Results in Ramsey theory and extremal graph theory

Robert Harry Kaye Balfour Metrebian

Supervisor: Prof. Béla Bollobás

Department of Pure Mathematics and Mathematical Statistics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Trinity College August 2023

To my mother

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text.

Chapter 2 is based on joint work with Vojtěch Dvořák.

Chapter 4 is based on joint work with Victor Souza.

Robert Harry Kaye Balfour Metrebian
August 2023

Acknowledgements

I would like to begin by thanking my supervisor, Professor Béla Bollobás. Béla has been a constant source of encouragement, and I could not have asked for a more generous person to be my supervisor. As well as all the beautiful mathematics that he has shown me, he has given me the opportunity to meet incredible people and visit amazing places. Béla has fostered a wonderful mathematical community and it has been a privilege to be a part of it. In addition, I would like to thank Béla and his wife Gabriella for being such welcoming hosts, both in Cambridge and abroad.

I would also like to thank my graduate advisor, Professor Imre Leader, who has been very helpful to me throughout my studies. His enthusiasm for combinatorics is infectious and has been rubbing off on me ever since I first met him.

I would like to express my gratitude to my collaborators, Vojtěch Dvořák and Victor Souza, and my other academic siblings, Marius Tiba, Peter van Hintum, Adva Mond, Jan Petr, Julien Portier, Agnijo Banerjee and Lawrence Hollom. It has been a pleasure to be surrounded by so many brilliant mathematicians, and I have learned a lot from all of you.

I am extremely grateful to Trinity College, Cambridge, not only for funding my studies, but also for providing such a stimulating working environment, and for giving me the opportunity to meet so many wonderful people, both inside and outside mathematics, over the last few years.

Some of my research was carried out on visits to the University of Memphis and to IMPA in Rio de Janeiro, and I would like to thank both institutions for their hospitality. These trips would have been impossible without financial support from Trinity College, Cambridge and the University of Memphis, and I owe further thanks to both.

I would not have been able to make it through my PhD without occasional distractions from my friends. However many people I list, there will be some I feel I have unfairly omitted, but I would like to thank in particular Andrew, Rakesh, Maraia, Lillian, Tim, Savvas, Stefan, Angus and Bianca for making my postgraduate years so enjoyable and for keeping me sane during the dark days of lockdown.

Finally, I would like to say thank you to my mother. You have put a huge amount of time and energy into supporting and encouraging me throughout my educational journey, and I am eternally grateful for that and for everything that you have done for me.

Abstract

In this thesis, we study several combinatorial problems in which we aim to find upper or lower bounds on a certain quantity relating to graphs. The first problem is in Ramsey theory, while the others are in extremal graph theory.

In Chapter 2, which is joint work with Vojtěch Dvořák, we consider the Ramsey number $R(F_n)$ of the fan graph F_n, a graph consisting of n triangles which all share a common vertex. Chen, Yu and Zhao showed that $\frac{9}{2}n - 5 \leq R(F_n) \leq \frac{11}{2}n + 6$. We build on the techniques that they used to prove the upper bound of $\frac{11}{2}n + 6$, and adopt a more detailed approach to examining the structure of the graph. This allows us to improve the upper bound to $\frac{31}{6}n + 15$.

In Chapter 3, we work on a problem in graph colouring. Petruševski and Škrekovski recently introduced the concept of odd colouring, and the odd chromatic number of a graph, which is the smallest number of colours in an odd colouring of that graph. They showed that planar graphs have odd chromatic number at most 9, and this bound was improved to 8 by Petr and Portier. We consider the odd chromatic number of toroidal graphs, which are graphs that embed into a torus. By using the discharging method, along with detailed analysis of a remaining special case, we show that toroidal graphs have odd chromatic number at most 9.

In Chapter 4, which is joint work with Victor Souza, we consider a problem in the hypercube graph Q_n. Huang showed that every induced subgraph of the hypercube with $2^{n-1} + 1$ vertices has maximum degree at least $\lceil \sqrt{n} \rceil$, which resolved a major open problem in computer science known as the Sensitivity Conjecture. Huang asked whether analogous results could be obtained for larger induced subgraphs. For induced subgraphs of Q_n with $p2^n$ vertices, we find a simple lower bound that holds for all p, and substantially improve this bound in the range $\frac{1}{2} < p < \frac{2}{3}$ by analysing the local structure of the graph. We also find constructions of subgraphs achieving the simple lower bound asymptotically when $p = 1 - \frac{1}{r}$.

Table of contents

Chapter 1

Introduction

This thesis consists of four chapters, including this introduction. Each of the remaining chapters is devoted to a different combinatorial problem involving graphs. In Chapter 2 we demonstrate an improved bound for a problem in graph Ramsey theory. In Chapter 3 we prove a bound for a graph colouring problem on the torus. In Chapter 4 we investigate induced subgraphs of the hypercube graph with small maximum degree.

1.1 Ramsey numbers of fans

Chapter 2 is joint work with Vojtěch Dvořák and is adapted from [31].

Let G and H be graphs. The *Ramsey number* $R(G, H)$ is the smallest positive integer N such that if we colour the edges of K_N, the complete graph on N vertices, with two colours, the colouring must contain a copy of G in the first colour or a copy of H in the second colour. When G and H are the same graph, we simply denote this as $R(G)$.

Ramsey [69] proved in 1930 that $R(K_n)$ exists for all n. This result became known as Ramsey's theorem, and it immediately implies the existence of $R(G, H)$ for all graphs G and H. Erdős and Szekeres [34] proved in 1935 that $R(K_n)$ is at most $(4 - o(1))^n$. Despite considerable efforts over almost 90 years, only improvements to the $o(1)$ term were made until Campos, Griffiths, Morris and Sahasrabudhe [19] recently proved that $R(K_n) \leq (4 - c)^n$ for an effective constant c and sufficiently large n. In 1947, Erdős [32] used the probabilistic method to find a lower bound of the form $(\sqrt{2} + o(1))^n$, and the best known lower bound is still of this form.

In Chapter 2, we are instead concerned with the Ramsey numbers of *fans*. The fan graph F_n is a graph consisting of n copies of K_3 which all have a single vertex in common. Ramsey numbers of fans were first investigated by Li and

Rousseau [56] in 1996; they showed that $4n + 1 \leq R(F_n) \leq 8n - 2$. The lower and upper bounds were both improved over the following decades, and in 2021 Chen, Yu and Zhao [24] proved that $\frac{9}{2}n - 5 \leq R(F_n) \leq \frac{11}{2}n + 6$.

In Chapter 2, by building on the techniques of Chen, Yu and Zhao [24], we improve the upper bound on $R(F_n)$ from $\frac{11}{2}n + 6$ to $\frac{31}{6}n + 15$. Our general approach is as follows. We call the two colours black and white, and assume that we have a graph G on at least $\lceil \frac{31}{6}n + 15 \rceil$ vertices with no F_n in either colour. We find the vertex v with the most edges of a single colour, which is black without loss of generality, and examine the subgraph induced by the vertices joined to v by black edges. This subgraph cannot contain a white F_n or a black matching of n edges, and we prove that it must instead contain a large clique A of one colour or the other.

We use the existence of the clique A to prove that G contains further cliques in the other colour which are related to A in a suitable way. We then use the relationships between the cliques to find a monochromatic copy of F_n, which is a contradiction. There are several different cases, of varying degrees of difficulty, depending on the number of vertices joined to v by black edges, as well as other factors specific to the graph in question.

1.2 Odd colourings on the torus

Chapter 3 is adapted from [61].

In 1976, Appel and Haken, assisted by Koch [5, 6], proved the celebrated Four-Colour Theorem, which states that all planar graphs G admit a proper vertex-colouring with at most 4 colours.

The ideas behind Appel and Haken's proof built on the work of many previous researchers. They found a large set of configurations which is both *unavoidable*, meaning that every planar graph contains a configuration in the set, and *reducible*, meaning that any graph which contains a configuration in the set can be reduced to a smaller graph in such a way that if the smaller graph has a proper 4-colouring then so does the original graph. The reducibility of the set was proved by computer, but the unavoidability was shown by hand using a technique called the *method of discharging*, invented by Heesch [49], in which charge is distributed across the vertices and faces of a graph and then redistributed according to a set of rules.

Instead of allowing all proper colourings, we can impose some conditions on which colourings are allowed. For example, Petruševski and Škrekovski [68] recently introduced the concept of *odd colouring*. An odd colouring of a graph G is a proper colouring of $V(G)$ with the property that, for every non-isolated vertex v of G, there is some colour that appears an odd number of times in the neighbourhood of v. The smallest number k such that G admits an odd colouring using k colours is called the *odd chromatic number* of G and denoted $\chi_o(G)$. Petruševski and Škrekovski [68] used the discharging method to prove that every planar graph G satisfies $\chi_o(G) \leq 9$, and Petr and Portier [67] improved this bound to 8.

In Chapter 3, we will consider odd colourings of graphs that embed into the torus; these graphs are called *toroidal graphs*. It is straightforward to show using Euler's formula that the minimum degree of any toroidal graph is at most 6. The discharging rules used by Petruševski and Škrekovski [68] and by Petr and Portier [67] do not allow us to obtain any useful results about odd colourings of toroidal graphs, so we use a new set of discharging rules to show that if G is a minimal toroidal graph with $\chi_o(G) > 9$ then the minimum degree of G cannot be 5 or less. It follows that $\delta(G) = 6$.

Toroidal graphs with $\delta(G) = 6$ were fully classified by Altshuler [4] in 1973. We use this classification to show that all toroidal graphs with $\delta(G) = 6$ admit an odd colouring with at most 9 colours and therefore cannot be a mimimal toroidal graph with $\chi_o(G) > 9$. This implies that all toroidal graphs have odd chromatic number at most 9. Note that at least 7 colours are sometimes required, since K_7 can be embedded in the torus.

Our proof also shows that every graph that embeds into the real projective plane has odd chromatic number at most 9.

1.3 Induced subgraphs of the hypercube

Chapter 4 is based on joint work with Victor Souza.

The hypercube graph Q_n has vertex set $\{0,1\}^n$, with two vertices being adjacent if and only if they differ in exactly one coordinate. The independence number of Q_n is clearly $2^{n-1} = \frac{1}{2}|Q_n|$. In 1988, Chung, Füredi, Graham and Seymour [26] showed that for every n there is an induced subgraph of Q_n with $2^{n-1}+1$ vertices, one more than the independence number, and maximum degree $\lceil \sqrt{n} \rceil$.

In 1992, Gotsman and Linial [42] showed that if Chung, Füredi, Graham and Seymour's upper bound of $\lceil \sqrt{n} \rceil$ was tight, then it would have major consequences in theoretical computer science. More specifically, a certain measure of the complexity of a Boolean function, the *sensitivity*, would be polynomially related to many other measures of complexity, such as the degree of the function as a real polynomial. The question of whether such a polynomial relation exists was an important open problem in computer science and became known as the Sensitivity Conjecture.

In 2019, Huang [51] used a beautiful spectral argument to show that Chung, Füredi, Graham and Seymour's upper bound of $\lceil \sqrt{n} \rceil$ is indeed tight, proving the Sensitivity Conjecture. Huang asked whether the smallest possible maximum degree could be determined for induced subgraphs of Q_n of other sizes.

In Chapter 4, we take the first steps towards resolving Huang's problem. We consider induced subgraphs of Q_n of size $p2^n$ where p is fixed and $\frac{1}{2} < p < 1$. We begin by showing that the maximum degree must be at least $\frac{2p-1}{p}n$ using a double-counting argument. We then use Hamming codes to construct induced subgraphs for which this lower bound is tight: these exist when $p = 1 - \frac{1}{2^r}$ and n is a multiple of $2^r - 1$. We show that for other values of n we can still obtain a maximum degree that is asymptotically equal to $\frac{2p-1}{p}n$. Next, we find constructions based on partitions of the infinite grid $\mathbb{Z}^d$, which allow the bound $\frac{2p-1}{p}n$ to be attained asymptotically (though in a slightly weaker sense than for the Hamming code construction) when p is of the form $1 - \frac{1}{r}$.

The remainder of the chapter is devoted to proving that the lower bound of $\frac{2p-1}{p}n$ is not tight when $\frac{1}{2} < p < \frac{2}{3}$, and finding a better lower bound for this range of p. In fact, we find several lower bounds, each better than the last, by employing more and more complicated counting arguments on small configurations of vertices. Our final lower bound, which we believe is almost certainly not tight, does not have a simple closed-form expression and must be calculated with a computer.

Chapter 2

Ramsey numbers of fans

2.1 Introduction

This chapter is joint work with Vojtěch Dvořák, adapted from [31].

In 1930, Ramsey [69] proved the fundamental result that for every n, there exists N such that if we colour the edges of the complete graph K_N in two colours, then the colouring must contain a monochromatic copy of K_n. We denote the smallest N with this property by $R(n)$ or $R(K_n)$. This result became known as Ramsey's theorem, and the numbers $R(n)$ as Ramsey numbers. For a general graph G, the Ramsey number $R(G)$ is the smallest N such that every colouring of the edges of K_N with two colours contains a monochromatic copy of G. More generally, for two graphs G and H, we write $R(G, H)$ for the smallest N such that every 2-colouring of the edges of K_N contains either a copy of G in the first colour or a copy of H in the second colour. Ramsey's theorem immediately implies that $R(G, H)$ exists for all graphs G and H.

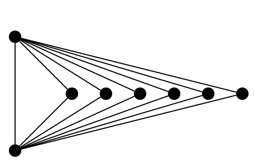
The study of Ramsey numbers began in earnest in 1935 with the work of Erdős and Szekeres [34], who found the first non-trivial upper bounds for $R(K_m, K_n)$, including the bound $R(K_n) \leq (1 + o(1))4^{n-1}/\sqrt{\pi n}$ for the diagonal case. In 1947, Erdős [32] famously proved a lower bound of the form $R(K_n) \geq (\sqrt{2} + o(1))^n$ using the probabilistic method. Ever since, the problem of improving the bounds on $R(K_n)$ has received more attention than almost any other problem in combinatorics. Nonetheless, the lower bound has resisted virtually all attempts to improve it, and the upper bound of Erdős and Szekeres remained the best known for over 50 years until Thomason [81] improved it by a polynomial factor using quasirandomness in 1988. Conlon [27], extending the techniques of Thomason, made a superpolynomial improvement in 2009, and recently Sah [77] built on the methods of Thomason and Conlon to obtain a further superpolynomial improvement. However, these bounds are all still of the form $(4 - o(1))^n$. Very recently, Campos, Griffiths, Morris and Sahasrabudhe [19] made a remarkable breakthrough by improving the upper bound by an exponential factor: they proved that $R(K_n) \leq (4 - c)^n$ for an effective constant c and sufficiently large n.

More progress has been made in the off-diagonal case $R(K_m, K_n)$ when m is small. In 1980, Ajtai, Komlós and Szemerédi [2, 3] proved that $R(K_3, K_n) = O(n^2/\log n)$, and Kim [53] showed in 1995 that this is the correct order of magnitude. The constants have since been improved in each case: Shearer [78] proved in 1983 that $R(K_3, K_n) < (1 + o(1))n^2/\log n$, and the lower bound of $R(K_3, K_n) > (\frac{1}{4} - o(1))n^2/\log n$ was obtained independently using the triangle-free process by Fiz Pontiveros, Griffiths and Morris [39] in 2020, and by Bohman and Keevash [11] in 2021. For $m = 4$, a breakthrough was made very recently by Mattheus and Verstraete [60], who proved that $R(K_4, K_n) = \Omega(n^3/\log^4 n)$. This is within a polylogarithmic factor of the best known upper bound of $(1 + o(1))(n^3/\log^2 n)$, due to Li, Rousseau and Zang [57].

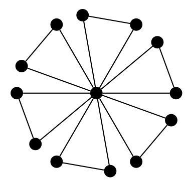
Around 1970, Ramsey numbers began to be investigated for classes of graphs other than complete graphs, and in the following years many results in this area were published. For example, the exact values for $R(C_m, C_n)$, the Ramsey number of two cycles, were found in a series of papers in the early 1970s, by Chartrand and Schuster [22], Bondy and Erdős [14], Rosta [73, 74], and Faudree and Schelp [37]. Many more results in graph Ramsey theory from this period are given in a contemporary survey by Burr [17].

Burr, Erdős and Spencer [18] showed in 1975 that $R(mK_3, nK_3) = 3m + 2n$ for $m \geq n, m \geq 2$, where mK_3 is the union of m disjoint copies of K_3. In particular, $R(nK_3) = 5n$, except for $n = 1$ where we have $R(K_3) = 6$. Instead of taking the copies of K_3 to be disjoint, we could join them together. The *book* B_n is formed by n triangles which all share an edge. In 1978, Rousseau and Sheehan [75] showed that $R(B_m, B_n) \leq 2(m + n + 1)$ for all m, n such that $2(m + n) + 1 \geq \frac{1}{3}(n - m)^2$, and therefore $R(B_n) \leq 4n + 2$. They also showed that this bound on $R(B_n)$ is tight when $4n + 1$ is a prime power. Combined with known bounds on gaps between primes (e.g. [7, 50]), this implies that $R(B_n) = (4 - o(1))n$. More generally, a graph formed from n copies of K_{k+1} which share a common K_k is also called a book and denoted $B_n^{(k)}$, so $B_n = B_n^{(2)}$. Conlon [28] and Conlon, Fox and Wigderson [29] recently generalised Rousseau and Sheehan's result by proving that $R(B_n^{(k)}) = (2^k + o_k(1))n$. Note that the word "book" is used in several other ways in the graph theory literature, for example to mean a graph consisting of 4-cycles sharing a common edge.

After nK_3 and B_n, it is natural to consider a graph consisting of n copies of K_3 which all have a single vertex in common. This graph is called the *fan* F_n.



B_n



F_n

It is also sometimes referred to as the *friendship graph*; this name originates from a result of Erdős, Rényi and Sós [33], who proved that the F_n are the only finite graphs on more than one vertex with the property that every two vertices have exactly one common neighbour. The graph F_n has $2n+1$ vertices: one vertex v, called the *centre* of the fan, and $2n$ other vertices $v_1, \dots, v_{2n}$ such that for $i = 1, \dots, n$, $vv_{2i-1}v_{2i}$ is a triangle. Each of the n edges $v_{2i-1}v_{2i}$ is called a *blade* of the fan. In this chapter we will be concerned with the Ramsey numbers of fans.

While the Ramsey numbers of nK_3 and B_n were studied in the 1970s, it appears that Ramsey numbers related to F_n were not investigated until 1996, when Li and Rousseau [56] proved that $R(F_1, F_n) = 4n+1$ for $n \geq 2$ (note that $F_1 = K_3$), and that $4n+1 \leq R(F_m, F_n) \leq 4(m+n)-2$ for $m \geq 1$ and $n \geq 2$, implying the initial bounds $4n+1 \leq R(F_n) \leq 8n-2$. The next advance was made in 2009 by Lin and Li [58], who showed that $R(F_2, F_n) = 4n+1$ for $n \geq 2$, and that $R(F_m, F_n) \leq 4m+2n$ for $m \geq n \geq 2$, implying that $R(F_n) \leq 6n$ for $n \geq 2$. Shortly afterwards, Lin, Li and Dong [59] showed that $R(F_m, F_n) = 4n+1$ if n is sufficiently large in terms of m, and Zhang, Broersma and Chen [87] later quantified this by proving that $R(F_m, F_n) = 4n+1$ for $n \geq \max\{m^2 - \frac{1}{2}m, \frac{11}{2}m+4\}$. They also showed that $R(F_m, F_n) \geq 4n+2$ for $m \leq n < \frac{1}{2}m(m-1)$, slightly improving the lower bound for $R(F_n)$.

Prior to our result, the best upper and lower bounds for $R(F_n)$ were due to Chen, Yu and Zhao [24], who made a significant advance in 2021 by showing that $\frac{9}{2}n - 5 \leq R(F_n) \leq \frac{11}{2}n+6$. They noted that their lower bound implies $R(B_n) < R(F_n)$ for $n \geq 15$. This, together with the observation that

$$|V(B_n)| < |V(F_n)| < |V(nK_3)|,$$

led them to believe that

$$R(B_n) \leq R(F_n) \leq R(nK_3)$$

should hold for sufficiently large n, but they were unable to show that $R(F_n) \leq R(nK_3) = 5n$.

As our main result, we make a further improvement to the upper bound, decreasing it from $5.5n$ to about $5.167n$.

Theorem 2.1.1. *For every $n \geq 1$, we have*

$$R(F_n) \leq \tfrac{31}{6}n + 15.$$

We are almost certain that $\frac{31}{6}n$ is not the true asymptotic magnitude of $R(F_n)$, and hence we make no attempts to optimise the additive constant in the expression above.

Our approach builds on the ideas of Chen, Yu and Zhao [24]: we aim to find large monochromatic cliques in the graph and then "cover" them in a suitable way. The first crucial new idea in this chapter is that of controlling the degrees of vertices in each colour: for that, we use Lemma 2.2.4. The proof of this lemma is essentially analogous to the proof of a key lemma of Chen, Yu and Zhao, but using this more general version turns out to be very beneficial. In fact, using the techniques of Chen, Yu and Zhao, this lemma alone can be used to obtain $R(F_n) \leq \frac{16}{3}n + O(1)$.

To go further, we must also introduce a different, more global approach in the later parts of the proof: assuming that the graph contains no F_n of either colour, we find several large, suitably related monochromatic cliques and exploit these relations to construct a monochromatic F_n, giving a contradiction.

The rest of the chapter is organised as follows. In Section 2.2, we introduce our notation and state several basic results and lemmas that we will use. In Section 2.3, we give a short, non-technical overview of our proof. In Section 2.4, we go through the technical details of the proof. Finally, in Section 2.5, we briefly outline further directions of research.

2.2 Preliminaries and notation

We use standard graph theoretic notation throughout. For a simple graph G, we denote its vertex set by $V(G)$ and its edge set by $E(G)$. For $A \subset V(G)$, we write $G[A]$ for the induced subgraph on A, and we denote $V(G) \setminus A$ by $\overline{A}$. On the other hand, for a graph H, we will write $\overline{H}$ to mean the complement of H.

For $v \in V(G)$, we write $N(v) = \{w \in V(G) \,|\, vw \in E(G)\}$. More generally, for $S \subset V(G)$, we use $N(S)$ to mean $\bigcup_{v \in S} N(v)$.

Throughout, instead of a 2-colouring, we will consider a graph G on $\lfloor \frac{31}{6}n + 15 \rfloor$ vertices in the usual graph-theoretic sense, and we show that we can always find F_n inside G or $\overline{G}$. This will be done by contradiction: assume from now on that neither G nor $\overline{G}$ contains F_n. We will examine G more and more thoroughly until we are able to reach the desired contradiction.

As the role of the colours is interchangeable, we will often refer to non-edges as *white edges* and to edges as *black edges*. We also sometimes refer to an independent set of vertices as a white clique and to a clique in the usual sense as a black clique. Accordingly, we write $N_W(v) = \{w \in V(G) \setminus \{v\} \mid vw \notin E(G)\}$ for the white neighbourhood of v, and similarly $N_W(S) = \bigcup_{v \in S} N_W(v)$ for $S \subset V(G)$. If $\overline{G}$ contains a copy of some graph H, we will sometimes say that G contains a white H. In diagrams, black and white edges will be represented by solid and dashed lines respectively.

For a graph H, denote the size of the largest matching of H by $\nu(H)$, and the number of odd sized components of H by $q(H)$. Let us recall the following deficient forms of the fundamental theorems of Hall and Tutte (see, e.g., [12]).

Theorem 2.2.1 (Hall)**.** *Let H be a bipartite graph on parts X and Y. For any non-negative integer d, $\nu(H) \geq |X| - d$ if and only if $|N(S)| \geq |S| - d$ for every $S \subset X$.*

Theorem 2.2.2 (Tutte)**.** *Let H be a graph of order N. For any non-negative integer d, $\nu(H) \geq \frac{N-d}{2}$ if and only if $q(H - S) \leq |S| + d$ for every $S \subset V(H)$.*

The value $|S| - |N(S)|$ is known as the *deficiency* of S. For a matching M from X to Y, we will also refer to $|X| - |M|$ as the deficiency of M, where $|M|$ is the number of edges in M. Theorem 2.2.1 therefore states that there exists a matching from X to Y of deficiency at most d if and only if every $S \subset X$ has deficiency at most d.

Note that one can trivially extend one direction of Theorem 2.2.1 to non-integer values as follows.

Corollary 2.2.3. *Let H be a bipartite graph on parts X and Y. For any non-negative real number r, if $\nu(H) < |X| - r$, then there exists a subset $S \subset X$ with $|N(S)| < |S| - r$.*

Proof. If r is an integer, the result follows immediately from Theorem 2.2.1.

If r is not an integer, as $\nu(H) < |X| - r$, we also have $\nu(H) < |X| - \lfloor r \rfloor$. Theorem 2.2.1 then guarantees that there exists a subset $S \subset X$ with $|N(S)| < |S| - \lfloor r \rfloor$. But as $|N(S)|$ is an integer and r is not an integer, this implies

$$|N(S)| \leq |S| - \lfloor r \rfloor - 1 < |S| - r,$$

as required. □

Chen, Yu and Zhao [24, Lemma 1.2] showed that for any integers m, n, N with $N = 4n + m + \lfloor \frac{6n}{m} \rfloor + 1$, and any graph H on N vertices, there is a copy of F_n or K_m inside H or $\overline{H}$. By applying their argument more generally, we obtain the following result, which we will use throughout our proof.

Lemma 2.2.4. *Let H be a graph on $3n - c + 4$ vertices, where $0 < c < \frac{5}{8}n$, such that H does not contain nK_2 and $\overline{H}$ does not contain F_n. Then there is a copy of K_{2n-2c} in H or $\overline{H}$.*

Proof. If $|H| \geq 3n$, then by the result of Lin and Li [58] that $R(nK_2, F_n) = 3n$, there is a copy of nK_2 in H or a copy of F_n in $\overline{H}$. We can therefore assume that $|H| \leq 3n - 1$.

If $\nu(H) \geq n$ then H contains nK_2, so we have $\nu(H) \leq n - 1$. Theorem 2.2.2 with $d = n - c + 4$ now implies that there exists $S \subset V(H)$ with

$$q(H - S) \geq |S| + d + 1 = |S| + n - c + 5.$$

Denote the components of $H - S$ by $C_1, \dots, C_l$, where C_1 has minimal size among the components. If $l \geq 2n - 2c$, then $\overline{H}$ contains a copy of K_{2n-2c}, with one vertex in each C_i. So we may assume that $l \leq 2n - 2c - 1$. We also have that

$$l \geq q(H - S) \geq |S| + n - c + 5.$$

Since C_1 has minimal size, we have that

$$|C_1| \leq \frac{|H| - |S|}{l} \leq \frac{3n - 1}{n - c + 5} < \frac{3n}{n - c} < 8,$$

where the final inequality follows from $c < \frac{5}{8}n$. Writing C for $\bigcup_{i=2}^{l} C_i$, we now have

$$|C| = |H| - |S| - |C_1| \geq 3n - c + 4 - (l - n + c - 5) - 8 = 4n - 2c - l + 1.$$

Since $l \leq 2n - 2c - 1$, this implies that $|C| \geq 2n + 2$.

Now, for every $i, 1 \leq i \leq l$, pick a vertex v_i in C_i, and let $T = \{v_2, \dots, v_l\}$. Writing C' for $C \setminus T$, we claim that C' contains a copy of K_{2n-2c} or a white matching of at least $n - l + 2$ edges. Indeed, suppose that such a matching does not exist. Then removing a white matching of maximal size from C' leaves a black clique consisting of at least $(4n - 2c - l + 1) - (l - 1) - 2(n - l + 1) = 2n - 2c$ vertices, as claimed.

If C' contains a K_{2n-2c} then we are done, so suppose instead that it contains a white matching M of at least $n - l + 2$ edges. We claim that C contains a

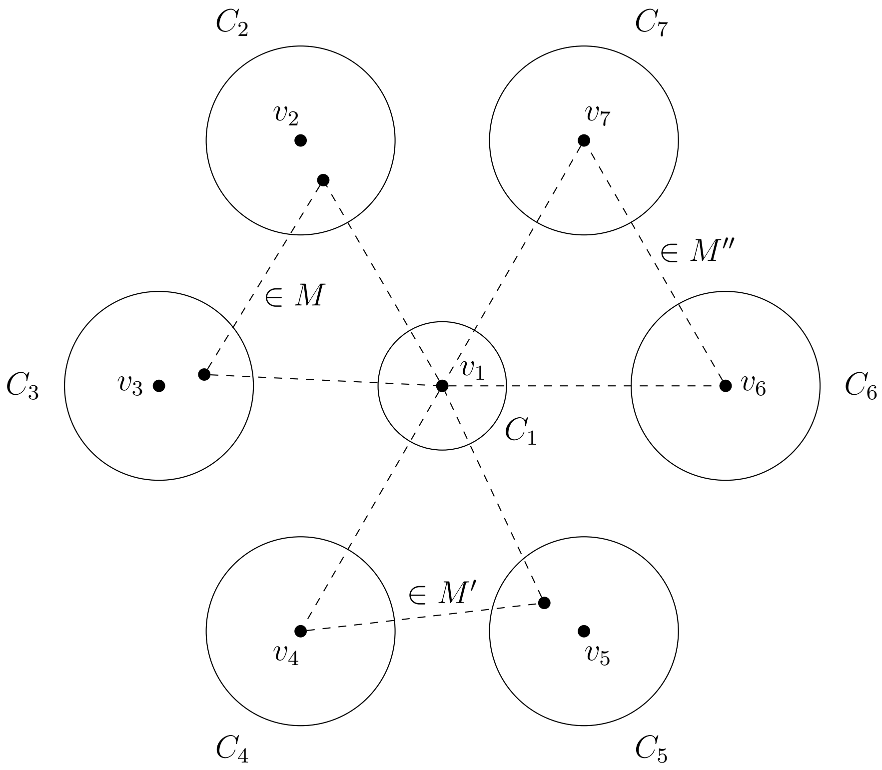


Fig. 2.1: Construction of a white fan with centre v_1

white matching of at least n edges, producing a copy of F_n in $\overline{H}$ with centre v_1, which is a contradiction.

Denote $C' \setminus V(M)$ by X. Every vertex $w \in X$ is joined by white edges to the $l-2$ vertices of T that are not in the same component as w. This allows us to greedily construct a white matching M' between T and X with at least $\min\{l-2, |X|\}$ edges. If $|M'| \geq l-2$, then

$$|M \cup M'| \geq (n-l+2) + (l-2) = n,$$

so $M \cup M'$ is the desired white matching. If instead $|M'| = |X|$, then every vertex of C' is contained in $M \cup M'$. The vertices of T form a white clique, so we can pair up all but at most one of the vertices of $V(T) \setminus V(M')$ into a white matching M''. Combining this matching with $M \cup M'$, we obtain a white matching in C with at least $|C|-1$ vertices. But $|C| \geq 2n+2$, so $M \cup M' \cup M''$ contains at least n edges, as claimed. $\square$

Note that there is nothing special about the bound $c < \frac{5}{8}n$ in the statement of Lemma 2.2.4. For any fixed $\epsilon > 0$, we could prove a version of the lemma for $0 < c < (1-\epsilon)n$ and graphs on $3n-c+k$ vertices, where the additive constant k is dependent on ϵ.

Next we shall prove a simple lemma.

Lemma 2.2.5. *Let H be a graph on $2k$ vertices. Suppose that $V(H)$ is the disjoint union of A and B, each of size k, where $H[A]$ is a clique and $H[B]$ is an empty graph. Then there is a copy of $F_{\lceil \frac{3}{4}k-2 \rceil}$ in H or $\overline{H}$.*

Proof. The role of the colours is interchangeable, so without loss of generality, we have

$$D = \max_{v \in A} |N(v) \cap B| \geq \max_{w \in B} |\overline{N(w)} \cap A|.$$

Moreover, clearly $D \geq \frac{k}{2}$. Let $z \in A$ be such that $|N(z) \cap B| = D$.

If there is a matching of deficiency at most $D - \frac{k}{2}$ from $N(z) \cap B$ to $A \setminus \{z\}$, then H contains a black fan with centre z using at least $D - (D - \frac{k}{2}) = \frac{k}{2}$ vertices of B and at least $k-2$ vertices of $A \setminus \{z\}$ (note that we cannot replace $k-2$ with $k-1$ here because we do not know the parity of k). As this is at least $\lceil \frac{3}{2}k - 2 \rceil$ non-central vertices in total, we know H contains $F_{\lceil \frac{3}{4}k-2 \rceil}$ with centre z.

