PROBLEMS IN EXTREMAL GRAPH THEORY

BY

ANDRÉ KÜNDGEN

Vordiplom, Universität Trier, 1991
M.S., University of South Carolina, Columbia, 1993

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1999

Urbana, Illinois

Abstract

We consider several extremal problems in graph theory and computational geometry.

A *cut* in a graph G is the set of all edges between some set of vertices S and its complement $\overline{S} = V(G) - S$. A *cut-cover* of G is a collection of cuts whose union is $E(G)$ and the *total size* of a cut-cover is the sum of the number of edges of the cuts in the cover. The cut-cover size of a graph G, denoted by $\mathrm{cs}(G)$, is the minimum total size of a cut-cover of G.

We give general bounds on $\mathrm{cs}(G)$, find sharp bounds for classes of graphs such as 4-colorable graphs and random graphs. We give sharp bounds for the sum of the cut-cover sizes of a graph and its complement. We also determine the cut-cover size of the complete graph on n vertices and establish a close connection between this problem and an isoperimetric problem due to Ahlswede and Katona.

Consider an *art gallery* formed by a polygon on n vertices with m pairs of vertices joined by interior diagonals, the *interior walls*. Each interior wall has an arbitrarily placed, arbitrarily small doorway. We determine the minimum number of guards that suffice to guard all art galleries with n vertices and m interior walls. We also determine the minimum number of guards that suffice for galleries with convex rooms of specified minimum size. The proofs lead to linear time guard placement algorithms in most cases.

What is the maximum number of edges in a multigraph on n vertices if every k-set spans at most r edges? We asymptotically determine this maximum for almost all k and r as n tends to infinity, thus giving a generalization of Turán's theorem. We find exact answers in many cases, even when edges of negative multiplicity are allowed.

Acknowledgement

I want to thank Jeff Erickson and Tibor Szabó for serving on my thesis committee and giving me feedback on a number of issues, including the thesis.

I am grateful to Paul Weichsel for having been on “all my committees” and for the great job he is doing as the Associate Chair, which made my teaching experiences here smooth and enjoyable.

I am also grateful for the many stimulating discussions that I had with the students here in Illinois, especially Maria Axenovic, Ya-Chen Chen, Tao Jiang and Dhruv Mubayi.

This thesis was greatly influenced by the countless hours that I spent in Doug West’s courses and seminars and in informal discussions in his office or on the road. I benefited from him as a source for problems, his untiring efforts to improve my exposition, and his helpfulness in many mathematical discussions and in the job search. Thanks Doug!

I also want to thank my advisor Zoltán Füredi for the many deep insights he has been willing to share with me in his courses and in personal communications in his office. I learned a lot from working with him on joint problems and I am grateful that I was able to visit him in Europe.

I also want to thank the many other people that contributed to the thesis in some fashion, and helped making my years as a graduate student at Illinois successful and enjoyable, especially: Kequan Ding, Harold Diamond, Luis Goddyn, Jerry Griggs, Alexander Kostochka, Jenő Lehel, Tony Peressini, Bruce Reznick, David Sumner and all staff members.

Most of all I want to thank my parents for encouraging me in my life in many ways and my wife Radhika for being supportive and caring during the sometimes trying days of graduate school.

Table of Contents

1. Introduction **1**
1.1. Covering a graph efficiently **1**
1.2. Guarding art galleries **2**
1.3. A generalization of Turán's theorem **2**
1.4. Basic definitions **3**
2. Covering a graph with cuts of minimum total size **7**
2.1. Covering the edges of a graph **7**
2.2. Covering a graph efficiently **7**
2.3. The cut-cover problem: Definitions and main results **8**
2.4. Upper and lower bounds **11**
2.5. Extremal cut-cover problems **12**
2.6. Algorithmic aspects **16**
2.7. Cut-covers in random graphs **18**
2.8. Complementary graphs **22**
2.9. More bounds and exact values **23**
2.10. Cut-covers and L_∞-representations **27**
2.11. Open questions **29**
3. Guarding art galleries **31**
3.1. The basic art gallery problem **31**
3.2. Art galleries with interior walls **31**
3.3. Art galleries with convex rooms **36**
3.4. Complexity **37**
4. Turán problems for weighted Graphs **39**
4.1. Preliminaries, density results in graphs **39**
4.2. The density problem **41**
4.3. Definitions and main results **43**
4.4. Proof of the general upper bound **46**
4.5. Construction of dense graphs for integer weights **48**
4.6. Some exact values of $\mathrm{ex}_{\mathbb{Z}}(n,k,r)$ **52**
4.7. Proof of the lower bound **52**
4.8. More on the linear term **56**
4.9. The asymptotic density for nonnegative integer weights **60**
References **62**
Vita **71**

1. Introduction

Extremal graph theory is the branch of combinatorics in which we are interested in determining the extreme value of some graph parameter over a given class of graphs.

Extremal graph theory, as a graph theoretical discipline, was established in a labor camp in 1940. Under difficult circumstances Paul Turán proved his famous Theorem [113,114]: among all graphs on n vertices, the *Turán graph* $T_{n,k-1}$, i.e. the complete $(k-1)$-partite graph with part sizes as equal as possible, has the maximum number of edges without containing a complete graph on k vertices. Since then extremal graph theory has experienced tremendous growth; the book by Bollobás [14] collected many early results.

The results in this thesis are all, in some sense, in the field of extremal graph theory, with the results in Chapter 4 being a generalization of Turán's Theorem. In the following three sections we briefly summarize the main results. To assist readers unfamiliar with the basic definitions in graph theory all relevant terms have been collected in the last section of this chapter.

1.1. Covering a graph efficiently

Covering the edges of a graph G by members of a given family of subgraphs $\mathcal{H}$, like cliques, matchings, trees, or cycles, is one of the basic themes in graph theory. In a sense, an efficient cover of a graph is one in which few edges are covered more than once. So, the most efficient way of covering $E(G)$ is by decomposing it into subgraphs from $\mathcal{H}$. However, not all graphs can be partitioned into specified types of subgraphs, even when it is possible to cover them with these subgraphs. For example, every complete graph on $n \geq 3$ vertices can be covered with triangles, but it can only be partitioned if $\mathrm{e}(K_n) = \binom{n}{2}$ is divisible by 3 and if every vertex has even degree.

In this spirit we will define the *covering size* of a graph G (with respect to a family of subgraphs $\mathcal{H}$), denoted by $\mathrm{cs}(G, \mathcal{H})$ to be the minimum total number of edges in a cover of G by graphs in $\mathcal{H}$.

So $\mathrm{cs}(G, \mathcal{H}) = \mathrm{e}(G)$ if and only if G can be decomposed into subgraphs from $\mathcal{H}$, and in this sense $\mathrm{cs}(G, \mathcal{H}) - \mathrm{e}(G)$ measures how far a graph is from being decomposable into graphs from $\mathcal{H}$.

Determining the length of a shortest cycle cover in a graph G without cut-edges, in our terminology $\mathrm{cs}(G, \{\text{cycles in } G\})$, is a widely investigated question that is related to some long standing conjectures.

In Chapter 2 we will study the problem of covering a graph with cuts, so we define $\mathrm{cs}(G) := \mathrm{cs}(G, \{\text{cuts in } G\})$. For a partition of the vertex set $V = S \cup \overline{S}$ the *cut* induced by S is the set of all edges joining vertices in S and $\overline{S}$. We will determine $\mathrm{cs}(K_n)$ and therefore also solve the extremal problem of determining the maximum value of $\mathrm{cs}(G)$ over all graphs on n vertices. We will also prove that the maximum of $\mathrm{cs}(G)$ over all k-colorable graphs on n vertices is achieved by the Turán graph $T_{n,k}$, and we will determine $\mathrm{cs}(T_{n,k})$ when $k > 200$ and when $k > 8$ and n is sufficiently large.

We will show that $2\mathrm{e}(G) - \mathrm{Cut}(G) \leq \mathrm{cs}(G) \leq 2\mathrm{e}(G) - \mathrm{Cut}'(G)$, where $\mathrm{Cut}(G)$ is the maximum size of a cut in G and $\mathrm{Cut}'(G)$ is the maximum size of a cut that is induced by an independent set of vertices. We show that the lower bound is achieved when G is 4-colorable – this also includes the case of planar graphs, in which case the problem is equivalent to finding the shortest cycle cover of the dual graph. On the other hand we will show that the upper bound is achieved by the random graph with edge probability $p \geq 6(\log^{1/3} n)/n^{1/3}$. We will also use these bounds to prove the following result for a graph and its complement: $2\binom{n}{2} - \mathrm{e}(T_{n,4}) \leq \mathrm{cs}(G) + \mathrm{cs}(\overline{G}) \leq \mathrm{cs}(K_n)$. Both bounds are achievable.

In Section 2.10 we will establish a connection between the problem of determining $\mathrm{cs}(G)$ and the following difficult isoperimetric problem due to Ahlswede and Katona [4]:

Find n points in the hypercube Q_k with minimum average pairwise distance. Equivalently, find n vectors in $\{0,1\}^k$ whose average pairwise Hamming distance is minimal. Equivalently, find n subsets of $[k] = \{1, \ldots, k\}$ whose average symmetric difference is as small as possible.

As an immediate corollary of our determination of $\mathrm{cs}(K_n)$, we will be able to solve this problem *exactly* when $k \geq n-1$. We will also generalize the original question as a graph representation problem and pose a number of problems that could be worth pursuing in the future.

1.2. Guarding art galleries

Consider an *art gallery* P formed by a polygon on n vertices with m pairs of vertices joined by nonintersecting interior diagonals, the *interior walls*. Each interior wall has an arbitrarily placed, arbitrarily small doorway. In general the goal is to determine the minimum number of guards that suffice to guard P, denoted by $g(P)$. A set of points S *guards* P if for every point x in the interior of P there is at least one point $y \in S$ such that the line segment between x and y does not intersect P. In this case we say that y *sees* x.

This problem can be interpreted as a geometric covering problem: The set $P(y)$ of all the points that a guard y can see is star-shaped (or star-convex) with center at y, i. e. the line segment from y to any point $x \in P(y)$ is completely contained in $P(y)$. We have to determine the minimum number of star-shaped sets with which we can cover the interior of the art gallery.

Determining the minimum number of guards that suffice to guard a given art gallery is NP-hard, even when there are no interior walls ($m = 0$). So we content ourselves with solving the natural extremal problem in Chapter 3: Determine the minimum number of guards $g(n,m)$ that suffice to guard *each* art gallery with n vertices and m interior walls. For $m = 0$ this was done by Chvátal [32], and we will extend Fisk's elegant proof [51] that "$\lfloor n/3 \rfloor$ are always sufficient and sometimes necessary" to the general case. We will show that $g(n,m) = \min\{\lfloor (2n-3)/3 \rfloor, \lfloor (2n+m-2)/4 \rfloor, \lfloor (2m+n)/3 \rfloor\}$.

If we restrict ourselves to art galleries where every room is convex and has at least r walls, then the answer improves to $\min\{m, \lfloor (n+m)/r \rfloor\}$. The proofs lead to $O(n)$ time guard placement algorithms in most cases.

1.3. A generalization of Turán's theorem

What is the maximum number of edges in a graph on n vertices if every k-set of vertices induces a subgraph of size at most $\binom{k}{2} - 1$? This question has been settled completely by Turán [113,114] (see also [97]), and variations of this problem have stimulated combinatorics ever since. In particular, if $\binom{k}{2} - 1$ is replaced by a parameter r, then the answer is given asymptotically by a theorem of Erdős, Simonovits and Stone [47,48] as

$$\frac{c-2}{c-1}\binom{n}{2} + \mathrm{o}(n^2),$$

where c is the smallest chromatic number among k-vertex graphs with $r+1$ edges. The exact answer to this question is difficult, and very little is known in the case when $c = 2$, since the question is directly related to the notorious problem of determining the maximum number of edges in a graph not containing a specified bipartite subgraph B. However, in a recent paper Griggs, Simonovits and Thomas [65] determine the exact answer in the case when the error term above is at most linear.

In Chapter 4 we attack the above problem for multigraphs – that is, when multiple edges are allowed. First we study the relaxed problem where we allow edges with negative multiplicities. A few cases of this problem have been solved by Bondy and Tuza [17], but the problem had already been suggested in earlier work of Brown, Erdős and Simonovits. We prove that for the relaxed problem the answer is of the form $m_{k,r}\binom{n}{2} + \mathrm{O}(n)$, and we determine $m_{k,r}$. We obtain exact answers for more than half the pairs (k, r), and for all (k, r) when n is small. Furthermore, in the case when $m_{k,r} = 0$ we determine the exact answer for n sufficiently large. Thus the relaxed problem is asymptotically completely solved, and we give a construction we conjecture to be optimal. Most of our constructions do not require negative weights, and we are thus able to asymptotically solve the original problem, when k is fixed, for all but finitely many r. We get exact answers in many cases. The results of Chapter 4 can also be found in [56].

Since our results are tight and fairly general, we expect that they will be valuable tools in the study of Turán type problems. In fact, using our results for the case $(k, r) = (4, 20)$, de Caen and Füredi [35] recently proved that the number of edges in a 3-uniform hypergraph not containing the Fano plane (the projective plane of order 2) is at most $\frac{3}{4}\binom{n}{3} + O(n^2)$. This is best possible and settles an old conjecture of Erdős and Sós [108].

1.4. Basic definitions

In this section we review basic terminology used throughout this thesis. Unfortunately notation in graph theory is far from standardized. We encourage readers that are unfamiliar with the basic graph theoretical concepts to consult the book of West [117], whose notation we usually follow.

For every $x \in \mathbb{R}$, $\lfloor x \rfloor$ and $\lceil x \rceil$ are the biggest integer not greater than x and the smallest integer not less than x, respectively. For every set S and integer $k \geq 0$, we define $\binom{S}{k}$ to be the set of all k-element subsets, also called *k-sets*, of S. If S is finite, then $\left|\binom{S}{k}\right| = \binom{|S|}{k}$.

We denote the vertex set and the edge set of a graph G by $V(G)$ and $E(G) \subset \binom{V(G)}{2}$, respectively. The *order* of G is $n(G) := |V(G)|$, which we usually abbreviate to n. The *size* of G is $\mathrm{e}(G) := |E(G)|$.

An edge $e \in E(G)$ corresponds to a 2-set of vertices $\{u, v\}$ which we usually abbreviate to $e = uv$. The vertices u and v are the *endpoints* of e, and we say that u and v are *incident* to e and that *e joins u and v*. When two vertices are joined by an edge, we say they are *adjacent* or that they are *neighbors*. Note that we will not consider graphs with *loops*, so in our graphs every edge has two distinct endpoints. In a graph every pair of vertices is joined by at most one edge. If we want to specifically allow that a pair of vertices can be joined by several (copies of the same) edges, then we speak of a *multigraph.* The number of edges joining u to v in a multigraph is the *multiplicity* of the edge uv.

The set of neighbors of a vertex v, also called its *neighborhood*, will be denoted by $N(v)$. The number of edges incident to v is its *degree* $d(v)$, so that for a graphs $d(v) = |N(v)|$. In an *even graph* every vertex has even degree. The minimum degree of a graph is denoted by $\delta(G)$ and the maximum degree by $\Delta(G)$. If every vertex has degree d, then G is *d-regular.* Since summing the degrees of each vertex counts every edge twice, we have $\sum_{v \in V(G)} d(v) = 2\mathrm{e}(G)$. A *matching* is a set of edges that are pairwise vertex disjoint. A matching *saturates* a vertex if there is an edge in the matching incident to the vertex.

The complement $\overline{G}$ of a graph G is the graph with $V(\overline{G}) = V(G)$ such that u and v are adjacent in $\overline{G}$ if and only if they are not adjacent in G. A *subgraph* of a graph G is a graph H such that $V(H) \subset V(G)$ and $E(H) \subset E(G)$, and we say that G *contains* H. If also $V(H) = V(G)$, then H is a *spanning subgraph* of G. If $E(H) = E(G) \cap \binom{V(H)}{2}$, then H is an *induced subgraph* of G and we say that $A = V(H)$ *induces* H, denoted by $H = G|_A$. For $S, T \subset V(G)$ (where we assume that $S \cap T = \emptyset$) we denote the set of edges from S to T by

$$[S, T] := \{uv \in E(G) : u \in S, v \in T\}.$$

If $T = \overline{S} = V(G) - S$, then we say that $[S, \overline{S}\,]$ is the *cut* induced by S.

Graphs G and H are *isomorphic* when there is a bijection $f : V(G) \to V(H)$, such that $uv \in E(G)$ if and only if $f(u)f(v) \in E(H)$. We sometimes say that H is a *copy* of G and we often use the same notation for a graph and its isomorphism class.

The *union* of G and H, denoted by $G \cup H$, is the graph with vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H)$. If $V(G)$ and $V(H)$ are disjoint we also write $G + H$ and call this the *disjoint union* of G and H. The disjoint union of k copies of G is denoted by kG.

A *complete graph* or *clique* is a graph in which every pair of vertices is adjacent, that is $E(G) = \binom{V(G)}{2}$. We denote the complete graph on n vertices by K_n, and its complement $\overline{K}_n$ has no edges. The order of the biggest complete subgraph of G is $\omega(G)$, the *clique number* of G. An *independent* or *stable set* of vertices is a subset $S \subset V(G)$ whose vertices are pairwise nonadjacent. The size of a biggest independent set of G is denoted by $\alpha(G)$.

A graph G is a *complete k-partite graph* if $V(G)$ can be partitioned into k parts, so that two vertices are adjacent if and only if they are in distinct parts. The *Turán graph* $T_{n,k}$ is the complete k-partite graph on n vertices where the part-sizes are as equal as possible, that is are all $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$. A graph is *k-partite*, or *k-colorable*, if it is a subgraph of some complete k-partite graph. The *chromatic number* $\chi(G)$ is the smallest k so that G is k-colorable, and we say that G is *$\chi(G)$-chromatic.* 2-colorable graphs are also called *bipartite* graphs and the complete bipartite graph, with parts of sizes m and n, is denoted by $K_{m,n}$.

A *path* on n vertices, denoted by P_n, is a graph G whose vertices can be ordered $v_1, v_2, \ldots, v_n$ so that $E(G) = \{v_i v_{i+1} : 1 \le i < n\}$. The vertices v_1 and v_n are the *endpoints* of the path, and G is a *v_1, v_n-path* of *length* $n - 1$. The *distance* $d(u, v)$ between vertices u and v is the minimum length of a u, v-path. A *cycle* on n vertices, denoted by C_n, is a graph G whose vertices can be ordered $v_1, v_2, \ldots, v_n$ so that $E(G) = \{v_i v_{i+1} : 1 \le i < n\} \cup \{v_n v_1\}$. Graphs isomorphic to $C_3 = K_3$ are frequently called *triangles.*

A graph is *connected* if it contains a uv-path for every pair of vertices $\{u, v\}$. G is k-*connected* [k-*edge-connected*] if the removal of any set of $k-1$ vertices [edges] results in a connected subgraph of G. A maximal connected subgraph of a graph is a *component*. A *cut-edge* is an edge whose removal increases the number of components of G. An edge is a cut-edge if and only if it is in no cycle. An *acyclic* graph, or *forest*, is a graph containing no cycle. A connected forest is a *tree*, so that every forest is a disjoint union of trees. A vertex of degree 1 is a *leaf*. A *star* is any graph isomorphic to $K_{1,n-1}$, that is a tree with at most one vertex that is not a leaf – the *center* of the star.

The hypercube Q_k is the graph with $V(Q_k) = \{0,1\}^k$ (all 0,1-vectors of length k), in which two vertices are adjacent if they differ in exactly one coordinate. The *Hamming distance* between two 0,1-vectors is exactly their distance as vertices in Q_k: the number of coordinates in which they differ.

For $x, y \in \mathbb{R}^2$, a *polygonal xy-path* is a finite sequence of closed line segments in $\mathbb{R}^2$, with the first one beginning at x, every other segment beginning at the endpoint of the previous one, the last one ending at y, and no point appearing in more than one segment except for common endpoints of consecutive segments. A *drawing* of a multigraph G is a function that maps each vertex $v \in V(G)$ to a distinct point $f(v) \in \mathbb{R}^2$ and each edge uv to a polygonal $f(u)f(v)$-path. A point that is contained in two polygonal paths, other than a common endpoint, is a *crossing*. A multigraph G is *planar* if it has a drawing without crossings and a *plane graph* (G, f) (usually abbreviated to G) is a planar graph together with a particular such drawing. A *region* $U \subset \mathbb{R}^2$ is an open set containing an xy-path for all $x, y \in U$. The *faces* of a plane graph are the maximal regions that are disjoint from the image of the drawing.

A *dual graph* (G^*, f^*) of a given plane graph (G, f) is a plane graph with the following properties: $f^*(V(G^*))$ contains exactly one point from every face of G, and every edge $e \in E(G)$ corresponds to an edge $e^* \in E(G^*)$ connecting the faces on opposite sides of $f(e)$, with $f^*(e^*)$ obtained by connecting the images of its endpoints via polygonal paths in the corresponding faces of (G, f) to a common point on $f(e)$. All duals of a given plane graph are isomorphic and thus we usually speak of "the" dual graph G^*.

To express the growth rate of a function we will frequently use the convenient "Oh" terminology, following the common abuse of notation: Let f and g be real-valued functions. If $\limsup_{x\to\infty} |f(x)/g(x)| < \infty$, then we write $f = \mathrm{O}(g)$ or equivalently $g = \Omega(f)$. We write $f = \Theta(g)$ if $f = \mathrm{O}(g)$ and $f = \Omega(g)$. If $\limsup_{x\to\infty} |f(x)/g(x)| = 0$, then we write $f = \mathrm{o}(g)$, or equivalently $g = \omega(f)$. We say that f is *asymptotic* to g if $\lim_{x\to\infty} f(x)/g(x) = 1$, and we write $f \sim g$. Note that $f \sim g$ is equivalent to $f = (1 + \mathrm{o}(1))g$.

COMPLEXITY: For an introduction to complexity and for good references on early results we refer to the book of Garey and Johnson [59]. The efficiency of an algorithm can be judged by its worst-case running time, over all possible inputs of size n, as a function of n. In complexity theory, the *time complexity* of a decision problem (that is a problem with a YES-NO answer) is the minimum worst-case running time over all algorithms that solve it, again as a function of the input size. Optimization problems can also be regarded as decision problems in that we have to decide whether the optimum is at most k for a given k.

Simply put, a decision problem is solvable in polynomial time if there is an algorithm solving the problem whose number of steps is bounded by a polynomial function of the size of the input. In other words, the complexity is $\mathrm{O}(n^k)$ for some constant k. The class of all decision problems that are solvable in polynomial time is denoted by P. A problem is in NP, the class of all *nondeterministic polynomial-time* solvable problems, if there is a polynomial-time algorithm to verify YES answers. For example, deciding whether a graph is 3-colorable is a problem in NP, since given a partition of $V(G)$ into 3 parts it can be verified in time $O(n^2)$ if it is a 3-coloring, by comparing the colors of the endpoints of every edge. Clearly $P \subset NP$, and it is widely believed that $P \neq NP$. A problem is *NP-hard* if a polynomial-time algorithm for it could be used to construct a polynomial-time algorithm for every problem in NP. If it also belongs to NP, then it is *NP-complete*. If some NP-hard problem can be shown to be in P, then $P = NP$. However there are so many NP-complete problems for which polynomial-time algorithms seem hard to find, that it is generally believed that $P \neq NP$.

2. Covering a graph with cuts of minimum total size

2.1. Covering the edges of a graph

Covering the edges of a graph G with members of a given family of subgraphs $\mathcal{H}$, like cliques, matchings, trees, or cycles, is a common problem in graph theory. Erdős, Goodman and Pósa [44] showed that the edges of every graph on n vertices can be decomposed into $\lfloor n^2/4 \rfloor$ cliques, and the balanced complete bipartite graph $T_{n,2} = K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ shows that this is best possible. It can also be desirable to minimize parameters other than the number of subgraphs used in the cover. Győri and Kostochka [69], Chung [30] and Kahn [81] independently proved the stronger result that every graph has a decomposition into cliques whose *order-sum* (sum of the orders of the cliques in the cover) is at most $\lfloor n^2/2 \rfloor$.

A host of similar question have been investigated since then. For example, Tuza [115] showed that the edges of every n-vertex graph can be covered using at most $n - \lfloor \lg n \rfloor + 1$ complete bipartite subgraphs. This settles (up to the $+1$) a conjecture of Bermond and is basically best possible, since Rödl showed that there are graphs requiring $n - c \log n$ complete bipartite subgraphs. Correspondingly, Chung, Erdős and Spencer [31] (see also [115]) proved that every sufficiently large graph can be covered with complete bipartite graphs of order-sum at most $(1+\varepsilon)n^2/2\log n$, and they show that this is best possible, up to a factor of $1/e$. Much earlier Katona and Szemerédi [82] had already proved that the order-sum of a cover of $E(K_n)$ with complete bipartite graphs is always at least $n \lg n$ and there are covers with order-sum $n\lceil \lg n \rceil$. Reznick, Tiwari and West [101] and Kratzke, Reznick and West [84] have studied the number of complete bipartite subgraphs needed to cover $E(G)$ in terms of the eigenvalues of the adjacency matrix of G.

In his survey paper with the suggestive title "Covering the edges of a graph by ... " Pyber [99] defines $\text{cov}(G, \mathcal{H})$ to be the minimum number of graphs in $\mathcal{H}$ that are required to *cover* $E(G)$. Similarly $\text{cov}^*(G, \mathcal{H})$ denotes the minimum number of subgraphs needed to *partition* or *decompose* $E(G)$. Hence if we let $\mathcal{H}$ be the set of cliques in G, then by [44]

$$\text{cov}(G, \mathcal{H}) \leq \text{cov}(T_{n,2}, \mathcal{H}) = \lfloor n^2/4 \rfloor.$$

They actually showed the stronger result that every graph can be decomposed into few triangles and edges: if $\mathcal{H}' = \{$subgraphs of G isomorphic to K_2 or $K_3\}$, then

$$\text{cov}^*(G, \mathcal{H}') \leq \lfloor n^2/4 \rfloor.$$

However, not all graphs can be partitioned into specified subgraphs, even if it is possible to cover them with these subgraphs. For example, every complete graph on $n \geq 3$ vertices can be covered with triangles, but it can be partitioned into triangles if and only if $n \equiv 1, 3 \pmod 6$. The condition is necessary because the existence of a decomposition requires that the number of edges is divisible by 3, and every degree is even. The sufficiency of this condition is one of the oldest results in combinatorial design theory; these decompositions are equivalent to *Steiner Triple Systems* (see [8] for a good introduction to design theory).

2.2. Covering a graph efficiently

An efficient cover of a graph is one in which few edges are covered more than once. In this spirit we will define the *covering size* of a graph G (with respect to a family of subgraphs $\mathcal{H}$), denoted by $\mathrm{cs}(G, \mathcal{H})$, to be the minimum total number of edges in a cover of G with graphs in $\mathcal{H}$, that is

$$\mathrm{cs}(G, \mathcal{H}) := \min\{\sum e(H) : \bigcup E(H) = E(G), H \in \mathcal{H}\}.$$

So $\mathrm{cs}(G, \mathcal{H}) = \mathrm{e}(G)$ if and only if G can be decomposed into subgraphs from $\mathcal{H}$, and in this sense $\mathrm{cs}(G, \mathcal{H}) - \mathrm{e}(G)$ measures how far a graph is from being decomposable into graphs from $\mathcal{H}$. If all members of $\mathcal{H}$ have k edges, e.g. if $\mathcal{H}$ is the set of all spanning trees of the connected graph G, then $\mathrm{cs}(G, \mathcal{H}) = k \cdot \mathrm{cov}(G, \mathcal{H})$. In less trivial cases the determination, or even just the estimation of $\mathrm{cs}(G, \mathcal{H})$ leads to challenging problems.

Determining the length of a *shortest cycle-cover* in a graph G without cut-edges, $\mathrm{scc}(G)$, is a widely investigated question that is connected to some long standing conjectures. In our terminology $\mathrm{scc}(G) = \mathrm{cs}(G, \{\text{cycles in } G\})$. Zhang's book on cycle-covers [119] devotes an entire chapter to this problem. The proofs of most of the cycle-cover results we mention can be found there. Recently Fan [50] proved that $\mathrm{scc}(G) \leq \mathrm{e}(G) + n(G) - 1$, thus settling a conjecture of Itai and Rodeh [74]. The conjecture of Alon and Tarsi [5] that $\mathrm{scc}(G) \leq \frac{7}{5}\mathrm{e}(G)$, achieved by the Petersen graph, still remains open. It has been shown by Jamshy and Tarsi [79] that this conjecture implies the famous cycle double cover conjecture: the edges of a graph without cut-edges can be covered by cycles so that every edge is in exactly 2 cycles.

