solved, even asymptotically. Our goal is to outline the ideas behind known some bounds, summarise the rest of them and mention a few recent developments.

For the purpose of this section it is convenient to change to the "complementary" notation that was preferred by many early writers on extremal numbers. They define the Turán number $T(n,k,r)$ to be the minimum number of edges in an r-graph $\mathcal{G}$ on n vertices such that any subset of k vertices contains at least one edge of $\mathcal{G}$. Note that $\mathcal{G}$ has this property if and only if the "complementary" r-graph of r-sets that are non edges of $\mathcal{G}$ is $K_k^{(r)}$-free. Thus,

$$T(n,k,r) + \mathsf{ex}(n, K_k^{(r)}) = \binom{n}{r}.$$

They also define the density $t(k,r) = \lim_{n\to\infty} \binom{n}{r}^{-1} T(n,k,r)$. Thus,

$$t(k,r) + \pi(K_k^{(r)}) = 1.$$

We start with the lower bound on $t(k,r)$, which is equivalent to an upper bound on $\pi\big(K_k^{(r)}\big)$. A trivial averaging argument gives $t(k,r) \geqslant \binom{k}{r}^{-1}$. In general, the best known bound is

$$t(k,r) \geqslant \binom{k-1}{r-1}^{-1},$$

due to de Caen [13]. This follows from his exact bound on

$$T(n,k,r) \geqslant \frac{n-k+1}{n-r+1}\binom{k-1}{r-1}^{-1}\binom{n}{r}.$$

This, in turn, is deduced from a hypergraph generalisation of a theorem of Moon and Moser [45] that relates the number of cliques of various sizes in a graph. Suppose that $\mathcal{G}$ is an r-graph on n vertices and let N_k be the number of copies of $K_k^{(r)}$ in $\mathcal{G}$. Then the inequality is

$$N_{k+1} \geqslant \frac{k^2 N_k}{(k-r+1)(k+1)}\left(\frac{N_k}{N_{k-1}} - \frac{(r-1)(n-k)+k}{k^2}\right), \qquad (1)$$

provided that $N_{k-1} \neq 0$. Given this inequality, the bound on $T(n,k,r)$ follows from some involved calculations; the main step is to show by induction on k that

$$N_k \geqslant N_{k-1}\frac{r^2\binom{k}{r}}{k^2\binom{n}{r-1}}(|E(\mathcal{G})| - F(n,k,r)),$$

where

$$F(n,k,r) = \left(r^{-1}(n-r+1) - \binom{k-1}{r-1}^{-1}(n-k+1)\right)\binom{n}{r-1}.$$

Inequality (1) is proved by the following double counting argument. Let P be the number of pairs (S,T) where S and T are each sets of k vertices, such that S spans a $K_k^{(r)}$, T does not span one, and $|S \cap T| = k-1$. For an upper bound on P, enumerate the N_{k-1} copies of $K_{k-1}^{(r)}$ and let a_i be the number of $K_k^{(r)}$'s containing the i-th copy. Since, $\sum_{i=1}^{N_{k-1}} a_i = kN_k$, we have

$$P = \sum_{i=1}^{N_{k-1}} a_i(n-k+1-a_i) \leqslant (n-k+1)kN_k - N_{k-1}^{-1}k^2N_k^2.$$

For a lower bound, enumerate the copies of $K_k^{(r)}$ as $B_1, \dots, B_{N_k}$ and let b_i be the number of $K_{k+1}^{(r)}$'s containing the i-th copy. For each B_j, there are $n-k-b_j$ ways to choose an $x \notin B_j$ such that $B_j \cup \{x\}$ does not span a $K_{k+1}^{(r)}$. Given such an x, there is some $C \subseteq B_j$ of size $k-1$ such that $C \cup \{x\}$ is not an edge. Then, for each $y \in B_j \backslash C$, the pair $(B_j, B_j \cup x \backslash y)$ is counted by P. This gives

$$P \geqslant \sum_{j=1}^{N_k}(n-k-b_j)(k-r-1) = (k-r-1)((n-k)N_k - (k+1)N_{k+1}).$$

Combining with the lower bound and rearranging provides us with the required inequality.

