Graph Parameters

A Dissertation

Presented for the

Doctor of Philosophy

Degree

University of Szeged

József Balogh

May, 2002

Dedication

To my family.

Abstract

The interval number of a graph G, denoted by $i(G)$, is the least natural number t such that G is the intersection graph of sets, each of which is the union of at most t intervals. Here, in the first chapter we settle a conjecture of Griggs and West about bounding $i(G)$ in terms of e, that is the number of edges in G. Namely, it will be shown that $i(G) \leq \lceil 1/2\sqrt{e} \rceil + 1$. It is also observed that the edge bound induces $i(G) \leq \sqrt{3/2\gamma(G)} + o(1)$, where $\gamma(G)$ is the genus of G.

Most known bounds on $i(G)$ are grossly excessive when G has more than half of the possible edges. A plausible remedy is to develop bounds on $i(G)$ that are monotone decreasing in G. In Chapter 2, we bound $i(G)$ in terms of $e(\bar{G})$, the number of edges in the complement of G. We prove that $i(G) \leq \lceil \frac{1}{2}\sqrt{e(\bar{G})} \rceil + O(n/\log n)$.

In Chapter 3 we examine further properties of the interval number of graphs. Namely, we characterize some graphs, whose interval number is maximal possible for the degree bound, we determine up to a 1/2 factor the interval number of split graphs, and we claim that the main reason for the unusually high interval number is a "large" induced bipartite graph.

In Chapter 3 we define the game domination number. The game domination number of a (simple, undirected) graph is defined by the following game. Two players, A and D, orient the edges of the graph alternately

until all edges are oriented. Player D starts the game, and his goal is to decrease the domination number of the resulting digraph, while A is trying to increase it. The *game domination number* of the graph G, denoted by $\gamma_g(G)$, is the domination number of the directed graph resulting from this game. This is well-defined if we suppose that both players follow their optimal strategies.

A *dominating set* of a digraph $\vec{G}$ is a set S of vertices such that for every vertex $v \notin S$ there exists some $u \in S$ with $\vec{uv} \in E(\vec{G})$. The *domination number* $\gamma(\vec{G})$ of $\vec{G}$ is defined as the cardinality of the smallest dominating set.

We define a "domination parameter" of an undirected graph G as the domination number of one of its orientations, determined by the following two player game. Players A and D orient the unoriented edges of the graph G alternately with D playing first, until all edges are oriented. Player D (frequently called the *Dominator*) is trying to minimize the domination number of the resulting digraph, while player A (*Avoider*) tries to maximize the domination number. This game gives a unique number depending only on G, if we suppose that both A and D play according to their optimal strategies. We call this number the *game domination number of* G and denote it by $\gamma_g(G)$.

We determine the game domination number for several classes of graphs and provide general inequalities relating it to other graph parameters.

An extremal jump is a discrete step between measures guaranteed in certain situations. It has been known for some time that the density of a graph jumps; recent work on hereditary graph properties has shown that properties with “large” or “small” speeds jump, but it was unknown whether there is a clean jump for properties with speed in a middle range. In Chapter 5, generalizations of the theorems of Dilworth, Ramsey, and Turán's are applied to answer this in the affirmative. In particular, we find a strict lower bound for the penultimate range of the speed hierarchy for hereditary properties of graphs.

This dissertation shall examine problems that have their roots in questions about properties and parameters of graphs. Although its origins may be humble, each question develops into a deep theory about its subject.

Chapter 1 and 2 is joint work with András Pluhár, Chapter 3 with Pascal Ochem and András Pluhár, Chapter 4 and 5 with Béla Bollobás, Chapter 4 with Noga Alon and Tamás Szabó, Chapter 5 with David Weinreich.

Acknowledgements

First of all, I would like to say thanks, to my advisor, Professor Péter Hajnal, who helps my steps in mathematics since my freshman year. I would like to say thanks to Professor Béla Bollobás, who invited me to Memphis to work with him. He not only pushed me in mathematics by posing interesting open questions, giving nice papers to study, and supporting my travels to conferences and workshops, but helped me to find my place in the "real world".

Thanks are also due to my coauthors as well:

to (above mentioned) Béla Bollobás,
to Professor András Pluhár, who showed me the problems on the interval number of graphs,
to Professor Tamás Szabó, who as a graduate student in Memphis contributed greatly to Chapter 4,
to Professor David Weinreich, who, as a graduate student in Memphis, contributed greatly to Chapter 5,
to Alon Noga and Pascal Ochem.

My family and friends endured many ups and downs as I pursued my degree. Their emotional and moral support helped make this happen, but they missed me for several years.

Detailed abstract of the dissertation

The main topics of my dissertation are *extremal graph theory*, and *bootstrap percolation*.

Extremal graph theory is a classical, well studied part of combinatorics. A central question of extremal graph theory concerns the maximum number of edges in a graph (network) on n vertices, containing no subgraph isomorphic to a fixed smaller graph. Extremal graph theory was one of the favorite research topics of Paul Erdős, and from the 1950s, many of the best known graph theorists worked in this area. In the last ten years much research has been devoted to *global* questions concerning *induced subgraphs*, going beyond the study of specific problems. In applications, such a question might appear in the following way: how can a computer network be built, if certain substructures cannot occur (for example it cannot contain five computers with each of the ten pairs connected). A property in this sense is a set of possible computer networks, which satisfy some restrictions. The size of a graph property is the cardinality of graphs in the property. Our goal, which is a living important topic in recent mathematical research, was to give estimates for the size of these sets. This type of question has suddenly become very important in the age of the internet. The significance of our results is that we show for some important classes of graph properties that their sizes fall into one of a number of narrow ranges.

In the last fifteen years, Erdős, Frankl, Rödl, Prömel, Steger, Bollobás, Thomason and many others studied these questions and obtained numerous substantial results. With Bollobás and Weinreich, we investigated the possible size and structure of a hereditary property, when the possible size is not very large, i.e., $2^{o(n^2)}$, solving three conjectures of Scheinerman and Zito. In particular, we gave a full description of the hereditary properties in the range below n^n, and we classified all possible sizes. Erdős, Frankl and Rödl proved various monotone results for properties having size in the higher ranges, and their work was continued by Prömel and Steger. These results imply that for every monotone property there is a natural number k such that the size of the property is $2^{(1-1/k+o(1))n^2}$. Further, in a series of papers, Prömel and Steger showed that the error term can be sharpened for some natural properties. Recently, with Bollobás and Simonovits, we managed to extend these results to all monotone properties. In particular we have given essentially sharp estimates: for all monotone properties, there is an integer k and a positive constant c such that the cardinality of a property is between $2^{(1-1/k)\binom{n}{2}}$ and $2^{(1-1/k)\binom{n}{2}+n^{2-c}}$. Our results are proved by applying techniques of probabilistic combinatorics, exploiting properties of the automorphism groups of graphs and hypergraphs, and making use of a number of classical combinatorial theorems including the famous Szemerédi's Regularity Lemma.

As a graduate student in Hungary, I lectured on combinatorics to undergraduates. With one of them, Pete, I developed a geometric approach to *bootstrap percolation* which, when used with probabilistic tools, gives some of the main results in this area. Bootstrap percolation originated in the 1990s in statistical physics and biology, where it was used as a model of cellular automata. With our new method, Pete and I could simplify the proof of a fundamental result of Aizenman and Lebowitz in *finite size* scaling bootstrap percolation. In Memphis, I continued my work on these finite models of bootstrap percolation (which tend to be less amenable to standard techniques than the infinite models). The speed of the transition is always an important question, and with Bollobás we have proved the existence of sharp threshold functions. Furthermore, we have given good bounds on the threshold function of the bootstrap percolation on the hypercube.

I am interested in looking at several questions left open in my dissertation on hereditary and monotone graph properties and on bootstrap percolation. There is no doubt that the two main areas of my research so far, the theory of hereditary graph properties and bootstrap percolation, are far from being exhausted and will remain fertile areas for a long time. In fact, it is very likely that answering a number of major questions would transform our understanding of these fields.

Contents

Chapter 1

A Sharp Edge Bound on the Interval Number of a Graph

1.1 Introduction and Results

The interval number of a graph G, denoted by $i(G)$, is the least natural number t such that G is the intersection graph of sets, each of which is the union of at most t intervals. Here we settle a conjecture of Griggs and West about bounding $i(G)$ in terms of e, that is the number of edges in G. Namely, it will be shown that $i(G) \leq \lceil 1/2\sqrt{e}\,\rceil + 1$. It is also observed that the edge bound induces $i(G) \leq \sqrt{3/2\gamma(G)} + o(1)$, where $\gamma(G)$ is the genus of G.

In this chapter we concerned with representing graphs as special intersection graphs. That is, we assign a set to each vertex of G so that v is adjacent to w if and only if the common part of the assigned sets is not empty. A *t-interval representation* is such an assignment, in which each set consists of at most t closed intervals. The *interval number* of G, denoted by $i(G)$, is the least integer t for which a t-representation of G exists. We need one more definition: a representation is *displayed* if each set of the representation has an open interval, which is disjoint from the other sets. Such an interval is called *displayed segment.*

Several interesting aspects of the interval number had been investigated before, here we wish to pursue only one of these in depth. More than one and a half decade ago Griggs and West conjectured that $i(G) \leq \lfloor 1/2\sqrt{e} \rfloor + 1$ for any graph G, where e is the number of edges in G. They also showed that this would be the best possible because the bound is attained for the complete bipartite graphs $K_{2m,2m}$, where m is a positive integer. In fact their conjecture can be restated as $i(G) \leq \lceil 1/2\sqrt{e} \rceil + 1$, since the two forms coincide for the previous graphs. In a series of papers it was proved that $i(G) \leq \lceil c\sqrt{e} \rceil + 1$ for $c = 1$, $c = \sqrt{2}/2$ and $c = 2/3$, see [50], [75] and [60]. The main goal of this paper is to conclude this process by showing that the right constant is really $c = 1/2$.

We also mention some connected problems, which are worth further study.

Theorem 1.1. *Every graph with e edges has a displayed interval representation with at most $\lceil 1/2\sqrt{e} \rceil + 1$ intervals for each vertex.*

(In [75] the form $i(G) \leq \lceil 1/2(\sqrt{e} + \sqrt{e^{-1}}) \rceil$ is mentioned, which can be a little better than the version of Theorem 1.1 when e is not a square of an even integer.)

There is some correspondence between the interval number of a graph G

and $\gamma(G)$, the *genus* of G. It was shown in [69] that $i(G) \leq \lceil \sqrt{3\gamma(G)} \rceil + 3$. As a consequence of Theorem 1.1 this result can be improved:

Theorem 1.2. *For a graph G, $i(G) \leq \lceil \sqrt{3/2\gamma(G)} \rceil + 4$, provided the genus is big enough.*

1.2 Important Lemmas

The proof of Theorem 1.1 is based on two previous results, which bound $i(G)$ in terms of the arboricity and the number of vertices of G, see [69] and [49]. However, we need improvements on both, partly in order to "smooth out" the subsequent calculations. The arboricity of a graph G, denoted by $a(G)$, is the size of the smallest partition of the edge set of G into forests. We shall use $e(G)$ and $v(G)$ to designate the number of edges and vertices of a graph G, respectively.

Lemma 1.3. *[69] Every graph with e edges has a displayed interval representation with at most $a(G) + 1$ intervals for each vertex.*

Following the proof of Lemma 1.3 one easily reads out a little more.

Lemma 1.4. *Let us assign pairwise disjoint intervals to each vertex of a graph G, and pick a vertex v. Assigning at most $a(G)$ additional intervals to a vertex $w \neq v$ and no more to v, we can get a displayed interval representation of G.*

Another variation of Lemma 1.3 says that the arboricity can be replaced by the maximal density in G. Although its proof is very similar to Lemma 1.3 we shall sketch it for the sake of compactness.

Lemma 1.5. *A graph G with e edges has a displayed interval representation with at most* $\max_{\emptyset \neq H \subset G} \lceil e(H)/v(H) \rceil + 1$ *intervals for each vertex.*

Now we are ready to prove Lemma 1.5.

Proof. Let

$$k = \max_{\emptyset \neq H \subset G} \lceil e(H)/v(H) \rceil + 1.$$

First we direct the edges of G in such a way that the out-degree of any vertex v is not more than k. Indeed, one verifies that the subsets of $E(G)$, containing at most one circuit of G, are independent sets of a matroid M. (In fact, M is just the Nash-Williams sum of the circuit matroid of G and the uniform matroid $U_{e,1}$, where $e = e(G)$.) According to the formulas of Nash-Williams or Edmonds, a matroid $M = (S, \mathcal{I})$ can be partitioned into l independent sets iff $l \geq \max_{\emptyset \neq X \subset S} \lceil |X|/rank(X) \rceil$ (see [58] and [29]). In our case this implies that $E(G) = E_1 \cup ... \cup E_k$, where the edges of E_i span at most one circuit for $i = 1, ..., k$. It suffices to show the edges of E_i can be directed such that all out-degrees become at most one. If a component, spanned by E_i, contains no circuit, then it must be a tree. Pick any vertex v from it and direct all edges of the component towards v as a root. When a component does have a circuit, pick v in such a way that $f = (u, v)$ undirected edge is in this unique circuit. Direct f from v to u, and direct all other edges towards the vertex v.

Having the desired direction of G we assign an interval I_v to each vertex v, such that $I_v \cap I_w = \emptyset$ for $v \neq w$. Furthermore, we assign a small interval to x inside I_y for a directed edge (x, y). Doing this we use up at most

$k+1$ intervals per vertex, and represent all edges of G. ∎

Another basic fact what we need is a bound on the interval number by the number of vertices.

Theorem 1.6. *[49] If a graph G has $n > 1$ vertices, then $i(G) \leq \lceil 1/4(n+1) \rceil$, and this bound is the best possible.*

Alas, the representation used in the proof of Theorem 1.6 is not displayed. On the other hand, as it was first noted in [75], the induction part of the proof would preserve the displayed feature. Furthermore, in the basis case only K_3 has no displayed t-representation, where $t = \lceil 1/4(n+1) \rceil = 1$. Here we do not wish to pursue the possibilities to improve on Theorem 1.6. Instead of this, we simply bypass the problem augmenting the previous bound to a level at which the argument of [49] could be repeated word by word, and it yields a displayed representation.

Lemma 1.7. *Every graph with n vertices has a displayed interval representation with at most $\lceil 1/4(n+1+\epsilon) \rceil$ intervals for each vertex, where $\epsilon > 0$.*

1.3 Proof of Theorem 1.1

The representation of the desired type is made up in two steps. First a dense subgraph of G is represented by the means of Lemma 1.7, then an appropriate extension is given by the line of Lemma 1.4. We need some notations and a new lemma in order to formalize this. If Q is a spanned subgraph of G, then G/Q stands for the contraction of Q in G such that the loops are deleted. More specifically $G/Q = (V(G/Q), E(G/Q))$, where

$$V(G/Q) = \{V(G) \setminus V(Q)\} \cup \{q\}$$

and

$$E(G/Q) = E(G \setminus Q) \cup \{(u,q) : \ \forall\, (u,v) \in G, \ u \notin V(Q), \ v \in V(Q)\}.$$

It is very convenient to use the notion of $\hat{\imath}(G)$, the displayed interval number of G, which is defined similarly as $i(G)$ except that the representations must be displayed (see [9]).

Lemma 1.8. *Let Q be a spanned subgraph of G. Then*

$$\hat{\imath}(G) \leq \max\{\hat{\imath}(Q), a(G/Q) + 1\}.$$

Proof. Let us take a displayed *$\hat{\imath}(G)$*-representation $\mathcal{R}$ of Q, then assign pairwise disjoint intervals to each $v \in V(G) \setminus V(Q)$, which do not meet with the elements of $\mathcal{R}$. By Lemma 1.4, the edges of G/Q can be represented such that at most $a(G/Q)$ new intervals are used for a vertex $v \in V(G) \setminus V(Q)$, and no new interval for q. Consequently the set $E(G) \setminus E(Q)$ can also be represented this way, since the representation $\mathcal{R}$ is displayed. (One can put a small interval, assigned to v, into the displayed segment of the interval corresponds to u in $\mathcal{R}$, if $u \in V(Q)$, $v \in V(G) \setminus V(Q)$ and $(u,v) \in E(G)$.) ∎

Let us set $k = \lceil 1/2\sqrt{e(G)}\, \rceil$, and let Q be a maximal subgraph of G such that $e(Q)/v(Q) \geq k - 1/2$. If $Q = \emptyset$, i.e. the maximal density of G

is less than k, then Lemma 1.5 implies Theorem 1.1, so we may assume that $Q \neq \emptyset$. Because of the maximality of Q, $e(H)/v(H) < k - 1/2$ for every $H \subset G/Q$. On the other hand $e(H)/(v(H) - 1) \leq k$ also holds. In fact, if $q \in H$, $e(H)/(v(H) - 1) > k$ would contradict the maximality of Q. If $q \notin H$, then $e(H)/(v(H) - 1) > k$ would imply $v(H) > 2k$, since $1/2(v(H) - 1)v(H) \geq e(H)$. With this we have

$$\frac{e(H)}{v(H) - 1} = \frac{e(H)}{v(H)} \frac{v(H)}{v(H) - 1} < (k - \frac{1}{2})(1 + \frac{1}{2k}) < k.$$

In other words $a(G/Q) \leq k = \lceil 1/2\sqrt{e(G)} \rceil$. Finally, we need to estimate $\hat{\imath}$(Q) in order to conclude the proof by Lemma 1.8. This is done by using Lemma 1.7.

$$\hat{\imath}(Q) \leq \lceil \frac{1}{4}(v(Q) + 1 + \epsilon) \rceil \leq \lceil \frac{1}{4}(\frac{e(Q)}{k - 1/2} + 1 + \epsilon) \rceil \leq$$
$$\leq \lceil \frac{1}{2}\sqrt{e(G)} + \frac{1}{2\sqrt{e(G)} - 2} + \frac{3}{4} + \epsilon \rceil \leq \lceil \frac{1}{2}\sqrt{e(G)} \rceil + 1,$$

provided that $e(G) > 9$. One can also see that there are displayed $(\lceil 1/2\sqrt{e(G)} \rceil + 1)$-representations for Q, if $e(G) \leq 9$. ∎

1.4 An Application and Remarks

There are several ways to get Theorem 1.2 by the application of the tools, developed in this paper. Now we reduce it to Theorem 1.1.

Proof. We may assume that all vertex of G has degree at least $\sqrt{3/2\gamma(G)}$. (An analogous idea appears in [75], since working with displayed representations all edges incident to a low-degree vertex v can be represented by intervals all assigned to v.) For the easier notations, let us use n, e and γ instead of $v(G)$, $e(G)$ and $\gamma(G)$ for the rest. It follows from our assumption that $n\sqrt{3/2\gamma} \leq 2e$. On the other hand, Euler's formula implies

$$3n + 6\gamma - 6 \geq e.$$

Plugging the estimation for n into this, we get

$$\frac{6e}{\sqrt{3/2\gamma}} + 6\gamma - 6 \geq e,$$

which gives

$$6\gamma + 30\sqrt{\gamma} \geq e,$$

if γ is big enough. By Theorem 1.1

$$i(G) \leq \lceil \frac{1}{2}\sqrt{e} \, \rceil + 1 \leq \lceil \frac{1}{2}\sqrt{6\gamma + 30\sqrt{\gamma}} \, \rceil + 1 \leq \lceil \sqrt{3/2\gamma} \, \rceil + 4,$$

and we are done. ∎

Since $\gamma(K_n) = \lceil 1/12(n-3)(n-4) \rceil$, one cannot get essentially better result by using Theorem 1.1 alone. However, we believe that the right constant is one.

Conjecture 1.9. *Every graph with γ genus has a displayed interval representation with at most $\lceil \sqrt{\gamma} \, \rceil + 3$ intervals for each vertex.*

One may get closer to this by taking into account that the interval number is *not* a monotone function. We can formulate another conjecture along this line which looks interesting in its own right. Let us recall that $\overline{G}$ is the complement of a graph G and $e(\overline{G})$ is the number of edges in $\overline{G}$.

Conjecture 1.10. *Every graph G has a displayed interval representation with at most $\lceil 1/2\sqrt{e(\overline{G})}\,\rceil + 1$ intervals for each vertex.*

In fact Conjecture 1.10 and Lemma 1.5 would imply Theorem 1.1. Furthermore Theorem 1.1 and Conjecture 1.10 immediately give almost the same bound as Theorem 1.6, which makes our claim very desirable, and probably hard. There are non-trivial graph classes for which Conjecture 1.10 is proved, but not much is known in the general case.

