Clique partitions and coverings of graphs

by

Michael S. Cavers

An essay presented
to the University of Waterloo
in fulfillment of the
essay requirement for the degree of
Masters of Mathematics

Waterloo, Ontario, Canada

December 20th, 2005

Abstract

This essay looks at the clique partition and covering numbers of graphs and introduces some new results.

In the first part, two cornerstone theorems in clique partitions are presented and proved. Namely, the Erdős-Goodman-Pósa Theorem and the de Bruijn-Erdős Theorem. The problem of partitioning the complement of a clique is looked at, and asymptotic results are provided. The existence of Steiner systems and projective planes can be used to construct clique partitions of complements of graphs. Using Steiner systems, upper bounds on the clique partition number of the complement of paths, cycles, and perfect matchings are given. For the first time, we provide an upper bound for the clique partition number of the complement of a forest and the complement of graphs with bounded maximum degree.

In the second part, the clique covering number of the complement of small graphs is analyzed. We provide results by Kohayakawa of induced paths and cycles in the Kneser graph. Bounds on the clique covering number of the complement of a path and cycle are given. Also, the clique covering number of the complement of a perfect matching is determined, a result of Gregory and Pullman. By using the Erdős-Ko-Rado Theorem, we are able to generalize this result to obtain a lower bound on the clique covering number for complete multipartite graphs. The clique covering number of the complement of the union of cycles and paths is looked at, and we demonstrate that in general, a lower bound of $\log_2 n$ cannot be obtained. Finally, for the first time, we obtain bounds on the clique covering number of the complement of a forest.

Acknowledgments

I would like to thank my supervisor Jacques Verstraëte for all of his valuable advice and feedback throughout the process of preparing this essay. I would also like to thank Nick Wormald for being an official reader of this essay. Finally, I would like to acknowledge Kevin Vander Meulen for introducing this problem to me.

Contents

Part I

Introduction

1.1 Edge-partitioning and covering problems

A clique partition of a graph G is a collection of complete subgraphs of G (called cliques) that partition the edge set of G. Similarly, a clique cover of G is a collection of cliques that cover the edge set of G (that is, every edge appears in at least one clique of the clique cover). In this essay, we study the problem of finding clique partitions and covers of minimum size of a graph. We analyze the clique partition number of a graph in Chapter II, the clique covering number of a graph in Chapter III, and state some open problems in Chapter IV. Monson, Pullman, and Rees [21] survey recent results regarding clique partitions and coverings of graphs.

The first result on clique partitions and covers is due to Hall [18], who proved that the edge set of any graph G on n vertices can be covered using at most $\lfloor n^2/4 \rfloor$ cliques, none of which need to be larger than a triangle. Later, Erdős, Goodman and Pósa [14] showed that Hall's result holds for clique partitions as well. In Chapter II we look at this result among others dealing with the clique partition number of a graph. In particular, Erdős, Goodman and Pósa [14] showed that the edge set of any graph G with n vertices can be partitioned using at most $\lfloor n^2/4 \rfloor$ triangles and edges, and the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ gives equality. We present the theorem and its proof in Section 2.2.

In Section 2.3 we look at Steiner systems. A rich family of clique partitions of K_n is provided by the existence of Steiner systems. We use the terminology from Cameron and Van Lint [9]. The complete subgraphs of K_n induced by each block in a Steiner system form a clique partition of K_n.

In Section 2.4 we present and prove the de Bruijn-Erdős [8] Theorem. It states that for $n \geq 3$, if $\mathcal{C}$ is a non-trivial clique partition of K_n, then $|\mathcal{C}| \geq n$. Further, equality holds if and only if $\mathcal{C}$ consists of one clique on $n-1$ vertices and $n-1$ copies of K_2 incident with a single vertex of K_n; or $\mathcal{C}$ consists of n copies of K_{q+1}, $n = q^2 + q + 1$ and each vertex of K_n is a vertex of exactly $q+1$ cliques of $\mathcal{C}$. This second condition is equivalent to the existence of a projective plane of order q. There are several proofs of the de Bruijn-Erdős Theorem in the literature. In this essay, we follow the proof given by Conway, which may be found in Batten and Beutelspacher [4].