We consider next the upper bound on $t(k,r)$, which is equivalent to a lower bound on $\pi(K_k^{(r)})$. The best general construction is due to Sidorenko [52] and implies the bound

$$t(k,r) \leqslant \left(\frac{r-1}{k-1}\right)^{r-1}.$$

For comparison with the lower bound, notice that

$$\left(\frac{r-1}{k-1}\right)^{r-1}\binom{k-1}{r-1} = \prod_{i=1} r-1\frac{k-1}{k-1}\frac{r-1}{r-i}.$$

If k is large enough, compared to r, then the ratio of the bounds is approximately $(r-1)^{r-1}((r-1)!)^{-1}$, which is exponential in r but independent of k. In order to explain the construction, we shall rephrase it using the following, simple to follow, fact.

Suppose there is a lorry driver who needs to follow a certain closed route. There are several petrol stations along the route, and the total amount of fuel in these stations is sufficient for the route. Then there exists some starting point from which the route can be completed. Indeed, imagine that the driver starts with enough fuel to drive around the route and consider the journey starting from an arbitrary point, in which she still picks up all the fuel at any station, even though she doesn't need it. Then the point at which the fuel reserves are lowest during this route can be used as a starting point for another route which satisfies the requirements.

The construction is to divide n vertices into $k-1$ roughly equal parts $A_1, \dots, A_{k-1}$, and say that a set B of size r is an edge of $\mathcal{G}$ if there is some j such that

$$\sum_{i=1}^{s} |B \cap A_{j+1}| \geqslant s+1$$

for each $1 \leqslant s \leqslant r-1$ (taking addition modulo $k-1$ in the subscript). To interpret this in the lorry driver framework, consider any set K of size k, imagine that each element of K represents a unit of fuel, and that it takes $\frac{k}{k-1}$ units of fuel to drive from A_i to A_{i+1}. Then K contains enough fuel for a complete circuit, so the lorry driver puzzle tells us that there is some starting point from which a complete circuit is possible. Let B be the set of the first r elements of K that are encountered on this circuit (breaking ties arbitrarily). Since $r \geqslant (r-1)\frac{k}{k-1}$, the lorry can advance distance $r-1$ using just the fuel from B. This implies that B is an edge, as $\lceil s\frac{k}{k-1} \rceil = s+1$ for $1 \leqslant s \leqslant r-1$. Thus, any set of size k contains an edge, as required.

It is not so obvious how to estimate the number of edges in the construction without tedious calculations, so we will give a simple combinatorial argument here. It is convenient to count edges together with an order of the vertices in each edge, thus counting each edge $r!$ times. We can form an ordered edge $B = x_1, \dots, x_r$ using the following three steps:

1. choose the starting index j;

2. assign each x_l to one of the parts A_{j+1}, for $1 \leqslant i \leqslant r-1$;

3. choose a vertex for each x_l within its assigned part.

Clearly there are $k-1$ choices in step (1) and $\left(\frac{n}{k-1}\right)^r + O(n^{r-1})$ choices in step (3). In step (2) there are $(r-1)^r$ ways to assign the parts if we ignore the required inequalities on the intersection sizes (that is, that there should be enough fuel for the lorry). We claim that, given any assignment, there is exactly one cyclic permutation that satisfies the required inequalities. More precisely, if we assign b_i of the x_l's to A_{j+1} for $1 \leqslant i \leqslant r-1$ where, as before, each of the x_l's is a unit of fuel, but now it takes one unit of fuel to advance from A_i to A_{i+1}, and the lorry is required to always have a spare unit of fuel. A valid starting point for the lorry is equivalent to a shifted sequence satisfying the required inequalities. As in the solution to the original puzzle, we imagine that the driver starts with enough fuel to drive around the route and consider the journey starting from any arbitrary point. Then the point at which the fuel reserves are lowest during this route is a starting point for a route where there is always one spare unit of fuel. Furthermore, this is the unique point at which the fuel reserves are lowest, and so it gives the unique cyclic permutation satisfying the required inequalities. We deduce that there are $(r-1)^{r-1}$ valid assignments in step (2). Putting everything together, the number of edges is