In this chapter we will study the dual problem of covering a graph with cuts. That is, we investigate the parameter

$$\mathrm{cs}(G) := \mathrm{cs}(G, \{[S, \overline{S}\,] : S \subset V(G)\}).$$

Recall that a *cut* $[S, \overline{S}\,]$ consists of all the edges in $E(G)$ with one endpoint in S and the other in $\overline{S} = V(G) - S$. A *cut-cover* of a graph G is a collection of cuts $\mathcal{C} = \{[S_1, \overline{S}_1\,], [S_2, \overline{S}_2\,], \ldots, [S_k, \overline{S}_k\,]\}$ whose union is $E(G)$. The *total size* of $\mathcal{C}$ is the sum of the sizes $|[S_i, \overline{S}_i\,]|$ of the cuts in $\mathcal{C}$. Thus the *cut-cover size* of G, denoted by $\mathrm{cs}(G)$, is the minimum total size of a cover of $E(G)$ with cuts. We notice immediately that the *star cuts* form a double cover:

$$\{[v, V(G) - v] : v \in V(G)\}.$$

It will turn out that in a vast number of cases, namely for the graphs which we will call type B, a minimum size cut-cover can be obtained by using only some of these star cuts, see Theorem 2.3.7.

2.3. The cut-cover problem: Definitions and main results

It is well-known that the minimum number of cuts (or equivalently bipartite subgraphs) needed to cover the edges of a graph G with chromatic number $\chi(G)$ is $\lceil \lg \chi(G) \rceil$ (see, e.g., [60,71,96], see also [100]), where lg denotes the base 2 logarithm. To obtain such a covering we label the vertices in the j-th color class by the binary expansion of $j - 1$, thus associating with each vertex a $\{0,1\}$-vector of length $\lceil \lg \chi(G) \rceil$. From this labeling we can construct the desired cut-cover by letting the i-th cut $[S_i, \overline{S}_i]$ consist of all the edges between vertices whose labels differ in the i-th coordinate. In other words, $S_i := \{v \in V(G) : i\text{th coordinate of } v = 1\}$. These $\lceil \lg \chi(G) \rceil$ cuts cover all the edges of G, since adjacent vertices have different colors and therefore different labels. To see that this way of covering the graph with cuts is best possible, notice that we can extract a labeling of the vertices with binary vectors of length k from a cover with k cuts. Adjacent vertices must receive different labels, so that the labeling is a good-coloring with at most 2^k colors. We will frequently use the equivalent labeling to express a cut-cover.

In the cover by $\lceil \lg \chi(G) \rceil$ cuts the sum of the sizes of the cuts could be as big as $\lfloor n^2/4 \rfloor \lceil \lg \chi(G) \rceil$. When minimizing the total size, however, other ways of cutting the graph can be more efficient. It is the aim of this chapter to give upper and lower bounds on the minimum total size of a cut-cover. Some of the results will appear in [55].

Note that if G is a proper subgraph of H, then $\mathrm{cs}(G) < \mathrm{cs}(H)$ since every cut-cover for H gives rise to a cut-cover for G. Thus we immediately get the trivial bounds that $\mathrm{e}(G) \le \mathrm{cs}(G) \le \mathrm{cs}(K_n)$. Equality in the lower bound holds for all bipartite graphs. The cut-cover size of the complete graph will be determined in Section 2.5:

Theorem 2.3.1.

$$\mathrm{cs}(K_n) = \begin{cases} (n-1)^2 & n \neq 4, 8 \\ (n-1)^2 - 1 & n = 4, 8. \end{cases}$$

For complete graphs with at least 8 vertices the optimal cut-cover is unique, up to isomorphism. For $n > 8$, cover K_n with stars by taking $n-1$ cuts so that the S_i are distinct sets of size one. For $n = 8$, cover K_8 with $K_{4,4}$'s by taking 3 cuts such that $|S_i| = 4$, $|S_i \cap S_j| = 2$ for $i \neq j$ and $|S_1 \cap S_2 \cap S_3| = 1$.

For odd cycles we have $\mathrm{cs}(C_{2k+1}) = 2k+2$. Indeed, every cut in a cycle has even size, so $\mathrm{cs}(C_{2k+1})$ must also be even. Together with the trivial lower bound, this fact yields the lower bound. There are many different covers achieving this value.

This observation implies that when G has c disjoint odd cycles, the trivial lower bound can be improved to to $\mathrm{e}(G) + c$. Hence, in a sense, $\mathrm{cs}(G)$ measures how "nonbipartite" a graph is. To state the improved bounds that are the main subject of investigation in this chapter, we need to define the following parameters:

$$\mathrm{Cut}(G) := \max\{|[S, \overline{S}]| : S \subset V\},$$
$$\mathrm{Cut}'(G) := \max\{|[S, \overline{S}]| : S \subset V,\ S \text{ stable set}\}.$$

A vertex set S is *stable*, or *independent*, if the subgraph induced by S has no edges. A *stable cut* is a cut in which one of the partition sets forms a stable set. We denote the size of a maximum stable set in G by $\alpha = \alpha(G)$. If we denote the minimum and maximum degree in G by $\delta(G)$ and $\Delta(G)$, then we get

$$\alpha(G)\delta(G) \le \mathrm{Cut}'(G) \le \alpha(G)\Delta(G), \tag{2.3.1}$$

although a maximum stable cut need not have its stable set be maximum.

We will prove the following bounds in Section 2.4

Theorem 2.3.2.

$$2\mathrm{e}(G) - \mathrm{Cut}(G) \le \mathrm{cs}(G), \tag{2.3.2}$$

$$\mathrm{cs}(G) \le 2\mathrm{e}(G) - \mathrm{Cut}'(G). \tag{2.3.3}$$

There are a number of graphs for which Theorem 2.3.2 suffices to compute $\mathrm{cs}(G)$. For example in the case of the typically uncooperative Petersen graph P, we get $\mathrm{Cut}(P) = \mathrm{Cut}'(P) = 12$, so that $\mathrm{cs}(P) = 18$. Note that every graph can be embedded in a graph that is only slightly bigger for which $\mathrm{Cut} = \mathrm{Cut}'$.

Definition 2.3.3. Given disjoint graphs G and H, we define their *join* $G \vee H$ to be the graph obtained by adding the edges to make every vertex from $V(G)$ adjacent to every vertex in $V(H)$. We have $n(G \vee H) = n(G) + n(H)$, $\mathrm{e}(G \vee H) = \mathrm{e}(G) + \mathrm{e}(H) + n(G)n(H)$ and $\chi(G \vee H) = \chi(G) + \chi(H)$.

Proposition 2.3.4. *For every graph G and $m \ge \max\{\Delta(G), 1\}$, $G \vee \overline{K}_m$ contains G, $\chi(G \vee \overline{K}_m) = \chi(G) + 1$, $\mathrm{Cut}'(G \vee \overline{K}_m) = \mathrm{Cut}(G \vee \overline{K}_m) = nm$, and $\mathrm{cs}(G \vee \overline{K}_m) = 2\mathrm{e}(G) + nm$.*

Thus it is unlikely that there will be any forbidden subgraph characterizations for $\mathrm{cs}(G)$. There are, however, many cases when equality holds in (2.3.2) or (2.3.3). A graph is type A if equality holds in (2.3.2), otherwise it is type A'. Similarly it is type B if equality holds in (2.3.3) and type B' otherwise. A graph is type AB if it is both type A and type B and so on. Bipartite graphs are trivially type AB.

Theorem 2.3.5. *If G is 4-colorable, then it is of type A, that is*

$$\mathrm{cs}(G) = 2\mathrm{e}(G) - \mathrm{Cut}(G).$$

In Proposition 2.9.1 of Section 2.9 we will see that Theorem 2.3.5 is best possible in the sense that for every fixed number $k > 4$ and every one of the 4 possible types AB, $A'B$, AB' and $A'B'$ there are infinitely many graphs with chromatic number k and the specified type. In Section 2.7 we will determine the cut-cover size for large classes of random graphs by first proving

Theorem 2.3.6. *Let G be a graph with edge-density $d = \mathrm{e}(G)/\binom{n}{2}$. If G satisfies the conditions given below, then it is of type $A'B$.*

$$\alpha(G) \le d^2 n/100, \tag{2.3.4}$$

$$\text{If } |S_1|, |S_2| \ge dn/10, \text{ then } |[S_1, S_2]| \ge \frac{5}{6} d |S_1||S_2|, \tag{2.3.5}$$

$$\text{If } |S| \le n/2, \text{ then } |[S, \overline{S}]| \ge \frac{20}{43} dn|S|. \tag{2.3.6}$$

Almost all random graphs fulfill these 3 requirements. The probability space $\mathcal{G}(n,p)$ is defined for $0 \le p \le 1$, where p may depend on n. The ground set of $\mathcal{G}(n,p)$ is the set of all $2^{\binom{n}{2}}$ graphs with $V(G) = \{1, \ldots, n\}$, and the probability of a graph G is given by $\mathrm{Prob}(G_p = G : G_p \in \mathcal{G}(n,p)) = p^{\mathrm{e}(G)}(1-p)^{\mathrm{e}(\overline{G})}$. We say that for a graph property Q and a sequence of probabilities $p(n)$ *almost every graph* in $\mathcal{G}(n,p)$ has property Q if

$$\mathrm{Prob}(G_p \text{ has property } Q : G_p \in \mathcal{G}(n,p)) \to 1 \quad \text{as} \quad n \to \infty.$$

Theorem 2.3.7. *Almost every graph in $\mathcal{G}(n,p)$ is of type $A'B$ when the probability $p = p(n)$ satisfies $p \geq 6((\log n)/n)^{1/3}$, and of type AB when $p \leq (1-\varepsilon)/n$.*

In Section 2.8 we prove the following bounds on the sum of the cut-cover size of a graph and its complement:

Theorem 2.3.8. *For every n-vertex graph G*

$$2\binom{n}{2} - t(n,4) \leq \mathrm{cs}(G) + \mathrm{cs}(\overline{G}) \leq \mathrm{cs}(K_n),$$

and the bounds are best possible.

In Section 2.9 we will prove further bounds and exact results for special types of graphs.

Proposition 2.3.9. *If G is a complete k-partite graph on n vertices, then*

$$\mathrm{cs}(K_{k-1} \vee \overline{K}_{n-k+1}) \leq \mathrm{cs}(G) \leq \mathrm{cs}(T_{n,k}). \tag{2.3.7}$$

Furthermore

$$\mathrm{cs}(K_{k-1} \vee \overline{K}_{n-k+1}) = (k-1)(n-1), \tag{2.3.8}$$

except that $\mathrm{cs}(K_4) = 8$ and $\mathrm{cs}(K_8) = 48$. Also for all but finitely many pairs (n,k) with $n \geq k > 8$

$$\mathrm{cs}(T_{n,k}) = 2t(n,k) - \lceil \frac{n}{k} \rceil \left(n - \lceil \frac{n}{k} \rceil\right). \tag{2.3.9}$$

However, $\mathrm{cs}(T_{n,4}) = 2t(n,4) - \lceil \frac{n}{2} \rceil \lfloor \frac{n}{2} \rfloor$.

In Section 2.10 we give a geometric formulation for $\mathrm{cs}(G)$, and in Section 2.11 we pose several open questions.

2.4. Upper and lower bounds

Before solving the extremal problem of determining $\mathrm{cs}(K_n)$ in the next section, we will prove the simple bounds and show that many graphs achieve at least one of the bounds.

Theorem 2.3.2.

$$2\mathrm{e}(G) - \mathrm{Cut}(G) \leq \mathrm{cs}(G), \tag{2.3.2}$$

$$\mathrm{cs}(G) \leq 2\mathrm{e}(G) - \mathrm{Cut}'(G). \tag{2.3.3}$$

Proof. For the upper bound we need an efficient covering. Given S as in the definition of $\mathrm{Cut}'(G)$, we simply take a covering of $E(G)$ by stars centered in $\overline{S}$. Since S is a stable set this covers all edges in $E(G)$, and furthermore the edges in $[S, \overline{S}]$ are all covered exactly once, while all edges within $\overline{S}$ are covered exactly twice.

For the lower bound we consider the labeling mentioned in Section 2.3. That is with a given optimal covering by k cuts we identify a labeling of the vertex set with binary vectors of length k as follows: let the i-th entry of the label of v be 1 if $v \in S_i$ and 0 otherwise. The number of times an edge is cut is exactly the number of coordinates in which the two labels differ. Let the *weight* of a label be the number of ones it contains. If we now define S_{odd} to be the set of vertices with odd weight, then the edges that are covered once must be contained in $[S_{\mathrm{odd}}, \overline{S_{\mathrm{odd}}}]$, so that

$$\mathrm{cs}(G) \geq 2\mathrm{e}(G) - |[S_{\mathrm{odd}}, \overline{S_{\mathrm{odd}}}]| \geq 2\mathrm{e}(G) - \mathrm{Cut}(G). \quad \square$$

Facts 2.4.1. If $N(u) = N(v)$, so in particular u and v are not adjacent, then

- in a maximum cut u and v can be assumed to be on the same side of the cut, since otherwise the vertex with the fewer crossing edges can be moved to the other side.
- in an optimal cut-cover u and v can be assumed to be on the same side in every cut, since otherwise the vertex with the bigger total number of crossing edges can be moved to the side of the other vertex in every cut.

Proposition 2.3.4. *For every graph G and $m \geq \max\{\Delta(G), 1\}$, $G \vee \overline{K}_m$ contains G, $\chi(G \vee \overline{K}_m) = \chi(G)+1$, $\mathrm{Cut}'(G \vee \overline{K}_m) = \mathrm{Cut}(G \vee \overline{K}_m) = nm$, and $\mathrm{cs}(G \vee \overline{K}_m) = 2\mathrm{e}(G) + nm$.*

Proof. Since vertices in $V(\overline{K}_m)$ all have the same neighborhood, we can assume that in a maximum cut they are on the same side, say $V(\overline{K}_m) \subset S$. Furthermore all vertices in $V(G)$ have at least as many neighbors in $V(\overline{K}_m)$ as in $\overline{S}$, since $m \geq \Delta$, so that we can move them to $\overline{S}$ without decreasing the size of the cut. Thus $[V(\overline{K}_m), V(G)]$ is a maximum cut with size nm. It is a stable cut, since $V(\overline{K}_m)$ is a stable set. □

Theorem 2.3.5. *If G is 4-colorable, then it is of type A, that is*

$$\mathrm{cs}(G) = 2\mathrm{e}(G) - \mathrm{Cut}(G).$$

Proof. To see that the lower bound can be achieved it suffices to construct a covering such that all the edges in the maximum cut $[S, \overline{S}]$ are covered once and all other edges are covered twice. Denote the color classes by V_i $(1 \leq i \leq 4)$ and let $V_i \cap S = V_i^1$ and $V_i \cap \overline{S} = V_i^0$. We define the required covering by defining the equivalent labeling as suggested in Figure 2.1; that is we give all vertices in V_i^j the label as indicated. Then $E(G)$ is covered by 3 cuts, determined by $S_1 = V_1^1 \cup V_2^1 \cup V_3^0 \cup V_4^0$, $S_2 = V_1^1 \cup V_2^0 \cup V_3^1 \cup V_4^0$ and $S_3 = V_1^1 \cup V_2^0 \cup V_3^0 \cup V_4^1$ respectively.

V_1^0 000	V_2^0 011	V_3^0 101	V_4^0 110	$\overline{S}$
111 V_1^1	100 V_2^1	010 V_3^1	001 V_4^1	S
V_1	V_2	V_3	V_4	

Figure 2.1

Edges within S (or $\overline{S}$) are covered twice, since the labels on the vertices of these edges all have odd (respectively even) weights and must thus differ by 2. Edges between S and $\overline{S}$ are covered once, because the weights of the vertices involved have different parity (and hence are covered 1 or 3 times), but since each V_i is a stable set no edge will be covered three times. □

2.5. Extremal cut-cover problems

In this section we will solve the extremal problem of determining the maximum cut-cover size among all n-vertex graphs, $f(n)$. By monotonicity we see that $f(n) = \mathrm{cs}(K_n)$

and that only K_n achieves equality, hence it remains to determine $\mathrm{cs}(K_n)$. This problem had already been suggested by Tarsi [111], and he gives the values of $f(n)$, for $n \leq 8$, without providing a proof. Tarsi investigated the problem of determining the maximum length of a shortest circuit cover, taken over all regular matroids that admit a nowhere zero k-flow. The cocycle matroid of a graph has a nowhere zero k-flow exactly when the graph is k-colorable. Since the class of regular matroids includes both the cycle and cocycle matroids, his question includes that of finding short cycle-covers in graphs with nowhere zero k-flows, as well as short cut-covers in k-colorable graphs.

In [78] Jamshy and Tarsi determine $f(n)$ for $n \geq 9$ by a nice inductive argument. However, they do not establish the base case $n = 9$, so that the proof is incomplete and would still require some amount of case analysis. They also mention that the result had been independently obtained by Alon (personal communication) and Jaeger, Khelladi and Mollard [77]. While the proof in [77] is complete, it involves a 4 page long case analysis. We will give a simpler proof in this section.

A cut $[S, \overline{S}]$ in K_n is isomorphic to a complete bipartite graph $K_{s,n-s}$, where $|S| = s$. A cover is a collection of cuts $\mathcal{C} = \{C_1, C_2, \ldots, C_k\}$ whose union is $E(K_n)$. So the i-th cut C_i is of the form $K_{s_i,n-s_i}$ and we will denote the smaller partite set (with $s_i \leq n/2$ vertices) by S_i.

Theorem 2.5.1. *$f(n) = (n-1)^2$ for $n \geq 1$, except that $f(n) = (n-1)^2 - 1$ when $n \in \{4, 8\}$.*

Proof [86]. For the upper bound, a covering with $n-1$ stars (that is $|S_i| = 1$) yields $f(n) \leq (n-1)^2$. Also $f(4) \leq 8$ by using two $K_{2,2}$'s and $f(8) \leq 48$ by using three $K_{4,4}$'s.

For the lower bound, let $\mathcal{C}$ be a cover of minimum total size, and let k be the number of cuts in $\mathcal{C}$. Again, encode the vertices $v \in V(K_n)$ as binary k-vectors by letting the i-th coordinate of v be 1 when $v \in S_i$ and 0 otherwise. Observe that the number of times any edge uv is covered is exactly the number of coordinates in which u and v differ. So if we define S_{odd} to be the set of vertices of odd weight, as in the proof of (2.3.2), then $C = [S_{\text{odd}}, \overline{S_{\text{odd}}}]$ is exactly the set of edges that appear in an odd number of cuts C_i. Again C is equivalent to a complete bipartite graph $K_{s,n-s}$ with $s \leq n/2$.

Let $X \subset C$ be the set of edges in K_n that are covered once, and let $X_i = X \cap C_i$ be the set of edges that are covered only in C_i. Each X_i is a matching, since if $uv, vw \in X_i$, then u and w would be in the same partite set in every C_j, and thus uw would not be covered. Also X_i is a maximum matching in C_i, since otherwise the vertices in S_i that are not saturated could be moved to the other partite set in C_i, thus decreasing the total size of the cover.

Let $x = |X|$. Since each X_i is a maximum matching in C_i, we have $|X_i| = s_i$ and $x = \sum s_i$. Since $X_i \subset C$, also $s_i \leq s$. Therefore,

$$f(n) = \sum s_i(n - s_i) \geq \sum s_i(n - s) = x(n - s). \tag{2.5.1}$$

On the other hand, all edges in $K_n - C$ are covered at least twice, while those in X are covered once and those in $C - X$ are covered at least three times. Thus

$$f(n) \geq 2\binom{n}{2} + s(n - s) - 2x. \tag{2.5.2}$$

Since the lower bound in (2.5.1) is increasing in x and the lower bound in (2.5.2) is decreasing in x, we get the smallest maximum when the two expressions are equal, which occurs when

$$x = \frac{n(n-1)+s(n-s)}{n-s+2}.$$

By using this value for x, we obtain a lower bound independent of x. For fixed n, we define $g(s)$ to be the value of this lower bound.

$$f(n) \geq \frac{n-s}{n-s+2}(n^2 - n + sn - s^2) = g(s).$$

By showing that g is concave down ($g''(s) < 0$) on $[1, n/2]$, we conclude that $g(s)$ attains a minimum at one of the extremes $s = 1$ or $s = \lfloor n/2 \rfloor$. The value of s cannot be 0, since otherwise every edge would be covered at least twice and therefore one of the cuts could be omitted from the cover.

$$g''(s) = -2(n-s+2)^{-3}[2n(n-1)+2s(n-s)+(n-s+2)^2(n-s)+2(n-s+2)(n-2s)].$$

We compute that $g(1) = (n-1)^2$, $g(n/2) = n^2(5n-4)/(4n+16)$, and $g((n-1)/2) = (n+1)(n-1)(5n+1)/(4n+20)$.

For the even case, $g(1) \leq g(n/2)$ when

$$0 \leq n^2(5n-4) - (n-1)^2(4n+16) = n^3 - 12n^2 + 28n - 16$$
$$= (n-2)(n-5-\sqrt{17})(n-5+\sqrt{17}),$$

which occurs when $n \leq 2$ or $n \geq 10$. In the odd case, $g(1) \leq g((n-1)/2)$ when

$$\begin{aligned} 0 &\leq (n+1)(n-1)(5n+1) - (n-1)^2(4n+20) \\ &= (n-1)(n-3)(n-7), \end{aligned}$$

which occurs when $n \neq 5$.

So we know that $f(n) \geq g(1) = (n-1)^2$, except possibly when $n = 4, 5, 6, 8$. For $n = 4, 5, 8$ we compute quickly that $f(4) \geq g(2) = 8$ and $f(8) \geq g(4) = 48$ and $f(5) \geq g(2) = 15.6$ (but $(5-1)^2 = 16$).

For $n = 6$ we have $(n-1)^2 = 25$, but $f(6) \geq g(3) = 23.4$. The inequalities (2.5.1) and (2.5.2) imply that a value of 24 is only possible when $s = 3$ and $x = 8$. Since the X_i must be maximum matchings, the 8 edges in X are partitioned by the C_i's so that $\{s_1, \ldots, s_k\}$ with $k \leq 4$ is one of $\{3, 3, 2\}$ (yielding 26 edges), $\{3, 3, 1, 1\}$ (28 edges), $\{3, 2, 2, 1\}$ (30 edges) and $\{2, 2, 2, 2\}$ (32 edges). For $k \geq 5$, $5k$ edges is certainly the fewest we can have. □

We can prove the following bound for $\mathrm{cs}(G)$ in terms of $\mathrm{e}(G)$:

Theorem 2.5.2.

(2.5.3) $$\mathrm{cs}(G) \leq 2\mathrm{e}(G) - \left\lceil \sqrt{2\mathrm{e}(G) + 1/4} - 1/2 \right\rceil,$$

with equality for $G = K_k$ and for $G = K_k + K_2$ if $k \neq 4, 8$. Furthermore,

(2.5.4) $$\max\{\mathrm{cs}(G) : \mathrm{e}(G) = e\} = 2e - \sqrt{2e} - \mathrm{O}(e^{1/4}).$$

Proof. Let S be a stable set in G so that $[S,\overline{S}\,]$ is maximum and, subject to this constraint, has the maximum number of vertices, say $|S| = s$. Therefore, with $n(G) = n$,

$$\text{(2.5.5)} \qquad \mathrm{Cut}'(G) = |[S,\overline{S}\,]| \geq \mathrm{e}(G) - \binom{n-s}{2}$$

since $G|_{\overline{S}}$ can contain at most $\binom{n-s}{2}$ edges and S is a stable set. On the other hand, every vertex $v \in \overline{S}$ must have an edge going to S, since otherwise we can move v to S and thus increase $|S|$ without decreasing the size of the cut. But then

$$\text{(2.5.6)} \qquad \mathrm{Cut}'(G) = |[S,\overline{S}\,]| \geq n - s.$$

If we let $n - s = x$, then (2.5.5) is decreasing in x while (2.5.6) is increasing. Hence the minimum lower bound is achieved if the right-hand sides are equal,

$$x = \mathrm{e}(G) - \binom{x}{2}.$$

Solving this quadratic equation and discarding the negative solution we find that $x = -\frac{1}{2} + \sqrt{2\mathrm{e}(G) + 1/4}$, so that $\mathrm{Cut}'(G) \geq \lceil x \rceil$ and the inequality follows from (2.3.3).

To see that we can achieve equality suppose that $\mathrm{e}(G) = \binom{k}{2}$. Then we obtain

$$-\frac{1}{2} + \sqrt{2\mathrm{e}(G) + 1/4} = -\frac{1}{2} + \sqrt{(k - 1/2)^2} = k - 1,$$

so that $\mathrm{cs}(G) \leq (k-1)^2$, which by Theorem 2.5.1 is achieved for $G = K_k$, as long as $k \neq 4, 8$. Furthermore, if $\mathrm{e}(G) = \binom{k}{2} + 1$, then $\lceil -\frac{1}{2} + \sqrt{2\mathrm{e}(G) + 1/4} \rceil = k$, so that $\mathrm{cs}(G) = (k-1)^2 + 1$ can be achieved by $K_k + K_2$, as long as $k \neq 4, 8$.

The upper bound for (2.5.4) follows by (2.5.3). To establish the lower bound suppose that $e = \binom{m}{2} + \binom{t}{2} + r$ with $m > \binom{t}{2} + r$ and $t > r$ and let $G = K_m + K_t + rK_2$. So we take a clique with as close to e edges as possible, and then another maximal clique on the remaining edges. Thus by Theorem 2.5.1

$$\mathrm{cs}(G) \geq ((m-1)^2 - 1) + ((t-1)^2 - 1) + r = 2e - (m + t + r).$$

From the choice of m and t we obtain that $\sqrt{2e} \geq m - 1$ and $\sqrt{2\sqrt{2e}} \geq \sqrt{2(m-1)} \geq t - 1 \geq r$, so that

$$\mathrm{cs}(G) \geq 2e - \sqrt{2e} - 2\sqrt{2\sqrt{2e}} - 2 = 2e - \sqrt{2e} - \mathrm{O}(e^{1/4}). \quad \square$$

When $\mathrm{e}(G)$ is not of the form $\binom{k}{2}$ or $\binom{k}{2} + 1$, then (2.5.3) can probably be improved a bit. It seems difficult to determine which graphs on e edges minimize $\mathrm{Cut}'(G)$. A natural conjecture would be that this extreme value can be achieved by a disjoint union of cliques. However there are cases when other graphs are better; for example if $e = 26$ then identifying K_7 and K_4 at an edge yields a graph with $\mathrm{Cut}'(G) = 9$, whereas

$$\mathrm{Cut}'(K_6 + K_5 + K_2) = 10 = \mathrm{Cut}'(K_7 + K_3 + 2K_2),$$

which is the best among all disjoint unions of cliques. The following observation is due to Tao Jiang: If $e = \binom{m}{2} + \binom{t}{2} + r$ as above then $\mathrm{Cut}'(G) \le m + t$, where G is obtained from $K_m + K_t$ by adding a matching with r edges between K_m and K_t. Thus

$$\sqrt{2e + 1/4} - 1/2 \le \max\{\mathrm{Cut}'(G) : \mathrm{e}(G) = e\} \le \sqrt{2e} + (8e)^{1/4},$$

giving an even sharper bound on $\mathrm{Cut}'(G)$.

Another extremal problem is that of determining the maximum cut-cover size among all k-colorable graphs on n vertices. We will prove Theorem 2.3.9, which addresses this situation, in Section 2.9. The problem of determining the maximum cut-cover size among all k-colorable graphs on e edges is investigated in [77]. Using an averaging argument it is shown that for every k-colorable graph G

$$\mathrm{cs}(G) \le \mathrm{cs}(K_k)\mathrm{e}(G)/\binom{k}{2}.$$

2.6. Algorithmic aspects

The algorithmic aspect of the shortest cycle-cover problem is well-studied. Recently Thomassen [112] proved that determining $\mathrm{scc}(G)$ is NP-complete, thus settling a conjecture of Itai, Lipton, Papadimitriou, and Rodeh [75]. However, for some classes of graphs, like planar graphs [13,66], $\mathrm{scc}(G)$ can be computed in polynomial time:

A *postman tour* in a connected graph G is a closed walk that covers $E(G)$. The *Chinese postman problem* is to find a postman tour of minimum length $\ell(G)$. Edmonds and Johnson [39] showed that a Chinese postman tour can be found in polynomial time. It is immediate that $\ell(G) \le \mathrm{scc}(G)$, since every cycle-cover of a connected graph can be used to obtain a postman tour of length equal to the size of the cover. Equality need not hold since we may not be able to decompose a Chinese postman tour into cycles, as in the case of the Petersen graph P, where $\ell(P) = 20 < 21 = \mathrm{scc}(P)$. However, Alspach, Goddyn and Zhang [6] showed that $\ell(G) = \mathrm{scc}(G)$ for every graph not containing a subdivision of the Petersen graph (for graphs with a nowhere-zero 4-flow this had been shown already by Jackson [76] and Zhang [120]). Thus for this big class of graphs, which includes planar graphs, $\mathrm{scc}(G)$ can be computed in polynomial time.