Chapter 2

The Interval Number of Dense Graphs

2.1 Introduction

The interval number of a graph G, denoted by $i(G)$, is the least natural number t such that G is the intersection graph of sets, each of which is the union of at most t closed intervals. Most known bounds on $i(G)$ are grossly excessive when G has more than half of the possible edges. A plausible remedy is to develop bounds on $i(G)$ that are monotone decreasing in G. Here we bound $i(G)$ in terms of $e(\bar{G})$, the number of edges in the complement of G. We prove that $i(G) \leq \lceil \frac{1}{2}\sqrt{e(\bar{G})}\, \rceil + O(n/\log n)$.

In this chapter we study special intersection representations of graphs. Let us recall the definitions and the main results from the previous chapter. We assign a set to each vertex of G so that v is adjacent to w if and only if the assigned set intersect. A *t-interval representation* is such an assignment in which each set consists of the union at most t closed intervals. The *interval number* of G, denoted by $i(G)$, is the least integer t such that a t-representation of G exists. A t-interval representation is *displayed* if the set assigned to each vertex contains an open interval (called a *displayed segment*) that is disjoint from the other assigned sets. Various aspects of the interval number have been studied. Sharp bounds are known in terms of the number of vertices ($v(G)$), number of edges ($e(G)$), the maximum degree ($\Delta(G)$), and the maximum density of a graph. The maximum density is defined as $\rho(G) = \max_{H \subset G} e(H)/v(H)$.

Theorem 2.1. *[50] $i(G) \leq \lceil (\Delta(G) + 1)/2 \rceil$.*

Theorem 2.2. *[49] $i(G) \leq \lceil (v(G) + 1)/4 \rceil$.*

Theorem 2.3. *[15] Every graph G has a displayed t-interval representation with $t \leq \lceil \frac{1}{2}\sqrt{e(G)} \rceil + 1$.*

Theorem 2.4. *[67, 15] Every graph G has a displayed t-interval representation with $t \leq \lceil \rho(G) \rceil + 1$. Moreover, G has an orientation in which the maximum outdegree is no more than $\lceil \rho(G) \rceil$.*

These bounds are attained for the complete bipartite graph $K_{2m,2m}$ with m a positive integer. (See [50] and [79] for $i(K_{m,m}) = \lceil (m + 1)/2 \rceil$.)

The message of all these bounds is that when a graph is sparse, its interval number is small. However, the interval number of a dense graph is not necessarily large; in particular $i(K_n) = 1$. A bound for very special graphs was proved in [79]: $i(\bar{T}) \leq 3$ when T is a tree. Several bounds in terms of the complement should be investigated.

Conjecture 2.5. *[15]* $i(G) \leq \frac{1}{2}\lceil\sqrt{e(\bar{G})}\rceil + 1$.

This bound would be sharp, since it is attained for the complete bipartite graph $K_{2m,2m}$ and for the graph obtained by adding an edge from each vertex of $K_{2m,2m}$ to each vertex of a clique of order t. These are the only extremal graphs known so far. Note that Conjecture 2.5 and Theorem 2.3 would essentially imply Theorem 1.6.

Here we take a step toward settling Conjecture 2.5. As a byproduct, we also obtain upper bounds in terms of $\Delta(\bar{G})$.

Theorem 2.6. $i(G) \leq \lceil \frac{1}{2}\sqrt{e(\bar{G})}\, \rceil + (3/2 + o(1))n/\log n$.

Theorem 2.7. $i(G) \leq \lceil(\Delta(\bar{G}) + 1)/2\rceil + 2^{n/(\Delta(\bar{G})+1)}$.

Theorem 2.8. $i(G) \leq \Delta(\bar{G}) + \frac{1}{2}\chi(\bar{G}) + 1$, *where* χ *denotes the chromatic number.*

Corollary 2.9. $i(G) \leq \frac{3}{2}(\Delta(\bar{G}) + 1)$.

Perhaps every bounds in terms of density has an analogue in terms of the complement.

Conjecture 2.10. $i(G) \leq \lceil\frac{1}{2}(\Delta(\bar{G}) + 1)\rceil$ *and* $i(G) \leq \lceil\rho(\bar{G})\rceil + 1$.

2.2 Proofs of the Results

It is costumary in this subject to speak of assigning individual intervals to a vertex instead of assigning the set that is their union. In our representations, we first establish disjoint displayed intervals in a unused portion of the line, assigning one to each vertex. We refer to these intervals as the *basic intervals*. Often we will represent an edge uv by assigning a small interval for u in the interior of the basic interval for v. This adds an interval for u to the representation, but does not alter the set assigned to v.

Observation 0: Given basic intervals for $V(G)$ and an orientation of a subgraph of G, one can represent the oriented edges using at most $d^+(x)$ additional intervals for each vertex x.

Proof. For each oriented edge xy, we assign a small interval for x within the basic interval for y. (This technique first appeared in [67], although it was observed earlier by Trotter.) ∎

We next consider optimal proper colorings of $\bar{G}$. With every proper coloring of $\bar{G}$ with $\chi(\bar{G})$ colors, we associate a $\chi(\bar{G})$-dimensional vector in which the components are the sizes of the color classes. A *lexicographically minimal* coloring of $\bar{G}$ is a proper coloring with $\chi(\bar{G})$ colors for which the vector of sizes is lexicographically minimal. Within a lexicographically minimal coloring, let c_1, c_2, c_3 count the color classes with one element, two element, or at least three elements, respectively; these are the *parameters* of the coloring.

Lemma 2.11. *For a lexicographically minimal coloring of $\bar{G}$ with parameters c_1, c_2, c_3,*

$$e(\bar{G}) \geq \binom{c_1}{2} + 2\binom{c_2}{2} + 3\binom{c_3}{2} + c_1c_2 + c_1c_3 + 2c_2c_3.$$

Proof. In a lexicographically minimal coloring, every vertex has neighbor in each color class that is at least as big as its own; otherwise, its color can be changed to that of the larger class to improve the size vector. The expression claimed is the result of summing these requirements. ∎

We shall partition the edge set of G and represent the edges in these classes separately, almost independently. First we consider the part of G that is induced by the color classes of size at most two in a lexicographically minimal coloring of $\bar{G}$.

Lemma 2.12. *Given basic intervals, one can represent the subgraph H of G induced by the vertices in color classes of size at most two in a lexicographically minimal coloring of $\bar{G}$ using at most $p = \lceil \frac{c_2(c_1+c_2)}{c_1+2c_2} \rceil$ additional intervals per vertex.*

Proof. First we show that $\rho(H_1) \leq p$ for all $H_1 \subset H$. Let $d_1 \leq c_1$ be the number of one-element classes in H that lie in H_1. Let $d_2 \leq c_2$ be the number of two-element classes in H that contribute at least one vertex to H_1, with l of them contributing exactly one vertex and $d_2 - l$ contributing both. Lexicographic minimality of the coloring implies that a vertex from one of the incomplete classes in H_1 cannot be connected by two edges in G to one of the complete classes. Similarly, the number of edges induced by the union of the complete 2-element classes is at most $2\binom{d_2-l}{2} + (d_2 - l) = (d_2 - l)^2$. Now we can bound $e(H_1)$, much like in the proof of Lemma 2.11. We compute

$$e(H_1) \leq ld_1 + d_1(d_2-l) + l(d_2-l) + (d_2-l)^2 = d_1d_2 + (d_2-l)d_2 =$$

$$d_2(d_1+d_2-l) \leq d_2(d_1+2d_2-l)\frac{d_1+d_2}{d_1+2d_2} = (d_1+2(d_2-l)+l)\frac{d_2(d_1+d_2)}{d_1+2d_2} \leq v(H_1)p.$$

That is $\lceil \rho(H) \rceil \leq p$, so Theorem 2.4 implies Lemma 2.12. ∎

Scheinerman [69] introduced a generalization of the displayed representation method. Consider a clique-partition of the vertices of a graph G. View cliques as vertex sets. For each clique Q in the partition, establish $2^{|Q|}$ disjoint intervals in an unused portion of the line. For each subset of Q, assign the corresponding interval to each member of that subset.

This *generalized displayed representation* increases the number of intervals used x by 2^{k-1}, where k is the size of the clique containing x. Given such a system, we can represent the edges between the cliques P and Q cheaply; for a vertex $x \in P$ we just use a small interval within the system of identical intervals for the neighborhood of x in Q. This way of representing all edges of a (P,Q) pair uses one interval for each vertex of P and none for Q; we call this method a *Scheinerman representation.*

Lemma 2.13. *If k is the maximum size of a color class in a lexicographically minimal coloring of $\bar{G}$ with parameters c_1, c_2, c_3, then*

$$i(G) \leq \left(\frac{c_2(c_1+c_2)}{c_1+2c_2}(c_1+c_2) + (c_1+c_2)c_3 + \binom{c_3}{2} \right) \frac{1}{c_1+c_2+c_3} + 2^{k-1} + 3.$$

Proof. The color classes in the coloring of $\bar{G}$ yield a clique partition of G. We establish generalized displayed representations for the cliques with at least three vertices. The main problem is how to choose the ordered pairs of cliques for representation of edges between them in the Scheinerman representation.

If we process an ordered pair (P, Q), then we need not process the pair (Q, P). Also, nothing needs to be done for pairs of cliques that both have size at most 2; using Lemma 2.12, these edges are being represented in H using p intervals per vertex. We aim to orient the $(c_1 + c_2)c_3 + \binom{c_3}{2}$ leftover pairs so that each clique appears first not too often. We start the count at zero for a clique of size at least three and p for a clique of size at most two.

In effect, consider the graph J that is the join of K_{c_3} with c_1+c_2 vertices that are independent of each other but each incident to p loops. We orient J to minimize the maximum outdegree, so that each vertex has essentially the average outdegree. (Since $p < c_2$, the average exceeds p, so one can see that $\lceil \rho(J) \rceil \leq e(J)/v(J)$. Then the second part of Theorem 2.4 applies.)

This orientation yields the claimed formula. We wrote a $+3$ to ignore the ceiling on p and ρ, and account for the basic intervals for small cliques in case there are no large cliques. The 2^{k-1} intervals in the generalized displayed representation of large cliques can include the basic intervals for those vertices. ∎

Lemma 2.14. *If k is the maximum size of a clique in a lexicographically minimal coloring of $\bar{G}$, then $i(G) \leq \frac{1}{2}\sqrt{e(\bar{G})} + 2^{k-1} + 4$.*

Proof. We compare the key quantity in Lemma 2.13 with the number of edges in $\bar{G}$. Let

$$f = \left(\frac{c_2(c_1+c_2)}{c_1+2c_2}(c_1+c_2) + (c_1+c_2)c_3 + \binom{c_3}{2} \right) \frac{1}{c_1+c_2+c_3}.$$

Observe that

$$\frac{2c_2(c_1+c_2)^2}{c_1+2c_2} = (c_1+c_2)^2 - \frac{c_1(c_1+c_2)^2}{c_1+2c_2}.$$

Thus

$$2f = \left(\frac{2c_2(c_1+c_2)^2}{c_1+2c_2} + 2(c_1+c_2)c_3 + c_3^2 - c_3 \right) \frac{1}{c_1+c_2+c_3} \leq$$

$$\left((c_1+c_2)^2 + 2c_1c_3 + 2c_2c_3 + c_3^2 - \frac{c_1(c_1+c_2)^2}{c_1+2c_2} \right) \frac{1}{c_1+c_2+c_3} =$$

$$c_1+c_2+c_3 - \frac{c_1(c_1+c_2)^2}{(c_1+c_2+c_3)(c_1+2c_2)}.$$

To obtain an useful bound on $(2f)^2$, we need to show that

$$\frac{2c_1(c_1+c_2)^2}{c_1+2c_2} - \frac{c_1^2(c_1+c_2)^4}{(c_1+c_2+c_3)^2(c_1+2c_2)^2} \leq c_1(c_1+c_2).$$

To this end it suffices to consider $c_3 = 0$, get rid of fractions, and compare terms. As a result,

$$(2f)^2 \leq (c_1+c_2+c_3)^2 - c_1(c_1+c_2).$$

Rearranging the bound on $e(\bar{G})$ from Lemma 2.11 yields

$$e(\bar{G}) \geq (c_1+c_2+c_3)^2 - c_1(c_1+c_2) + \frac{1}{2}(c_1-c_3)^2 - \frac{1}{2}(c_1+2c_2+c_3).$$

It is easy to check by Lemma 2.11 that $c_1 + 2c_2 + c_3 \leq 3\sqrt{e(\bar{G})}$. From here we obtain

$$(2f)^2 \leq e(\bar{G}) + \frac{3}{2}\sqrt{e(\bar{G})} \leq e(\bar{G}) + 2\sqrt{e(\bar{G})} + 1,$$

and thus $2f \leq 1 + \sqrt{e(\bar{G})}$. The result now follows from Lemma 2.13.∎

When k is large, this bound is not very good; what can we do with the large cliques? We cut them into small pieces to obtain the main theorem. Hence, we are ready to prove Theorem 2.6.

Proof. If the size of a color class in the lexicographically minimal coloring of $\bar{G}$ exceeds s, a value to be chosen later, then we cut it into sets of sizes between $s/2$ and s. This yields at most $2n/s$ new sets; we consider one piece from each color class to be an old set. Let G' be the graph induced by the small classes and the old pieces. Representing it as in Lemma 2.13 uses at most $\frac{1}{2}\sqrt{e(\bar{G})} + 2^{s-1} + 3$ intervals per vertex.

There are at most $2n/s$ new pieces; we apply the Scheinerman representation. We establish generalized displayed representations for the cliques that are the new pieces. We represent the edges between old and new pieces using at most $2n/s$ additional intervals for each vertex in an old piece, and we represent the edges between new pieces using at most n/s additional intervals for each vertex in a new piece. This yields

$$i(G) \leq \frac{1}{2}\sqrt{e(\bar{G})} + 2^{s-1} + 2 + 2n/s,$$

which implies Theorem 2.6 by setting $s = (\log n - \log\log n - \log\log\log n)/\log 2$.∎

Theorem 2.15. *[51] If G is an n-vertex graph with $\Delta(G)$ less than an integer l, then $V(G)$ can be partitioned into l independent sets whose size differ by at most 1.*

Theorem 2.7 follows from Theorem 2.15:

Proof. Applying the method of Theorem 2.6 to a coloring of $\bar{G}$, described in Theorem 2.15, yields Theorem 2.7. ∎

Similarly, we can prove Theorem 2.8.

Proof. An optimal coloring of $\bar{G}$ yields a partition of $V(G)$ into $\chi(\bar{G})$ cliques in G. We represent each clique using one interval per vertex.

In order to represent the edges from a vertex x to a clique Q, we can place a small interval for x into the clique for Q and cut a hole in the intervals for each vertex of Q that is not adjacent to x. When we do this with many vertices, the interval for a vertex $z \in Q$ breaks into at most $\Delta(\bar{G})$ parts.

As usual, we distribute the ordered pairs of cliques uniformly, so that we insert an interval for vertex x into at most $\frac{1}{2}\chi(\bar{G})$ other cliques. This gives the claimed bound. ∎

Now we can easily prove Corollary 2.9:

Proof. We need to show that $i(G) \leq \frac{3}{2}(\Delta(\bar{G}) + 1)$. This immediately follows from Theorem 2.8, since $\chi(\bar{G}) \leq \Delta(\bar{G}) + 1$. ∎

Chapter 3

Further results on the Interval Number of a Graph

The interval number of a graph G is the least natural number t such that G is the intersection graph of sets, each of which is the union of at most t intervals, denoted by $i(G)$. Griggs and West showed that $i(G) \leq \lceil \frac{1}{2}(d+1) \rceil$. We describe the extremal graphs of that inequality when d is even. For three special perfect graph classes we give bounds on the interval number in terms of the independence number. Finally we show that a graph needs to contain complete bipartite subgraphs in order to have interval number larger than the random graph on the same number of vertices.

3.1 Introduction

One way is to represent a graph G is the intersection representation. That is, one assigns a set to each vertex of G such that v is adjacent to w if and only if their assigned sets meet. A *t-interval representation* is an assignment, where each set consists of at most t closed intervals. The *interval number* of G, denoted by $i(G)$, is the least integer t for which a t-representation of G exists.

Furthermore a representation is *displayed* if each set of the representation has an open interval disjoint from the other sets. Such an interval is called *displayed segment.*

3.1.1 The Degree Bound

Theorem 3.1. *[50] If G is a graph with maximum degree d, then $i(G) \leq \lceil \frac{1}{2}(d+1) \rceil$.*

This upper bound is sharp, because for a triangle-free graph G one also has the lower bound $i(G) \geq \lceil (e(G)+1)/v(G) \rceil$ (see also in [50]). Thus the equality is attained for example the d-regular, triangle-free graphs. The original proof of Theorem 3.1 was greatly simplified in [82] and [60]. However, the following statement of [60] turned out to be false.

Theorem 3.2. *[60] If a graph G has no d-regular, K_3-free component, then $i(G) \leq \lceil \frac{1}{2} d \rceil$.*

In the next section give a counterexample that helps to understand the following definitions. First of all, let us call a graph G *even* if all its degrees are even. An *even cut* of G means the following operation.

We select a vertex x such that the remaining vertices of G can be divided into sets A and B, and

- there are no edges between A and B,
- the degree of x in A is even (of course the same holds for the degree of x in B).

Then we double the vertex x, (making vertices x and x') and consider two disjoint graphs; one is induced by the vertex set $A \cup \{x\}$, while the other by the set $B \cup \{x'\}$. We call the graphs arising by the repeated execution of this operation an *even decomposition* of G, and x as a *pivot vertex* of the cut.

A connected graph H is *rich* if it contains such a triangle T_H that after deleting the edges of T_H the remaining graph has at most two non-trivial components. (A non-trivial component consist of at least two vertices; here, because of the nature of decomposition it consists of at least three vertices.)

Note that the triangle T_H is not necessarily unique, but it does not make any difference for us.

Finally for an even graph G a *rich decomposition* is such a decomposition in which the arising components (blocks) are all rich. With these notions we can spell out the correct characterization of extremal graphs.

Theorem 3.3. *Let G be a connected graph of maximal degree $d = 2k$.*

Then G has a k-interval representation if and only if either G is not $2k$-regular or G has a rich decomposition.

3.1.2 Perfect Graphs and Forbidden Subgraphs

Most of the classical results on the interval number were upper bounds for a general graph G, in terms of some monotone increasing functions of G. (See for example [50], [15], [69] and [71].) It is also possible to derive bounds which are better for denser graphs. See in [16] that $i(G) \le \lceil \sqrt{e(\bar{G})/2} \rceil + o(n)$, where $e(\bar{G})$ and n are the number of edges in the complement of G and the size of vertex set of G, respectively.

Another interesting direction is to consider the interval number of perfect graphs. We refer to [48] as a vast collection of facts on perfect graphs, but also try to spell out the used facts on those.

A graph G is *triangulated* or *chordal* if for $n \ge 4$ it does not contain an induced C_n, a circuit of length n. A special case of chordal graphs are the *split* graphs, for which the set of vertices can be partitioned in such a way that one class is an independent set, while the other is a clique.

G is a *comparability* graph if there is a partial order $\mathcal{P}$ on its vertices such that an edge $(u, v) \in E(G)$ iff u and v is comparable by $\mathcal{P}$. Finally $\omega(G)$ and $\alpha(G)$ stands for the clique number of G and the independence number of G, respectively.

It was shown in [9] that $i(G) \le \lceil \omega(G)/2 \rceil + 1$ for a chordal graph G. An even better result can be found in [68], stating $i(G) \le (1 + o(1))\omega(G)/\log_2 \omega(G)$. The following results have the similar flavor stating bounds on the interval number for perfect graphs such that the denser the graph the smaller the bound.

Theorem 3.4. *If G is a chordal graph, then $i(G) \leq \lceil \alpha(G)/2 \rceil$.*

Theorem 3.5. *If G is a split graph, then $i(G) \leq (1+o(1))\alpha(G)/\log_2 \alpha(G)$, and there is a sequence of split graphs G_k such that $\alpha(G_k) = k$ and $i(G_k) > (1/2 + o(1))k/\log_2 k$, where k goes to infinity.*

Theorem 3.6. *If G is a comparability graph, then $i(G) \leq \alpha(G)$.*

Trotter and Harary showed that for the complete bipartite graph $K_{m,n}$ holds the equality $i(K_{m,n}) = \lceil (mn+1)/(m+n) \rceil$. For complete multipartite graphs (see [52]), one have $i(K_{n_1,n_2,\ldots}) \leq i(K_{n_1,n_2}) + 1$ with equality for certain values, where $n_1 \geq n_2 \geq \ldots$ One checks that in both cases $i(G) = \lceil (\alpha(G) + 1)/2 \rceil$.