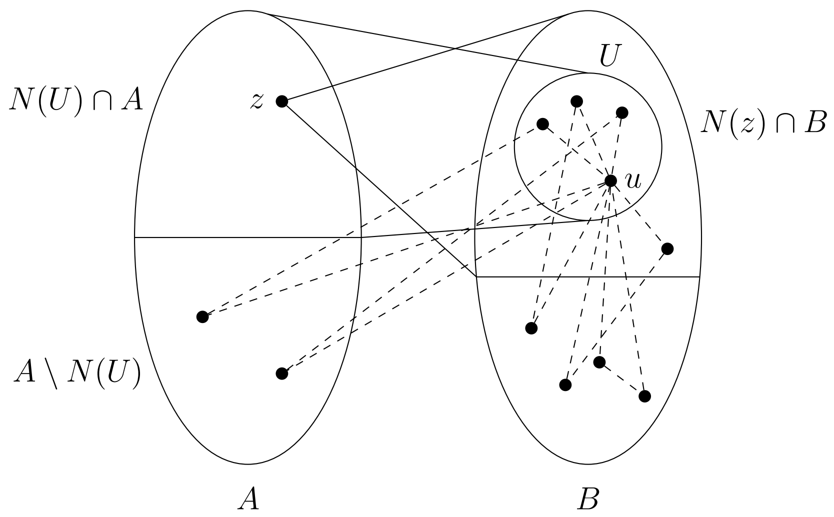


Fig. 2.2: Construction of a white fan with centre u

So assume no such matching exists. Then, by Corollary 2.2.3, there exists $U \subset N(z) \cap B$ with

$$|N(U) \cap A| = |N(U) \cap (A \setminus \{z\})| + 1 < |U| - (D - \tfrac{k}{2}) + 1.$$

Now, as $D = \max_{v \in A} |N(v) \cap B| \geq \max_{w \in B} |\overline{N(w)} \cap A|$, and as U is non-empty, we may pick a vertex $u \in U$, for which we have

$$|N(U) \cap A| \geq |N(u) \cap A| \geq k - D.$$

Hence we get

$$|U| > |N(U) \cap A| + (D - \tfrac{k}{2}) - 1 \geq (k - D) + (D - \tfrac{k}{2}) - 1 \geq \tfrac{k}{2} - 1.$$

Since

$$D \geq |U| > |N(U) \cap A| + D - \tfrac{k}{2} - 1,$$

we also get $|N(U) \cap A| < \frac{k}{2} + 1$, and hence $|A \setminus N(U)| > \frac{k}{2} - 1$. The bounds on the sizes of U and $A \setminus N(U)$, combined with the observation that there are no edges between these two sets, now guarantee a white fan centred at u, with at least $\frac{k}{2} - \frac{3}{2}$ non-central vertices in $A \setminus N(U)$ and at least $k - 2$ non-central vertices in $B \setminus \{u\}$. This fan contains at least $\lceil \frac{3}{2}k - \frac{7}{2} \rceil$ non-central vertices in total, so $\overline{H}$ contains $F_{\lceil \frac{3}{4}k - 2 \rceil}$. □

Suppose that G does not contain a copy of F_n, and suppose that G has a clique A such that $|A| > n$ and every vertex of A has degree more than $2n$ in G. Let v be a vertex of A with degree $d(v)$. We now construct sets $S(v, A)$ and $C(v, A)$, in a slightly more general way than Chen, Yu and Zhao [24].

Let M be a maximal matching in $G[N(v) \setminus A]$, and let M' be a matching of largest size between the independent set $N(v) \setminus (A \cup V(M))$ and $A \setminus \{v\}$. Write m and m' for the number of edges in M and M' respectively. The edges of M and M' form the blades of a fan centred at v, and we can pair up all but at most one of the remaining vertices of $A \setminus \{v\}$ into additional blades. We must therefore have $2m + m' + |A| - 2 \leq 2n - 2$, so $m' \leq 2n - |A| - 2m$.

Note that

$$|N(v) \setminus (A \cup V(M))| = d(v) + 1 - |A| - 2m,$$

so M' has deficiency at least $d(v) + 1 - 2n$. Theorem 2.2.1 now implies that there exists a set $S(v, A) \subset N(v) \setminus (A \cup V(M))$ with

$$|S(v, A)| \geq |N(S(v, A)) \cap (A \setminus \{v\})| + d(v) + 1 - 2n,$$

that is,

$$|S(v, A)| \geq |N(S(v, A)) \cap A| + d(v) - 2n.$$

Moreover, we can insist that $S(v, A)$ has minimal size among all the sets satisfying the inequality above. If there are multiple possible choices of $S(v, A)$, we choose one arbitrarily. Note that since $S(v, A)$ is contained in $N(v) \setminus (A \cup V(M))$ and M is a maximal matching, $S(v, A)$ is an independent set. For convenience, we write $C(v, A) = N(S(v, A)) \cap A$, so we have

$$|S(v, A)| \geq |C(v, A)| + d(v) - 2n.$$

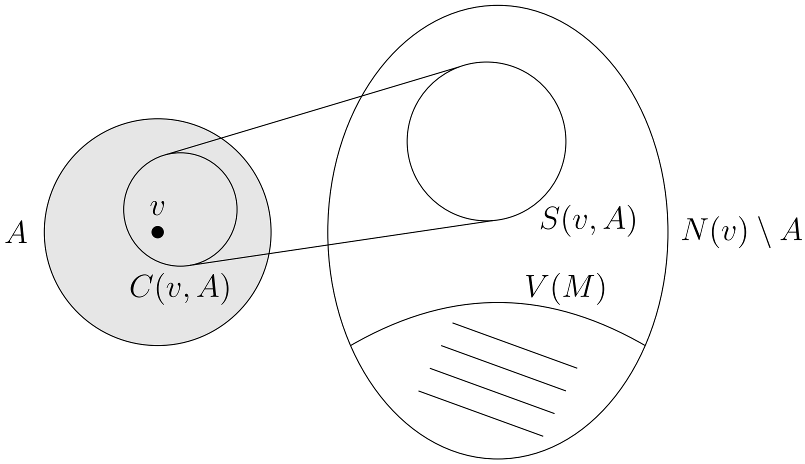


Fig. 2.3: Construction of $S(v, A)$ and $C(v, A)$

We can apply the same argument when A is a white clique. In this case, we consider white edges instead of edges, white degree instead of degree, and so on. We still denote the resulting sets by $S(v, A)$ and $C(v, A)$; it will be clear from the context whether we are working with white or black edges.

Note the following property, which follows directly from the fact that $|S(v, A)| \leq d(v) + 1 - |A|$, combined with the inequality above relating $|S(v, A)|$ and $|C(v, A)|$:

Observation 2.2.6. *We have* $|C(v, A)| \leq 2n + 1 - |A|$.

We need the notion of *coverability* (again analogous to a concept introduced by Chen, Yu and Zhao [24]).

Definition 2.2.7. *Let A be a monochromatic clique such that $n < |A| < 2n+1$. For $t \geq 1$, we say A is t-coverable if t is the smallest integer for which there exists a sequence $v_1, \dots, v_t$ of vertices of A with the following properties:*

- $\bigcup_i C(v_i, A) = A$.
- *For $i = 2, \dots, t$, we have $v_i \notin \bigcup_{j<i} C(v_j, A)$.*
- *For $i = 1, \dots, t$, we have*

$$|C(v_i, A) \setminus \bigcup_{j<i} C(v_j, A)| \geq |C(z, A) \setminus \bigcup_{j<i} C(v_j, A)|$$

for any vertex z of A with $z \notin \bigcup_{j<i} C(v_j, A)$.

So, for example, A is 2-coverable if there exist $v_1, v_2 \in A$ where $|C(v_1, A)|$ is maximal over all $|C(v, A)|$, and v_2 is such that $v_2 \notin C(v_1, A)$ and $C(v_1, A) \cup C(v_2, A) = A$. It is clear from the definition that t must exist.

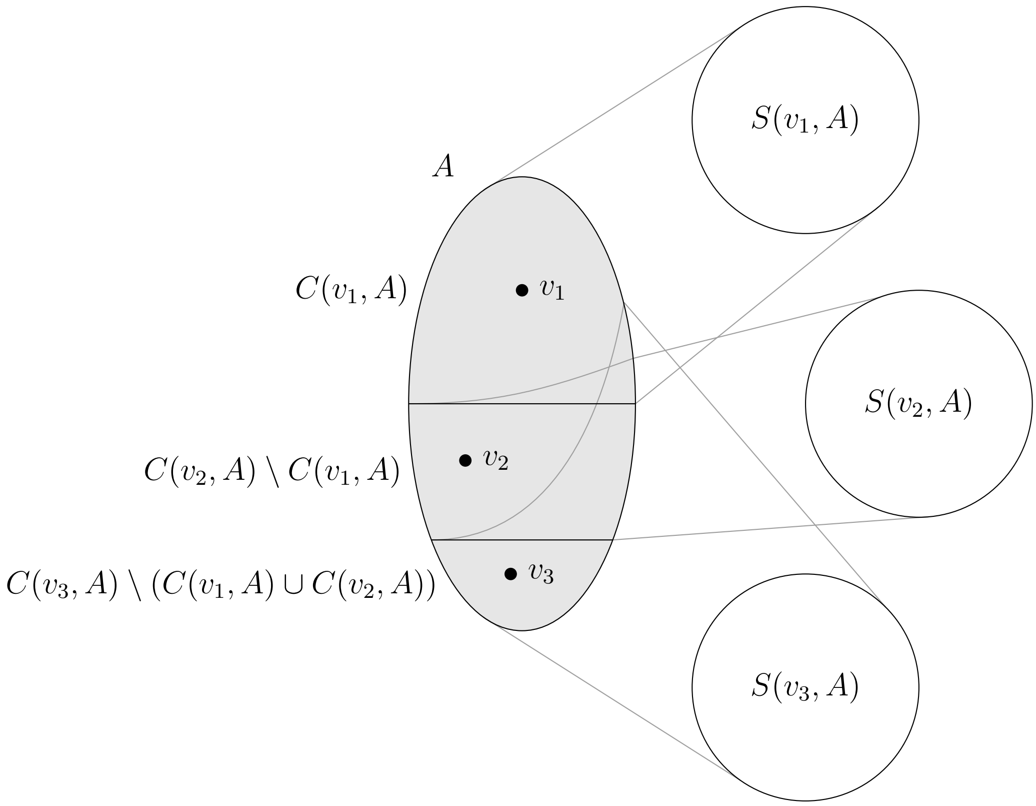


Fig. 2.4: Example where A is 3-coverable

Note the following simple properties.

Observation 2.2.8. *We have that:*

- *For any $j_1 < j_2$,*

$$|C(v_{j_1}, A) \setminus \bigcup_{i<j_1} C(v_i, A)| \geq |C(v_{j_2}, A) \setminus \bigcup_{i<j_1} C(v_i, A)|.$$

- *For any $j_1 \neq j_2$, the sets $S(v_{j_1}, A)$ and $S(v_{j_2}, A)$ are disjoint.*
- *If $|A| > \frac{k-1}{k}(2n+1)$ and A is t-coverable, then $t \geq k$.*

The final point above follows from Observation 2.2.6, and implies that no clique satisfying the conditions of Definition 2.2.7 is 1-coverable.

2.3 Overview of the rest of the proof

The rest of the proof is quite technical, so we first summarise the general strategy. There are five cases.

Call a monochromatic clique A *big* if $|A| \geq \frac{7}{6}n + 5$ and call it *significant* if $|A| \geq n + 1$.

In subsections 2.4.1 and 2.4.2, we handle the easier cases when either some vertex has very unbalanced degrees (that is, a much larger degree in one colour than the other) or a significant clique of either colour is t-coverable for some $t \geq 4$. Lemma 2.2.4 and the strategy of Chen, Yu and Zhao [24] suffice to tackle these cases.

The next three cases, where no vertices have very unbalanced degrees and all significant cliques are 2-coverable or 3-coverable, form the heart of the proof.

In subsection 2.4.3, there is still a vertex with slightly unbalanced degrees, forcing the existence of a very large (and in particular big) 3-coverable clique, and in subsection 2.4.4, the degrees are balanced but we assume there is some big 3-coverable clique. The proofs of these cases follow a very similar argument. Both times, we start with the clique A (black without loss of generality) and its 3-covering v_1, v_2, v_3. We then argue that there must be a large black clique T disjoint from A in $N_W(v_3)$ which satisfies certain properties: otherwise, we would find a white F_n centred at v_3. Then we take any $z \in T$ and argue that $S(v_1, A) \cup S(v_2, A)$ must contain a large white clique C with at least one vertex in each of $S(v_1, A)$ and $S(v_2, A)$; else we would find a black F_n centred at z. Finally, we conclude that there must be a white F_n centred at some $a \in C$.

In subsection 2.4.5, we consider the final case where all vertices have balanced degrees and every big clique is 2-coverable. We start by using Lemma 2.2.4 to find two significant cliques A and B of the same colour, without loss of generality black, with A moreover being big. We consider a 2-covering v_1, v_2 of A and a 2- or 3-covering $\{w_i\}$ of B. We then show there must exist i such that $S(v_1, A)$ and $S(w_i, B)$ intersect. Finally we fix some a in this intersection and find a white F_n centred at it.

2.4 Proof of Theorem 2.1.1

Now we prove Theorem 2.1.1. Let G be a graph with at least $\left\lfloor \frac{31}{6}n + 15 \right\rfloor$ vertices, and suppose for contradiction that neither G nor $\overline{G}$ contains a copy of F_n. Throughout, denote

$$d = \max\left\{\max_v |N(v)|, \max_w |N_W(w)|\right\}.$$

That is, d is the larger one of the maximum degree and non-degree in our graph G.

As discussed in Section 2.3, we consider five separate cases.

2.4.1 $d \geq \frac{11}{4}n + 5$

If $d > 3n$, contradiction follows immediately from the result of Lin and Li [58] that $R(nK_2, F_n) = 3n$: consider the neighbourhood of a vertex of degree d in some colour.

If $\frac{11}{4}n + 5 \leq d \leq 3n$, by applying Lemma 2.2.4 to the neighbourhood of a vertex of degree d in some colour, we find that our graph contains a monochromatic clique A with size at least $2d - 4n - 8$. We assume without loss of generality that A is a black clique. Since $2d - 4n - 8 > \frac{3}{2}n + 1$, we know by Observation 2.2.8 that this clique is t-coverable for some $t \geq 4$.

Now v_t is the centre of a white fan with blades in the sets $S(v_i, A)$ for $i = 1, \dots, t-1$. The fan contains all but at most one vertex in each $S(v_i, A)$, so the total number of vertices in the fan is is at least

$$|S(v_1, A)| + \ldots + |S(v_{t-1}, A)| - (t-1) + 1.$$

Since for $i = 1, \dots, t-1$, we have

$$\begin{aligned} |S(v_i, A)| &\geq |C(v_i, A)| + d(v_i) - 2n \\ &\geq |C(v_i, A)| + (\tfrac{31}{6}n + 13 - d) - 2n \\ &\geq |C(v_i, A)| + \tfrac{19}{6}n - d + 13, \end{aligned}$$

and $\sum_{i=1}^{t-1} |C(v_i, A)| \geq \frac{t-1}{t}|A| \geq \frac{3}{4}|A|$ by Observation 2.2.8, the number of vertices in this fan is at least $\frac{3}{4}|A| + (t-1)(\frac{19}{6}n - d + 12) + 1$.

By our earlier observations that $|A| \geq 2d - 4n - 8$ and $d \leq 3n$, it follows that

$$\tfrac{3}{4}|A| + (t-1)(\tfrac{19}{6}n - d + 12) + 1 \geq \tfrac{13}{2}n - \tfrac{3}{2}d + 31 \geq 2n + 31.$$

Hence the fan has more than n blades and contradiction follows.

2.4.2 $d < \frac{11}{4}n + 5$ and some significant clique is t-coverable for some $t \geq 4$

Call this significant clique A, and recall $|A| \geq n + 1$. We assume without loss of generality that A is a black clique.

Again, v_t is the centre of a white fan containing all but at most one vertex in each $S(v_i, A)$ for $i = 1, \ldots, t-1$. The number of vertices in the fan is therefore at least

$$|S(v_1, A)| + \ldots + |S(v_{t-1}, A)| - (t-1) + 1.$$

Since

$$\begin{aligned} |S(v_i, A)| &\geq |C(v_i, A)| + d(v_i) - 2n \\ &\geq |C(v_i, A)| + (\tfrac{31}{6}n + 13 - d) - 2n \\ &\geq |C(v_i, A)| + \tfrac{5}{12}n + 8 \end{aligned}$$

for $i = 1, \ldots, t-1$, and $\sum_{i=1}^{t-1} |C(v_i, A)| \geq \frac{t-1}{t}|A| \geq \frac{3}{4}|A|$ by Observation 2.2.8, the number of vertices in this fan is at least

$$\tfrac{3}{4}|A| + (t-1)(\tfrac{5}{12}n + 7) + 1 \geq 2n + 22.$$

The contradiction follows.

2.4.3 $\frac{8}{3}n + 6 \leq d < \frac{11}{4}n + 5$ and every significant clique is 2- or 3-coverable

By applying Lemma 2.2.4 to the neighbourhood of a vertex of degree d in some colour, there exists a monochromatic clique A such that $|A| \geq \frac{4}{3}n + 4$, which is black without loss of generality. By Observation 2.2.8, A is not 2-coverable, so as it is significant it must be 3-coverable. Let v_1, v_2, v_3 be its 3-covering. Note also that Observation 2.2.8 tells us that $|A| < \frac{3}{2}n + 1$.

Claim 2.4.1. *The degrees of v_1, v_2, v_3 are all at least $\frac{5}{2}n + 5$.*

Proof. Assume not and suppose that some v_i has degree less than $\frac{5}{2}n + 5$. Then it has white degree at least $\frac{8}{3}n + 8$. Now by Lemma 2.2.4, $N_W(v_i)$ contains a clique of size at least $\frac{4}{3}n + 8$ in some colour; call this clique B. But if B is white, then by Lemma 2.2.5 applied to A and B, there is a copy of F_n in G or $\overline{G}$, a contradiction. Hence the clique is black. But now A and B are disjoint, and each of the white cliques $S(v_1, A)$, $S(v_2, A)$, $S(v_3, A)$ is disjoint from A and contains at most one vertex from B. Moreover $S(v_1, A)$, $S(v_2, A)$, $S(v_3, A)$

are mutually disjoint sets too by Observation 2.2.8. So we have

$$\begin{aligned}|G| &\geq |A| + |B| + \sum_i |S(v_i, A)| - 3\\ &\geq |A| + |B| + \sum_i (|C(v_i, A)| + d(v_i) - 2n) - 3\\ &\geq |A| + |B| + \sum_i |C(v_i, A)| + 3(\tfrac{31}{6}n + 13 - d - 2n) - 3\\ &\geq |A| + |B| + |A| + 3(\tfrac{5}{12}n + 7) - 3\\ &> \tfrac{31}{6}n + 15,\end{aligned}$$

which is a contradiction. □

So we in fact have $|S(v_i, A)| \geq |C(v_i, A)| + \frac{1}{2}n + 5$. Next we show two simple results that will be useful later.

Claim 2.4.2. *We have* $\frac{1}{3}|A| \leq |C(v_1, A)| \leq \frac{2}{3}n$ *and* $|C(v_2, A)| \geq \frac{1}{2}|A \setminus C(v_1, A)| \geq \frac{1}{3}n$.

Proof. This follows immediately from Observations 2.2.6 and 2.2.8. □

Claim 2.4.3. *We have* $|S(v_1, A)| + |S(v_2, A)| > \frac{17}{9}n + 10$.

Proof. Using Observation 2.2.8 and Claim 2.4.1, we have

$$\begin{aligned}|S(v_1, A)| + |S(v_2, A)| &\geq |C(v_1, A)| + |C(v_2, A)| + 2(\tfrac{1}{2}n + 5)\\ &\geq \tfrac{2}{3}|A| + n + 10 > \tfrac{17}{9}n + 10\end{aligned}$$

as required. □

Now we get to the heart of the argument.

Claim 2.4.4. *There exists a black clique* T *in* $N_W(v_3) \setminus (S(v_1, A) \cup S(v_2, A))$ *such that* $|T| > |N_T| + \frac{5}{12}n + 6$, *where* $N_T = N_W(T) \cap (S(v_1, A) \cup S(v_2, A))$.

Proof. Set $T' = N_W(v_3) \setminus (S(v_1, A) \cup S(v_2, A))$. We form a white fan centred at v_3, and we show that if there is no black clique $T \subset T'$ such that $|T| > |N_T| + \frac{5}{12}n + 6$ then this fan has at least n blades, which is a contradiction. Let M be a maximal white matching within T', and add blades consisting of the edges of M. Next, take a maximal white matching M' from the black clique $T' \setminus V(M)$ to $S(v_1, A) \cup S(v_2, A)$ and add blades consisting of M'. Finally, add all but at most one of the remaining vertices of $S(v_i, A)$ for $i = 1, 2$ by pairing them up together within each set.

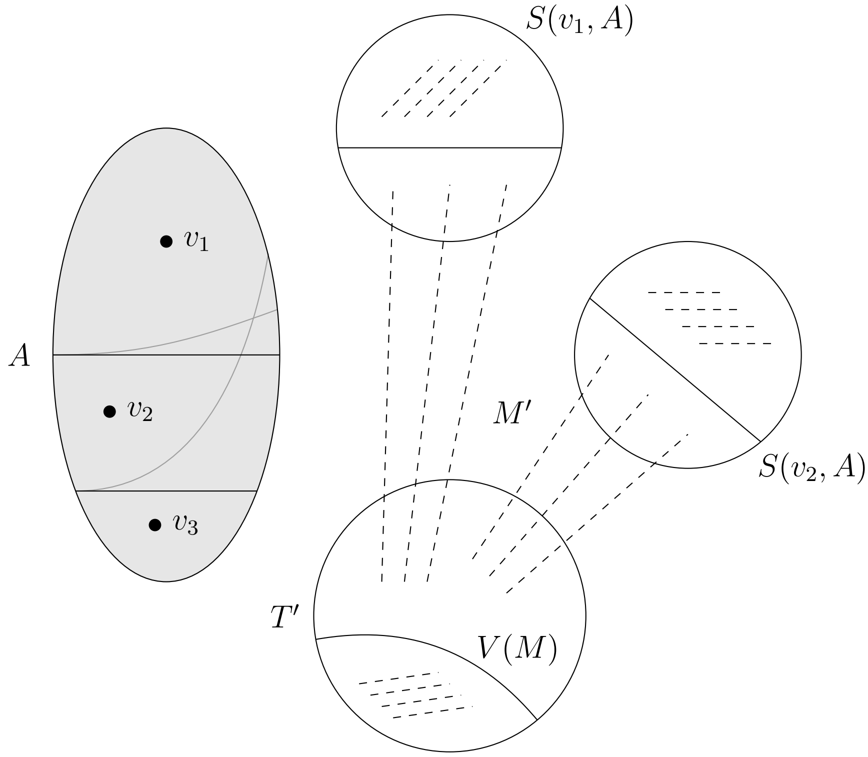


Fig. 2.5: Construction of a white fan with centre v_3 in Claim 2.4.4. Only the blades of the fan are shown.

Note that we have $|N_W(v_3)| \geq \frac{31}{6}n + 13 - (\frac{11}{4}n + 5) = \frac{29}{12}n + 8$. The blades of our fan contain all of the vertices of $N_W(v_3)$ except for $T' \setminus (V(M) \cup V(M'))$ and at most two vertices of $S(v_1, A) \cup S(v_2, A)$. But $|T' \setminus (V(M) \cup V(M'))|$ is the deficiency of M', and this is at most $\frac{5}{12}n + 6$ by Theorem 2.2.1 and our assumption that there is no black clique $T \subset T'$ with $|T| > |N_T| + \frac{5}{12}n + 6$. This implies that the blades of the fan contain at least $\frac{29}{12}n + 8 - (\frac{5}{12}n + 6) - 2 = 2n$ vertices, producing the desired contradiction. □

Now denote by C the largest white clique that can be obtained as follows: start with $S(v_1, A) \cup S(v_2, A)$, remove a set U consisting of $|N_T|$ arbitrary vertices, and then remove a maximal black matching between $S(v_1, A) \setminus U$ and $S(v_2, A) \setminus U$.

Claim 2.4.5. *We have* $|C| \geq |S(v_1, A)| + |S(v_2, A)| - |N_T| - 2n + 2|T| - 6$.

Proof. Assume that instead $|C| < |S(v_1, A)| + |S(v_2, A)| - |N_T| - 2n + 2|T| - 6$. Pick any $z \in T$, and form a black F_n centred at z as follows. First note that a maximal matching between $S(v_1, A) \setminus N_T$ and $S(v_2, A) \setminus N_T$ must contain at least $n + 3 - |T|$ edges, else the remaining vertices of $S(v_1, A) \cup S(v_2, A)$ form a white clique C' larger than C which also satisfies our assumptions. We can therefore take a matching M between $S(v_1, A) \setminus N_T$ and $S(v_2, A) \setminus N_T$ with

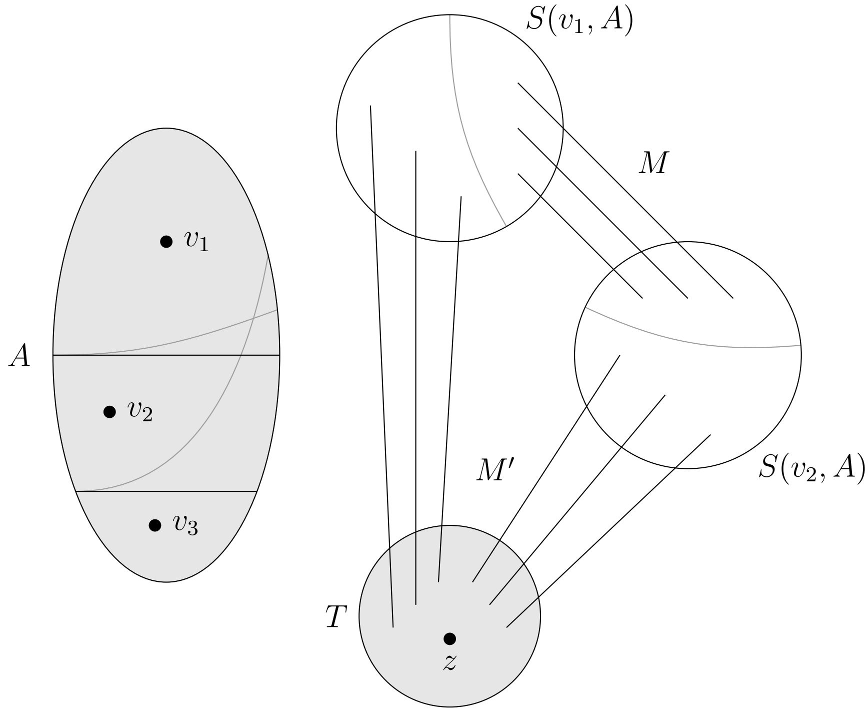


Fig. 2.6: Construction of a black fan with centre z. Only the blades of the fan are shown.

exactly $n+3-|T|$ edges and add it to our fan. After this, add a maximal matching M' between T and $(S(v_1,A) \cup S(v_2,A)) \setminus (N_T \cup V(M))$. If M' consists of fewer than $|T|-3$ edges, then we have

$$|(S(v_1,A) \cup S(v_2,A)) \setminus (N_T \cup V(M))| \leq |T| - 4,$$

and therefore, by Claim 2.4.4,

$$\begin{aligned} |S(v_1,A)| + |S(v_2,A)| &\leq |T| - 4 + |N_T| + 2(n+3-|T|) \\ &= 2n + 2 - (|T| - |N_T|) \\ &< 2n + 2 - (\tfrac{5}{12}n + 6) = \tfrac{19}{12}n - 4. \end{aligned}$$

But this is impossible, since by Claim 2.4.3 we know that $|S(v_1,A)| + |S(v_2,A)| > \frac{17}{9}n + 10$. So M' has at least $|T|-3$ edges, and we have found a fan F_n, which is a contradiction. $\square$

Consider the vertices erased from $S(v_i,A)$ for $i=1,2$ when we obtain C. We erase all vertices of $U \cap S(v_i,A)$, of which there are at most $|U| = |N_T|$, together with half the vertices of the maximal matching between $S(v_1,A) \setminus U$ and $S(v_2,A) \setminus U$. By Claim 2.4.5, we erase at most $|N_T| + 2n + 6 - 2|T|$ vertices from $S(v_1,A) \cup S(v_2,A)$, so the maximal matching has at most $2n+6-2|T|$

vertices. We therefore erase at most $|N_T| + n + 3 - |T|$ vertices from each of $S(v_1, A)$ and $S(v_2, A)$.

Recall that Claim 2.4.1 implies that $|S(v_i, A)| \geq |C(v_i, A)| + \frac{1}{2}n + 5$. Combined with Claims 2.4.2 and 2.4.4, this gives

$$|S(v_1, A)|, |S(v_2, A)| \geq \tfrac{5}{6}n + 5 > \tfrac{7}{12}n + 5 > |N_T| + n + 3 - |T|.$$

The white clique C therefore contains a vertex $a_1 \in S(v_1, A)$ and a vertex $a_2 \in S(v_2, A)$.

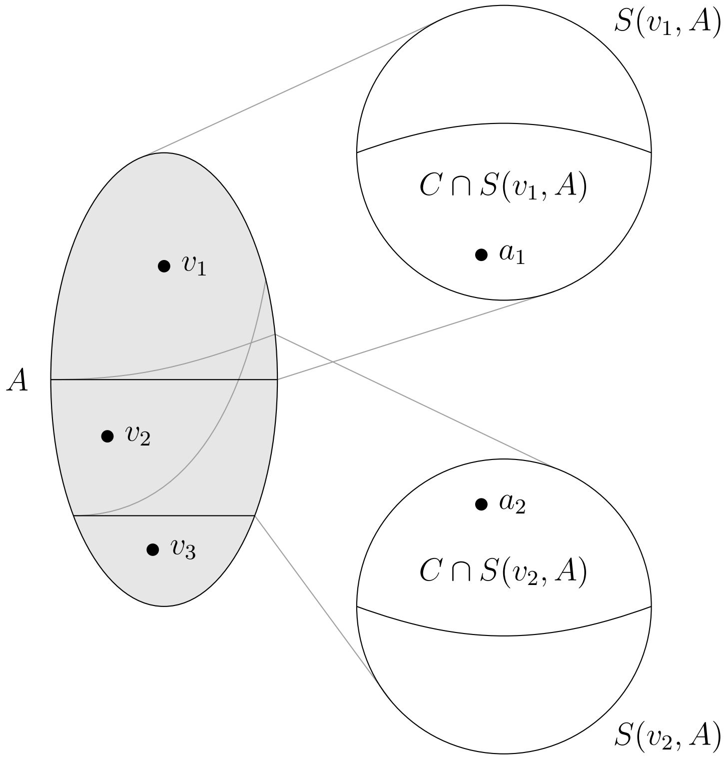


Fig. 2.7: Set-up for construction of a white fan centred at a_1 or a_2

Note that by Claims 2.4.3, 2.4.4 and 2.4.5, we have

$$\begin{aligned}|C| &\geq |S(v_1, A)| + |S(v_2, A)| - |N_T| - 2n + 2|T| - 6\\ &> \tfrac{17}{9}n + 10 + 2(|T| - |N_T|) - 2n - 6\\ &> \tfrac{13}{18}n + 16.\end{aligned}$$

Consequently, we must have $|C \cap S(v_2, A)| > \frac{1}{6}n$ or $|C \cap S(v_1, A)| > \frac{5}{9}n + 16$. We treat these cases separately.

First, assume that $|C \cap S(v_2, A)| > \frac{1}{6}n$. We will construct a white fan centred at a_1 and show that it has at least n blades. We begin by claiming that

$$|S(v_1, A)| > |A \setminus C(v_1, A)| = |A| - |C(v_1, A)|.$$

Indeed this holds, since by Claim 2.4.2 we have

$$\begin{aligned}|S(v_1, A)| - |A| + |C(v_1, A)| &= (|S(v_1, A)| - |C(v_1, A)|) + 2|C(v_1, A)| - |A| \\ &\geq \tfrac{1}{2}n + 5 + \tfrac{2}{3}|A| - |A| \\ &> \tfrac{1}{2}n + 5 - \tfrac{1}{3}(\tfrac{3}{2}n + 1) > 0.\end{aligned}$$

So up to at most two vertices, we can use all the vertices of $S(v_1, A), A \setminus C(v_1, A)$ and $C \cap S(v_2, A)$ in our fan, by first taking blades with a vertex in $S(v_1, A)$ and the other in $A \setminus C(v_1, A)$, and then pairing up all but at most one of the remaining vertices in $S(v_1, A)$ and all but at most one of the vertices in $C \cap S(v_2, A)$. But now the number of vertices in the fan is at least

$$\begin{aligned}&|S(v_1, A)| + |A \setminus C(v_1, A)| + |C \cap S(v_2, A)| - 2 \\ &\quad = |A| + (|S(v_1, A)| - |C(v_1, A)|) + |C \cap S(v_2, A)| - 2 \\ &\quad > \tfrac{4}{3}n + 4 + \tfrac{1}{2}n + 5 + \tfrac{1}{6}n - 2 = 2n + 7,\end{aligned}$$

so our fan has at least n blades, which is a contradiction.