$$(r!)^{-1}(k-1)(r-1)^{r-1}\left(1+O\Big(\frac{1}{n}\Big)\right)\left(\frac{n}{k-1}\right)^{r} \sim \left(\frac{r-1}{k-1}\right)^{r-1}\binom{n}{r},$$

as required.

Having discussed the general case, we will now summarise some better bounds that have been found in specific cases. One natural case to focus on is $t(r+1,r) = 1-\pi\big(K_{r+1}^{(r)}\big)$. For large r, a construction of Sidorenko [53] gives the best known upper bound, which is

$$t(r+1,r) \leqslant (1+o(1))\frac{\log r}{2r}.$$

Other known bounds are effective for small r; these are

$$t(r+1,r) \leqslant \frac{1+2\ln r}{r}$$

by Kim and Roush [31], and

$$t(2s+1,2s) \leqslant \frac{1}{4}+2^{-2s}$$

by de Caen, Kreher and Wiseman [14].

On the other hand, the known lower bounds are very close to the bound $t(r_1,r) \geqslant \frac{1}{r}$ discussed above in the general case. Improvements to the second order term were given by Chung and Lu [11], who showed that

$$t(r+1,r) \geqslant \frac{1}{r}+\frac{1}{r(r+3)}+O(r^{-3})$$

when r is odd, and by Lu and Zhao [39], who obtained some improvements when r is even, the best of which is

$$t(r+1,r) \geqslant \frac{1}{r}+\frac{1}{2r^3}+O(r^{-4})$$

when r is of the form $6k+4$. Thus, the known upper and lower bounds are separated by a factor of $\big(\frac{1}{2}+o(1)\big)\log r$. As a first step towards closing this gap, de Caen [15] conjectured that $r\cdot t(r+1,r)\to\infty$ as $r\to\infty$.

For $K_4^{(3)}$, we only know that

$$\frac{5}{9} \leqslant \mathsf{ex}(n,K_4^{(3)}) \leqslant 0.561666.$$

The lower bound is not particularly hard to come by; take three vertex sets V_1, V_2 and V_3, each of size $\frac{n}{3}$. We let an edge uvw be in $\mathcal{G}$ if $uv \in V_i$ and $w \in V_{i+1}$, for

$i = 1, 2$, or if $u \in V_1$, $v \in V_2$ and $w \in V_3$. It is straightforward to check that this contains no $K_4^{(3)}$ and that its Turán density is $\frac{5}{9}$. The upper bound, on the other hand, is much more difficult to obtain, using a combinatorial technique known as *flag algebras*.

For $K_5^{(4)}$ the following construction is due to Giraud [24]. Suppose M is an $m \times m$ matrix with entries equal to 0 or 1. We define a 4-graph $\mathcal{G}$ on $n = 2m$ vertices corresponding to the rows and columns of M. Any 4-set of rows or 4-set of columns is an edge. Also, any 4-set of 2 rows and 2 columns inducing a 2×2 submatrix with even sum is an edge. We claim that any 5-set of vertices of $\mathcal{G}$ contains an edge. This is clear if we have at least 4 rows or at least 4 columns, so suppose without loss of generality that we have 3 rows and 2 columns. Then, in the induced 3×2 submatrix, we can choose 2 rows whose sums have the same parity, that is, a 2×2 submatrix with even sum, which is an edge. To count edges in $\mathcal{G}$, note first that we have $2\binom{m}{4}$ from 4-sets of rows and 4-sets of columns. Also, for any pair i, j of columns, we can divide the rows into two classes O_{ij} and E_{ij} according to whether the entries in columns i and j have odd or even sum. Then, the number of 2×2 submatrices using columns i and j with even sums is

$$\binom{|O_{ij}|}{2} + \binom{|E_{ij}|}{2} \geqslant \binom{\frac{m}{2}}{2}.$$

Furthermore, for some values of m, there is a construction that achieves equality for every pair i, j: take a Hadamard matrix, that is, a matrix with entries equal to 1 or -1, in which every pair of columns is orthogonal, then replace the -1 entries with 0.