Note that $K_{m,n}$ and the complete multipartite graphs are comparability graphs. We believe a common generalization of these results, that is

Conjecture 3.7. *If G is a comparability graph, then $i(G) \leq \lceil (\alpha(G) + 1)/2 \rceil$.*

Practically all known perfect graph classes can be characterized by forbidden induced subgraphs [48].

One may ask, what is the reason of high interval number in terms of induced subgraphs? For the random graph $G_{n,1/2}$ the equality $i(G_{n,1/2}) = (1/2 + o(1))n/\log_2 n$ holds almost surely [69]. If G contains the complete bipartite graph $K_{k,k}$ as an induced subgraph, then $i(G) \geq \lceil (k + 1)/2 \rceil$.

The following result roughly states that the big induced complete bipartite graphs are responsible for the unusually high interval number.

Theorem 3.8. *Let k be a positive integer. If a graph G does not contain $K_{k,k}$ as an induced subgraph, then* $i(G) \leq (1 + o(1))n/\log_2 n$.

Remark. We do not know matching lower bound here. Standard use of the probabilistic method (see e. g. [40]) shows the existence of a bipartite $K_{r,r}$-free graph G which has $2n$ vertices and $n^{2-2/r}$ edges. Applying the formula $i(G) \geq \lceil (e(G)+1)/v(G) \rceil$ to G, which is triangle-free of course, we get that $i(G) = \Omega(n^{1-2/r})$.

3.2 Counterexamples to Theorem 3.2

To construct a minimal counterexample G for Theorem 3.2, we need copies of triangle-free graphs H and Q. Except for one special vertex, all vertices are of degree six in both H and Q. The special vertex is of degree four in H, and of degree two in Q. Let us take a graph on five vertices h_1, h_2, h_3, h_4 and q such that q is connected to all the others, h_1 is connected to h_2 and h_3 is connected to h_4. To build up G, we take four disjoint copies of H and one copy of Q, and identify their special vertices (one-to-one) with h_i, $i = 1, \cdots, 4$ and q, respectively.

The interval number of the resulting graph is more than 3; since the proof is rather straightforward we give only a sketch. (See a similar argument in [50].) So let us suppose on contrary that G has a 3-representation. Consider the intervals assigned a copy of H (or Q) in that representation;

it is easy to see that it must start and end with intervals assigned to the special vertex, and exactly three intervals were used up for every vertex.

Now one can check that at least four intervals for q are needed to represent the leftover edges.

Remark. This example led to the right formulation of Theorem 3.3, namely the need of the notion of rich decompositions. The first counterexample to Theorem 3.2 was the following. Take three disjoint copies of a triangle-free graph H, such that there is a vertex $h \in v(H)$ of degree 2, while all the other vertices are of degree 4. Then connect the special vertices to each other. Now one easily verifies that for the resulting graph G, $i(G) = 3$.

3.3 Proofs

For the sake of completeness we repeat the proof of Theorem 3.1 and the first part that of Theorem 3.2 given in [60]. Note, that the statement of Theorem 3.1 has been slightly rephrased in order to make its connection to Lemma 3.10 more apparent.

Theorem 3.9 (modified version of Theorem 3.1). *There is a displayed interval representation for the graph G such that at most $\lceil (d(v)+1)/2 \rceil$ intervals are assigned to each vertex v, where $d(v)$ designates the degree of the vertex v.*

3.3.1 Proof of Theorem 3.1

A *walk* W in G is a sequence of vertices $W = \{v_1, v_2, ..., v_l\}$ such that there is an edge between v_i and v_{i+1} for each $i = 1, 2, ..., l-1$, and all these edges are different from each other. Let us partition the edges of G into minimal number of edge disjoint walks $\{W_i\}_{i=1}^{j}$. Now represent the walk $W_i = (v_1^i, v_2^i, ..., v_{n(i)}^i)$ for $1 \leq i \leq j$, assigning an I_p^i interval to the vertex v_p^i such that two intervals meet if and only if the corresponding vertices are next to each other in the walk W_i.

This procedure leads to a displayed interval representation of G. Since a vertex v can be an endvertex of the walks at most two times, if v is represented by l intervals, then $d(v) \geq 2(l-2) + 2 = 2l - 2$. Hence

$$\lceil \frac{1}{2}(d(v)+1) \rceil \geq \lceil \frac{1}{2}(2l-2+1) \rceil = \lceil l - \frac{1}{2} \rceil = l.$$

□

3.3.2 Proof of Theorem 3.3

First we show that if a connected graph G has maximum degree $2k$ and $i(G) = k+1$, then G must be $2k$-regular.

Let G be a graph with $i(G) = k+1$ and maximal degree $2k$. Let us choose among all partitions of the edge set into a minimum number of edge disjoint walks a partition $\{W_i\}_{i=1}^{j}$ which also minimizes the size of the set Q of vertices occurring $k+1$ times in the walks $\{W_i\}_{i=1}^{j}$. The representation is the same as in the proof of Theorem 3.1.

If $Q = \emptyset$, we are done. For an $x \in Q$ there exists a $p \in \{1, ..., j\}$ such that $x = v_1^p$, $x = v_{n(p)}^p$ and $x \notin W_l$ for all $l \neq p$. The last statement follows from the minimality of j, since in case of $x = v_s^l \in W_l$ we could replace the walks W_p and W_l by the walk

$$W^* = (v_1^l, v_2^l, ..., v_s^l, v_2^p, ..., v_{n(p)}^p, v_{s+1}^l, ..., v_{n(l)}^l).$$

For any vertex $y = v_s^p \neq x$ from W_p, we can transform the walk W_p into the walk

$$W_p^* = (v_s^p, v_{s-1}^p, ..., v_1^p, v_{n(p)-1}^p, v_{n(p)-2}^p, ..., v_s^p).$$

That is, by the minimality of Q, y occurs in the walks $\{W_i\}_{i\neq p} \cup \{W_p^*\}$ $k+1$ times. Then again, all neighbors of y are in W_p. That is the vertex set of W_p is a $2k-$regular component of G. But G is connected, so in fact G itself is $2k$-regular.

In order to finish the proof of Theorem 3.3, we need to show that for a $2k$-regular graph G the assumptions $i(G) \leq k$ and G has a rich decomposition are equivalent. This is done by the following lemma.

Lemma 3.10. *If G is an even graph, then there is a interval representation of G assigning at most $d(v)/2$ intervals to each vertex v if and only if G has a rich decomposition.*

3.3.2.1 Proof of Lemma 3.10

First we assume that G has rich decomposition. It is enough to prove Lemma 3.10 for the blocks of the rich decomposition of G, since putting together their representation we get the stated representation for G.

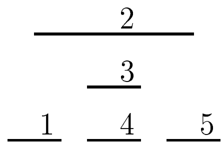
Having a rich component H, we can represent the edges of the triangle T_H by using three overlapping intervals, one for each vertices of T_H. However, it is crucial to put such two intervals to the two sides of that sub-representation which would fall into different non-trivial components after deleting the edges of T_H. Now to represent the non-trivial components of $H \setminus T_H$, we use the ideas of Theorem 3.1. Since such a component is an even graph, it has an Eulerian circuit. We make a walk out of that circuit by starting it and finishing it with a vertex of T_H. Then we represent these (at most two) walks, and identify one of their end interval with the appropriate interval representing T_H.

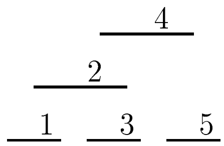
To prove the other direction, we assume that G is connected and that G has an above mentioned representation. Then there is a point which is the element of at least three intervals. (Imagine that we draw the intervals one by one. If there is no such a point, then every new interval may represent at most two units of the degree sum of G (or with other words, at most one edge). But the first drawn interval does not represent any edges, which gives the claim.) We may also assume that G does not contain $K_4 \setminus e$, i. e. a clique minus an edge on four vertices, since then G would be itself a rich component.

Let us scan the representation from the left, and stop at the first point where three intervals meet. This corresponds to a T_G triangle of G. If the deletion of the edges of T_G leaves one or two connected non-trivial component, then we are ready or we can proceed by induction. Indeed, in the second case one of the components contains exactly one vertex of

T_G, and this vertex is suits to be a pivot vertex for an even cut. (Note that after executing an even cut operation, there must be triangles in each non-trivial components.)

In case of three non-trivial components, because of our previous assumptions, there are two possible arrangements for the intervals around the ones representing T_G.





In the first case the vertices $\{1, 2, 5\}$, $\{3\}$ and $\{4\}$ are in different components, while in the second the $\{1, 2\}$, $\{3\}$ and $\{4, 5\}$ are. In either cases the vertex 3 can be a pivot for a cut, and we can finish the proof of the statement by induction. $\square$

One applies Lemma 3.10 to a $2k$-regular graph G, and gets Theorem 3.3. $\square$

Remark. Lemma 3.10 holds for general simple graphs; one just have to write $\lceil d(v)/2 \rceil$ instead of $d(v)/2$ in the statement. The way of proof is standard: join a new vertex x to G, and connect it to all vertices of odd degree.

3.3.3 An application

It was shown in [71] that $i(G) \leq 3$ for any planar graph G. By the help of Theorem 3.3 we can point out an interesting special case.

Theorem 3.11. *Let G be a planar graph with maximal degree at most* 4. *Then $i(G) \leq 2$.*

Proof of Theorem 3.11. One can assume that G is connected. If G is not 4-regular, then we are ready by Theorem 3.3, that is at the rest of the proof we can assume that G is 4-regular.

Since G is planar, the Euler formula $n(G) + l(G) = e(G) + 2$ holds, where $n(G)$, $l(G)$ and $e(G)$ are the number of vertices, domains and edges, respectively. G must contain a triangle, since otherwise the Euler formula would give that $2n(G) \geq e(G) + 6$, while the 4-regularity implies that $2n(G) = e(G)$. Our aim is to show that G has rich decomposition. Let us assume that the deletion of the edges of any triangle leads to exactly three non-trivial components. The numbers of vertices in the non-trivial components vary, let T be a triangle such that one of these components, say H, is the smallest possible.

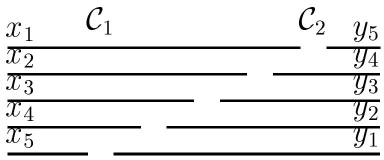
Of course $2n(H) \geq e(H) + 6$, since H is planar and H cannot contain a triangle. On the other hand one vertex of H has degree 2, all the others have degree 4, giving us the equality $2n(H) = e(G) + 1$, which is a contradiction. It means that there must be a triangle such that the deletion of its edges results in less than three non-trivial components, and now Theorem 3.3 implies our claim. $\square$

3.4 Proof of Theorem 3.4

First we prove a slightly stronger lemma. For a graph G let $\chi(G)$ be the *chromatic number* of G, that is the smallest number k such that the vertex set of G can be partitioned into k independent sets.

Lemma 3.12. *If G is an induced C_4-free graph, then $i(G) \leq \lceil \chi(\bar{G})/2 \rceil$.*

Proof of Lemma 3.12. The idea of the proof is to partition the vertex set of G into $\chi(\bar{G})$ cliques, then represent the edges in the cliques and among the cliques. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two arbitrary cliques of that partition. Then the vertices of $\mathcal{C}_1 = x_1, ..., x_l$ can be ordered in such a way that $N_2(x_1) \supset N_2(x_2) \supset ... \supset N_2(x_l)$, where $N_2(x)$ denotes the vertices of $\mathcal{C}_2$ neighboring to x, since G is C_4-free. To represent the edges in and between $\mathcal{C}_1$ and $\mathcal{C}_2$ we need to draw two piles of intervals containing only one interval per vertex.



Of course we can use both ends of these intervals (the left ends of those assigned to $\mathcal{C}_1$ and the right ends of those assigned to $\mathcal{C}_2$). One can think of this as a generalization of the proof of Theorem 3.1, where the "degree" of each clique is not more than $\chi(\bar{G}) - 1$. That is in the worst case we need to use up at most $\lceil \chi(\bar{G})/2 \rceil$ piles, and therefore the same number of intervals per vertex. □

To finish the proof of Theorem 3.4, we recall some basic facts about perfect graphs, see the details in [48]. Chordal and comparability graphs are perfect. A graph G is perfect iff for every induced subgraph $H \subset G$ it holds that $\omega(H) = \chi(H)$. The complement of a perfect graph is perfect (this is the celebrated Perfect Graph Theorem).

A chordal graph G is induced C_4-free, and since it is perfect we have $\chi(\bar{G}) = \alpha(G)$ and Lemma 3.12 applies, giving Theorem 3.4. □

3.5 Proof of Theorem 3.5

For any k positive integer, we define the *universal split graph* G_k.

The vertex set of G_k is indexed by the numbers $\{1, ..., k\} = A$, and the non-empty subsets of $\{1, ..., k\}$. There are no edges inside A, while $V(G_k) \setminus A$ is a clique. Furthermore a vertex $i \in A$ is connected to a vertex indexed by $S \subset A$ if and only if $i \in S$. A graph H is *essential* if $N(x) \neq N(y)$ for $x \neq y \in V(H)$. G_k contains every connected essential split graph H of independence number k as an induced subgraph, and also $\alpha(G_k) = k$. To conclude the proof it is enough to show that $(1/2 - o(1))k/\log_2 k \leq i(G_k) \leq (1 + o(1))k/\log_2 k$. (If a split graph H is not essential, then we may represent the biggest induced essential subgraph H^* of it, and take a copy of the intervals of $x \in V(H^*)$ for a $y \notin V(H^*)$ if $N(x) = N(y)$.)

First we show that $i(G_k) \geq (1/2 - o(1))k/\log_2 k$. Without loss of generality we may assume that in a t-representation of G_k there are exactly

t intervals corresponding to each vertex of A, and all of those intervals degenerated, that is those are single points of the line, which set is denoted by $\mathcal{A}$.

Clearly, $|\mathcal{A}| = kt$. Let I_x be the set of intervals representing vertex x. If $x \notin A$, then $\mathcal{A} \cap I_x$ consists of at most t consecutive segments of $\mathcal{A}$. This means, the number of such sets is not more than $\sum_{l=1}^{t} \binom{kt+1}{2l}$. For vertices $x \neq y$ we have $\mathcal{A} \cap I_x \neq \mathcal{A} \cap I_y$ implying the inequality

$$2^k - 1 \leq \sum_{l=1}^{t} \binom{kt+1}{2l} \leq t \binom{kt+1}{2t}.$$

Applying the inequality $\binom{a}{b} \leq (ea/b)^b$ it reduces to $t \geq (1/2 - o(1))k/\log_2 k$. Note, that a very similar argument is used in [68]. It is essential that G is induced C_4-free, since $i(G^*) = \Omega(k/\log k)$, where G^* is the same as G, except that A is a clique in it.

For the upper bound let us divide A into blocks of size $t = \log_2 k - \log_2 \log_2 k$. Then for each subset of a block we secure an unused part of the line, and place small disjoint intervals for each element of the subset. This way we use up not more than $2^t \leq k/(2\log_2 k)$ intervals per vertex. To represent the edges between A and a vertex x of the clique on $V(G_k) \setminus A$, we use one interval per blocks, putting a long interval under the intervals corresponding to the subset $A \cap N(x)$. Finally we represent the edges of the clique by one interval per vertex. Since this procedure needs only $\max(2^t, k/t) + 1$ intervals per vertex, we are ready. $\square$

3.6 Proof of Theorem 3.6

Indeed the proof closely follows those of Theorem 3.4, and of the following lemma.

Lemma 3.13. *Let us suppose that a graph G can be partitioned into k cliques such that the subgraphs induced by any two of those cliques are comparability graphs. Then $i(G) \le k$.*

Proof of Lemma 3.13. We imitate the proof of Lemma 3.12. However, the structure of the neighborhoods is more complicated here. Still, the vertices of a $\mathcal{C}_1 = x_1, ..., x_l$ can be ordered. (It is a clique, so the partial order assumed is in fact a linear order.) Now $N_2(x) = N_2^<(x) \cup N_2^>(x)$, where $N_2^<(x)$ is the set of those vertices in $\mathcal{C}_2$ that are smaller than x in the partial order of the induced subgraph, while $N_2^>(x)$ is the subset of $\mathcal{C}_2$ consisting of the vertices bigger than x. It is easy to see that

$$N_2^<2(x_1) \supset N_2^<(x_2) \supset ... \supset N_2^<(x_l) \text{ and } N_2^>2(x_1) \subset N_2^>(x_2) \subset ... \subset N_2^>(x_l).$$

We may represent only one type of those neighborhoods by drawing two piles, although the free ends of the piles are re-usable again. Taking this, and the previous considerations into account, we get Lemma 3.13. □

Lemma 3.13 implies Theorem 3.6, since the subgraphs of a comparability graph are also comparability graphs. □

3.7 Proof of Theorem 3.8

We need a result of Erdős and Hajnal, see [34]. A set $L \subset V(G)$ is *homogeneous* if L spans either a clique or the vertices of L are independent. If a graph G on n vertices does not contain a fixed graph H as an induced subgraph, then it contains a large homogeneous set, with size at least $\exp(c\sqrt{\log n}/2)$, where $c < 1/|V(H)|$. If n is big enough, then $\exp(c\sqrt{\log n}/2) >> \log^{2k} n$ for any k.

We use repeatedly this theorem, always cutting off a big homogeneous set L from the remaining part of $V(G)$. If L is a clique, we take about $\log_2 n - \log_2 \log_2 n$ vertices of it, otherwise (i. e. if L is an independent set) subset of size between $\log_2^{2k+1} n$ and $2\log_2^{2k+1} n$ is taken. This procedure can go on until the leftover graph has only $o(n/\log_2 n)$ vertices.

Summing up, the procedure gives a partition of the vertices of G:

$$V(G) = K_1 \cup \cdots \cup K_\ell \cup E_1 \cup \cdots \cup E_r \cup A,$$

where

- $|A| = o(n/\log n)$,
- each K_i is a clique of size $\log n - 2\log\log n$,
- each E_j is an independent set such that $\log^{2k+1} n \leq |E_j| \leq 2\log^{2k+1}$.

Now we represent the edges of G step by step, starting by the ones that having an endpoint in A, then those having both ends in $\cup_i K_i$, finally the edges having one endpoint in $\cup_j E_j$.

First let us build up a displayed interval system for the vertices of A, and represent on them the edges with at least one endpoint on A. This contribute to the interval number with at most $o(n/\log n)$ as the size of A is $o(n/\log_2 n)$.

For the cliques, we construct a Scheinerman-type of displayed system, see [69] or [16]. It means that for a clique Q of size q one places 2^{q-1} intervals per vertex in such a way that for every subset P of Q there exists an interval I_P of the line where exactly some intervals representing P meet. The use of such system is that one can represent all edges between a vertex x and Q by only one additional interval of x, placing that interval into $I_{N(x)\cap Q}$. Note that in our case $q = \log_2 n - 2\log_2\log_2 n$ that is $2^{q-1} = n/2(\log_2 n)^2$.

Note that when we represent the edges among the cliques, and pick a pair (K_i, K_j) of those, then we need to use up the additional one intervals per vertex only for one of those. That is for all vertices of $\cup_i K_i$ the number of additional intervals will be smaller than $\ell/2$, where ℓ is the number of cliques.

This way for a vertex in a clique we will use up at most

$$n/2(\log n)^2 + \ell/2 + 1 = (1/2 + o(1))n/(\log_2 n)$$

intervals per edges.

Finally for a vertex in an independent set we use no more than $n/(\log_2 n - 2\log_2\log_2 n) = (1 + o(1))n/\log_2 n)$ intervals to represent its edges to the cliques.

The number of the edges between two independent sets E_i and E_j is bounded by $cm^{2-1/k}$ $(c < 2)$, where $m = \max\{|E_i|, |E_j|\}$, because G is $K_{k,k}$-free, see [57].

With this we can bound the number of edges among the independent sets of the partition. Since $\log_2^{2k+1} n \leq |E_j| \leq 2\log_2^{2k+1} n$, we have that $r \leq n/(\log_2 n)^{2k+1}$ and $m \leq 2(\log_2 n)^{2k+1}$. Thus the number of edges is no more than

$$\binom{r}{2} 2m^{2-1/k} \leq \frac{4n^2}{(\log_2 n)^{2+1/k}}.$$

In order to represent the edges among all E_j we use the edge-bound theorem from [15], which states that for a graph G, having e edges, holds the inequality $i(G) \leq \lceil \sqrt{e}/2 \rceil + 1$.

Plugging the estimation above in the formula we get a representation of that part of G using only $o(n/\log_2 n)$ intervals per vertex.