A natural generalization is to consider the clique partition number of $K_n \backslash K_m$, the complement of a clique of size m, for $m \geq 2$. Using the edge chromatic number of a graph, Pullman and Donald [23] compute the clique partition number of $K_n \backslash K_m$, for $\frac{n}{2} \leq m \leq n$. This result is presented in Section 2.5.1. Wallis [26] determined that for $m \leq \sqrt{n}$, the number of cliques required to partition the edge set of $K_n \backslash K_m$ is asymptotic to n. This result follows from a theorem in Section 2.6. To try and fill in the gap, Erdős, Faudree and Ordman [13] use affine planes to get asymptotic results for other values of m, which can be found in Section 2.5.2. A consequence is that the clique partition number of $K_n \backslash K_m$ is asymptotic to $\max\{m^2, n\}$, for $m = o(n)$.

The existence of Steiner systems and projective planes can be used to construct clique partitions of complements of graphs. The idea is to fit most of the graph into few blocks to get good clique partitions of the complement. In Section 2.6, we prove a result of Wallis [28] where if H is a graph with at most $\sqrt{n}$ vertices, the complement of H in K_n can be partitioned into $O(n)$ cliques. It is more difficult to determine clique partitions of complements of spanning subgraphs of K_n. In Sections $2.7 - 2.8$, we consider matchings, paths, cycles and forests. Gregory, McGuinness and Wallis [15] prove that for n sufficiently large, the complement of a perfect matching on n vertices can be partitioned into about $n \log_2 \log_2 n$ cliques. The details can be found in Section 2.7. In an attempt to generalize these results, we look at the complement of a spanning forest of K_n in Section 2.8. Using projective planes we show the existence of a clique partition of the complement of a forest using about $n(\log_2 n)^{\log_2 3}$ cliques.

Finally, we present a general method for finding upper bounds for any graph with bounded maximum degree. We use Steiner systems and a probabilistic argument to prove the existence of a clique partition of $\bar{G}$ using $n^{3/2}\sqrt{\Delta}(\log_2 n)^2$ cliques, where $\Delta = o(n/(\log_2 n)^4)$ is the maximum degree of G. This is done in Section 2.9. We conjecture that the complement of any graph on n vertices with $o(n^2)$ edges can be partitioned into $o(n^2)$ cliques.

In Chapter III, we analyze the clique covering number of the complement of graphs with very few edges, such as a cycle, path, and perfect matching. Finding an induced graph H in the Kneser graph gives rise to a clique cover of $\bar{H}$. Thus, by looking at induced subgraphs in the Kneser graph we may obtain upper bounds on the clique covering numbers of graphs. In Section 3.1, we show the relationship between clique coverings and intersection graphs. Using this result, Gyárfás [17] gives a lower bound of $\log_2(n+1)$ for a graph with n vertices that has no isolated vertices and no equivalent vertices.

In Section 3.2, we present a result of Gregory and Pullman [16]. In particular, it is shown that the clique covering number of the complement of a perfect matching on n vertices is asymptotic to $\log_2 n$. This is accomplished by using the famous Erdős-Ko-Rado Theorem on set intersections. We use the terminology from set theory found in Bollobás [6]. For the first time, we generalize this result to give a lower bound on the clique covering number of $K_s(t)$, the complete s-partite graph with parts of size t. In particular, we show for s sufficiently large and fixed t, covering the edges of $K_s(t)$ requires at least $\log_b(st)$ cliques, where $b = b(t) \in (1, 2]$, which when $t = 2$, gives Gregory and Pullman's [16] result.