Next consider the case $|C \cap S(v_2, A)| \leq \frac{1}{6}n$, $|C \cap S(v_1, A)| > \frac{5}{9}n + 16$. If $|S(v_2, A)| \geq |A \setminus C(v_2, A)|$, we can finish the argument as above and obtain a fan with at least n blades, now with a_2 as the centre: the number of vertices in the fan is at least

$$\begin{aligned}&|S(v_2, A)| + |A \setminus C(v_2, A)| + |C \cap S(v_1, A)| - 2 \\ &\quad = |A| + (|S(v_2, A)| - |C(v_2, A)|) + |C \cap S(v_1, A)| - 2 \\ &\quad > \tfrac{4}{3}n + 4 + \tfrac{1}{2}n + 5 + \tfrac{5}{9}n + 16 - 2 > \tfrac{43}{18}n.\end{aligned}$$

So assume $|S(v_2, A)| < |A \setminus C(v_2, A)|$. We construct a white fan centred at a_2 and show that it has at least n blades. First add $|S(v_2, A)| - 1$ blades with one vertex in $S(v_2, A)$ and one in $A \setminus C(v_2, A)$, and then pair up all but at most one vertex of $C \cap S(v_1, A)$. Using Claim 2.4.2, we find that the number of vertices in the fan is at least

$$\begin{aligned}2(|S(v_2, A)| - 1) + |C \cap S(v_1, A)| - 1 &> 2(|C(v_2, A)| + \tfrac{1}{2}n + 4) + \tfrac{5}{9}n + 15 \\ &\geq \tfrac{5}{3}n + 8 + \tfrac{5}{9}n + 15 > \tfrac{20}{9}n,\end{aligned}$$

which is a contradiction. Thus we have shown that if $\frac{8}{3}n+6 \le d < \frac{11}{4}n+5$ and every significant clique is 2- or 3-coverable, then G contains a monochromatic F_n.

2.4.4 $d < \frac{8}{3}n + 6$ and there is a 3-coverable big clique

By assumption, there exists a monochromatic (without loss of generality black) clique A such that $\frac{7}{6}n + 5 \le |A| < \frac{3}{2}n + 1$ and A is 3-coverable, with 3-covering v_1, v_2, v_3. As before, the upper bound on $|A|$ comes from Observation 2.2.8.

Note that v_1, v_2, v_3 all have black degree at least $\frac{5}{2}n + 7$. So we have $|S(v_i, A)| \ge |C(v_i, A)| + \frac{1}{2}n + 7$.

As in the previous subsection, we begin by proving some simple results.

Claim 2.4.6. *The following inequalities hold:*

- $\frac{1}{3}|A| \le |C(v_1, A)| \le \frac{5}{6}n$.
- $|C(v_2, A)| \ge \frac{1}{2}|A \setminus C(v_1, A)| \ge \frac{1}{6}n$.
- $|C(v_3, A) \setminus (C(v_1, A) \cup C(v_2, A))| \ge \frac{1}{6}n$.

Proof. The first two results follow immediately by Observations 2.2.6 and 2.2.8. If we did not have $|C(v_3, A) \setminus (C(v_1, A) \cup C(v_2, A))| \ge \frac{1}{6}n$, we would have

$$\begin{aligned} |S(v_1, A)| + |S(v_2, A)| &\ge |C(v_1, A)| + |C(v_2, A)| + 2(\tfrac{1}{2}n + 7) \\ &\ge n + 4 + n + 14 > 2n + 2 \end{aligned}$$

and hence we would have a white F_n centred at v_3 with blades inside $S(v_1, A)$ and $S(v_2, A)$. But that is a contradiction. □

Claim 2.4.7. $|S(v_1, A)| + |S(v_2, A)| > \frac{16}{9}n + 16$.

Proof. Using Observation 2.2.8, we have

$$\begin{aligned} |S(v_1, A)| + |S(v_2, A)| &\ge |C(v_1, A)| + |C(v_2, A)| + 2(\tfrac{1}{2}n + 7) \\ &\ge \tfrac{2}{3}|A| + n + 14 > \tfrac{16}{9}n + 16 \end{aligned}$$

as required. □

Now we get to the key parts of the argument. The next two claims are analogous to Claims 2.4.4 and 2.4.5.

Claim 2.4.8. *There exists a black clique T in $N_W(v_3) \setminus (S(v_1, A) \cup S(v_2, A))$ such that $|T| > |N_T| + \frac{1}{2}n + 5$, where $N_T = N_W(T) \cap (S(v_1, A) \cup S(v_2, A))$.*

Proof. We proceed as in the proof of Claim 2.4.4. Let $T' = N_W(v_3) \setminus (S(v_1, A) \cup S(v_2, A))$, and assume that there is no black clique $T \subset T'$ with $|T| > |N_T| + \frac{1}{2}n + 5$. We form a white fan centred at v_3 and show that it has at least n blades, giving a contradiction. Begin by adding blades consisting of the edges of a maximal white matching M within T'. Next, add the edges of a maximal white matching M' from the black clique $T' \setminus V(M)$ to $S(v_1, A) \cup S(v_2, A)$, and finally add all but one of the vertices within each of $S(v_1, A)$ and $S(v_2, A)$ by pairing them up within each set.

We have that $|N_W(v_3)| \geq \frac{31}{6}n + 13 - (\frac{8}{3}n + 6) = \frac{5}{2}n + 7$. The blades of our fan contain all the vertices of $N_W(v_3)$ except for $T' \setminus (V(M) \cup V(M'))$ and at most two vertices of $S(v_1, A) \cup S(v_2, A)$. Since there is no black clique $T \subset T'$ with $|T| > |N_T| + \frac{1}{2}n + 5$, Theorem 2.2.1 implies that M' has deficiency at most $\frac{1}{2}n + 5$. But the deficiency of M' is $|T' \setminus (V(M) \cup V(M'))|$, so the blades of the fan contain at least $\frac{5}{2}n + 7 - (\frac{1}{2}n + 5) - 2 = 2n$ vertices, a contradiction. □

Now denote by C the largest white clique that can be obtained as follows. Start with $S(v_1, A) \cup S(v_2, A)$. Then remove a set U consisting of $|N_T|$ vertices. Finally, remove a maximal black matching between $S(v_1, A) \setminus U$ and $S(v_2, A) \setminus U$.

Claim 2.4.9. *We have* $|C| \geq |S(v_1, A)| + |S(v_2, A)| - |N_T| - 2n + 2|T| - 6$.

Proof. As in the proof of Claim 2.4.5, we suppose that $|C| < |S(v_1, A)| + |S(v_2, A)| - |N_T| - 2n + 2|T| - 6$ and construct a black fan centred at an arbitrary $z \in T$. We first take a matching from $S(v_1, A) \setminus N_T$ to $S(v_2, A) \setminus N_T$ with exactly $n + 3 - |T|$ edges, which must exist since otherwise $S(v_1, A) \cup S(v_2, A)$ contains a white clique larger than C satisfying the same conditions as C. We then add a maximal matching M' between T and $(S(v_1, A) \cup S(v_2, A)) \setminus (N_T \cup V(M))$. If M' has fewer than $|T| - 3$ edges, then we have

$$|(S(v_1, A) \cup S(v_2, A)) \setminus (N_T \cup V(M))| \leq |T| - 4,$$

and now Claim 2.4.8 implies that

$$\begin{aligned}
|S(v_1, A)| + |S(v_2, A)| &\leq |T| - 4 + |N_T| + 2(n + 3 - |T|) \\
&= 2n + 2 - (|T| - |N_T|) \\
&< \tfrac{3}{2}n - 3.
\end{aligned}$$

But Claim 2.4.7 tells us that $|S(v_1, A)| + |S(v_2, A)| > \frac{16}{9}n + 16$, so M' must have at least $|T| - 3$ edges, and we have a black F_n centred at z, which is a contradiction. □

Recall that in the previous subsection, Claim 2.4.5 implied that when obtaining C we erased at most $|N_T| + n - 3 - |T|$ vertices from each of $S(v_1, A)$ and $S(v_2, A)$. The same is true in this case, using Claim 2.4.9 in place of Claim 2.4.5. The argument is identical word-for-word, so we will not repeat it here.

Recall that $|S(v_i, A)| \geq |C(v_i, A)| + \frac{1}{2}n + 7$. Together with Claims 2.4.6 and 2.4.8, this implies that

$$|S(v_1, A)|, |S(v_2, A)| \geq \tfrac{2}{3}n + 7 > \tfrac{1}{2}n > |N_T| + n + 3 - |T|.$$

The white clique C therefore contains a vertex $a_1 \in S(v_1, A)$ and a vertex $a_2 \in S(v_2, A)$.

Note that by Claims 2.4.7, 2.4.8 and 2.4.9, we have

$$\begin{aligned}
|C| &\geq |S(v_1, A)| + |S(v_2, A)| - |N_T| - 2n + 2|T| - 6 \\
&> \tfrac{16}{9}n + 16 + 2(|T| - |N_T|) - 2n - 6 \\
&> \tfrac{7}{9}n + 20.
\end{aligned}$$

So we must have $|C \cap S(v_2, A)| > \frac{1}{3}n$ or $|C \cap S(v_1, A)| > \frac{4}{9}n + 20$. We treat the two cases separately.

First, assume $|C \cap S(v_2, A)| > \frac{1}{3}n$. We will construct a white fan centred at a_1 and show that it has at least n blades. We claim that

$$|S(v_1, A)| > |A \setminus C(v_1, A)| = |A| - |C(v_1, A)|.$$

Indeed this holds, since by Claim 2.4.6 we have

$$\begin{aligned}
|S(v_1, A)| - |A| + |C(v_1, A)| &= (|S(v_1, A)| - |C(v_1, A)|) + 2|C(v_1, A)| - |A| \\
&\geq \tfrac{1}{2}n + 7 + \tfrac{2}{3}|A| - |A| \\
&> \tfrac{1}{2}n + 7 - \tfrac{1}{3}(\tfrac{3}{2}n + 1) > 0.
\end{aligned}$$

So up to at most two vertices, we can use all the vertices of $S(v_1, A)$, $A \setminus C(v_1, A)$ and $C \cap S(v_2, A)$ in our fan, by first taking blades with one vertex in $S(v_1, A)$ and the other in $A \setminus C(v_1, A)$, and then pairing up all but at most one of the remaining vertices in $S(v_1, A)$ and all but at most one of the vertices in $C \cap S(v_2, A)$. But

$$\begin{aligned}
&|S(v_1, A)| + |A \setminus C(v_1, A)| + |C \cap S(v_2, A)| - 2 \\
&\qquad = |A| + (|S(v_1, A)| - |C(v_1, A)|) + |C \cap S(v_2, A)| - 2 \\
&\qquad > \tfrac{7}{6}n + 5 + \tfrac{1}{2}n + 7 + \tfrac{1}{3}n - 2 = 2n + 10,
\end{aligned}$$

and so our fan has at least n blades, which is a contradiction.

Next consider the case $|C \cap S(v_1, A)| > \frac{4}{9}n + 20$ and $|C \cap S(v_2, A)| \leq \frac{1}{3}n$. Here we consider two subcases.

If $|C(v_2, A)| \geq \frac{5}{18}n$, we construct a white fan centred at a_2 and show that it has at least n blades. First form as many blades as possible with one vertex in $S(v_2, A)$ and the other in $A \setminus C(v_2, A)$, and then pair up all but at most one of the vertices in $C \cap S(v_1, A)$. If $|S(v_2, A)| > |A \setminus C(v_2, A)|$, then we can also pair up all but at most one of the remaining vertices in $S(v_2, A)$, and we get a contradiction as in the previous case, but with v_1 and v_2 interchanged: the number of vertices in the fan is at least

$$
\begin{aligned}
&|S(v_2, A)| + |A \setminus C(v_2, A)| + |C \cap S(v_1, A)| - 1 \\
&\qquad = |A| + (|S(v_2, A)| - |C(v_2, A)|) + |C \cap S(v_1, A)| - 2 \\
&\qquad > \tfrac{7}{6}n + 5 + \tfrac{1}{2}n + 7 + \tfrac{4}{9}n + 20 - 2 > \tfrac{19}{9}n.
\end{aligned}
$$

If instead $|S(v_2, A)| \leq |A \setminus C(v_2, A)|$, then our white fan has at least

$$
\begin{aligned}
&2(|S(v_2, A)| - 1) + |C \cap S(v_1, A)| - 1 \\
&\qquad \geq 2|C(v_2, A)| + 2(\tfrac{1}{2}n + 7) - 2 + \tfrac{4}{9}n + 19 \\
&\qquad \geq \tfrac{5}{9}n + n + 14 + \tfrac{4}{9}n + 17 = 2n + 31
\end{aligned}
$$

vertices, so once again it has at least n blades and we reach a contradiction.

If $|C(v_2, A)| < \frac{5}{18}n$, it follows by Observation 2.2.8 that

$$|C(v_3, A) \setminus (C(v_1, A) \cup C(v_2, A))| < \tfrac{5}{18}n.$$

Hence $|C(v_1, A) \cup C(v_2, A)| \geq \frac{8}{9}n$, and then Claims 2.4.8 and 2.4.9 give

$$
\begin{aligned}
|C| &\geq |S(v_1, A)| + |S(v_2, A)| - |N_T| - 2n + 2|T| - 6 \\
&\geq |C(v_1, A)| + |C(v_2, A)| + 2(\tfrac{1}{2}n + 7) + 2(|T| - |N_T|) - 2n - 6 \\
&\geq \tfrac{8}{9}n + n + 14 + 2(\tfrac{1}{2}n + 5) - 2n - 6 = \tfrac{8}{9}n + 18.
\end{aligned}
$$

So as $|C \cap S(v_2, A)| \leq \frac{1}{3}n$, we have $|C \cap S(v_1, A)| \geq \frac{5}{9}n + 18$.

Now we form a white fan and show that it has at least n blades. Once again, we pick a_2 as the centre. Recall that by Claim 2.4.6,

$$|C(v_3, A) \setminus (C(v_1, A) \cup C(v_2, A))| \geq \tfrac{1}{6}n.$$

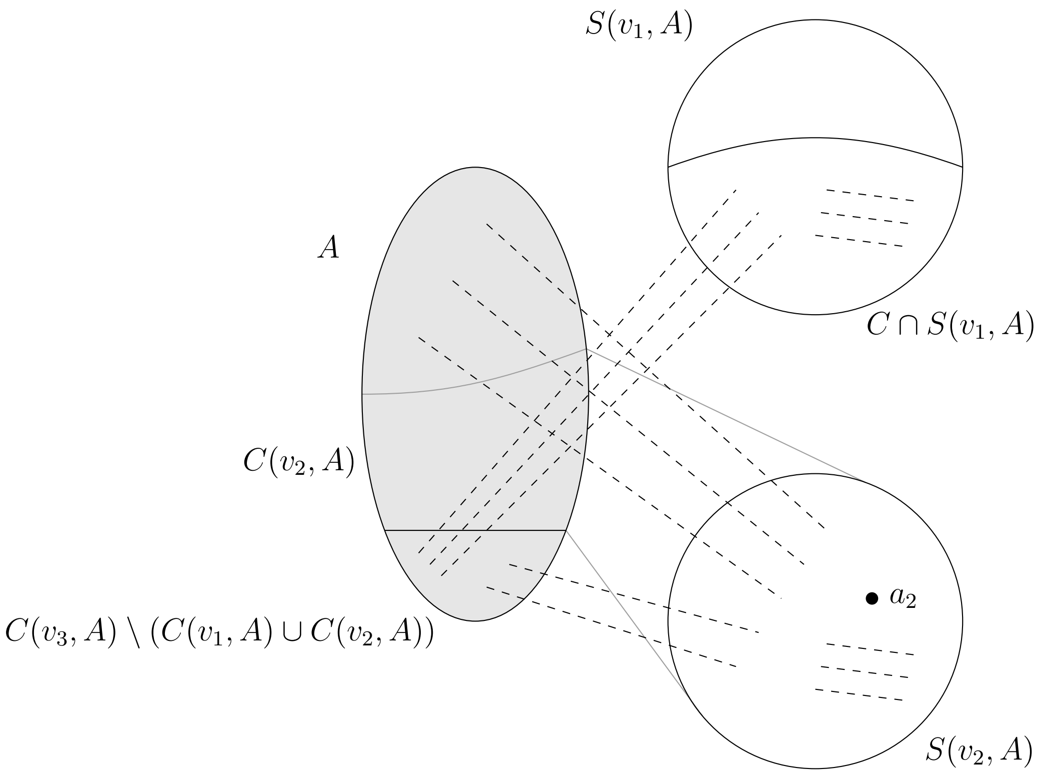


Fig. 2.8: Construction of a white fan centred at a_2 when $|C(v_2, A)| < \frac{5}{18}n$. Only the blades of the fan are shown.

First form $\lfloor \frac{1}{6}n \rfloor$ blades by pairing up vertices of $C \cap S(v_1, A)$ and $C(v_3, A) \setminus (C(v_1, A) \cup C(v_2, A))$. Next, pair all but at most one of the remaining vertices of $C \cap S(v_1, A)$ with each other into blades, and then form as many blades as possible with one vertex in $S(v_2, A)$ and the other in $A \setminus C(v_2, A)$. If some vertices of $S(v_2, A)$ remain, pair all of those except at most one with each other into blades. This means our fan either contains all but at most two vertices of $S(v_2, A), A \setminus C(v_2, A), C \cap S(v_1, A)$, in which case we reach a contradiction as before, or by Claim 2.4.6 it contains at least

$$\begin{aligned}
&\tfrac{1}{6}n - 1 + |C \cap S(v_1, A)| - 1 + 2(|S(v_2, A)| - 1) \\
&\qquad \geq \tfrac{1}{6}n + \tfrac{5}{9}n + 16 + 2|C(v_2, A)| + 2(\tfrac{1}{2}n + 7) - 2 \\
&\qquad \geq \tfrac{13}{18}n + 28 + 2(\tfrac{1}{6}n) > \tfrac{19}{18}n
\end{aligned}$$

vertices, once again giving the desired contradiction. Therefore if $d < \frac{8}{3}n + 6$ and G contains a 3-coverable big clique, then G contains a monochromatic F_n.

2.4.5 $d < \frac{8}{3}n + 6$ and every big clique is 2-coverable

We now come to the final case. We start by proving a simple but important claim.

Claim 2.4.10. *There exist two disjoint cliques A and B of the same colour such that $\frac{4}{3}n + 1 > |A| \geq |B| \geq n + 1$ and $|A| + |B| = \left\lfloor \frac{7}{3}n + 18 \right\rfloor$.*

Proof. First, note that by applying Lemma 2.2.4 to the neighbourhood of a vertex of degree d in some colour, there is a monochromatic clique C of order at least $2d - 4n - 8$ in G. Since $2d \geq |V(G)| - 1 \geq \frac{31}{6}n + 13$, this clique must have size at least $\frac{7}{6}n + 5$. Hence C is big, and so by our assumption it is 2-coverable. Let $S(u_1, C), S(u_2, C)$ be a 2-covering of C. We have

$$\begin{aligned} |S(u_1, C)| + |S(u_2, C)| &\geq |C| + d(u_1) + d(u_2) - 4n \\ &\geq 2d - 8n - 8 + 2\left(\tfrac{31}{6}n + 13 - d\right) = \tfrac{7}{3}n + 18. \end{aligned}$$

We may assume that $|S(u_1, C)| \geq |S(u_2, C)|$. We can now clearly pick cliques $A \subset S(u_1, C)$ and $B \subset S(u_2, C)$, of the opposite colour to C, with $|A| \geq |B|$ and $|A| + |B| = \left\lfloor \frac{7}{3}n + 18 \right\rfloor$. Since $S(u_1, C)$ and $S(u_2, C)$ are disjoint, so are A and B. Note that A is a big clique, so it is 2-coverable. By Observation 2.2.8, we must have $|A| < \frac{4}{3}n + 1$. The result then follows. □

Without loss of generality, assume A and B are black cliques. Note that B need not be big, but it is significant, so by assumption it is not t-coverable for any $t \geq 4$. Hence B is either 2-coverable or 3-coverable.

Now let v_1, v_2 be the covering of A, and let w_1, w_2 (and possibly also w_3) be the covering of B.

Claim 2.4.11. *There exists i such that $S(w_i, B) \cap S(v_1, A) \neq \emptyset$.*

Proof. Assume not. Then all the sets $S(w_i, B)$ as well as $S(v_1, A)$ are disjoint independent sets. A and B are disjoint cliques, and hence each of these can share at most one vertex with each of the independent sets. So G has at least

$$|A| + |B| + |S(v_1, A)| + \sum_i |S(w_i, B)| - 8$$

vertices. But now

$$\begin{aligned} |S(v_1, A)| &\geq |C(v_1, A)| + d(v_1) - 2n \\ &\geq \tfrac{1}{2}|A| + \tfrac{31}{6}n + 13 - d - 2n \\ &\geq \tfrac{1}{2}|A| + \tfrac{1}{2}n + 7. \end{aligned}$$

Similarly, we obtain

$$|S(w_i, B)| \geq |C(w_i, B)| + \tfrac{1}{2}n + 7$$

for each i, and therefore

$$\sum_i |S(w_i, B)| \geq |B| + 2(\tfrac{1}{2}n + 7).$$

Putting these together,

$$\begin{aligned}|A| + |B| + |S(v_1, A)| + \sum_i |S(w_i, B)| - 8 \\ \geq \tfrac{3}{2}(|A| + |B|) + \tfrac{1}{2}|B| + \tfrac{3}{2}n + 13 \geq \tfrac{11}{2}n + 39 > |V(G)|\end{aligned}$$

which is a contradiction. □

We now fix $S(w_i, B)$ such that $S(w_i, B) \cap S(v_1, A) \neq \emptyset$, and let $a \in S(v_1, A) \cap S(w_i, B)$. We will consider two cases, namely $|B \setminus C(w_i, B)| \geq \frac{1}{3}n$ and $|B \setminus C(w_i, B)| < \frac{1}{3}n$. In each case, we will construct a white fan centred at a and show that it has at least n blades. We will need a simple claim.

Claim 2.4.12. *The following inequalities hold:*

- $|S(w_i, B)| \geq \frac{1}{2}n + 7$.
- $|S(v_1, A)| \geq |A \setminus C(v_1, A)| + \frac{1}{2}n + 7$.
- $|S(v_1, A)| + |A \setminus C(v_1, A)| \geq \frac{5}{3}n + 15$.

Proof. For the first inequality, we showed in the proof of Claim 2.4.11 that

$$|S(w_i, B)| \geq |C(w_i, B)| + \tfrac{1}{2}n + 7 \geq \tfrac{1}{2}n + 7.$$

We also showed that $|S(v_1, A)| \geq |C(v_1, A)| + \frac{1}{2}n + 7$ using the same argument. Observation 2.2.8 gives that $|C(v_1, A)| \geq \frac{1}{2}|A|$, implying the second inequality. For the final inequality, note that $|A| \geq \frac{1}{2}\left\lfloor \frac{7}{3}n + 18 \right\rfloor \geq \frac{7}{6}n + 8$, giving

$$\begin{aligned}|S(v_1, A)| + |A \setminus C(v_1, A)| &= |S(v_1, A)| - |C(v_1, A)| + |A| \\ &\geq \tfrac{1}{2}n + 7 + \tfrac{7}{6}n + 8 = \tfrac{5}{3}n + 15.\end{aligned}$$

□

First, assume that $|B \setminus C(w_i, B)| \geq \frac{1}{3}n$. We start by picking $\left\lfloor \frac{1}{3}n \right\rfloor$ blades with a vertex in $S(w_i, B)$ and the other in $B \setminus C(w_i, B)$; by Claim 2.4.12, this can be done. Including a, we have used at most $\frac{1}{3}n + 1$ vertices of $S(v_1, A)$, so Claim 2.4.12 tells us that there are still more than $|A \setminus C(v_1, A)|$ vertices in $S(v_1, A)$ that we have not yet added to the fan. We can therefore keep adding blades with a vertex in $S(v_1, A)$ and the other in $A \setminus C(v_1, A)$ until we run out of vertices in $A \setminus C(v_1, A)$. Finally, we pair all the remaining vertices of $S(v_1, A)$ into blades, except possibly one vertex.

The fan now contains all the vertices of $A \setminus C(v_1, A)$, all but at most one of the vertices of $S(v_1, A)$, and $\lfloor \frac{1}{3}n \rfloor$ vertices of $B \setminus C(w_i, B)$. We have counted at most one vertex twice, since A is disjoint from $S(v_1, A)$ and B, and B shares at most one vertex with $S(v_1, A)$. Thus, by Claim 2.4.12, we have used a total of at least

$$|S(v_1, A)| + |A \setminus C(v_1, A)| + \tfrac{1}{3}n - 3 \geq \tfrac{5}{3}n + 15 + \tfrac{1}{3}n - 3 = 2n + 12$$

vertices. So our fan has at least n blades and we have reached a contradiction.

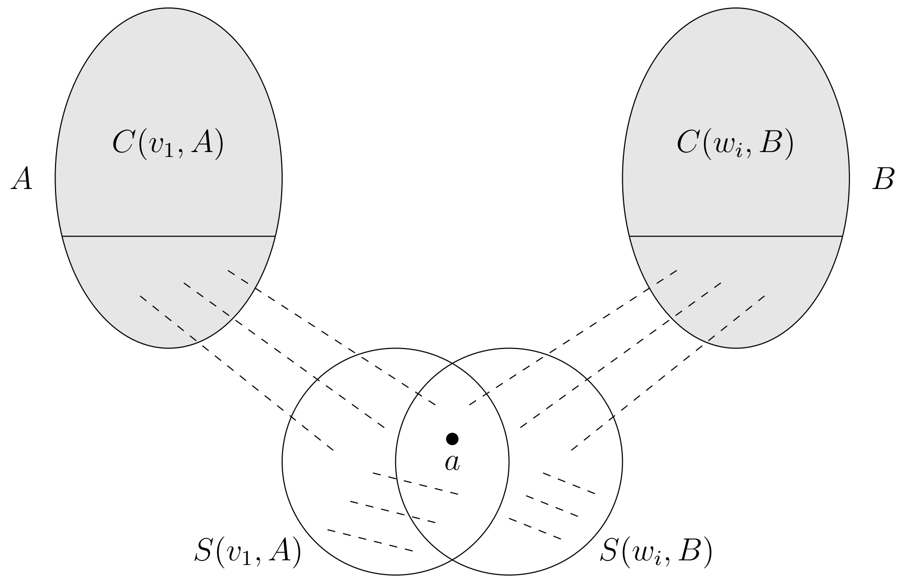


Fig. 2.9: Construction of a white fan centred at a when $|B \setminus C(w_i, B)| < \frac{1}{3}n$. Only the blades of the fan are shown. Potential single-vertex intersections $A \cap S(w_i, B)$ and $B \cap S(v_1, A)$ are not shown.

Next, assume that instead $|B \setminus C(w_i, B)| < \frac{1}{3}n$. Note that as $S(v_1, A) \cap S(w_i, B)$ is a white clique, it must have fewer than $\frac{4}{3}n + 1$ vertices, or else by Observation 2.2.8 it would not be 2-coverable, contradicting our assumption about big cliques. Since $S(v_1, A)$ and $S(w_i, B)$ intersect each of B and A in at most one vertex, we have

$$\begin{aligned}
&|S(v_1, A) \cup S(w_i, B)| + |A \setminus C(v_1, A)| + |B \setminus C(w_i, B)| \\
&\qquad \geq |S(v_1, A)| + |S(w_i, B)| - |S(v_1, A) \cap S(w_i, B)| + |A| + |B| \\
&\qquad\quad - |C(v_1, A)| - |C(w_i, B)| - 2 \\
&\qquad \geq 2(\tfrac{1}{2}n + 7) + \tfrac{7}{3}n + 17 - (\tfrac{4}{3}n + 1) - 2 = 2n + 28.
\end{aligned}$$

Thus if we can show we can find a white fan centred at a using all but at most two of the vertices in

$$S(v_1, A) \cup S(w_i, B) \cup (A \setminus C(v_1, A)) \cup (B \setminus C(w_i, B)),$$

then it has at least n blades and we are done.

We start by creating blades with one vertex in $B \setminus C(w_i, B)$ and the other in $S(w_i, B)$. We eventually run out of vertices in $B \setminus C(w_i, B)$, as

$$|B \setminus C(w_i, B)| < \tfrac{1}{3}n < \tfrac{1}{2}n + 7 \leq |S(w_i, B)|,$$

using Claim 2.4.12 for the last inequality. Next, we create blades with one vertex in $S(v_1, A)$ and the other in $A \setminus C(v_1, A)$. Since we know that

$$|S(v_1, A)| \geq |A \setminus C(v_1, A)| + \tfrac{1}{2}n + 7$$

by Claim 2.4.12, and we used at most $\frac{1}{3}n + 1$ vertices of $S(v_1, A)$ in the previous step, we will run out of vertices in $A \setminus C(v_1, A)$ first. Finally, we can use all but at most one of the remaining vertices in $S(v_1, A)$ by pairing them up, and we can use all but at most one of the remaining vertices in $S(w_i, B) \setminus S(v_1, A)$ by pairing them up. The result follows, finishing the proof of Theorem 2.1.1.

2.5 Conclusion

In this chapter, through controlling the degrees of the vertices as well as taking a more global approach, we have reduced the bound on $R(F_n)$ from $(5+\frac{1}{2})n+O(1)$ to $(5 + \frac{1}{6})n + O(1)$. This is still far from the lower bound of $\frac{9}{2}n + O(1)$, which we suspect is much closer to the correct magnitude.

We expect that with some more care the methods in our proof could likely be improved to give an upper bound of $(5 + \delta)n + O(1)$ for some $\delta < \frac{1}{6}$, but we encounter more obstacles as we approach $5n$. For example, our proof repeatedly makes use of monochromatic cliques of order significantly larger than n. However, in a graph of order $5n+O(1)$ in which every vertex has degree around $\frac{5}{2}n$ and there is no monochromatic F_n, Lemma 2.2.4 only guarantees the existence of a clique of order $n+O(1)$. It therefore seems unlikely that present methods could bring the upper bound close to $\frac{9}{2}n$, or even verify the conjecture of Chen, Yu and Zhao [24] that $r(F_n) \leq R(nK_3) = 5n$.

Chapter 3

Odd colourings on the torus

3.1 Introduction

The subject of graph colouring originates with the statement of the Four-Colour Problem by Guthrie in 1852: can every map on the plane be coloured with four colours in such a way that no two neighbouring regions share the same colour? This problem was a major focus of research in combinatorics for a century, beginning in earnest with Cayley's article in 1879 [21] and Kempe's claimed proof [52] later the same year. Kempe's proof contained an error which was discovered only in 1890 by Heawood [48]; however, as Heawood noted, Kempe's argument did show that five colours suffice.

Kempe observed (although not using modern terminology) that the problem of colouring a map is equivalent to finding a *proper colouring* of a graph G, namely a colouring of the vertices of G such that adjacent vertices have different colours. This graph is formed by placing a vertex in every face and joining two vertices if the corresponding faces are adjacent. Maps drawn on the plane correspond to *planar graphs*. The smallest number of colours in any proper colouring of G is the *chromatic number* of G, denoted $\chi(G)$, so the Four-Colour Problem corresponds to asking whether every planar graph G has $\chi(G) \leq 4$.

In modern language, Kempe's general approach is as follows. He considers a minimal planar graph G with $\chi(G) \geq 5$ and finds a set of configurations that is *unavoidable*, meaning that one of these configurations must appear somewhere in any planar graph. We will purposefully keep the definition of "configuration" vague, but it should be thought of as a particular arrangement of vertices, some of which have their degrees specified. Kempe attempts to show that his set of unavoidable configurations is also *reducible*: this means that any graph H containing such a configuration can be reduced to a graph with fewer vertices or edges, which by induction we can assume to be 4-colourable, in such a way that the 4-colouring can be extended to a 4-colouring of H. The existence of a set that is both unavoidable and reducible then immediately implies that no such minimal counterexample can exist, and so $\chi(G) \leq 4$ for all planar G.

While Kempe's proof was flawed, the ideas behind it proved to be useful. Other approaches towards solving the Four-Colour Problem were tried: for example, Birkhoff [9] attempted to use algebraic methods, inventing the chromatic polynomial. However, constructing a set of configurations that is both unavoidable and reducible was ultimately the more successful method.