We have similar results for the cut-cover problem. It is well-known that the related problem of determining $\mathrm{Cut}(G)$ is NP-complete [61]. This has been sharpened by Yannakakis [118] who showed that determining $\mathrm{Cut}(G)$ is NP-complete for graphs of maximum degree $\Delta \le 3$. Thus Theorem 2.3.5 implies that determining $\mathrm{cs}(G)$ is NP-complete even for graphs with maximum degree $\Delta(G) \le 3$, because by Brook's Theorem [18] all graphs with $\Delta(G) \le 3$ are 4-colorable.

On the other hand

Theorem 2.6.1. *Every planar graph G is of type A and a minimum size cut-cover can be found in polynomial time.*

Proof 1. If we assume the 4 Color Theorem (every loopless planar graph is 4-colorable), then Theorem 2.3.5 implies that every planar graph is type A. For planar graphs, Hadlock [70,9] proved that a maximum (weighted) cut can be found in polynomial time.

Robertson, Sanders, Seymour and Thomas [102] observe that their proof of the 4 Color Theorem could be turned into an $\mathrm{O}(n^2)$ algorithm for finding a proper 4-coloring. $\square$

The following proof and the observation that the problem is connected to T-joins were also provided by A. Kostochka:

Proof 2. We can also prove Theorem 2.6.1 without the help of the 4 Color Theorem, by considering the dual graph. If G is a (loopless) plane graph, then the dual graph G^* has no cut-edge and it can be found in linear time. A *bond* in a graph is a minimal nonempty cut. Every cut can be decomposed into bonds and a bond in G corresponds to a cycle in G^* (see, e.g. [117, 7.1.12]), so that

$$\begin{aligned}\mathrm{cs}(G) &= \mathrm{cs}(G, \{\text{cuts in } G\}) = \mathrm{cs}(G, \{\text{bonds in } G\}) \\ &= \mathrm{cs}(G^*, \{\text{cycles in } G^*\}) = \mathrm{scc}(G^*) = \ell(G^*).\end{aligned}$$

So by the result of Edmond and Johnson [39], $\mathrm{cs}(G)$ can be determined in polynomial time. Furthermore, since a bond corresponds to a cycle in the dual, a cut in G corresponds to an edge-disjoint union of cycles in G^*, i.e. an even subgraph of G^*. Notice that if H is an even subgraph of G^*, then in $H' = G^* - E(H)$ the degree of every vertex will have the same parity as in G^*, i.e. H' is a *parity subgraph* of G^*. Thus

$$\begin{aligned}2e(G) - \mathrm{Cut}(G) &= 2e(G^*) - \max\{e(H) : H \text{ is an even subgraph of } G^*\} \\ &= e(G^*) + \min\{e(H') : H' \text{ is a parity subgraph of } G^*\} = \ell(G^*),\end{aligned}$$

where the last equality is easy (see [119, 8.1.4]). Thus G is type A. $\square$

The algorithms in [70] and [39] are based on the idea that it basically suffices to find a smallest parity subgraph in G^*. This can be done by finding a minimum weight perfect matching in an auxiliary graph: The graph is the complete graph whose vertex set consists of the odd degree vertices in G^* with the weight of each edge being the distance between the two vertices in G^*. The fastest algorithm for this problem currently requires $\mathrm{O}(n^{5/2}(\log n)^{3/2}\alpha(n^2,n)^{1/2})$ steps [58], where α denotes the (very slowly growing) inverse of the Ackerman function. This essentially determines the running time for the algorithms obtained by either approach. One way to obtain a shortest cycle cover in G^* in polynomial time from the parity subgraph is by an algorithm of Fleischner and Frank [52] which avoids the 4 Color Theorem.

2.7. Cut-covers in random graphs

In this section we will prove Theorems 2.3.6 and 2.3.7 and thereby determine the cut-cover size of a wide range of random graphs.

Theorem 2.3.6. *Let G be a graph with edge-density $d = \mathrm{e}(G)/\binom{n}{2}$. If G satisfies the conditions given below, then it is of type $A'B$.*

$$\alpha(G) \le d^2n/100, \tag{2.3.4}$$

$$\text{If } |S_1|, |S_2| \ge dn/10, \text{ then } |[S_1,S_2]| \ge \frac{5}{6}d|S_1||S_2|, \tag{2.3.5}$$

$$\text{If } |S| \le n/2, \text{ then } |[S,\overline{S}\,]| \ge \frac{20}{43}dn|S|. \tag{2.3.6}$$

Proof. Note that (2.3.4) immediately implies that $n \ge 100$, so that this theorem only holds for large graphs. By (2.3.4) and (2.3.6),

$$\mathrm{Cut}'(G) \le \alpha(G)(n-\alpha(G)) \le \frac{d^2n}{100}\left(n - \frac{d^2n}{100}\right) < \frac{10}{43}dn(n-1) \le \mathrm{Cut}(G),$$

so that G cannot be type AB.

Next we show that

$$\text{If } U \subset V(G) \text{ and } |U| \ge dn/5+1, \text{ then } |E(U)| \ge \frac{5}{6}d\binom{|U|}{2}. \tag{2.7.1}$$

Indeed, let $t = \lfloor |U|/2 \rfloor \ge dn/10$. Then (2.7.1) is implied by (2.3.5):

$$2\binom{|U|-2}{t-1}|E(U)| = \sum_{T\subset U, |T|=t} |[T,U-T]| \ge \binom{|U|}{t}\frac{5}{6}dt(|U|-t).$$

To see that G is type B let $\{[S_i,\overline{S_i}] : 1 \le i \le k\}$ be an optimal cover. We can assume without loss of generality that $|S_i| \le \frac{n}{2}$ for all $1 \le i \le k$. We let $s = \sum_{i=1}^{k} |S_i|$ and claim that

$$s < 2.15n. \tag{2.7.2}$$

Indeed, $dn^2 > 2e(G) > \mathrm{cs}(G) = \sum_{i=1}^k |[S_i, \overline{S}_i]| \geq \sum_{i=1}^k \frac{20}{43} dn|S_i| = \frac{20}{43} dns$.

Furthermore we define, for $0 \leq j \leq k$,

$$E_j := \{e \in E(G) : e \text{ is covered in exactly } j \text{ cuts}\}, \qquad E_{\geq j} := \bigcup_{j \leq \ell \leq k} E_\ell,$$

$$W_j := \{v \in V(G) : v \text{ is in exactly } j \text{ of the } S_i\}, \qquad W_{\geq j} := \bigcup_{j \leq \ell \leq k} W_\ell.$$

In the labeling equivalent to the covering, the vertices in W_j are exactly those of weight j. Therefore an edge between a vertex in W_j and $W_{j'}$ with $j \leq j'$ is covered in exactly $j' - j + 2m$ ways for some $0 \leq m \leq j$. The edges in $E_1 \cap [S_i, \overline{S}_i]$ form a graph with the property that the neighborhood of any vertex is a stable set in G. Indeed, if the edges uv and uw are only established in $[S_i, \overline{S}_i]$, then v and w must be in the same partite set for every $[S_\ell, \overline{S}_\ell]$ so that vw would not be covered, and thus there can be no such edge in $E(G)$. So we can use (2.3.4) to conclude that the maximum degree in this graph is at most $\alpha \leq d^2n/100$. Thus $|E_1 \cap [S_i, \overline{S}_i]| \leq |S_i| d^2 n/100$ and

$$|E_1| = \sum_{i=1}^k |E_1 \cap [S_i, \overline{S}_i]| \leq \sum_{i=1}^k |S_i| \frac{d^2 n}{100} = s\frac{d^2 n}{100} < 0.0215 d^2 n^2. \tag{2.7.3}$$

Also $|E_0| = 0$, so that $2e(G) > \mathrm{cs}(G) = \sum_{j=1}^k j|E_j|$ implies that

$$|E_{\geq 3}| < |E_1| < 0.0215 d^2 n^2 \text{ and } |E_4| < |E_1|/2 < 0.011 d^2 n^2. \tag{2.7.4}$$

It is also important that in our labeling the set of vertices with a fixed label is a stable set, so that for example $|W_0| \leq d^2 n/100$.

To show that G is type B we give a two-step proof that W_1 is large and in a third step we show that in an optimal cover $W_{\geq 2} = \emptyset$. This will finish our proof, since W_0 is a stable set, the edges in $[W_0, W_1]$ are covered once, and all other edges are covered twice.

Step 1 : $|W_1| > dn/10$. Note that $n = \sum_{j=0}^k |W_j|$ so that

$$2.15n > s = \sum_{i=1}^k |S_i| = \sum_{j=0}^k j|W_j| = 2n + \sum_{j=0}^k (j-2)|W_j| \geq 2n - 2|W_0| - |W_1| + |W_{\geq 3}|.$$

But then

$$\begin{aligned} |W_2| &= n - |W_0| - |W_1| - |W_{\geq 3}| > n - |W_0| - |W_1| - (0.15n + 2|W_0| + |W_1|) \\ &= 0.85n - 3|W_0| - 2|W_1| \geq 0.85n - 3\frac{d^2 n}{100} - 2|W_1|, \end{aligned}$$

so that if we assume that $|W_1| \leq dn/10$, then $|W_2| > 0.62n$.

Next we define $A_i = S_i \cap W_2$ and $B_i = \overline{S}_i \cap W_2$. Notice that $|B_i| \geq |W_2| - |S_i| > 0.12n$. The edges in $[A_i, B_i]$ are covered either two times or four times, since they are between vertices in W_2. If $v \in B_1$, then without loss of generality $v \in A_2, A_3$ and $v \in B_i$, for $i > 3$, so the label of v is $011000\ldots$. If the edge $uv \in [A_1, B_1]$ is covered twice, then $u \in A_1$, so that the label of u is $110000\ldots$ or $101000\ldots$. However each of these labels induces an independent set, so that at most $2d^2n/100$ edges at v in $[A_1, B_1]$ are covered twice.

But now we can conclude that $|A_1| \leq dn/6$, since otherwise:

$$|E_4| \geq |[A_1, B_1]| - |B_1| 2\frac{d^2n}{100} \geq \frac{5}{6}d|A_1||B_1| - |B_1|\frac{d^2n}{50}$$
$$> |B_1|\big(\frac{5}{6}d\frac{1}{6}dn - \frac{d^2n}{50}\big) > 0.12n\frac{107}{900}d^2n > 0.011d^2n^2,$$

which is a contradiction to (2.7.4). Since we can argue in the same fashion for $i > 1$, we can now assume that $|A_i| \leq dn/6$ for all i.

Again we observe that if an edge uv in $[A_i, B_i]$ is covered twice, then its vertices u and v have a 1 in the same position somewhere, that is $u, v \in A_\ell$ for some ℓ. Also $\sum_{i=1}^k |A_i| = 2|W_2|$, since each vertex in W_2 is in exactly 2 A_i, so that the number of edges covered twice within W_2 is at most

$$\sum_{i=1}^{k} \binom{|A_i|}{2} \leq \frac{2|W_2|}{dn/6}\binom{dn/6}{2} < \frac{dn}{6}|W_2|.$$

All other edges in W_2 are covered 4 times, so using (2.7.1) and the fact that $n \geq 100$ we get a contradiction to (2.7.4) again:

$$|E_4| > |E(W_2)| - \frac{dn}{6}|W_2| \geq \frac{5}{6}d\binom{|W_2|}{2} - \frac{dn}{6}|W_2|$$
$$= \frac{d|W_2|}{6}\big(\frac{5}{2}(|W_2| - 1) - n\big) \geq 0.1dn(1.55n - 2.5 - n) > 0.05dn^2.$$

<u>Step 2 : $|\overline{W_1}| < dn/10$.</u> Suppose to the contrary that $|\overline{W_1}| \geq dn/10$. From an argument similar to the argument above we see that only few edges from W_1 to $\overline{W_1}$ can be covered twice: These edges are in $[W_1, W_3]$ and for every vertex $v \in W_3$, say with label $1110000\ldots$, there are only three labels possible for a vertex $u \in W_1$ such that the edge uv is covered exactly twice: $1000000\ldots$, $0100000\ldots$ and $0010000\ldots$, so that

$$|\overline{E_2}| \geq |[W_1, \overline{W_1}]| - 3\frac{d^2n}{100}|W_3| \geq \frac{5}{6}d|W_1||\overline{W_1}| - \frac{3d^2n}{100}|\overline{W_1}|$$
$$= (n - |W_1|)d\big(\frac{5}{6}|W_1| - \frac{3dn}{100}\big).$$

This concave function in $|W_1|$ is minimized at one of the endpoints of the interval on which it is defined: $|W_1| = dn/10$ or $|W_1| = n - dn/10$. Both of the values obtained are still greater than $0.043d^2n^2$, which is by (2.7.4) an upper bound on $|\overline{E_2}| = |E_1| + |E_{\geq 3}|$.

<u>Step 3 : $|W_{\geq 2}| = 0$.</u> If $|W_{\geq 2}| > 0$, then we will be able to obtain a better cover by moving the vertices in $W_{\geq 2}$ to W_1. We leave the vertices in $W_{\leq 1}$ unchanged, but change the labels of the vertices in $W_{\geq 2}$ so as to obtain a cover by stars. That is for every vertex in $W_{\geq 2}$ we introduce a new coordinate, and make its label in that coordinate 1, all other

coordinates will be zero. Vertices in $W_{\le 1}$ will receive zeros in the new coordinates, so that the number of times the edges in $E(W_{\le 1})$ are covered does not change. All other edges are now covered at most twice.

Before this change, the only edges not in $E(W_{\le 1})$ that were covered at most twice were contained in $E(W_{\ge 2})$, $[W_0, W_2]$, $[W_1, W_2]$ or $[W_1, W_3]$, which adds up to at most

$$\binom{|W_{\ge 2}|}{2} + |W_0||W_2| + 2\frac{d^2n}{100}|W_2| + 3\frac{d^2n}{100}|W_3| \le |W_{\ge 2}|\big(\frac{1}{2}|W_{\ge 2}| + \frac{3d^2n}{100}\big)$$
$$< 0.08dn|W_{\ge 2}|.$$

However from (2.3.6) we know that $|[W_{\ge 2}, W_{\le 1}]| \ge \frac{20}{43}dn|W_{\ge 2}| > 2(0.08dn|W_{\ge 2}|)$, so that more of the edges that we changed were covered 3 or more times, than once or twice. $\square$

Theorem 2.3.7. *Almost every graph in $\mathcal{G}(n,p)$ is of type $A'B$ when the probability $p = p(n)$ satisfies $p \ge 6((\log n)/n)^{1/3}$, and of type AB when $p \le (1-\varepsilon)/n$.*

Proof. The proof requires only a few basic results about random graphs, most of which can be found in the classical book of Bollobás [15].

A graph is *unicyclic* if it contains exactly one cycle. Trees are bipartite and thus type AB. Unicyclic graphs are either bipartite or a bipartite graph plus an edge – in either case Cut = Cut′, so that unicyclic graphs are also type AB. Thus the second statement of the Theorem follows immediately:

Theorem [cf. 15,V.7(i)]**.** *If $p = o(1/n)$, then almost every graph in $\mathcal{G}(n,p)$ is a forest.*

Corollary [cf. 15,V.8]**.** *Suppose $p = c/n$ with $0 < c < 1$. Then almost every graph in $\mathcal{G}(n,p)$ is such that every component is a tree or a unicyclic graph.*

To prove that almost all graphs in $\mathcal{G}(n,p)$ (for $p \ge 6((\log n)/n)^{1/3}$) are type $A'B$ it suffices to check (2.3.4)-(2.3.6). Theorem [cf. 15,II.8] implies that if $252n^{-1}\log n < p \le 1/2$, then almost every graph in $\mathcal{G}(n,p)$ has $|\mathrm{e}(G) - p\binom{n}{2}| < \sqrt{7pn^{-1}\log n}\binom{n}{2}$. Thus almost every graph with edge probability p has $|d-p| < (7p\log n/n)^{1/2}$ as long as $1 - (252\log n)/n > p > (252\log n)/n$. Therefore, since $p \ge 6(\log n/n)^{1/3}$,

$$\Big|\frac{d}{p} - 1\Big| = \frac{|d-p|}{p} < \sqrt{\frac{7\log n}{np}} \le \sqrt{7/6}\Big(\frac{\log n}{n}\Big)^{1/3} = o(1).$$

To establish (2.3.4) we will apply

Theorem [cf. 15,XI.22(ii)]**.** *If $2.27/n < p \le 1/2$, then almost every graph in $\mathcal{G}(n,p)$ has independence number at most $2\log(pn)/p$.*

For $p \le 1/2$ it suffices to show that $2(\log pn)/p \le p^2n(1-o(1))/100 (\le d^2n/100)$, or equivalently $200\log pn \le p^3n(1-o(1))$. Because $p \ge 6(\log n/n)^{1/3}$ the right-hand side eventually exceeds $200\log n$, and thus the left-hand side. For $p > 1/2$ we simply use the values for $p = 1/2$, since then $d \ge 1/2 - o(1)$, so that $d^2n/100 > n/500$ for almost every graph. Furthermore for almost every graph with $p \ge 1/2$ we get $\alpha(G) \le 4\log(n/2)$, which grows slower.

To prove (2.3.5) we need

Theorem [cf. 15,II.11]**.** *Let $0 < p \le 1/2$. Then almost every graph in $\mathcal{G}(n,p)$ is such that*

$$\big|\,|[S_1,S_2]| - p|S_1||S_2|\,\big| < (7p|S_1|^{-1}\log n)^{1/2}|S_1||S_2|,$$

whenever S_1, S_2 are disjoint sets of vertices satisfying $(252/p)\log n < |S_1| \le |S_2|$.

Note that for n sufficiently large, $d \ge \frac{5}{6}p \ge 5(\log n/n)^{1/3}$. Therefore

$$dn/10 \ge n^{2/3} > (252/6)(n\log^2 n)^{1/3} \ge (252\log n)/p$$

for almost every graph. So for $p \le 1/2$ we can apply Theorem II.11 to prove (2.3.5):

$$\begin{aligned}|[S_1,S_2]| &> p|S_1||S_2| - (7p\log n/n^{2/3})^{1/2}|S_1||S_2|\\ &\ge p|S_1||S_2|(1-o(1)) \ge d|S_1||S_2|(1-o(1)).\end{aligned}$$

For $1/2 < p \le 1 - 6(\log n/n)^{1/3}$ we can argue similarly with $\overline{G}$ and for other $p \ge 1 - o(1)$ we can interpolate again.

If $|S| \ge n^{2/3}$, then the statement that we just proved implies (2.3.6). Erdős and Rényi [45] proved that if $\varepsilon > 0$ and $\log n/n = \mathrm{o}(p)$ then for almost every graph in $\mathcal{G}(n,p)$, $\delta > (1-\varepsilon)pn$. Hence (2.3.6) also follows for $|S| < n^{2/3}$:

$$\begin{aligned}|[S,\overline{S}\,]| &\ge \delta(G)|S| - 2\binom{|S|}{2}\\ &> pn(1-\varepsilon)|S| - |S|n^{2/3} = pn|S|(1-\varepsilon-o(1)). \quad \square\end{aligned}$$

2.8. Complementary graphs

In this section we determine the maximum and minimum value that $\mathrm{cs}(G) + \mathrm{cs}(\overline{G})$ can take when G is an n-vertex graph. Results of this kind are frequently referred to as "Nordhaus-Gaddum type results".

Theorem 2.3.8. *For every n-vertex graph G*

$$2\binom{n}{2} - t(n,4) \le \mathrm{cs}(G) + \mathrm{cs}(\overline{G}) \le \mathrm{cs}(K_n),$$

and the bounds are best possible.

Proof. The upper bound is immediate, since every cut-cover for K_n yields a simultaneous cover for G and $\overline{G}$, and this is sharp for $G = K_n$. For the lower bound observe that by the proof of (2.3.2), for some maximum cuts $[S_1,\overline{S}_1\,]$ in G and $[S_2,\overline{S}_2\,]$ in $\overline{G}$

$$\begin{aligned}\mathrm{cs}(G) + \mathrm{cs}(\overline{G}) &\ge \big(2\mathrm{e}(G) - |[S_1,\overline{S}_1\,]|\big) + \big(2\mathrm{e}(\overline{G}) - |[S_2,\overline{S}_2\,]\big)\\ &= 2\mathrm{e}(K_n) - (|[S_1,\overline{S}_1\,]| + |[S_2,\overline{S}_2\,]|) \ge 2\binom{n}{2} - t(n,4),\end{aligned}$$

since the graph with edge-set $[S_1, \overline{S}_1] \cup [S_2, \overline{S}_2]$ is 4-colorable. To show that the lower bound in Theorem 2.3.8 is optimal we construct the following example. Take blow-ups of the self-complementary path on 4 vertices as follows: Partition the n vertices into 4 parts V_i, $1 \le i \le 4$, that are in size as equal as possible. Now let G be the subgraph of K_n formed by taking all edges in $[V_1, V_2]$, $[V_2, V_3]$, $[V_3, V_4]$ and all edges induced by V_1 and V_4 as indicated in Figure 2.2.

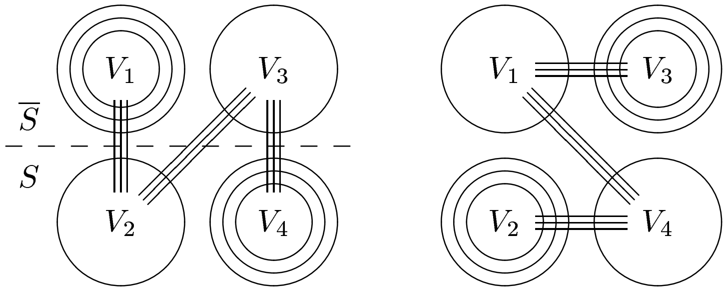


Figure 2.2: G and $\overline{G}$

For G we can take the cover that consists of stars centered at the vertices in $V_1 \cup V_4$ and the single cut $[V_1 \cup V_2, V_3 \cup V_4]$. Every edge in $[V_1 \cup V_3, V_2 \cup V_4]$ is covered once and every other edge twice. Since $\overline{G}$ has the same structure as G, we obtain

$$\begin{aligned} \mathrm{cs}(G) + \mathrm{cs}(\overline{G}) &\le (2\mathrm{e}(G) - |[V_1 \cup V_3, V_2 \cup V_4]|) + (2\mathrm{e}(\overline{G}) - |[V_1 \cup V_2, V_3 \cup V_4]|) \\ &= 2\binom{n}{2} - t(n, 4). \quad \square \end{aligned}$$

Remark. While the lower bound is achieved for a number of graphs and the upper bound gives the exact answer only for $G = K_n$, the latter is optimal in the following sense: for almost every graph in $\mathcal{G}(n, p)$, with fixed probability p and $q = 1 - p$, it can be seen from Theorem 2.3.7 and the facts that $\alpha(G) = \frac{2\log n}{\log(1/p)}(1+o(1))$ and that G is almost pn-regular, that

$$\mathrm{cs}(G) + \mathrm{cs}(\overline{G}) = n^2 - \left(\frac{2p}{\log(1/q)} + \frac{2q}{\log(1/p)}\right) n \log n(1 + o(1)).$$

2.9. More bounds and exact values

Proposition 2.3.9. *If G is a complete k-partite graph on n vertices, then*

$$\mathrm{cs}(K_{k-1} \vee \overline{K}_{n-k+1}) \le \mathrm{cs}(G) \le \mathrm{cs}(T_{n,k}). \tag{2.3.7}$$

Furthermore

$$\mathrm{cs}(K_{k-1} \vee \overline{K}_{n-k+1}) = (k-1)(n-1), \tag{2.3.8}$$

except that $\mathrm{cs}(K_4) = 8$ *and* $\mathrm{cs}(K_8) = 48$. *Also for all but finitely many pairs (n, k) with $n \ge k > 8$*

$$\mathrm{cs}(T_{n,k}) = 2t(n,k) - \lceil \frac{n}{k} \rceil \left(n - \lceil \frac{n}{k} \rceil\right). \tag{2.3.9}$$

However, $\mathrm{cs}(T_{n,4}) = 2t(n,4) - \lceil \frac{n}{2} \rceil \lfloor \frac{n}{2} \rfloor$.

Remark. Specifically, we will show that (2.3.9) holds if

- $k \leq 3$,
- $k > 200$, or
- $k > 8$ and $n > 2(k-1)^3$ or $k|n$.

Proof. For $k = 1$, $G = \overline{K}_n$ and all statements are true, so that we can assume $k > 1$. Let G be a complete k-partite graph with part sizes $n_1 \leq n_2 \leq \cdots \leq n_k$. By 2.4.1, we can assume that the cuts in an optimal cut-cover will not cut through any of the partite sets. Thus for every cut $[S_i, \overline{S}_i]$ in the cover we get that

$$|S_i| = n_{i_1} + n_{i_2} + \cdots + n_{i_j}, \qquad |[S_i, \overline{S}_i]| = |S_i|(n - |S_i|),$$

so that the size of the cover can be viewed as a function in the k variables n_i:

$$|\mathcal{C}| = \sum |[S_i, \overline{S}_i]| = f(n_1, n_2, \ldots, n_k).$$

If we keep all but two of the coordinates $i < j$ fixed, then by combining like terms we obtain

$$f(n_1, n_2, \ldots, n_k) = an_in_j + bn_i + cn_j + d,$$

for some $a, b, c, d \geq 0$. Furthermore, $a > 0$ (since the edges between the two parts need to be covered at least once) and $b \geq c$ (since otherwise we could swap the roles of n_i and n_j in the cover and not increase its total size). When $n_i > 1$, we can decrease n_i by one and increase n_j by one, thus defining another complete k-partite graph G' on n vertices. The same cover as the one defined by f now shows that

$$\begin{aligned} cs(G') &\leq f(n_1, n_2, \ldots, n_i - 1, \ldots, n_j + 1, \ldots, n_k) \\ &= a(n_i - 1)(n_j + 1) + b(n_i - 1) + c(n_j + 1) + d \\ &= [an_in_j + bn_i + cn_j + d] + a(n_i - n_j - 1) + (c - b) < cs(G). \end{aligned}$$

With this observation, (2.3.7) follows and we observe that the inequalities are strict, except when $G = T_{n,k}$ or $G = K_{k-1} \vee \overline{K}_{n-k+1}$.

To compute the exact values, note that for $k = n$ the values were given in Theorem 2.5.1, so we can assume that $n > k > 1$. Clearly for $G = K_{k-1} \vee \overline{K}_{n-k+1}$ we get that $\mathrm{Cut}'(G) = \max\{(k-1)(n-k+1), n-1\} = (k-1)(n-k+1)$, so that $\mathrm{cs}(G) \leq (k-1)(n-1)$. Furthermore, K_k is a subgraph of G so that for $k \neq 4, 8$,

$$\mathrm{cs}(G) \geq \mathrm{cs}(K_k) + \mathrm{e}(G) - \mathrm{e}(K_k) = (k-1)(n-1). \tag{2.9.1}$$

For $k = 4$ we get that $K_3 \vee \overline{K}_{n-3}$ is type B from Proposition 2.3.4, since $n - 3 \geq 2 = \Delta(K_3)$. For $k = 8$ the same proof as in (2.9.1) yields a lower bound that is only off by one:

$$\mathrm{cs}(K_7 \vee \overline{K}_{n-7}) \geq \mathrm{cs}(K_8) + \mathrm{e}(K_7 \vee \overline{K}_{n-7}) - \mathrm{e}(K_8) = 7(n-1) - 1. \tag{2.9.2}$$

If this inequality were sharp, then (in a given optimal cover) every induced K_8 must be covered optimally, that is by 3 $K_{4,4}$'s. In this covering of K_8 the edges involving a fixed vertex are covered a total of 12 times, so that it follows that

$$\text{(2.9.3)} \qquad \mathrm{cs}(K_7 \vee \overline{K}_{n-7}) = \mathrm{cs}(K_8) + 12(n-8) = 12n - 48.$$

Combining (2.9.2) and (2.9.3) we get $12n - 48 = 7n - 8$, or $n = 8$.