Putting together all these bounds on the representations of the subgraphs of G we get that $i(G) \leq n/\log_2 n + o(n/\log_2 n)$, which proves Theorem 3.8. $\square$

Chapter 4

Game Domination Number

4.1 Introduction

The game domination number of a (simple, undirected) graph is defined by the following game. Two players, A and D, orient the edges of the graph alternately until all edges are oriented. Player D starts the game, and his goal is to decrease the domination number of the resulting digraph, while A is trying to increase it. The *game domination number* of the graph G, denoted by $\gamma_g(G)$, is the domination number of the directed graph resulting from this game. This is well-defined if we suppose that both players follow their optimal strategies.

We determine the game domination number for several classes of graphs and provide general inequalities relating it to other graph parameters.

A *dominating set* of a digraph $\vec{G}$ is a set S of vertices such that for every vertex $v \notin S$ there exists some $u \in S$ with $\vec{uv} \in E(\vec{G})$. The *domination* *number* $\gamma(\vec{G})$ of $\vec{G}$ is defined as the cardinality of the smallest dominating set.

We define a "domination parameter" of an undirected graph G as the domination number of one of its orientations, determined by the following two player game. Players A and D orient the unoriented edges of the graph G alternately with D playing first, until all edges are oriented. Player D (frequently called the *Dominator*) is trying to minimize the domination number of the resulting digraph, while player A (*Avoider*) tries to maximize the domination number. This game gives a unique number depending only on G, if we suppose that both A and D play according to their optimal strategies. We call this number the *game domination number of* G and denote it by $\gamma_g(G)$.

As the domination number of any orientation of a graph is at least as large as the domination number of the graph itself, we clearly have $\gamma(G) \leq \gamma_g(G)$. Also $\gamma_g(G) \leq DOM(G)$, where $DOM(G)$ denotes the maximal domination number among all orientations of G. This parameter was examined in [28].

Similar orientation games with different goals for the players were introduced and discussed in [1], [8], [20] and [27].

In Section 4.2 we determine the game domination number for several classes of graphs including complete graphs, complete bipartite and tripartite graphs, paths and cycles. Then we obtain sharp lower and upper bounds for the game domination number of trees in terms of the smallest degree that is at least three.

Finally, in Section 4.4 we prove several inequalities, relating the game domination number to other graph parameters such as the number of vertices and edges, independence number and 2-domination number. We establish a Nordhaus-Gaddum type upper bound for the sum of the game domination number of a graph and its complement.

For additional results on related domination parameters we refer the reader to two excellent books [45] and [46].

4.2 Examples

In this section we determine the game domination number for a few classes of graphs. These elementary examples enable the reader to gain a feel for the parameter; also, some of the examples will be needed in the sequel.

Example 4.2.1. *For the complete graph K_n on $n \geq 4$ vertices, we have $\gamma_g(K_n) = 2$.*

Proof. Let us see first why $\gamma_g(K_n) \geq 2$, i.e. why one vertex cannot dominate the oriented graph. Player A can clearly avoid a source in K_4, and if $n \geq 5$ then there exists a collection of n edge-disjoint paths of length 2, one centered at each vertex (see [20]). Whenever D orients one of these edges from the central vertex, A can orient the other edge of the corresponding path towards the central vertex. Thus the in-degree of each vertex becomes at least one.

On the other hand, $\gamma_g(K_n) \leq 2$ since the dominator can pick two vertices, u and v, say and then reply to each move $\vec{wu}$ by $\vec{vw}$, and to each move $\vec{wv}$ by $\vec{uw}$. This strategy ensures that $\{u, v\}$ becomes a dominating set of the resulting digraph. $\square$

Note, that in Example 4.2.1 the dominator made use of very few edges. The same idea can be applied for much sparser graphs.

Example 4.2.2. *Let G be a graph on $n \geq 4$ vertices containing all but one edges of a copy of $K_{2,n-2}$. Then $\gamma_g(G) = 2$. Also, if G contains a set S of s vertices such that every vertex not in S has at least 2 neighbors in S, then $\gamma_g \leq s$.* $\square$

Example 4.2.3. *Let $K_{n,m}$ denote a complete bipartite graph with $n \leq m$ vertices in the two partite sets, then*

$$\gamma_g(K_{n,m}) = \begin{cases} \lceil (m+1)/2 \rceil & \text{if } n = 1 \\ 2 & \text{if } n = 2 \\ 3 & \text{if } n = 3, 4 \text{ or } 5 \\ 4 & \text{otherwise.} \end{cases}$$

$\square$

Example 4.2.4. *If G is a complete k-partite graph ($k \geq 3$) with at least three vertices in each partite class, then $\gamma_g(G) = 3$.* $\square$

Now we turn to some sparser graphs that (as expected) have larger game domination numbers. They also show that the game domination number can be much larger than the domination number. The reader is encouraged to verify the statement of the next example.

Example 4.2.5. *For the three dimensional cube Q_3 and the Petersen graph P we have $\gamma_g(Q_3) = 3$ and $\gamma_g(P) = 4$.* $\square$

Example 4.2.6. *For a path P_n on n vertices we have $\gamma_g(P_n) = \lceil n/2 \rceil$.*

Proof. The vertices of the path can be partitioned into $\lceil n/2 \rceil$ sets of disjoint edges and possibly one single vertex. Each of these sets can be dominated by one of its vertices regardless of the orientation of the edges, showing $\gamma_g(P_n) \leq \lceil n/2 \rceil$.

For the lower bound, A would like to prevent D creating many vertices that dominate both their neighbors. Although A cannot do this, he can easily achieve that no even numbered vertex (with the vertices of the path labeled with $1, 2, \ldots, n$ from left to right) dominates three vertices (including itself). Indeed, whenever D orients an edge out of an even vertex, A immediately orients the other edge in.

This strategy results in an oriented path, where to dominate the $\lceil n/2 \rceil$ odd vertices we must choose $\lceil n/2 \rceil$ vertices, as no vertex will dominate two odd vertices. $\square$

Example 4.2.7. *For a cycle C_n on n vertices, $\gamma_g(C_n) = \lfloor \frac{n}{2} \rfloor$.*

Proof. We show first that D can achieve an orientation with domination number $\lfloor \frac{n}{2} \rfloor$. Indeed, following his second move D can make sure that there is a vertex dominating both of its neighbors. The remaining $n-3$ vertices can be partitioned into $\lceil \frac{n-3}{2} \rceil$ independent edges (with possibly one single vertex), and these vertices will be dominated by $\lceil \frac{n-3}{2} \rceil$ vertices regardless of the orientation. Thus $\gamma_g(C_n) \leq n/2$.

On the other hand player A can force the dominating set to be as big as $\lfloor \frac{n}{2} \rfloor$ using the same idea as in case of paths: he labels the vertices by $1, 2, \ldots, n$, and ensures that no even vertex dominates both of its neighbors. Then, to dominate the $\lceil \frac{n}{2} \rceil$ odd vertices we need at least $\lfloor \frac{n}{2} \rfloor$ vertices, giving $\gamma_g \geq \lfloor n/2 \rfloor$. $\square$

Our next example is a family of less natural graphs, this result will be used later.

Example 4.2.8. *Let G be a "lollipop" on n vertices formed by an even cycle with a tail (a single path) attached to one of its vertices. Then $\gamma_g(G) = \lfloor \frac{n}{2} \rfloor$.*

Proof. To prove that $\gamma_g(G) \leq \lfloor \frac{n}{2} \rfloor$, write v for the vertex of degree 3, and u for its neighbor on the path. The dominator D starts the game with $\vec{vu}$, and in his second move also orients an edge away from v. Thus v dominates 3 vertices (including itself), and as the cycle is even, the remaining vertices can be partitioned into a matching (with possibly a singleton), showing that $\gamma_g(G) \leq \lfloor \frac{n}{2} \rfloor$. The lower bound can be shown as in the previous examples. $\square$

First it seems that by adding edges to a graph we cannot increase its game domination number. Indeed, this is clearly the case if we add an even number of edges to a graph. However, rather surprisingly, this does not hold if we add exactly one edge to our graph.

Example 4.2.9. *Let G be obtained from the complete bipartite graph $K_{t,4}$ ($t \geq 6$) as follows. Let $K_{t,4} = K(M, N)$ with $M = \{e, f, g, \ldots, z\}$, $N = \{a, b, c, d\}$), and $G = K_{t,4} + ab + cd - dz$. Then $\gamma_g(G) = 2$, while $\gamma_g(G + dz) = 3$.*

Proof. Player D has an easy strategy to finish the game with a two-element dominating set: he starts with $\vec{az}$, then he ensures that every vertex in M is dominated by at least one vertex of $\{a, d\}$ and one vertex of $\{b, c\}$. Whenever A plays ab or cd, the dominator orients the other so that either $\{a, d\}$ or $\{b, c\}$ becomes the dominating set. Hence $\gamma_g(G) \leq 2$.

What about $\gamma(G + dz)$? Playing the game on $G + dz$, the Avoider can force the Dominator to be the first to orient an edge in N. Clearly the only way D could end up with a dominating set of size 2 is to use two independent vertices of N: $\{a, c\}, \{a, d\}, \{b, c\}$ or $\{b, d\}$. The strategy of A will be to try to dominate some of these pairs from vertices of M, making them impossible to become dominating sets. He cannot "kill" all of the four possible dominating pairs, but at least two disjoint pairs he can. Whenever D orients the first edge of N, A can orient the other so that none of the four possible pairs could dominate the graph. We leave it to the reader to verify that D cannot do any better by orienting an edge of N before A "kills" two of the possible dominating pairs. $\square$

The "jump" can be larger then one, as a little modification of the previous example shows us.

Example 4.2.10. *If G is the same as in the previous example, $(k-1)(G+dz)+G$ has game domination number $2k$ and adding only one edge to the graph, $k(G + dz)$ has game domination number $3k$.*

We believe, that this is the biggest possible jump: if $\gamma_g(G) \leq 2k$ then $\gamma_g(G + ab) \leq 3k$.

4.3 Trees

First we shall derive a sharp lower bound for the game domination number of trees, then we look for upper bounds in terms of different parameters.

Theorem 4.3.1. *For any tree T on n vertices*

$$\gamma_g(T) \geq \left\lceil \frac{n}{2} \right\rceil.$$

Proof. We apply induction on the number of vertices. We clearly need $\lceil \frac{n}{2} \rceil$ vertices to dominate if $n \leq 3$.

To proceed with the induction, if $n \geq 4$ we need to find either a vertex with at least two leaves attached to it or a leaf attached to a vertex of degree two. One of these always exists, we might for example consider a vertex v and the longest path starting from v. This path ends in a leaf and the previous vertex has either degree two, or another leaf adjacent with it. The two cases are essentially the same, and we shall discuss only one in detail.

Suppose that there are two leaves u and v attached to a vertex w. Player A could play the game in $T - \{u, v\}$ according to his strategy resulting in a domination number at least $\lceil \frac{n-2}{2} \rceil$. Whenever D orients one of the edges adjacent to u or v, A immediately orients the other edge from the leaf. Thus two of the three vertices u, v, w are needed in the dominating set. At least another $\lceil \frac{n-2}{2} \rceil - 1$ vertices are needed from the rest of the graph, giving no dominating set smaller than $\lceil \frac{n}{2} \rceil$. $\square$

Theorem 4.3.2. *Let T denote a tree on n vertices that is not a path, and let d denote the smallest degree in T that is at least three. Then*

$$\gamma_g(T) \leq \min\left\{ \left\lfloor \frac{n}{2} + \frac{n-2}{2(d-1)} \right\rfloor , \left\lfloor \frac{2}{3}n \right\rfloor \right\},$$

where the $\lfloor \frac{2}{3}n \rfloor$ bound takes over the other only if $d = 3$.

Proof. We need to provide a strategy for player D resulting in a digraph with small domination number. Suppose that the tree has k vertices of degree at least d. An easy counting argument shows that $k \leq \lfloor \frac{n-2}{d-1} \rfloor$. We orient the edges so that the digraph we obtain has a small dominating set containing these k vertices.

Note, that the remaining $n - k$ vertices of T can be partitioned into paths attached to the k vertices of large degree. We claim that these vertices can be dominated by $\lfloor \frac{n-k}{2} \rfloor$ vertices in addition to the k vertices of degree at least d. Indeed, a path of even length $2l$ contains l independent edges and a dominating set of size l regardless of the orientation. Thus all D needs to take care of are the odd paths. But a path of length $2l+1$ starting from a dominating vertex v can easily be dominated by another l vertices if D takes the first move (he dominates the first vertex from v), and by $l+1$ vertices otherwise. As D starts the game, he is able to make the first move in at least half of the odd paths dominating the vertices on odd paths by at most half of them.

Thus we have a dominating set of the resulting digraph of size

$$\gamma_g(T) \leq k + \left\lfloor \frac{n-k}{2} \right\rfloor = \left\lfloor \frac{n+k}{2} \right\rfloor \leq \left\lfloor \frac{n}{2} + \frac{n-2}{2(d-1)} \right\rfloor.$$

If $d = 3$, then we do not put every vertex of degree three or larger into the dominating set. Instead, we partition the vertices of T into stars of at least two vertices (the existence of such a partition is obvious by induction). Player D can easily dominate a star $K_{1,r}$ with $\lfloor \frac{2}{3}(r+1) \rfloor$ vertices even if A starts the orientation, unless $r = 3$. In $K_{1,3}$ two vertices will dominate if D starts the game and three if A does. Thus the strategy of D is to make use of a star-partition: he starts in a $K_{1,3}$ (if there is any) then plays in the same star as A, except if he chooses another $K_{1,3}$. Then D does the same, ensuring that at least half of the three-stars will be dominated by two vertices, which in average gives a dominating set of size at most $5/8 < 2/3$ of the vertices in these stars. This strategy provides a domination number at most $\lfloor \frac{2}{3} n \rfloor$ in the resulting digraph. $\square$

Note, that both bounds in the previous theorem are sharp, for example if T is constructed from a path of k vertices with $d-1$ or $d-2$ leaves attached to each vertex such that each of the k vertices has degree $d \geq 4$. Then $\gamma_g(T) = \lfloor \frac{n}{2} + \frac{n-2}{2(d-1)} \rfloor$ as A can always orient edges from leaves to the central vertices.

For $d = 3$ consider a tree of three levels. The first level has only one vertex of degree $k+2$, the next level has k vertices of degree 3 and two leaves, and the third level contains $2k$ leaves. It is easy to check, that this tree has game domination number $2n/3$, showing that the second part of the theorem is also sharp.

Let us spell out one of the inequalities in the previous proof, which gives a slightly stronger version of the theorem.

Corollary 4.3.3. *If T is a tree on n vertices with $k \geq 1$ vertices of degree at least 3, then*

$$\gamma_g(T) \leq \left\lfloor \frac{n+k}{2} \right\rfloor .$$

$\square$

We summarize our results for trees in the following inequalities. Note, that the general upper bound could have been improved a lot for special trees.

Corollary 4.3.4. *For any tree T we have*

$$\left\lceil \frac{1}{2}n \right\rceil \leq \gamma_g(T) \leq \left\lfloor \frac{2}{3}n \right\rfloor .$$

$\square$

Corollary 4.3.5. *For any connected G we have*

$$\gamma_g(G) \leq \left\lfloor \frac{2}{3}n \right\rfloor .$$

$\square$

4.4 Inequalities

First we shall give a lower bound for the game domination number of a graph in terms of its maximal degree. This corresponds to the easiest basic inequality on the domination number: $\gamma(G) \geq \frac{n}{\Delta+1}$.

During our game A orients half of the edges and he might succeed in decreasing the out-degree of each vertex to about $\Delta/2$. This prompts us to make the following conjecture.

Conjecture 4.4.1. *For any graph G with n vertices and maximal degree Δ, we have $\gamma_g(G) \geq \frac{2n}{(1+o(1))\Delta}$.*

We have not been able to prove this conjecture, but the following somewhat weaker result is an improvement on the trivial lower bound $\gamma_g(G) \geq \gamma(G) \geq \frac{n}{\Delta+1}$.

Theorem 4.4.2. *If G is a graph with n vertices and maximal degree Δ, then*

$$\gamma_g(G) \geq \left\lceil \frac{4n}{3\Delta+7} \right\rceil .$$

Proof. The goal of A is to ensure that the out-degree of any vertex is at most $\frac{3\Delta+3}{4}$. We can add edges to the graph, until it becomes a $2k$-regular multigraph G', with $2k = \Delta$ or $2k = \Delta + 1$ depending on the parity of Δ.

As shown by Tarsi [78], this graph G' has a k-system, i.e. n edge-disjoint k-stars, one centered at each vertex of the graph. Corresponding to these, there are n edge-disjoint stars in G, each with at most k edges, centered at different vertices, and at any vertex v, at most k of the incident edges do not belong to the star centered at v.

The strategy of A is to orient an edge of the same star, in which D made his previous move, into the central vertex. Hence each vertex will have out-degree at most $\lfloor \frac{3\Delta+3}{4} \rfloor$, and dominate at most $\lfloor \frac{3\Delta+7}{4} \rfloor$ of the n vertices. □

Note, that when we provided strategies for D to obtain upper bounds for the game domination number we usually did that by finding a small set S of vertices dominating every other vertex at least twice, and thus ensuring that at least one of those edges can be oriented by D out of S, making S a dominating set of the resulting digraph. The concept of multiple domination was introduced by Fink and Jacobson in [42]. They call a set S *k-dominating* if every vertex of $V - S$ is adjacent to at least k vertices in S. The *k-domination number*, $\gamma_k(G)$, is the minimal cardinality of a k-dominating set. Our argument above shows that $\gamma_g(G) \leq \gamma_2(G)$. This gives the following immediate bounds, by some of the results on 2-domination numbers in [42] and [76].

Theorem 4.4.3. *For any graph G:* $\gamma_g(G) \leq \gamma_2(G) \leq \beta_2(G) \leq 2\beta_0(G)$.

In the inequality above, $\beta_0(G)$ denotes the independence number of G, and $\beta_2(G)$ is the 2-independence number, i.e. the maximal cardinality of a set I of the vertices such that the graph spanned by I has maximal degree at most 1. The complete graph shows that the inequality is sharp: $\gamma_g(K_n) = 2\beta_0(K_n) = 2$.

Theorem 4.4.4. *If the minimal degree* $\delta(G) \geq 3$, *then* $\gamma_g(G) \leq \gamma_2(G) \leq n/2$.

For $G = tK_4$ we have $\gamma_g(G) = 2t$, $v(G) = 4t$, so both inequalities are sharp in Theorem 4.4. We have seen the game domination number of trees to fall between $n/2$ and $2n/3$. Clearly the proof implies that this upper bound holds for any connected graph, as player D can concentrate his attention on a spanning tree of the graph (if player A moves outside of the spanning tree, D continues to orient tree edges according to his strategy). The following theorem improves the upper bound for graphs with minimal degree at least two.

Theorem 4.4.5. *If a graph G has minimal degree at least 2, then $\gamma_g(G) \leq \lfloor \frac{n}{2} \rfloor$.*

Proof. Our goal is to find a large 1-factor in the graph and use those edges to dominate the pairs of vertices by one of them regardless of their orientation. If G has a complete matching, this gives us a dominating set.

Suppose first that $n = 2k + 1$ odd, and there is a matching of size k, containing the edges $(v_1, u_1), \ldots (v_k, u_k)$, leaving only one more vertex for v to dominate. From the minimal degree condition, v is connected to a vertex of the matching, say v_1 . If u_1 is also connected to v, then the proof is done, as the resulting triangle vv_1u_1 can easily be dominated by one vertex if D starts the game. Otherwise u_1 is connected to another vertex of the matching, say v_2. Following this algorithm we build an alternating path $vv_1u_1v_2u_2 \ldots u_i$ until u_i is connected to a previous vertex on this alternating path. If this vertex is v or u_j, then we end up with an odd cycle and a matching (we might need to change the matching edges along the alternating path up to the cycle) and finish with a dominating set of size at most $n/2$ as before in Example 4.2.7. Finally, if u_i is attached to a vertex v_j on the path, then we have an even cycle with an odd path attached to it, and some independent edges of the original matching, and we can easily get the desired dominating set by Example 4.2.8.

We shall call a component odd (even), if its order is odd (even). Now we suppose that a maximal matching of G covers all but $t \geq 2$ vertices. By the extended version of Tutte's theorem there is a set S of s vertices such that after deleting S from the graph we shall get $s + t$ odd components. Choose S to be maximal among all such sets. Note, that S might be empty, if G was not connected and had exactly t odd components. We shall use this S to dominate the graph.