Through the 20$^{\text{th}}$ century, more and more sets of configurations in the plane were found to be either unavoidable or reducible. Much of the early work on reducibility was done by Birkhoff [10], and his work was extended by Franklin [40] in 1922. In general, proving that a configuration is reducible requires showing that it can be reduced no matter how the neighbours of the vertices in the configuration are coloured. The number of possible colourings of these vertices quickly becomes very large and difficult for a human to check.

Kempe gave the first non-trivial example of an unavoidable set when he showed that any planar graph must contain a vertex of degree at most 5. Cayley [21] noted that we can add edges to any planar graph until all faces are triangles; it therefore suffices to prove that all such *triangulated graphs* are 4-colourable, since adding edges cannot reduce the chromatic number of a graph. We can thus consider sets that are unavoidable in triangulated graphs: finding such a set that consists only of reducible configurations would prove that $\chi(G) \leq 4$ for all triangulated graphs and therefore for all planar graphs. The first advance on unavoidable sets was made in 1904 by Wernicke [85], who proved that any triangulated graph contains a vertex of degree less than 5 or a vertex of degree 5 that is adjacent to a vertex of degree 5 or 6. Further progress was made by Franklin [40] and Lebesgue [55], but these ideas were not sufficient to solve the problem.

The most important breakthrough was made by Heesch, who developed a much more powerful technique for finding unavoidable sets which would become known as the *method of discharging*. Heesch published his work in 1969 [49] but had been working on it for the previous 20 years. We will discuss the method of discharging in more detail in Section 3.2 and use it in Section 3.4. Finally, Appel and Haken, assisted by Koch [5, 6], proved that $\chi(G) \leq 4$ for all planar graphs G, using a very complex discharging argument. The proof was announced in 1976 and published the following year. The unavoidable set contained 1936 configurations, and a computer was required to prove that these were all reducible. The Four-Colour Problem therefore became the Four-Colour

Theorem. See Wilson's book [86] for further background, more details on the techniques used in the proof, and other results on graph colouring.

The Four-Colour Problem can be modified by imposing constraints on the colouring. For example, we can insist that the colours of the vertices in the neighbourhood of each vertex satisfy some condition. One such constraint was recently introduced by Petruševski and Škrekovski [68]: a *proper odd vertex-colouring* of a simple graph G, often referred to as an *odd colouring* for short, is a proper colouring c of $V(G)$ with the property that, for every non-isolated vertex v of G, there exists a colour i such that $|c^{-1}(i) \cap N(v)|$ is odd: in other words, for every v, there is some colour that appears an odd number of times in the neighbourhood of v. The *odd chromatic number* of a graph G, denoted $\chi_o(G)$, is the smallest number k such that G admits an odd colouring using k colours.

Odd colouring on graphs is a special case of a more general notion of odd colouring for hypergraphs, introduced by Cheilaris, Keszegh and Pálvölgyi [23] in 2013. An odd colouring of a hypergraph H is a colouring of the vertices of H such that for every edge e, some colour appears an odd number of times in the vertices of e. Note that the notion of odd colouring for a graph G does not arise from setting $H = G$; instead, H has the same vertex set as G, but the edges of H are the edges of G (forcing the colouring to be proper) and the neighbourhoods of the vertices of G. Now odd colourings of H as a hypergraph correspond directly to odd colourings of G as a graph.

Cheilaris, Keszegh and Pálvölgyi used odd colourings to find lower bounds for *conflict-free colourings* of hypergraphs; a conflict-free colouring of a hypergraph H is a colouring of the vertices of H such that for every edge e, some colour appears exactly once in the vertices of e. Conflict-free colourings are of interest in modelling frequency assignments on cellular networks: see, for example, [35, 79]. Conflict-free colourings can be defined for graphs, analogously to odd colourings: a proper colouring of a graph G is *conflict-free* if, for every vertex v, some colour appears exactly once in its neighbourhood. Clearly any conflict-free colouring is an odd colouring.

Note that while we trivially have $\chi_o(G) \geq \chi(G)$, there is no upper bound for $\chi_o(G)$ in terms of $\chi(G)$. To see this, fix an integer $n \geq 2$ and consider a bipartite graph H with vertex classes $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_{i,j} : 1 \leq i < j \leq n\}$, and edges $y_{i,j} \sim x_i$ and $y_{i,j} \sim x_j$ for each i, j. Clearly $\chi(H) = 2$, but for each i and j, x_i and x_j must have different colours, since otherwise no colour appears an odd number of times in $N(y_{i,j})$. Hence $\chi_o(H) \geq n + 1$.

One way in which odd colouring differs fundamentally from standard graph colouring is that a proper subgraph of G can have a larger odd chromatic number than G itself. For example, the star $K_{1,3}$ has odd chromatic number 2, but its proper subgraph $K_{1,2}$ has odd chromatic number 3.

Just as for proper colourings, we can ask for the minimum number of colours required to properly odd-colour every planar graph. It is easily seen that 4 colours will no longer suffice: the cycle C_5 requires 5 colours. Petruševski and Škrekovski [68] used the discharging method to prove that every planar graph G satisfies $\chi_o(G) \leq 9$, and conjectured that in fact $\chi_o(G) \leq 5$ for all planar graphs, which would be best possible. Caro, Petruševski and Škrekovski [20] then used a result of Aashtab, Akbari, Ghanbari and Shidani [1], derived from the Four-Colour Theorem, to prove that $\chi_o(G) \leq 8$ if G is a connected planar graph that has even order or contains a vertex of degree 2 or any odd degree. Petr and Portier [67] used the discharging method to cover the remaining cases, proving that $\chi_o(G) \leq 8$ for all planar graphs G. Fabrici, Lužar, Rindošová and Soták [36] gave an independent proof of the same bound; in fact they proved the more general result that every planar graph has a conflict-free colouring with at most 8 colours.

A major difference between the original Four-Colour Problem and the equivalent problem for odd colourings is that we cannot only work with triangulated graphs, because adding edges can decrease the odd chromatic number of a graph.

Another way in which the Four-Colour Problem can be modified is by considering graphs embedded in surfaces other than the plane. We will use the word "surface" to refer to compact surfaces with no boundary. The classification theorem for such surfaces was first proved rigorously by Brahana [15] although it had been stated previously multiple times without full justification. This classification is as follows:

- T_g, the torus with g holes, $g \geq 0$. These are orientable surfaces. Note that T_0 is the sphere and T_1 is the standard torus.

- S_g, a sphere with g discs cut out, $g \geq 1$, and a *cross-cap* added in place of each disc. In other words, every point on the boundary of each disc is identified with the point opposite it. These are non-orientable surfaces. S_1 is the real projective plane and S_2 is the Klein bottle.

Here g is the *genus* of the surface.

Note that the plane can be turned into a sphere by the addition of a single point, so for our purposes the plane and the sphere are the same. Considering chromatic numbers of graphs on other compact surfaces is therefore a generalisation of the Four-Colour Problem. We can define the chromatic number of a surface S, $\chi(S)$, to be the maximum of the chromatic numbers of all graphs that embed into S. It is not immediately clear why such a maximum should exist, but Heawood [48] proved that

$$\chi(S) \leq \left\lfloor \frac{7 + \sqrt{49 - 24E}}{2} \right\rfloor$$

for surfaces other than the sphere, where E is the *Euler characteristic* of the surface, equal to $2 - 2g$ for orientable surfaces and $2 - g$ for non-orientable surfaces.

Heawood's bound gives that $\chi(T) \leq 7$ for the torus T, and Heawood also found an embedding of the complete graph K_7 in the torus, proving that $\chi(T) = 7$. In fact, Heawood's bound is exact for all surfaces except the Klein bottle, which has Heawood bound 7 but chromatic number 6. This was proved across many papers by different authors, finishing with Ringel and Youngs [72] in 1968.

In this chapter, we will consider odd colourings of graphs that can be embedded into a torus; we call these graphs *toroidal graphs*. Since K_7 embeds in the torus, at least 7 colours are sometimes required to odd-colour a toroidal graph. We will demonstrate an upper bound on $\chi_o(G)$ for toroidal graphs:

Theorem 3.1.1. *Let G be a toroidal graph. Then $\chi_o(G) \leq 9$.*

This implies that the maximum possible value of $\chi_o(G)$ for toroidal graphs must be 7, 8 or 9.

Our proof mainly builds on the discharging techniques used by Petruševski and Škrekovski [68] in their proof that $\chi_o(G) \leq 9$ for planar graphs. Our application of the discharging method is rather more sensitive and leaves a special case that must be dealt with separately.

We begin with an overview of the discharging method in Section 3.2. In Section 3.3, we make some preliminary observations. We apply the discharging method in Section 3.4. We then tackle the remaining special case in Section 3.5. In Section 3.6, we discuss the implications of our result and opportunities for further research.

3.2 Overview of the discharging method

Before we embark on the proof of Theorem 3.1.1, we will give an introduction to the method of discharging in a simpler setting. Following Wilson's book [86], we will use it to give a proof of the following result, which was mentioned in the previous section:

Theorem 3.2.1 (Wernicke [85]). *Let G be a triangulated planar graph. Then either G contains a vertex of degree at most 4, or G contains a vertex of degree 5 that is adjacent to a vertex of degree 5 or 6.*

Let the number of vertices, edges and faces of a graph G be V, E and F respectively. The set of faces of G, which we call $F(G)$, includes the exterior, infinite face (note that if we draw the graph on the sphere, this face is just like all the others). Euler's formula states that, for a connected planar graph G, we have

$$V - E + F = 2.$$

Euler's formula is essential for almost all colouring results for planar graphs, because it implies that every planar graph contains a vertex of degree at most 5. We begin by noting that every face of G has at least 3 edges, so by double-counting we obtain $3F \leq 2E$. This implies that

$$V - \tfrac{1}{3}E \geq 2,$$

and so

$$\sum_{v \in G} d(v) = 2E \leq 6V - 12.$$

It immediately follows that some v has $d(v) \leq 5$.

The idea behind the discharging method is that we assign "charge" to the vertices (and possibly faces) of a graph G according to their degrees (and possibly sizes) in such a way that the total charge across the whole graph is negative. We then redistribute the charge between the vertices and faces according to some set of rules which preserve charge. The aim is to find some set $\mathcal{S}$ of configurations such that if none of those configurations is present in G, then after the redistribution of charge, every vertex and face has non-negative charge. This is a contradiction, since the total charge is negative and charge is conserved by the rules. Hence $\mathcal{S}$ is an unavoidable set.

We will now prove Theorem 3.2.1 using the discharging method.

Proof of Theorem 3.2.1. We begin by assigning charge $d(v) - 6$ to every vertex $v \in G$. The total charge is now

$$\sum_{v \in G} d(v) - 6V \leq -12,$$

by Euler's formula as above.

We use a single discharging rule: every vertex of degree at least 7 sends charge $\frac{1}{5}$ to each of its neighbours of degree 5.

If G has a vertex of degree at most 4, then we are done, so as G is planar we can assume it has minimum degree 5. The vertices of degree 5 begin with charge -1. If a vertex of degree 5 has a neighbour of degree 5 or 6 then we are done, so we may assume that if $v \in G$ has degree 5 then every neighbour of v has degree at least 7. By the discharging rule, v then receives charge $\frac{1}{5}$ from each of its 5 neighbours and finishes with charge 0. The vertices of degree 6 begin with charge 0 and are unaffected by the discharging rule, so they also finish with charge 0.

We are left with the vertices of degree at least 7. Let w be such a vertex, and suppose $d(w) = d \geq 7$. We consider the neighbours of w in the order in which they appear around w. If two neighbours of w appear consecutively in order around w, then they must be adjacent, because G is a triangulated graph. Since we can assume no two vertices of degree 5 are adjacent, it follows that at most half of the neighbours of w have degree 5. The vertex w therefore finishes with charge at least

$$(d-6) - \tfrac{1}{2}d\left(\tfrac{1}{5}\right) = \tfrac{9}{10}d - 6 \geq \tfrac{63}{10} - 6 > 0.$$

Hence every vertex of G ends with charge at least 0. But this is a contradiction, since G began with total charge -12 across all vertices, and charge is conserved. Theorem 3.2.1 follows. □

In general we will assign charge to the faces of the graph as well as the vertices. We need to define the *size* of a face f, which we will denote $d(f)$. Note that the boundary of f is not necessarily a cycle: for example, it could consist of a cycle together with some vertices of degree 1 in the interior of the cycle, each joined to a vertex in the cycle. We must therefore take care with the definition of $d(f)$. Consider the boundary of f as a closed walk. We define $d(f)$ to be the number of vertices appearing in this walk, counting with multiplicity; when f is a cycle, this is of course equal to the number of vertices of f.

There are several standard ways in which charge may be assigned to the vertices and faces of a planar graph such that the total charge is a negative constant. One way is to assign charge $d(v) - 6$ to each vertex, as we did for the proof of Theorem 3.2.1, and charge $2d(f) - 6$ to each face. The total charge across all the vertices and faces of G is now

$$\begin{aligned}\sum_{v \in G}(d(v) - 6) + \sum_{f \in F(G)}(2d(f) - 6) &= 2E - 6V + 4E - 6F \\ &= -6(V - E + F) = -12,\end{aligned}$$

by noting that summing $d(f)$ over all faces of the graph counts every edge exactly twice, and then applying Euler's formula. This is the distribution of initial charges that was used by Petruševski and Škrekovski [68] in their proof that $\chi_o(G) \leq 9$ for planar graphs, and by Petr and Portier [67] in improving this upper bound to 8. It is also the distribution that we will use, but the fact that we are working with toroidal graphs rather than planar graphs will cause difficulties, because the total charge turns out to be 0 rather than negative. This will be discussed further in Section 3.4.

3.3 Outline of the proof and preliminaries

For convenience, we will refer to a proper odd colouring with at most 9 colours as a *nice colouring*. From now on, we assume that there exists a toroidal graph which does not have a nice colouring, and $G = (V, E)$ will always denote a minimal such graph. Clearly G is connected.

In our proof, we will often be faced with a situation where we have a colouring c of $G \setminus \{v\}$, and we would like to find a colour that we can use at v to extend c to a nice colouring of G. There are two reasons why we might not be able to use a particular colour at v. Firstly, for each $w \in N(v)$, we cannot use $c(w)$ at v, since the resulting colouring would not be proper. We refer to such colours as *forbidden at v by properness*. Secondly, for each $w \in N(v)$, there can be at most one colour $b(w)$ such that using $b(w)$ at v would result in each colour being used an even number of times in $N(w)$. We refer to such colours as *forbidden at v by oddness*. There are therefore at most $2d(v)$ colours forbidden at v in total.

We will say that a colouring is *proper at v* if v is coloured differently from all its neighbours, and *odd at v* to mean that some colour appears an odd number of times in $N(v)$. We will also use this terminology for partial colourings as long as they are defined at all the vertices of $N(v)$. Note that if v has odd degree, then any colouring is automatically odd at v.

For a graph G on a surface S with Euler characteristic x, we have a version of Euler's formula:

$$V - E + F \geq x.$$

The torus has Euler characteristic 0, so for any toroidal graph we have

$$V - E + F \geq 0.$$

Equality holds if and only if every face of G is homeomorphic to a disk (see [63], Chapter 3). Note that, unlike for planar graphs, G being connected is not sufficient for equality to hold.

As for planar graphs, we have $3F \leq 2E$ by double-counting. It follows that $V \geq \frac{1}{3}E$, and so

$$\sum_{v \in G} d(v) = 2E \leq 6V,$$

so $\delta(G) \leq 6$. In order for equality to hold, G must be 6-regular and every face of G must have 3 edges. In addition, equality must hold in Euler's formula, so every face of G must be homeomorphic to a disk. Therefore $\delta(G) = 6$ iff G is 6-regular and all its faces are triangles, in other words, iff G is a 6-regular triangulation of the torus.

In fact, for our minimal G, we must have $\delta(G) \geq 5$. This was proved by Petruševski and Škrekovski [68] in the context of planar graphs, but still holds for toroidal graphs. For completeness, we reproduce their proof here.

Claim 3.3.1. *G has minimum degree $\delta(G) \geq 5$.*

Proof. Suppose that G contains a vertex v of degree 1 or 3. Let c be a nice colouring of $G \setminus \{v\}$, which exists by minimality. We would like to extend c to a nice colouring of G. There are now at most 3 colours forbidden at v by properness and at most 3 forbidden by oddness, so there is some colour left over that can be used at v. Since v has odd degree, the resulting colouring is odd at v and therefore nice, which is a contradiction.

Now suppose instead that G contains a vertex v of degree 2 or 4. Let w be an arbitrary neighbour of v, and let G' be the graph constructed from G by removing v and then joining w to every vertex in $N_G(v)$ to which it is not already adjacent. The graph G' is toroidal, and so by minimality it has a nice colouring c. Now we return to G and colour the vertices of $G \setminus \{v\}$ according to c. There are at most 4 colours forbidden at v by properness and 4 forbidden by oddness, leaving a colour that can be used at v. Since c is a proper colouring of G', the colour $c(w)$ appears exactly once in $N(v)$, so the resulting colouring is odd at v. We therefore have a nice colouring of G, which is again a contradiction. □

The rest of the proof will be structured as follows. In Section 3.4, we use the discharging method and the minimality of G to show that G cannot contain a vertex of degree 5. This leaves the case where G is a 6-regular triangulation. In Section 3.5, we use the classification of 6-regular triangulations of the torus by Altshuler [4] to show that every such triangulation admits a nice colouring, finishing the proof.

As is standard, we will refer to a vertex of degree d as a d-vertex, and a vertex of degree at least d as a d^+-vertex. We would like to define $d(f)$ as we did for planar graphs, but in a toroidal graph the boundary of a face f need not be connected. However, it can still be considered as a disjoint union of closed walks, and we define $d(f)$ to be the total number of vertices appearing in this union of walks, counting with multiplicity. We call a face of size k a k-face, and a face of size at least k a k^+-face.

3.4 Application of the discharging method

In this section, we will deal with the case $\delta(G) = 5$. We will use the fact that the torus is locally homeomorphic to the plane, and therefore we will draw subgraphs of G as if they were on the plane, although this may not always faithfully represent the embedding in the torus. However, the order of the edges around each vertex will be unambiguous and represented correctly.

Throughout the proof, there will be some occasions on which two vertices that are given different names could in fact be the same, or where two named vertices which are not defined to be adjacent could in fact be adjacent. However, this will never affect our arguments.

First, we will need a simple observation made by Petruševski and Škrekovski [68].

Claim 3.4.1. *Suppose $v \in G$ is a 5-vertex. Then v has at most one neighbour of odd degree.*

Proof. Suppose that v has two neighbours x and y of odd degree. Let c be a nice colouring of $G \setminus \{v\}$; we would like to extend c to v. There are at most 5 colours forbidden at v by properness, and since x and y have odd degree, they cannot forbid colours at v by oddness, so there are at most 3 colours forbidden by oddness. This means there is a colour left over that can be used at v, and since v has odd degree, the resulting colouring is nice, which is a contradiction. □

We will now introduce our discharging rules. We begin by assigning to every vertex $v \in G$ a charge $d(v) - 6$ and to every face f a charge $2d(f) - 6$. For planar graphs, Euler's formula would now give total charge -12 as discussed before, but for our toroidal G we only obtain that the total charge across all vertices and faces of the graph is at most 0:

$$\sum_{v\in G}(d(v)-6)+\sum_{f\in F(G)}(2d(f)-6) = 2E-6V+4E-6F$$
$$= -6(V-E+F) = 0.$$

The discharging rules used by Petruševski and Škrekovski [68] are not sufficient when the total charge is 0, so we will need our own set of rules, which are as follows. If a vertex appears with multiplicity greater than 1 in the closed walk around the boundary of a face, we consider each appearance to be a separate vertex when applying the rules.

(R1) Every 5^+-face sends charge 1.1 to each incident 5-vertex.

(R2) Every 4-face sends charge 1 to each incident 5-vertex, unless its incident vertices are, in order, two adjacent 5-vertices and two adjacent 6^+-vertices, in which case it sends charge $\frac{3}{4}$ to each 5-vertex.

(R3) If u and v are 6^+-vertices on a 4^+-face f that are adjacent along an edge of f and also both incident to a 3-face uvw, and w is a 5-vertex, then f sends charge $\frac{1}{2}$ to w.

(R4) Suppose v is a 7^+-vertex with at least one neighbouring 5-vertex. Let a *block* be a maximal set of 5-vertices in $N(v)$ that appear consecutively in order around v. Now v distributes its charge evenly between the blocks, and within each block the charge is distributed evenly between the vertices. For example, an 8-vertex with 4 neighbouring 5-vertices in blocks of size 2, 1 and 1 sends charge $\frac{2}{3}$ to each vertex in a block of size 1, and $\frac{1}{3}$ to each of the vertices in the block of size 2.

Claim 3.4.2. *After discharging, all faces and* 6^+*-vertices have non-negative charge.*

Proof. The only way in which a vertex can lose charge is by (R4), and by definition no vertex can end with negative charge after the application of this rule. Therefore a vertex can only end with negative charge if it began with negative charge, and by Claim 3.3.1, this is only the case for 5-vertices.

We now turn to faces. When considering the total charge given out by a face f, we can imagine the charge $\frac{1}{2}$ distributed by (R3) as being split into two $\frac{1}{4}$ charges, one given to u and one to v. Each 6^+-vertex incident to f is on at most two edges for which (R3) applies. Therefore f gives out charge at most $\frac{1}{2}$ for each incident 6^+-vertex, and at most 1.1 for each incident 5-vertex.

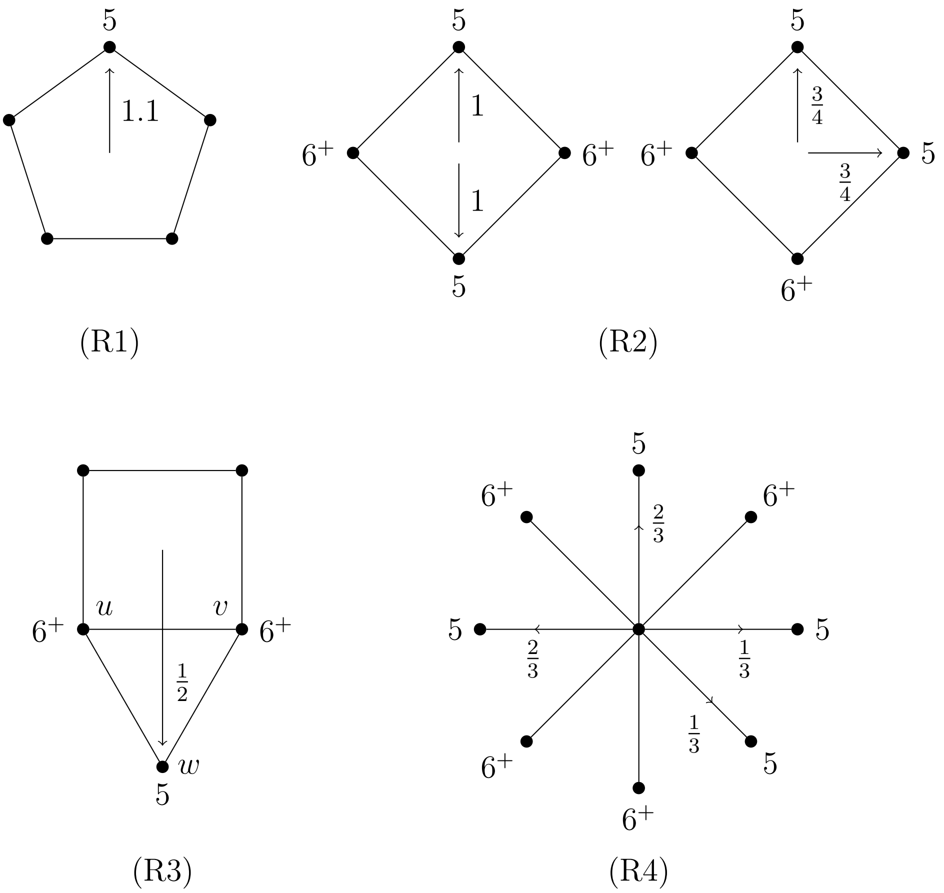


Fig. 3.1: Examples of the discharging rules. The numbers next to vertices indicate their degrees, and the arrows indicate movement of charge.

If f is a 7^+-face, then the initial charge $2d(f) - 6$ is greater than $1.1d(f)$, so trivially f must finish with non-negative charge. If f is a 6-face, then it has initial charge 6, but it cannot have more than four incident 5-vertices: otherwise, one 5-vertex would be adjacent to two others, contradicting Claim 3.4.1. Note that this is still true even if the boundary of f consists of two disjoint 3-cycles. Therefore, by (R1) and (R3), f gives out charge at most $4.4 + 1 < 6$, so it finishes with positive charge.

If f is a 5-face, then it has initial charge 4, and by Claim 3.4.1 it has at most three incident 5-vertices. If it has exactly three, then the two remaining vertices are not adjacent on f and therefore (R3) does not apply; thus f gives out charge at most 3.3. If instead f has at most two incident 5-vertices, then it gives out charge at most $2.2 + \frac{3}{2} < 4$, so in either case f finishes with positive charge.

If f is a 4-face, then it begins with charge 2 and has at most two incident 5-vertices by Claim 3.4.1. First suppose f has exactly two 5-vertices. If they are adjacent, then (R2) implies that they each receive charge $\frac{3}{4}$. The remaining two vertices of f are 6^+-vertices, so the greatest additional charge that f can give out is $\frac{1}{2}$, by (R3). Thus f gives out charge at most 2, as required. If instead the two 5-vertices are not adjacent, then they each receive charge 1 by (R2), but (R3) does not apply and so f again gives out total charge 2.

Now consider the case where f is a 4-face with exactly one incident 5-vertex. This vertex receives charge 1, and (R3) applies to at most two edges of f, so it gives out total charge at most 2. If instead every vertex of f is a 6^+-vertex, then it gives out charge at most $\frac{1}{2}$ for each vertex, so once again the total charge given out is at most 2.

Finally, if f is a 3-face then it begins with charge 0 but cannot give out charge, so we are done. □

Our graph G begins with total charge at most 0 and the rules preserve charge, so if G has minimum degree 5, Claim 3.4.2 implies that some 5-vertex must finish with charge at most 0. If instead $\delta(G) = 6$, then as noted earlier, G is a 6-regular triangulation, and thus every vertex and face start and end with charge 0. The discharging method therefore does not help when $\delta(G) = 6$, and we will have to treat this case separately.

For the remainder of this section, we will restrict ourselves to the case where $\delta(G) = 5$. Let v be a 5-vertex that has charge at most 0 after the discharging process: in other words, v receives total charge at most 1 during discharging (recall that v cannot give out charge). Let the neighbours of v be v_1, v_2, v_3, v_4 and v_5 in anticlockwise order.

Claim 3.4.3. *There are five distinct faces around v.*

Proof. If v is not in 5 distinct faces, then some face f appears twice around v. This face clearly must be a 4^+-face. By the discharging rules, f sends charge at least $\frac{3}{4}$ to v for each appearance of v on the boundary of f, so f sends total charge at least $\frac{3}{2}$ to v, a contradiction. □

We will make use the following lemma to eliminate the cases in the remainder of the proof.

Lemma 3.4.4. *Suppose that the edges v_1v_2, v_2v_3 and v_3v_4 are all present in G, and that v_2 and v_3 are 6-vertices with a common neighbour $x \neq v$. Let c be a nice colouring of $G \setminus \{v\}$ in which $c(v_1)$, $c(v_2)$, $c(v_3)$ and $c(v_4)$ are all distinct. Then v_2 and v_3 do not forbid two distinct colours at v by oddness unless at least one is already forbidden by properness.*

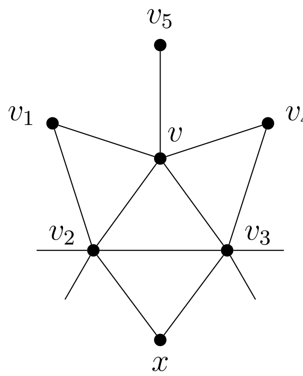


Fig. 3.2: The setting of Lemma 3.4.4. All edges from v, v_2 and v_3 are shown.

Proof. Suppose that a colouring c exists as above, and let $c(v_i) = i$ for $i = 1, 2, 3, 4$. Suppose further that v_2 and v_3 forbid colours $2'$ and $3'$ respectively at v by oddness, where $2' \neq 3'$ and $2', 3' \notin \{1, 2, 3, 4\}$. Now the colours of the neighbours of v_2 in $G \setminus \{v\}$ must be $1, 1, 3, 3, 2'$ in some order. In particular, $c(x) \in \{1, 3, 2'\}$. Similarly, consideration of the colours around v_3 shows that $c(x) \in \{2, 4, 3'\}$. But now there is no possible colour for x and we have a contradiction. □

We will need one more structural lemma.

Lemma 3.4.5. *G does not contain two adjacent 5-vertices which have two common neighbours.*

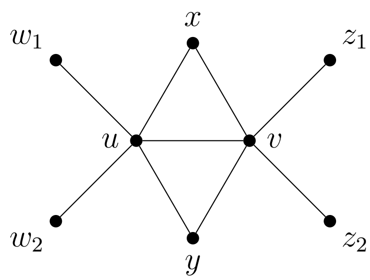


Fig. 3.3: The setting of Lemma 3.4.5. All edges from u and v are shown.

Proof. Suppose to the contrary that u and v are adjacent 5-vertices with common neighbours x and y. Let the two remaining neighbours of u be w_1 and w_2, and let the two remaining neighbours of v be z_1 and z_2. Note that we could have some $w_i = z_j$.

Consider the graph $G' = G/\{u, v\}$ formed by contracting the edge uv to form a single vertex $\{u, v\}$. By the minimality of G, this graph has a nice colouring, which we will call c. We will use c to colour the vertices of $G \setminus \{u, v\}$, and show that we can always choose colours at u and v to produce an odd colouring of G. Note that any colouring of G will be odd at u and v, since they both have odd degree.

Suppose that we colour u with $c(\{u,v\})$. The resulting partial colouring is both proper and odd at each w_i that is not the same as some z_j, since c is an odd colouring of G'. Note that u has degree 5 so cannot forbid a colour at v by oddness. For every remaining neighbour t of v, let $b(t)$ be the colour that t forbids at v by oddness, if it exists. There are at most 5 colours forbidden at v by properness, and at most 4 forbidden by oddness. Therefore there is always a colour available at v to produce an odd colouring of G unless the 9 colours $c(\{u,v\}), c(x), c(y), c(z_1), c(z_2), b(x), b(y), b(z_1), b(z_2)$ all exist and are all distinct. Similarly, we could colour v with $c(\{u,v\})$ instead of u, and define b as above; there is no ambiguity in the definitions of $b(x)$ and $b(y)$ since the neighbourhoods of x and y have the same multisets of colours in each case. This partial colouring extends to an odd colouring of G unless $c(\{u,v\}), c(x), c(y), c(w_1), c(w_2), b(x), b(y), b(w_1), b(w_2)$ exist and are all distinct.

Suppose that the two sets of 9 colours above are indeed distinct. Now we assign the colour $b(x)$ to u and $b(y)$ to v. This results in a proper colouring of G by distinctness. This colouring is odd at each w_i and z_j, also by distinctness. Finally, $c(\{u,v\})$ and $b(y)$ each appear an even number of times in $N_{G'}(x)$, since they are distinct from each other and from $b(x)$. They therefore appear an odd number of times in $N_G(x)$. Similarly, $c(\{u,v\})$ and $b(x)$ appear an odd number of times in $N(y)$. Thus G has a nice colouring, which is a contradiction. □

Proposition 3.4.6. *The five faces around v are all 3-faces.*

Proof. First note that, by (R1), a 5-face would send charge greater than 1 to v, which is impossible. As noted previously, a 4-face sends charge at least $\frac{3}{4}$, so v can be incident to at most one 4-face. Let the 4-face be v_5vv_1w. Now Lemma 3.4.5 implies that v_2, v_3 and v_4 are all 6^+-vertices.

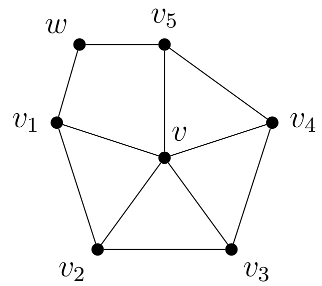


Fig. 3.4: v and its surrounding faces

Let c be a nice colouring of $G \setminus \{v\}$, which must exist by the minimality of G. We now divide into 3 cases.

1. v has a neighbouring 5-vertex.
2. All the neighbours of v are 6^+-vertices, and $|c(N(v))| = 4$.
3. All the neighbours of v are 6^+-vertices, and $|c(N(v))| = 5$.