The upper bound for (2.3.9) is given by (2.3.3), since $\mathrm{Cut}'(T_{n,k}) = \lceil n/k \rceil (n - \lceil n/k \rceil)$. For $k \leq 4$, $T_{n,k}$ is 4-colorable, so that $\mathrm{cs}(T_{n,k})$ can be easily computed from Theorem 2.3.5. If $k > 200$, then Theorem 2.3.6 applies. For the case $k > 8$, observe again that we can assume that the cuts in an optimal cut-cover will not cut through any of the partite sets – thus we essentially have a cut-cover of a weighted K_k where edge ij has weight $n_i n_j$. The cover achieving (2.3.9) corresponds to the cover of K_k with stars, which is the only cover achieving $\mathrm{cs}(K_k)$. Thus every other cover of K_k has total size at least $(k-1)^2 + 1$, so to show (2.3.9) it suffices to check that

$$\text{(2.9.4)} \qquad 2t(n,k) - \lceil \frac{n}{k} \rceil \left(n - \lceil \frac{n}{k} \rceil \right) \leq ((k-1)^2 + 1) \lfloor n/k \rfloor^2.$$

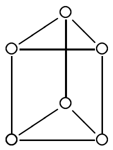
If k divides n, or $n > 2(k-1)^3$ then this is indeed the case. More detailed computations are possible to provide more cases of equality. $\square$

Observe that it suffices to compute all relevant parameters on the blocks of G, since if $G = G_1 \cup G_2$ with $|V(G_1) \cap V(G_2)| \leq 1$, then $\mathrm{cs}(G) = \mathrm{cs}(G_1) + \mathrm{cs}(G_2)$ and the same holds for $\mathrm{e}(G)$, $\mathrm{Cut}(G)$ and $\mathrm{Cut}'(G)$. Furthermore we will define $G_1 \leftrightarrow G_2$ to be any graph obtained by identifying G_1 and G_2 at any one of their vertices. By the previous remark $G_1 \leftrightarrow G_2$ is type A (B) exactly when G_1 and G_2 are both of type A (B).

To be able to construct many graphs whose cut-cover size is easily computed we define the cartesian product $G \square H$ of the graphs G and H to be the graph with vertex set

$$V(G \square H) = V(G) \times V(H) = \{(v,w) : v \in V(G), w \in V(H)\},$$

and (v_1, w_1) adjacent to (v_2, w_2) if $v_1 = v_2$ and w_1 adjacent to w_2, or if $w_1 = w_2$ and v_1 adjacent to v_2 (see Figure 2.3).



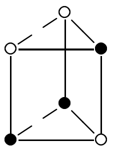


Fig. 2.3: $K_3 \square K_2$ Fig. 2.4: $\mathrm{Cut}(K_3 \square K_2) = 7$

The following facts summarize all relevant properties of the cartesian product:

$$\text{(2.9.5)} \qquad n(G \square H) = n(G) n(H)$$

$$\text{(2.9.6)} \qquad \mathrm{e}(G \square H) = \mathrm{e}(G) n(H) + n(G) \mathrm{e}(H),$$

$$\text{(2.9.7)} \qquad \chi(G \square H) = \max\{\chi(G), \chi(H)\},$$

$$\text{(2.9.8)} \qquad \mathrm{Cut}(G \square H) = \mathrm{Cut}(G) n(H) + n(G) \mathrm{Cut}(H),$$

$$\text{(2.9.9)} \qquad \mathrm{Cut}'(G \square H) \leq \mathrm{Cut}'(G) n(H) + n(G) \mathrm{Cut}'(H),$$

$$\text{(2.9.10)} \qquad \mathrm{cs}(G \square H) = \mathrm{cs}(G) n(H) + n(G) \mathrm{cs}(H),$$

(2.9.11) $G \square H$ is type A exactly when G and H are type A,

(2.9.12) If $G \square H$ is type B then G and H are type B.

For example to prove (2.9.10) we observe that $G\Box H$ contains $n(H)$ copies of G and each one of them contributes at least $\mathrm{cs}(G)$ towards the total sum in an optimal cut-cover. Combining this with a similar argument for the copies of H, we get that the left-hand side is no smaller than the right-hand side. A cover achieving this can be obtained from optimal covers of G and H, by putting all copies of a given vertex in $V(G)$ in the same partition as in the original cut of G (similarly for the cuts from H). The other proofs are similar and will be omitted.

Note that K_2 and K_3 are both type AB, but $K_3\Box K_2$ is type AB', since the cut illustrated in Figure 2.4 is, up to isomorphism, the only cut of size at least 7.

To see that Theorem 2.3.5 is indeed best possible we define, for a given type T,

$$\chi(T) = \min\{\chi(G) : G \text{ is type } T\}.$$

Hence $\chi(AB) = 1$, $\chi(AB') = 3$ (achieved by $K_3\Box K_2$), $\chi(A'B) = 5$ (achieved by K_5) and $\chi(A'B') = 5$ (achieved by $K_5\Box K_2$). We will show that for all types T there are infinitely many graphs G with $\chi(G) = k \geq \chi(T)$ in a strong sense:

Proposition 2.9.1. *For every type T and graph G with $\chi(G) \geq \chi(T) - 1$, there exists a graph G' of type T containing G as an induced subgraph so that $\chi(G') = \chi(G) + 1$ and $n(G') \leq n(G) + \Delta(G) + \chi(T)$.*

Proof. For $T = AB$, we can choose $G' = G \vee \overline{K}_{\Delta+1}$. For $T = A'B$ we let $G' = (G \vee \overline{K}_{\Delta}) \leftrightarrow K_5$. If $T = AB'$ and $\chi(G) \geq 3$, then we let $G' = (G \vee \overline{K}_{\Delta}) \leftrightarrow K_4$. If $T = AB'$ and $\chi(G) = 2$ then we let $G' = G \leftrightarrow (K_3\Box K_2)$, which works if $\Delta(G) \geq 2$ – otherwise $\Delta(G) = 1$ and G is a matching plus isolated vertices so that we can extend one of the matching edges to a $K_3\Box K_2$. Finally, in the case $T = A'B'$ we note that $\Delta(G) \geq \chi(G)$ unless G contains a $K_{\Delta+1}$. When $\chi(G) = 4$ we can let $G' = G \leftrightarrow (K_5\Box K_2)$, except if $\Delta(G) \leq 3$. In this case G must have a component that is a K_4 and we just extend this component to a $K_5\Box K_2$. For $\chi(G) > 4$ we can let $G' = (G \vee K_1) \leftrightarrow (K_5\Box K_2)$, except if $\Delta(G) = 4$. In this case G must have a component that is a K_5 and we just extend this component by a vertex v to a K_6 and let $G' = (G + v) \leftrightarrow (K_3\Box K_2)$. $\Box$

2.10. Cut-covers and L_∞-representations

The cut-cover problem can also be viewed as a geometric problem. This connection had already been observed in [77]. Indeed, there is a correspondence between the cut-cover problem and the following isoperimetric problem proposed by Ahlswede and Katona [4]:

Find n points in the hypercube $Q_k(= \{0,1\}^k)$ such that their average pairwise graph distance (Hamming distance) is minimal.

As noted before, the encoding of vertices as binary k-vectors obtained from a cover by k cuts has the property that the number of times any edge uv is covered is exactly the number of coordinates in which u and v differ, i.e. their Hamming distance. So the total size of a cover of K_n is nothing but the sum of the pairwise distances between the vertices in the corresponding n-vertex subset of Q_k. Thus the problem of Ahlswede and Katona is equivalent to determining $f_k(n)/\binom{n}{2}$, where $f_k(n)$ is the minimum total size of a cover of the edges of K_n with at most k cuts.

Recent progress on this problem is described in Ahlswede and Althöfer [2], where $f_k(n)$ is asymptotically determined if $n = c^k$, and Althöfer and Sillke [7], where a general upper bound on $f_k(n)$ is derived that shows that the only optimal configuration in the case $n = 2^{k-1}$ is to take all points which have value ε in coordinate j, for some $\varepsilon \in \{0,1\}$ and some $1 \leq j \leq k$. Since then Fu [53] found a simpler proof of a bound that is slightly stronger than that in [7].

Notice that, with the notation from Section 2.5, $f(n) = \min\{f_k(n) : k \geq \lceil \lg n \rceil\}$. We restrict the value of k in the definition of $f(n)$ to $k \geq \lceil \lg n \rceil$ since there are no covers with fewer cuts — equivalently, Q_k has only $2^k \geq n$ points. So Theorem 2.5.1 and its proof show that for $k \geq n-1$, $f_k(n) = f_{n-1}(n) = f(n)$. Apart from the result of Althöfer and Sillke these are the only known nontrivial *exact* solutions for this isoperimetric problem. The proof also gives us, besides the value of $f(n)$, a description of the point sets achieving the optimum. For $n > 7$ this is unique up to isomorphism.

In general the exact solution of this problem may be hopeless, although there is a very nice conjecture suggested by Erickson and Harris [49] as to what an extremal configuration may look like: For given k and n determine a suitable point $x \in [0,1]^k$, not necessarily with integer coordinates, and a radius $0 < r < k$. The configuration is a *discrete ball* centered at x with radius r: take all the points in $\{0,1\}^k$ that are of distance at most r from x. This construction includes the previously known good constructions, like Hamming balls centered at $(0,0,\ldots,0)$ or sub-cubes (for example the Althöfer-Sillke cube has $x = (0, 1/2, 1/2, \ldots, 1/2)$ and $r = (k-1)/2$), as well as some new constructions which improve on both of these (they are essentially products of cubes and spheres). For recent progress on this problem see [89].

Even for general graphs we can use the equivalent labeling to define, and solve, a geometric graph representation problem:

For $x \in \mathbb{R}^d$ we define $\|x\|_\infty = \max\{|x_i| : 1 \leq i \leq d\}$ and $\|x\|_1 = \sum_{1 \leq i \leq d} |x_i|$. An L_∞-representation (in $\mathbb{R}^d$) of a graph G is an assignment

$$f : V(G) \to \mathbb{R}^d,$$

such that $\|f(u) - f(v)\|_\infty \geq 1$ whenever $uv \in E(G)$. There is no restriction if $uv \notin E(G)$. For a given L_∞-representation (G, f) we define

$$L_1(G, f) := \sum_{uv \in E(G)} \|f(u) - f(v)\|_1.$$

Therefore the average L_1-distance between adjacent vertices in the L_∞-representation is $L_1(G,f)/\mathrm{e}(G)$.

Theorem 2.10.1.

$$\mathrm{cs}(G) = \inf\{L_1(G,f) : f \textit{ is an } L_\infty\textit{-representation of } G\}$$

Proof. The $\{0,1\}$-labeling f associated with any cut-cover is an L_∞-representation of G and $L_1(G,f)$ counts the total size of the cut-cover. Thus $\mathrm{cs}(G) \geq \inf\{L_1(G,f) :$ f is an L_∞-representation of $G\}$.

For the reverse inequality it suffices to show that for every L_∞-representation f we can find a $\{0,1\}$-representation f^*, maybe higher-dimensional, such that $L_1(G,f^*) \leq L_1(G,f)$.

We denote the value of the i-th coordinate in $f(v)$ by $f(v)_i$. Among L_∞-representations let f' be one with a maximum number of integer coordinates, subject to $L_1(G, f') \leq L_1(G, f)$. We claim that all $f'(v)_i \in \mathbb{Z}$. Indeed, let $U_c(i) = \{v \in V : f'(v)_i = \lfloor f'(v)_i \rfloor + c\}$ and suppose that $U_c(i) \neq \emptyset$ for some $1 \leq i \leq d$ and $c > 0$. We can partition $V = \bigcup U_{c_j}(i)$ with $0 \leq c_1 < c_2 \cdots < c_k < 1$, and define $x^+ := 1 - c_k$ and $x^- := c_{k-1} - c_k$ where we set $c_0 = 0$. Now f^x defined by

$$f^x(u)_j = \begin{cases} f'(u)_j + x & j = i, u \in U_{c_k}(i) \\ f'(u)_j & \text{otherwise} \end{cases}$$

is an L_∞-representation as long as $x^- \leq x \leq x^+$. Furthermore if we let

$$\mathrm{e}_1 = |\{uv \in E(G) : u \in U_{c_k}, v \notin U_{c_k}, f(u)_i > f(v)_i\}|,$$
$$\mathrm{e}_2 = |\{uv \in E(G) : u \in U_{c_k}, v \notin U_{c_k}, f(u)_i < f(v)_i\}|,$$

then $L_1(G, f^x) = L_1(G, f') + \mathrm{e}_1 x - \mathrm{e}_2 x$. Hence $L_1(G, f^x) \leq L_1(G, f')$ for either $x = x^+$ or $x = x^-$. But if $x = x^+$ or $k = 1$ then we increased the number of integer coordinates. However if $x = x^-$ and $k > 1$ then we decreased the number of sets in our partition, and by iterating this process we must eventually be in the previous case.

Now we are in the position to obtain f^*. We can assume that, for all $1 \leq i \leq d$, $\min\{f'(v)_i : v \in V\} = 0$ since we can shift the coordinates appropriately. Set $k = \max\{f'(v)_i : v \in V, 1 \leq i \leq d\}$ and for $0 \leq j \leq k$ let $s(j)$ be the k-dimensional vector whose first j coordinates are 1, all other coordinates 0. Define

$$f^* : V \to \{0,1\}^{kd}, v \mapsto (s(f'(v)_1), s(f'(v)_2), \ldots s(f'(v)_d))\,.$$

If $\|f^*(v) - f^*(u)\|_\infty < 1$, then $f^*(v) = f^*(u)$, so that already $f'(u) = f'(v)$ and thus f^* is an L_∞-representation. Finally $\|f^*(v) - f^*(u)\|_1 = \sum |f'(v)_i - f'(u)_i| = \|f'(v) - f'(u)\|_1$, so that $L_1(G, f^*) = L_1(G, f') \leq L_1(G, f)$. □

2.11. Open questions

Since the investigation of $\mathrm{cs}(G)$ has just started, there are still many basic questions that need to be answered – a few of which we will mention.

Question 2.11.1. Can the bounds of Theorem 2.3.2 be improved? Find improvements for special classes of graphs, like triangle-free graphs.

Problem 2.11.2. Determine $\max\{\mathrm{cs}(G) : \mathrm{e}(G) = e\}$ and $\min\{\mathrm{Cut}'(G) : \mathrm{e}(G) = e\}$.

Question 2.11.3. For any given probability function $p(n)$: Is almost every graph in $\mathcal{G}(n, p(n))$ type B?

Question 2.11.4. Is there a threshold $t(n)$ such that almost every graph in $\mathcal{G}(n, p(n))$ is type A if $p(n) < t(n)$ and type A' if $p(n) > t(n)$? Determine $t(n)$.

Problem 2.11.5. Determine $\mathrm{cs}(T_{n,k})$.

Remark 2.11.6. We can define L_q-representations for G, $L_p(G, f)$ and

$$L_{p,q}(G) = \inf\{L_p(G, f) : f \text{ is an } L_q\text{-representation of } G\}$$

for all parameters $1 \leq p, q \leq \infty$. For $p < p'$ we have $L_{p,q}(G) \geq L_{p',q}(G)$ and $L_{q,p}(G) \leq L_{q,p'}(G)$ for all q, since $\|x\|_p \geq \|x\|_{p'}$ for all $x \in \mathbb{R}^d$. Furthermore, $L_{p,q}(G) = 0$ when $p > q$ and $L_{p,p}(G) = \mathrm{e}(G)$.

Problem 2.11.7. Investigate $L_{p,q}(G)$.

The *depth* of an edge in a cycle cover is the number of times the edge is covered and the *edge-depth* of a cover is the maximum depth of all the edges in $E(G)$. A graph has a cycle double cover if and only if it has a cycle cover of edge-depth at most 2, since the edges that are covered only once form an even graph. Zhang [121] conjectures that every 2-edge-connected graph has a shortest cycle-cover of edge depth at most 3, but currently it is not even known if there is a constant c such that every 2-edge-connected graph has a shortest cycle-cover of edge depth at most c. Zhang gives a suitable subdivision of the Peterson graph to show that $c \geq 3$ is needed, but conjectures that if a graph is 3-edge-connected then $c = 2$ is sufficient. Every 4-edge-connected graph has a shortest cycle-cover of edge-depth at most 2, since these graphs admit a nowhere-zero 4-flow.

We can ask the same question for the cut-cover problem

Question 2.11.8. Is there a constant c such that for every graph there is an optimal cut-cover so that no edge is covered more than c times?

For Question 2.11.8, $c \geq 3$ since every optimal cover of K_8 contains an edge that is covered 3 times. We are not aware of any graph for $c = 4$ – obviously such a graph would have to be type $A'B'$. In this direction we have

Proposition 2.11.9. *For every $e \in E(G)$*

$$1 \leq \mathrm{cs}(G) - \mathrm{cs}(G - e) \leq n - 1.$$

Proof. $\mathrm{cs}(G - e) \leq \mathrm{cs}(G) - 1$ since every cover for G also covers $G - e$, and this is best possible since, for example, cut-edges are covered only once in every optimal covering.

To show $\mathrm{cs}(G) \leq \mathrm{cs}(G - e) + n - 1$ we note that every optimal cover has at most $n - 1$ cuts. Indeed in every cut there is at least one edge covered only in this cut, and taking one such edge from every cut must result in an acyclic graph (since no cycle crosses a cut just once) – but acyclic graphs have at most $n - 1$ edges. Thus we can take any optimal cover for $\mathrm{cs}(G - e)$ and use it as a cover for G. If it does cover e it can do so at most $n - 1$ times, and if it doesn't then we add one more cut: a star centered at an endpoint of e. Note that this upper bound is also best possible, since for $n \neq 4, 8$

$$\mathrm{cs}(K_n) = (n - 1)^2 = \mathrm{cs}(K_{n-2} \vee \overline{K}_2) + (n - 1). \quad \square$$

Remark 2.11.10. This does not settle Question 2.11.8, since the sharp drop in $\mathrm{cs}(G)$ can result from having a different cover (i.e. one fewer star). Probably Proposition 2.11.9 can be sharpened:

Question 2.11.11. Is $\mathrm{cs}(G) \leq \mathrm{cs}(G - e) + \Delta(G)$?

Question 2.11.12. Is $\mathrm{cs}(G) \leq \mathrm{cs}(G - uv) + \min\{d(u), d(v)\}$?

Both questions can be answered in the affirmative if $G - e$ is type B.

3. Guarding art galleries

3.1. The basic art gallery problem

The original art gallery problem, posed by Klee and solved by Chvátal [32], is to find the smallest number of guards necessary to cover any simple polygon, the art gallery, not necessarily convex, on n vertices. Here a covering by g guards means that one can find g points in the interior of the polygon such that every point in the interior is covered by some guard, that is for each point in the interior the line segment between it and some guard does not intersect the polygon. The *comb polygons* in Figure 3.1 show that $\lfloor n/3 \rfloor$ guards are sometimes necessary – if n is not divisible by 3 simply take a comb on $3\lfloor n/3 \rfloor$ vertices and subdivide one or two of its edges. Chvátal also showed that $\lfloor n/3 \rfloor$ guards always suffice. For more information on the history of this problem and related problems, see [98] and [107]. The results in this chapter are the content of [88].

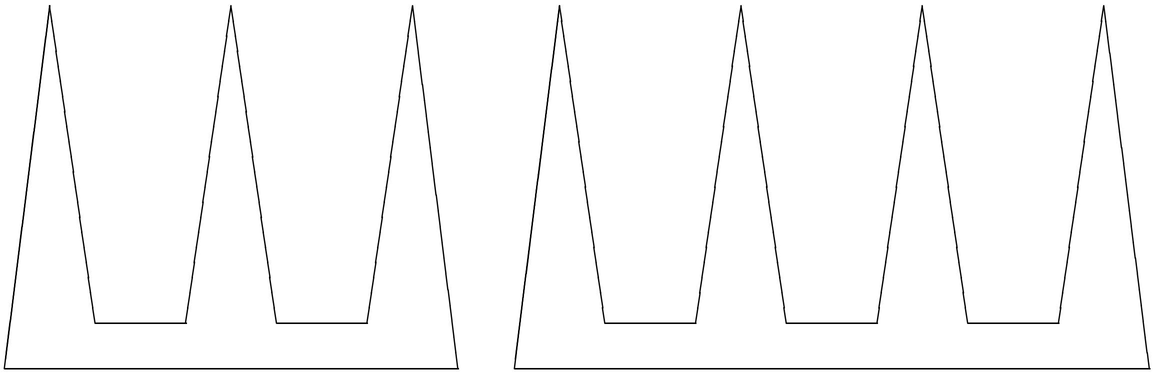


Figure 3.1: Comb polygons

3.2. Art galleries with interior walls

Hutchinson [72] generalized the basic art gallery problem by allowing interior walls. Throughout this chapter an *art gallery (with interior walls)* will be a simple polygon on n vertices with some pairs of vertices joined by nonintersecting interior diagonals, the *interior walls*. Also suppose that in the interior of each of the walls there is an arbitrarily placed, arbitrarily small opening, the *doorway*. Figure 3.2 is an example of an art gallery on $n = 15$ vertices and 8 interior walls that requires 9 guards. Hutchinson now asked to find the minimum number of guards that suffice to cover any such art gallery on n vertices.

To motivate our proofs we will give Fisk's [51] elegant proof of Chvátal's result:

First triangulate the polygon. The resulting plane graph, with vertex set the corner points of the polygon, has all its vertices on the outside face. Graphs that can be embedded in the plane in such a way are called *outerplanar*. It is well known that outerplanar graphs are 3-colorable, which can be easily seen by cutting along a chord and applying induction. Since each triangle in the triangulation must have vertices of all three colors, putting a guard at each vertex in the smallest color class produces a covering set of $\lfloor n/3 \rfloor$ guards.

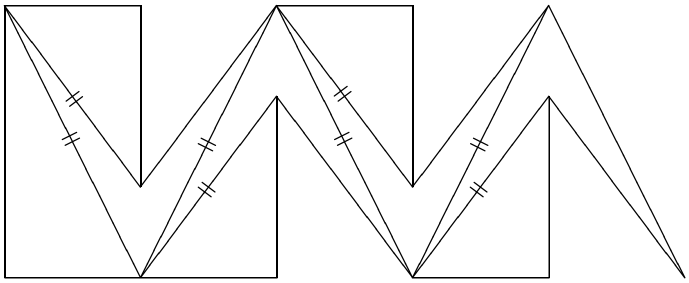


Figure 3.2: Art gallery with interior walls

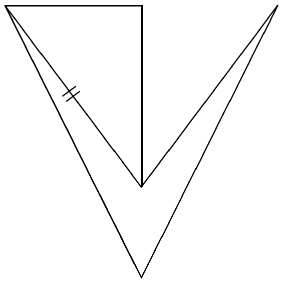


Figure 3.3: V

We will now answer Hutchinson's question with an argument in the spirit of Fisk's proof (see also [87]).

Lemma 3.2.1. $\lfloor (2n-3)/3 \rfloor$ *guards suffice to cover any art gallery on* n *vertices and there are galleries with* $\lfloor 2n/3 \rfloor - 2$ *interior walls where this many guards are required.*

Proof. We may assume that the interior walls triangulate the art gallery, since adding extra interior walls cannot make it easier to guard the gallery. This outerplanar graph can now be 3-colored. From such a coloring, we get a labeling of the edges of the graph by assigning to each edge the color *not used* on its endpoints. Now each triangle has each color appearing on one of its incident edges. Placing a guard into the doorway, for interior walls, or just next to the wall, for exterior walls, we can see that each set of labels corresponds to a set of guards that covers the whole art gallery. Since n-vertex outerplanar triangulations have $2n-3$ edges, which can be seen by induction or by Euler's formula, taking the least frequent color suffices.

Art galleries of the type in Figure 3.2 achieve this bound. They are obtained by starting with a small gallery, and attaching $k = \lfloor n/3 \rfloor - 1$ of the V-shaped galleries in Figure 3.3. If we attach a V with the leftmost wall to an already existing art gallery and put the doorway exactly in the center we increase the number of vertices by 3, the number of interior walls by 2 and the number of guards needed by 2. Note that even if a guard from the smaller gallery can be placed in the doorway of the interior wall connecting the V to the smaller gallery that guard still can not see the other branch of the V or the triangle, so that those will require one additional guard each.

When $n = 3k+3$ our starting gallery is a triangle (as in Figure 3.2), when $n = 3k+4$, it is any quadrilateral and when $n = 3k+5$, a V. This produces art galleries with $\lfloor 2n/3 \rfloor - 2$ (that is $2k$, $2k$ and $2k+1$ respectively) interior walls. The galleries also require exactly one more guard than they have interior walls. $\square$

This settles the problem when the number of interior walls is unspecified. But what happens when we have a specified number of interior walls, say m? This question was suggested by J.Griggs. If $m \geq \lfloor 2n/3 \rfloor - 2$, then Lemma 3.2.1 shows that the answer is still $\lfloor (2n-3)/3 \rfloor$, since adding additional interior walls in the art galleries provided does not make guarding any easier.

Theorem 3.2.2. *The minimum number of guards that suffice to cover all art galleries* *with n vertices and m interior walls, $g(n,m)$, is*

$$\min\left\{\left\lfloor\frac{2n-3}{3}\right\rfloor, \left\lfloor\frac{2m+n}{3}\right\rfloor, \left\lfloor\frac{2n+m-2}{4}\right\rfloor\right\},$$

or more precisely:

$$g(n,m) = \begin{cases} \lfloor\frac{2n-3}{3}\rfloor, & \text{for } m \geq \lfloor\frac{2}{3}n\rfloor - 2 \\ \lfloor\frac{2m+n}{3}\rfloor, & \text{for } m < \lfloor\frac{2}{5}n\rfloor \\ \lfloor\frac{2n+m-2}{4}\rfloor, & \text{otherwise.} \end{cases}$$

Lemma 3.2.1 proves the first part of the statement, and the other two parts will be Lemmas 3.2.3 and 3.2.4.

Lemma 3.2.3. *Always $g(n,m) \leq \lfloor(2m+n)/3\rfloor$, and for $m < \lfloor 2n/5\rfloor$ there are galleries with m interior walls where this many guards are required.*

Proof. The bound can be easily established by induction, but there is also a Fisk type argument. Before triangulating the art gallery, assign each vertex v a weight of $d(v)-1$, where $d(v)$ is the degree of v in the outerplanar graph determined by the gallery – i.e. the number of walls meeting at v. For example in the art gallery without interior walls each vertex has weight 1. Now triangulate the gallery, also using the interior walls that are already present, 3-color the triangulation, and find the color class of smallest total weight W. It will suffice to find a guard set with W guards, since the total weight on all vertices is

$$\sum_{v\in V(G)} (d(v)-1) = 2|E(G)| - n = 2m+n.$$

To do this simply put $d(v)-1$ guards at each vertex v in the color class of total weight W by putting one guard on every interior angular bisector of walls that meet v, as shown in Figure 3.4, close to v. Note that when we place the guards we ignore the chords that are not walls and were introduced in the triangulation step. Since every triangle has a vertex of our chosen color, and each triangle is covered by a guard at the vertex associated with it, we are done.

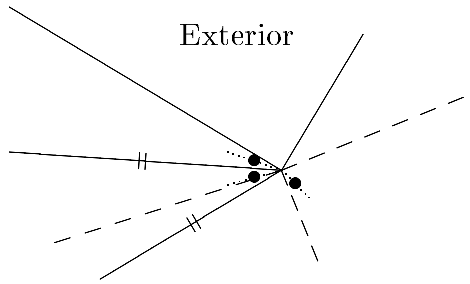


Figure 3.4: Guard placement

Call the art gallery in Figure 3.5 an E. Notice that an E has 7 vertices, 1 interior wall and requires 3 guards since no guard can cover more than one of the 3 alcoves. If we attach an E with the vertical wall to an already existing art gallery and put the doorway exactly in the center we increase the number of vertices by 5, the number of interior walls by 2 and the number of guards necessary by 3. We will put the doorway in the middle of the vertical wall, so that a guard there can't cover either alcove.

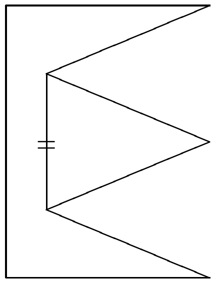


Figure 3.5: E

To achieve the bound, let $m = 2k + \epsilon$ with $\epsilon \in \{0, 1\}$.