First note, that there exists a complete matching in every even component of $G - S$ (as there is a matching covering all but t vertices of the graph). We claim that if U is an odd component of $G - S$ and $u \in U$, then $U - \{u\}$ also has a complete matching. Otherwise it has a cut set K with k vertices leaving at least $k + 2$ odd components in $U - \{u\}$, since U had an odd number of vertices. But then $S \cup \{u\} \cup K$ would be a cut-set of order $s + k + 1$ giving $s + t + k + 1$ odd components, contradicting the maximality of S.

The strategy of D is simple: he ensures that the $s + t$ odd components of $G - S$ will be dominated "efficiently", i.e. by less then half of their vertices plus the vertices of S. To show how this can be done we need to distinguish three types of odd components in $G - S$ depending on the numbers of edges connecting them to S.

First, there may be some isolated components, which must be odd components of G with all but one vertex covered by the matching. It is easy to see that these can be dominated by half of their vertices if D manages to start the game in at least half of them. He starts in one, and starts another isolated component every time A does so, achieving his goal.

Second, there are odd components attached to S with at least two edges, but these can be dominated from S by orienting one of those edges out of S, and using the complete matching on the remaining part of it.

Third, there are odd components attached to S by only one edge between S and a vertex v in the odd component. In this case either D is able to orient the bridge from S toward v or if A has done this, he starts the game in the component and succeeds exactly like above when we had an almost complete matching missing only vertex v. Note, that in that argument we did not use the fact that v has degree at least two (that might not hold here), this was only needed for the vertices of the matching.

Even if we have to choose every vertex of S into the dominating set, it cannot be larger than

$$s + \frac{n-s-s-t}{2} = \frac{n-t}{2} < \left\lfloor \frac{n}{2} \right\rfloor,$$

completing the proof. □

This theorem is sharp as the examples of cycles with n vertices show. The upper bound for the game domination number can be further strengthened if the minimal degree is larger using a probabilistic argument.

Theorem 4.4.6. *For every graph $G = (V, E)$ with n vertices and minimum degree $\delta \geq 2$ and for every real number p between 0 and 1, $\gamma_g(G) \leq np + 2n(1-p)^\delta + 1 + n\delta p(1-p)^\delta$. Therefore, $\gamma_g(G) \leq (1+o(1))\frac{n\ln(\delta+1)}{\delta+1}$, where the $o(1)$-term tends to zero as δ tends to infinity, and the above the estimate is tight, up to the $o(1)$ error term.*

Proof. By Theorem 4.3 it suffices to prove that there is a set S of at most $np + 2n(1-p)^{\delta+1} + n\delta p(1-p)^{\delta}$ vertices of G, such that each vertex not in S has at least two neighbors in S. To prove the existence of such an S let X be a random set of vertices of G obtained by choosing each vertex $v \in V$, randomly and independently, to be a member of X with probability p. Let us fix arbitrarily some set $N(v)$ of precisely δ neighbors of each vertex $v \in V$, let Y denote the set of all vertices v such that neither v nor any member of $N(v)$ lies in X, and Z will denote the set of vertices u such that precisely one member of $N(u)$ is in X. The expected cardinalities of X, Y and Z are, respectively, np, $n(1-p)^{\delta+1}$ and $n\delta p(1-p)^{\delta}$. Moreover, by adding to X two arbitrarily chosen neighbors of u for each $u \in Z$, we obtain a set S of cardinality $|X| + 2|Y| + |Z|$ such that each vertex not in S has at least 2 neighbors in S. By linearity of expectation the expected cardinality of S is $np + 2n(1-p)^{\delta+1} + n\delta p(1-p)^{\delta}$ and hence there is such a set of cardinality at most this quantity. For large δ we can choose $p = \frac{\ln\delta + \ln\ln\delta}{\delta}$ and check that for this choice of p the resulting set S is of cardinality at most $\frac{n\ln(\delta+1)}{\delta+1} + O(n\frac{\ln\ln\delta}{\delta})$, as needed.

The tightness of the estimate follows easily from the fact that the game domination number is always at least as large as the domination number of the graph and the well-known fact that there are undirected graphs with n vertices and domination number $(1+o(1))\frac{n\ln(\delta+1)}{\delta+1}$ where the $o(1)$-term tends to zero as δ tends to infinity (see for example the discussion following Theorem 2.2 on page 7 in Alon and Spencer [7]). $\square$

In 1956 Nordhaus and Gaddum [59] established sharp bounds on the sum and product of the chromatic numbers of a graph and its complement. Similar results have been found for several parameters, including the following due to Jaeger and Payan [53].

Theorem 4.4.7. *If G is a graph of order n, then $\gamma(G) + \gamma(\bar{G}) \leq n + 1$ and this bound is sharp.*

We establish a sharp Nordhaus-Gaddum-type inequality for the game domination number of a graph and its complement.

Theorem 4.4.8. *For a graph G with n vertices, $\gamma_g(G) + \gamma_g(\bar{G}) \leq n + 2$. Furthermore, the bound is sharp.*

Proof. If the minimum degrees of G and $\bar{G}$ are at least two then by Theorem 4.4.5, $\gamma_g(G)+\gamma_g(\bar{G}) \leq n$. Hence, we may assume that $\delta(\bar{G}) \leq 1$. Then $\delta(G) \geq 2$, otherwise we have a vertex of degree at least $n-2$ in the complement as well, and the dominator D can use these two vertices to dominate almost half of both the graph and its complement, which results in dominating sets whose sum of sizes is at most $n + 2$ even if we have chosen every remaining vertex into the corresponding dominating set.

Suppose now that $\delta(\bar{G}) \leq 1$ and $\delta(G) \geq 2$. Thus there is a vertex v in G with degree at least $n-2$. If another vertex u of G has degree at least $n-2$, then $\gamma_g(G) \leq 3$ as using u, v and possibly one more vertex as dominating set, D can dominate G. On the other hand, if there exists an edge in $\bar{G}$, then $\gamma_g(\bar{G}) \leq n-1$, otherwise we need n vertices to dominate $\bar{G}$, but two vertices suffice to dominate G, providing the desired bound in either case.

We remain with the case when $d(v) \geq n-2$ and $d(u) \leq n-3$ for every $u \neq v$. Then by Theorem 4.4.5 player D can dominate in $\bar{G}$ with $\gamma_g(\bar{G}) \leq \lceil \frac{n-1}{2} \rceil + 1$ vertices by adding v to the dominating set. Also, using the star of G centered at v, player D can easily dominate G by $\lfloor (n-2)/2 \rfloor + 2$ vertices. Hence $\gamma_g(G) + \gamma_g(\bar{G}) \leq n+2$.

The complete graph K_n shows that this bound is sharp: $\gamma_g(K_n) = 2$ and $\gamma_g(\bar{K}_n) = n$. □

We believe that the inequality in Theorem 4.8 can be strengthened for connected graphs.

Conjecture 4.4.9. *If both G and $\bar{G}$ are connected graphs with n vertices, then $\gamma_g(G) + \gamma_g(\bar{G}) \leq \frac{2}{3}n + 3$.*

If true, this inequality is sharp, as shown by a tree of order n with game domination number $\frac{2}{3}n$ (see Section 4.3).

Chapter 5

A jump to the Bell number for hereditary graph properties

5.1 Introduction

An extremal jump is a discrete step between measures guaranteed in certain situations. It has been known for some time that the density of a graph jumps; recent work on hereditary graph properties has shown that properties with "large" or "small" speeds jump, but it was unknown whether there is a clean jump for properties with speed in a middle range. In this chapter, generalizations of the theorems of Dilworth, Ramsey, and Turán's are applied to answer this in the affirmative. In particular, we find a strict lower bound for the penultimate range of the speed hierarchy for hereditary properties of graphs.

Extremal graph theory concerns itself with the intrinsic structure of graphs, in this sense it is the central field of study in graph theory. Most of the results in the field concern themselves with forcing behavior. For instance, perhaps the first question of extremal graph theory was: what is the minimum number of edges that forces a graph to contain a triangle. This was answered by Mantel in 1901 and considerably extended to general complete graphs by Turán's Theorem of 1941. Turán showed that a graph that contains no K_{p+1} as a subgraph has at most $t(n,p) = (1 - 1/p + o(1))\binom{n}{2}$ edges. In the latter result, Turán showed not only the number of edges necessary, but provided the unique graph having that number of edges, the Turán graph $T(n,p)$. The two aspects, measure and structure, tend to go hand in hand in extremal graph theory.

One of the most remarkable facts of extremal graph theory, hinted at by Turán's Theorem but explicitly shown by Erdős and Stone in 1946 [41], is that the measures studied tend to "jump" in discrete steps. They showed that graphs with $t(n,p) + \epsilon n^2$ edges contain not only K_{p+1} but $K_{p+1}(t)$, a complete $(p+1)$-partite graph with classes of order t, where $t \to \infty$ as $n \to \infty$. That is, a graph containing a few (in particular, ϵn^2) more edges than guarantees a K_{p+1} jumps to contains $K_{p+1}(t)$ as well. This is a structural result. The work of Erdős, Stone, and Simonovits [41, 39] gives a similar result for measures: For every $0 < \alpha, \epsilon \leq 1$ and $m \geq 2$, a graph having $n > n(m,\epsilon)$ vertices and at least $(\alpha+\epsilon)\binom{n}{2}$ edges contains a subgraph on m vertices with $(1 - 1/\ell)\binom{m}{2}$ edges, where $\ell = \ell(\alpha) \in \mathbb{N}$ and $1 - 1/\ell > \alpha + \epsilon$.

Studying graph properties, Bollobás and Thomason [22] found similar jumps in a different measure. A *graph property* is an infinite collection of labeled graphs closed under isomorphism. A property is *hereditary* if it is further closed under taking induced subgraphs. The *speed* of a graph property $\mathcal{P}$, denoted $|\mathcal{P}^n|$, is a function of n giving the number of graphs in the property on n vertices. In 1997, Bollobás and Thomason showed that the speed of a graph property also jumps, and in precisely the same places as for the number of edges. That is, the only speeds that occur for hereditary graph properties are of the form $|\mathcal{P}^n| = 2^{(1-1/\ell+o(1))\binom{n}{2}}$. They also presented necessary and sufficient structural characteristics for the properties that evidence each type of speed.

Scheinerman and Zito [70] were the first to notice that these jumps also occur with speeds at lower levels. They saw that there some functional types that are not allowed as the speed of any hereditary property, and that there are discrete jumps, for example, from polynomial to exponential speeds. The present authors extended their work in [11], where many jumps were noted, precise functions that occur were enumerated, and even structural characteristics of properties with each type of speed were described. For example, we found that there is a jump from the speed $n^{(1-1/k+o(1))n}$ to $n^{(1-1/(k+1)+o(1))n}$ for all integers $k > 1$. The results discussed so far may be summarized in the following theorem.

Theorem 5.1. *Let $\mathcal{P}$ be a hereditary property of graphs. Then one of the following is true:*

1. *there exists $N, k \in \mathbb{N}$ and a collection $\{p_i(n)\}_{i=0}^{k}$ of polynomials such that for all $n > N$, $|\mathcal{P}^n| = \sum_{i=0}^{k} p_i(n) i^n$,*

2. *there exists* $k \in \mathbb{N}, k > 1$ *such that* $|\mathcal{P}^n| = n^{(1-1/k+o(1))n}$,

3. $n^{(1+o(1))n} \leq |\mathcal{P}^n| \leq n^{o(n^2)}$,

4. *there exists* $k \in \mathbb{N}, k > 1$ *such that* $|\mathcal{P}^n| = 2^{(1-1/k+o(1))n^2/2}$.

□

The existence of jumps within and between the first two cases, proven by the authors in [11], are clear from the statement of the theorem. A jump to the third was shown in [13]. The last case and the jump from case 3 to 4 are shown by Bollobás and Thomason in [22]. Case 3 remains puzzling, however, for a number of reasons.

In this "penultimate range," where properties have speeds $n^{(1+o(1))n} \leq |\mathcal{P}^n| \leq 2^{n^{2-\epsilon}}$, the jumps are not clear and properties do not necessarily behave as nicely as in every other part of the hierarchy. In [12], the authors describe the unusual behavior of properties in the penultimate range, and show that there exist properties which have speeds that oscillate between extremes. The question is raised as to whether any bounds can be given for properties in this range; that is whether the jump to or from this range is a clean one. An upper bound for speeds of monotone, rather than hereditary, properties in the range is given in [13] (monotone properties are closed under taking arbitrary subgraphs, rather than induced subgraphs); for hereditary properties the upper bound remains unknown.

In this paper, we shall show that the jump from speeds of the type $n^{(1-1/k+o(1))n}$ to the penultimate range is in fact clean, and provide a sharp lower bound for hereditary properties in this range. This bound will also hold for monotone properties, and so for that class the bounds are completely settled.

We shall approach this question by taking a tour through classical graph theory. We shall start with three important results of combinatorics: Dilworth's Theorem on posets, Ramsey's Theorem on substructures of graphs, and the results of Turán, Erdős, Simonovits, and Stone in extremal graph theory mentioned above. These will be generalized and then applied to hereditary properties of graphs. Definitions and notation will be introduced as needed.

5.2 Dilworth's Theorem and hypergraphs

To prove the main results of this paper, we shall need some Ramsey-type results on hypergraphs, extending the classical theorem of Dilworth. We shall also need more Ramsey-type results on graphs, which will be discussed in the next section.

Our definitions are standard, but for completeness shall be given below. A *hypergraph* $\mathcal{H}$ is a pair $\mathcal{H} = (V, \mathcal{E})$, with *vertex set* V and with *edge set* $\mathcal{E}$ consisting of subsets of V. For $x \in V$, the *degree* of x is $d(x) = |\{F : x \in F \in \mathcal{E}\}|$. Clearly, a hypergraph defines a poset on the set of edges, with the order given by inclusion. Viewed this way, we may define the *complement* $\overline{\mathcal{H}}$ of a hypergraph $\mathcal{H} = (V, \mathcal{E})$ by taking the complement of each of the edges over the base set, i.e. $\overline{\mathcal{H}} = (V, \overline{\mathcal{E}})$, where $\overline{\mathcal{E}} = \{V \setminus E : E \in \mathcal{E}\}$. In a hypergraph, we shall allow the empty edge but multiple edges shall not occur.

With this perspective, a *chain* in a hypergraph $\mathcal{H} = (V, \mathcal{E})$ is a collection $\mathcal{F} \subset \mathcal{E}$ such that, for all pairs $A, B \in \mathcal{F}$, either $A \subset B$ or $B \subset A$. An *antichain* in $\mathcal{H}$ is a collection $\mathcal{F} \subset \mathcal{E}$ such that, for all $A, B \in \mathcal{F}$, $A \neq B$, we have $A \not\subset B$. Note that the complement of a chain or an antichain is again a chain or an antichain, respectively.

A basic tool in the theory of posets, Dilworth's Theorem, can be restated for hypergraphs as follows.

Theorem 5.2. *A hypergraph containing at least $km + 1$ edges contains a chain containing $k + 1$ or an antichain containing $m + 1$ elements.* □

While chains have only one form allowed by their definition, antichains are a rich class of sets with very little prescribed form. We wish to extend Dilworth's Theorem to describe the structure of some large antichains that must exist in any hypergraph with no large chain.

Let $H = (V, \mathcal{E})$ be a hypergraph and $\mathcal{F} = \{A_1, \dots, A_k\} \subset \mathcal{E}$.

- $\mathcal{F}$ is a *k-star* in $\mathcal{H}$ if there exists $F = \{x_1, \dots, x_k\} \subset V$ such that $x_i \in A_j$ if and only if $i = j$.

- $\mathcal{F}$ is a *k-costar* in $\mathcal{H}$ if there exists $F = \{x_1, \dots, x_k\} \subset V$ such that $x_i \notin A_j$ if and only if $i = j$.

- $\mathcal{F}$ is a *k-skewchain* in $\mathcal{H}$ if there exists $F = \{x_1, \dots, x_{k-1}\} \subset V$ such that $x_i \in A_j$ if and only if $i < j$.

In each case, we call F a *representing set* for $\mathcal{F}$. We may also refer to the pair $(\mathcal{F}, F)$ as the star, costar, or skewchain, where usage should be clear from context. Note that stars and costars are antichains, whereas a skewchain can be either a chain or antichain. In fact, $(\mathcal{F}, F)$ is a skewchain if and only if the trace of $\mathcal{F}$ on F is a chain of the same length. Also note that the complement of a star is a costar, while the complement of a skewchain is also a skewchain.

With these definitions, and motivated by Dilworth's Theorem, we define the number $f(k, l, m)$ to be the smallest number such that every hypergraph with at least $f(k, l, m)$ edges contains a k-star, l-costar, or m-skewchain. A priori, it is not clear that $f(k, l, m)$ is well defined, but the following theorem tells us that this is in fact the case and gives a bound on its growth.

Note that $f(k, l, m)$ is symmetric in k and l. To see this, note that if $\mathcal{H}$ does not contain a k-star, l-costar, or m-skewchain, then $\overline{\mathcal{H}}$ cannot contain an l-star, k-costar, or m-skewchain. Hence $f(k, l, m) - 1 \geq f(l, k, m) - 1$, and the inequality is also true with k and l reversed.

Theorem 5.3. *The function $f(k, l, m)$ is well defined for all $k, l, m \in \mathbb{N}$. In fact, for $k, l, m > 2$,*

$$f(k, l, m) \leq 2(m-1)(m-2)f(k-1, l, m)f(k, l-1, m) + 1 \tag{5.1}$$

Proof. Note that $f(k, l, m) = 1$ if and only if $\min\{k, l, m\} = 1$. When $\min\{k, l\} = 2$, the function $f(2, l, m) = f(k, 2, m) = m$, as any non-nested pair of edges is both a 2-star and a 2-costar, and $f(k, l, 2) = 2$, since repeated edges are not allowed. Hence to prove the theorem it suffices to show (5.1) holds when $\min\{k, l, m\} > 2$.

Fix $k, l, m \geq 3$. Let $\mathcal{G} = (V, \mathcal{F})$ be a hypergraph with a number of edges at least as large as the right hand side of (5.1). If it contains a chain of length m we are done, since this would also be an m-skewchain. Otherwise, by Theorem 5.2, $\mathcal{G}$ contains an antichain hypergraph $\mathcal{H} = (V, \mathcal{E})$ with

$$|\mathcal{E}| > 2(m-2)f(k-1, l, m)f(k, l-1, m)$$

Pick any $x \in V$ that is not in every edge, but is in at least one edge. If $|\mathcal{E}| > d(x) \geq |\mathcal{E}|/2$, we shall show that $\mathcal{H}$ contains a k-star, l-costar, or m-skewchain (and, therefore, so does $\mathcal{G}$). If $0 < d(x) < |\mathcal{E}|/2$, then consider the complement of $\mathcal{H}$ and, in this, $|\overline{\mathcal{E}}| > d(x) > |\overline{\mathcal{E}}|/2$. If we can find an l-star, k-costar, or m-skewchain in the complement of $\mathcal{H}$, then $\mathcal{H}$ contains a k-star, l-costar, or m-skewchain. Since $f(k, l, m)$ is symmetric in k and l, it does not matter whether we are looking for an l-star or a k-costar or for a k-star or l-costar. Therefore, without loss of generality, we may say $|\mathcal{E}| > d(x) \geq |\mathcal{E}|/2$, and showing that $\mathcal{H}$ contains a k-star, l-costar, or m-skewchain will prove the result.

We shall partition the edge set to close in on one of the desired structures, establishing a collection of sets according to our choice of x. Let $\mathcal{E}_x = \{E \in \mathcal{E} : x \in E\}$. Since $d(x) \geq \mathcal{E}/2$,

$$|\mathcal{E}_x| > (m-2)f(k-1, l, m)f(k, l-1, m).$$

Pick some $A \in \mathcal{E}$ such that $x \notin A$ and let

$$\mathcal{F}_{x,A} = \{E \cap A : E \in \mathcal{E}_x\}.$$

For each $B \in \mathcal{F}_{x,A}$, define

$$\mathcal{E}_B = \{E \in \mathcal{E}_x : E \cap A = B\}.$$

Note that for all $B \in \mathcal{F}_{x,A}$, $A \setminus B \neq \emptyset$, since $\mathcal{E}$ is an antichain.

Case 1. $|\mathcal{E}_B| \geq f(k-1, l, m)$ for some $B \in \mathcal{F}_{x,A}$.