In Section 3.3, the clique covering number of the complement of a cycle and path are analyzed. For m sufficiently large, de Caen, Gregory, and Pullman [11] give a lower bound of $\log_2 m$ and an upper bound of $2\log_2 m$ on the clique covering number of the complement of a cycle and path on m vertices. They conjecture that the clique covering number is $\log_2 m$ asymptotically. Alles and Poljak [1] improve the upper bound to $1.695\log_2 m$, and Kohayakawa [19] improves the upper bound to $1.459\log_2 m$, for m sufficiently large. This is done by finding a long induced path in the Kneser graph. In particular, Kohayakawa [19] uses a bipartite graph to give a recursive construction for an induced path in the Kneser graph. Kohayakawa [19] made a conjecture regarding the length of a longest induced path in the Kneser graph, which if true, would imply de Caen, Gregory, and Pullman's [11] conjecture of the existence of a clique cover of the complement of a cycle and path with size asymptotic to $\log_2 m$. We follow Vander Meulen [25] for most of the results in Section 3.3 regarding induced cycles and paths in the Kneser graph.

In Section 3.4, bounds on the clique covering number of the complement of graphs whose maximum degree is two are obtained. de Caen, Gregory and Pullman [11] prove that for n sufficiently large, an upper bound on the order of $\log_2 n$ can be obtained. They ask if an upper bound of $(1 + o(1))\log_2 n$ can be obtained, and for the first time we demonstrate that in general, it cannot. We demonstrate with the complete multipartite graph whose parts each have size three, and show this graph requires at least $1.088 \log_2 n$ cliques in a minimum size clique covering.

In Section 3.5, we use the techniques of Section 3.4 to obtain bounds on the clique covering number of the complement of a forest. We show an upper bound of $10.3 \log_2 n$ for the clique covering number of the complement of a forest, and a lower bound of $\log_2 k$, where k is the size of the largest induced matching of the forest.

Finally, in Section 4.1, we list some open problems and present some conjectures dealing with clique partitions and coverings of graphs.

1.2 Notation

We use the usual definitions from graph theory found in Bondy and Murty [12]. We often write G_n to emphasize that G has n vertices. The *complement* $\bar{G}$ of G_n is the graph on vertex set V, with edge set $E(K_n) \backslash E(G)$. If G has n vertices, then we also use $K_n \backslash G$ to denote the complement of the graph G. If $N > n$ and $V(G) \subset V(K_N)$, then $K_N \backslash G$ denotes the graph with vertex set $V(K_N)$ and edge set $E(K_N) \backslash E(G)$. Let G be a non-empty graph. If $v \in V(G)$, then the set of vertices adjacent to v in G is denoted by $\Gamma(v)$. The union $G \cup H$ of two graphs G and H has vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H)$.

We use the usual definitions from set theory found in Bollobás [6]. Sets with k elements will be called *k-sets*; subsets with k elements are *k-subsets*. Let A be a set. By $A^{(k)}$ we denote the set of all k-subsets of A. For convenience, we use $[n]$ to denote the set $\{1, 2, \ldots, n\}$.

The *Kneser graph* $K_{n:k}$ is the graph whose vertex set consists of the k-subsets of $[n]$, and where two vertices are adjacent if and only if they are disjoint. Note that $K_{n:k}$ is an $\binom{n-k}{k}$-regular graph. For example, $K_{5:2}$ is isomorphic to the Petersen graph. If $n < 2k$, then $K_{n:k}$ is the empty graph, and if $n = 2k$, then $K_{n:k}$ is a matching of size $\binom{2k}{k}$.

Part II

Clique Partitions

2.1 Preliminary results

Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. We call a complete subgraph of G a *clique* of G. A *clique partition* of G is a set of cliques of G, which together contain each edge of G exactly once. The smallest cardinality of any clique partition of G is called the *clique partition number* of G, and is denoted by $\mathrm{cp}(G)$. This number exists as the edge set of G forms a clique partition for G. A clique partition of G with size $\mathrm{cp}(G)$ is referred to as a *minimum clique partition* of G. Note that any minimum clique partition does not contain any cliques of size one, and also the clique partition number of the empty graph is 0.

Orlin [22] first noted that a clique partition of G gives rise to a clique partition of $G - v$. Namely, delete the vertex v from each clique in the partition of G. If this produces a clique with a single vertex, then delete this vertex from the clique partition.