Start the construction with a comb on $n - 5k - \epsilon$ vertices. Since $n - 5k - \epsilon \geq 3 + \epsilon$, such a comb exists. If m is odd, attach one triangular room to a wall parallel to the long horizontal wall in the comb. This adds 1 vertex, 1 interior wall and 1 guard, since the guard that is needed to cover the triangle cannot cover any prong of the comb. (If $n - 5k - 1 \in \{4, 5\}$, then start with a V and subdivide any wall if necessary). In either case add k E's, starting from an end of the comb. The resulting art gallery has n vertices and m interior walls. The number of guards required is

for $\underline{m = 2k}$:

$$\left\lfloor \frac{n - 5k}{3} \right\rfloor + 3k = \left\lfloor \frac{n + 4k}{3} \right\rfloor = \left\lfloor \frac{2m + n}{3} \right\rfloor.$$

for $\underline{m = 2k + 1}$:

$$\left\lfloor \frac{n - 5k - 1}{3} \right\rfloor + 1 + 3k = \left\lfloor \frac{n + 4k + 2}{3} \right\rfloor = \left\lfloor \frac{2m + n}{3} \right\rfloor. \quad \square$$

Lemma 3.2.4. *Always $g(n, m) \leq \lfloor (2n + m - 2)/4 \rfloor$, and for $\lfloor (2n - 4)/5 \rfloor \leq m \leq \lfloor (2n - 5)/3 \rfloor$ there are galleries with m interior walls where this many guards are required.*

Proof. For the construction let $k = \lfloor (2n - 3m - 5)/4 \rfloor \geq 0$. Start with the art gallery with $n - 5k \geq 3$ vertices and $m - 2k$ interior walls constructed in Lemma 3.2.1. It is a straightforward computation to check that indeed $\lfloor 2(n - 5k)/3 \rfloor - 2 \leq m - 2k$ (use the fact that for every integer x, $x \geq \lfloor y \rfloor \iff x + 1 > y$), so that the construction from Lemma 3.2.1 applies. Now add k E's. The number of guards needed for this art gallery is

$$\begin{aligned} 3k + \lfloor \frac{2(n - 5k) - 3}{3} \rfloor &= \lfloor \frac{2n - \lfloor (2n - 3m - 5)/4 \rfloor - 3}{3} \rfloor \\ &\geq \lfloor \frac{8n - 2n + 3m + 5 - 12}{12} \rfloor = \lfloor \frac{2n + m - 2}{4} \rfloor, \end{aligned}$$

unless $2n+m-2$ is divisible by 4. But in that case our inequality was not sharp and we can gain an additional $1/12$ from that term.

We now prove the upper bound by induction on n, with the added requirement that when $2n+m-2$ is divisible by 4 we can place one of the guards arbitrarily at an exterior wall.

The bound holds for $m=0$, since $g(n,0)=\lfloor n/3 \rfloor \le \lfloor (n-1)/2 \rfloor$ for $n \ge 3$. This is the only feasible value for the base case $n=3$. The added requirement also holds when $n=3$: we can place the guard where we want since triangles are convex.

Consider an art gallery A with $n>3$ vertices and $m>0$ walls, and let $g(A)$ be the number of guards needed to cover A. Cut the art gallery along an interior wall xy (see Figure 3.6). This splits the gallery into two parts sharing two vertices. For $i \in \{1,2\}$, let n_i and m_i be the number of vertices and interior walls in the i-th part. Hence $n_1+n_2=n+2$ and $m_1+m_2=m-1$, so that

$$\begin{aligned} g(A) &\le g(n_1,m_1)+g(n_2,m_2) \le \left\lfloor \frac{2n_1+m_1-2}{4} \right\rfloor + \left\lfloor \frac{2n_2+m_2-2}{4} \right\rfloor \\ &\le \frac{2(n_1+n_2)+(m_1+m_2)-4}{4} = \frac{2n+m-1}{4}. \end{aligned}$$

Therefore $g(A) \le \lfloor (2n+m-2)/4 \rfloor$, unless equality holds everywhere, which implies that $(2n_1+m_1-2)$ and $(2n_2+m_2-2)$ are divisible by 4. However in that case we can invoke the stronger hypothesis and require guards to be on either side of the separating wall xy right next to the doorway. Now replacing these two guards by a single guard in the doorway yields the claim.

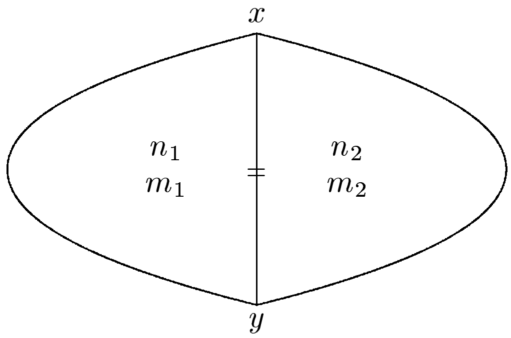


Figure 3.6

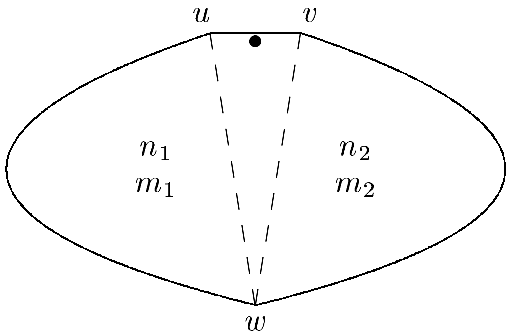


Figure 3.7

To show the additional requirement suppose that $2n+m-2$ is divisible by 4 and that a guard must be placed near exterior wall uv. Triangulate the gallery, also using the interior walls that are already present, and let w be the third vertex in the triangle containing uv. Removing the triangle uvw, which is already covered, splits the gallery into two parts (see Figure 3.7), satisfying $n_1+n_2=n+1$ and $m_1+m_2 \le m$. Thus

$$\begin{aligned} g(A) &\le g(n_1,m_1)+g(n_2,m_2)+1 \le \left\lfloor \frac{2n_1+m_1-2}{4} \right\rfloor + \left\lfloor \frac{2n_2+m_2-2}{4} \right\rfloor + 1 \\ &\le \frac{2(n_1+n_2)+(m_1+m_2)}{4} \le \frac{2n+m+2}{4}. \end{aligned}$$

This also works in the degenerate case when one of the $n_i = 2$. So we are done unless equality holds everywhere, since $2n+m-2$ is divisible by 4. This implies that $(2n_1+m_1-2)$ and $(2n_2+m_2-2)$ are divisible by 4. Also $m_1+m_2 = m$ so that neither one of the chords uw or vw is an interior wall. Again we invoke the stronger induction hypothesis and require guards near the chords uw and vw close to w. Replacing these two guards by a single guard right at w finishes the proof. □

3.3. Art galleries with convex rooms

In [33] Czyzowicz et al. study art galleries that consist of polygons on n vertices that are subdivided, not necessarily along chords, into k *convex* regions and show that these can be covered with $\lfloor 2(n+k)/3 \rfloor$ guards. This result is independent of Theorem 3.2.2, since our problem allows rooms of arbitrary shape but requires the interior walls to be chords.

It would be a common special case to study art galleries with n vertices and m interior walls such that all $k = m+1$ interior rooms are convex. Notice that we are not requiring the polygon itself to be convex. However doing so does not change the answer, since our construction achieving the upper bound can easily be built to meet this additional requirement. Since it will pose no additional difficulty, we will also require each room to have at least $r \geq 3$ walls, with $r = 3$ being the general case.

Summing the sizes of the rooms yields $n+2m$, since the interior walls are counted twice. Hence for art galleries such that all rooms have size at least r, we have $n+2m \geq (m+1)r$, or equivalently $n \geq m(r-2)+r$.

Theorem 3.3.1. *The minimum number of guards that suffice to cover all art galleries with $m > 0$ interior walls and $n \geq m(r-2)+r$ vertices, such that all rooms are convex with at least r walls, $g_r^*(n,m)$, is*

$$\min\left\{m, \left\lfloor \frac{n+m}{r} \right\rfloor\right\},$$

or more precisely:

$$g_r^*(n,m) = \begin{cases} m, & \text{for } m \leq \lfloor \frac{n}{r-1} \rfloor \\ \lfloor \frac{n+m}{r} \rfloor, & \text{for } m > \lfloor \frac{n}{r-1} \rfloor. \end{cases}$$

Proof. The bounds are straightforward. For the first bound, place a guard in each doorway. Since the rooms are convex every room can be covered by at least one of the guards.

For the second bound, one can first provide a labeling of the walls with r labels, such that each room has every label on one of its walls. We will do so by induction on m with the result being trivial for $m = 0$. For $m > 0$ cut along an interior wall and apply induction on both parts. To combine both labelings into one it may be necessary to swap labels in one of the parts if they disagree on the separating wall. Now placing one guard at each wall suggested by the label used least frequently establishes the bound.

For the construction, let D_r be the art gallery with one central room and $r-1$ side rooms of size r. The side rooms share a wall only with the central room. Figure 3.8 shows D_3, D_4 and D_5. Each D_r has $r-1$ interior walls, $r(r-2)+2$ vertices and requires $r-1$ guards.

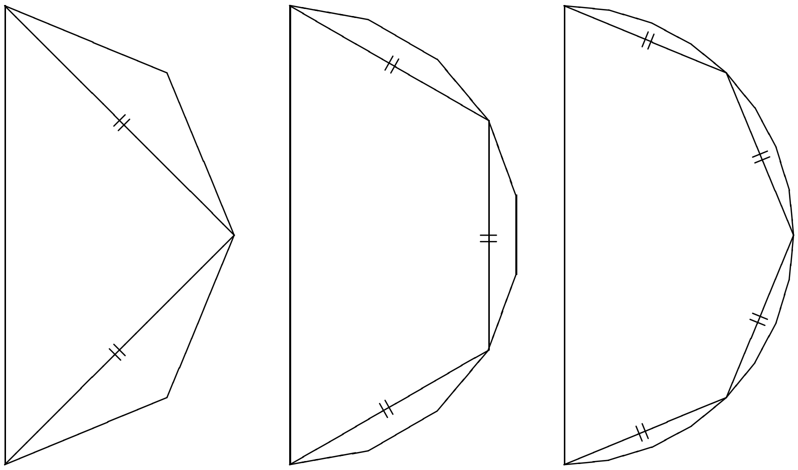


Figure 3.8: D_3, D_4 and D_5

If $m \le \lfloor \frac{n}{r-1} \rfloor$, our gallery will be similar to a D_r:

Take a center room of size $n - m(r-2) \ge r$, and attach m side rooms of size r each. This is possible, since when $m \le \lfloor \frac{n}{r-1} \rfloor$ we have $n - m(r-2) \ge m$, and this gallery requires m guards.

If on the other hand $m > \lfloor \frac{n}{r-1} \rfloor$, then let $k = \lceil \frac{m(r-1)-n}{r} \rceil > 0$. We will form our gallery by taking a gallery on $n' = n - kr(r-2)$ vertices and $m' = m - kr \ge 0$ interior walls and then attaching k D_r's to it. Always attaching the next D_r with the free wall of its center room to any other room we obtain a gallery on n vertices and m walls.

It can be readily checked that $n' \ge m'(r-2) + r$ and $n' \ge m'(r-1)$ so that we can take a gallery from the first case to start out with. So we need

$$m' + k(r-1) = m - k = \lfloor m - \frac{m(r-1)-n}{r} \rfloor = \lfloor \frac{m+n}{r} \rfloor$$

guards to cover this gallery. $\square$

3.4. Complexity

In implementing Fisk's proof, Avis and Toussaint [10] obtained an $O(n \log n)$ algorithm to place $\lfloor n/3 \rfloor$ guards to guard an n-vertex art gallery. This approach can be improved to obtain a linear time algorithm, since Chazelle ([27,28]) showed that an n-vertex polygon can be triangulated in time $O(n)$.

However, so far this complex algorithm has resisted implementation. The fastest algorithms that have been successfully implemented in practice are randomized algorithms with an expected running time of $O(n \log^* n)$. See for example Seidel [106] or Devillers [37]; or Manocha and Narkede [93] for an implementation of Seidel's algorithm. Here $\log^*$ is recursively defined to be $\log^* x = 1$, for $x \le 2$, and $\log^* x = 1 + \log^* \lg x$, for $x > 2$. Note that this "inverse of the tower function $2^{2^{2^{\cdot^{\cdot^{\cdot}}}}}$" grows very slowly and is negligible for practical purposes.

First triangulate the polygon in time $\mathrm{O}(n)$. Since outerplanar triangulations are chordal graphs, one can find a vertex elimination scheme (see for example [104]) and then use this to obtain a 3-coloring of the vertices ([62]), both in $\mathrm{O}(n+e) = \mathrm{O}(n)$ time. Now placing the guards just requires $\mathrm{O}(n)$ time and the algorithm is linear.

From here it is easy to see how the upper bound arguments in Lemma 3.2.1 and 3.2.3 can be used to find linear time algorithms for these problems too. The only problem could be that the existence of interior walls does not necessarily make triangulation easier. However we can triangulate each room separately. This is still possible in linear time, since they have a total of $n + 2m < 3n$ vertices. Finding the rooms and therefore also the weak dual can be done in linear time. See [29] or [80] to find the rotation scheme from which this can be done.

A straightforward implementation of Lemma 3.2.4 results in an $\mathrm{O}(n^2)$ algorithm, due to the stronger statement, even for $m = 0$. However this case can be implemented in linear time even with the stronger statement. For $n = 3$ just add the guard where requested. For $n \geq 4$ apply the basic algorithm and just add in the extra guard at the required place if necessary. This will work since $\lfloor n/3 \rfloor + \epsilon \leq \lfloor (n-1)/2 \rfloor$, for $n \geq 4$ with $\epsilon = 1$ when n is odd and 0 otherwise.

This makes a faster algorithm plausible, however it is still an open question whether a linear time algorithm can be obtained in this case. A Fisk type proof for the upper bound in Lemma 3.2.4 would certainly yield a fast algorithm.

In the case of convex rooms the situation is easier. The first bound trivially leads to an $\mathrm{O}(m)$ algorithm. The second bound can be implemented in time $\mathrm{O}(m) = \mathrm{O}(n)$ as well, since the labeling can be found in linear time: Find the weak dual of the art gallery, then starting at any vertex conduct a Breadth-First Search on this dual tree to determine the order in which the rooms will be labeled. In the first room label the edges using each one of the r labels at least once. On each consecutive room one wall is already labeled, so label the remaining walls accordingly to assure that every label is being used.

It is important to note that although our algorithms give fast guard placements for given classes of art galleries efficiently they do not necessarily give the best possible answer for a specific art gallery. This problem is known to be NP-hard even when the art gallery has no interior walls ([1,99]).

4. Turán problems for weighted Graphs

4.1. Preliminaries, density results in graphs

In this section we survey a few basics of Turán theory to indicate the context of our results. We will give the main results in the next two sections, and follow them by the proofs. We also use some part of the theory later in the proofs.

In 1963 Erdős [40] proposed the following problem: Given the parameters n, k and r, determine the maximum number $\mathrm{ex}(n,k,r)$ of edges a graph G on n vertices can have if every k-vertex subgraph has at most r edges. He reviewed all cases $k \le 5$.

Let the *Turán graph* $T_{n,p}$ be the balanced complete p-partite graph on n vertices, i.e., the n vertices are grouped into p parts as evenly as possible and two vertices are joined if and only if they belong to distinct parts. Let $t(n,p)$ denote the number of edges in $T_{n,p}$. We have $t(n,2) = \lfloor n^2/4 \rfloor$, and in general

$$t(n,p) = (1 - \frac{1}{p})\binom{n}{2} + \mathrm{O}(n).$$

Turán's theorem [113,114] states that in the case $(k,r) = (k, \binom{k}{2} - 1)$, i.e., if G does not contain K_k as a subgraph, then G can have at most $t(n,k-1)$ edges and that the only graph having this many edges is $T_{n,k-1}$. The case $\mathrm{ex}(n,3,2) = \lfloor n^2/4 \rfloor$ was proved earlier by Mantel [94]. Dirac [38] generalized Turán's theorem, for $r = \binom{k}{2} - t$ with $1 \le t \le k/2$

$$\mathrm{ex}(n,k,\binom{k}{2} - t) = t(n, k-t).$$

The question of Erdős is still unanswered, but one can briefly summarize the main results as follows. There are three types of answers, Gol'berg and Gurvich [64] call them *zones*. The *linear zone* consists of pairs $U^1 := \{(k,r) : 0 \le r < k-1\}$, the *polynomial zone* $U^2 := \{(k,r) : k-1 \le r < \lfloor k^2/4 \rfloor\}$ and finally $U^3 := \{(k,r) : \lfloor k^2/4 \rfloor \le r \le \binom{k}{2}\}$ is called the *quadratic zone*.

Trivially, $\mathrm{ex}(n,k,r) = r$ for $0 \le r < (k-1)/2$. For $0 \le r < k-1$ we have

$$\mathrm{ex}(n,k,r) = v - 1 + \lfloor \frac{u-1}{u}(x-v) \rfloor, \tag{4.1.1}$$

where $u = \lfloor (k-1)/(k-r-1) \rfloor$ and $v = k-1-u(k-r-2)$. We have $1 \le u \le v \le r+1$. Here an example of an extremal (k,r)-dense graph is the disjoint union of one star on v vertices and as many stars on u vertices as possible, plus one remainder star on at most u vertices. Note that (4.1.1) covers the case $0 \le r < (k-1)/2$ too, since in this case $u = 1$.

In the polynomial range for $(k,r) \in U^2$ we have

$$\Omega(n^{1+1/(k+1)}) \le \mathrm{ex}(n,k,r) \le \mathrm{O}(n^{2-1/k}).$$

As Erdős [40] already pointed out, determining $\mathrm{ex}(n,k,k-1)$ is in fact equivalent to determining $\mathrm{ex}(n, C_{2t})$, the maximum number of edges of an n vertex graph containing no cycle of length $2t$, where $t = \lfloor k/2 \rfloor$. Furthermore, this is basically equivalent to the problem of determining $\mathrm{ex}(n, \{C_3, C_4, \dots, C_{2t}\})$, the maximum number of edges of an n vertex graph of girth at least $2t+1$. Erdős used probabilistic arguments to prove a lower bound $\mathrm{ex}(n, C_{2t}) = \Omega(n^{1+1/2t})$. An upper bound due to Bondy and Simonovits [16] states $\mathrm{ex}(n, C_{2t}) = \mathrm{O}(n^{1+1/t})$. This is conjectured to give the correct order of magnitude. However, this is supported by a construction only for $t = 2, 3$ and 5 (see, e.g., in Benson [12], Wenger [116], or Bollobás [14]). In the remaining cases the best lower bound is given by the so-called "Ramanujan graphs" of Lubotzky, Phillips and Sarnak [92], Margulis [95] and Imrich [73]. Their construction was slightly improved by Lazebnik, Ustimenko and Woldar [90], who borrowed ideas from the theory of Lie algebras to prove that $cn^{1+\frac{2}{3k-2}} < \mathrm{ex}(n, C_{2k})$.

In the polynomial range a result of Kővári, T. Sós and Turán [83] implies that

$$\mathrm{ex}(n, k, \lfloor k^2/4 \rfloor - 1) \le \mathrm{ex}(n, K_{\lfloor k/2 \rfloor, \lceil k/2 \rceil}) = \mathrm{O}(n^{2-2/k}).$$

Finally, in the quadratic zone, the Erdős-Stone-Simonovits [47,48] theorem implies that

$$\mathrm{ex}(n,k,r) = (1 - \frac{1}{q})\binom{n}{2} + \mathrm{o}(n^2), \text{ for } t(k,q) \le r < t(k, q+1). \tag{4.1.2}$$

Even more, it was recently summarized by Griggs, Simonovits, Thomas [65] (in a crude form also in Gol'berg and Gurvich [64]) that for each q the range in the above inequality can be again cut into three parts where the error term is constant, linear, or polynomial. Let $r = t(k,q) + a$ (where $0 \le a < t(k, q+1) - t(k,q)$). Then for $n > n_0(k)$

$$\mathrm{ex}(n,k,r) = \begin{cases} t(n,q) + a, & \text{for } 0 \le a < \lfloor \frac{1}{2} \lceil \frac{k}{q} \rceil \rfloor, \\ t(n,q) + \Theta(n), & \text{for } \lfloor \frac{1}{2} \lceil \frac{k}{q} \rceil \rfloor \le a \le \lceil \frac{k}{q} \rceil - 2, \\ t(n,q) + \mathrm{o}(n^2), & \text{for } \lceil \frac{k}{q} \rceil - 2 < a. \end{cases}$$

They also present a method to determine the linear error term precisely, and in the third zone the error term is super-linear. An interesting example of their results for this case is the description of the maximal graph for $k = 6$, $r = 11$. It is obtained from a complete bipartite graph $K[A, B]$ where the parts A and B do not necessarily have equal sizes, but $|A|, |B| = \frac{1}{2}n + \mathrm{o}(n)$ and one can add an arbitrary graph of girth 5 on A. Obviously, this example carries at most 11 edges in every 6 vertices, and it has $\frac{1}{4}n^2 + \Theta(n^{3/2})$ edges. (For C_4-free graphs see Brown [19] and Erdős, Rényi, Sós [46].)

The density problem $\mathrm{ex}(n,k,r)$ was investigated in several more papers, for example in [42] Erdős, Faudree, Jagota and Łuczak deal with the case when $k = \Theta(n)$ and $r = \mathrm{O}(n^2)$.

A strongly related problem, the restricted multicoloring of the edges of the complete graph K_n, was investigated by Ahlswede, Cai and Zhang [3]. The resulting formulas are very similar. For recent developments see Axenovich and Kündgen [11].

Density problems in graphs are often difficult. Brown, Erdős and Sós [24,25] proposed an even more difficult version – they investigated h-uniform hypergraphs, i.e., families of h-element sets. They introduced $\mathrm{ex}_h(n,k,r)$, the maximum number of edges in an h-uniform hypergraph of order n such that no k vertices induce more than r edges. For example, a $(4,2)$-free family of triples is a (partial) Steiner family, we have $\mathrm{ex}_3(n, 4, 1) \le n(n-1)/6$. Also see, Gurvich [67].

A celebrated result of Ruzsa and Szemerédi [105] deals with $(6,3)$-free triple systems

$$\mathrm{ex}_3(n,6,2) = \mathrm{o}(n^2) \quad \text{but} \quad \lim_{n\to\infty} \mathrm{ex}_3(n,6,2)/n^{2-c} = \infty \text{ for all } c > 0.$$

Some generalizations and a proof involving Szemerédi's regularity lemma can be found in Erdős, Frankl, Rödl [43].

Since many problems in combinatorics have weighted versions it seems natural to study weighted generalizations. In what follows we deal with multigraphs, where we usually do not have any constraints on the multiplicity of the edges. We even allow negative (integer) multiplicities, when the solution proves to be simpler.

This chapter is organized as follows. In the next section we generalize Erdős' density problem and mention a few more preliminaries. In Section 4.3 we state our theorems and we determine the limits in the case of arbitrary integer weights. In Section 4.4 a recursive upper bound is given. In Section 4.5 a construction is given with special attention to the (in some cases not exactly known) linear error term. In Section 4.6 we summarize our results by providing exact answers for "slightly more than 50% of the cases". In Section 4.7 we continue the investigation of the upper bound to obtain a lower bound which is conjectured to be exact. In Section 4.8 we solve our density question for sufficiently large n whenever the limit is 0. In Section 4.9 we conclude this chapter by proposing further questions and applying the results for multigraphs.

4.2. The density problem

Definition 4.2.1. A *weighting* of a graph G from a set of weights $W \subset \mathbb{R}$ is a function $w : E(G) \to W$ that assigns a weight to each edge. The *weight of a subgraph* H of G is just the total weight on all of its edges, that is

$$w(H) := \sum \{w(e) : e \in E(H)\}.$$

The *weight of a vertex set* A is the weight of the subgraph induced by A, $w(A) := w(G|_A)$. For example for every vertex v $w(\{v\}) = 0 = w(\emptyset)$. We will also define the *weight of a vertex* v by $w(v) := \sum\{w(uv) : uv \in E(G)\}$, but we will only use this notion in Sections 4.7 and 4.8. A weighted graph (G, w) is (k,r)*–dense* if every set of k vertices has weight at most r. In this language we define the weighted Turán numbers as

$$\mathrm{ex}_W(n,k,r) := \sup\{w(G) : |V(G)| = n,\ (G,w) \text{ is } (k,r)\text{–dense}\}.$$

Since every non-edge can be considered to be an edge of weight 0, we will always assume that $0 \in W$, so that it will suffice to consider $G = K_n$. From now on, when we refer to a graph we mean the corresponding $\{0,1\}$-weighting, where edges are weight 1 and non-edges are weight 0. Multigraphs can be viewed as $\mathbb{Z}^+$-weightings of K_n, where we define $\mathbb{Z}^+$ as the set $\{0,1,2,3,\dots\}$.

It is an easy observation that $\mathrm{ex}_W(n,k,r)/\binom{n}{2}$ is a monotone nonincreasing sequence in n, that is for $k \le n' < n$, we have

$$\frac{\mathrm{ex}_W(n,k,r)}{\binom{n}{2}} \le \frac{\mathrm{ex}_W(n',k,r)}{\binom{n'}{2}}. \tag{4.2.1}$$

Indeed, for any (k, r)-dense weighting w of K_n, we have

$$\frac{w(K_n)}{\binom{n}{2}} = \frac{w(K_n)\binom{n-2}{n'-2}}{\binom{n'}{2}\binom{n}{n'}} = \frac{\sum\{w(A) : |A| = n', A \subset V(K_n)\}}{\binom{n'}{2}\binom{n}{n'}} \leq \frac{\mathrm{ex}_W(n', k, r)}{\binom{n'}{2}}.$$

Hence the *asymptotic density*

$$\mathrm{ex}_W(k, r) := \lim_{n\to\infty} \mathrm{ex}_W(n, k, r)/\binom{n}{2}$$

exists for all k and $r \geq 0$. From the monotonicity we also get the *counting bound*

$$\mathrm{ex}_W(n, k, r) \leq \frac{\mathrm{ex}_W(k, k, r)}{\binom{k}{2}}\binom{n}{2} \leq \frac{r}{\binom{k}{2}}\binom{n}{2}. \tag{4.2.2}$$

If W is dense in $\mathbb{R}$ then the counting bound is always achieved and equalities hold in (4.2.2). We focus our interest on the case when all weights as well as r are integers. For example, from the reformulation of the Erdős-Stone-Simonovits theorem, (4.1.2), we obtain

$$\mathrm{ex}_{\{0,1\}}(k, r) = 1 - \frac{1}{q}, \text{ for } t(k, q) \leq r < t(k, q + 1).$$

Thus the possible limit points form a well-ordered set when $W = \{0, 1\}$. Brown, Erdős and Simonovits [20,21] proved that the same is true in the case of multigraphs with maximum edge multiplicity 2, i.e. $W = \{0, 1, 2\}$, (and also for digraphs [22,23]). However, more recent results of Rödl and Sidorenko [103] about the Turán number of multigraphs with edge multiplicities 4 or more state that in general the limit points of $\mathrm{ex}_{[m]}(n, G)/\binom{n}{2}$ are not well-ordered. This makes it unlikely that the set $\mathrm{ex}_W(k, r)$ is well-ordered for $W = [m] := \{0, 1, 2, \ldots, m\}$ for any given $m \geq 4$, but this problem, as well as many others remain open.

Our main result in this chapter is to determine $\mathrm{ex}_{\mathbb{Z}}(k, r)$ for all (k, r). For fixed k it takes all nonnegative rationals with denominator at most $k - 1$. Furthermore, we will give constructions of graphs with positive and negative integer weights (i.e., functions $w : E(K_n) \to \mathbb{Z}$) and show that these constructions give the exact answer for $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$ at least half the pairs (k, r). We also show that they are asymptotically optimal, and in contrast to (4.1.2) our error term is always at most $\mathrm{O}(n)$. The problem of $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$ has been studied by Bondy and Tuza [17], and more recently by Kuchenbrod [85] (up to $k \leq 6$), but it had already been suggested in earlier works of Brown, Erdős and Simonovits [20–23].