Note that we must have $|c(N(v))| \geq 4$, otherwise there would be at most 8 colours forbidden at v and we could extend c to a nice colouring of G. Henceforth, if the edge v_iv_{i+1} is present, we will refer to the face adjacent to vv_iv_{i+1} along v_iv_{i+1} as the *external face along* v_iv_{i+1}.

Case 1: v has a neighbouring 5-vertex.
By Claim 3.4.1, v_1 and v_5 cannot both be 5-vertices, so without loss of generality, v_1 is a 5-vertex and v_5 is a 6^+-vertex. Now v_1 does not forbid a colour at v by oddness, so we must have $|c(N(v))| = 5$, otherwise there would be at most 8 colours forbidden at v.

Note that a vertex of degree $d \geq 7$ can have at most $\left\lfloor \frac{d}{2} \right\rfloor$ neighbouring blocks in the terminology of (R4). Since the total charge given out is $d - 6$, this implies that a 7-vertex gives out charge at least $\frac{1}{3}$ to every block, and an 8^+-vertex gives out charge at least $\frac{1}{2}$.

If either v_3 or v_4 is a 7^+-vertex, then by (R4) it sends charge at least $\frac{1}{3}$ to v, since v is a singleton block. This is a contradiction because v already receives charge at least $\frac{3}{4}$ from the 4-face. Hence v_3 and v_4 are both 6-vertices.

By (R3), the external face along v_3v_4 cannot be a 4^+-face, otherwise it would send charge $\frac{1}{2}$ to v. Hence it is a 3-face, and so v_3 and v_4 have a common neighbour that is not v. Now we can apply Lemma 3.4.4 to v_2, v_3, v_4, v_5, which implies that v_3 and v_4 cannot forbid distinct colours at v by oddness that are not already in $c(N(v))$. Hence at most 8 colours are forbidden at v and we can extend c to v. This completes Case 1.

Case 2: All the neighbours of v are 6^+-vertices, and $|c(N(v))| = 4$.
Since v does not have a neighbouring 5-vertex, the 4-face vv_1wv_5 gives charge 1 to v by (R2), so v cannot receive any further charge.

The external faces along v_1v_2, v_2v_3, v_3v_4 and v_4v_5 must all be 3-faces, otherwise they would give charge $\frac{1}{2}$ to v by (R3). Denote the common neighbour of v_2 and v_3 on this external 3-face by x, and the corresponding common neighbour of v_3 and v_4 by y.

All the neighbours of v must be 6-vertices, since any neighbouring 7^+-vertex would send some charge to v by (R4). Lemma 3.4.4 now applies to either

$\{v_1, v_2, v_3, v_4\}$ or $\{v_2, v_3, v_4, v_5\}$, giving a contradiction, unless $c(v_2) = c(v_4)$. Let $c(v_i) = i$ for $i = 1, 2, 3, 5$.

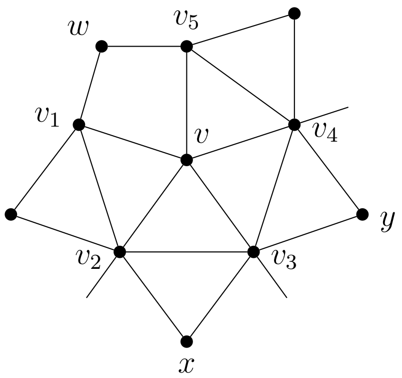


Fig. 3.5: The setting of Case 2. All edges from v, v_2, v_3 and v_4 are shown.

Consider the colours in $N(v_2) \setminus \{v\}$. v_2 must forbid a colour $2' \notin \{1, 2, 3, 5\}$ at v by oddness, so the vertices of $N(v_2) \setminus \{v\}$ must have colours $1, 1, 3, 3, 2'$ in some order. Hence $c(x) \in \{1, 2'\}$. Similarly $c(y) \in \{5, 4'\}$, where $4'$ is defined analogously to $2'$.

But now $N(v_3) \setminus \{v\}$ has two vertices of colour 3 as well as one of colour 1 or $2'$ and one of colour 5 or $4'$. There is therefore no way for v_3 to forbid a new colour $3'$ at v by oddness, and so there are at most 8 colours forbidden at v. This completes Case 2.

Case 3: All the neighbours of v are 6^+-vertices, and $|c(N(v))| = 5$.
First, note that as in Case 2, the face vv_1wv_5 sends charge 1 to v, so it cannot receive any further charge.

The external faces along v_1v_2, v_2v_3, v_3v_4 and v_4v_5 must all be 3-faces, as in Case 2. For $i = 1, 2, 3, 4$, let x_i be the common neighbour of v_i and v_{i+1} on this external 3-face. Let $c(v_i) = i$, and let the colour that v_i forbids at v by oddness be i', if it exists.

Once again, all the neighbours of v are 6-vertices, since any 7^+-vertex would send charge to v.

Lemma 3.4.4 applies to both $\{v_1, v_2, v_3, v_4\}$ and $\{v_2, v_3, v_4, v_5\}$, and there must be at least 4 colours forbidden by oddness at v. Thus the only possibility is that these are $1', 2', 4', 5'$, while v_3 does not forbid a colour by oddness at v that is not already forbidden in some other way.

The colours of the vertices of $N(v_2) \setminus \{v\}$ must be $1, 1, 3, 3, 2'$, and so $c(x_2) \in \{1, 2'\}$. Similarly $c(x_3) \in \{5, 4'\}$. Let c' be the restriction of c to $G \setminus \{v, v_2\}$. We will extend c' to a nice colouring of G.

First, assign colour 2 to v. This produces a proper partial colouring that is odd at v_4 and v_5. The colouring is also odd at v_2, because $c(v_2) = 2$ and so

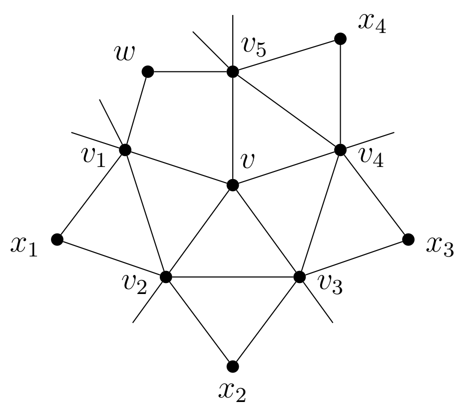


Fig. 3.6: The setting of Case 3. All edges from v and each v_i are shown.

no neighbour of v_2 other than v can have colour 2. In addition, the colours $2, 4, c(x_2), c(x_3)$ are all distinct and all already used in $N(v_3)$, so v_3 cannot forbid a colour at v_2 by oddness. Therefore there are at most 8 colours forbidden at v_2: the 4 colours of its neighbours $(1, 2, 3, 2')$, and at most 4 colours forbidden by oddness, since neither v nor v_3 forbids a colour by oddness. Hence there is a colour left over that we can use to colour v_2. This finishes the proof of Proposition 3.4.6. □

We now know that all the faces around v are 3-faces. Lemma 3.4.5 therefore implies that all the v_i are 6^+-vertices. As before, let c be a nice colouring of $G \setminus \{v\}$. Once again, we must have $|c(N(v))| \geq 4$, so we first rule out the case $|c(N(v))| = 4$.

Proposition 3.4.7. *We have* $|c(N(v))| = 5$.

Proof. Suppose that $|c(N(v))| = 4$. Without loss of generality, we have $c(v_i) = i$ for $i = 1, 2, 3, 5$ and $c(v_4) = 2$. Since all 9 colours must be forbidden at v by either properness or oddness, every v_i forbids a colour by oddness, which we call i' as before. Thus every v_i has even degree.

By Lemma 3.4.4, if v_1 and v_2 are both 6-vertices and the external face along v_1v_2 is a 3-face, then one of v_1 and v_2 does not forbid a distinct colour by oddness, which is a contradiction. Hence either the external face along v_1v_2 is a 4^+-face, giving charge $\frac{1}{2}$ to v by (R3), or one of v_1 and v_2 is an 8^+-vertex, giving charge at least $\frac{1}{2}$ to v by (R4). Similarly, either the external face along v_4v_5 is a 4^+-face or one of v_4 and v_5 is an 8^+-vertex. Together these give a total charge of at least 1 to v. This implies that v cannot receive any further charge: in particular, the external faces along v_2v_3 and v_3v_4 are 3-faces, and v_3 has degree 6. Let x be the common neighbour of v_2 and v_3 on this external face, and let y be the corresponding common neighbour of v_3 and v_4.

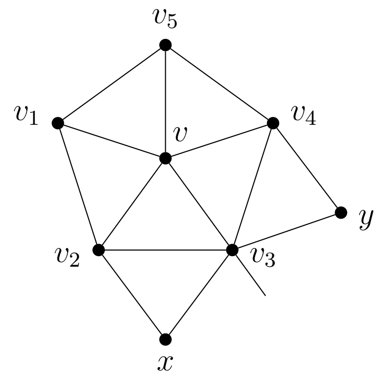


Fig. 3.7: The case $|c(N(v))| = 4$. All edges from v and v_3 are shown.

Suppose first that v_2 and v_4 are both 6-vertices. Then the colours in $N(v_2) \setminus \{v\}$ are $1, 1, 3, 3, 2'$ in some order, so $c(x) \in \{1, 2'\}$. Similarly $c(y) \in \{5, 4'\}$. But now v_3 cannot forbid a new colour $3'$ by oddness. So at least one of v_2 and v_4 must be an 8^+-vertex. Without loss of generality it is v_2.

Note that v_2 must send charge exactly $\frac{1}{2}$ to v, since otherwise v would receive total charge greater than 1. The only way this can happen is if v_2 has degree exactly 8 and the neighbours of v_2 are alternately 5-vertices and 6^+-vertices. This implies that x is a 5-vertex.

Now consider the restriction c' of c to $G \setminus \{v, v_3\}$. We will extend this to a nice colouring of G. First we assign the colour 3 to v. This produces a proper partial colouring that is odd at v_1 and v_5; it is also odd at v_3, because v must be the only neighbour of v_3 with colour 3. Since v_3 forbids a colour by oddness at v in the colouring c, we must have $|c(N(v_3) \setminus \{v\})| \leq 3$, and hence after colouring v we have used at most 4 colours in $N(v_3)$. In addition, v and x both have odd degree, so there at most 4 colours forbidden by oddness at v_3. There is therefore a colour left over to use at v_3, and we are done. □

We are now ready to finish the proof for the case $\delta(G) = 5$. Let c be a nice colouring of $G \setminus \{v\}$. Since we know $|c(N(v))| = 5$, the v_i must together forbid the remaining 4 colours by oddness. Therefore, without loss of generality, we have that for $i = 1, 2, 4, 5$, v_i forbids a colour $i' \notin \{1, 2, 3, 4, 5\}$, where all the i' are distinct. This implies that all the v_i except possibly v_3 have even degree.

If the external face along v_5v_1 is a 3-face and v_1 and v_5 are both 6-vertices, then Lemma 3.4.4 applied to $\{v_4, v_5, v_1, v_2\}$ gives a contradiction, since $1'$ and $5'$ are distinct from each other and all the other colours. Therefore either the external face along v_5v_1 is a 4^+-face, or one of v_1 and v_5 is an 8^+-vertex. In either case, charge at least $\frac{1}{2}$ is sent to v.

We can similarly apply Lemma 3.4.4 to $\{v_3, v_4, v_5, v_1\}$ and $\{v_5, v_1, v_2, v_3\}$. This shows that either the external face along v_4v_5 is a 4^+-face or one of v_4 and v_5 is an 8^+-vertex, and that either the external face along v_1v_2 is a 4^+-face or one of v_1 and v_2 is an 8^+-vertex.

In particular, if both v_1 and v_5 are 6-vertices, then this implies that each of the edges v_1v_2 and v_4v_5 must either have an external 4^+-face or contain an 8^+-vertex. Together, these send charge at least 1 to v, implying that the external face along v_1v_5 cannot be a 4^+-face. But this is a contradiction, since we have just shown that if v_1 and v_5 are both 6-vertices then the external face along v_5v_1 is a 4^+-face. Hence at least one of v_1 and v_5 is an 8^+-vertex.

Suppose without loss of generality that v_1 is an 8^+-vertex. We still require that either the external face along v_4v_5 is a 4^+-face or one of v_4 and v_5 is an 8^+-vertex. In either case, v receives total charge at least 1. As in the proof of Proposition 3.4.7, in order for the total charge to be exactly 1, we must have that v_1 has degree exactly 8, v_2 is a 6-vertex, and the external face along v_1v_2 is a 3-face. Let x be the common neighbour of v_1 and v_2 on this face. Again, as in the proof of Proposition 3.4.7, the neighbours of v_1 must alternate between 5-vertices and 6^+-vertices, so x has degree 5.

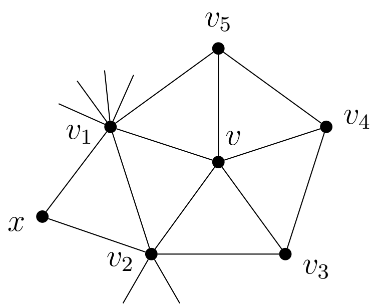


Fig. 3.8: The case $|c(N(v))| = 5$, where v_1 is an 8-vertex without loss of generality. All edges from v, v_1 and v_2 are shown.

Now we restrict c to a colouring c' of $G \setminus \{v, v_2\}$, and extend this to a nice colouring of G. First we assign colour 2 to v. This produces a proper partial colouring which is odd at v_4 and v_5. The vertices of $N(v_2) \setminus \{v\}$ have colours $1, 1, 3, 3, 2'$, so after colouring v, we have that v is the only neighbour of v_2 with colour 2, and hence the colouring is odd at v_2. In addition, there are 4 colours forbidden at v_2 by properness, and since v and x have odd degree, there are at most 4 colours forbidden by oddness. Hence there is a colour left over, which we use to colour v_2. This completes the proof for the case $\delta(G) = 5$.

3.5 The case $\delta(G) = 6$

We are now left with the case $\delta(G) = 6$, which, as discussed earlier, corresponds to the case where G is a 6-regular triangulation of the torus. Such triangulations were classified by Altshuler [4]. We will use the notation of Balachandran and Sankarnarayanan [8]. For $m, n \geq 1$ and $0 \leq t < n$, we define the graph $H = T(m, n, t)$ on vertex set $V(H) = \{(i, j) : 1 \leq i \leq m, 1 \leq j \leq n\}$ as follows:

- $(i, j) \sim (i, j + 1)$ for all i, j,
- $(i, j) \sim (i + 1, j), (i + 1, j - 1)$ for $1 \leq i < m$ and all j,
- $(m, j) \sim (1, j - t), (1, j - t - 1)$ for all j.

The addition in the second coordinate is modulo n above and throughout this section.

In other words, we begin with a grid graph of dimensions $(m + 1) \times (n + 1)$ and triangulate it. We then identify the top and bottom rows, and we identify the leftmost and rightmost columns with a shift of t vertices.

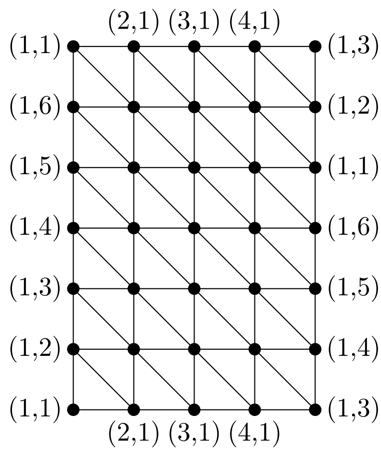


Fig. 3.9: A diagram of $T(4, 6, 4)$. The first row and column are shown twice.

Note that two graphs $T(m, n, t)$ and $T(m', n', t')$ with different parameters can be isomorphic, and that this construction does not always produce a simple graph, although we are only concerned with the cases where it does.

Theorem 3.5.1 (Altshuler). *Every 6-regular triangulation of the torus is isomorphic to $T(m, n, t)$ for some m, n, t.*

To finish the proof of Theorem 3.1.1, it therefore suffices to prove that every $T(m, n, t)$ that is a simple graph admits an odd colouring with at most 9 colours. We do not claim that this is optimal; indeed, we believe that with more care it should be possible to show that 7 colours will always suffice. We will consider three cases: $m \geq 3$, $m = 2$ and $m = 1$. Note that we must have $n \geq 3$, or $T(m, n, t)$ would not be simple.

We will find it useful to partition the 9 available colours into three classes $C_1 = \{1, 2, 3\}$, $C_2 = \{4, 5, 6\}$, $C_3 = \{7, 8, 9\}$. In diagrams, the colours of C_1 will be represented by three shades of red, with 1 being darkest and 3 lightest. Similarly, C_2 will be represented by three shades of blue, and C_3 by three shades of green.

The case $m \geq 3$

We begin by using only one colour class in each column: in column i we will use the class C_r where $r \equiv i \pmod 3$, unless $m \equiv 1 \pmod 3$, in which case we will use C_2 in column m. This ensures that the same class of colours is not used in two neighbouring columns. We call a column *bad* if its two neighbouring columns use the same class; otherwise it is *good*. If $m \equiv 0 \pmod 3$ then there are no bad columns. If $m \equiv 1 \pmod 3$ then columns 1 and $m - 1$ are bad, and if $m \equiv 2 \pmod 3$ then columns 1 and m are bad.

Next, we apply the same construction within each column: we colour (i, j) with the colour s in the correct class that satisfies $s \equiv j \pmod 3$, unless $n \equiv 1 \pmod 3$, in which case we use the colour with $s \equiv 2 \pmod 3$ at (i, n). We call the resulting colouring c; note that this initial colouring does not depend on the value of t. Again, we call a row *bad* if its two neighbouring rows use the same set of colours; otherwise it is *good*. The same classification of bad rows applies as for bad columns, according to the value of n modulo 3.

This construction clearly produces a proper colouring, since $n \geq 3$. In addition, for any vertex (i, j) with $i \neq m$ that is not in a bad column, the colours of $(i + 1, j)$ and $(i + 1, j - 1)$ appear exactly once in its neighbourhood. For $i = m$, the same is true for the colours of $(1, j - t)$ and $(1, j - t - 1)$ if column m is good. If (i, j) is in a bad column but not in a bad row, then the colours of $(i, j + 1)$ and $(i, j - 1)$ appear exactly once in its neighbourhood. Therefore the only vertices at which the colouring c may not be odd are those that are both in a bad column and a bad row. These only exist if neither m nor n is 0 (mod 3), and in that case there are exactly 4 of them. We call these *bad vertices.*

First consider the case where m and n are both 1 (mod 3). Columns 1 and $m - 1$ and rows 1 and $n - 1$ are bad. Now, for each bad vertex $v = (i, j)$, its neighbour $w = (i + 1, j - 1)$ is in a good column. Therefore the two neighbours of w in column $i + 2$ use colours from the one class that is not used at v or w.

Hence there is one colour in this class that does not appear in $c(N(w))$, and we now recolour w with that colour. We do this for all four bad vertices, creating a new colouring c'. Let S be the set consisting of the four recoloured vertices. By construction, c' is a proper colouring, since none of the four vertices of S are adjacent.

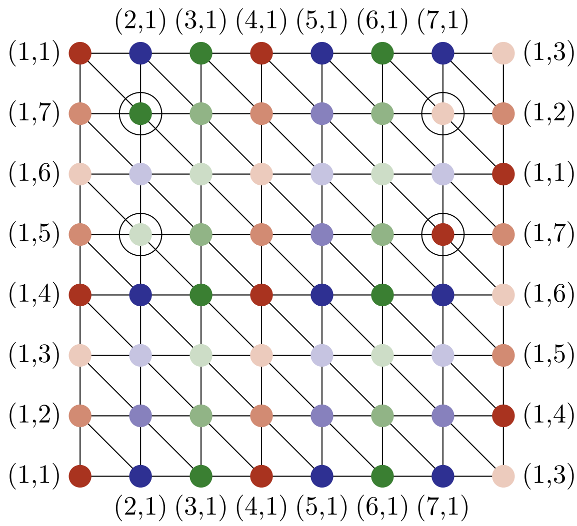


Fig. 3.10: The case $m \equiv 1 \pmod 3$, $n \equiv 1 \pmod 3$, illustrated by $T(7, 7, 5)$. The vertices of S are circled.

We now need to check that c' is an odd colouring. First, observe that c is odd at any vertex that has no neighbours in S, and therefore c' is also odd at all such vertices. Note that in c, every vertex has an even number of neighbours in each colour class C_i. Therefore any vertex v that is adjacent to exactly one vertex of S has an odd number of neighbours in some colour class, and so some colour in this class must appear an odd number of times in $N(v)$. Hence c' is odd at v.

There is no vertex adjacent to more than two vertices in S, so we are left to consider vertices with exactly two neighbours in S. There are two types of such vertices. First, there are vertices v in bad columns with one neighbour in S in each adjacent column. But then v has exactly three neighbours in its own colour class in c', so c' is odd at v. Secondly, there are vertices $v = (i, j)$ in good columns where both $(i, j+1)$ and $(i, j-1)$ are in S. But in this case column $i - 1$ is bad, and the two neighbours of v in this column have distinct colours, each of which only appears once in $N(v)$. Thus c' is an odd colouring.

Next we have the case $m \equiv 1 \pmod 3$, $n \equiv 2 \pmod 3$. The bad rows are rows 1 and n, so $u = (2, n)$ is adjacent to the two bad vertices in column 1 and $w = (m, n)$ is adjacent to the two bad vertices in column $m - 1$. We recolour each of these vertices similarly to the previous case: we assign colour 9 to u

since $9 \notin c(N(u))$, and recolour w with whichever colour in C_1 does not appear in $c(N(w))$. This produces a proper colouring c'. The proof that c' is an odd colouring proceeds just as in the $n \equiv 1 \pmod 3$ case, though it is simpler since only a vertex in a bad column can be adjacent to both u and w.

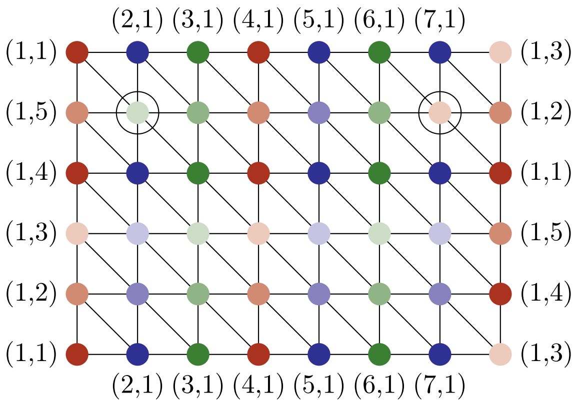


Fig. 3.11: The case $m \equiv 1 \pmod 3$, $n \equiv 2 \pmod 3$, illustrated by $T(7, 5, 3)$. The two recoloured vertices u and w are circled; u is on the left.

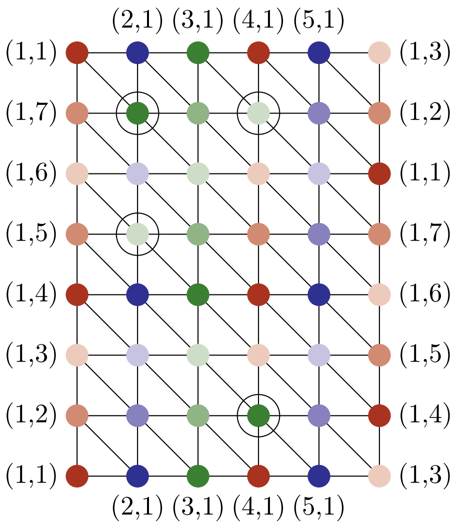


Fig. 3.12: The case $m \equiv 2 \pmod 3$, $n \equiv 1 \pmod 3$, illustrated by $T(5, 7, 5)$. The vertices of S are circled.

We now move on to $m \equiv 2 \pmod 3$, with bad columns 1 and m. Suppose that $n \equiv 1 \pmod 3$, so that rows 1 and $n-1$ are bad. We recolour vertices as follows: $c'((2, n)) = c'((m-1, 2)) = 7$, $c'((2, n-2)) = c'((m-1, n)) = 9$. Elsewhere $c' = c$. Let the set of recoloured vertices be S. This creates a proper colouring, and the proof that c' is odd runs the same as in the previous cases, but with one additional detail: if $m = 5$ then it is possible for a vertex v in column 3, which is good, to have two neighbours in S, one in column 2 and one in column 4. However, in this case, v has exactly one neighbour in colour class C_1 and one in class C_2, so c' is odd at v.

Finally, we have the case $m \equiv 2 \pmod 3$, $n \equiv 2 \pmod 3$. Now rows 1 and n are bad. We recolour $c'((2, n)) = c'((m - 1, 1)) = 9$ and otherwise leave c unchanged. Once again, c' is proper, and the proof that it is odd is the same as in the previous case. This completes the proof that $T(m, n, t)$ has a nice colouring for $m \geq 3$.

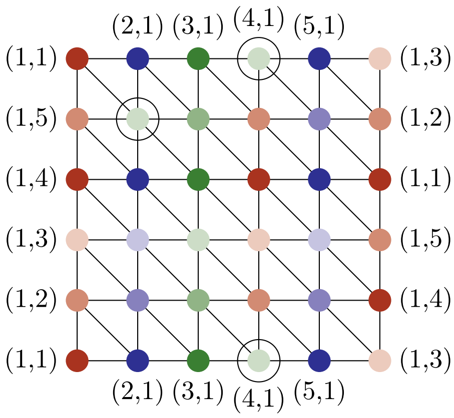


Fig. 3.13: The case $m \equiv 2 \pmod 3$, $n \equiv 2 \pmod 3$, illustrated by $T(5, 5, 3)$. The two recoloured vertices are circled; note that $(4, 1)$ is shown twice.

The case $m = 2$

We begin by colouring the vertices as in the $m \geq 3$ case: we use colours $\{1, 2, 3\}$ for column 1 and $\{4, 5, 6\}$ for column 2, and (i, j) is assigned the colour in the corresponding class that is equivalent to $j \pmod 3$, unless $n \equiv 1 \pmod 3$ in which case $(1, n)$ and $(2, n)$ receive colours 2 and 5 respectively. Let this colouring be c.

In the terminology of the $m \geq 3$ case, both columns are bad and there are at most two bad rows. If bad rows exist, then we can choose two good vertices u and w such that every bad vertex is adjacent to at least one of u and w. We then recolour by using colours 7 and 8 at u and w respectively to create a new colouring c'. This colouring is trivially proper. It is also odd, since c is odd at every vertex that is not adjacent to u or w, and for every vertex that is adjacent to u or w, either 7 or 8 appears exactly once in its neighbourhood in c'.

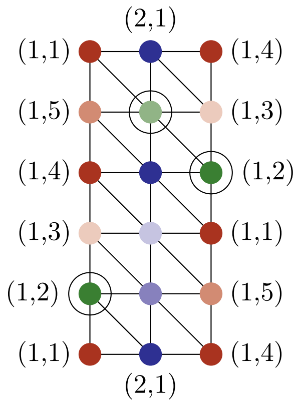


Fig. 3.14: $T(2,5,2)$ coloured as above. The two recoloured vertices are circled; note that $(1,2)$ is shown twice.

The case $m = 1$

Let $H = T(1,n,t)$. We will refer to vertex $(1,j)$ as just j for simplicity. Thus j is adjacent to $j \pm 1$, $j \pm t$ and $j \pm (t+1)$, working modulo n. We can now see that $T(1,n,t)$ is isomorphic to $T(1,n,n-(t+1))$, and so $t < \frac{n}{2}$ without loss of generality. Let $r = \lceil \frac{n}{t} \rceil$, so that by the above we have $r \geq 3$. Note also that $t \geq 2$, otherwise the graph $T(1,n,t)$ would not be simple.

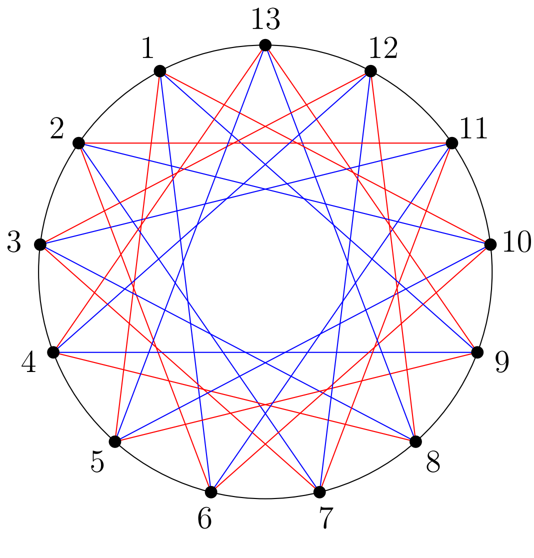


Fig. 3.15: The graph $T(1,13,4)$. Edges $j \sim (j \pm 4)$ are shown in red and edges $j \sim (j \pm 5)$ in blue.

Now we partition the vertices of $T(1,n,t)$ into intervals. For $1 \leq k \leq \lfloor \frac{n}{t} \rfloor$, let $I_k = \{j : (k-1)t+1 \leq j \leq kt\}$ so that I_k has length t. If n is not a multiple of t, then we let $I_r = \{j : kt+1 \leq j \leq n\}$ consist of the remaining vertices. There are r intervals in total.

As in the case $m \geq 3$, we split the colours into classes $C_1 = \{1, 2, 3\}$, $C_2 = \{4, 5, 6\}$ and $C_3 = \{7, 8, 9\}$. For each C_i we now define the set of vertices S_i at which C_i will be used.

- If $r \equiv 0 \pmod 3$, then $S_i = \bigcup_{k \equiv i \pmod 3} I_k$.
- If $r \equiv 1 \pmod 3$, then the S_i are defined as for the $0 \pmod 3$ case except that I_r is part of S_2 instead of S_1.
- If $r \equiv 2 \pmod 3$, then the S_i are defined as for the $0 \pmod 3$ case except that I_{r-1} is part of S_2 and I_r is part of S_3.

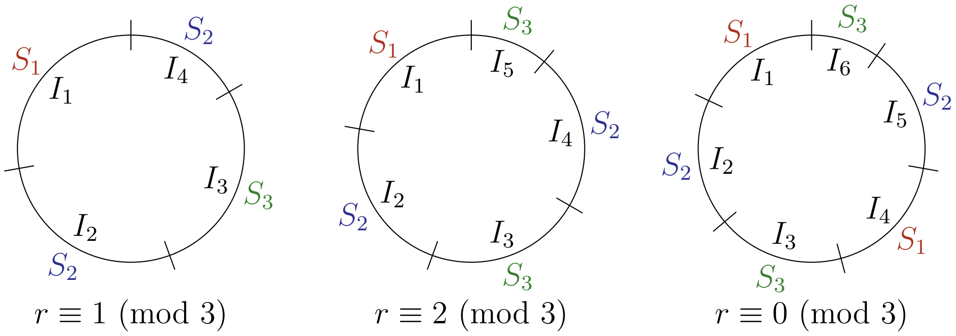


Fig. 3.16: The partition of $T(1, n, t)$ into intervals I_i and subsets S_i. As in Fig. 3.15, the graph is depicted as a circle.

We now consider the induced graphs $H[S_i]$. The construction above ensures that the only edges between two vertices in the same interval are edges of the form $\{j, j+1\}$. In addition, the only edges between two distinct intervals in the same S_i are edges of the form $\{kt, (k+1)t+1\}$ between the last vertex of one interval and the first vertex of another. This means that the induced graphs $H[S_i]$ are unions of disjoint paths, with the exception of the case $r = 4$, when $H[S_2]$ is a cycle. A path clearly has a proper odd 3-colouring, so if $r \neq 4$ we can use the colours of C_i to colour the vertices of each S_i such that a proper odd colouring is induced on $H[S_i]$. This produces a nice colouring of H.

We are left with the case $r = 4$, where $3t + 1 \leq n \leq 4t$. We begin by colouring the vertices of S_1 and S_3 using C_1 and C_3 respectively as in the previous case. Recall that S_2 consists of the union of $I_2 = \{j : t+1 \leq j \leq 2t\}$ and $I_4 = \{j : 3t+1 \leq j \leq n\}$. We properly 3-colour the cycle $H[S_2]$ and apply the resulting colouring in H, giving no regard to oddness for the time being.

For $t + 2 \leq j \leq 2t$, vertex j is adjacent to exactly two vertices in I_1, and these are adjacent to each other and thus receive different colours from C_1. The colouring is therefore odd at j for $t + 2 \leq j \leq 2t$. Since $t \geq 2$, the same argument can be applied to vertex $t+1$ and interval I_3, showing that the colouring is odd at $j=t+1$ and therefore at every vertex in I_2.