Conversely, a clique partition of $G-v$ gives rise to a clique partition of G. Namely, add the edges adjacent to v to the clique partition of $G - v$ to get a clique partition of G.

Lemma 2.1.1 *Given any graph G,*

$$\mathrm{cp}(G) - \deg(v) \leq \mathrm{cp}(G - v) \leq \mathrm{cp}(G).$$

This lemma can be generalized as follows.

Lemma 2.1.2 *If H is an induced subgraph of G then,*

$$\mathrm{cp}(G) - (|E(G)| - |E(H)|) \leq \mathrm{cp}(H) \leq \mathrm{cp}(G).$$

Also, if H is a subgraph of G, then for any $n \geq |V(G)|$

$$\mathrm{cp}(K_n \backslash H) \leq \mathrm{cp}(K_n \backslash G) + |E(G)| - |E(H)|.$$

Proof. By taking a clique partition of H and adding the edges $E(G)\backslash E(H)$ we get a clique partition of G, so the first inequality follows. If H is an induced subgraph of G, then $\text{cp}(H) \leq \text{cp}(G)$ follows by applying Lemma 2.1.1 to each vertex of $G\backslash H$.

If H is a subgraph of G, then to get a clique partition of $K_n\backslash H$, we take a clique partition of $K_n\backslash G$ and add the edges of G that do not appear in H. □

Monson [20] lists other results that the effect vertex and edge deletion have on the clique partition number of a graph.

2.2 The Erdős-Goodman-Pósa Theorem

How large in absolute terms can $\text{cp}(G)$ be for a simple graph on n vertices?

Erdős, Goodman and Pósa [14] showed that the edge set of any simple graph G with n vertices can be partitioned using at most $\lfloor n^2/4 \rfloor$ triangles and edges, and that the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ gives equality. This is the first fundamental result in clique partitions of graphs. Note that their proof can be adapted to show that $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is the only graph that gives $\text{cp}(G) = \lfloor n^2/4 \rfloor$.

Theorem 2.2.1 *Let G be a simple graph with n vertices. Then G has a clique partition of size at most $\lfloor n^2/4 \rfloor$ consisting of edges and triangles. In particular,*

$$\text{cp}(G) \leq \lfloor n^2/4 \rfloor,$$

with equality if and only if $G \cong K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

Proof. First note that given any positive integer n,

$$\lfloor n^2/4 \rfloor = \lfloor (n-1)^2/4 \rfloor + \lfloor n/2 \rfloor.$$

We will prove the theorem by induction on the number of vertices n of G. If $n = 1, 2$, then the theorem is true. Assume that it is true for all graphs with less than n vertices, and let G be a simple graph with n vertices. Suppose that the minimum degree of G is $\delta = \lfloor n/2 \rfloor + r$, where r is an integer, and let x be a vertex of smallest degree in G.

Case 1: $r \le 0$. Then $\deg(x) \le \lfloor n/2 \rfloor$. By induction, $G - x$ can be partitioned into $\lfloor (n-1)^2/4 \rfloor$ edges and triangles. Hence, by Lemma 2.1.1,

$$\begin{aligned} \mathrm{cp}(G) &\le \mathrm{cp}(G-x) + \deg(x) \\ &\le \lfloor (n-1)^2/4 \rfloor + \lfloor n/2 \rfloor \\ &= \lfloor n^2/4 \rfloor. \end{aligned}$$

Furthermore, this clique partition of G consists of the edges and triangles from the clique partition of $G - x$ and the edges incident to x. For equality to hold in the theorem, we need $\mathrm{cp}(G - x) = \lfloor (n-1)^2/4 \rfloor$ and $\deg(x) = \lfloor n/2 \rfloor$. By induction, $G - x \cong K_{\lfloor (n-1)/2 \rfloor, \lceil (n-1)/2 \rceil}$. Since $\deg(x) = \lfloor n/2 \rfloor$ and $\mathrm{cp}(G) = \lfloor n^2/4 \rfloor$, x may only be adjacent to vertices on one side of $G - x$ (otherwise we form a triangle which may be included in our clique partition to give $\mathrm{cp}(G) < \lfloor n^2/4 \rfloor$). Hence, $G \cong K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