This shows that

$$t(5,4) \leqslant \lim_{m\to\infty} \binom{2m}{4}^{-1} \left(2\binom{m}{4} + 2\binom{\frac{m}{2}}{2}\binom{m}{2}\right) = \frac{5}{16};$$

equivalently

$$\pi(K_5^{(4)}) \geqslant \frac{11}{16} = 0.6875.$$

Sidorenko [53] conjectured that equality holds.

Markström [43] gave an upper bound

$$\pi(K_5^{(4)}) \leqslant \frac{1753}{2380} = 0.73655\ldots$$

This was achieved by extensive computer search to find all extremal 4-graphs for $n \leqslant 16$.

4.2 Hamiltonian cycles

Let $\mathcal{H}$ be an r-graph. An *l-tight Hamiltonian cycle* in $\mathcal{H}$, where $0 \leqslant l \leqslant r-1$ and $(k-l) \mid |V(\mathcal{H})|$, is a spanning subgraph whose vertices can be cyclically order in such a way that the edges are segments of that ordering and every two consecutive edges intersect in exactly l vertices. More formally, it is a graph $([n], E)$ with

$$E = \left\{ \{i(r-l)+1, i(r-l)+2, \ldots, i(r-l)+r\} : 0 \leqslant i < \frac{n}{r-l} \right\},$$

where addition is made modulo n. We denote an l-tight Hamiltonian cycle in a r-graph $\mathcal{H}$ on n vertices by $C_n^{(r,l)}$, and call it *tight* if it is $(r-1)$-tight.

A natural question that arises in the study of Hamiltonian cycles is the estimation of their extremal number in r-graphs. Katona and Kierstead [30] were the first to study sufficient conditions for the appearance of a $C_n^{(r,r-1)}$ in r-graphs. They showed that for all integers r and n with $r \geqslant 2$ and $r-1 \leqslant n$, the inequality

$$ex\left(n, C_n^{(r,r-1)}\right) \geqslant \binom{n-1}{r} + \binom{n-2}{r-2}$$

holds. In the same paper, they proved that the bound is not tight for $r = 3$ by showing that for all integers n and q with $q \geqslant 2$ and $n = 3q+1$,

$$ex\left(n, C_n^{(3,2)}\right) \geqslant \binom{n-1}{3} + n - 1.$$

A few years later, Tuza [60] gave a construction for general r and tight Hamiltonian cycles, improving the lower bound to

$$ex\left(n, C_n^{(r,r-1)}\right) \geqslant \binom{n-1}{r} + \binom{n-1}{r-2},$$

if a Steiner system $S(r-2, 2r-3, n-1)$ exists, where a Steiner system $S(t,b,v)$ is a b-graph on v vertices such that every t-element vertex subset is contained in precisely one edge.

An interesting approach to forbid Hamiltonian cycles in hypergraphs is to prohibit certain substructures in the *link* of a fixed vertex. For a vertex v in an r-graph $\mathcal{H} = (V, E)$, we define the *link of v in $\mathcal{H}$* to be the $(r-1)$-graph $\mathcal{H}(v) = (V \setminus \{v\}, E_v)$, where $\{x_1, \ldots, x_{r-1}\} \in E_v$ if and only if $\{v, x_1, \ldots, x_{r-1}\} \in E$. The structure of interest in this case is a generalisation of a path for hypergraphs.