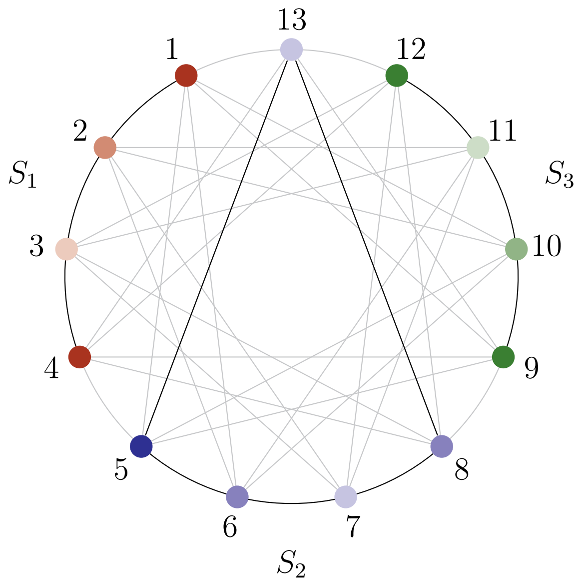


Fig. 3.17: $T(1,13,4)$ coloured as below. Edges within the same S_i are shown in black and edges between different S_i in grey.

If $n \geq 3t+2$, we can repeat this argument to show that the colouring is odd at every vertex of I_4, and therefore we have found a nice colouring of H. If instead $n = 3t+1$ then the argument breaks down at vertex $3t+1$. However, we know that the cycle $H[S_2]$ can be properly 3-coloured in such a way that there are at most two vertices at which the colouring is not odd: indeed we made heavy use of the necessary constructions in the case $m \geq 3$. Clearly we can ensure that vertex $3t+1$ is not one of these vertices. Thus the colouring of $H[S_2]$ is odd at $3t+1$, and so the resulting colouring of H is also odd at $3t+1$ and therefore at every vertex of S_2. Since we still only used colours from C_i in each S_i, the colouring is also odd at all vertices of S_1 and S_3. This completes the proof of Theorem 3.1.1.

3.6 Conclusion

First, we note that the proof of Theorem 3.1.1 also works for the real projective plane. Since the projective plane has Euler characteristic 1, Euler's formula tells us that any graph G on the projective plane has $\delta(G) \leq 5$, and that the total charge over all vertices and faces of G in the discharging process is at most -6. Thus the discharging method alone suffices, and there is no special case such as the one dealt with in Section 3.5. We did not use any other properties of the torus that do not also hold for the real projective plane: like the torus, the projective plane is locally homeomorphic to the Euclidean plane, and orientation did not matter anywhere in our discharging proof. We therefore have the following result.

Theorem 3.6.1. *Let G be a simple graph that embeds in the real projective plane. Then G has a proper odd colouring with at most 9 colours.*

For a surface S, let $\chi_o(S)$ denote the maximum value of $\chi_o(G)$ over all simple graphs that can be embedded in S. We have therefore shown that for the torus T,

$$7 \leq \chi_o(T) \leq 9.$$

We believe that in fact 7 colours will always suffice.

Conjecture 3.6.2. Let G be a toroidal graph. Then $\chi_o(G) \leq 7$.

Our main result that $\chi_o(T) \leq 9$ was later obtained independently by Tian and Yin [84], who also used the discharging method but with a different set of rules. In addition, they showed, again using discharging, that if G is a toroidal graph without 3-cycles (in other words, with girth at least 4) then $\chi_o(G) \leq 7$ [82], and that if G is a toroidal graph without two 3-cycles sharing a common edge then $\chi_o(G) \leq 8$ [83].

Improving the upper bound beyond 9 will likely require a new approach. Recall that the two proofs that 8 colours suffice for planar graphs, by Petr and Portier [67] and Fabrici, Lužar, Rindošová and Soták [36], both use the Four Colour Theorem: in fact, the bound of 8 arises specifically as twice the bound of 4 for ordinary proper colourings.

The same methods applied to the torus would only produce an upper bound of 14 colours for odd colourings. Indeed, we obtain a bound of $\chi_o(S) \leq 2\chi(S)$ for a general surface S. Any improvement on this bound would be of interest. In particular, it is natural to ask the following:

Question 3.6.3. *Does $\chi(S) = \chi_o(S)$ for every surface other than the plane?*

It seems very plausible that the answer is yes, but proving this appears to be completely out of reach with current methods.

Chapter 4

Induced subgraphs of the hypercube with small maximum degree

4.1 Introduction

This chapter is based on joint work with Victor Souza.

The hypercube graph Q_n has vertex set $\{0,1\}^n$, with two vertices being adjacent if and only if they differ in exactly one coordinate. Equivalently, the vertices can be considered as the subsets of $[n] = \{1, 2, \dots, n\}$, with two vertices adjacent if and only if their symmetric difference has size 1. Q_n is the Cartesian product of n copies of P_2.

Hypercube graphs have been widely studied: see, for example, the survey of Harary, Hayes and Wu [44]. Many of the basic properties are relatively simple to determine due to the structure of the graph; one that is related to the problem considered in this chapter is the independence number, which is clearly equal to $2^{n-1} = \frac{1}{2}|Q_n|$. Nevertheless, there are a great number of intriguing questions that can be asked about Q_n, some of which have been a focus of research for decades. For example, Ruskey and Savage [76] asked in 1993 whether, for all $n \geq 2$, every matching in Q_n extends to a Hamiltonian cycle. This problem remains open, although some progress has been made: for instance, Fink [38] showed in 2007 that every perfect matching in Q_n extends to a Hamiltonian cycle, proving a conjecture of Kreweras [54].

Subgraphs of Q_n are interesting objects of study in their own right. One such subgraph is the middle levels graph, which is the subgraph of Q_{2n+1} consisting only of the subsets of $[2n+1]$ of size n and $n+1$. Havel [47] and Buck and Wiedemann [16] conjectured in the 1980s that the middle levels graph is Hamiltonian. This became known as the middle levels conjecture, and it was proved by Mütze [64] in 2016.

Another problem on subgraphs of Q_n, more closely related to the one we consider in this chapter, is that of determining their largest eigenvalue λ_1; recall that the largest eigenvalue of a graph G is always at least the average degree and at most the maximum degree. In 2012, Fink (communicated by Bollobás, Lee and Letzter [13]) asked how large λ_1 could be for a subgraph of Q_n with exactly m vertices. Bollobás, Lee and Letzter [13] showed that for fixed i, the Hamming ball H_n^i of radius i, which is the graph induced by all subsets of $[n]$ with size at most i, gives the maximum possible value of λ_1 over all induced subgraphs of Q_n of its size when n is sufficiently large. In addition, they showed that Hamming balls of radius $o(n)$ give λ_1 within a factor $1+o(1)$ of the maximum, but the problem remains open for larger radii. Hamming balls also appear as extremal examples in a much older result: in 1966, Harper [45, 46] proved that H_n^i has the smallest *vertex-boundary* of all subsets of Q_n of size $|H_n^i|$. The vertex-boundary of $S \subset V(Q_n)$ is the set of vertices of $Q_n \setminus S$ which are adjacent to at least one vertex in S.

Subsets of Q_n can be considered as Boolean functions from $\{0,1\}^n$ to $\{0,1\}$, and so the hypercube graph is closely linked to various problems in theoretical computer science, and indeed some results originally stated in a computer-scientific setting have interpretations in the hypercube. In 1988, inspired by some of these results, Chung, Füredi, Graham and Seymour [26] considered the following question: if G is an induced subgraph of Q_n with more than 2^{n-1} vertices, at least how large must $\Delta(G)$, the maximum degree of G, be? They proved that it must be at least $\frac{1}{2}\log n - \frac{1}{2}\log\log n + \frac{1}{2}$, and found a construction that produces, for every n, an induced subgraph of Q_n with $2^{n-1}+1$ vertices and maximum degree $\lceil\sqrt{n}\rceil$.

The construction of Chung, Füredi, Graham and Seymour is as follows. We associate the vertices of Q_n with subsets of $[n]$. First, we partition $[n]$ into k intervals $I_1, \dots, I_k$, where $\left|k-\sqrt{n}\right| < 1$ and $\left||I_i|-\sqrt{n}\right| < 1$ for each i; it is easy to see that this is always possible. Next, let S be the subset of Q_n consisting of all even-sized sets that contain some I_i, $1 \le i \le k$, along with all odd-sized sets that do not contain any I_i. It is clear that the maximum degree of $Q_n[S]$, the subgraph of Q_n induced by S, is equal to $\max\{k, |I_1|, \dots, |I_k|\} = \lceil\sqrt{n}\rceil$. The same is true for the maximum degree of $Q_n[S^{\mathsf{c}}]$. Chung, Füredi, Graham and Seymour showed that $|S| = 2^{n-1} + (-1)^{n+k+1}$, and so either S or S^{c} is a set of size exactly $2^{n-1}+1$ that induces a subgraph of Q_n with maximum degree $\lceil\sqrt{n}\rceil$.

In 1992, Gotsman and Linial [42] demonstrated an equivalence between Chung, Füredi, Graham and Seymour's problem on the hypercube and concepts from computer science. More specifically, they showed that lower bounds on the maximum degree of induced subgraphs of Q_n with more than 2^{n-1} vertices imply upper bounds on the degree of a Boolean function (as a real polynomial) in terms of a measure called the *sensitivity*. For $x \in \{0,1\}^n$, denote by x^i the point derived from x by flipping the i-th coordinate. For a Boolean function $f : \{0,1\}^n \to \{0,1\}$ and $x \in \{0,1\}^n$, the *local sensitivity* of f at x, denoted $s(f,x)$, is the number of i such that $f(x^i) \neq f(x)$. The sensitivity of f, denoted $s(f)$, is the maximum of $s(f,x)$ over all x.

As Gotsman and Linial observed, their result implied that if Chung, Füredi, Graham and Seymour's upper bound of $\lceil \sqrt{n} \rceil$ was tight, then the sensitivity would be polynomially related to many other measures of complexity of a Boolean function, including the degree. Soon afterwards, Nisan and Szegedy [66] asked explicitly whether such a polynomial relation existed; this problem became known as the Sensitivity Conjecture and was one of the major open problems in computer science.

Despite the importance of the problem, progress towards the Sensitivity Conjecture was slow. This made it all the more remarkable when Huang [51] resolved Chung, Füredi, Graham and Seymour's hypercube problem in 2019 with a short and beautiful spectral argument, thereby proving the Sensitivity Conjecture:

Theorem 4.1.1 (Huang)**.** *Let G be an induced subgraph of the hypercube Q_n with at least $2^{n-1}+1$ vertices. Then $\Delta(G) \geq \sqrt{n}$, and the inequality is tight when n is a perfect square.*

Huang defined $g(n,k)$ as the smallest m such that every induced subgraph of Q_n with m vertices has maximum degree at least k. Theorem 4.1.1 states that $g(k^2,k) = 2^{k^2-1}+1$, and Huang asked whether $g(n,k)$ could be determined asymptotically in other cases. Equivalently, given m as a function of n, we would like to find the smallest possible value of $\Delta(G)$ among all induced subgraphs G of Q_n with m vertices.

In this chapter, we take some steps towards resolving this problem for induced subgraphs whose size is a constant proportion of the hypercube: in other words, $|G| = p2^n$ for constant $p \in (\frac{1}{2}, 1)$. In Section 4.2, we demonstrate a simple lower bound of $\frac{2p-1}{p}n$ for the maximum degree, which holds in a more general setting. In Section 4.3, we use Hamming codes to show that this lower bound is tight for some values of n when p is of the form $1 - \frac{1}{2^r}$. In Section 4.4 we examine what happens as n goes to infinity, and show that when $p = 1 - \frac{1}{2^r}$

the lower bound of $\frac{2p-1}{p}n$ is asymptotically tight for all n as $n \to \infty$. In Section 4.5 we find constructions that show the bound of $\frac{2p-1}{p}n$ is also asymptotically tight, in a slightly weaker sense, for all p of the form $1 - \frac{1}{r}$. In Section 4.6, we find a better lower bound than $\frac{2p-1}{p}n$ for the minimum degree of an induced subgraph when $\frac{1}{2} < p < \frac{2}{3}$. Finally, in Section 4.7, we discuss some of the many avenues for future research on this problem and others closely related to it.

4.2 Induced subgraphs of regular graphs

Let G be a graph. As usual, we denote the sets of vertices and edges of G by $V(G)$ and $E(G)$ respectively. If $S \subseteq V(G)$, then $G[S]$ denotes the subgraph of G induced by S. For subsets $S, S' \subset V(G)$, let $e(S)$ be the number of edges in $G[S]$, and let $e(S, S')$ be the number of edges that have one endpoint in S and the other in S'. The degree of a vertex $v \in V(G)$ is denoted $d(v)$. We write $d_S(v)$ for the number of neighbours of v in the set S.

Our first proposition gives a lower bound for the maximum degree of an induced subgraph of a regular graph.

Proposition 4.2.1. *Let G be a k-regular graph and $S \subset V(G)$ be a set with $|S| = p|G|$ and $e(S^{\mathsf{c}}) = \gamma e(G)$ where $p, \gamma \in [0, 1]$. Then the maximum degree of $G[S]$ satisfies*

$$\Delta(G[S]) \geq \left(\frac{2p - 1 + \gamma}{p}\right) k.$$

Moreover, we have equality if and only if $G[S]$ is regular.

Proof. First, note that

$$e(G) = e(S) + e(S, S^{\mathsf{c}}) + e(S^{\mathsf{c}}) = \tfrac{1}{2}k|G|.$$

Furthermore,

$$\sum_{v \in S^{\mathsf{c}}} d(v) = e(S, S^{\mathsf{c}}) + 2e(S^{\mathsf{c}}) = k|S^{\mathsf{c}}| = (1-p)k|G|.$$

Combining these two observations, we get

$$\begin{aligned} e(S) &= e(G) - e(S, S^{\mathsf{c}}) - e(S^{\mathsf{c}}) \\ &= e(G) + e(S^{\mathsf{c}}) + (p-1)k|G| = (p - \tfrac{1}{2} + \tfrac{1}{2}\gamma)k|G|. \end{aligned} \tag{4.1}$$

On the other hand, we can bound $e(S)$ from above using the maximum degree within $G[S]$:

$$e(S) = \frac{1}{2}\sum_{v\in S} d_S(v) \leq \frac{p|G|}{2}\Delta(G[S]). \tag{4.2}$$

Finally, relations (4.1) and (4.2) imply that

$$\Delta(G[S]) \geq \left(\frac{2p-1+\gamma}{p}\right)k,$$

as desired.

In order for equality to hold, we need equality in (4.2), so $d_S(v) = \Delta(G[S])$ for every $v \in S$; in other words, $G[S]$ must be regular. □

It will not always be possible to have control on $e(S^{\mathsf{c}})$, so we state a corollary of Proposition 4.2.1.

Corollary 4.2.2. *Let G be a k-regular graph and $S \subseteq V(G)$ be a set with $|S| = p|G|$ where $\frac{1}{2} \leq p \leq 1$. Then the maximum degree in $G[S]$ satisfies*

$$\Delta(G[S]) \geq \left(\frac{2p-1}{p}\right)k.$$

Moreover, we have equality if and only if $G[S]$ is regular and S^{c} is an independent set.

Note that if instead $p < \frac{1}{2}$, then there is no non-trivial lower bound for $\Delta(G[S])$, since the independence number of the hypercube graph Q_n is $\frac{1}{2}|Q_n|$.

For convenience, we will state the application of Corollary 4.2.2 to Q_n as a further corollary.

Corollary 4.2.3. *Let S be a subset of $V(Q_n)$ with $|S| = p2^n$. Then the maximum degree of $Q_n[S]$ satisfies*

$$\Delta(Q_n[S]) \geq \left(\frac{2p-1}{p}\right)n.$$

Equality holds if and only if $Q_n[S]$ is regular and S^{c} is an independent set.

This result gives us an initial lower bound for $\Delta(Q_n[S])$ in terms of p. We are interested in the minimum possible value of $\Delta(Q_n[S])$ among all $S \subset V(Q_n)$ with $|S| = p2^n$, and so we make the following definition:

$$\Delta_n(p) = \min_{\substack{S\subset V(Q_n)\\ |S|=p2^n}} \Delta(Q_n[S]).$$

Note that $\Delta_n(p)$ is only defined if p is of the form $\frac{m}{2^r}$ where m is an integer. We will extend the definition to all $0 \leq p \leq 1$ in Section 4.4.

Since we would like to investigate the behaviour of $\Delta_n(p)$ as $n \to \infty$, we will also define

$$D_n(p) = \frac{\Delta_n(p)}{n}.$$

With this notation, Corollary 4.2.3 states that for all n and p for which it is defined, we have

$$D_n(p) \geq \frac{2p-1}{p},$$

which is non-trivial for $\frac{1}{2} < p < 1$.

4.3 Binary codes and precise constructions

We call a subset $C \subset V(Q_n)$ of the hypercube a *binary code*. We identify the vertices of the hypercube with the finite-dimensional vector space $\mathbb{F}_2^n$, which we endow with its canonical basis $e_1, \ldots, e_n$. A *binary linear code* is a linear subspace $C \subseteq \mathbb{F}_2^n$. A code $C \subseteq \mathbb{F}_2^n$ is said to be *perfect* 1*-error-correcting* if for every $x \in C$ there is no $i \in [n]$ such that $x + e_i \in C$, and for each $x \in \mathbb{F}_2^n \setminus C$, there is precisely one $i \in [n]$ such that $x + e_i \in C$. The name comes from coding theory. If a message consists of blocks of n binary digits corresponding to elements $x \in C$, then a single transmission error in a block can be detected and corrected.

We define the *Hamming distance* between two points $x, y \in \mathbb{F}_2^n$ to be

$$d(x, y) = |\{i \in [n] : x_i \neq y_i\}| = |x \triangle y|.$$

A perfect binary 1-error-correcting code partitions the hypercube into closed balls of radius 1: if we define $B_1(x) = \{y \in \mathbb{F}_2^n : d(x, y) \leq 1\}$, then

$$\mathbb{F}_2^n = \coprod_{x \in C} B_1(x).$$

Since each $B_1(x)$ has size $n+1$, we must have $n + 1 \mid 2^n$, which implies that $n = 2^r - 1$ for some $r \geq 1$. Hamming [43] showed that such codes do exist for every r. These are the *Hamming codes*, which are linear codes $C_r \subseteq \mathbb{F}_2^{2^r-1}$ with $\dim C_r = 2^r - r - 1$.

To construct such codes, let $n = 2^r - 1$ and observe that there are n nonzero elements in $\mathbb{F}_2^r$. Let H_r be an $r \times n$ matrix with all columns distinct and nonzero; in other words, every nonzero vector in $\mathbb{F}_2^r$ appears exactly once as a column of H_r. Now we define

$$C_r = \{x \in \mathbb{F}_2^n : \ H_r x = 0\}.$$

Clearly H_r has rank r, so rank-nullity gives $\dim C_r = n - r = 2^r - r - 1$. To see that C_r is perfect 1-error-correcting, first suppose that there is some $x \in C_r$ and $i \in [n]$ such that $x + e_i \in C_r$, so

$$0 = H_r(x + e_i) = H_r e_i.$$

This is absurd, as all the columns of H_r are non-zero.

If there exists $x \notin C_r$ and distinct i and j such that $x+e_i = y$ and $x+e_j = z$ are both in C_r, then we have $y + e_i = z + e_j$. But then $H_r e_i = H_r e_j$, which is also absurd as all the columns of H_r are distinct. This implies that the balls $B_1(x)$ for $x \in C_r$ are all disjoint. It remains to show that their union is $\mathbb{F}_2^r$.

We know that $|B_1(x)| = n + 1 = 2^r$ for each x. Since $\dim C_r = 2^r - r - 1$, we have $|C_r| = 2^{2^r - r - 1}$. But now, because the balls are disjoint, the union of all the $B_1(x)$ for $x \in C_r$ has size $|C_r||B_x| = 2^{2^r - r - 1} 2^r = 2^{2^r - 1} = 2^n$, so the balls perfectly pack $\mathbb{F}_2^n$, as desired.

Note that C_r is an independent set, and that every vertex $x \notin C_r$ has exactly 1 neighbour in C_r and therefore exactly $n-1$ neighbours in ${C_r}^{\mathsf{c}}$. In other words, if we set $S = {C_r}^{\mathsf{c}} \subset V(Q_n)$, then $Q_n[S]$ is regular and S^{c} is independent. We therefore have the following:

Observation 4.3.1. *Let $n = 2^r - 1$, where $r \geq 1$ is an integer, and let $S \subset V(Q_n)$ be the complement of the Hamming code C_r. Then S achieves equality in Corollary 4.2.3, with $p = 1 - \frac{1}{2^r}$.*

The next lemma will allow us to extend this construction to higher values of n while keeping p constant.

Lemma 4.3.2. *Let S be a subset of $V(Q_n)$ with $|S| = p2^n$ and $\Delta(Q_n[S]) = qn$. Then for every integer $k \geq 1$ there exists a subset $S^{(k)} \subset V(Q_{kn})$ with $|S^{(k)}| = p2^{kn}$ and $\Delta(Q_{kn}[S^{(k)}]) = qkn$. Furthermore, if S^{c} is independent then so is ${S^{(k)}}^{\mathsf{c}}$.*

Proof. We identify Q_n with $\mathbb{F}_2^n$ as before. We define

$$S^{(k)} = \left\{(x_1, \ldots, x_k) \in (\mathbb{F}_2^n)^k : \ x_1 + \cdots + x_k \in S\right\}.$$

This is a subset of $\mathbb{F}_2^{kn}$, and we have

$$\frac{|S^{(k)}|}{2^{kn}} = \frac{|S|}{2^n} = p,$$

so $|S^{(k)}| = p2^{kn}$.

Let $x \in S^{(k)}$, with $x = (x_1, \ldots, x_k)$ as an element of $(\mathbb{F}_2^n)^k$. Let $\overline{x} = x_1 + \cdots + x_k$, so that by the definition above we have $\overline{x} \in S$. Now consider a neighbour $x + e_i$ of x, where e_i is a basis vector in $\mathbb{F}_2^{kn}$. We have $i = un + v$ for some $0 \leq u \leq k - 1$ and $1 \leq v \leq n$. The vector e_i is then of the form $(0, \ldots, \overline{e}_v, \ldots, 0)$ in $(\mathbb{F}_2^n)^k$, where $\overline{e}_v$, a basis vector in $\mathbb{F}_2^n$, is in the $(u + 1)$-th position.

We now have that $x + e_i \in S^{(k)}$ if and only if $\overline{x} + \overline{e}_v \in S$. There are $d_S(\overline{x})$ values of v for which this holds, and k values of i corresponding to each v. Therefore x has precisely $kd_S(\overline{x})$ neighbours in $S^{(k)}$. Note that the mapping from x to $\overline{x}$ is surjective onto $\mathbb{F}_2^n$, so taking the maximum over all $x \in S^{(k)}$, we obtain that

$$\Delta(Q_{kn}[S^{(k)}]) = qkn$$

as desired.

Now suppose that S^{c} is independent. To see that ${S^{(k)}}^{\mathsf{c}}$ is also independent, suppose that we have $x \in {S^{(k)}}^{\mathsf{c}}$ and $i = un + v$ such that $x + e_i \in {S^{(k)}}^{\mathsf{c}}$. Then $\overline{x}$ and $\overline{x} + \overline{e}_v$ are both in S^{c}, which is a contradiction. $\square$

This lemma gives us some more cases where equality can be achieved in Corollary 4.2.3.

Corollary 4.3.3. *Let $n = k(2^r - 1)$, where $k, r \geq 1$ are integers, and let $p = 1 - \frac{1}{2^r}$. Then there exists $S \subset V(Q_n)$ with $|S| = p2^n$ and $\Delta(Q_n[S]) = \frac{2p-1}{p}n$.*

Proof. Apply Lemma 4.3.2 to the construction in Observation 4.3.1. $\square$

While the construction in Corollary 4.3.3 may be new, we would not be surprised if it is already known in the context of error-correcting codes.

4.4 Behaviour in the limit

In the previous section, we have shown that the lower bound of $D_n(p) = \frac{2p-1}{p}$ can be achieved when $p = 1 - \frac{1}{2^r}$ and n is of the form $k(2^r - 1)$. It is natural, then, to ask how $D_n(p)$ behaves for the same value of p when n is not a multiple of $2^r - 1$. In particular, we would like to know whether a limit exists as $n \to \infty$, not just in the case $p = 1 - \frac{1}{2^r}$ but for general values of p. The following proposition tells us that it does.

Proposition 4.4.1. *Let p be a rational of the form $\frac{b}{2^r}$ for some integers $b, r \geq 0$, so that $D_n(p)$ is defined for $n \geq r$. Then*

$$D_n(p) \to \inf_{n \geq r} D_n(p)$$

as $n \to \infty$.

Proof. Let $a = \inf_{n \geq r} D_n(p)$. It suffices to show that for every $\epsilon > 0$ there exists N such that $D_n(p) < a + \epsilon$ for all $n \geq N$.

By the definition of a, there exists some m such that $D_m(p) < a + \frac{\epsilon}{2}$. In other words, there exists $S \subset Q_m$ such that $|S| = p2^m$ and $\Delta(Q_m[S]) < (a + \frac{\epsilon}{2})m$. Now, by Lemma 4.3.2, for every $k \geq 1$ there exists $S^{(k)} \subset Q_{km}$ with $|S^{(k)}| = p2^{km}$ and $\Delta(Q_{km}[S^{(k)}]) < (a + \frac{\epsilon}{2})km$.

Now consider a general value of n. Let us write $n = km - r$ for some integers q and $0 \leq r \leq m - 1$. The hypercube Q_{km} can be partitioned into 2^r subcubes isomorphic to Q_n, each corresponding to a particular choice of the first r coordinates. Let Q' be one of these subcubes, and consider the set $S' = S^{(k)} \cap Q'$. Since $|S^{(k)}| = p|Q_{km}|$, we can choose Q' such that $|S'| \geq p|Q'|$. For any vertex $x \in S'$, we clearly have $d_{S'}(x) \leq d_{S^{(k)}}(x)$. Therefore

$$\Delta(Q'[S']) \leq \Delta(Q_{km}[S^{(k)}]) < \left(a + \frac{\epsilon}{2}\right) km.$$

Identifying Q' with Q_n, we now have $S' \subset V(Q_n)$ with

$$\begin{aligned} \Delta(Q_n[S']) < \left(a + \frac{\epsilon}{2}\right) km &= \left(a + \frac{\epsilon}{2}\right)(n + r) \\ &= \left(a + \frac{\epsilon}{2} + \frac{r}{n}\left(a + \frac{\epsilon}{2}\right)\right) n. \end{aligned}$$

If

$$n \geq \frac{2m}{\epsilon}\left(a + \frac{\epsilon}{2}\right),$$

then we have

$$\frac{r}{n}\left(a + \frac{\epsilon}{2}\right) < \frac{m}{n}\left(a + \frac{\epsilon}{2}\right) \leq \frac{\epsilon}{2},$$

and therefore

$$\Delta(Q_n[S']) < (a + \epsilon)n.$$

We know that $|S'| \geq p2^n$, so we can remove vertices from S' to get $S'' \subset V(Q_n)$ with $|S''| = p2^n$ and

$$\Delta(Q_n[S'']) < (a + \epsilon)n.$$

Therefore $D_n(p) < a + \epsilon$ for large enough n, and so $D_n(p) \to a$ as $n \to \infty$, as claimed. $\square$

For a rational p of the form $\frac{m}{2^r}$, we can now define

$$D(p) = \lim_{n\to\infty} D_n(p).$$

Corollary 4.2.3 implies that

$$D(p) \geq \frac{2p-1}{p},$$

and together with Corollary 4.3.3 and Proposition 4.4.1 this tells us that for $p = 1 - \frac{1}{2^r}$, we have

$$D(p) = \frac{2p-1}{p} = 1 - \frac{1}{2^r - 1}.$$

We would like to extend the definition of $D(p)$ to all $p \in [0,1]$. For p not of the form $\frac{m}{2^r}$, we will define

$$D(p) = \lim_{q\to p^+} D(q) = \lim_{n\to\infty} \left(\frac{1}{n} \min_{\substack{S\subset V(Q_n)\\ |S|\geq p2^n}} \Delta(Q_n[S]) \right),$$

where the limit on the left is taken over q of the form $\frac{m}{2^r}$. The function $D(q)$ is clearly increasing, so the limit on the left exists.

That the limit on the right exists and is equal to the limit on the left can be seen as follows. The minimum on the right is always attained by a set S of size $\lceil p2^n \rceil$, and for any $q > p$ this is smaller than $q2^n$ for large enough n. Therefore, for every $q > p$ of the form $\frac{m}{2^r}$, the quantity inside the limit on the right is bounded above by $D_n(q)$ for large enough n, and so the limit itself, if it exists, is bounded above by $D(q)$. But for any given n, the quantity inside the limit is bounded below by $D(\lceil p2^n \rceil / 2^n)$, which is at least $\lim_{q\to p^+} D(q)$. This implies that the limit on the right exists and is equal to $\lim_{q\to p^+} D(q)$, as desired.

Note that we clearly still have

$$D(p) \geq \frac{2p-1}{p}$$

by Corollary 4.2.3.

As we will see in Section 4.5, for $p = 1 - \frac{1}{r}$ it turns out to be possible to control $\Delta(Q_n[S_n])$ for certain sets S_n such that $|S_n|/2^n \to p$ as $n \to \infty$, but $p2^n - |S_n|$ grows without bound. This allows us to bound $D(q)$ above for all $q < p$, but it does not tell us anything about $D(p)$ itself. It will therefore be useful for us to make another definition: we will set

$$D^-(p) = \lim_{q \to p^-} D(q).$$

As before, the limit above exists since $D(q)$ is increasing. In addition, we still have the bound from Corollary 4.2.3:

$$D^-(p) \geq \frac{2p-1}{p}.$$

Note that knowing $D^-(p)$ does not give us any information about the behaviour of

$$\min_{\substack{S \subset V(Q_n) \\ |S|=(p-o(1))2^n}} \Delta(Q_n[S]).$$

In particular, it does not say anything about how small $\Delta(Q_n[S])$ can be for sets S of size $\lfloor p2^n \rfloor$ or sets whose size is within a constant of $p2^n$.

4.5 Tilings of $\mathbb{Z}^m$ and asymptotic constructions

In this section, we will prove the following result.

Theorem 4.5.1. *Let $p = 1 - \frac{1}{r}$, where $r \geq 2$ is an integer. Then*

$$D^-(p) = \frac{2p-1}{p} = 1 - \frac{1}{r-1}.$$

To do this, we will construct induced subgraphs of Q_n of size $(p - o(1))2^n$ which have maximum degree $\left(\frac{2p-1}{p} + o(1)\right) n$, so that in some sense equality in Corollary 4.2.3 holds asymptotically for $p = 1 - \frac{1}{r}$.

We will make use of Chernoff's standard concentration inequality for binomial random variables [25]: see, for instance, Mitzenmacher and Upfal [62] for the simple form given here.

Proposition 4.5.2 (Chernoff bound)**.** *Let $X \sim \mathrm{Bin}(n, \frac{1}{2})$. Then*

$$\mathbb{P}\left(\left|X - \frac{n}{2}\right| \geq t\right) \leq 2\exp\left(-\frac{2t^2}{3n}\right).$$

We will construct our induced subgraphs of Q_n from various tilings of $\mathbb{Z}^m$. We consider the usual graph structure on $\mathbb{Z}^m$, where two vertices are connected by an edge when they differ only in one coordinate and exactly by one. Also, we denote by e_i the standard basis on $\mathbb{Z}^m$. We say that $\mathcal{T} \subseteq \mathbb{Z}^m$ is a *k-covering t-tiling* if all of the following hold:

1. $\mathcal{T}$ is an independent set in $\mathbb{Z}^m$,
2. every vertex $v \in \mathbb{Z}^m \setminus \mathcal{T}$ has exactly k neighbours in $\mathcal{T}$,
3. there are t distinct translations $\mathcal{T}^{(i)}$ of $\mathcal{T}$ such that $\mathbb{Z}^m = \mathcal{T}^{(1)} \cup \cdots \cup \mathcal{T}^{(t)}$ is a partition.

Next, we construct some families of tilings.

Proposition 4.5.3. *For each $m \geq 1$, there is a 1-covering $(2m+1)$-tiling of $\mathbb{Z}^m$ and a 2-covering $(m+1)$-tiling of $\mathbb{Z}^m$.*

Proof. We will show that the following sets are such tilings:

$$\mathcal{T}_m = \{(y_1, \ldots, y_m) \in \mathbb{Z}^m : \ y_1 + 2y_2 + \cdots + my_m \equiv 0 \pmod{2m+1}\},$$
$$\mathcal{T}'_m = \{(y_1, \ldots, y_m) \in \mathbb{Z}^m : \ y_1 + 2y_2 + \cdots + my_m \equiv 0 \pmod{m+1}\}.$$

First, these sets are indeed independent, since changing any coordinate by one breaks the congruence condition.