Case 2: $r > 0$. Let $\Gamma(x) = \{y_1, y_2, \ldots, y_\delta\}$, and let H be the subgraph of G induced by $\Gamma(x)$. We claim that H has at least r independent edges. We prove this claim by contradiction. Suppose H has at most $r - 1$ independent edges. Without loss of generality, assume this set of edges is

$$\{y_1y_2, y_3y_4, \ldots, y_{s-1}y_s\},$$

for some maximal even integer $s \le 2r - 2$. By hypothesis, $\deg(y_{s+1}) \ge \delta$. Note that $\Gamma(y_{s+1}) \subseteq V(G) \backslash \{y_{s+1}, y_{s+2}, \ldots, y_\delta\}$. Hence,

$$\begin{aligned} \deg(y_{s+1}) &\le n - (\delta - s) \\ &\le (2r-2) + (n - \delta) \\ &= \lceil n/2 \rceil + r - 2 \\ &\le \lfloor n/2 \rfloor + r - 1 \\ &< \lfloor n/2 \rfloor + r = \delta. \end{aligned}$$

This contradicts that $\deg(y_{s+1}) \ge \delta$. Thus H has at least r independent edges. Without loss of generality, assume they are

$$\{y_1y_2, y_3y_4, \ldots, y_{2r-1}y_{2r}\}.$$

For convenience, define $G' = (G - x)\backslash\{y_1y_2, y_3y_4, \ldots, y_{2r-1}y_{2r}\}$. By induction, G' can be partitioned into at most $\lfloor (n-1)^2/4 \rfloor$ edges and triangles. These edges and triangles along with the triangles induced by $\{x, y_{2i-1}, y_{2i}\}$, for $i = 1, 2, \ldots, r$, and the edges $\{xy_{2r+1}, xy_{2r+2}, \ldots, xy_\delta\}$ give a partition of G into at most

$$\lfloor (n-1)^2/4 \rfloor + r + (\delta - 2r) = \lfloor (n-1)^2/4 \rfloor + r + \lfloor n/2 \rfloor + r - 2r = \lfloor n^2/4 \rfloor$$

edges and triangles.

For equality to hold in the theorem, we need $\mathrm{cp}(G') = \lfloor (n-1)^2/4 \rfloor$. Hence, by induction, $G' \cong K_{\lfloor (n-1)/2 \rfloor, \lceil (n-1)/2 \rceil}$. We may assume that G is not the complete graph, since $n > 2$ and $\mathrm{cp}(G) > 1$. Thus, there is a vertex, say z, which is not adjacent to x (as x is of minimum degree). But then z is adjacent to at most $\lceil (n-1)/2 \rceil$ vertices in G, as z belongs to one of the parts of G'. So, as $r > 0$,

$$\deg(z) \leq \lceil (n-1)/2 \rceil = \lfloor n/2 \rfloor < \lfloor n/2 \rfloor + r = \delta,$$

contradicting that δ is the minimum degree of G. $\square$

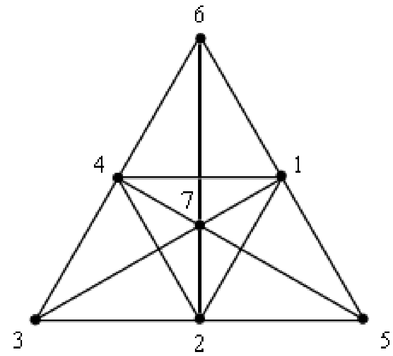
2.3 Steiner systems, projective planes and clique partitions

A rich family of clique partitions of K_n is provided by the existence of Steiner systems. In defining a Steiner system, we use the text by Cameron and Van Lint [9].