An *l-tight r-uniform t-path*, denoted by $P_t^{(r,l)}$, is an r-graph on t vertices, where $(r-l) \mid (t-l)$, such that there exists an ordering of the vertices, say $\{x_1, \ldots, x_t\}$, so that the edges are segments of that ordering and every two consecutive edges intersect in exactly l vertices. Observe that a $P_t^{(r,l)}$ has $\frac{t-l}{r-l}$ edges.

For arbitrary r and l, Glebov, Person and Weps [25] gave the exact extremal number and the extremal graphs of l-tight Hamiltonian cycles, which rely on the extremal number and graphs of $P(r,l) = P^{(r-1,l-1)}_{\lfloor \frac{r}{r-l} \rfloor (r-l)+l-1}$. In particular, the proved the following result.

Theorem 4.2.1. *For any $r \geqslant 2$ and $l \in \{0,\dots,r-1\}$, there exists an n_0 such that for any $n \geqslant n_0$ and $(r-l) \mid n$,*

$$ex\left(n, C_n^{(r,l)}\right) = \binom{n-1}{r} + \mathsf{ex}(n-1, P(r,l))$$

holds. Furthermore, any extremal graph on n vertices contains an $(n-1)$-clique and a vertex whose link is $P(r,l)$-free.

For $r = 3$ and $l = 1$, Theorem 4.2.1 states that $\binom{n-1}{3} + 1$ hyperedges ensure the existence of a 1-tight Hamiltonian cycle $C_n^{(3,1)}$ for n large enough.

In the same paper, Glebov *et al.* proved an even stronger statement, namely, that with one more edge we find a Hamiltonian cycle that is l-tight in the neighbourhood of one vertex and is $(r-1)$-tight on the rest.

Using a result by Győri, Katona and Lemons [27] stating that

$$(1+o(1))\binom{n-1}{r-2} \leqslant ex\left(n-1, P_{2r-2}^{(r-1,r-2)}\right) \leqslant (r-1)\binom{n-1}{r-2},$$

they obtained lower and upper bounds for $l = r-1$, showing that

$$\binom{n-1}{r} + (1+o(1))\binom{n-1}{r-2} \leqslant ex\left(n, C_n^{(r,l)}\right) \leqslant \binom{n-1}{r} + (r-l)\binom{n-1}{r-2}.$$

The upper bound actually holds for $l \neq r-1$ as well.

Aside from the above result, which depends only on the extremal number of a certain path, there are also results which rely solely on the minimum vertex degree of a hypergraph, much like Dirac's theorem does for graphs.

The *degree* of $\{x_1,\dots,x_i\}$, $1 \leqslant i \leqslant r-1$, in an r-graph $\mathcal{H}$ is the number of edges the set is contained in, and is senoted by $\deg(x_1,\dots,x_i)$. Let

$$\delta_d(\mathcal{H}) = \min\{\deg(x_1,\dots,x_d)\colon \{x_1,\dots,x_d\} \subseteq V(\mathcal{H})\},$$

for $0 \leqslant d \leqslant r-1$. The number $\delta_1(\mathcal{H})$ is the *minimum vertex degree* of $\mathcal{H}$. Note that $\delta_0(\mathcal{H}) = |E(\mathcal{H})|$.

For every d, r, l and n with $0 \leqslant d \leqslant r-1$ and $(r-l) \mid n$, we define the number $h_d^l(r,n)$ to be the smallest integer h such that every r-graph $\mathcal{H}$ on n vertices satisfying $\delta_d(\mathcal{H}) \geqslant h$ contains an l-tight Hamiltonian cycle. Note that

$$h_0^l(r,n) = ex\left(n, C_n^{(r,l)}\right) + 1.$$

Katona and Kierstead [30] showed that

$$h_{r-1}^{r-1}(r,n) \geqslant \left\lfloor \frac{n-r+3}{2} \right\rfloor$$

by giving an extremal construction. Rödl, Ruciński and Szemerédi [49] proved that this bound is tight for $r = 3$. For $r \geqslant 4$, the same authors [50] showed that