We apply our results to determine $\mathrm{ex}_{\mathbb{Z}^+}(k, r)$ for infinitely many pairs (k, r). The study of $\mathrm{ex}_{\mathbb{Z}^+}(n, k, r)$, i.e., to study multigraphs without upper bound on the multiplicities on the edges was proposed by de Caen and Székely [36], where they needed the case $(k, r) = (4, 15)$. However the problem in general is more difficult, several cases remain open. The case $W = \mathbb{Z}^+$ is particularly interesting, since it is a generalization of Turán's theorem for multigraphs.

Finally, let us mention another application of Theorem 4.3.1. In [35] the following old conjecture of Erdős and Sós is proved. Any set of $\frac{3}{4}\binom{n}{3} + cn^2$ triples from an n-set, where c is a suitable constant, must contain a copy of the Fano configuration (the projective plane of order two). This is an asymptotically sharp estimate as the following triple-system contains no Fano plane: Divide the n-element underlying set into two almost equal parts and consider all triples meeting both.

4.3. Definitions and main results

Let $\mathbb{Z}^+ = \{0, 1, 2, 3 \ldots\}$ denote the nonnegative integers. For a positive integer k and a real $m \geq 0$ let $R_k(m)$ be the region in $\mathbb{R}^2$ bounded by the inequalities $1 \leq x \leq k-1$ and $0 \leq y \leq mx$ (See Figure 4.1). This region contains exactly $|R_k(m) \cap \mathbb{Z}^2| = \sum_{i<k} \lfloor 1 + mi \rfloor$ integer lattice points. We will define $m_{k,r}$ to be the smallest slope m so that $R_k(m)$ contains at least $r+1$ lattice points,

$$m_{k,r} := \min\{m : |R_k(m) \cap \mathbb{Z}^2| > r\}. \tag{4.3.1}$$

For example, in the case $k = 11$, $r = 32$ the region $R_{11}(1/2)$ in Figure 4.1 contains 35 lattice points, 5 of them on its upper boundary, showing that $m_{11,30} = m_{11,31} = \cdots = m_{11,34} = 1/2$.

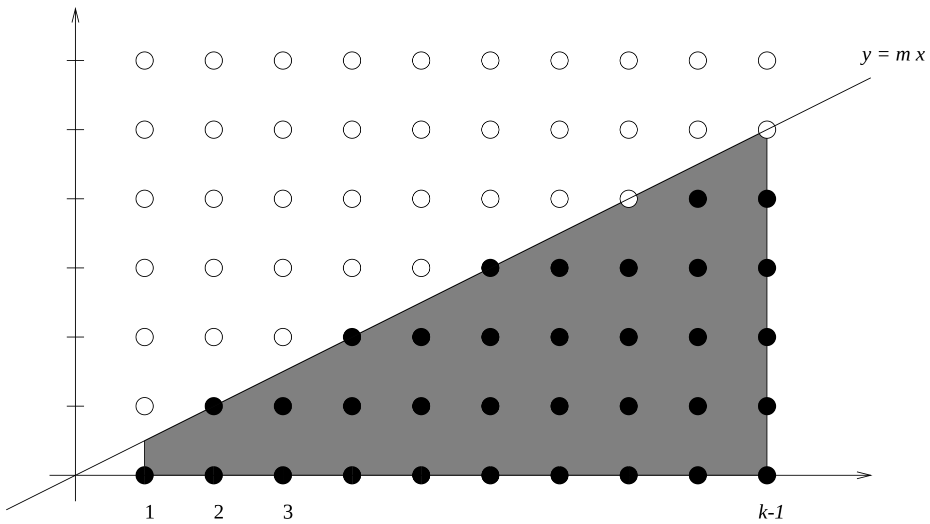


Figure 4.1: $R_k(m_{k,r})$ with $(k, r) = (11, 32)$. Here $\ell = 5$ and $a = 3$.

Our main result is that the slope m obtained from (4.3.1) yields the asymptotic density in the case when $W = \mathbb{Z}$:

$$\mathrm{ex}_{\mathbb{Z}}(k, r) = m_{k,r}. \tag{4.3.2}$$

Moreover, the region $R_n(m_{k,r})$ is intimately related to the maximal weight of a (k, r)-dense graph:

$$0 \leq |R_n(m_{k,r}) \cap \mathbb{Z}^2| - \mathrm{ex}_{\mathbb{Z}}(n, k, r) < n \tag{4.3.3}$$

holds for all $n \geq k$. Summarizing, we have

Theorem 4.3.1. *(Asymptotic)*

$$\mathrm{ex}_{\mathbb{Z}}(n, k, r) = m_{k,r} \binom{n}{2} + \mathrm{O}(n)$$

In other words, if we put $r+1$ lattice points in the region $0 < x < k$, $y \geq 0$, "as far down as possible", then the smallest slope m of any line through the origin that creates a region containing all these points gives the asymptotic density. Obviously, m is always a rational number and can be written in reduced form, $m = p/q$ with g.c.d.$(p,q) = 1$ and $q > 0$.

To further determine $\mathrm{ex}_{\mathbb{Z}}(n,k,r)$ let $\ell_k(p/q)$ denote the number of lattice points of $R_k(p/q)$ on the line $y = \frac{p}{q}x$, $\ell_k(p/q) = \lfloor (k-1)/q \rfloor$. Place the $r+1$ lattice points into $R_k(p/q)$ by using all points below the line $y = \frac{p}{q}x$. Then let $a_k(r)$ denote the number of points we have to use on the upper boundary. We have

$$1 \leq a_k(r) = r + 1 - (|R_k(p/q) \cap \mathbb{Z}^2| - \ell_k(p/q)) = r + 1 - \sum_{i<k} \lceil \frac{p}{q} i \rceil \leq \ell_k(p/q).$$

From now on, we abbreviate these two functions as ℓ and a.

In Section 4.4 we will give the upper bound for Theorem 4.3.1 by establishing

Theorem 4.3.2. *(Upper bound for all values)*

$$\mathrm{ex}_{\mathbb{Z}}(n,k,r) \leq \sum_{i<n} \lceil \frac{p}{q} i \rceil + \begin{cases} a-1, & \text{for } a \leq (\ell+1)/2, \\ \lfloor \frac{n-1}{q} \rfloor + a - \ell - 1, & \text{for } a > (\ell+1)/2. \end{cases}$$

The proof of Theorem 4.3.2 essentially uses the monotonicity of the density (4.2.1) in the following form.

$$\mathrm{ex}_{\mathbb{Z}}(n+1,k,r) \leq \lfloor \frac{\binom{n+1}{2}}{\binom{n}{2}} \mathrm{ex}_{\mathbb{Z}}(n,k,r) \rfloor = \mathrm{ex}_{\mathbb{Z}}(n,k,r) + \lfloor \frac{2}{n-1} \mathrm{ex}_{\mathbb{Z}}(n,k,r) \rfloor \tag{4.3.4}$$

In Section 4.5 we will give a recursive construction for the lower bound in Theorem 4.3.1. Furthermore, we will find the exact value for many cases in Section 4.6:

Theorem 4.3.3. *(Exact answer for half of the cases)*
If $2a \leq \ell + 1$ then

$$\mathrm{ex}_{\mathbb{Z}}(n,k,r) = \sum_{i<n} \lceil \frac{p}{q} i \rceil + a - 1.$$

Theorem 4.3.4. *(Exact answer for small n)*
If $2a \geq \ell + 1$ and $n \leq 2aq$ then

$$\mathrm{ex}_{\mathbb{Z}}(n,k,r) = \sum_{i<n} \lceil \frac{p}{q} i \rceil + \lceil \frac{n}{q} \rceil + a - \ell - 2.$$

Note that $\lfloor \frac{n-1}{q} \rfloor = \lceil \frac{n}{q} \rceil - 1$ so the upper bound for $\mathrm{ex}_{\mathbb{Z}}$ in Theorem 4.3.4 follows from Theorem 4.3.2. Theorem 4.3.4 gives exact answers only for small values of n, $n < 2k$, and in the case $a = (\ell+1)/2$ its range is included in $\{n : k \leq n < k + q\}$. This is the reason that in this case the two kinds of remainders in Theorem 4.3.2 coincide.

In Section 4.7 we will establish the lower bound that we believe to be optimal.

Theorem 4.3.5. *(Lower bound)*

$$\mathrm{ex}_{\mathbb{Z}}(n,k,r) \geq \sum_{i<n} \lceil \frac{p}{q} i \rceil + \mathrm{ex}_{\mathbb{Z}}(\lceil \frac{n}{q} \rceil, \ell+1, a-1).$$

Theorems 4.3.3 and 4.3.4 show that in the respective cases $2a \leq \ell+1$ and $n \leq 2aq$ equality holds. We conjecture that equality always holds, in other words that the linear term is itself of the form $\mathrm{ex}_{\mathbb{Z}}(n,k,r)$, for $r \leq k-2$. In Section 4.8 we give a conjecture for the linear term, which we can prove for large n (see Theorem 4.3.6). Proving both conjectures would completely settle the problem in the case when we allow integer weights.

Note that for $m_{k,r} \neq 0$ Theorem 4.3.1 yields the order of magnitude and the leading coefficient $\mathrm{ex}_{\mathbb{Z}}(n,k,r) = (1/2)m_{k,r}n^2 + \mathrm{O}(n)$ – and it gives very tight bounds. For the missing case, $m_{k,r} = 0$, we have the next result.

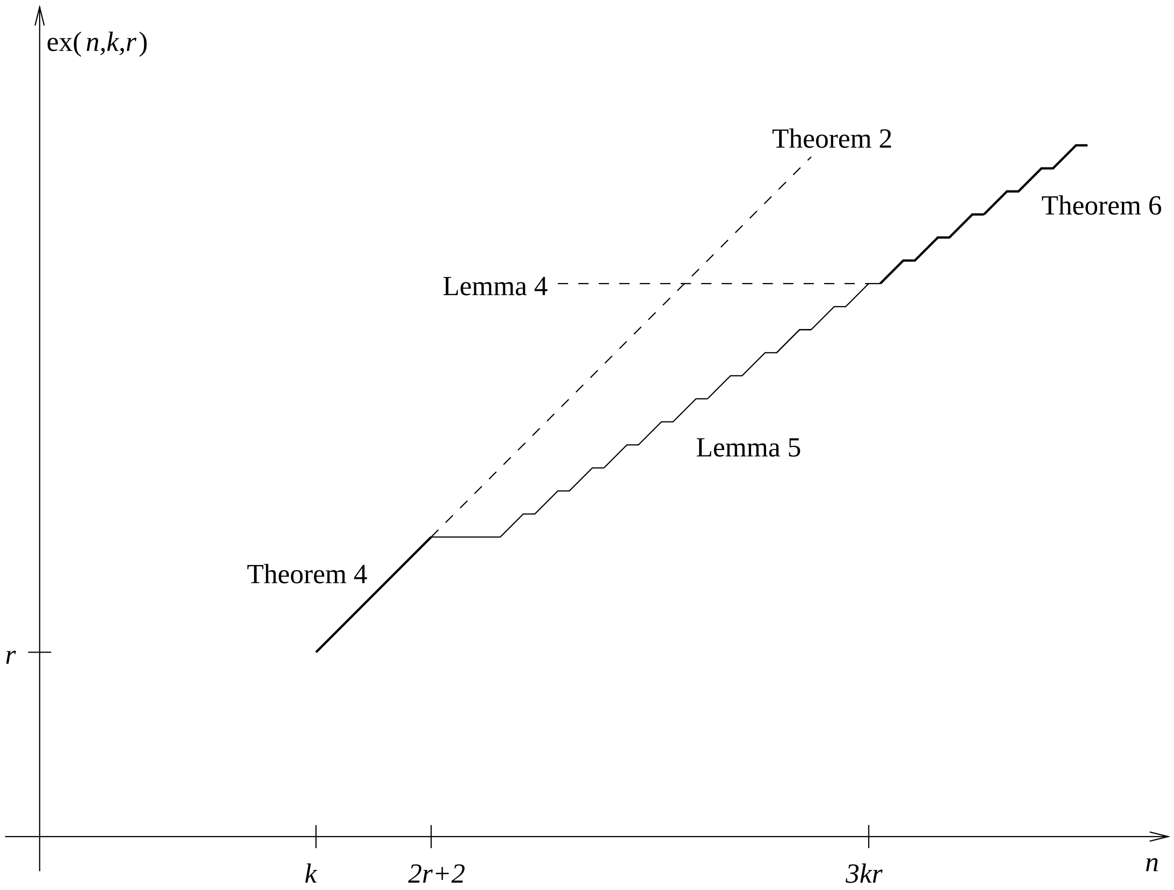


Figure 4.2: The linear case. $k = 12$, $r = 8$, $u = 3$.

Theorem 4.3.6. *(Exact value for $m_{k,r} = 0$ and large n)*
Suppose that $0 < r \leq k-2$ and $n \geq 3kr$. Then

$$\mathrm{ex}_{\mathbb{Z}}(n,k,r) = r + \lfloor \frac{u-1}{u}(n-k+1) \rfloor,$$

where $u = \lfloor (k-1)/(k-1-r) \rfloor$.

Finally in Theorem 4.9.1 in Section 4.9 we address the multigraph case $W = \mathbb{Z}^+$, by observing that many of the constructions previously given do not use negative weights and are therefore also optimal in this case.

4.4. Proof of the general upper bound

In this section we first study a recurrence relation. After that, the proof of Theorem 4.3.2 will be immediate. In the next statement p, q, a and ℓ are as in Section 4.3

Theorem 4.4.1. *Suppose that $f = f_{k,r}$ is an integer function defined by the starting value $f(k) = r \geq 0$ and for $n \geq k > 0$ by the recursion $f(n+1) = \lfloor \frac{n+1}{n-1} f(n) \rfloor$. Then*

$$f(n) = \sum_{i<n} \lceil \frac{p}{q} i \rceil + \begin{cases} a-1, & \text{for } a \leq (\ell+1)/2, \\ \lfloor \frac{n-1}{q} \rfloor + a - \ell - 1, & \text{for } a > (\ell+1)/2. \end{cases}$$

This Theorem, in a slightly different form, is due to Gessel and Indik [63]. (We are grateful to M. Simonovits for this information.) Our proof here is slightly more straightforward. Special cases appear in the literature in several places, even in papers related to Turán theory (e.g., Ahlswede, Cai, Zhang [3], Stechkin [109], Stechkin, Frankl [110]). Frequently it is stated and used in the form of inequalities, i.e., if for some sequence of integers $s_k, s_{k+1}, \ldots$ one has

$$s_k = r \text{ and } \frac{s_{n+1}}{\binom{n+1}{2}} \leq \frac{s_n}{\binom{n}{2}} \text{ for } n \geq k, \text{ then } s_n \leq f_{k,r}(n)$$

holds. For example the case $k=5$, $r=7$ (i.e., $f_{5,7}(n) = (n^2+n)/2$) was proved in [54]. Guy [68] also suggested that these and similar recurrences may be worth studying for their own sake.

Proof of Theorem 4.4.1. We will proceed by induction on n. Write $f(n)$ in the form $\sum_{i=1}^{n-1} \lceil \frac{p}{q} i \rceil + h(n)$. It suffices to establish

$$\lfloor \frac{2}{n-1} f(n) \rfloor = \lceil \frac{p}{q} n \rceil + \varepsilon, \tag{4.4.1}$$

where $\varepsilon = 1$ if q divides n and $2h(n) \geq \lfloor (n-1)/q \rfloor$ and $\varepsilon = 0$ otherwise. Simultaneously to (4.4.1) we use induction to prove that

$$0 \leq h(n) < \lfloor (n-1)/q \rfloor. \tag{4.4.2}$$

We establish the following flow of implications (Theorem 4.4.1 for $f(n)$) $\implies$ ((4.4.2) for $h(n)$) $\implies$ ((4.4.1) for $f(n)$) $\implies$ (Theorem 4.4.1 for $f(n+1)$). Here the first implication is obvious. For the base case $n = k$ we have

$$f(k) = r = \sum_{i<k} \lceil \frac{p}{q} i \rceil + a - 1 = \sum_{i<k} \lceil \frac{p}{q} i \rceil + a - 1 + \lfloor \frac{k-1}{q} \rfloor - \ell$$

thus (4.4.2) holds for $h(k)$.

Next we consider $2f(n)/(n-1)$. Observe that

$$2\ f(n) = \sum_{i=1}^{n-1} (\lceil \frac{p}{q} i \rceil + \lceil \frac{p}{q}(n-i) \rceil) + 2h \geq (n-1) \lceil \frac{p}{q} n \rceil$$

In the case when q does not divide n we also conclude that

$$\sum_{i=1}^{n-1} (\lceil \frac{p}{q} i \rceil + \lceil \frac{p}{q}(n-i) \rceil) + 2h \leq (n-1)(\lceil \frac{p}{q} n \rceil + 1) - 2\lfloor (n-1)/q \rfloor + 2h.$$

This inequality holds since for i or $n-i$ divisible by q we can drop the corresponding ceiling and save one. Thus

$$\lceil \frac{p}{q} n \rceil \leq \lfloor \frac{2}{n-1} f(n) \rfloor \leq \lceil \frac{p}{q} n \rceil + \lfloor 1 + 2 \frac{h - \lfloor (n-1)/q \rfloor}{n-1} \rfloor \leq \lceil \frac{p}{q} n \rceil.$$

In the case when q does divide n we get that

$$2\ f(n) = (n-1)(\lceil \frac{p}{q} n \rceil + 1) - \lfloor (n-1)/q \rfloor + 2h,$$

so that for $2h < \lfloor (n-1)/q \rfloor$ again

$$\lfloor \frac{2}{n-1}\ f(n) \rfloor = \lceil \frac{p}{q} n \rceil + \lfloor 1 + \frac{2h - \lfloor (n-1)/q \rfloor}{n-1} \rfloor = \lceil \frac{p}{q} n \rceil,$$

and for $2h \geq \lfloor (n-1)/q \rfloor$

$$\lfloor \frac{2}{n-1}\ f(n) \rfloor = \lceil \frac{p}{q} n \rceil + \lfloor 1 + \frac{2h - \lfloor (n-1)/q \rfloor}{n-1} \rfloor = \lceil \frac{p}{q} n \rceil + 1.$$

In the case $a \leq (\ell+1)/2$ we have $2h = 2a - 2 \leq \ell - 1 = \lfloor (k-1)/q \rfloor - 1 < \lfloor (n-1)/q \rfloor$ so that $\varepsilon = 0$ for all n. Thus $h(n) = a - 1$ for every $n \geq k$.

In the case $a > (\ell+1)/2$ we have $2h(k) > \lfloor (k-1)/q \rfloor$, and by induction we obtain that $2h(n) > \lfloor (n-1)/q \rfloor$, because $\lfloor (n-1)/q \rfloor + \varepsilon = \lfloor n/q \rfloor$, since $\varepsilon = 0$ unless n/q is an integer. $\square$

Proof of Theorem 4.3.2. This is immediate from Theorem 4.4.1. For the base case we have $\mathrm{ex}_{\mathbb{Z}}(k,k,r) \leq r = f_{k,r}(r)$. Then $\mathrm{ex}_{\mathbb{Z}}(n,k,r) \leq f(n)$ is implied by the monotonicity of $\mathrm{ex}_W(n,k,r)/\binom{n}{2}$, see (4.3.4). $\square$

Unfortunately, this upper bound does not always give the exact value of $\mathrm{ex}_{\mathbb{Z}}(n,k,r)$.

4.5. Construction of dense graphs for integer weights

In this section a recursive construction is given which yields the right lower bound up to a $O(n)$ error term. To get the best possible result we formulate it in a slightly more general way. One of our technical difficulties is that we do not know $ex_{\mathbb{Z}}(n,k,r)$ for all $n > k \geq r+2 \geq 2$ (i.e., when it is linear in n according to Theorem 4.3.6), so we are going to introduce a linear error term function h.

The main idea is to obtain a structure similar to that of $T_{n,k}$ with the right properties. We distribute the vertices as equally as possible into q classes. The weights on the edges are determined by a symmetric q by q matrix $M(p/q)$, that is $w(uv) = M_{i,j}$ if u is in the i-th class and v in the j-th. In the Turán case $T_{n,k}$, $M = M((k-1)/k)$ is the k by k matrix with all entries 1 except the main diagonal which consists of zeros. Other examples for $M(p/q)$ are in Figure 4.3.

$$
\begin{bmatrix} 0 \end{bmatrix} \qquad
\begin{bmatrix} -1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad
\begin{bmatrix}
-2 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & -1 & -1 & 1 & 1 & 1 & 1 & 1 \\
0 & -1 & -1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

Figure 4.3: $M(0/1)$, $M(1/3)$ and $M(3/8)$.

These constructions will asymptotically yield the correct density p/q. However in some cases we can improve the construction by adding a small forest-like construction to the class with the highest negative weight, that is in Figure 4.3 in the top left corner. This will result in adding a linear number of edges.

To capture the behavior of this linear term we will need the following

Definition 4.5.1. A function $h : \mathbb{Z}^+ \to \mathbb{Z}^+$ is called a *linear weighting function* if it has the following properties:

(L1) *Monotonicity:* $h(x) \leq h(x+1)$,

(L2) *Lipschitz:* $h(x+1) \leq h(x)+1$, and

(L3) For every $n \in \mathbb{Z}^+$ there exists a weighting w of K_n with integer weights such that $w(K_n) = h(n)$ and $w(A) \leq h(|A|)$ for every vertex set A.

Notice that this forces $h(1) = 0 = h(0)$ so that we only need to specify h for $x > 1$. The zero function is an example of a linear weighting function and we will give other examples in Sections 4.6 and 4.7. We call these weighting functions linear since $h(n) = O(n)$.

We will now give a recursive construction establishing

Lemma 4.5.2. *For every linear weighting function with $h(\ell+1) \leq a-1$*

$$ex_{\mathbb{Z}}(n,k,r) \geq \sum_{i<n} \lceil \frac{p}{q} i \rceil + h(\lceil \frac{n}{q} \rceil).$$

Using the zero function for h, which corresponds to the pure matrix construction without adding a linear term, we immediately obtain the lower bound for Theorem 4.3.1, finishing its proof. Notice that the difference between the lower bound from Lemma 4.5.2 and the upper bound in Theorem 4.3.2 is less than n/q. The reader may want to follow the proof in the matrix structure with the examples provided in Figure 4.3, while assuming that h is the zero function.

Proof of Lemma 4.5.2. It suffices to consider the case where $0 \leq p < q$, since in general when $\frac{p}{q} = j + \frac{p'}{q}$ and $r = a - 1 + \sum_{i<k} \lceil \frac{p}{q} i \rceil = a - 1 + j\binom{k}{2} + \sum_{i<k} \lceil \frac{p'}{q} i \rceil = r' + j\binom{k}{2}$, we get that

$$\begin{aligned} \mathrm{ex}_{\mathbb{Z}}(n,k,r) &= \mathrm{ex}_{\mathbb{Z}}(n,k,r' + j\binom{k}{2}) = \mathrm{ex}_{\mathbb{Z}}(n,k,r') + j\binom{n}{2} \\ &\geq \sum_{i<n} \lceil \frac{p'}{q} i \rceil + h(\lceil \frac{n}{q} \rceil) + j\binom{n}{2} = \sum_{i<n} \lceil \frac{p}{q} i \rceil + h(\lceil \frac{n}{q} \rceil). \end{aligned}$$

We will also use the fact that $\lceil x \rceil = \lceil \lceil xq \rceil / q \rceil$, for $x \in \mathbb{R}$ and q a positive integer.

Our recursive construction of a weight function w for K_n will satisfy the requirement that for every set A of z vertices ($0 \leq z \leq n$) we get $w(A) \leq \sum_{i<z} \lceil \frac{p}{q} i \rceil + h(\lceil \frac{z}{q} \rceil)$, with equality for $z = n$. Hence the fact that w is indeed (k,r)-dense immediately follows by considering the case $z = k$:

$$\sum_{i<k} \lceil \frac{p}{q} i \rceil + h(\lceil \frac{k}{q} \rceil) = (r - a + 1) + h(\ell + 1) \leq r.$$

We build the construction according to the size of the denominator q. In the base case $q = 1$, we get that $p = 0$, so that it suffices to give a weighting with $w(A) \leq h(z)$ and equality for $z = n$, but such a weighting is guaranteed by the definition of h. The corresponding matrix is $M(0/1)$.

For $q > 1$ we get that $p > 0$ since g.c.d.$(q, 0) = q > 1$. So let $q = p(t+2) - p'$ with $0 \leq p' < p$ and $0 \leq t = \lceil \frac{q}{p} \rceil - 2$. Let $n' = \lceil \frac{np}{q} \rceil$. Since g.c.d.$(p, p') = 1$ we can start with a weight function w' for $K_{n'}$ with $w'(A) \leq \sum_{i<z} \lceil \frac{p'}{p} i \rceil + h(\lceil \frac{z}{p} \rceil)$ for all z-sets A and equality for $z = n'$. From this we obtain the required weight function w on K_n, by partitioning $V(K_n) = V(K_{n'}) \cup V'$ and defining

$$w(uv) = \begin{cases} 0 & u, v \in V', \\ 1 & u \in V', v \in V(K_{n'}), \\ w'(uv) - t & u, v \in V(K_{n'}). \end{cases}$$

As an example $M(1/3)$ was constructed from $M(0/1)$ and $M(3/8)$ from $M(1/3)$ in this fashion, with $t = 1$ in both cases.

To see that this works, take any $A \subset V(K_n)$ of size z and let $|A \cap V(K_{n'})| = x$. We obtain

$$\begin{aligned}
w(A) &= x(z-x) - t\binom{x}{2} + w'(A \cap V(K_{n'})) \\
&\leq x(z-x) - t\binom{x}{2} + \sum_{i=0}^{x-1} \lceil \frac{p'}{p} i \rceil + h(\lceil \frac{x}{p} \rceil) \\
&= x(z-x) + 2\binom{x}{2} + \sum_{i=0}^{x-1} \lceil -(t+2)i + \frac{p'}{p} i \rceil + h(\lceil \frac{x}{p} \rceil) \\
&= x(z-1) + \sum_{i=0}^{x-1} \lceil -\frac{q}{p} i \rceil + h(\lceil \frac{x}{p} \rceil) \\
&= \sum_{i=0}^{x-1} \lceil z - 1 - \frac{q}{p} i \rceil + h(\lceil \frac{x}{p} \rceil).
\end{aligned}$$

Next we observe that

$$\sum_{i=0}^{z-1} \lceil \frac{p}{q} i \rceil = \sum_{i=0}^{\lceil \frac{p}{q}(z-1) \rceil - 1} \lceil z - 1 - \frac{q}{p} i \rceil,$$

since both sums count the number of lattice points in $R_z(p/q)$ that are below the line $y = \frac{p}{q}x$, except that the first sum is counting vertical first, while the second one goes horizontal first. Furthermore $\lceil z - 1 - \frac{q}{p} i \rceil = 0$ exactly when $\lceil \frac{p}{q}(z-1) \rceil \leq i \leq \lceil \frac{p}{q} z \rceil - 1$. Hence if $\lceil z/q \rceil \geq \lceil x/p \rceil$, then by the monotonicity of h and taking only the positive terms in the sum:

$$w(A) \leq \sum_{i=0}^{\lceil \frac{p}{q}(z-1) \rceil - 1} \lceil z - 1 - \frac{q}{p} i \rceil + h(\lceil \frac{z}{q} \rceil) = \sum_{i=0}^{z-1} \lceil \frac{p}{q} i \rceil + h(\lceil \frac{z}{q} \rceil).$$

If $\lceil z/q \rceil < \lceil x/p \rceil$, then $x > p\lceil z/q \rceil = p\lceil \lceil zp/q \rceil / p \rceil \geq \lceil zp/q \rceil$, so that

$$\begin{aligned}
w(A) &\leq \sum_{i=0}^{\lceil \frac{p}{q}(z-1) \rceil - 1} \lceil z - 1 - \frac{q}{p} i \rceil + \sum_{i=\lceil \frac{p}{q} z \rceil}^{x-1} \lceil z - 1 - \frac{q}{p} i \rceil + h(\lceil \frac{x}{p} \rceil) \\
&\leq \sum_{i=0}^{z-1} \lceil \frac{p}{q} i \rceil + (x - \lceil \frac{p}{q} z \rceil)(-1) + h(\lceil \frac{x}{p} \rceil).
\end{aligned}$$

Hence we are done if $x - \lceil \frac{p}{q} z \rceil \geq h(\lceil \frac{x}{p} \rceil) - h(\lceil \frac{z}{q} \rceil)$. However $h(\lceil \frac{x}{p} \rceil) - h(\lceil \frac{z}{q} \rceil) \leq \lceil \frac{x}{p} \rceil - \lceil \frac{z}{q} \rceil$ by the Lipschitz property of h. So we obtain the desired result if $x - \lceil \frac{x}{p} \rceil \geq \lceil \frac{p}{q} z \rceil - \lceil \frac{z}{q} \rceil = \lceil \frac{p}{q} z \rceil - \lceil \lceil \frac{zp}{q} \rceil / p \rceil$. This is true since the function $x - \lceil \frac{x}{p} \rceil$ is nondecreasing on integers and we assumed that $x > \lceil zp/q \rceil$.