Then $\mathcal{E}_B$ contains either a $(k-1)$-matching, an l-costar, or an m-skew-star. In the latter two cases we are done, as $\mathcal{E}_B \subset \mathcal{E}$. Otherwise $\mathcal{E}_B$ contains a $(k-1)$-star, say $\S$, with $\S = \{S_1, \ldots, S_{k-1}\}$ and representing set $S = \{x_1, \ldots, x_{k-1}\}$. As $k \geq 3$, and any element in B is in all elements of $\mathcal{E}_B$, we have $x_i \notin B$, so $x_i \notin A$. However, as noted above, there is some $y \in A \setminus B$. This means that adding A to $\S$ and y to S in the $(k-1)$-star yields a k-star, as desired.

Case 2. $|\mathcal{E}_B| < f(k-1, l, m)$ for all $B \in \mathcal{F}_{x,A}$.

Then, since the set $\{\mathcal{E}_B : B \in \mathcal{F}_{x,A}\}$ is a partition of $\mathcal{E}_x$,

$$|\mathcal{F}_{x,A}| > (m-2) f(k, l-1, m).$$

If $\mathcal{F}_{x,A}$ contains a chain $\S$ with order (at least) $m-1$, then, similarly to the situation in Case 1, $\mathcal{E}$ contains an m-skewchain consisting of a collection of edges the set of whose intersections with A is $\S$ together with A itself.

Otherwise, by Theorem 5.2, $\mathcal{F}_{x,A}$ contains an antichain $\mathcal{C}$ with $|\mathcal{C}| > f(k, l-1, m)$.

By the induction statement, $\mathcal{C}$ contains either an $(l-1)$-costar, a k-star, or an m-skewchain. In the latter two cases we simply take, for each set B in the k-star (respectively, m-skewchain), an edge F from $\mathcal{H}$ with $F \cap A = B$, and the collection we get is a k-star (respectively, m-skewchain) in $\mathcal{H}$, with the same representing set as for $\mathcal{F}_{x,A}$. If $\mathcal{C}$ contains an $(l-1)$-costar $\S$ with edges $\{C_1, \ldots, C_{l-1}\}$ and representing set $\{x_1, \ldots, x_{l-1}\}$, then for each set

C_i, there is a $C'_i \in \mathcal{E}_x$ with $C_i = C'_i \cap A$ for all $1 \leq i \leq l-1$. As $C'_i \in \mathcal{E}_x$ for all i, $\{C'_1, \ldots, C'_{l-1}, A\}$ is an l-costar in $\mathcal{H}$ with representing set $\{x_1, \ldots, x_{l-1}, x\}$.

Thus, in all cases, a hypergraph with $2(m-1)(m-2)f(k-1,l,m)f(k,l-1,m)+1$ edges contains a k-star, an l-costar, or an m-skewchain.

We shall apply this theorem in section 5.6 to find certain structures in graphs.

5.3 Ramsey Theory

Theorem 5.3 guarantees "large" regular substructures in any hypergraph that is large enough. In this sense, it falls into the vast field of Ramsey Theory. Ramsey-type theorems have to do with finding certain substructures in large graphs. Before we discuss Ramsey results, let us review some definitions.

Given a graph G, the graph H is isomorphic to an induced subgraph of G if the vertices of H can be mapped to a subset of those of G so that edges are mapped to edges and non-edges to non-edges. We write $H \leq G$, and say H is an *induced subgraph* of G. A vertex set $U \subset V(G)$ induces H if $H \leq G$ and the vertices of H can be mapped to U so that edges and non-edges are preserved. We write $H = G[U]$. Slightly differently, given two disjoint sets $U, V \subset V(G)$, the *induced bipartite graph* $G[U,V]$ has vertex set $U \cup V$ and edge set consisting of those edges of G with one end in U and the other end in V. The *bipartite complement* of $G[U,V]$ has the same vertex set as $G[U,V]$ but has edge set $\{uv : u \in U, v \in V, uv \notin E(G)\}$.

The first Ramsey result was the following.

Theorem 5.4. *There is a number $R(n)$ such that any graph on $R(n)$ vertices contains either K_n or $\overline{K_n}$ as an induced subgraph.* $\square$

Hence any "large" graph contains an arbitrarily large clique or independent set. Much work has been done to determine precise bounds on the function $R(n)$, but we shall not investigate much in this area. Instead, we are interested in finding subgraphs other than the complete graph. A Ramsey-type result along these lines for bipartite graphs was obtained by Kővári, Sós, and Turán [57] in response to a question of Zarankiewicz about matrices.

Theorem 5.5. *Let t be fixed. There is a function $H_t(n) = O(n^{2-1/t})$ such that any bipartite graph on n vertices with at least $H_t(n)$ edges contains $K_{t,t}$ as a subgraph.*

$\square$

The function $H_t(n)$ has been studied in great detail; here we shall only need the following bound, also established in [57].

Lemma 5.6. *For $1 \le t < n$,*

$$H_t(n) \le \frac{1}{2}(t-1)^{1/t} n^{2-1/t} + \frac{1}{2}(t-1)n < 2n^{2-1/t}$$

$\square$

In turn simple calculations give the following corollary.

Corollary 5.7. *There is a number $H(t)$ such that any bipartite graph on $H(t)$ vertices contains either a $K_{t,t}$ or an independent set containing t vertices from each partition.*

Proof. $H(t) = H_t(n)$ with $n^2 > 2n^{2-1/t}$ will do. $\square$

Combining Ramsey's Theorem and the result above, we get the following.

Corollary 5.8. *There is a number $n(t, r)$ such that if G is an $R(r)$-partite graph with $n(t, r)$ vertices then G either contains the Turán graph $T(tr, r)$ or an independent set of tr vertices that intersects r of the sets of the partition in t vertices each.*

Proof. Let G be an $R(r)$-partite graph on $H(H(\ldots(t)\ldots))$ vertices, where the dots signify composition $\binom{R(r)}{2}$ times. Applying Corollary 5.7 to each pair of partite sets, we obtain an $R(r)$-partite graph with t vertices in each partite set and such that each pair of partite sets is either completely connected or disconnected. The graph H obtained by contracting each partite set to a point is a graph with $R(r)$ vertices and thus contains either a K_r or $\overline{K_r}$, corresponding to $T(tr, r)$ or an independent set that spans r sets of the partition and contains t vertices from each set it intersects, respectively. $\square$

As mentioned earlier, Ramsey Theory deals with any sort of large substructure that can be guaranteed as a subgraph of an arbitrary graph. We now consider instead a very specific case of both parent and child graphs. Recall that P_n is a path on n vertices. In the next lemma, we show that paths contain highly structured induced path forests. The result could be viewed as a statement about colorings of the path, saying that multicolored paths contain induced path forests in which each color appears many times. Or, as stated below, it can be viewed as a statement about words and sentences.

Recall that a *word* is a sequence of letters, where each letter is chosen from a given *alphabet*. A *sentence* can be formed from the word by removing letters and leaving a space wherever consecutive letters have been removed. The *words of the sentence* are then *blocks* of consecutive letters that remain between spaces. This can also be phrased in terms of sequences, colors, subsequences, and blocks, respectively.

For the next result, we fix numbers ℓ, n, and p, and we define $m(\ell, n, p)$ to be the minimal number, if it exists, such that, for any sequence of positive integers $\{a_1, a_2, \ldots, a_p\}$ with $\sum_{i=1}^{p} a_i = n$, every word of length $m(\ell, n, p)$ from an alphabet of size ℓ (containing ℓ letters) contains a sentence with p words, the i^{th} word having length a_i, such that each letter of the alphabet that appears in the sentence appears at least $\lfloor p/\ell \rfloor$ times.

While such a condition sounds quite restrictive, the following lemma says that this number does in fact exist, and gives an inductive bound on its size.

Lemma 5.9. *The function $m(\ell, n, p)$ is well-defined for all $l, n, p \in \mathbb{N}$. Furthermore, for $l > 1$,*

$$m(\ell, n, p) \leq (p-1)m(\ell-1, n, p) + n + p - 1. \tag{5.2}$$

Proof. As is implied by the statement, we proceed by induction on ℓ. Clearly $m(1,n,p) = n + p - 1$. So suppose $\ell \geq 2$ and $m(\ell-1,n,p)$ exists. Let σ be a word at least as long as given by the right hand side of (5.2) that uses ℓ different letters. Let $\{a_1, a_2, \ldots, a_p\}$ be a set of positive integers with $\sum_{i=1}^{p} a_i = n$. If σ contains a word of length $m(\ell-1,n,p)$ with only $\ell - 1$ different letters, then, by induction, this word contains a sentence with the desired characteristics.

So assume every subword of σ of length $m(\ell-1,n,p)$ contains all ℓ colors. We may construct a sentence with the desired properties with a greedy algorithm. We consider our alphabet to be $[\ell]$ and, without loss of generality, assume the first letter of σ is 1. Take a word σ_1 of length a_1, and throw out the next entry. If 1 does not appear $\lfloor p/\ell \rfloor$ times in σ_1, skip forward to the next 1 entry. As noted above, we need to go forward at most $m(\ell-1,n,p)$ entries. Starting with that 1, take a word σ_2 of length a_2. If the letter 1 has not yet appeared in the sentence $\lfloor p/\ell \rfloor$ times, repeat the process, and repeat for each letter that has not appeared $\lfloor p/\ell \rfloor$ times in the sentence we have picked thus far. We can ensure that our sentence has each letter appearing at least $\lfloor p/\ell \rfloor$ times, as σ has length

$$(p-1)m(\ell-1,n,p) + n - (p-1) =$$

$$(a_1+1)+m(\ell-1,n,p)+(a_2+1)+m(\ell-1,n,p)+\cdots+(a_{p-1}+1)+m(\ell-1,n,p)+(a_p).$$

$\square$

5.4 (ℓ, d)-graphs

We briefly leave the classical realm to discuss some technical definitions, but shall return to extremal graph theory by the end of this section. As in the previous section, we consider here graphs rather than hypergraphs, and look at substructures in the graph.

Some notation will be useful here. Given two sets of vertices of G, $U, V \subset V(G)$, the maximum degree between them, $\Delta(U, V) = \max\{|\Gamma(u) \cap V|, |\Gamma(v) \cap U| : v \in V, u \in U\}$, where $\Gamma(u)$ is the neighborhood of u. With $\overline{\Gamma}(u) = V(G) \setminus (\Gamma(u) \cup \{u\})$, $\overline{\Delta}(U, V) = \max\{|\overline{\Gamma}(u) \cap V|, |\overline{\Gamma}(v) \cap U| : v \in V, u \in U\}$. Note $\Delta(U, U)$ is simply the maximum degree in $G[U]$, also denoted $\Delta(G[U])$.

A graph H is an *(ℓ, d)-graph* if $V(H)$ admits a partition $V_1, V_2, \ldots, V_\ell$ such that, for each pair i, j (not necessarily distinct) either $\Delta(V_i, V_j) \leq d$ or $\overline{\Delta}(V_i, V_j) \leq d$. We call $V_1, V_2, \ldots, V_\ell$ an *(ℓ, d)-partition.*

Given an (ℓ, d)-partition $V_1, V_2, \ldots, V_\ell$ of H, for each $x \in V(H)$ and $i \in [\ell]$, either $|\Gamma(x) \cap V_i| \leq d$ or $|\overline{\Gamma}(x) \cap V_i| \leq d$. In the former case we say that x is *sparse* with respect to V_i, in the opposite case x is *dense* with respect to V_i. Likewise we may refer to a class V_i as dense or sparse with respect to another class.

Given any (ℓ, d)-partition with some empty classes, we may obtain another (ℓ, d)-partition with fewer empty classes by subdividing any of the classes. Similarly, in some cases we may be able to unify a collection of classes to obtain a new (ℓ, d)-partition. In this sense, (ℓ, d)-partitions are not unique. However, if we choose ℓ to be minimal so that a graph H admits an (ℓ, d)-partition, perhaps (in a sense described below) even allowing d to grow slightly to reflect the graph's structure more accurately, then the partition will in fact be unique if H is large enough. In this case, the (ℓ, d)-partition will actually reflect the dense or sparse portions of the graph.

Precisely, we have the following: Let H be an (ℓ, d)-graph which admits a partition π such that each class has at least $2\ell d + 2d + 1$ vertices. Let ℓ' be minimal so that H admits an $(\ell', \ell d)$-partition π'. In this partition, we have the following condition:

$(*)$ if $i, j \in [\ell']$ and $i \neq j$, then there exists a $k \in [\ell']$ such that $\Delta(V_i, V_k) \leq \ell d$ and $\overline{\Delta}(V_j, V_k) \leq \ell d$ or vice-versa.

This is clearly the case, as if π does not satisfy $(*)$, then there are some classes, without loss of generality, $V_1, V_2 \in \pi$, such that for all $k \in [\ell]$, V_k is dense with respect to both V_1 and V_2 or is sparse with respect to both. But then $V_1 \cup V_2, V_3, \ldots, V_\ell$ is an $(\ell - 1, 2d)$-partition of H. In this way, given an (ℓ, d)-graph H we may join classes that "act the same" in the original partition to obtain an $(\ell', \ell d)$-partition of H satisfying $(*)$ and also having each class with at least $2\ell d + 2d + 1$ elements.

Condition $(*)$ and this minimum number of vertices in a class in fact creates a much stronger condition. The following facts about (ℓ, d)-partitions will imply that a partition satisfying $(*)$ is in fact unique.

Let $V_1, \ldots, V_\ell$ be an (ℓ, d)-partition of H and $x, y \in V(H)$.

1. If $x, y \in V_1$, then $|\Gamma(x)\Delta\Gamma(y)| \leq 2dl$, as, for each V_i, $|(\Gamma(x) \cap V_i) \cup (\Gamma(y) \cap V_i)| \leq 2d$ or $|(\overline{\Gamma}(x) \cap V_i) \cup (\overline{\Gamma}(y) \cap V_i)| \leq 2d$.

2. On the other hand, suppose $x \in V_1$ and $y \in V_2$. Suppose further that there exists a non-empty set V_i such that x is dense with respect to V_i and y is sparse with respect to V_i, or vice-versa. Then, since $|\overline{\Gamma}(x) \cap V_i| \leq d$ while $|\Gamma(y) \cap V_i)| \leq d$ (or vice-versa), $|\Gamma(x)\Delta\Gamma(y)| \geq |V_i| - 2d$.

Hence, considering Facts 1 and 2, if condition $(*)$ holds and there exists a partition $V_1, \ldots, V_\ell$ such that $|V_i| - 2d > 2dl$ whenever $V_i \neq \emptyset$, then this partition is uniquely defined by the graph itself, as the conditions in the two facts would be mutually exclusive.

Motivated by the above, we define a *strong* (ℓ, d)-graph to be one which admits an (ℓ, d)-partition, each of whose classes contains at least $2d\ell+2d+1$ vertices. A *strong* (ℓ, d)-partition is defined similarly. By the discussion above, a strong graph admits a unique partition with a minimal number of sets. Any strong (ℓ, d)-partition with no empty sets may always be obtained from the unique partition by subdividing some of the classes. Note that a strong (ℓ, d)-graph contains at least $2d\ell^2 + 2d\ell + l$ vertices.

Note for each H that there is a minimal ℓ (for fixed $d \geq 1$) such that H is an (ℓ, d)-graph (every graph on n vertices is an (n, d)-graph). Similarly, for fixed ℓ there is a minimal d so that H is an (ℓ, d)-graph.

Given a strong (ℓ, d)-graph H, consider its minimal partition. We obtain a graph $\varphi(H)$ from the partition by replacing $H[V_i, V_j]$ with its bipartite complement for every pair with $\overline{\Delta}(V_i, V_j) \leq \ell d$. Note that the maximal degree of $\varphi(H)$ is thus at most ℓd.

We are now able to show that strong (ℓ, d)-graphs not only produce a unique minimal partition, but that the unique partition is preserved by any subgraph that would still be strong.

Lemma 5.10. *Let H be a strong (ℓ, d)-graph with unique partition $V_1, \ldots, V_{\ell'}$. Let $F \leq \varphi(H)$ with $F \cap V_i \geq 2d\ell + 2d + 1$ for all i such that $V_i \neq \emptyset$. Let $G = H[V(F)]$. Then G is a strong (ℓ, d)-graph with unique partition induced by $V_1, \ldots, V_{\ell'}$. In particular, $\varphi(G) = F$.*

Proof. We shall construct an (ℓ, d)-partition of G as follows. Pick two vertices of G, say x and y. Although we do not know what the partition of F or H was, we can say that they either came from the same class of the partition of H or they did not. If they came from the same class, without loss of generality say V_1, then their neighborhoods can differ by at most $2d\ell$ vertices. This maximum is achieved if for each class V_j (including $j = 1$) on which V_1 is dense (respectively, sparse) they have disjoint sets of d neighbors which they are not adjacent to (respectively, are adjacent to) (i.e. if $\Delta(V_1, V_j) \leq d$ then x and y are *adjacent* to different sets from V_j, if $\overline{\Delta}(V_1, V_j) \leq d$ then there are disjoint sets from V_j that x and y are not adjacent to). If they come from different classes say V_1 and V_2, then there is a set, say V_3 which differentiates V_1 from V_2 (otherwise the partition would not be minimal). Without loss of generality say $\Delta(V_1, V_3) \leq d$ and $\overline{\Delta}(V_2, V_3) \leq d$. Then x and y may act the same on at most $2d$ vertices in

V_3. Since $|V_3| \geq 2d\ell+2d+1$, they act differently on at least $2d\ell+1$ vertices in G. Hence, by comparing the neighborhoods of any pair of vertices in G we may reconstruct the partition, uniquely. It will agree with the unique partition of H and each part will have at least $2d\ell+2d+1$ vertices, so G is a strong (ℓ, d)-graph. $\square$

Lemma 5.11. *If H is a strong (ℓ, d)-graph, then $Aut(H) \subseteq Aut(\varphi(H))$, and thus the number of distinct labelings of H is at least as much as the number of distinct labelings of $\varphi(H)$.*

Proof. We shall show that any automorphism of H is also an automorphism of $\varphi(H)$. Let Ψ be an automorphism of H and π be the unique partition of H. Let $x, y \in V(G)$. The vertices x and y are in the same class of π if and only if $\Psi(x)$ and $\Psi(y)$ are in the same class of π, as otherwise the orders of the neighborhoods of $\Psi(x)$ and x or $\Psi(y)$ and y would be different. More precisely, Ψ preserves classes. Let $x, \Psi(x) \in V_i$ and $y, \Psi(y) \in V_j$. Since Ψ is an automorphism, $xy \in E(H)$ if and only if $\Psi(x)\Psi(y) \in E(H)$. Without loss of generality, assume $xy \in E(H)$. If V_i is sparse with respect to V_j, we have $xy \in E(\varphi(H))$ and $\Psi(x)\Psi(y) \in E(H)$. If V_i is dense with respect to V_j, then both xy and $\Psi(x)\Psi(y) \notin E(H)$. Since our choice of x and y was arbitrary, in either case Ψ is an automorphism of $\varphi(H)$. $\square$

Note that the converse of the statement in the proof above is generally not true. Two vertices that have different neighborhoods in H may have identical neighborhoods in $\varphi(H)$.

Also note that if H is not strong, the lemma may not be true at all. In particular, if a graph is an (ℓ, d)-graph, it is also an $(\ell+1, d)$-graph, with $(\ell+1, d)$-partition obtained by breaking up any class of the (ℓ, d)-partition. As a trivial example, consider K_3 as a $(2,1)$-graph. If we make a non-trivial $(2,1)$-partition into V_1 and V_2, with $|V_1| = 2$ and say $1 = \Delta(V_1, V_1) \leq 1$ and $0 = \overline{\Delta}(V_1, V_2) \leq 1$, then $\varphi(K_3) = K_2 \dot{\cup} K_1$, which has 3 different labelings while K_3 has only one. In fact, for this reason, $\varphi(H)$ is only defined for strong (ℓ, d)-graphs. The lemma applies when we view K_3 as a strong $(1,1)$-graph. Then with the unique $(1,1)$-partition $|\mathrm{Aut}(\varphi(K_3))| = |\mathrm{Aut}(K_3)|$.

5.5 Hereditary properties of graphs

The terminology of (ℓ, d)-graphs may seem a bit awkward, but in fact (ℓ, d)-graphs are key in understanding the structure of complicated properties of graphs.

In [11], the present authors showed that partitions like those in an (ℓ, d)-graph provide an easy way to bound the number of graphs in a property. In fact, in a certain range, where $|\mathcal{P}^n|$ is roughly factorial in n, they provide the best way to bound the speed. These results will be presented below, but as usual we need a few more definitions.