A Steiner system $\mathcal{S}(n, k)$ is a pair $(X, \mathcal{B})$, where X is a set of *points* of cardinality n, and $\mathcal{B}$ is a collection of k-element subsets of X called *blocks*, with the property that any two points are contained in precisely one block. We assume that X and $\mathcal{B}$ are non-empty, and that $n \geq k \geq 2$. As any two points of a Steiner system $\mathcal{S}(n, k)$ appear in exactly one block, we can think of the blocks of $\mathcal{S}(n, k)$ as cliques of K_n, such that every edge of K_n appears in exactly one clique. Hence, we get a clique partition of K_n into $\binom{n}{2}/\binom{k}{2}$ cliques of size k.

A *projective plane of order* q is a Steiner system $\mathcal{S}(q^2 + q + 1, q + 1)$. We call the blocks of a projective plane *lines*. Alternatively, a projective plane of order q consists of $q^2 + q + 1$ points, where any two points determine a line; any two lines determine a point; every point has $q + 1$ lines on it; and every line has $q + 1$ points. Projective planes are known to exist for all prime power orders, but no plane of non-prime power order is known.

For example, the projective plane of order 2 provides a clique partition of K_7 into seven cliques, each of size three. This is called the Fano plane and corresponds to the Steiner system $\mathcal{S}(7, 3)$.



The cliques are the triangles induced by the sets

$$\{1,2,4\},\{2,3,5\},\{3,4,6\},\{4,5,7\},\{1,5,6\},\{2,6,7\},\{1,3,7\},$$

which correspond to the lines of the Fano plane.

An *affine plane of order q* is a Steiner system $\mathcal{S}(q^2, q)$. Note that an affine plane can be created from a projective plane by deleting all of the points of a particular block. We call the blocks of an affine plane *lines*, and we say two lines are *parallel* if they are equal or disjoint. Then if L is a line, and p is a point, there is a unique line parallel to L which contains p. Hence, parallelism is an equivalence relation on the set of lines, and the lines in any parallel class partition the set of points. So, any parallel class has q lines, and there are $q + 1$ parallel classes. It is straightforward to prove that there exists an affine plane of order q if and only if there exists a projective plane of order q.

2.4 The de Bruijn-Erdős Theorem

A clique partition is said to be *trivial* if it consists of a single clique, and *non-trivial* otherwise.

For $n \geq 3$, de Bruijn and Erdős [8] prove that if $\mathcal{C}$ is a non-trivial clique partition of K_n, then $|\mathcal{C}| \geq n$. They further showed that equality holds if and only if $\mathcal{C}$ consists of one clique on $n - 1$ vertices and $n - 1$ copies of K_2 incident with a single vertex of K_n; or $\mathcal{C}$ consists of n copies of K_{q+1}, $n = q^2 + q + 1$ and each vertex of K_n is a vertex of exactly $q+1$ cliques of $\mathcal{C}$. This second condition is equivalent to the existence of a projective plane of order q. There are several proofs in the literature. We follow the proof given by Conway, which may be found in Batten and Beutelspacher [4].

Theorem 2.4.1 *If $n \geq 3$ and $\mathcal{C}$ is a non-trivial clique partition of K_n, then $|\mathcal{C}| \geq n$, with equality if and only if*

(i) *$\mathcal{C}$ consists of one clique on $n-1$ vertices and $n-1$ copies of K_2 incident with a single vertex of K_n, or*

(ii) *$\mathcal{C}$ consists of n copies of K_{q+1}, $n = q^2+q+1$ and each vertex of K_n is a vertex of exactly $q+1$ cliques of $\mathcal{C}$.*

Proof. Clearly if (i) or (ii) is satisfied, then we get $|\mathcal{C}| = n$.