Consider $x = (x_1, \ldots, x_m) \in \mathbb{Z}^m \setminus \mathcal{T}_m$. We have

$$x_1 + 2x_2 + \cdots + mx_m \equiv k \pmod{2m+1}$$

for some $k \in \{\pm 1, \pm 2, \ldots, \pm m\}$. Suppose $k = \pm i$. Then x has precisely one neighbour in $\mathcal{T}_m$, namely $x \mp e_i$. It is also clear that the translations $\mathcal{T}_m + ke_1$ are all distinct and disjoint for $0 \leq k \leq 2m$ and that they cover $\mathbb{Z}^m$: indeed, they correspond to the sets

$$\{(y_1, \ldots, y_m) \in \mathbb{Z}^m : \ y_1 + 2y_2 + \cdots + my_m \equiv k \pmod{2m+1}\}.$$

Hence $\mathcal{T}_m$ is a 1-covering $(2m+1)$-tiling of $\mathbb{Z}^m$.

Now consider $x = (x_1, \ldots, x_m) \in \mathbb{Z}^m \setminus \mathcal{T}'_m$. We have

$$x_1 + 2x_2 + \cdots + mx_m \equiv k \pmod{m+1}$$

for some $k \in \{1, 2, \ldots, m\}$. Then x has precisely two neighbours in $\mathcal{T}'_m$, namely $x - e_k$ and $x + e_{m+1-k}$. Again, the translations $\mathcal{T}'_m + ke_1$ are all distinct and disjoint for $0 \leq k \leq m$, and furthermore, they cover $\mathbb{Z}^m$. Therefore $\mathcal{T}'_m$ is a 2-covering $(m+1)$-tiling of $\mathbb{Z}^m$. $\square$

The 1-covering $(2m+1)$-tiling $\mathcal{T}_m$ has been discovered previously in a different context: it was used by Golomb and Welch [41] to construct error-correcting codes in the Lee metric. We believe that the 2-covering $(m+1)$-tiling $\mathcal{T}'_m$ is new.

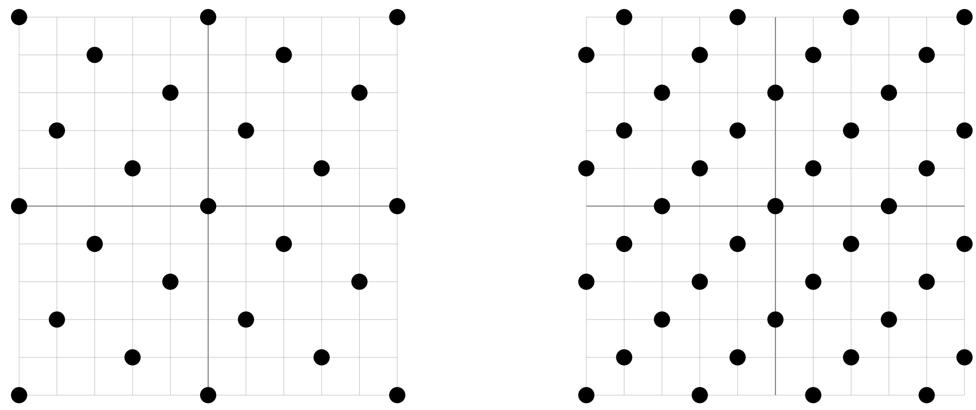


Fig. 4.1: Two tilings of $\mathbb{Z}^2$: the 1-covering 5-tiling $\mathcal{T}_2$ and the 2-covering 3-tiling $\mathcal{T}_2'$

Theorem 4.5.4. *Suppose that there exists a k-covering t-tiling of $\mathbb{Z}^m$. Then for every n there is a subset $S \subset V(Q_n)$ with $|S| = \left(1 - \frac{1}{t} - O\left(\frac{1}{n}\right)\right) 2^n$ and*

$$\Delta(Q_n[S]) = \left(1 - \frac{k}{2m} + O\left(\sqrt{\frac{\log n}{n}}\right)\right) n.$$

Proof. We identify the vertices of Q_n with $\mathcal{P}([n])$. Divide the ground set $[n]$ into m subsets $I_1, \dots, I_m$ with sizes as equal as possible, that is, with $|I_i| = \left\lfloor \frac{n}{m} \right\rfloor$ or $\left\lfloor \frac{n}{m} \right\rfloor + 1$ for every i. The *profile* of a vertex $x \in Q_n$ is defined by

$$v(x) = (|x \cap I_1|, \dots, |x \cap I_m|) \in \mathbb{Z}^m.$$

First, observe that if x and y are adjacent in Q_n, then $v(x)$ and $v(y)$ are adjacent in $\mathbb{Z}^m$. For a neighbour y of x, we have $v(y) = v(x) + e_i$ if and only if $y = x \cup j$ for some $j \in I_i \setminus x$, and $v(y) = v(x) - e_i$ if and only if $y = x \setminus j$ for some $j \in x \cap I_i$. Therefore x has $|I_i| - |x \cap I_i|$ neighbours with $v(y) = v(x) + e_i$ and $|x \cap I_i|$ neighbours with $v(y) = v(x) - e_i$.

In order to control the degrees, we would like to ensure that for each of the $2m$ choices of $\pm e_i$, the number of neighbours of x with profile $v(x) \pm e_i$ is roughly the same. To do so, we restrict Q_n to vertices where $v(x)$ is typical. Indeed, let $c_i = \sqrt{\frac{3}{2}|I_i| \log n}$ and define the set

$$B_n = \left\{ x \in Q_n : \left| |x \cap I_i| - \tfrac{1}{2}|I_i| \right| \le c_i \text{ for all } i \in [m] \right\}.$$

Let $x \in Q_n$ be chosen uniformly at random. Then B_n consists of all the vertices in Q_n such that $|x \cap I_i|$ is close to its expected value for every i. We can use the Chernoff bound (Proposition 4.5.2) to give a lower bound for $|B_n|$. Indeed, note that for each $i \in [m]$, $|x \cap I_i|$ is a $\mathrm{Bin}\big(|I_i|, \frac{1}{2}\big)$ random variable, and they are all independent. Hence,

$$\begin{aligned}\frac{|B_n{}^{\mathsf{c}}|}{2^n} &= \mathbb{P}\Big(\exists i \in [m] : \ \big||x \cap I_i| - \tfrac{1}{2}|I_i|\big| > c_i\Big)\\ &\leq \sum_{i\in[m]} \mathbb{P}\Big(\big||x \cap I_i| - \tfrac{1}{2}|I_i|\big| \geq c_i\Big)\\ &\leq 2\sum_{i\in[m]} \exp\left(\frac{-2c_i^2}{3|I_i|}\right)\\ &\leq \frac{2m}{n}.\end{aligned}$$

Therefore, we have

$$|B_n| \geq \left(1 - O\left(\frac{1}{n}\right)\right)2^n.$$

Given a vertex $x \in B_n$ and $u \in \mathbb{Z}^m$, denote by $d_u(x)$ the number of vertices $y \in Q_n$ that are neighbours of x with $v(y) = v(x) + u$. Hence $d_u(x) = 0$ unless $u = \pm e_i$ for some $i \in [m]$, and in this case we have the bounds

$$\begin{aligned} d_{+e_i}(x) &= |I_i| - |x \cap I_i| \geq \frac{1}{2}|I_i| - c_i \geq \frac{n}{2m} - O\Big(\sqrt{n \log n}\Big),\\ d_{-e_i}(x) &= |x \cap I_i| \geq \frac{1}{2}|I_i| - c_i \geq \frac{n}{2m} - O\Big(\sqrt{n \log n}\Big).\end{aligned}$$

Now let $\mathcal{T}$ be a k-covering t-tiling of $\mathbb{Z}^m$ with translates $\mathcal{T}^{(1)}, \ldots, \mathcal{T}^{(t)}$ that partition $\mathbb{Z}^m$. For each $i \in [t]$, define

$$S_i = \left\{x \in B_n : \ v(x) \notin \mathcal{T}^{(i)}\right\}.$$

We then have

$$\sum_{i\in[t]} |S_i| = (t-1)|B_n|,$$

and so we can choose $j \in [t]$ so that $S = S_j$ satisfies

$$|S| = |S_j| \geq \left(1 - \frac{1}{t}\right)|B_n| \geq \left(1 - \frac{1}{t} - O\left(\frac{1}{n}\right)\right)2^n.$$

We now need to check that the maximum degree of $Q_n[S]$ is not too large. For $x \in S$, we have $v(x) \notin \mathcal{T}^{(j)}$, and since $\mathcal{T}^{(j)}$ is a k-covering t-tiling, there are exactly k neighbours of $v(x)$ in $\mathcal{T}^{(j)}$, say $v(x) + u_1, \ldots, v(x) + u_k$, where $u_1, \ldots, u_k \in \{\pm e_1, \ldots, \pm e_m\} \subset \mathbb{Z}^m$. It follows that all vertices $y \in Q_n$ with

$v(y) = v(x) + u_i$ are not in S. Therefore we have

$$d_S(x) \le n - \sum_{i\in[k]} d_{u_i}(x) \le n - \frac{kn}{2m} + O\Big(\sqrt{n\log n}\Big),$$

where we use our earlier bounds on $d_{\pm e_i}(x)$. In other words, we have

$$\Delta(Q_n[S]) \le \left(1 - \frac{k}{2m} + O\left(\sqrt{\frac{\log n}{n}}\right)\right) n,$$

as we wanted. □

We note that in the construction above, most vertices $x \in S$ do not have any neighbours outside B_n. This allows us to bound $d_S(x)$ from below as well as from above using the same argument as in the proof of Theorem 4.5.4: for such an x we have

$$\left| d_S(x) - \left(1 - \frac{k}{2m}\right) n \right| = O\left(\sqrt{n\log n}\right).$$

However, it is possible for a vertex $x \in S$ to have many neighbours outside B_n. In the worst case, $\big||x \cap I_i| - \frac{1}{2}|I_i|\big| = \lfloor c_i \rfloor$ for every $i \in [m]$, and around half of the neighbours of x are not in B_n. As a result, the minimum degree of $Q_n[S]$ is much smaller than the maximum degree, and we obtain the following lower bound:

$$\delta(Q_n[S]) \ge (m-k)\left(\frac{n}{2m} + O\left(\sqrt{n\log n}\right)\right) = \left(\frac{m-k}{2m} + O\left(\sqrt{\frac{\log n}{n}}\right)\right) n.$$

We can now finish the proof of Theorem 4.5.1, establishing that equality holds in Corollary 4.2.3 in an asymptotic sense for $p = 1 - \frac{1}{r}$.

Proof of Theorem 4.5.1. Let $p = 1 - \frac{1}{r}$. Proposition 4.5.3 tells us that there exists a 2-covering r-tiling $\mathcal{T}'_{r-1}$ of $\mathbb{Z}^{r-1}$. Now, by Theorem 4.5.4, there exists $S \subset V(Q_n)$ with

$$|S| = \left(1 - \frac{1}{r} - O\left(\frac{1}{n}\right)\right) 2^n$$

and

$$\Delta(Q_n[S]) = \left(1 - \frac{1}{r-1} + O\left(\sqrt{\frac{\log n}{n}}\right)\right) n.$$

Taking $n \to \infty$, we see that for every $\epsilon > 0$, $D(p-\epsilon) \le 1 - \frac{1}{r-1}$. Hence

$$D^{-}\left(1 - \frac{1}{r}\right) = 1 - \frac{1}{r-1}$$

as claimed. $\square$

Note that we have not used the 1-covering $(2m+1)$-tiling $\mathcal{T}_m$ of $\mathbb{Z}^m$ constructed in Proposition 4.5.3. However, the constructions from $\mathcal{T}_m$ and $\mathcal{T}'_{2m}$ are essentially the same. Recall that

$$\mathcal{T}'_{2m} = \left\{(y_1, \ldots, y_{2m}) \in \mathbb{Z}^{2m} : \ y_1 + 2y_2 + \cdots + 2my_{2m} \equiv 0 \pmod{2m+1}\right\}.$$

As a subset of $V(\mathbb{Z}^{2m})$, this is clearly isomorphic to

$$\begin{aligned} &\big\{(y_1, \ldots, y_{2m}) \in \mathbb{Z}^{2m} : \ y_1 + 2y_2 + \cdots + my_m \\ &\qquad - (m+1)y_{m+1} - \cdots - 2my_{2m} \equiv 0 \pmod{2m+1}\big\} \\ = &\big\{(y_1, \ldots, y_{2m}) \in \mathbb{Z}^{2m} : \ (y_1 + y_{2m}) + 2(y_2 + y_{2m-1}) \\ &\qquad + \cdots + m(y_m + y_{m+1}) \equiv 0 \pmod{2m+1}\big\}. \end{aligned}$$

Call this set $\mathcal{T}''_{2m}$. Now construct sets S from $\mathcal{T}_m$ and S' from $\mathcal{T}''_{2m}$ as in Theorem 4.5.4. Let $[n]$ be divided into subsets $I_1, \ldots, I_m$ when constructing S, and $I'_1, \ldots, I'_{2m}$ when constructing S'. Call the resulting profile functions v and v' respectively. Now if we choose the subsets so that $I_i = I'_i \cup I'_{2m+1-i}$ for each $i \in [m]$, then for $x \in Q_n$ we have $v(x) \in \mathcal{T}_m$ if and only if $v'(x) \in \mathcal{T}'_{2m}$.

This does not mean that S and S' are identical, as they could arise from different translates of $\mathcal{T}_m$ and $\mathcal{T}''_{2m}$, and the two sets B_n are not the same. However, the overall structure of S and S', and the environment around a typical vertex, are very similar.

Thus, while the 1-covering tilings are not strictly necessary for our proof, they give a slightly simpler description of the set S when p is of the form $1 - \frac{1}{r}$ with r odd. For example, when $p = \frac{2}{3}$, S is simply

$$\left\{x \in Q_n : \ |x| \not\equiv i \pmod 3 \text{ and } |x - \tfrac{1}{2}n| \le \sqrt{\tfrac{3}{2}n \log n}\right\}$$

for some $i \in \{0, 1, 2\}$.

4.6 Better lower bounds

In this section, we will show that for $\frac{1}{2} < p < \frac{2}{3}$, we have

$$D(p) > \frac{2p-1}{p},$$

that is, equality cannot hold in Corollary 4.2.3 for these values of p. By examining the structure of S in increasing detail, we will obtain a series of successively better lower bounds for $D(p)$ when $\frac{1}{2} < p < \frac{2}{3}$. We begin with the simplest of these bounds.

Proposition 4.6.1. *Let* $\frac{1}{2} < p < \frac{2}{3}$. *Then*

$$D(p) \geq \sqrt{\frac{2p-1}{2p}}.$$

Remark. Note that

$$\sqrt{\frac{2p-1}{2p}} > \frac{2p-1}{p}$$

precisely when $\frac{1}{2} < p < \frac{2}{3}$, so Proposition 4.6.1 improves on the lower bound in Corollary 4.2.3 for this range of p. The bound is still valid for $p > \frac{2}{3}$, but does not give any new information there.

Proof of Proposition 4.6.1. It suffices to prove the proposition for p of the form $\frac{m}{2^r}$. Let $S \subset V(Q_n)$ with $|S| = p2^n$ where $\frac{1}{2} < p < \frac{2}{3}$. Let $\Delta(Q_n[S]) = qn$. For convenience, for $u \in S$, we denote $\frac{d_S(u)}{n}$ by δ_u, so that $\delta_u \leq q$ for all $u \in S$.

Fix a vertex $v \in S$. Let C be the set of vertices of Q_n at distance exactly 2 from v, that is, the vertices that differ from v in 2 coordinates. Each $w \in C$ has exactly 2 common neighbours with v. We partition C into three subsets C_1, C_2, C_3 as follows:

- $w \in C_1$ if both common neighbours of v and w are in S,
- $w \in C_2$ if v and w have one common neighbour in S and one in S^c,
- $w \in C_3$ if both common neighbours of v and w are in S^c.

Each vertex in $N(v) \cap S$ has at most $qn - 1$ neighbours in S apart from v. Therefore at most $d_S(v)(qn-1)$ vertices of C_2 are in S. Hence

$$\begin{aligned}|C_2 \cap S^c| &\geq d_S(v)(n - d_S(v)) - d_S(v)(qn-1)\\ &= d_S(v)((1-q)n - d_S(v) + 1)\\ &= \delta_v n((1-q-\delta_v)n + 1).\end{aligned}$$

Note that this already implies that $q > \frac{2p-1}{p}$. Indeed, if not, then $\delta_v \le q < \frac{1}{2}$, and so $|C_2 \cap S^{\mathsf{c}}| > 0$ as long as we choose a vertex v with $d_S(v) > 0$. But now every vertex in $C_2 \cap S^{\mathsf{c}}$ has a neighbour in $N(v) \cap S^{\mathsf{c}}$, so $e(S^{\mathsf{c}}) > 0$. Thus S^{c} is not an independent set, so equality cannot hold in Corollary 4.2.3.

Each vertex in $C_2 \cap S^{\mathsf{c}}$ has exactly one neighbour in $N(v) \cap S^{\mathsf{c}}$. For each edge $e = xy$ within S^{c}, there are at most $2(n-2)$ choices of v for which e can be found in this way, namely the neighbours of x and y that are in S. By double-counting, we therefore have

$$\begin{aligned} e(S^{\mathsf{c}}) &\ge \frac{1}{2n} \sum_{v \in S} \delta_v n((1-q-\delta_v)n + 1) \\ &\ge \frac{1}{2} \sum_{v \in S} \delta_v n(1-q) - \frac{n}{2} \sum_{v \in S} \delta_v^2 \\ &= (1-q)e(S) - \frac{n}{2} \sum_{v \in S} \delta_v^2 \end{aligned}$$

From the proof of Proposition 4.2.1, we have that

$$e(S) + e(S, S^{\mathsf{c}}) + e(S^{\mathsf{c}}) = n2^{n-1}$$

and

$$e(S, S^{\mathsf{c}}) + 2e(S^{\mathsf{c}}) = (1-p)n2^n.$$

Combining these, we obtain

$$e(S) - e(S^{\mathsf{c}}) = (2p-1)n2^{n-1}, \tag{4.3}$$

and so

$$e(S) \ge (2p-1)n2^{n-1} + (1-q)e(S) - \frac{n}{2} \sum_{v \in S} \delta_v^2.$$

Therefore

$$\begin{aligned} qe(S) &\ge (2p-1)n2^{n-1} - \frac{n}{2} \sum_{v \in S} \delta_v^2 \\ \frac{q}{2} \sum_{v \in S} d_S(v) &\ge (2p-1)n2^{n-1} - \frac{n}{2} \sum_{v \in S} \delta_v^2 \\ \sum_{v \in S} \delta_v(q + \delta_v) &\ge (2p-1)2^n. \end{aligned}$$

But $\delta_v \le q$ for all $v \in S$, so

$$\sum_{v \in S} \delta_v(q + \delta_v) \le 2q^2 p 2^n.$$

It follows that

$$q \geq \sqrt{\frac{2p-1}{2p}},$$

which implies the desired result. $\square$

For the remainder of this section, we will often be concerned with the number of subgraphs $H \subset V(Q_n)$ isomorphic to some fixed graph F, some of whose vertices have been marked, such that the isomorphism between H and F sends the vertices of $H \cap S$ to the marked vertices of F. We will represent this number of subgraphs with a small diagram depicting the graph F with the marked vertices in black and the unmarked vertices in white.

For example, we may wish to consider the number of subgraphs of Q_n isomorphic to $C_4 \equiv Q_2$ which contain exactly 2 vertices of S in the configuration where these 2 vertices are adjacent. This number is denoted by the symbol .

By carefully counting configurations of 3 and 4 vertices such as the one above, we will prove the following result.

Proposition 4.6.2. *Let* $\frac{1}{2} < p < \frac{2}{3}$. *Then*

$$D(p) \geq \frac{1 - 2p + \sqrt{(2p-1)(1-p)}}{2 - 3p}.$$

Proof. As in the previous proposition, we may assume p is of the form $\frac{m}{2^r}$. Let $S \subset V(Q_n)$ be such that $|S| = p2^n$ with $\frac{1}{2} < p < \frac{2}{3}$, and let $\Delta(Q_n[S]) = qn$.

We have that

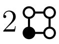
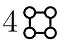
$$e(S^{\mathsf{c}}) = \frac{1}{2} \sum_{v \in S^{\mathsf{c}}} d_{S^{\mathsf{c}}}(v).$$

Also, we have


$$= \frac{1}{2} \sum_{v \in S^{\mathsf{c}}} d_{S^{\mathsf{c}}}(v)(d_{S^{\mathsf{c}}}(v) - 1) = \frac{1}{2} \sum_{v \in S^{\mathsf{c}}} d_{S^{\mathsf{c}}}(v)^2 - e(S^{\mathsf{c}})$$
$$\geq \frac{2}{(1-p)2^n} e(S^{\mathsf{c}})^2 - e(S^{\mathsf{c}})$$

by the Cauchy-Schwarz inequality.

Consider an edge in S^{c}. It is part of $n-1$ different copies of Q_2, so by considering the possible arrangements of vertices of S in such a Q_2 and double-counting, we obtain that

$$e(S^{\mathsf{c}}) = \frac{1}{n-1} \Big($$  $$+ 2$$  $$+ 4$$  $$\Big). \quad (4.4)$$

Every copy of P_3 in Q_n is contained within exactly one copy of Q_2, so we have that

$$= + 4 ,$$

and hence

$$\begin{aligned} e(S^{\mathsf{c}}) &\geq \frac{1}{n-1} (+) \\ &\geq \frac{1}{n-1} \left(+ \frac{2}{(1-p)2^n} e(S^{\mathsf{c}})^2 - e(S^{\mathsf{c}}) \right), \end{aligned}$$

with equality in the first line if and only if $= 0$. Rearranging, we find that

$$e(S^{\mathsf{c}}) \geq \frac{1}{n} \left(+ \frac{2}{(1-p)2^n} e(S^{\mathsf{c}})^2 \right).$$

We would now like to find an expression for . First, note that

$$= 2 + 2 .$$

We also have that

$$= + 4 . \tag{4.5}$$

Combining these, we obtain that

$$= \tfrac{1}{2} - \geq \tfrac{1}{2} - , \tag{4.6}$$

with equality if and only if $= 0$. We therefore have

$$\begin{aligned} &\geq \frac{1}{2} \sum_{v \in S} d_S(v)(n - d_S(v)) - \frac{1}{2} \sum_{v \in S} d_S(v)(d_S(v) - 1) \\ &= \sum_{v \in S} d_S(v) \left(\frac{1}{2}(n+1) - d_S(v) \right) \\ &= (n+1)e(S) - \sum_{v \in S} d_S(v)^2, \end{aligned} \tag{4.7}$$

giving

$$e(S^{\mathsf{c}}) \geq \frac{1}{n} \left((n+1)e(S) - \sum_{v \in S} d_S(v)^2 + \frac{2}{(1-p)2^n} e(S^{\mathsf{c}})^2 \right).$$

Recall from (4.3) that $e(S) - e(S^{\mathrm{c}}) = (2p-1)n2^{n-1}$. Substituting this in and rearranging, we obtain

$$\left(\frac{2(2p-1)}{1-p} - \frac{1}{n}\right) e(S) \\ \geq \frac{2}{(1-p)n2^n} e(S)^2 - \frac{1}{n}\sum_{v \in S} d_S(v)^2 + \frac{p(2p-1)n2^{n-1}}{1-p}.$$

This would be a quadratic inequality in $e(S)$ were it not for the presence of the term $\frac{1}{n}\sum_{v\in S} d_S(v)^2$. We would like to bound this term from above, but it could get quite large if the values of $d_S(v)$ are far from being equal. Fortunately, this forces the maximum degree of $Q_n[S]$ to be large, which is what we are aiming to prove.

By the Cauchy-Schwarz inequality, we have

$$p2^n \sum_{v\in S} d_S(v)^2 \geq \left(\sum_{v \in S} d_S(v)\right)^2,$$

with equality if and only if all the $d_S(v)$ are equal (to qn). Let

$$p2^n \sum_{v\in S} d_S(v)^2 = \eta \left(\sum_{v \in S} d_S(v)\right)^2 = 4\eta e(S)^2,$$

so that $\eta \geq 1$ with equality if and only if all the degrees are equal. η is a measure of how far $Q_n[S]$ is from being regular. We now have

$$\left(\sum_{v \in S} d_S(v)\right)^2 = \frac{p2^n}{\eta}\sum_{v\in S} d_S(v)^2 \leq \frac{pqn2^n}{\eta}\sum_{v\in S} d_S(v),$$

and so

$$q \geq \frac{\eta}{pn2^n}\sum_{v\in S} d_S(v). \tag{4.8}$$

In other words, if $\frac{1}{n}\sum_{v\in S} d_S(v)^2$ is large, then there is a corresponding gain in the lower bound for q compared to what is given by (4.2).

Our inequality in $e(S)$ becomes

$$\left(\frac{4\eta}{pn2^n} - \frac{2}{(1-p)n2^n}\right) e(S)^2 + \left(\frac{2(2p-1)}{1-p} - \frac{1}{n}\right) e(S) \\ - \frac{p(2p-1)n2^{n-1}}{1-p} \geq 0.$$

Changing variables to $x = \frac{e(S)}{n2^{n-1}}$ for convenience, we have

$$(2\eta(1-p)-p)x^2 + (2p(2p-1)-o(1))x - p^2(2p-1) \geq 0. \tag{4.9}$$

The x^2 coefficient is positive since $\eta \geq 1$ and $p > \frac{1}{2}$, and the constant coefficient is negative. Hence the quadratic on the left-hand side has two roots, one positive and one negative, and x must be at least the positive root.

The change of variables applied to (4.8) gives

$$q \geq \frac{\eta x}{p}.$$

We will now show that the right-hand side is minimised when $\eta = 1$, producing a lower bound for q.

Let x_0 be the positive root of (4.9), considered as a function of η. By implicit differentiation, we obtain that

$$2(1-p)\left(2\eta x_0 \frac{\mathrm{d}x_0}{\mathrm{d}\eta} + {x_0}^2\right) - 2px_0\frac{\mathrm{d}x_0}{\mathrm{d}\eta} + (2p(2p-1)-o(1))\frac{\mathrm{d}x_0}{\mathrm{d}\eta} = 0,$$

and so

$$\frac{\mathrm{d}x_0}{\mathrm{d}\eta} = \frac{-2(1-p){x_0}^2}{4(1-p)\eta x_0 - 2px_0 + 2p(2p-1) - o(1)}.$$

Therefore

$$\frac{\mathrm{d}(\eta x_0)}{\mathrm{d}\eta} = x_0 + \eta\frac{\mathrm{d}x_0}{\mathrm{d}\eta} = x_0\left(\frac{2(1-p)\eta x_0 - 2px_0 + 2p(2p-1) - o(1)}{4(1-p)\eta x_0 - 2px_0 + 2p(2p-1) - o(1)}\right). \tag{4.10}$$

But

$$\begin{aligned} 2(1-p)\eta x_0 - 2px_0 + 2p(2p-1) &\geq 2(1-p)x_0 - 2px_0 + 2p(2p-1) \\ &= 2(2p-1)(p-x_0), \end{aligned}$$

and

$$x_0 \leq x = \frac{e(S)}{n2^{n-1}} = \frac{1}{n2^n}\sum_{v\in S} d_S(v) \leq \frac{1}{n2^n} np2^n = p,$$

so as $p > \frac{1}{2}$, the numerator on the right-hand side of (4.10) is positive for all sufficiently large n. Since x_0 is positive, the denominator is also positive, and therefore ηx_0 is an increasing function of η for $\eta \geq 1$. In other words,

$$q \geq \frac{\eta x}{p} \geq \frac{\eta x_0}{p} \geq \frac{x_0(1)}{p}.$$

All that is left is to find the value of $x_0(1)$. Substituting $\eta = 1$ into (4.9) and solving, we find that

$$x_0(1) = \frac{-p(2p-1) + p\sqrt{(2p-1)(1-p)}}{2-3p} - o(1),$$

and hence

$$q \geq \frac{1 - 2p + \sqrt{(2p-1)(1-p)}}{2-3p} - o(1).$$

Taking the limit as $n \to \infty$ gives the desired result. $\square$

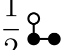
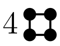
Remark. As we would hope, the lower bound on $D(p)$ obtained in Proposition 4.6.2 is strictly better than that in Proposition 4.6.1.

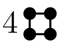
In the rest of this chapter, we will improve the lower bound on $D(p)$ further by examining configurations of vertices in copies of Q_3 rather than Q_2. The technique is generally the same as that used to prove Proposition 4.6.2, but more involved, and the resulting bound does not have a simple closed-form expression. As usual, we assume p is of the form $\frac{m}{2^r}$ and consider a set $S \subset V(Q_n)$ with $|S| = p2^n$, $\frac{1}{2} < p < \frac{2}{3}$. Let $\Delta(Q_n[S]) = qn$ as before.

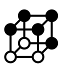
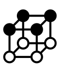
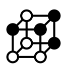
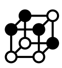
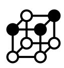
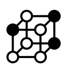
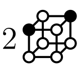
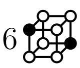
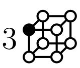
Recall from (4.4) that

$e(S^{\mathrm{c}}) = \frac{1}{n-1}$ ( $+ 2$  $+ 4$ ).

Combining this with (4.5) and (4.6), we find that

$e(S^{\mathrm{c}}) = \frac{1}{n-1}$ ($\frac{1}{2}$  $-$  $+$  $+$  $+ 4$ ). (4.11)

We would like to bound  $+ 4$  from below. In order to do this, we will consider the possible arrangements of vertices of S in subgraphs of Q_n isomorphic to Q_3. Every copy of $K_{1,3}$ in Q_n is contained within exactly one copy of Q_3. Therefore, by considering the configurations that contain  and double-counting, we obtain that

 $=$  $+ 4$  $+ 2$  $+$  $+ 3$ 
$+ 3$  $+ 4$  $+ 2$  $+ 6$  $+ 3$ . (4.12)

Note that all of the configurations that contain  also contain at least one copy of  or :

Config.	#	#	#	Config.	#	#	#
	1	1	1		3	3	0
	4	0	1		4	2	0
	2	2	0		2	4	0
	1	1	0		6	6	0
	3	2	0		3	3	0

Recall that our aim is to bound $+4$ from below. Each Q_2 in Q_n is contained in $n-2$ copies of Q_3, so we have

$$(n-2)\,(\;+4\;) = 5\; + 4\; + 2\; + \; + 2\; + 3\; + 2\; + 4\; + 6\; + 3\; .$$

Comparing the coefficients in the above and (4.12), it follows that

$$(n-2)\,(\;+4\;) \geq \frac{1}{2}\;,$$

with equality if and only if $=0$ and every Q_3 that contains or is of the form . This is because is the configuration for which

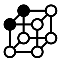
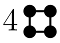
$$\frac{\#\; + 4\#\;}{\#\;}$$

is smallest.

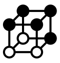
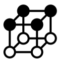
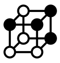
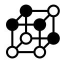
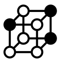
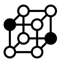
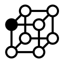
However, this is an extremely strong condition, since every copy of is in $n-2$ copies of Q_3 and all of these must be of the form . But now the vertex in S in our cycle is adjacent to $n-2$ other vertices in S, so $qn \geq n-2$. Indeed, for any Q_2 of the form , at most qn of the $n-2$ copies of Q_3 containing it can be of the form . The same is true for , which is the next best configuration for us in terms of minimising $+4$ as it has the next smallest value of $(\#\; + 4\#\;)/\#\;$ after . Hence the best we can hope for is that, for every copy of , qn of the copies of Q_3 containing it are of the form , and the other $n-2-qn$ are configurations for which

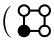
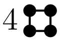
$$\frac{\#\; + 4\#\;}{\#\;} = 1.$$

For some $r \leq q$, we therefore have

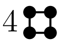
$$4$$  $$+\,3$$  $$\leq 2rn\,($$  $$+\,4$$  $$)$$

and

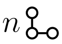
 $$+\,4$$  $$+\,2$$  $$+$$  $$+\,3$$  $$+\,2$$  $$+\,6$$  $$+\,3$$ 

$$\leq ((1-r)n-2)\,($$  $$+\,4$$  $$)\,.$$

Combining these, we have

$$((1+q)n-2)\,($$  $$+\,4$$  $$) \geq ((1+r)n-2)\,($$  $$+\,4$$  $$) \geq$$  $$.$$

Having obtained a bound on  $+\,4$ , we would like to extend this to a bound on  $+$  $+\,4$ . We have that

$$(1+q)n\,($$  $$+$$  $$+\,4$$  $$) \geq$$  $$+\,(1+q)n$$  $$, \qquad (4.13)$$

so we will consider the right-hand side above in more detail. From now on, we will ignore the $o(1)$ terms that frequently appear, as they will vanish when we take the limit as $n \to \infty$, and therefore they do not affect the eventual lower bound on $D(p)$. We have already done so by ignoring the -2 in $(1+q)n-2$ above.