Finally we need to check for equality when $z = n$. In that case $x = n' = \lceil \frac{np}{q} \rceil$ and thus, following the first chain of inequalities, we obtain:

$$
\begin{aligned}
w(K_n) &= \sum_{i=0}^{n'-1} \lceil n - 1 - \frac{q}{p} i \rceil + h(\lceil n'/p \rceil) \\
&= \sum_{i=0}^{\lceil \frac{np}{q} \rceil - 1} \lceil n - 1 - \frac{q}{p} i \rceil + h(\lceil \lceil \frac{np}{q} \rceil / p \rceil)) \\
&= \sum_{i=0}^{\lceil \frac{p}{q}(n-1) \rceil - 1} \lceil n - 1 - \frac{q}{p} i \rceil + h(\lceil n/q \rceil) \\
&= \sum_{i=0}^{n-1} \lceil \frac{p}{q} i \rceil + h(\lceil n/q \rceil).
\end{aligned}
$$

This finishes the proof of Lemma 4.5.2 $\square$

The construction yields the Turán graph $T_{n,q}$ when $p = q - 1$. For $p = 1$ we get a graph that is almost bipartite: $V = V_1 \cup V_2$, $|V_1| = \lceil n/q \rceil$ and all V_1V_2 edges have weight 1, all edges within V_2 have weight 0 and all edges within V_1 have weight $2 - q$, as indicated for example by $M(1/3)$. The constructions given by Figure 4.3 are all optimal in many cases. For example from Theorem 4.3.3 we will see that $M(3/8)$ yields an optimal construction when $(k, r) = (9, 17), (10, 21), (11, 25), (12, 30), \ldots$.

4.6. Some exact values of $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$

Recall that a *star* is a graph $K_{1,n}$ where we will allow $n \geq 0$ in the context of this chapter. For each star we call the vertex of degree n its *center*, the other vertices are the *leaves*. If $n = 1$ we designate only one of the two vertices to be the center and the other to be the leaf.

So far the zero function is the only linear weighting function we have seen. In this section we will explore other linear weighting functions, thus obtaining quick proofs for Theorems 4.3.3 and 4.3.4. Consider $h_a(x) = \min\{x - 1, a - 1\}$. The first two properties (L1) and (L2) of a linear weighting function are easily checked and the required weighting is given by a star on $\min\{x, a\}$ vertices, with the remaining vertices being isolated. Notice that h_1 is the zero function.

Proof of Theorem 4.3.3. Using $h = h_a$ the lower bound is immediate from Lemma 4.5.2, since $a - 1 \geq h(\lceil \frac{n}{q} \rceil) \geq h(\lceil \frac{k}{q} \rceil) = h(\ell + 1) \geq h(a) = a - 1$.

The upper bound for $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$ follows from Theorem 4.3.2. $\square$

In the case when $a \leq \ell \leq 2a - 1$ we let

$$
h_{a,\ell}(x) = \begin{cases} x - 1, & \text{for } x \leq a, \\ a - 1, & \text{for } a \leq x \leq \ell + 1, \\ x + a - \ell - 2, & \text{for } \ell + 1 \leq x \leq 2a, \\ 3a - \ell - 2, & \text{for } 2a \leq x. \end{cases}
$$

So $h_{a,\ell}(x) = h_a(x)$ for $x \le \ell+1$ and $h_{a,\ell}(x) = 2a-\ell-1+h_a(x-a)$ for $x \ge \ell+1$. Since the monotonicity and Lipschitz property of $h_{a,\ell}$ are again straight forward we only have to verify condition (L3) to see that $h_{a,\ell}$ is a linear weighting function. For this we will use two stars on a vertices and connect their centers by an edge of weight $a-\ell \le 0$. For $x > 2a$ the remaining vertices are isolated, and for $x < 2a$ we take the obvious subgraph. It can be seen that this satisfies condition (L3).

Proof of Theorem 4.3.4. The upper bound is exactly Theorem 4.3.2. The lower bound follows from Lemma 4.5.2 with $h = h_{a,\ell}$. It suffices to notice that $h_{a,\ell}(\ell+1) = a-1$ and that $2a \ge \lceil n/q \rceil \ge \lceil k/q \rceil = \ell+1$. □

4.7. Proof of the lower bound

In this section we prove Theorem 4.3.5. We use the construction of Section 4.5 to obtain a probably optimal lower bound. To be able to apply Lemma 4.5.2 we show that for fixed values of r and k, with $0 \le r \le k-2$, $\mathrm{ex}_{\mathbb{Z}}(x,k,r)$ is a linear weighting function. However $\mathrm{ex}_{\mathbb{Z}}(x,k,r)$ is only defined for $x \ge k$, hence we will let

$$h(x) = h(x,k,r) = \begin{cases} x-1, & \text{for } x \le r+1, \\ r, & \text{for } r+1 \le x \le k, \\ \mathrm{ex}_{\mathbb{Z}}(x,k,r), & \text{for } k \le x. \end{cases}$$

Observe that $\mathrm{ex}_{\mathbb{Z}}(k,k,r) = r$, so that h is well-defined. In the rest of this section we show that h is a linear weighting function. Once we know this (via Lemmas 4.7.1, 4.7.2 and 4.7.4) the proof of Theorem 4.3.5 is immediate:

Proof of Theorem 4.3.5. Since $a-1 \le (\ell+1)-2$ we have that $h(x,\ell+1,a-1)$ is a linear weighting function. Furthermore $h(\ell+1,\ell+1,a-1) = a-1$, so that we can apply Lemma 4.5.2. Because $\lceil n/q \rceil \ge \lceil k/q \rceil = \ell+1$ we get that in the interesting range $h(x) = \mathrm{ex}_{\mathbb{Z}}(x,\ell+1,a-1)$ and the inequality follows. □

In the rest of this section (and in the next section, too), we suppose that $r \le k-2$. We obtain $p=0$, $q=1$, $a=r+1$ and $\ell = k-1$. First, we consider the case $2(r+1) \le k$. Since this is the the case of Theorem 4.3.3, we have $h(n) = h_r(n)$ and there is nothing to prove.

Also for $2(r+1) > k$ and $n \le 2(r+1)$ Theorem 4.3.4 can be applied and we get that $h(n) = h_{r+1,k-1}(n)$. Therefore all 3 properties (L1)–(L3) hold in this range.

Thus from now on we can assume that $n > 2r+2 > k$ and $h(n) = \mathrm{ex}_{\mathbb{Z}}(n,k,r)$.

Notice that Theorem 4.3.2 yields (for all $0 < r+1 < k \le n$)

$$\mathrm{ex}_{\mathbb{Z}}(n,k,r) \le n+r-k < n-1. \tag{4.7.1}$$

Hence the Lipschitz property of h is an immediate consequence of (4.3.4). We obtain

$$\mathrm{ex}_{\mathbb{Z}}(n+1,k,r) \le \mathrm{ex}_{\mathbb{Z}}(n,k,r) + \lfloor \frac{2\mathrm{ex}_{\mathbb{Z}}(n,k,r)}{n-1} \rfloor \le \mathrm{ex}_{\mathbb{Z}}(n,k,r)+1.$$

For the other two properties (L1), (L3) we need to prove two Lemmas. For a given weighting w let $h_w(x) = \max\{w(A) : |A| = x\}$ and recall that the weight of a vertex was defined by $w(v) = w(\{uv : u \in V\})$.

Lemma 4.7.1. *If $0 < r+1 < k < 2r+2$, $k \le n$, $\mathrm{ex}_{\mathbb{Z}}(n,k,r) > \mathrm{ex}_{\mathbb{Z}}(n-1,k,r)$ and w is a weighting of K_n achieving $\mathrm{ex}_{\mathbb{Z}}(n,k,r)$ then every vertex has positive weight. Furthermore if $n > 2r+2$, then $0 = h_w(1) \le h_w(2) \le \cdots \le h_w(k) = r$.*

Proof. For a given vertex v the positivity of $w(v)$ follows from

$$\mathrm{ex}_{\mathbb{Z}}(n,k,r) > \mathrm{ex}_{\mathbb{Z}}(n-1,k,r) \ge w(K_n - v) = \mathrm{ex}_{\mathbb{Z}}(n,k,r) - w(v).$$

So now assume that $n > 2r+2$. The first equality is trivial and the second one follows from the choice of w. For the chain of inequalities, suppose that i is maximal such that $h_w(i) > h_w(i+1)$ and let A be a set achieving $h_w(i)$. Then for every vertex $v \notin A$ we have $w(\{uv : u \in A\}) \le h_w(i+1) - h_w(i)$, so that

$$\begin{aligned} i &\le \sum_{u \in A} w(u) \le 2h_w(i) + (n-i)(h_w(i+1) - h_w(i)) \\ &= 2h_w(i+1) + (n-i-2)(h_w(i+1) - h_w(i)) \\ &\le 2r + (n-i-2)(-1) = i + (2r+2-n) < i, \end{aligned}$$

a contradiction. □

Lemma 4.7.1 will enable us to prove

(4.7.2) *If $2r+2 < n < m$ and $h(n) > h(m)$, then also $h(n-1) > h(m)$.*

But then the monotonicity of h is straight forward. If $h(n+1) < h(n)$ then we can apply (4.7.2) repeatedly and obtain that $h(n+1) < h(2r+2)$. This is a contradiction since we know that $h(n+1) \ge h_{r+1,k-1}(n+1) \ge h_{r+1,k-1}(2r+2) = h(2r+2)$.

To prove (4.7.2) we may assume that $h(n-1) < h(n)$, since otherwise there is nothing to show. Let w be any weighting achieving $h(n)$ and extend this to a weighting w' of K_m by making all new edges weight 0. Notice that for any set A of k vertices in K_m

$$w'(A) = w(A \cap K_n) \le h_w(|A \cap K_n|) \le r$$

by Lemma 4.7.1. Hence w' is (k,r)-dense contradicting $h(n) > h(m)$.

The rest of the section is devoted to the proof of property (L3). It is clear that in general we may not be able to take just any weighting achieving $\mathrm{ex}_{\mathbb{Z}}(n,k,r)$ for this purpose. For example we need $w(uv) = w(\{u,v\}) \le h(2) = 1$, so that we can't have multiple edges.

We call a maximal vertex set A with the property $w(A) \ge |A| - 1$ a *component*. Notice that components need not be disjoint, but every vertex is in some component since $w(\{v\}) = 0$.

Lemma 4.7.2. *If $0 < r+1 < k < 2r+2 < n$, $\mathrm{ex}_{\mathbb{Z}}(n,k,r) > \mathrm{ex}_{\mathbb{Z}}(n-1,k,r)$ and w achieves $\mathrm{ex}_{\mathbb{Z}}(n,k,r)$ then every component A has $|A| \le r+1$ and $w(A) = |A| - 1$.*

Proof. Let A be a component. Then (4.7.1) implies $|A| < k$ and by Lemma 4.7.1 we get that $|A| \le r+1$. Next notice that for every vertex $v \notin A$

$$w(\{uv : u \in A\}) = w(A+v) - w(A) \le (|A+v| - 2) - w(A) = |A| - w(A) - 1,$$

so that for $w(A) \geq |A|$ we would obtain

$$\begin{aligned} w(A) + w(\{uv : u \in A, v \notin A\}) &\leq w(A) + (n - |A|)(|A| - w(A) - 1) \\ &= (n - |A|)(|A| - 1) - w(A)(n - |A| - 1) \\ &\leq (n - |A|)(|A| - 1) - |A|(n - |A| - 1) \\ &= 2|A| - n < 0. \end{aligned}$$

But that is impossible, since then $w(K_n \backslash A) > w(K_n) = \mathrm{ex}_{\mathbb{Z}}(n, k, r)$. Depending on $|K_n \backslash A|$ this would be a contradiction either to the monotonicity of $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$ or Lemma 4.7.1. □

Definition 4.7.3. A weighting w is called *simple* if
(S1) $w(e) \leq 1$ for every edge e,
(S2) the edges of weight 1 induce a disjoint union of stars, and
(S3) $w(v) > 0$, unless v is *isolated* (i.e., only incident to edges of weight zero).
A set of of vertices A is called *simple* if
(A1) A induces a star with center c,
(A2) For $u \in A$ and $v \notin A$ we have $w(uv) = 0$ unless $u = c$ and $w(cv) < 0$, and
(A3) $w(c) > 0$ for $|A| > 1$, $w(c) = 0$ otherwise.

In a simple weighting the only edges of negative weight are between the centers of stars. For example the weightings provided for h_a and $h_{a,\ell}$ are simple. Simple sets are always disjoint, so that a weighting is simple if and only if every vertex is in some simple set.

Lemma 4.7.4. *If $0 < r + 1 < k \leq n$, then $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$ can be achieved by a simple weighting w. Moreover h_w has the Lipschitz property and is nondecreasing.*

Lemma 4.7.4 will enable us to verify (L3) for h by choosing w to be the simple weighting guaranteed. Notice that h was chosen to increase as much as possible initially. Therefore we get that $h_w(x) \leq h(x)$ for $x \leq k$, since h_w also has the Lipschitz property, is nondecreasing and satisfies $h_w(k) \leq r = h(k)$. Thus for all sets A of size at most k we have $w(A) \leq h_w(|A|) \leq h(|A|)$. For $|A| \geq k$ we trivially get $w(A) \leq \mathrm{ex}_{\mathbb{Z}}(|A|, k, r) = h(|A|)$.

Proof of Lemma 4.7.4. For a given simple weighting w we see that $h_w(x + 1) \leq h_w(x) + 1$ by taking a maximum weight set of size $x+1$ and removing a leaf of one of the stars. If the set consisted only of centers then all internal edges were nonpositive and we can remove any vertex. Similarly $h_w(x + 1) \geq h_w(x)$ since we can take a set achieving $h_w(x)$ and add an isolated vertex or a leaf of some star. If only centers remain, then adding in any one of those will increase $h_w(x)$ by Property (S3). Thus it only remains to construct such a simple weighting w.

If $2r + 2 \leq k - 1$ or $n \leq 2r + 2$ we can use the simple weightings from h_{r+1} or $h_{r+1,k-1}$ respectively, so we may assume that $n > 2r + 2 > k$.

In the case when $\mathrm{ex}_{\mathbb{Z}}(n, k, r) = \mathrm{ex}_{\mathbb{Z}}(n - 1, k, r)$ we use the simple weighting w for K_{n-1} and add in an isolated vertex to obtain the simple weighting w'. We get $h_{w'}(k) = \max\{h_w(k), h_w(k - 1)\} = h_w(k) \leq r$, so that $h_{w'}$ achieves $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$.

So we may also assume that $\mathrm{ex}_{\mathbb{Z}}(n, k, r) > \mathrm{ex}_{\mathbb{Z}}(n - 1, k, r)$. Hence, by Lemma 4.7.1, every vertex has positive weight and $h_w(x) \leq r$ for $x \leq k$ for every weighting achieving $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$. Take a weighting w achieving $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$ with a minimum number of vertices not in any simple set. Let v be such a vertex, since if there is no such vertex then w is simple and we are done. Let A be a component containing v, so by Lemma 4.7.2 $w(A) = |A| - 1$. We will create a new weighting w' by making A simple: Replace w on A by a star centered at v and for $y \notin A$ define $w'(xy) = 0$ if $x \in A - v$ and $w'(vy) = \sum\{w(uy) : u \in A\} = w(A + y) - w(A) \leq (|A + y| - 2) - (|A| - 1) = 0$. Hence $w'(K_n) = w(K_n)$ and v is now in the simple set A. Furthermore if S used to be a simple set, then $S \setminus A$ still is: if the center of S was in A then $S \subset A$, since A was a component. Hence $S \setminus A$ is still a star. Furthermore all new edges from $S \setminus A$ are nonpositive, and all leafs still have weight 0 on all outgoing edges.

Let K be a set on k vertices. If $v \notin K$, then

$$w'(K) = w'(K \setminus A) = w(K \setminus A) \leq h_w(|K \setminus A|) \leq r.$$

If on the other hand $v \in K$, then let $K' = K \cup A$ and observe that by (4.7.1)

$$w'(K) + |K'| - k = w'(K') = w(K') \leq |K'| - k + r.$$

In either case $w'(K) \leq r$, so that $h_{w'}(k) \leq r$, in contradiction to our choice of w. $\square$

4.8. More on the linear term

The aim of this section is the proof of Theorem 4.3.6. Theorem 4.3.5 suggests that to solve the problem of $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$ for all values it will be crucial to study the purely linear term. Lemma 4.7.4 tells us what good constructions will essentially look like, and it also "simplifies" the proof of upper bounds. For example

Theorem 4.8.1. *If $r \leq k - 2$ then $\mathrm{ex}_{\mathbb{Z}}(n, k, r) \leq \lfloor \frac{r}{r+1} n \rfloor$ with equality when $r = k - 2$.*

Proof. Since $k \geq r + 1$ the inequality is clear for $n = k$. For $n > k$ there is nothing to show when $\mathrm{ex}_{\mathbb{Z}}(n, k, r) = \mathrm{ex}_{\mathbb{Z}}(n - 1, k, r)$. Otherwise we use Lemma 4.7.4 to obtain a simple weighting achieving $\mathrm{ex}_{\mathbb{Z}}(n, k, r)$ and observe that every simple set is also a component. Hence by Lemma 4.7.2 every one of the stars has at most $r + 1$ vertices. Therefore the most edges we can have is taking disjoint stars on $r+1$ vertices without negative weights and this is reflected in the formula. Furthermore for $r = k - 2$ this weighting is achievable. $\square$

We conjecture that the weighting function g given by a disjoint union of one star on v vertices and as many stars on u vertices $(1 \leq u \leq v)$ as possible, plus one remainder star on at most u vertices, is essentially optimal. (All other edges have weight 0, i.e., this is a simple weighting as it was defined in the previous section). Let

$$g(x, u, v) := \begin{cases} x - 1, & \text{for } x \leq v, \\ v - 1 + \lfloor \frac{u-1}{u}(x - v) \rfloor, & \text{for } v \leq x. \end{cases}$$

Lemma 4.8.2. *For the case $0 < r + 1 < k < 2r + 2$*

$$\mathrm{ex}_{\mathbb{Z}}(n, k, r) \geq \max\{h_{r+1,k-1}(n), g(n, u, v)\},$$

where $u = \lfloor (k-1)/(k-r-1) \rfloor$ *and* $v = k - 1 - u(k-r-2)$.

Proof. It was shown in Section 4.6 with a construction that $h_{r+1,k-1}(n)$ is a lower bound. Concerning the second term, consider the simple weighting defined before the Lemma. Since $v - u = k - 1 - u(k-r-1) \geq 0$, we get $v \geq u$. Also we have $u \geq 1$. Finally, one can check that this is indeed (k,r)-dense by observing that for $x \geq v$

$$g(x,u,v) = v - 1 + \lfloor \frac{u-1}{u}(x-v) \rfloor = r + \lfloor \frac{u-1}{u}(x-k+1) \rfloor,$$

and that $k \geq v$, so that $g(k,u,v) = r$. □

We believe that equality always holds. We were able to prove, with some case analysis, that for $k < 11$ equality holds in Lemma 4.8.2 for all $n \geq k$. In the rest of this section we show that Lemma 4.8.2 also gives the optimum when n is sufficiently large.

Proof of Theorem 4.3.6. If $k \geq 2r+2$, that is when $u = 1$, then the result follows from Theorem 4.3.3. So we may assume that $0 < r+1 < k < 2r+2$ and hence $u \geq 2$.

By Lemma 4.7.4 there exists an optimal simple weighting w. As in Lemma 4.7.1 let $h(x)$ denote the maximum possible weight an x-element set can have, where we dropped the subscript, since the underlying weight function is now fixed. For example $h(0) = h(1) = 0$, $h(2) = 1$. We already know from Lemma 4.8.2 that

$$h(n) = \mathrm{ex}_{\mathbb{Z}}(n,k,r) \geq r + \lfloor \frac{u-1}{u}(n-k+1) \rfloor. \tag{4.8.1}$$

The proof of the upper bound for $h(n)$ consists of two steps. First, we will prove in Lemma 4.8.3 that the weighting w contains at least k vertices in *good u-stars*, that is stars on u vertices without negative edges at the center vertex. Then Lemma 4.8.4 below will finish the proof of Theorem 4.3.6. □

Lemma 4.8.3. *If* $0 < r+1 < k < 2r+2$, $n \geq 3kr$ *and* w *is a simple weighting satisfying (4.8.1) then at least* k *vertices are in good* u*-stars.*

Proof. Let $\mathcal{A}$ be the set of components (in the sense of Lemma 4.7.2) given by w. Every component $A \in \mathcal{A}$ induces a star and we denote its center by v_A. Furthermore, let n_A denote the absolute value of the total sum of negative weights of the edges adjacent to the center v_A. The total contribution of the edges with at least one vertex in A to the weight of K_n is exactly $w(v_A) = (|A|-1) - n_A$. For components on more than one vertex this should be strictly positive, hence

$$n_A \leq |A| - 2 \quad \text{for} \quad |A| \geq 2. \tag{4.8.2}$$

From Lemma 4.7.4 we know that the function $h(x)$ is nondecreasing with $h(k) \leq r$, so that every star has at most r edges, that is $|A| \leq r+1$. Now divide $\mathcal{A}$ into four parts:

$\mathcal{A}_0 = \{A \in \mathcal{A} : |A| < u\}$ (small stars, including isolated vertices),
$\mathcal{A}_1 = \{A \in \mathcal{A} : |A| = u \text{ and } n_A > 0\}$ (bad u-stars),
$\mathcal{A}_2 = \{A \in \mathcal{A} : |A| = u \text{ and } n_A = 0\}$ (good u-stars), and
$\mathcal{A}_3 = \{A \in \mathcal{A} : |A| > u\}$ (big stars).

Let $V_i = \cup\{A : A \in \mathcal{A}_i$ be the set of all the vertices in the corresponding part. Our first aim is to prove that there are only few big stars:

$$|V_3| \le r(k-1). \tag{4.8.3}$$

Indeed, split $\mathcal{A}$ into r parts $\mathcal{B}_1, \ldots, \mathcal{B}_r$ such that each $\mathcal{B}_i$ contains no negative edge. This can be done, since by (4.8.2) $n_A \le r-1$. (We can define these parts recursively, putting the first member of $\mathcal{A}$ into $\mathcal{B}_1$. So suppose we have already split some members of $\mathcal{A}$. For the next member, call it A, we can always find some $\mathcal{B}_i$ such that A is not joined to $\cup\mathcal{B}_i$ by a negative edge, since at most $r-1$ negative edges are incident to v_A). We claim that the members of $\mathcal{B}_i \cap \mathcal{A}_3$ cover at most $k-1$ points, which immediately gives (4.8.3).

Indeed, suppose to the contrary that there exist $A_1, \ldots, A_t \in \mathcal{B}_i \cap \mathcal{A}_3$ such that $|A_1 \cup \cdots \cup A_t| \ge k$. Consider a minimal such union, i.e., $|A_1 \cup \cdots \cup A_{t-1}| < k$. Then there exists a k-set K containing $A_1 \cup \cdots \cup A_{t-1}$ and the center of A_t, which is contained in $A_1 \cup \cdots \cup A_t$. Thus $w(K) = k - t$. This is at most r, implying that $t - 1 \ge k - r - 1$ and hence

$$|A_1 \cup \cdots \cup A_{t-1}|/(t-1) \le (k-1)/(k-r-1).$$

This implies that $|A_i| \le u$ (for some $i \le t-1$), a contradiction.

Next we observe that for all stars $A \notin \mathcal{A}_3$ we have $w(A) \le (u-1)A/u$. Therefore we get that $w(B_i) \le g(|B_i|, u, v)$, where B_i stands for the vertex set $\cup\mathcal{B}_i$. Taking into account the small members of $\mathcal{B}_i$, we actually obtain

$$\begin{aligned} w(B_i) &\le g(|B_i \setminus V_0|, u, v) + \frac{u-2}{u-1}|B_i \cap V_0| \\ &\le \frac{u-1}{u}|B_i| + (k-r-1)/u - |B_i \cap V_0|/u(u-1), \end{aligned}$$

since $r - \frac{u-1}{u}(k-1) \le (k-r-1)/u$.

Adding up these inequalities for all $1 \le i \le r$ and considering the fact that every member in $\mathcal{A}_1$ is adjacent to a negative edge we get

$$h(n) = w(K_n) \le \frac{u-1}{u}n + r(k-r-1)/u - |V_0|\frac{1}{u(u-1)} - \frac{1}{2}|V_1|/u.$$

This upper bound and the lower bound (4.8.1) yields $|V_0 \cup V_1| \le 2kr - k + r$. Together with (4.8.3) we obtain that $|V_2| \ge n - 3kr + k \ge k$, the desired k vertices in good u-stars. $\square$

Lemma 4.8.4. *If $0 < r+1 < k < 2r+2$, $k \le n$ and w is an optimal simple weighting containing at least k vertices in good u-stars, then for all $0 \le x \le n$*

$$h_w(x) \le r + \lfloor \frac{u-1}{u}(x-k+1) \rfloor.$$

Proof. We proceed by induction on x. For $x = 0$ a simple computation yields that $0 \le r + \lfloor (u-1)(-k+1)/u \rfloor$. For $x > 0$ it is sufficient to prove the upper bound for numbers of the form $x = k + ju$ where j is a (not necessarily positive) integer, i.e.,

$$h(k+ju) \le r + j(u-1). \tag{4.8.4}$$

Indeed, the Lipschitz property of h (implied by Lemma 4.7.4), implies the remaining values.

So consider a set K on $k + ju$ elements with $w(K) = h(k + ju)$. We may suppose that K meets the minimum number of stars, say with vertex sets $A_1, \dots, A_c$. Abbreviate the center of A_i by v_i and observe that each v_i is contained in K (for $1 \le i \le c$). Indeed, if $v_i \notin K$, but $y \in A_i \cap K$ then $w(K - y) = w(K)$, and the result follows by induction. Thus we can also assume that $A_1 \cup \dots \cup A_{c-1} \subset K \subset A_1 \cup \dots \cup A_c$.

First we prove (4.8.4) for $j \le 0$. Observe that by the choice of K and the fact that there are at least k vertices in good u-stars we can find $|j|$ good u-stars $A'_1, \dots A'_{|j|}$ disjoint from K. As $w(K) + |j|(u-1) = w(K \cup A'_1 \cup \dots \cup A'_{|j|}) \le r$ we obtain (4.8.4).

To prove (4.8.4) for positive j, suppose to the contrary, that $w(K) > r + j(u-1)$. By hypothesis we have $h(k + (j-1)u) \le r + (j-1)(u-1)$. Then the Lipschitz property implies that $h(k + ju) \le h(k + (j-1)u) + u$, hence we have $w(K) = r + j(u-1) + 1$ and $h(k + (j-1)u + i) = r + (j-1)(u-1) + i$ for $0 \le i \le u$.