The first set of definitions is standard. Given a graph G and collection of vertices $v_1, \ldots, v_t \in V(G)$, we say that the disjoint sets $U_1, \ldots, U_m \subset V(G)$ are *distinguished by* $V = \{v_1, \ldots, v_t\}$ if, for each i, every vertex of

U_i has the same neighborhood in V and for each $i \neq j$, $x \in U_i$, $y \in U_j$ implies x and y have different neighborhoods in V. We say V *distinguishes* U_i. The set V is a *minimal* distinguishing set if no proper subset of V distinguishes the same sets.

The following definition is new, and may seem odd at first. For a property $\mathcal{P}$, we shall call k *good* if for every $m > 0$ there is a $G \in \mathcal{P}$ such that G contains a set of vertices which distinguish m sets, each of order at least k. Clearly 0 is good for all $\mathcal{P}$. Let $k_{\mathcal{P}}$ be the infimum of all bad k. This definition allows us to connect hereditary properties to (ℓ, d)-graphs in the following surprising result from [11, Lemma 27].

Lemma 5.12. *If $\mathcal{P}$ is a hereditary property with $k_{\mathcal{P}} < \infty$, then there exist absolute constants $\ell_{\mathcal{P}}$ and $c_{\mathcal{P}}$ such that for all $G \in \mathcal{P}$, the graph G contains an induced subgraph H such that H is an $(\ell_{\mathcal{P}}, k_{\mathcal{P}})$-graph and $|V(G \setminus H)| < c_{\mathcal{P}}$.* □

More importantly for computing speeds, we also showed the following [11, Theorem 28]. The hypothesis $k_{\mathcal{P}} < \infty$ was assumed in the surrounding text in that paper and therefore was omitted in the statement there.

Theorem 5.13. *Let $\mathcal{P}$ be a hereditary property with $k_{\mathcal{P}} < \infty$. Then $|\mathcal{P}^n| \geq n^{(1+o(1))n}$ if and only if for all m there exists a strong $(\ell_{\mathcal{P}}, k_{\mathcal{P}})$-graph H in $\mathcal{P}$ such that $\varphi(H)$ has a component of order at least m.* □

We shall put these ideas to even more use in the next section, where we deal with a particularly unsettling area of the established research on properties.

First, an easy pair of technical lemmas. Recall that P_n is a path on n vertices.

Lemma 5.14. *Let D be a constant. If G is connected and $\Delta(G) \leq D$, then, for any $n \leq \log_D |V(G)|$, we have $P_n \leq G$.*

Proof. Pick any $v \in V(G)$. The number of vertices at distance d from v is at most D^d, by the degree condition. Hence, for any $n \geq 1$, the number of vertices at distance less than n from v is at most $\sum_{i=0}^{n-1} D^i \leq D^n$, with equality only if $D = 1$ and $n = 1$. If $n \leq \log_D |V(G)|$, then $|V(G)| > D^n$ so there must be a $u \in V(G)$ with $d(u, v) \geq n$. An n vertex subpath of any u-v path is then an induced P_n in G. □

A *path forest* is a graph in which every component is a path. Let $p_i(n)$ be the number of labeled path forests with i components. As we noted above, $\mathcal{B}_n$ is the number of labeled graphs in which every component is a clique. It is clear that the number of labeled graphs with i components such that each component is a clique must be less than the number of such graphs with path components, and thus we have $\sum_{i=1}^{n} p_i(n) > \mathcal{B}_n$, but in fact the Bell number is dominated by any term in the sum that corresponds to path forests with fewer than $\sqrt{n}$ parts, as shall be shown in the following lemma. This is not too surprising, as $p_1(n)$ is the number of cyclic permutations of n, which clearly dominates the number of partitions of n.

Lemma 5.15. *If $n > c^2$, then $p_c(n) > \mathcal{B}_n$.*

Proof. How many path forests are there on $[n]$ with c parts? In order to form a labeled path forest, we could start with any permutation of $[n]$ and break it into c parts by splitting it at $c-1$ places. This does not necessarily yield a unique path forest, as any rearrangement of the paths would allow a different permutation of $[n]$ to give the same labeled path forest, as would reversing the direction of any non-trivial path. Hence each path forest can be represented by at most $c!2^c$ different permutations of $[n]$. Thus,

$$p_c(n) \geq \frac{n!\binom{n-1}{c-1}}{2^c c!} > n!,$$

the latter inequality coming from the standard bound on the binomial coefficient $\binom{a}{b} > (ea/b)^b$ and the choice of $n > c^2$. Since, clearly, $\mathcal{B}_n < n!$, the result is shown. $\square$

5.6 A lower bound on the penultimate range

It is not too surprising that there is a relationship between the speed of a property and the structure of graphs in that property. We can count the number of labelings of a graph roughly by grouping vertices into classes that can be distinguished from each other and then choosing labels for a group en masse. For example, if a graph G consists only of disjoint cliques, then the number of labelings of G is $\binom{n}{c_1,c_2,\ldots,c_m}$, where $c_1,\ldots,c_m$ are the orders of the cliques. Continuing the argument, consider the property, $\mathcal{P}_{cl}$, where each component of each graph in the property is a clique. Then a labeled graph in $\mathcal{P}$ on n vertices can be described by an unordered partition of $[n]$; in fact there is a one-to-one correspondence between graphs of $\mathcal{P}_{cl}^n$ and such partitions of $[n]$. Hence $|\mathcal{P}_{cl}^n| = \mathcal{B}_n$. Recall that the n^{th} Bell number, $\mathcal{B}_n$, is the number of partitions of $[n]$ and is asymptotically $\mathcal{B}_n \sim (n/\log n)^{n/2}$.

In fact, this property, $\mathcal{P}_{cl}$, and its complement $\mathcal{P}_{\overline{cl}}$ consisting of Turán graphs (and their induced subgraphs), are the only properties with speed exactly $\mathcal{B}_n$, as shall be shown below.

If the groupings of vertices are not cliques, then clearly such a count gives only a lower bound. Such lower bounds are instructive, however, in examining the speeds that occur.

We are now ready to prove our main results. We shall prove the lower bound in two parts, considering properties where $k_{\mathcal{P}} < \infty$ and where $k_{\mathcal{P}} = \infty$. The hard work of the former case has been done by the collection of lemmas and theorems in the proceeding sections.

Theorem 5.16. *Let $\mathcal{P}$ be a hereditary property with $|\mathcal{P}^n| \geq n^{(1+o(1))n}$. If $k_{\mathcal{P}} < \infty$, then, for n sufficiently large, $|\mathcal{P}^n| > \mathcal{B}_n$.*

Proof. Let $\ell_{\mathcal{P}}$ be given by Theorem 5.12. Let $c = (2\ell_{\mathcal{P}}k_{\mathcal{P}} + 2k_{\mathcal{P}} + 1)\ell_{\mathcal{P}}$ and assume $n > c^2$. By Theorem 5.13, for all m, $\mathcal{P}$ contains a strong $(\ell_{\mathcal{P}}, k_{\mathcal{P}})$-graph H such that $\varphi(H)$ has a component of order m. Since $\varphi(H)$ has bounded degree and m is arbitrarily large, Lemma 5.14 says $\mathcal{P}$ contains a graph H such that $\varphi(H)$ contains a path of length $m(\ell_{\mathcal{P}}, n, c)$. Color the vertices of this path according to the $(\ell_{\mathcal{P}}, k_{\mathcal{P}})$-partition of H. According to Lemma 5.9, this path contains any path-forest of total length

n and c components in a way that each class of the partition of H is intersected at least $c/\ell_{\mathcal{P}} = 2\ell_{\mathcal{P}}k_{\mathcal{P}} + 2k_{\mathcal{P}} + 1$ times. For each of these path forests F, Lemma 5.10 guarantees a graph $G_F \leq H$ such that $\varphi(G_F) = F$. Since $\mathcal{P}$ is hereditary, $G_F \in \mathcal{P}$ for all such path forests F. Now let $\mathcal{F}$ be the collection of all labeled path forests with n vertices and c components. With our choice of n, Lemma 5.15 says $|\mathcal{F}| > \mathcal{B}_n$, and, by Lemma 5.11, $|\mathcal{P}^n| > |\mathcal{F}|$, since each graph in $\mathcal{F}$ is the image of some unique graph in $\mathcal{P}^n$ under φ. □

The proof for when $k_{\mathcal{P}} = \infty$ involves case analysis of the structures that might occur and an application of the results of Section 5.3. Note that, both by a theorem of [11] and independently by the theorem below, $k_{\mathcal{P}} = \infty$ implies $|\mathcal{P}^n| \geq n^{(1+o(1))n}$.

Theorem 5.17. *Let $\mathcal{P}$ be a hereditary property. If $k_{\mathcal{P}} = \infty$ then, for n sufficiently large, $|\mathcal{P}^n| \geq \mathcal{B}_n$. Equality holds if and only if $\mathcal{P} = \mathcal{P}_{cl}$ or $\mathcal{P}_{\overline{cl}}$.*

Proof. Fix n. We shall show that $|\mathcal{P}^n| \geq \mathcal{B}_n$ and note the restrictive criteria for equality. Let k and r be large enough to guarantee that the Ramsey results we apply below hold, and let $m \geq f(r,r,r)$, where $f(r,r,r)$ is the function from Lemma 5.3.

By the definition of $k_{\mathcal{P}}$, for all k, m, there is a $G \in \mathcal{P}$ and $X \subseteq V(G)$ such that X distinguishes m sets each of order at least k. Let $G \in \mathcal{P}$ be such a graph for our choices of k and m. Let X be a distinguishing set for G and $V_1, \ldots, V_m$ be distinguished sets of order at least k. Let $\mathcal{H} = (X, \mathcal{E})$ be the hypergraph defined by $\mathcal{E} = \{\Gamma_X(V_i)\}$. That is, the vertices of $\mathcal{H}$ are

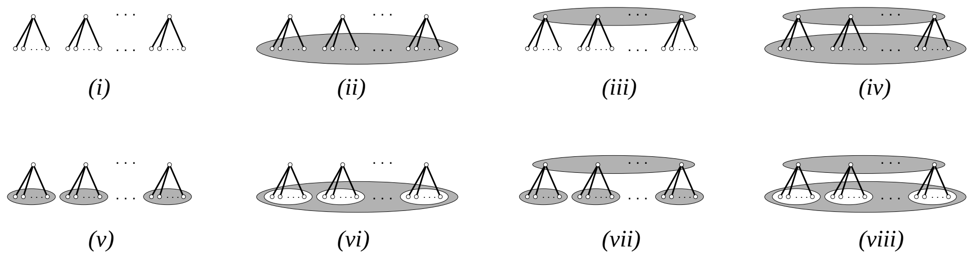


Figure 5.1: The eight possibilities for S if the hypergraph based on $G \in \mathcal{P}$ contains an r-star. The grey ovals indicate sets which induce a clique, while an empty oval within a grey oval represents an induced independent set within an otherwise fully connected group of vertices (i.e. a Turán graph). In each figure, the top vertices are $S_1 = \{v_1, \ldots, v_n\}$ and the bottom vertices are $S_2 = U_1 \cup U_2 \cup \cdots \cup U_n$.

the distinguishing vertices of G and the edges correspond to the subsets of X that create the distinguished partition. Note that $\mathcal{H}$ has no multiple edges, so $|\mathcal{E}| = m$. Hence, by Lemma 5.3, $\mathcal{H}$ contains an r-star, r-costar, or an r-skewchain. Consider the induced subgraph $S' \leq G$ corresponding to this r-star, r-costar, or r-skewchain.

The graph S' contains a set, X, of r vertices which distinguish r sets, $V_1, \ldots, V_r$, each of order at least k. As k and r were chosen large enough, Theorem 5.4 of Ramsey guarantees that each of $X, V_1, \ldots, V_r$ contains either a large clique or a large independent set. Similarly, Corollary 5.8 guarantees that among the distinguished sets, ignoring their internal structure, there is a large spanning independent set or a large Turán graph. Thus we may choose a distinguishing set $S_1 \subseteq X$ so that S_1 is either K_n or $\overline{K_n}$ and similarly choose distinguished sets $U_1, \ldots, U_n$ (subsets, with renumbering of the sets $V_1, \ldots, V_n$) that are either K_n or $\overline{K_n}$ uniformly and so that all pairs of distinguished sets ((U_i, U_j)) induce either $K_{n,n}$ or $\overline{K_{n,n}}$ uniformly.

Let $S \subseteq S'$ be the graph corresponding to our choices made above. Its distinguishing set is called S_1 and let $S_2 = V(S) \setminus S_1$. Then $|S_1| = n$ and $|S_2| = n^2$. Further, S_1 is either K_n or $\overline{K_n}$ and S_2 is one of K_{n^2}, $\overline{K_{n^2}}$, nK_n, or $\overline{nK_n} = T(n^2, n)$.

Based on the manner in which S is created as a subgraph of S', we know that the hypergraph based on the distinguishing relationship between S_1 and S_2 is either an n-star, n-costar, or n-skewchain. There are 24 different possible structures that can be described as above, and $\mathcal{P}$ must contain an arbitrarily large graph containing one of these structures. The 8 possibilities when the hypergraph based on S is an n-star are shown in Figure 5.1; there are similarly 8 possible structures if that hypergraph is a costar, and 8 more when it is a skewchain.

If $S_2 = nK_n$ or $\overline{nK_n}$, then $\mathcal{P}^n_{cl} \subseteq \mathcal{P}^n$ or $\mathcal{P}^n_{\overline{cl}} \subseteq \mathcal{P}^n$, respectively, and $|\mathcal{P}^n| \geq \mathcal{B}_n$, since $|\mathcal{P}^n_{cl}| = \mathcal{B}_n$ as noted earlier. Note that equality occurs if and only if the property is $\mathcal{P}_{cl}$ or $\mathcal{P}_{\overline{cl}}$.

Hence we need only consider the cases when $S_2 = K_{n^2}$ or $\overline{K_{n^2}}$. For the configurations based on stars, these are the top 4 structures shown in Figure 5.1. Considering costars and skewchains, then, there are 12 possibilities in total to be considered. We may cut this number in half by counting the number of labelings of the complementary property $\overline{\mathcal{P}} = \{\overline{G} : G \in \mathcal{P}\}$. Clearly $|\overline{\mathcal{P}}^n| = |\mathcal{P}^n|$. Each possible configuration based on a star is the complement of a configuration based on a costar, and the 4

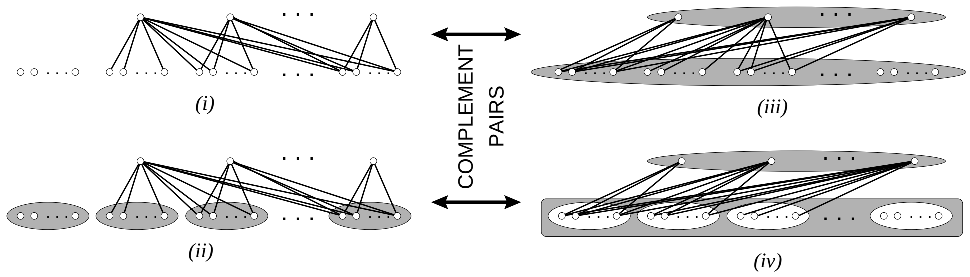


Figure 5.2: The four possibilities for S if a hypergraph based on $G \in \mathcal{P}$ contains an r-skewchain, but the property does not contain $\mathcal{P}_{cl}$ or its complement. The horizontal pairings indicate complementary pairs of graphs. The grey ovals indicate sets which induce a clique, while an empty oval within a grey oval represents an induced independent set within an otherwise fully connected group of vertices (i.e. a Turán graph). In each figure, the top vertices are $S_1 = \{v_1, \ldots, v_n\}$ and the bottom vertices are $S_2 = U_0 \cup U_1 \cup U_2 \cup \cdots \cup U_n$.

remaining configurations based on a skewchain may be paired as shown in Figure 5.2, so we need only consider the left partner of each pair in that figure. The 2 skewchains and 4 chains give us 6 cases to consider.

In each of the cases, we shall show that $|\mathcal{P}^n| > \mathcal{B}_n$ by finding a correspondence between subgraphs of S on n vertices and partitions of $[n]$.

We describe a function from partitions of $[n]$ to subgraphs of S as follows, which we will refer to as f. Call a class of a partition *non-trivial* if it has at least 2 elements; a *singleton* is a vertex forming a trivial class. First order the partition so that the smallest entries of the non-trivial sets in the partition occur in increasing order. If S is not like Figure 5.1(iv), $\overline{K_n} \leq S$ and so we map the singletons to isolated vertices. With the non-trivial classes, we proceed as follows: the i^{th} set in the partition labels v_i and U_i, with v_i receiving the smallest label in the set. For Figure 5.1(iv), the singleton labels become labels of vertices in S_2 after all of the numbers in non-trivial sets have been assigned. The graph associated with the partition is then the subgraph induced by labeled vertices. If we show that this mapping is a bijection, then we have the desired result. This will be our strategy. Given a graph that is the image of a partition of $[n]$ under f, we shall uniquely reconstruct that partition. In some cases, however, we shall need to modify f before applying this strategy, in others we shall enumerate the few exceptions that cannot be reconstructed and count them separately. These sub-strategies shall be made clear in the cases below.

Case 1. S is as in Figure 5.1(i).

In this case, the image of any partition is a star forest, and we can reconstruct the partition according to its components. As the smallest label in a set of the partition always gets mapped to the center of the star, no two partitions give the same star forest, so f is a bijection. Further, some labeled star forests (and hence subgraphs of S) are not the images of any partition (e.g. the star $K_{1,n-1}$ with center labeled n), so the inequality is strict.

Case 2. S is as in Figure 5.1(ii).

To reconstruct the partition, we need consider only non-isolated vertices, as isolated vertices must correspond to singletons in the original partition. For each non-isolated vertex, we consider the number of maximal cliques that it is a member of, where by *maximal clique* we mean a clique that is not a proper subset of any clique. If each of these vertices is in only one maximal clique, then this graph is the image of a partition with only one non-trivial set. Otherwise, all vertices that are members of only one maximal clique are the smallest elements of their set in the partition, and the clique partition of the graph that they induce is the partition that yielded the graph. Once again we may find subgraphs of S that are not the images of any partition. Indeed, any labeled subgraph of S with more than one maximal clique that gives the label n to some non-isolate that is in only one maximal clique (some element of that clique must be in more than one maximal clique) is not the image of any partition. Therefore, the inequality is strict.

Case 3. S is as in Figure 5.1(iii).

To reconstruct the partition, we do the same as in Case 2 in reverse. If every vertex is in only one maximal clique, then this graph is the image of a partition with at most one non-trivial class, this class having order 2. Otherwise, the non-trivial component of the graph has at least 4 vertices, and the vertices with degree greater than 1 induce a clique. This clique is S_1, and induces (by neighborhoods) a partition of the graph into stars. This, in turn, corresponds to the original partition. Again, many subgraphs of S are missed (e.g. for any $k > 1$, any labeling of the graph consisting of a k-clique with a pendant edge and $n-k-1$ isolated vertices), so the inequality is strict.

Case 4. S is as in Figure 5.1(iv).

Recall that in this case we use a slightly different definition of f to account for the lack of isolates in S. Here the isolates get mapped to vertices in S_2 after the labels in the non-trivial parts are mapped. The function f, defined either way, is not a bijection. However, it is "almost" a bijection; we shall isolate those configurations which do not have a unique preimage under f and show that enough subgraphs of S are not in the image of f to account for the overlap.

We again proceed by identifying the maximal cliques in the graph. Note that no vertex of S is in more than two maximal cliques, so in any induced subgraph no vertex is in more than two maximal cliques either. Also note that the function f will always yield a connected graph.

There are no isolated vertices in S, so we cannot immediately identify the trivial sets of the partition. If every vertex is in only one maximal clique, since the graph is a subgraph of S it must be a complete graph. The graph might be the image of either the discrete or indiscrete partition. We shall deal with this case below.

If the graph has more than one maximal clique, then the partition that yielded the graph must have at least one non-trivial part that is not all of $[n]$. If the partition has exactly one non-trivial part, then the graph will have exactly two maximal cliques. If the partition has at least two non-trivial parts, then the graph will have at least 4 maximal cliques.

Suppose the graph has exactly two maximal cliques. Consider the set, M, of all vertices appearing in both cliques. There must be a vertex u such that one of the two cliques is $M \cup \{u\}$, where the label on u is smaller than any label in M and $M \cup \{u\}$ corresponds to the non-trivial part of the partition that yielded the graph. If one of the cliques has more than one element outside of M, or has an element outside of M with a label larger than one appearing in M, then u, and hence the partition, is uniquely determined. Hence, the only way that $M \cup \{u\}$ is not uniquely determined is if $M = \{3, \ldots, n\}$. Then the partition that yielded the graph is either $\{\{1, 3, 4, \ldots, n\}\{2\}\}$ or $\{\{2, \ldots, n\}\{1\}\}$. This is the second case of two partitions yielding the same graph, and shall again be dealt with below.