We have

 $$+\,(1+q)n$$  $$= \sum_{v \in S^{\mathsf{c}}} \binom{d_{S^{\mathsf{c}}}(v)}{2}((1+q)n + d_S(v))$$
$$= \frac{1}{2} \sum_{v \in S^{\mathsf{c}}} d_{S^{\mathsf{c}}}(v)^2((2+q)n - d_{S^{\mathsf{c}}}(v)).$$

Writing $d_{S^{\mathsf{c}}}(v) = \beta_v n$, this becomes

 $$+\,(1+q)n$$  $$= \frac{1}{2} n^3 \sum_{v \in S^{\mathsf{c}}} \beta_v^2 (2+q-\beta_v). \qquad (4.14)$$

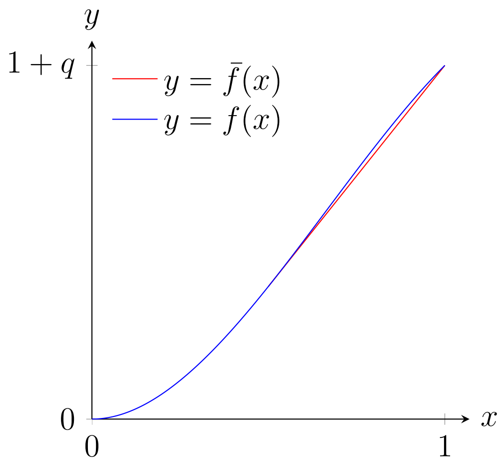
Claim 4.6.3. *Let* $\beta = \frac{1}{|S^{\mathsf{c}}|}\sum_{v \in S^{\mathsf{c}}} \beta_v$. *Then*

$$\sum_{v \in S^{\mathsf{c}}} \beta_v^2(2+q-\beta_v) \geq |S^{\mathsf{c}}|\beta^2(2+q-\beta).$$

Proof. For every $v \in S^{\mathsf{c}}$, we have $0 \leq \beta_v \leq 1$. Therefore if $f(x) = x^2(2+q-x)$ were convex for $0 \leq x \leq 1$, then we would immediately be done by Jensen's inequality. However, the second derivative of $f(x)$ is $2(2+q)-6x$, which is negative for $x > \frac{2+q}{3}$, so we cannot apply Jensen's inequality straight away.

Let $0 < x_0 < 1$ be such that the tangent to the curve $y = f(x)$ at $(x_0, f(x_0))$ passes through $(1, f(1)) = (1, 1+q)$. We will define a new function $\bar{f}$ as follows:

$$\bar{f}(x) = \begin{cases} f(x) & 0 \le x \le x_0 \\ \frac{x-x_0}{1-x_0} f(1) + \frac{1-x}{1-x_0} f(x_0) & x_0 < x \le 1 \end{cases}$$



In other words, $y = \bar{f}(x)$ is the aforementioned tangent for $x_0 < x \le 1$. This ensures that $\bar{f}$ is continuous and convex for $0 \le x \le 1$. Furthermore, $f(x) \ge \bar{f}(x)$ for all $0 \le x \le 1$. Now Jensen's inequality applies, giving

$$\sum_{v \in S^c} \beta_v^2 (2 + q - \beta_v) \ge \sum_{v \in S^c} \bar{f}(\beta_v) \ge |S^c| \bar{f}(\beta).$$

It remains to show that $\beta \le x_0$, since then $\bar{f}(\beta) = \beta^2(2+q-\beta)$, giving the desired result. First we will find x_0. We have that $f'(x_0)(1-x_0) = f(1) - f(x_0)$, so

$$(1-x_0)(2(2+q)x_0 - 3x_0^2) = 1 + q - (2+q)x_0^2 + x_0^3.$$

The right-hand side also has a factor of $1 - x_0$; indeed, there is a double root at 1. Factorising and rearranging, we find that $(1-x_0)^2(2x_0 - (1+q)) = 0$, and so $x_0 = \frac{1+q}{2}$.

Finally, note that

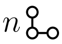
$$\beta = \frac{1}{(1-p)n2^n} \sum_{v \in S^c} d_{S^c}(v) = \frac{2e(S^c)}{(1-p)n2^n} = \frac{e(S) - (2p-1)n2^{n-1}}{(1-p)n2^{n-1}}.$$

Since $e(S) \le pqn2^{n-1}$, it suffices to show that

$$\frac{pq - 2p + 1}{1-p} \le \frac{1+q}{2}.$$

But this simplifies to $2pq - 4p + 2 \le 1 - p + q - pq$, which is equivalent to $(3p-1)(1-q) \ge 0$, so the Claim is proved. $\square$

We can now apply Claim 4.6.3 to (4.14), which tells us that

 $+ (1+q)n$  $\ge \frac{1}{2}n^3|S^{\mathsf{c}}|\beta^2(2+q-\beta)$

$$= \frac{1}{2}n^3(1-p)2^n\left(\frac{2e(S^{\mathsf{c}})}{(1-p)n2^n}\right)^2\left(2+q-\frac{2e(S^{\mathsf{c}})}{(1-p)n2^n}\right)$$

$$= \frac{2(2+q)n}{(1-p)2^n}e(S^{\mathsf{c}})^2 - \frac{4}{(1-p)^2 2^{2n}}e(S^{\mathsf{c}})^3.$$

Substituting into (4.13) and then into (4.11) tells us that $e(S^{\mathsf{c}})$ is at least

$\frac{1}{n-1}\Bigg(\frac{1}{2}$  $-$  $+ \frac{2(2+q)}{(1-p)(1+q)2^n}e(S^{\mathsf{c}})^2 - \frac{4}{(1-p)^2(1+q)n2^{2n}}e(S^{\mathsf{c}})^3\Bigg).$

$\frac{1}{2}$  $-$  is an expression that we encountered previously in the proof of Proposition 4.6.2. Recall from (4.7) that it is equal to $(n+1)e(S) - \sum_{v\in S} d_S(v)^2$. As before, we define η such that

$$p2^n \sum_{v\in S} d_S(v)^2 = \eta\left(\sum_{v\in S} d_S(v)\right)^2 = 4\eta e(S)^2,$$

so that $\eta \ge 1$ by the Cauchy-Schwarz inequality; since we are ignoring error terms, equality holds if and only if the $d_S(v)$ are close to all being equal in some asymptotic sense. Once again, we have that

$$q \ge \frac{\eta}{pn2^n}\sum_{v\in S} d_S(v).$$

Returning to the main inequality, we now have

$$e(S^{\mathsf{c}}) \ge \frac{1}{n-1}\Bigg((n+1)e(S) - \frac{4\eta}{p2^n}e(S)^2 + \frac{2(2+q)}{(1-p)(1+q)2^n}e(S^{\mathsf{c}})^2 - \frac{4}{(1-p)^2(1+q)n2^{2n}}e(S^{\mathsf{c}})^3\Bigg).$$

Substituting in $e(S^{\mathrm{c}}) = e(S) - (2p-1)n2^{n-1}$ from (4.3), and ignoring $o(1)$ terms, gives

$$\begin{aligned}
ne(S) - (2p-1)n^2 2^{n-1} \geq & -\frac{4}{(1-p)^2(1+q)n2^{2n}} e(S)^3 \\
& + \left(-\frac{4\eta}{p2^n} + \frac{2(2+q)}{(1-p)(1+q)2^n} + \frac{6(2p-1)}{(1-p)^2(1+q)2^n} \right) e(S)^2 \\
& - \left(\frac{2(2+q)(2p-1)n}{(1-p)(1+q)} + \frac{3(2p-1)^2 n}{(1-p)^2(1+q)} - n \right) e(S) \\
& + \frac{(2+q)(2p-1)^2 n^2 2^{n-1}}{(1-p)(1+q)} + \frac{(2p-1)^3 n^2 2^{n-1}}{(1-p)^2(1+q)}.
\end{aligned}$$

Setting $x = \frac{e(S)}{n2^{n-1}}$, as in the proof of Proposition 4.6.2, yields the following cubic inequality:

$$\begin{aligned}
\frac{1}{(1-p)^2(1+q)} x^3 + & \left(\frac{2\eta}{p} - \frac{2+q}{(1-p)(1+q)} - \frac{3(2p-1)}{(1-p)^2(1+q)} \right) x^2 \\
& + \left(\frac{2(2+q)(2p-1)}{(1-p)(1+q)} + \frac{3(2p-1)^2}{(1-p)^2(1+q)} \right) x \\
& - \frac{(2+q)(2p-1)^2}{(1-p)(1+q)} - \frac{(2p-1)^3}{(1-p)^2(1+q)} - 2p + 1 \geq 0. \qquad (4.15)
\end{aligned}$$

Call the left-hand side $h(x)$. Since the leading coefficient is positive and the constant coefficient is negative, $h(x)$ has at least one positive root. Considering $h'(x)$ as a function of x, p and q, we can show that it is at least 0 on the boundaries of the region in $\mathbb{R}^3$ defined by $x \geq 0$, $\frac{1}{2} \leq p \leq \frac{2}{3}$ and $0 \leq q \leq \frac{1}{2}$; moreover, it can be proved that there is no stationary point inside this region. (Note that since $D^-(\frac{2}{3}) = \frac{1}{2}$, we may assume that $q \leq \frac{1}{2}$.) These calculations are tedious and unenlightening, and are therefore omitted. Since $h'(x) \to \infty$ as $x \to \infty$ whatever the values of p, q and η, it follows that $h'(x) \geq 0$ for $x \geq 0$. Therefore $h(x)$ has exactly one positive root, which we will call x_0, and $x \geq x_0$ holds.

Again following the proof of Proposition 4.6.2, we know that

$$q \geq \frac{\eta x}{p} \geq \frac{\eta x_0}{p}.$$

But x_0 depends on q: indeed, we have

$$x_0(\eta, q) \leq \frac{pq}{\eta}.$$

It follows that $h(\frac{pq}{\eta}) \geq 0$. Making this substitution into (4.15) and rearranging gives a cubic inequality in q:

$$\begin{aligned}&\left(\frac{p^3}{(1-p)^2\eta^3} - \frac{p^2}{(1-p)\eta^2} + \frac{2p}{\eta}\right) q^3 \\ &+ \left(-\frac{2p^2}{(1-p)\eta^2} + \frac{2p^2}{(1-p)\eta} - \frac{3p^2(2p-1)}{(1-p)^2}\right) q^2 \\ &+ \left(\frac{p(2p-1)(2p+1)}{(1-p)^2\eta} - \frac{p(2p-1)}{1-p}\right) q - \frac{(2p-1)p^2}{(1-p)^2} \geq 0.\end{aligned}$$

Let the left-hand side be $g_{p,\eta}(q)$. The constant coefficient is negative, and the leading coefficient is

$$\begin{aligned}\frac{p^3}{(1-p)^2\eta^3} - \frac{p^2}{(1-p)\eta^2} + \frac{2p}{\eta} &> \frac{p^3}{(1-p)^2\eta^3} - \frac{2p^2}{(1-p)\eta^2} + \frac{p}{\eta} \\ &= \frac{p}{\eta}\left(\frac{p}{(1-p)\eta} - 1\right)^2 \geq 0.\end{aligned}$$

Hence $g_{p,\eta}(q)$ has at least one positive root. Define $q_{p,\eta}$ to be the smallest such root. Then we have the following lower bound for q:

$$q \geq \min_{\eta \geq 1} q_{p,\eta}.$$

There does not appear to be a simple closed-form expression, but the bound can be calculated and plotted using a computer. One might expect that the minimum of $q_{p,\eta}$ is always attained when $\eta = 1$, corresponding to $Q_n[S]$ being asymptotically close to regular; in this case the cubic formula gives a complicated closed-form expression for $q_{p,\eta}$. However, we rather surprisingly find that for $0.5 < p < 0.519026...$, the minimum is attained when $\eta > 1$. We believe that this is just a limitation of the method, and that the actual induced subgraph of size $p2^n$ with the smallest maximum degree will always be asymptotically close to regular.

We now have four lower bounds for $D(p)$ in the range $\frac{1}{2} < p < \frac{2}{3}$, each an improvement on the one before. Let the original lower bound of $\frac{2p-1}{p}$ from Corollary 4.2.3 be denoted $D_0(p)$, the bounds from Propositions 4.6.1 and 4.6.2 be denoted $D_1(p)$ and $D_2(p)$ respectively, and the new lower bound of $\min_{\eta \geq 1} q_{p,\eta}$ be denoted $D_3(p)$. Figure 4.2 shows these bounds in the range $\frac{1}{2} \leq p \leq 1$. The values for which $D_0(p)$ is asymptotically tight, that is, for which $D_0(p) = D^-(p)$, are marked up to $p = 1 - \frac{1}{20}$.

Note in particular how quickly the bound $D_3(p)$ rises for p just above $\frac{1}{2}$. For example, an induced subgraph of Q_n with $0.501 \cdot 2^n$ vertices must contain a vertex of degree greater than $0.1n$ when n is sufficiently large.

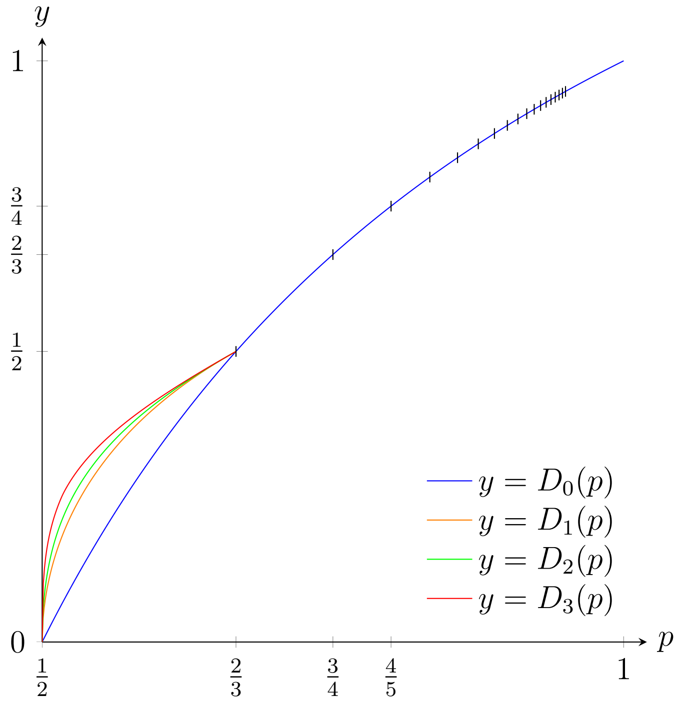


Fig. 4.2: The lower bounds for $D(p)$

4.7 Conclusion and open problems

We have shown that $D(p)$ displays more complex behaviour than one might expect. There are at least two different regimes: one for $\frac{1}{2} < p < \frac{2}{3}$ and another for $\frac{2}{3} < p < 1$. (One could count $D(p) = 0$ for $p < \frac{1}{2}$ as being a third regime.) At $p = \frac{2}{3}$, the boundary between these regimes, $D(p)$ is not smooth. However, there is much still to learn about this function. First of all, the lower bound $D_3(p)$, which was derived from careful examination of the possible configurations of vertices of S in copies of Q_3, is almost certainly not tight. It seems likely that the same techniques could be applied to configurations within copies of Q_4, Q_5 and ever larger subcubes to produce increasingly better lower bounds on $D(p)$ for $\frac{1}{2} < p < \frac{2}{3}$. However, the calculations would increase rapidly in difficulty and complexity as the number of possible configurations grows.

Ideally one would aim to completely characterise $D(p)$ for all p, but there are some interesting questions about its behaviour that could be easier to resolve. Since the lower bound of $\frac{2p-1}{p}$ from Corollary 4.2.3 is not tight for $\frac{1}{2} < p < \frac{2}{3}$, it is natural to wonder whether the same is true for $\frac{2}{3} < p < \frac{3}{4}$ and so on. One is reminded of the results of Razborov [70], Nikiforov [65] and Reiher [71] on the smallest possible density of cliques in a graph with a given edge density: in those cases, a certain lower bound can only be achieved when the edge density is of the form $1 - \frac{1}{r}$.

Question 4.7.1. *Is the bound*

$$D^{-}(p) \geq \frac{2p-1}{p}$$

tight for any values of p *that are not of the form* $1 - \frac{1}{r}$ *for integer* r*?*

We know that $D(p)$ is not smooth as a function of p in the range $\frac{1}{2} \leq p \leq 1$. The following question is therefore a natural one to consider:

Question 4.7.2. *Is* $D(p)$ *continuous for* $\frac{1}{2} \leq p \leq 1$*?*

For $p = 1 - \frac{1}{2^r}$ and $n = k(2^r - 1)$, Corollary 4.3.3 tells us that there exists an induced subgraph of Q_n with $p2^n$ vertices and maximum degree exactly $\frac{2p-1}{p}n$, matching the lower bound from Corollary 4.2.3. By analogy with Huang's result [51] for induced subgraphs of Q_n with $2^{n-1} + 1$ vertices, the following problem may be of interest.

Question 4.7.3. *Let* $n = k(2^r - 1)$ *and* $p = 1 - \frac{1}{2^r}$*. Suppose that* G *is an induced subgraph of* Q_n *with* $p2^n + 1$ *vertices. How small can* $\Delta(G) - \frac{2p-1}{p}n$ *be?*

We could also consider the problem of finding the smallest possible maximum degree of an induced subgraph of a given size for families of regular graphs other than the hypercube. Recall from Corollary 4.2.2 that we have a lower bound of $\frac{2p-1}{p}k$ for a general k-regular graph. Consider, for example, the cycle C_n, a 2-regular graph. We have

$$\min_{\substack{S \subset V(C_n) \\ |S| \geq m}} \Delta(C_n[S]) = \begin{cases} 0, & 0 \leq m \leq \left\lfloor \frac{n}{2} \right\rfloor \\ 1, & \left\lfloor \frac{1}{2}n \right\rfloor < m \leq \left\lfloor \frac{2}{3}n \right\rfloor \\ 2, & \left\lfloor \frac{2}{3}n \right\rfloor < m \leq n. \end{cases}$$

For $m \geq \frac{1}{2}n$, this exactly matches the lower bound obtained from Corollary 4.2.2, which after taking the ceiling function is $\left\lceil 4 - \frac{2n}{m} \right\rceil$. For $m < \frac{1}{2}n$, the trivial lower bound of 0 is tight.

However, for some families of regular graphs, the lower bound from Corollary 4.2.2 is far from the truth. For example, an induced subgraph of K_n with m

vertices has maximum degree $m-1$, but Corollary 4.2.2 only gives a lower bound of $(2-\frac{n}{m})(n-1)$, which is smaller when $m<n$.

The hypercube Q_n is the Cartesian product of n copies of P_2, and so it may be interesting to consider maximum degrees of induced subgraphs for families of regular graphs constructed in a similar way. Two such families that can be seen as generalisations of the hypercube are the Cartesian products of cycles, C_k^n, and the Cartesian products of complete graphs, K_k^n, which are known as Hamming graphs.

The generalisation to K_k^n of Chung, Füredi, Graham and Seymour's problem on the cube [26], namely the problem of determining the smallest possible maximum degree of an induced subgraph of K_k^n with $k^{n-1}+1=\alpha(K_k^n)+1$ vertices, has been studied in recent years. Dong [30] extended Chung, Füredi, Graham and Seymour's construction to show that a maximum degree of $\lceil\sqrt{n}\rceil$ is possible for all k. Tandya [80] then solved the problem completely by showing that the Hamming graphs for $k\geq 3$ do not behave like the hypercube: for all $k\geq 3$ and all n, there exists an induced subgraph of K_k^n with $k^{n-1}+1$ vertices and maximum degree 1. It would be of interest to determine whether the Hamming graphs continue to display different behaviour from Q_n when larger induced subgraphs are considered.

References

[1] A. Aashtab, S. Akbari, M. Ghanbari, and A. Shidani. "Vertex partitioning of graphs into odd induced subgraphs". *Discuss. Math. Graph Theory* 43 (2023), 385–399.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. "A dense infinite Sidon sequence". *European J. Combin.* 2.1 (1981), 1–11.

[3] M. Ajtai, J. Komlós, and E. Szemerédi. "A note on Ramsey numbers". *J. Combin. Theory Ser. A* 29.3 (1980), 354–360.

[4] A. Altshuler. "Construction and enumeration of regular maps on the torus". *Discrete Math.* 4 (1973), 201–217.

[5] K. Appel and W. Haken. "Every planar map is four colorable, Part I: discharging". *Illinois J. Math.* 21.3 (1977), 429–490.

[6] K. Appel, W. Haken, and J. Koch. "Every planar map is four colorable, Part II: reducibility". *Illinois J. Math.* 21.3 (1977), 491–567.

[7] R. C. Baker, G. Harman, and J. Pintz. "The difference between consecutive primes, II". *Proc. Lond. Math. Soc.* 83.3 (2001), 532–562.

[8] N. Balachandran and B. Sankarnarayanan. "The choice number versus the chromatic number for graphs embeddable on orientable surfaces". *Electron. J. Comb.* 28 (2021), #P4.50.

[9] G. D. Birkhoff. "A determinant formula for the number of ways of coloring a map". *Ann. of Math.* 14.1 (1912-13), 42–46.

[10] G. D. Birkhoff. "The reducibility of maps". *Amer. J. Math.* 35.2 (1913), 115–128.

[11] T. Bohman and P. Keevash. "Dynamic concentration of the triangle-free process". *Random Structures Algorithms* 58.2 (2021), 177–380.

[12] B. Bollobás. *Modern Graph Theory.* Springer, 1998.

[13] B. Bollobás, J. Lee, and S. Letzter. "Eigenvalues of subgraphs of the cube". *European J. Combin.* 70 (2018), 125–148.

[14] J. A. Bondy and P. Erdős. "Ramsey numbers for cycles in graphs". *J. Combin. Theory Ser. B* 14.1 (1973), 46–54.

[15] H. R. Brahana. "Systems of circuits on two-dimensional manifolds". *Ann. of Math.* 23.2 (1921), 144–168.

[16] M. Buck and D. Wiedemann. "Gray codes with restricted density". *Discrete Math.* 48.2-3 (1984), 163–171.

[17] S. A. Burr. "Generalized Ramsey theory for graphs – a survey". In: *Graphs and Combinatorics.* Ed. by R. Bari and F. Harary. Lecture Notes in Mathematics, vol. 406. Springer, 1974.

[18] S. A. Burr, P. Erdős, and J. Spencer. "Ramsey theorems for multiple copies of graphs". *Trans. Amer. Math. Soc.* 209 (1975), 87–99.

[19] M. Campos, S. Griffiths, R. Morris, and J. Sahasrabudhe. "An exponential improvement for diagonal Ramsey". arXiv:2303.09521.

[20] Y. Caro, M. Petruševski, and R. Škrekovski. "Remarks on odd colorings of graphs". *Discrete Appl. Math.* 321 (2022), 392–401.

[21] A. Cayley. "On the colouring of maps". *Proc. R. Geogr. Soc.* 1.4 (1879), 259–261.

[22] G. Chartrand and S. Schuster. "On the existence of specified cycles in complementary graphs". *Bull. Amer. Math. Soc.* 77.6 (1971), 995–998.

[23] P. Cheilaris, B. Keszegh, and D. Pálvölgyi. "Unique-Maximum and Conflict-Free Coloring for Hypergraphs and Tree Graphs". *SIAM J. Discrete Math.* 27.4 (2013), 1775–1787.

[24] G. Chen, X. Yu, and Y. Zhao. "Improved bounds on the Ramsey number of fans". *European J. Combin.* 96 (2021), 103347.

[25] H. Chernoff. "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations". *Ann. Math. Stat.* 23.4 (1952), 493–507.

[26] F. Chung, Z. Füredi, R. Graham, and P. Seymour. "On induced subgraphs of the cube". *J. Combin. Theory Ser. A* 49.1 (1988), 180–187.

[27] D. Conlon. "A new upper bound for diagonal Ramsey numbers". *Ann. of Math.* 170.2 (2009), 941–960.

[28] D. Conlon. "The Ramsey number of books". *Adv. Combin.* (2019), Paper no. 3.

[29] D. Conlon, J. Fox, and Y. Wigderson. "Ramsey numbers of books and quasirandomness". *Combinatorica* 42.3 (2022), 309–363.

[30] D. Dong. "On induced subgraphs of the Hamming graph". *J. Graph Theory* 96.1 (2021), 160–166.

[31] V. Dvořák and H. Metrebian. "A new upper bound for the Ramsey number of fans". *European J. Combin.* 110 (2023), 103680.

[32] P. Erdős. "Some remarks on the theory of graphs". *Bull. Amer. Math. Soc.* 53.4 (1947), 292–294.

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

[34] P. Erdős and G. Szekeres. "A combinatorial problem in geometry". *Compos. Math.* 2 (1935), 463–470.

[35] G. Even, Z. Lotker, D. Ron, and S. Smorodinsky. "Conflict-free colorings of simple geometric regions with applications to frequency assignment in cellular networks". *SIAM J. Comput.* 33.1 (2003), 94–136.

[36] I. Fabrici, B. Lužar, S. Rindošová, and R. Soták. "Proper conflict-free and unique-maximum colorings of planar graphs with respect to neighborhoods". *Discrete Appl. Math.* 324 (2023), 80–92.

[37] R. J. Faudree and R. H. Schelp. "All Ramsey numbers for cycles in graphs". *Discrete Math.* 8.4 (1974), 313–329.

[38] J. Fink. “Perfect matchings extend to Hamilton cycles in hypercubes”. *J. Combin. Theory Ser. B* 97.6 (2007), 1074–1076.

[39] G. Fiz Pontiveros, S. Griffiths, and R. Morris. “The triangle-free process and the Ramsey number $R(3,k)$”. *Mem. Amer. Math. Soc.* 263 (2020), no. 1274.

[40] P. Franklin. “The four colour problem”. *Amer. J. Math.* 44.3 (1922), 225–236.

[41] S. Golomb and L. Welch. “Perfect codes in the Lee metric and the packing of polyominoes”. *SIAM J. Appl. Math.* 18.2 (1970), 302–317.

[42] C. Gotsman and N. Linial. “The equivalence of two problems on the cube”. *J. Combin. Theory Ser. A* 61.1 (1992), 142–146.

[43] R. W. Hamming. “Error detecting and error correcting codes”. *Bell Syst. Tech. J.* 29.2 (1950), 147–160.

[44] F. Harary, J. P. Hayes, and H. Wu. “A survey of the theory of hypercube graphs”. *Comput. Math. Appl.* 15.4 (1988), 277–289.

[45] L. H. Harper. “Optimal assignments of numbers to vertices”. *J. Soc. Ind. Appl. Math.* 12.1 (1964), 131–135.

[46] L. H. Harper. “Optimal numberings and isoperimetric problems on graphs”. *J. Combin. Theory* 1.3 (1966), 385–393.

[47] I. Havel. “Semipaths in directed cubes”. In: *Graphs and other combinatorial topics*. Ed. by M. Fiedler. Teubner-Texte zur Mathematik 59. Teubner, 1983.

[48] P. J. Heawood. “Map-colour theorem”. *Q. J. Math.* 24 (1890), 332–338.

[49] H. Heesch. *Untersuchungen zum Vierfarbenproblem*. German. Bibliographisches Institut, 1969.

[50] G. Hoheisel. “Primzahlprobleme in der Analysis”. German. *Sitz. Preuss. Akad. Wiss.* 33 (1930), 580–588.

[51] H. Huang. “Induced subgraphs of hypercubes and a proof of the sensitivity conjecture”. *Ann. of Math.* 190.3 (2019), 949–955.

[52] A. B. Kempe. “On the geographical problem of the four colours”. *Amer. J. Math.* 2.3 (1879), 193–200.

[53] J. H. Kim. “The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$”. *Random Structures Algorithms* 7.3 (1995), 173–207.

[54] G. Kreweras. “Matchings and Hamiltonian cycles on hypercubes”. *Bull. Inst. Combin. Appl.* 16 (1996), 87–91.

[55] H. Lebesgue. “Quelques conséquences simples de la formule d’Euler”. French. *J. Math. Pures Appl. (9)* 19 (1940), 27–43.

[56] Y. Li and C. C. Rousseau. “Fan-complete graph Ramsey numbers”. *J. Graph Theory* 23.4 (1996), 413–420.

[57] Y. Li, C. C. Rousseau, and W. Zang. “Asymptotic upper bounds for Ramsey functions”. *Graphs Combin.* 17 (2001), 123–128.

[58] Q. Lin and Y. Li. “On Ramsey numbers of fans”. *Discrete Appl. Math.* 157.1 (2009), 191–194.

[59] Q. Lin, Y. Li, and L. Dong. "Ramsey goodness and generalized stars". *European J. Combin.* 31.5 (2010), 1228–1234.

[60] S. Mattheus and J. Verstraete. "The asymptotics of $r(4, t)$". arXiv:2306.04007.

[61] H. Metrebian. "Odd colouring on the torus". arXiv:2205.04398.

[62] M. Mitzenmacher and E. Upfal. *Probability and Computing*. Cambridge University Press, 2005.

[63] B. Mohar and C. Thomassen. *Graphs on Surfaces*. Johns Hopkins University Press, 2001.

[64] T. Mütze. "Proof of the middle levels conjecture". *Proc. Lond. Math. Soc.* 112.4 (2016), 677–713.

[65] V. Nikiforov. "The number of cliques in graphs of given order and size". *Trans. Amer. Math. Soc.* 363.3 (2011), 1599–1618.

[66] N. Nisan and M. Szegedy. "On the degree of Boolean functions as real polynomials". *Comput. Complexity* 4 (1994), 301–313.

[67] J. Petr and J. Portier. "The odd chromatic number of a planar graph is at most 8". *Graphs Combin.* 39 (2023), article 28.

[68] M. Petruševski and R. Škrekovski. "Colorings with neighborhood parity condition". *Discrete Appl. Math.* 321 (2022), 385–391.

[69] F. P. Ramsey. "On a problem of formal logic". *Proc. Lond. Math. Soc.* 30.1 (1930), 264–286.

[70] A. A. Razborov. "On the minimal density of triangles in graphs". *Combin. Probab. Comput.* 17.4 (2008), 603–618.

[71] C. Reiher. "The clique density theorem". *Ann. of Math.* 184.3 (2016), 683–707.

[72] G. Ringel and J. W. T. Youngs. "Solution of the Heawood map-coloring problem". *Proc. Natl. Acad. Sci. USA* 60.2 (1968), 438–445.

[73] V. Rosta. "On a Ramsey-type problem of J. A. Bondy and P. Erdős. I". *J. Combin. Theory Ser. B* 15.1 (1973), 94–104.

[74] V. Rosta. "On a Ramsey-type problem of J. A. Bondy and P. Erdős. II". *J. Combin. Theory Ser. B* 15.1 (1973), 105–120.

[75] C. C. Rousseau and J. Sheehan. "On Ramsey numbers for books". *J. Graph Theory* 2.1 (1978), 77–87.

[76] F. Ruskey and C. D. Savage. "Hamilton cycles that extend transposition matchings in Cayley graphs of S_n". *SIAM J. Discrete Math.* 6.1 (1993), 152–166.

[77] A. Sah. "Diagonal Ramsey via effective quasirandomness". *Duke Math. J.* 172.3 (2023), 545–567.

[78] J. B. Shearer. "A note on the independence number of triangle-free graphs". *Discrete Math.* 46.1 (1983), 83–87.

[79] S. Smorodinsky. "Conflict-free coloring and its applications". In: *Geometry – Intuitive, Discrete and Convex*. Ed. by I. Bárány, K. J. Böröczky, G. F. Tóth, and J. Pach. Bolyai Society Mathematical Studies, vol. 24. Springer, 2013.

[80] V. Tandya. “An induced subgraph of the Hamming graph with maximum degree 1”. *J. Graph Theory* 101.2 (2022), 311–317.

[81] A. Thomason. “An upper bound for some Ramsey numbers”. *J. Graph Theory* 12.4 (1988), 509–517.

[82] F. Tian and Y. Yin. “Every toroidal graph without 3-cycles is odd 7-colorable”. arXiv:2206.06052.

[83] F. Tian and Y. Yin. “Every toroidal graphs without adjacent triangles is odd 8-colorable”. arXiv:2206.07629.

[84] F. Tian and Y. Yin. “The odd chromatic number of a toroidal graph is at most 9”. *Inform. Process. Lett.* 182 (2023), 106384.

[85] P. Wernicke. “Über den kartographischen Vierfarbensatz”. German. *Math. Ann.* 58 (1904), 413–426.

[86] R. A. Wilson. *Graphs, Colourings and the Four-colour Theorem.* Oxford University Press, 2002.

[87] Y. Zhang, H. Broersma, and Y. Chen. “A note on Ramsey numbers for fans”. *Bull. Aust. Math. Soc.* 92.1 (2015), 19–23.