$$h_{r-1}^{r-1}(r,n) \sim \frac{1}{2}n.$$

Generalising the result, Markström and Ruciński [44] proved that

$$h_{r-1}^{l}(r,n) \sim \frac{1}{2}n$$

if $(r-l) \mid r, n$. Kühn, Mycroft and Osthus [36] proved that

$$h_{r-1}^{l}(r,n) \sim \frac{n}{\lceil \frac{r}{r-l} \rceil (r-l)}$$

if $(r-l) \nmid r$ and $(r-l) \mid n$. Rödl and Ruciński [51] gave the bounds

$$\left(\frac{5}{9} + o(1)\right)\binom{n-1}{2} \leqslant h_1^2(3,n) \leqslant \left(\frac{11}{12} + o(1)\right)\binom{n-1}{2}.$$

Finally, Glebov, Person and Weps [25] proved the following result.

Theorem 4.2.2. *For any $r \in \mathbb{N}$ there exists an n_0 such that every r-graph $\mathcal{H}$ on $n \geqslant n_0$ vertices with*

$$\delta_1(\mathcal{H}) \geqslant \left(1 - \frac{1}{22(1280r^3)}\right)\binom{n-1}{r-1}$$

contains a tight Hamiltonian cycle.

Note that Theorem 4.2.2 implies that

$$h_d^l(r,n) \leqslant \left(1 - \frac{1}{22(1280r^3)}\right)\binom{n-d}{r-d}$$

for all $l \in \{0, \ldots, r-1\}$ and all $1 \leqslant d \leqslant r-1$. This, in turn, implies the existence of a constant $c < 1$ such that, for all l and d, the inequality

$$h_d^l(r,n) \leqslant c\binom{n-d}{r-d}$$

holds, although this constant is probably far from the best possible.

Bibliography

[1] B. Albar and D. Gonçalves. *On triangles in K_r-minor free graphs.* Journal of Graph Theory 88.1 (2018): 154-173.

[2] N. Alon, M. Krivelevich and B. Sudakov. *Turán numbers of bipartite graphs and related Ramsey-type questions.* Combinatorics, Probability and Computing 12.5+ 6 (2003): 477-494.

[3] K. Appel and W. Haken. *Every planar map is four colorable. Part I: Discharging.* Illinois Journal of Mathematics 21.3 (1977): 429-490.

[4] K. Appel, W. Haken and J. Koch. *Every planar map is four colorable. Part II: Reducibility.* Illinois journal of Mathematics 21.3 (1977): 491-567.

[5] P. N. Balister, et al. *Connected graphs without long paths.* Discrete Mathematics 308.19 (2008): 4487-4494.

[6] B. Bollobás, P. A. Catlin and P. Erdös. *Hadwiger's conjecture is true for almost every graph.* Eur. J. Comb. 1.3 (1980): 195-199.

[7] J. A. Bondy and M. Simonovits. *Cycles of even length in graphs.* Journal of Combinatorial Theory, Series B 16.2 (1974): 97-105.

[8] W. G. Brown. *On graphs that do not contain a Thomsen graph.* Canad. Math. Bull 9.2 (1966): 1-2.

[9] S. A. Burr and P. Erdős. *On the magnitude of generalized Ramsey numbers for graphs.* Infinite and Finite Sets I, Colloquia Mathematica Soc. (1975): 214–240.

[10] M. Cera, et al. *Minor Extremal Problems Using Turán Graphs.* International Mathematical Forum. Vol. 2. No. 63. 2007.

[11] F. Chung, and L. Lu. *An upper bound for the Turán number $t_3(n, 4)$.* Journal of Combinatorial Theory, Series A 87.2 (1999): 381-389.

[12] V. Chvátal and P. Erdős. *A note on Hamiltonian circuits.* Discrete Mathematics 2.2 (1972): 111-113.

[13] D. de Caen. *Extension of a theorem of Moon and Moser on complete hypergraphs.* Ars Combin 16 (1983): 5-10.