Let $M = A_1 \cup \dots \cup A_c$. This set contains K, and we let $|M| = |K| + m$. As M can be obtained from K by adding m vertices of degree (and weight) one we have that $w(M) = w(K) + m$. Then, again, the Lipschitz property implies that

$$h(k + (j-1)u + i) = r + (j-1)(u-1) + i \quad \text{for all} \quad 0 \le i \le u + m. \tag{4.8.5}$$

Let q_i denote the sum of the absolute values of the weights of the edges from v_i to the other centers v_t ($1 \le t \le c$). We have that $w(M \backslash A_i) = w(M) - (|A_i| - 1) + q_i > w(M) - |A_i|$. This and (4.8.5) imply that $|A_i| > u + m$. Thus we can use the hypothesis that our result is true for $M \setminus A_i$. We obtain that

$$\begin{aligned} r + j(u-1) + m + 1 = h(|M|) = w(M) &= w(M \setminus A_i) + (|A_i| - 1) - q_i \\ &\le r + \lfloor \frac{u-1}{u}(k + ju + m - |A_i| - (k-1)) \rfloor + |A_i| - 1 - q_i. \end{aligned}$$

Multiplying by u, dropping the floors, and rearranging terms we get that

$$u + 1 + m + uq_i \le |A_i|.$$

Adding this up for all i we obtain

$$c(u+1) + cm + 2uQ \le \sum |A_i| = k + ju + m, \tag{4.8.6}$$

where Q is the absolute value of $w(\{v_1, \dots, v_c\})$. On the other hand, from (4.8.5) and from the structure of M we can calculate Q as follows:

$$r + j(u-1) + 1 + m = w(M) = |M| - c - Q = k + ju + m - c - Q.$$

Rearranging we have $Q + c = (k - r - 1) + j$. Subtracting $(u+1)$ times this equality from the two sides of (4.8.6) we obtain that

$$(c-1)m + Q(u-1) + (j-1) \le (k-1) - (u+1)(k-r-1).$$

However, here on the left-hand side all the three terms are nonnegative ($c, u, j \ge 1$ and $m, Q \ge 0$) and the right hand side is negative. This contradiction shows that our assumption $h(k + ju) > r + j(u-1)$ was wrong, finishing the proof of (4.8.4). $\square$

4.9. The asymptotic density for nonnegative integer weights

In this section we investigate $\mathrm{ex}_W(n,k,r)$ for $W=\mathbb{Z}^+$ and for $W=[i]:=\{0,1,2,\ldots,i\}$. Let $||M(p/q)||$ denote the absolute value of the smallest entry of the matrix $M(p/q)$.

Theorem 4.9.1. *Suppose that $r=r'+j\binom{k}{2}$ with $0\le r'<\binom{k}{2}$, $m(k,r')=p/q$ and $||M(p/q)||\le j$. Then*

$$\mathrm{ex}_{\mathbb{Z}^+}(k,r)=\mathrm{ex}_{[j+1]}(k,r)=j+p/q.$$

Proof. Indeed, the upper bound of Theorem 4.3.2 is also valid for these cases. The lower bound follows from the observation that the construction in Lemma 4.5.2 (using the zero function for h) has maximum weight 1 and minimum weight $-||M(p/q)||$ and we can add weight j to every edge. □

It can be checked that for $0\le p<q$

$$||M(p/q)||\le q-p-1. \tag{4.9.1}$$

This is best possible for example when $p=1$ or $p=q-1$, but in general we expect that the negative weight is only $\mathrm{O}(\log q)$.

Proof of (4.9.1). Induction on p. For $p=0$ and $q=1$ the construction has all weights 0 and the bound holds with equality. For $p>0$, let $q=(t+2)p-p'$ with $0\le p'<p$. We have a negative weight of at most $||M(p'/p)||\le p-p'-1$. Thus, for $0<p<q$, we have a negative weight of at most $p-p'-1+t\le p(t+2)-p'-p-1=q-p-1$. □

A (k,r)-dense graph G for $W=\mathbb{Z}^+$ has no edges of weight more than r. The following observation says that its maximum weight can be bounded by $\ell:=\lfloor r/\lfloor k^2/4\rfloor\rfloor$.

$$\mathrm{ex}_{\mathbb{Z}^+}(k,r)=\mathrm{ex}_{[\ell]}(k,r) \tag{4.9.2}$$

Proof. Indeed, let H be the graph formed by the of edges of weight at least $\ell+1$. This graph does not contain the complete bipartite graph $K_{\lfloor k/2\rfloor,\lceil k/2\rceil}$ as a subgraph. Hence a theorem of Kővári, T. Sós and Turán [83] implies that $|E(H)|=\mathrm{O}(n^{1-2/k})$. So the edges of high weight contribute only $\mathrm{o}(n^2)$ to the total weight of G. □

One can formulate the following consequence of (4.9.2).

$$\mathrm{ex}_{\mathbb{Z}^+}(k,r)=\mathrm{ex}_{[2i-1]}(k,r)\quad\text{for}\quad r\le i\binom{k}{2}.$$

We conjecture that $2i-1$ can be replaced by i or $i+1$.

For fixed k Theorem 4.9.1 immediately implies the value of $\mathrm{ex}_{\mathbb{Z}^+}(k,r)$ for all but finitely many r. For values $k-2<r<\lfloor k^2/4\rfloor$ (4.9.2) implies that $\mathrm{ex}_{\mathbb{Z}^+}(k,r)=0$, while $\mathrm{ex}_{\mathbb{Z}}(k,r)\ge 1/(k-1)$. So for these values of (k,r) negative weights help. The first such case is $(k,r)=(4,3)$, given by $M(1/3)$.

We close this section by reviewing $\mathrm{ex}_{\mathbb{Z}}(n,k,r)$ for $k\le 5$.

The cases for $k=3$ have been settled in [17] where it is shown that $\mathrm{ex}_{\mathbb{Z}}(n,3,1)=\lfloor n/2\rfloor$ and $\mathrm{ex}_{\mathbb{Z}}(n,3,2)=\lfloor n^2/4\rfloor$ which together with the trivial $\mathrm{ex}_{\mathbb{Z}}(n,3,0)=0$ shows that all these values can be obtained by nonnegative weights: A matching, $T_{n,2}$, the empty graph and for higher values of r we take translates of those.

For $k = 4$ and $r \leq 3$ we can have at most one edge of weight 2 in the positive weight case. $\mathrm{ex}_{\mathbb{Z}^+}(n, 4, 0) = 0$, $\mathrm{ex}_{\mathbb{Z}^+}(n, 4, 1) = 1$ and $\mathrm{ex}_{\mathbb{Z}^+}(n, 4, 2) = \lfloor 2n/3 \rfloor$ follows immediately. For $r = 3$ we also get that the weight 1 edges can not contain a C_4, so that $\mathrm{ex}_{\mathbb{Z}^+}(n, 4, 3) \leq \mathrm{ex}(n, C_4) = \frac{1}{2}n^{3/2} + \mathrm{O}(n^{4/3})$ by a result of Erdős, Rényi and Sós [46] and independently Brown [19]. Hence $\mathrm{ex}_{\mathbb{Z}^+}(4, r) = 0$ for all $r \leq 3$, but $\mathrm{ex}_{\mathbb{Z}}(4, 3) = 1/3$ which is the only case for $k = 4$. $\mathrm{ex}_{\mathbb{Z}}(4, 4) = 1/2 = \mathrm{ex}_{\mathbb{Z}^+}(4, 4)$ and $\mathrm{ex}_{\mathbb{Z}}(4, 5) = 2/3 = \mathrm{ex}_{\mathbb{Z}^+}(4, 5)$.

For $k = 5$, the cases where $\mathrm{ex}_{\mathbb{Z}}(5, r)$ is bigger than $\mathrm{ex}_{\mathbb{Z}^+}(5, r)$ are $r = 4, 5, 14$. The arguments for $r = 4, 5$ are similar to those above as the edges of weight 1 cannot contain a $K_{2,3}$, so that $\mathrm{ex}_{\mathbb{Z}^+}(5, r) = 0$, whereas $\mathrm{ex}_{\mathbb{Z}}(5, 4) = 1/4$ and $\mathrm{ex}_{\mathbb{Z}}(5, 5) = 1/3$. Finally, we have $\mathrm{ex}_{\mathbb{Z}}(5, 14) = 5/4$ while one can prove that $\mathrm{ex}_{\mathbb{Z}^+}(5, 14) = \mathrm{ex}_{\{0,1,2\}}(5, 14) = 7/6$. But this, and other results about $\mathrm{ex}_{\mathbb{Z}^+}$, are the subject of a paper by Füredi and Simonovits [57].

References

[1] A. Aggarwal, *The art gallery theorem: its variations, applications and algorithmic aspects*, PhD Thesis, The Johns Hopkins University, Baltimore, 1984.

[2] R. Ahlswede, I. Althöfer, *The asymptotic behaviour of diameters in the average*, J. Combin. Theory Ser. B **61** (1994), 167–177.

[3] R. Ahlswede, N. Cai, Z. Zhang, *Rich colorings with local constraints*, J. Combin. Inform. System Sci. **17** (1992), 203–216.

[4] R. Ahlswede, G. O. H. Katona, *Contributions to the geometry of Hamming spaces*, Discrete Math. **17** (1977), 1–22.

[5] N. Alon, M. Tarsi, *Covering multigraphs by simple circuits*, SIAM J. Algebraic Discrete Methods **6** (1985), 345–350.

[6] B. Alspach, L. Goddyn, CQ. Zhang, *Graphs with the circuit cover property*, Trans. Amer. Math. Soc. **344** (1994), 131–154.

[7] I. Althöfer, T. Sillke, *An "average distance" inequality for large subsets of the cube*, J. Combin. Theory Ser. B **56** (1992), 296–301.

[8] I. Anderson, *Combinatorial designs and tournaments*, Oxford Lecture Series in Mathematics and its Applications, 6, The Clarendon Press, Oxford University Press, New York, 1997.

[9] K. Aoshima, M. Iri, *Comments on F. Hadlock's paper: "Finding a maximum cut of a planar graph in polynomial time"*, SIAM J. Comput **6** (1977), no. 1, 86–87.

[10] D. Avis, G.T. Toussaint, *An efficient algorithm for decomposing a polygon into star-shaped polygons*, Pattern Recognition **13** (1981), no. 6, 395–398.

[11] M. Axenovich, A. Kündgen, *On a generalized Anti-Ramsey problem*, submitted (1999).

[12] C. T. Benson, *Minimal regular graphs of girths eight and twelve*, Canad. J. Math. **18** (1966), 1091–1094.

[13] J. Bermond, B. Jackson, F. Jaeger, *Shortest coverings of graphs with cycles*, J. Combin. Theory Ser. B **35** (1983), no. 3, 297–308.

[14] B. Bollobás, *Extremal Graph Theory*, Academic Press, New York/London, 1978.

[15] ———, *Random Graphs*, The International Series of Monographs on Computer Science, Academic Press, London, 1985.

[16] A. Bondy, M. Simonovits, *Cycles of even length in graphs*, Journal of Combinatorial Theory Ser. B **16** (1974), 97–105.

[17] J. A. Bondy, Zs. Tuza, *A weighted generalization of Turán's theorem*, J. Graph Theory **25** (1997), 267–275.

[18] R. L. Brooks, *On colouring the nodes of a network*, Proc. Cambridge Philos. Soc. **37** (1941), 194–197.

[19] W. G. Brown, *On graphs that do not contain a Thomsen graph*, Canad. Math. Bull. **9** (1966), 281–285.

[20] W. G. Brown, P. Erdős, M. Simonovits, *Extremal problems for directed graphs*, J. Combinatorial Theory, Ser. B **15** (1973), 77–93.

[21] ———, *Multigraph extremal problems*, Colloques Internationaux CNRS **260** (1978), 63–66.

[22] ———, *Inverse extremal digraph problems*, Colloq. Math. Soc. János Bolyai **37** (1984), North-Holland, Amsterdam - New York, 119–156.

[23] ———, *Algorithmic solution of extremal digraph problems*, Trans. Amer. Math. Soc. **292** (1985), 421–449.

[24] W. G. Brown, P. Erdős, V. T. Sós, *Some extremal problems on r-graphs*, New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich, 1971) (1973), Academic Press, New York, 53–63.

[25] ———, *On the existence of triangulated spheres in* 3*-graphs, and related problems*, Period. Math. Hungar. **3** (1973), 221–228.

[26] W. G. Brown, M. Simonovits, *Digraph extremal problems, hypergraph extremal problems, and the densities of graph structures*, Discrete Math. **48** (1984), 147–162.

[27] B. Chazelle, *Triangulating a simple polygon in linear time*, Discrete Comput. Geom. **6** (1991), no. 5, 485–524.

[28] ———, *Triangulating a simple polygon in linear time*, Proc. 31st Symp. Foundations of Computer Science **I** (1990), 220–230.

[29] N. Chiba, T. Nishizeki, S. Abe, T. Ozawa, *A linear algorithm for embedding planar graphs using PQ-trees*, J. Comput. System Sci **30** (1985), no. 1, 54–76.

[30] F. R. K. Chung, *On the decomposition of graphs*, SIAM J. Algebraic Discrete Methods **2** (1981), no. 1, 1–12.

[31] F. R. K. Chung, P. Erdős, J. Spencer, *On the decomposition of graphs into complete bipartite subgraphs*, Studies in pure mathematics (1983), Birkhäuser, Basel-Boston, 95–101.

[32] V. Chvátal, *A combinatorial theorem in plane geometry*, J. Combinatorial Theory Ser. B **18** (1975), 39–41.

[33] J. Czyzowicz, E. Rivera-Campo, N. Santoro, J. Urrutia, J. Zaks, *Tight bounds for the rectangular art gallery problem*, Lecture Notes in Comput. Sci. **570** (1992), 105–112.

[34] ———, *Guarding rectangular art galleries*, Discrete Appl. Math. **50** (1994), no. 2, 149–157.

[35] D. de Caen, Z. Füredi, *The maximum size of* 3*-uniform hypergraphs not containing a Fano plane*, manuscript (1999).

[36] D. de Caen, L. Székely, *On a problem of Erdős and Lovász*, Ars Combin. **35** (1993), 187–192.

[37] O. Devillers, *Randomization yields simple* $\mathrm{O}(n \log^* n)$ *algorithms for difficult* $\Omega(n)$ *problems*, Internat. J. Comput. Geom. Appl. **2** (1992), 97–111.

[38] G. Dirac, *Extensions of Turán's theorem on graphs*, Acta Math. Acad. Sci. Hungar. **14** (1963), 417–422.

[39] J. Edmonds, E. Johnson, *Matching, Euler tours and the Chinese postman*, Math. Programming **5** (1973), 88–124.

[40] P. Erdős, *Extremal problems in graph theory*, Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963) (1964), Publ. House Czechoslovak Acad. Sci., Prague, 29–36.

[41] P. Erdős, *Extremal problems in graph theory*, A Seminar on Graph Theory (1971), Holt, Rinehart and Winston, New York, 54–59.

[42] P. Erdős, R. Faudree, A. Jagota, T. Łuczak, *Large subgraphs of minimal density or degree*, J. Combin. Math. Combin. Comput. **22** (1996), 87–96.

[43] P. Erdős, P. Frankl, V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs and Combinatorics **2** (1986), 113–121.

[44] P. Erdős, A. W. Goodman, L. Pósa, *The representation of a graph by set intersections*, Canad. J. Math. **18** (1966), 106–112.

[45] P. Erdős, A. Rényi, *On the existence of a factor of degree one of a connected random graph*, Acta Math. Acad. Sci. Hungar. **17** (1966), 359–368.

[46] P. Erdős, A. Rényi, V. T. Sós, *On a problem of graph theory*, Studia Sci. Math. Hungar. **1** (1966), 215–235.

[47] P. Erdős, M. Simonovits, *A limit theorem in graph theory*, Studia Sci. Math. Hungar. **1** (1966), 51–57.

[48] P. Erdős, A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091.

[49] J. Erickson, M. Harris, *personal communication* (1999).

[50] G. Fan, *Proofs of two minimum circuit cover conjectures*, J. Combin. Theory Ser. B **74** (1998), no. 2, 353–367.

[51] S. Fisk, *A short proof of Chvátal's watchman theorem*, J. Combin. Theory Ser. B **24** (1978), no. 3, 374.

[52] H. Fleischner, A. Frank, *On circuit decomposition of planar Eulerian graphs*, J. Combin. Theory Ser. B **50** (1990), no. 2, 245–253.

[53] F. Fu, *On the improvements of Althfer-Sillke inequality*, J. Math. Res. Exposition **16** (1996), no. 3, 325–328.

[54] Z. Füredi, M. X. Goemans, D. J. Kleitman, *On the maximum number of triangles in wheel-free graphs*, Combin. Probab. Comput. **3** (1994), 63–75.

[55] Z. Füredi, A. Kündgen, *Covering a graph with cuts of minimum total size*, Discrete Math. (to appear).

[56] ———, *Turán problems for weighted Graphs*, submitted.

[57] Z. Füredi, M. Simonovits, *in preparation.*

[58] H. Gabow, R. Tarjan, *Faster scaling algorithms for general graph-matching problems*, J. Assoc. Comput. Mach. **38** (1991), 815–853.

[59] M. Garey, D. Johnson, *Computers and intractability. A guide to the theory of NP-completeness.*, W. H. Freeman and Co., San Francisco, 1979.

[60] M. Garey, D. Johnson, H. So, *An application of graph coloring to printed circuit testing*, IEEE Trans. Circuits and Systems **CAS-23** (1976), 591–599.

[61] M. Garey, D. Johnson, L. Stockmeyer, *Some simplified NP-complete graph problems*, Theoret. Comput. Sci. **1** (1976), 237–267.

[62] F. Gavril, *Algorithms for minimum coloring, maximum clique, minimum covering by cliques, and maximum independent set of a chordal graph*, SIAM J. Comput. **1** (1972), no. 2, 180–187.

[63] I. M. Gessel, R. Indik, *A recurrence associated with extremal problems*, preprint (1989).

[64] A. I. Gol'berg, V. A. Gurvich, *On the maximum number of edges for a graph with n vertices in which every subgraph with k vertices has at most l edges*, Soviet Math. Dokl. **35** (1987), 255–260.

[65] J. R. Griggs, M. Simonovits, G. R. Thomas, *Extremal graphs with bounded densities of small subgraphs*, J. Graph Theory **29** (1998), 185–207.

[66] M. G. Guan, H. Fleischner, *On the minimum weighted cycle covering problem for planar graphs*, Ars Combin. **20** (1985), 61–67.

[67] V. A. Gurvich, *Extremal homogeneous hypergraphs and a hypercombinatorial function*, Soviet Math. Dokl. **37** (1988), 146–151.

[68] R. K. Guy, *Sequences associated with a problem of Turán and other problems*, Combinatorial theory and its applications, II. (Proc. Colloq., Balatonfüred, 1969) (1970), North-Holland, Amsterdam, 553–569.

[69] E. Győri, A. V. Kostochka, *On a problem of G. O. H. Katona and T. Tarján*, Acta Math. Acad. Sci. Hungar. **34** (1979), 321–327.

[70] F. Hadlock, *Finding a maximum cut of a planar graph in polynomial time.*, SIAM J. Comput **4** (1975), 221–225.

[71] F. Harary, D. Hsu, Z. Miller, *The biparticity of a graph*, J. Graph Theory **1** (1977), 131–133.

[72] J. Hutchinson, *Art galleries with walls*, Problem #10478, Amer. Math. Monthly **102** (1995), no. 8, 746.

[73] W. Imrich, *Explicite construction of graphs without small cycles*, Combinatorica **4** (1984), 53–59.

[74] A. Itai, M. Rodeh, *Covering a graph by circuits*, Automata, languages and programming; Lecture Notes in Comput. Sci. **62** (1978), Springer, Berlin-New York, 289–299.

[75] A. Itai, R. Lipton, C. Papadimitriou, M. Rodeh, *Covering graphs by simple circuits*, SIAM J. Comput. **10** (1981), 746–750.

[76] B. Jackson, *Shortest circuit covers and postman tours in graphs with a nowhere zero 4-flow*, SIAM J. Comput. **19** (1990), 659–665.

[77] F. Jaeger, A. Khelladi, M. Mollard, *On shortest cocyle covers of graphs*, J. Combin. Theory Ser. B **39** (1985), 153–163.

[78] U. Jamshy, M. Tarsi, *Cycle covering of binary matroids*, J. Combin. Theory Ser. B **46** (1989), 154–161.

[79] U. Jamshy, M. Tarsi, *Short cycle covers and the cycle double cover conjecture*, J. Combin. Theory Ser. B **56** (1992), 197–204.

[80] R. Jayakumar, K. Thulasiraman, M. Swamy, *Planar embedding: linear-time algorithms for vertex placement and edge ordering*, IEEE Trans. Circuits and Systems **35** (1988), no. 3, 334–344.

[81] J. Kahn, *Proof of a conjecture of Katona and Tarján*, Period. Math. Hungar. **12** (1981), no. 1, 81–82.

[82] G. Katona, E. Szemerédi, *On a problem of graph theory*, Studia Sci. Math. Hungar **2** (1967), 23–28.

[83] T. Kővári, V. T. Sós, P. Turán, *On a problem of K. Zarankiewicz*, Colloquium Math. **3** (1954), 50–57.

[84] T. Kratzke, B. Reznick, D. West, *Eigensharp graphs: decomposition into complete bipartite subgraphs*, Trans. Amer. Math. Soc. **308** (1988), 637–653.

[85] J. Kuchenbrod, *Extremal problems on integer weighted graphs*, University of Kentucky Math Research Report **99-01**.

[86] A. Kündgen, *Covering cliques with spanning bicliques*, J. Graph Theory **27** (1998), no. 4, 223–227.

[87] ———, *Art galleries with walls*, Solution to Problem #10478, Amer. Math. Monthly **105** (1998), no. 3, 247–248.

[88] ———, *Art galleries with interior walls*, Discrete Comput. Geom. **22** (1999), 249–258.

[89] ———, *Minimum average distance subsets in the Hamming cube*, submitted (1999).

[90] F. Lazebnik, V. A. Ustimenko, A. J. Woldar, *A new series of dense graphs of high girth*, Bull. Amer. Math. Soc. **32** (1995), 73-79.

[91] D. T. Lee, A. K. Lin, *Computational complexity of art gallery problems*, IEEE Trans. Inform. Theory **32** (1986), no. 2, 276–282.

[92] A. Lubotzky, R. Phillips, P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), 261–277.

[93] D. Manocha, A. Narkede, *Fast Polygon Triangulation based on Seidel's Algorithm*, http://www.cs.unc.edu/ dm/CODE/GEM/chapter.html.

[94] W. Mantel, *Problem 28*, Wiskundige Opgaven **10** (1907), 60–61.

[95] G. A. Margulis, *Explicit construction of graphs without short cycles and low density codes*, Combinatorica **2** (1982), 71–78.

[96] D.W. Matula, *k-components, clusters and slicings in graphs*, SIAM J. Appl. Math. **22** (1972), 459–480.

[97] T. S. Motzkin, E. G. Straus, *Maxima for graphs and a new proof of a theorem of Turán*, Canad. J. Math. **17** (1965), 533–540.

[98] J. O'Rourke, *Art gallery theorems and algorithms*, The International Series of Monographs on Computer Science, The Clarendon Press, Oxford University Press, New York, 1987.

[99] L. Pyber, *Covering the edges of a graph by ...*, Sets, graphs and numbers (Budapest, 1991), Proc. Colloq. Math. Soc. János Bolyai **60** (1992), North-Holland, Amsterdam, 583–610.

[100] A. Rényi, *On random generating elements of a finite Boolean algebra*, Acta Sci. Math. Szeged **22** (1961), 75–81.

[101] B. Reznick, P. Tiwari, D. West, *Decomposition of product graphs into complete bipartite subgraphs*, Discrete Math. **57** (1985), 189–193.

[102] N. Robertson, D. Sanders, P. Seymour, R. Thomas, *A new proof of the four-colour theorem*, Electron. Res. Announc. Amer. Math. Soc. **2** (1996), 17–25 (electronic).

[103] V. Rödl, A. Sidorenko, *On the jumping constant conjecture for multigraphs*, J. Combinatorial Theory, Ser. A **69** (1995), 347–357.

[104] D.J. Rose, R.E. Tarjan, G.S. Lueker, *Algorithmic aspects of vertex elimination on graphs.*, SIAM J. Comput. **5** (1976), no. 2, 266–283.

[105] I. Ruzsa, E. Szemerédi, *Triple systems with no six points carrying three triangles*, Colloq. Math. Soc. János Bolyai **18** (1978), North-Holland, Amsterdam - New York, 939–945.

[106] R. Seidel, *A simple and fast incremental randomized algorithm for computing trapezoidal decompositions and for triangulating polygons*, Comput. Geom. **1** (1991), 51–64.

[107] T. Shermer, *Recent results in art galleries*, Proceedings IEEE **80(9)** (1992), 1384–1399.

[108] V. T. Sós, *Remarks on the connection of graph theory, finite geometry and block designs*, Colloquio Internazionale sulle Teorie Combinatorie (Roma, 1973), Tomo II, Atti dei Convegni Lincei, No. 17 (1976), Accad. Naz. Lincei, Rome, 223–233.

[109] B. S. Stečhkin, *On a surprising fact in extremal set theory*, J. Combin. Theory Ser. A **29** (1980), 368–369.

[110] B. S. Stečhkin, P. Frankl, *Local Turán property for k-graphs*, Mat. Zametki **29** (1981), 83–94. (Russian)

[111] M. Tarsi, *Nowhere zero flow and circuit covering in regular matroids*, J. Combinatorial Theory, Ser. B **39** (1985), 346–352.

[112] C. Thomassen, *On the complexity of finding a minimum cycle cover of a graph*, SIAM J. Comput. **26** (1997), 675–677.

[113] P. Turán, *On an extremal problem in graph theory*, Math. Fiz. Lapok **48** (1941), 436–452. (Hungarian)

[114] ______, *On the theory of graphs*, Colloquium Math. **3** (1954), 19–30.

[115] Zs. Tuza, *Covering of graphs by complete bipartite subgraphs: complexity of* 0-1 *matrices.*, Combinatorica **4** (1984), 111–116.

[116] R. Wenger, *Extremal graphs with no C^4's, C^6's or C^{10}'s*, Journal of Combinatorial Theory, Ser. B **52** (1991), 113–116.

[117] D. West, *Introduction to graph theory*, Prentice Hall, Inc., Upper Saddle River, NJ, 1996.

[118] M. Yannakakis, *Node- and edge-deletion NP-complete problems*, Conference Record of the Tenth Annual ACM Symposium on Theory of Computing (San Diego, Calif., 1978) (1978), 253–264.

[119] C.Q. Zhang, *Integer flows and cycle covers of graphs*, Marcel Dekker, Inc, New York, 1997.

[120] ———, *Minimum cycle coverings and integer flows*, J. Graph Theory **14** (1990), 537–546.

[121] ———, *Cycle cover theorems and their applications*, Contemp. Math. AMS **147** (1993), 405–418.

Vita

André Kündgen was born on May 8, 1969 in Trier, Germany. He received a Master of Science degree in mathematics from the University of South Carolina in 1993. In August 1993 he entered the Ph.D. program in mathematics at the University of Illinois at Urbana-Champaign. He received his Ph.D. in mathematics, under the direction of Prof. Zoltán Füredi, in May of 1999.

His publications include the following.

- *Covering cliques with spanning bicliques*, Journal of Graph Theory, **27** (1998), 223–227.
- *Art galleries with interior walls*, Discrete and Computational Geometry, **22** (1999), 249–258.
- *Covering a graph with cuts of minimum total size*, with Z. Füredi, to appear.
- *Turán problems for weighted graphs*, with Z. Füredi, submitted.
- *Minimal completely separating systems of k-sets*, with D. Mubayi, P. Tetali, submitted.
- *On a generalized Anti-Ramsey problem*, with M. Axenovic, submitted.
- *Perfection thickness of graphs*, with H. Asari, T. Jiang, D. West, to appear.
- *Minimum average distance subsets in the hypercube*, submitted.

PROBLEMS IN EXTREMAL GRAPH THEORY

André Kündgen, Ph.D.
Department of Mathematics
University of Illinois at Urbana-Champaign, 1999
Zoltán Füredi, Advisor

We consider several extremal problems in graph theory and computational geometry.

A *cut* in a graph G is the set of all edges between some set of vertices S and its complement $\overline{S} = V(G) - S$. A *cut-cover* of G is a collection of cuts whose union is $E(G)$ and the *total size* of a cut-cover is the sum of the number of edges of the cuts in the cover. The cut-cover size of a graph G, denoted by $\mathrm{cs}(G)$, is the minimum total size of a cut-cover of G.

We give general bounds on $\mathrm{cs}(G)$, find sharp bounds for classes of graphs such as 4-colorable graphs and random graphs. We give sharp bounds for the sum of the cut-cover sizes of a graph and its complement. We also determine the cut-cover size of the complete graph on n vertices and establish a close connection between this problem and an isoperimetric problem due to Ahlswede and Katona.

Consider an *art gallery* formed by a polygon on n vertices with m pairs of vertices joined by interior diagonals, the *interior walls.* Each interior wall has an arbitrarily placed, arbitrarily small doorway. We determine the minimum number of guards that suffice to guard all art galleries with n vertices and m interior walls. We also determine the minimum number of guards that suffice for galleries with convex rooms of specified minimum size. The proofs lead to linear time guard placement algorithms in most cases.

What is the maximum number of edges in a multigraph on n vertices if every k-set spans at most r edges? We asymptotically determine this maximum for almost all k and r as n tends to infinity, thus giving a generalization of Turán's theorem. We find exact answers in many cases, even when edges of negative multiplicity are allowed.