As noted above, if the graph has more than two maximal cliques, it has at least 4, and the partition that yielded it has at least two non-trivial parts. Any vertex corresponding to an element in a non-trivial part appears in two maximal cliques: S_1 or S_2 and the clique corresponding to its part in the partition. Hence if a graph has at least 4 maximal cliques and some vertex appears in only one maximal clique, then it corresponds to an isolate in the original partition, the maximal clique it is a member of is S_2, and the partition may be reconstructed according to the cliques that intersect S_2.

So let us assume that the graph under consideration has at least 4 maximal cliques and each vertex is in two maximal cliques. Thus the partition that yielded it has no singleton sets and is not the indiscrete partition.

If the graph contains at least 5 maximal cliques, then, since the graph is a subgraph of S, there are 2 non-intersecting cliques which partition

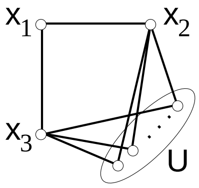


Figure 5.3: A troublesome configuration for an image under f.

the vertices (each of the other cliques intersect both, but not each other). These non-intersecting cliques are S_1 and S_2, and the partition may be reconstructed corresponding to the other maximal cliques.

Thus (except for the two cases deferred above) we must only consider graphs that have exactly 4 maximal cliques and every vertex in exactly two of those. Since there are more than two maximal cliques, the partition that yielded the graph must have at least two non-trivial parts, and since there are fewer than 5 maximal cliques it must have no more than two non-trivial parts. Hence the partition that yielded the graph must be a two-set partition of $[n]$ with no singletons (the latter condition as no vertex is in only one maximal clique). Therefore the graph has a clique with only two vertices. If only one clique has exactly two vertices, then these correspond to the smallest elements of their respective parts in the partition, and we may reconstruct the partition according to the other cliques they are in. Thus only case left to consider is that shown in Figure 5.3.

Assume here $n > 4$. Then $|U| > 1$ so the vertices of U must have come from S_2 and x_1 must then have come from S_1. Hence, the label on x_1 must be smaller than the label on either x_2 or x_3. If x_1 is only smaller than one of them, then the partition is known. For example, if label($x_2 <$ label(x_1) but label(x_3) $>$ label(x_1), then $\{\{x_1, x_3\}, \{x_2\} \cup U\}$ is the original partition. So assume that the label on x_1 is smaller than the labels on both x_2 and x_3. Then label(x_1)$=1$, since all the labels on U must be bigger than one of label(x_2) or label(x_3). Similarly, the label 2 can only occur on x_2 or x_3, so without loss of generality the label on x_2 is 2. Now if the label on x_3 is not 3, then some vertex in U is 3 and the label on x_3 is bigger than 3, so again $\{\{x_1, x_3\}, \{x_2\} \cup U\}$ must be the original partition. If the label on x_3 is in fact 3, then we cannot be sure whether the graph is the image of $\{\{x_1, x_3\}, \{x_2\} \cup U\}$ or $\{\{x_1, x_2\}, \{x_3\} \cup U\}$. These pairs and the two pairs mentioned earlier get mapped to the same graph.

So that this will not happen, we shall modify our encoding as follows. Map the partition $\{\{1, 2\}, \{3, \ldots, n\}\}$ to its image under f, but map $\{\{1, 3\}, \{2, 4, 5, \ldots, n\}\}$, which under f gets mapped to the same image, to the graph in Figure 5.3 so x_1 is labeled 3, x_2 is labeled 1, and x_3 is labeled 2. This latter graph is not the image under f of any partition, as the label on x_1 is not smaller than that on x_2 or x_3. Similarly, we map the indiscrete partition as usual but the discrete partition to the graph in Figure 5.3 so that x_1 is labeled n, x_2 is labeled $n-1$, and x_3 is labeled $n-2$. Map $\{\{1, 3, 4, \ldots, n\}\{2\}\}$ to its normal image under f but map $\{\{2, \ldots, n\}\{1\}\}$ to the same graph (an $n-2$ clique with two pendant vertices) with the pendant vertices labeled n and $n-1$. Once again, this latter graph is not the image of any partition, as may be seen by the arguments in the paragraph discussing that case. This new function is clearly invertible.

Again, we have missed many graphs, including any graph isomorphic to that in Figure 5.3 where $\{x_1, x_2, x_3\}$ is labeled from any set other than $\{1, 2, 3\}$ or $\{n-2, n-1, n\}$ (among others). Hence the inequality is strict.

Case 5. S is as in Figure 5.2(i).

We shall refer to the labeling described in the caption. To reconstruct the partition, we consider simply the degrees of the vertices. All vertices of degree 0 are singletons in the original partition. If a vertex has degree 1, it is either in U_1 or is v_p, where p is the number of parts in the partition (that is, the last of the vertices in S_1). In the latter case, the last non-singleton set in the partition must have exactly two elements.

We may determine which vertices belong to U_1 as follows. If any pair of vertices with degree 1 have a common neighbor, then they are both in U_1 and their common neighbor is v_1, and its neighbors of degree 1 constitute U_1. We may completely reconstruct the partition by considering the neighbors of v_1 of degree 2. These constitute U_2 and their common neighbor other than v_1 is v_2. We may continue in this fashion to reconstruct all sets U_i and thus determine the original partition.

If no pair of vertices with degree 1 has a common neighbor, then there can be at most two vertices of degree 1. If there is only one of degree 1, this must be the entirety of U_1, its neighbor is v_1, and we may proceed as above. Otherwise there are exactly two vertices with degree 1 (and the first and last non-trivial set in the partition each have exactly two elements). Call these vertices x and y, and their neighbors x' and y', respectively.

Then either the label on x is bigger than the label on x' and x' is v_1, or the label on y is bigger than the label on y', and y' is v_1. Once we have identified v_1, we may proceed as above to reconstruct the partition. Note that the possibilities for x and y above are exclusive. We may construct a class of labeled graphs that are not the images of any partition, and thus obtain a strict inequality, by considering the last case and, for example, labeling two vertices of degree 1 with labels 1 and 2.

Case 6. S is as in Figure 5.2(ii).

We proceed as in the previous case, but, for non-isolates, rather than consider the degree of each vertex we consider the number of maximal cliques it is in. With this change, the argument is identical, and the same type of example described there shows that the inequality is strict. $\square$

Taken together, these results give the following corollary, the goal of this paper.

Theorem 5.18. *If $|\mathcal{P}^n| \geq n^{(1+o(1))n}$, then $|\mathcal{P}^n| \geq \mathcal{B}_n$ for all n. Furthermore, equality is possible only if $\mathcal{P} = \mathcal{P}_{cl}$ or $\mathcal{P} = \mathcal{P}_{\overline{cl}}$.* $\square$

In the past, we have sought to give, in addition to bounds on the speed of properties, collections of minimal properties that "force" the speed to be in the range given. For the penultimate range, this type of result is only partially done.

Let G_1 be the infinite graph with structure given in Figure 5.1(i), i.e. an infinite forest of infinite stars. Similarly, define G_2, G_3, $\ldots$, G_6 as the infinite graphs corresponding to Figures 5.1(ii), 5.1(iii), 5.1(iv), 5.2(i), and 5.2(ii), respectively. Let $\mathcal{P}(G_i)$ be the property containing all finite induced subgraphs of G_i. Then the proof of Theorem 5.17 implies that under the hypotheses of that theorem, $\mathcal{P}$ contains one of $\{\mathcal{P}_{cl}, \mathcal{P}(G_1), \mathcal{P}(G_2), \ldots, \mathcal{P}(G_6)\}$ or its complement. However, these are not the minimal properties for the penultimate range, as it is only when $k_{\mathcal{P}} = \infty$ that we can guarantee the inclusion. It would seem that a characterization of minimal properties for $k_{\mathcal{P}} < \infty$ would not have a simple representation, although surely there is such a class. This class of minimal properties would have to be based on φ-transformations of path forests, and we would be happy to see such a result in the future.

The space provided by the strict inequality in Theorem 5.16, which is due to Lemma 5.15, does tell us that the very smallest of properties in this range, however, do in fact contain one of the properties listed above. In particular, if $\mathcal{B}_n \leq |P^n| < \sqrt{n}\mathcal{B}_n$, then $\mathcal{P}$ must contain one of these properties, the upper bound given by the bound on $p_c(n)$. This itself may be a jump, and further study is warranted.

In fact, it is unclear whether there are jumps within the penultimate range at any point between its bounds. This promises to be a rich area of research in the future.

Bibliography

[1] M. Aigner, E.Triesch and Zs.Tuza, Searching for acyclic orientations of graphs, Discrete Mathematics 144 (1995) 3-10.

[2] V. E. Alekseev, Hereditary classes and coding of graphs, *Probl. Cybern.* **39** (1982), 151–164 (in Russian).

[3] V. E. Alekseev, On the entropy values of hereditary classes of graphs, *Discrete Math. Appl.* **3** (1993), 191–199.

[4] N. Alon and B. Bollobás, Graphs with a small number of distinct induced subgraphs, *Disc. Math.* **75** (1989), 23–30.

[5] N. Alon and A. Hajnal, Ramsey graphs contain many distinct induced subgraphs, *Graphs and Combinatorics* **7** (1991), 1–6.

[6] N. Alon, L. Rónyai, and T. Szabó, Norm-graphs: variations and applications, *J. Combin. Theory Ser. B* **76** (1999), 280–290.

[7] N. Alon and J. H. Spencer, The probabilistic method, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, (1992), xvi+254 pp. ISBN 0-471-53588-5

[8] N. Alon and Zs.Tuza, The acyclic orientation game on random graphs, Random Structures and Algorithms 6 (1995) 261-268.

[9] T. Andreae, "On the Interval Number of a Triangulated Graph" *J. Graph Theory* **3** (1987), 273-280.

[10] J. Balogh, B. Bollobás and M. Simonovits, On the number of Ł-free graphs, submitted to JCT B.

[11] J. Balogh, B. Bollobás, and D. Weinreich, The speed of hereditary properties of graphs, *J. Combin. Theory Ser. B*, **79**, (2000) 131–156.

[12] J. Balogh, B. Bollobás, and D. Weinreich, The penultimate range of growth for graph properties, to appear in *European J. of Comb.*

[13] J. Balogh, B. Bollobás, and D. Weinreich, The size and speed of monotone properties of graphs, to appear in *Discrete Applied Mathematics*

[14] J. Balogh, B. Bollobás, and D. Weinreich, An extension of Dilworth's Theorem to hypergraphs, submitted to JCT B.

[15] J. Balogh and A. Pluhár, "A Sharp Edge Bound on the Interval Number of a Graph" *J. Graph Theory* **32** (1999), 153–159.

[16] J. Balogh and A. Pluhár, The interval number of Dense Graphs, To appear in Discrete Mathematics.

[17] B-PP) J. Balogh, P, Ochem, and A. Pluhár, On the Interval Number of Special Graphs. submitted to Journal of Graph Theory.

[18] B. Bollobás, *Extremal Graph Theory*, Academic Press, London (1978).

[19] B. Bollobás, *Modern Graph Theory*, Springer-Verlag, New York (1998).

[20] B. Bollobás and Tamás Szabó, The oriented cycle game, Discrete Mathematics 186 (1998) 55–67.

[21] B. Bollobás and A. Thomason, Projections of bodies and hereditary properties of hypergraphs, *J. London Math. Soc.* **27** (1995), 417–424.

[22] B. Bollobás and A. Thomason, Hereditary and monotone properties of graphs, in "The Mathematics of Paul Erdős II" (R. L. Graham and J. Nešetřil, eds.) *Algorithms and Combinatorics*, **14**, Springer-Verlag (1997), 70–78.

[23] B. Bollobás and A. Thomason, Generalized chromatic numbers of random graphs, *Random Structures and Algorithms* **6** (1995), 353–356.

[24] B. Bollobás and A. Thomason, Colouring random graphs by hereditary properties, to appear.

[25] B. Bollobás and A. Thomason, The structure of hereditary properties and colourings of random graphs, *Combinatorica*, **20** (2000) no. 2, 173–202.

[26] B. Bollobás and P. Erdős, On the structure of edge graphs, *Bull. London Math. Soc.* **5** (1973), 317–321.

[27] G. Chartrand, F. Harary, M. Schultz and D. Vanderjagt, Achievement and Avoidance of a Strong Orientation of a Graph, Congressus Numerantium 108 (1995) 193–203.

[28] G. Chartrand, D.W. VanderJagt, B.Q. Yue, Orientable Domination in Graphs, Congressus Numerantium 119 (1996) 51–63.

[29] J. Edmonds, "Minimum partition of a matroid into independent subsets", *J. Res. Nat. Bur. Stand.* **69B** (1965), 73–77.

[30] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183–190.

[31] P. Erdős, Some recent results on extremal problems in graph theory. Results. *Theory of Graphs (Internat. Sympos., Rome, 1966)*, 117–123.

[32] P. Erdős, On some new inequalities concerning extremal properties of graphs, *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, 77–81.

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

[34] P. Erdős and A. Hajnal, Ramsey-Type theorems. Combinatorics and complexity (Chicago IL, 1987) *Discrete Appl. Math.* **25** (1989) no. 1-2, 37-52.

[35] P. Erdős and A. Hajnal, On the number of distinct induced subgraphs of a graph, *Disc. Math.* **75** (1989), 145–154.

[36] P. Erdős, D. J. Kleitman, and B. L. Rothschild, Asymptotic enumeration of K_n-free graphs, in *Colloquio Internazionale sulle Teorie Combinatorie*, Atti dei Convegni Lincei (Rome) **17** Tomo II (1976), 19–27.

[37] P. Erdős and A. Rényi, On a new law of large numbers *Journ. Analyse Math.***22** (1970), 103–111.

[38] P. Erdős and M. Simonovits, Some extremal problems in graph theory, *Combinatorial Theory and Its Applications, I (Proc. Colloq., Balatonfüred, 1969)*, 377–390.

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

[40] P. Erdős and J. Spencer, *Probabilistic methods in combinatorics,* Academic Press, New York-London, 1974.

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

[42] J.F. Fink and M.S. Jacobson, n-domination in graphs, in Y.Alavi and A.J. Schwenk, editors, Graph Theory with Applications to Algorithms and Computer Science (Wiley, New York, 1985) 283-300.

[43] P. Frankl and V. Rödl, Hypergraphs do not jump, *Combinatorica* **4** 2/3 (1984), 149–159.

[44] P. Frankl and V. Rödl, Some Ramsey-Turán type results for hypergraphs, *Combinatorica* **8** 4 (1988), 323–332.

[45] T.W. Haynes, S.T. Hedetniemi, P.J. Slater editors, Domination in Graphs (Marcel Dekker, New York, 1998).

[46] T.W. Haynes, S.T. Hedetniemi, P.J. Slater, Fundamentals of Domination in Graphs (Marcel Dekker, New York, 1998).

[47] C. Hundack, H. J. Prömel, and A. Steger, Extremal graph problems for graphs with a color-critical vertex. Combinatorics, Probability and Computing, 2 (1993), no. 4, 465–477.

[48] M. C. Golumbic, *Algorithmic graph theory and perfect graphs*, Academic Press, New York-London-Toronto, Ont., 1980.

[49] J.R. Griggs, "Extremal Values of the Interval Number of Graph, II.", *Discrete Math.* **28** (1979), 37–47.

[50] J.R. Griggs and D.B. West, "Extremal Values of the Interval Number of Graph, I.", *SIAM J. Alg. Discrete Meth.* **1** (1980), 1-8.

[51] A. Hajnal and E. Szemerédi , "Proof of a conjecture of P. Erdős", *Colloquia Mathematica Societatis János Bolyai 4., Combinatorial Theory and its Applications II.* (1970), 601–623.

[52] L. B. Hopkins, W. T. Trotter and D. B. West, The Interval Number of a Complete Multipartite Graph, *Discrete Applied Math.* **8** (1984), 163-187.

[53] F. Jaeger and C. Payan, Relations du type Nordhaus-Gaddum pour le nombre d'absorption d'un graphe simple, C.R. Acad. Sci. Ser. A 274 (1972) 728–730.

[54] D. J. Kleitman and K. J. Winston: On the number of graphs without 4-cycles, *Discrete Mathematics* **41**, (1982) 167–172.

[55] Ph. G. Kolaitis, H. J. Prömel, and B. L. Rothschild, K_{l+1}-free graphs: Asymptotic enumeration and a 0-1-law, *Trans. Amer. Math. Soc.* **303** (1987) 637–671.

[56] J. Kollár, L. Rónyai, and T. Szabó, Norm-graphs and bipartite Turán numbers, *Combinatorica* **16** 3 (1996), 399–406.

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

[58] C.St.J.A. Nash-Williams, "Edge-disjoint spanning trees of finite graphs", *J. London Math. Soc.* **36** (1961), 445–450.

[59] E.A. Nordhaus and J.W. Gaddum, On complementary graphs, Amer. Math. Monthly 63 (1956) 175–177.

[60] A. Pluhár, "Remarks on the Interval Number of Graphs" *Acta Cybernetica* **12** (1995) 125–129.

[61] H. J. Prömel and A. Steger, Excluding induced subgraphs: Quadrilaterals, *Random Structures and Algorithms* **2** (1991), 55–71.

[62] H. J. Prömel and A. Steger, Excluding induced subgraphs II: Extremal graphs, *Discrete Applied Mathematics* **44** (1993), 283–294.

[63] H. J. Prömel and A. Steger, Excluding induced subgraphs III: A general asymptotic, *Random Structures and Algorithms* **3** (1992), 19–31.

[64] H. J. Prömel and A. Steger, The asymptotic structure of H-free graphs, *Graph Structure Theory* (N. Robertson and P. Seymour, eds), Contemporary Mathematics **147**, Amer. Math. Soc., Providence (1993) 167–178.

[65] H. J. Prömel and A. Steger, Almost all Berge graphs are perfect, *Combin., Prob., and Comp.* **1** (1992), 53–79.

[66] H. J. Prömel and A. Steger, The asymptotic number of graphs not containing a fixed color-critical subgraph. Combinatorica, 12 (1992), no. 4, 463–473.

[67] E.R. Scheinerman, “The Maximum Interval Number of Graphs with Given Genus”, *J. Graph Theory* **3** (1987), 441–446.

[68] E. R. Scheinerman, On the interval number of a chordal graph, *J. Graph Theory* **12** (1988), 311-316.

[69] E.R. Scheinerman, “On the Interval Number of Random Graphs”, *Discrete Math.* **82** (1990), 105–109.

[70] E. R. Scheinerman and J. Zito, On the size of hereditary classes of graphs, *J. Combinat. Theory (B)* **61** (1994), 16–39.

[71] E.R. Scheinerman and D.B. West, “The interval number of a planar graph: three intervals suffice.”, *Jour. Combinat. Theory B* **35** (1985), 224-239.

[72] M. Simonovits, A method for solving extremal problems in graph theory, stability problems, *Theory of Graphs (Proc. Colloq., Tihany, 1966*, 279–319.

[73] M. Simonovits: Extremal Graph Theory, Selected Topics in Graph Theory, (ed. by Beineke and Wilson) Academic Press, London, New York, San Francisco, (1983), 161–200.

[74] M. Simonovits, Paul Erdős’ influence on extremal graph theory. The mathematics of Paul Erdős, II, 148–192, Algorithms Combin., 14, Springer, Berlin, 1997.

[75] J.R. Spinrad, G. Vijayan and D.B. West, “An Improved Edge Bound on the Interval Number of a Graph”, *J. Graph Theory* **3** (1987), 447–449.

[76] C. Stracke and L. Volkmann, A new domination conception, J. Graph Theory 17 (1993) 315–323.

[77] E. Szemerédi, Regular partitions of graphs, in *Proc. Colloque Inter. CNRS* (J.-C. Bermond, J.-C. Fournier, M. las Vergnas, D. Sotteau, eds) (1978), 399–401.

[78] M. Tarsi, On the decomposition of graphs into stars, Discrete Mathematics 36 (1981) 299–304.

[79] W.T. Trotter, Jr. and F. Harary, “On Double and Multiple Interval Graphs”, *J. Graph Theory* **3** (1979), 205–211.

[80] P. Turán, Egy gráfelméleti szélsőérték feladatról. (On an extremal problem in graph theory). Mat. Fiz. Lapok, **48** (1941), 436–452.

[81] P. Turán, On the theory of graphs. Colloq. Math. **3** (1954), 19–30.

[82] D. B. West, A short proof of the degree bound on interval number, *Discrete Math.* **73** (1989), 307-310.