[14] D. de Caen, D. L. Kreher and J. Wiseman. *On constructive upper bounds for the Turán numbers* $T(n, 2r+1, 2r)$. Congressus Numerantum 65 (1988): 277-280.

[15] D. de Caen. *The current status of Turán's problem on hypergraphs.* Extremal problems for finite sets 3 (1991): 187-197.

[16] W. F. De la Vega. *On the maximum density of graphs which have no subcontraction to* K_3. Discrete Mathematics 46.1 (1983): 109-110.

[17] G. A. Dirac. *Some theorems on abstract graphs.* Proceedings of the London Mathematical Society 3.1 (1952): 69-81.

[18] G. A. Dirac. *Homomorphism theorems for graphs.* Mathematische Annalen 153.1 (1964): 69-80.

[19] P. Erdős and T. Gallai. *On maximal paths and circuits of graphs.* Acta Mathematica Academiae Scientiarum Hungarica 10.3-4 (1959): 337-356.

[20] P. Erdős, A. Rényi and V. T. Sós. *On a problem in the theory of graphs.* Publ. Math. Inst. Hungar. Acad. Sci 7 (1962): 215-235.

[21] P. Erdős and A. H. Stone. *On the structure of linear graphs.* Bull. Amer. Math. Soc 52.1087-1091 (1946): 1.

[22] P. Erdős and M. Simonovits. *A limit theorem in graph theory.* Studia Sci. Math. Hung. 1965.

[23] R. J. Faudree and R. H. Schelp. *Path Ramsey numbers in multicolorings.* Journal of Combinatorial Theory, Series B 19.2 (1975): 150-160.

[24] G. R. Giraud. *Remarques sur deux problemes extrémaux.* Discrete mathematics 84.3 (1990): 319-321.

[25] R. Glebov, Y. Person and W. Weps. *On extremal hypergraphs for Hamiltonian cycles.* European Journal of Combinatorics 33.4 (2012): 544-555.

[26] W. T. Gowers. *Lower bounds of tower type for Szemerédi's uniformity lemma.* Geometric & Functional Analysis GAFA 7.2 (1997): 322-337.

[27] E. Győri, G. Y. Katona and N. Lemons. *Hypergraph extensions of the Erdős-Gallai theorem.* Electronic Notes in Discrete Mathematics 36 (2010): 655-662.

[28] R. Halin. *Unterteilungen vollständiger Graphen in Graphen mit unendlicher chromatischer Zahl.* Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg. Vol. 31. No. 3-4. Springer-Verlag, 1967.

[29] L. K. Jørgensen. *Contractions to K_8.* Journal of Graph Theory 18.5 (1994): 431-448.

[30] G. Y. Katona and H. A. Kierstead. *Hamiltonian chains in hypergraphs.* Journal of Graph Theory 30.3 (1999): 205-212.

[31] K. H. Kim, and F. W. Roush. *On a problem of Turán.* Studies in pure mathematics. Birkhäuser, Basel, 1983. 423-425.

[32] A. V. Kostochka. *Lower bound of the Hadwiger number of graphs by their average degree.* Combinatorica 4.4 (1984): 307-316.

[33] T. Kovári, V. Sós and P. Turán. *On a problem of K. Zarankiewicz.* Colloquium Mathematicum. Vol. 1. No. 3. 1954.

[34] D. Kühn and D. Osthus. *Topological minors in graphs of large girth.* Journal of Combinatorial Theory, Series B 86.2 (2002): 364-380.

[35] D. Kühn and D. Osthus. *Complete minors in $K_{s,s}$-free graphs.* Combinatorica 25.1 (2004): 49-64.

[36] D. Kühn, R. Mycroft and D. Osthus. *Hamilton l-cycles in uniform hypergraphs.* Journal of Combinatorial Theory, Series A 117.7 (2010): 910-927.

[37] C. Kuratowski. *Sur le probleme des courbes gauches en topologie.* Fundamenta mathematicae 15.1 (1930): 271-283.

[38] F. Lazebnik, V. A. Ustimenko and A. J. Woldar. *Polarities and $2k$-cycle-free graphs.* Discrete Mathematics 197.198 (1999): 503-513.

[39] L. Lu and Y. Zhao. *An Exact Result for Hypergraphs and Upper Bounds for the Turán Density of K^r_{r+1}.* SIAM Journal on Discrete Mathematics 23.3 (2009): 1324-1334.

[40] W. Mader. *Homomorphieeigenschaften und mittlere Kantendichte von Graphen.* Mathematische Annalen 174.4 (1967): 265-268.

[41] W. Mader. *Homomorphiesätze für graphen.* Mathematische Annalen 178.2 (1968): 154-168.

[42] W. Mantel. *Problem 28.* Wiskundige Opgaven 10.60-61 (1907): 320.

[43] K. Markström. *Extremal hypergraphs and bounds for the Turán density of the 4-uniform K_5.* Discrete Mathematics 309.16 (2009): 5231-5234.

[44] K. Markström and A. Ruciński. *Perfect matchings (and Hamilton cycles) in hypergraphs with large degrees.* European Journal of Combinatorics 32.5 (2011): 677-687.

[45] J. W. Moon and L. Moser. *On a problem of Turán.* Magyar Tud. Akad. Mat. Kutató Int. Közl 7 (1962): 283-286.

[46] O. Ore. *A note on hamiltonian circuits.* American Mathematical Monthly 67 (1960): 55.

[47] N. Robertson, P. Seymour and R. Thomas. *Hadwiger's conjecture for K_6-free graphs.* Combinatorica 13.3 (1993): 279-361.

[48] K. F. Roth. *On certain sets of integers.* Journal of the London Mathematical Society 1.1 (1953): 104-109.

[49] V. Rödl, A. Ruciński and E. Szemerédi. *Dirac-type conditions for hamiltonian paths and cycles in 3-uniform hypergraphs.* Advances in Mathematics 227.3 (2011): 1225-1299.

[50] V. Rödl, A. Ruciński and E. Szemerédi. *An approximate Dirac-type theorem for k-uniform hypergraphs.* Combinatorica 28.2 (2008): 229-260.

[51] V. Rödl and A. Ruciński. *Dirac-type questions for hypergraphs—a survey (or more problems for Endre to solve).* An irregular mind. Springer, Berlin, Heidelberg, (2010). 561-590.

[52] A. Sidorenko. *What we know and what we do not know about Turán numbers.* Graphs and Combinatorics 11.2 (1995): 179-199.

[53] A. Sidorenko. *Upper bounds for Turán numbers.* Journal of Combinatorial Theory Series A 77 (1997): 134-147.

[54] M. Simonovits. *Extremal graph problems with symmetrical extremal graphs. Additional chromatic conditions.* Discrete Mathematics 7.3-4 (1974): 349-376.

[55] E. Szemerédi. *On sets of integers containing no k elements in arithmetic progression.* Acta Arith 27 (1975): 299-345.

[56] E. Szemerédi. *Regular partitions of graphs.* Problemes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), CNRS, Paris, 1978. MR 81i 5095.3.2: 3-4.

[57] A. Thomason. *An extremal function for contractions of graphs.* Mathematical Proceedings of the Cambridge Philosophical Society. Vol. 95. No. 2. Cambridge University Press, 1984.

[58] A. Thomason. *The extremal function for complete minors.* Journal of Combinatorial Theory, Series B 81.2 (2001): 318-338.

[59] P. Turán, Paul. *On an external problem in graph theory.* Mat. Fiz. Lapok 48 (1941): 436-452.

[60] Z. Tuza. *Steiner systems and large non-hamiltonian hypergraphs.* Le Matematiche 61.1 (2006): 179-183.

[61] K. Wagner. *Über eine Eigenschaft der ebenen Komplexe.* Mathematische Annalen 114.1 (1937): 570-590.

[62] K. Zarankiewicz. *Problem p 101.* Colloq. Math. Vol. 2. No. 301. 1951.