Let $\mathcal{C}$ be a clique partition of K_n with size $|\mathcal{C}| \leq n$. We will show that $|\mathcal{C}| = n$ and either (i) or (ii) is satisfied. Let V be the vertex set of K_n. For $v \in V$, let r_v denote the number of cliques of $\mathcal{C}$ that contain the vertex v. Since $n \geq 3$, $|\mathcal{C}| - r_v > 0$, for all vertices $v \in V$. Also, for any clique $K \in \mathcal{C}$, $|\mathcal{C}| - |V(K)| > 0$. This follows as taking any vertex not in $V(K)$ (one exists as $|\mathcal{C}| > 1$) gives rise to a set of $|V(K)|$ edges, no two which may share a clique in $\mathcal{C}$.

For any vertex $v \in V$ and clique $K \in \mathcal{C}$, define

$$\delta(v, K) = \begin{cases} 0 & \text{if } v \text{ is not a vertex of } K \\ 1 & \text{if } v \text{ is a vertex of } K \end{cases}$$

We note that for each $K \in \mathcal{C}$,

$$\sum_{v \in V} \delta(v, K) = |V(K)|,$$

and for every fixed $v \in V$,

$$\sum_{K \in \mathcal{C}} \delta(v, K) = r_v.$$

Now,

$$n = \sum_{v \in V} \frac{|\mathcal{C}| - r_v}{|\mathcal{C}| - r_v} = \sum_{v \in V} \sum_{K \in \mathcal{C}} \frac{1 - \delta(v, K)}{|\mathcal{C}| - r_v}.$$

Note that for any clique $K \in \mathcal{C}$ and any vertex v which is not a vertex of K, then $r_v \geq |V(K)|$. This follows as no two edges joining v to a vertex of K can appear in the same clique. Thus,

$$\frac{1-\delta(v,K)}{|\mathcal{C}|-r_v} \geq \frac{1-\delta(v,K)}{|\mathcal{C}|-|V(K)|} \tag{1}$$

for all cliques $K \in \mathcal{C}$ and vertices $v \in V$. So,

$$\begin{aligned} n &\geq \sum_{v\in V}\sum_{K\in\mathcal{C}} \frac{1-\delta(v,K)}{|\mathcal{C}|-|V(K)|} \\ &= \sum_{K\in\mathcal{C}}\sum_{v\in V} \frac{1-\delta(v,K)}{|\mathcal{C}|-|V(K)|} \\ &= \sum_{K\in\mathcal{C}} \frac{n-|V(K)|}{|\mathcal{C}|-|V(K)|} \\ &\geq \sum_{K\in\mathcal{C}} \frac{n}{|\mathcal{C}|} = n. \end{aligned}$$

The last inequality follows as $|\mathcal{C}| \leq n$ and $|\mathcal{C}| > |V(K)|$ imply

$$\frac{n-|V(K)|}{|\mathcal{C}|-|V(K)|} \geq \frac{n}{|\mathcal{C}|},$$

for any clique $K \in \mathcal{C}$. Thus we must have equality everywhere, implying that $|\mathcal{C}| = n$.

In addition, for each vertex $v \in V$ not in clique K, equality in equation (1) implies $r_v = |V(K)|$. We claim that any two cliques $J, K \in \mathcal{C}$ meet in at least one vertex. Suppose J, K do not meet in a vertex. Take $v \in V(J)$. Then v appears in $|V(K)|$ cliques of $\mathcal{C}$, as $r_v = |V(K)|$. But there are $|V(K)|$ edges from v to K, no two of which appear in the same clique of $\mathcal{C}$. As v also appears in clique J, we get that $r_v \geq |V(K)| + 1$, a contradiction. Hence, any two cliques must meet in at least one vertex.

Note, if two cliques shared a pair of vertices, there would be an edge which appears in two cliques, contradicting that $\mathcal{C}$ is a clique partition. Therefore, any two cliques meet in exactly one vertex. Also, any two vertices determine a clique.

Now let $J, K \in \mathcal{C}$ be any two different cliques from $\mathcal{C}$, and say J, K share vertex v.

First suppose $V(K_n) = V(J \cup K)$. Then $r_v = 2$, implying that all other cliques of $\mathcal{C}$ have size 2. This implies that either J or K is a clique on $n-1$ vertices. Hence, $\mathcal{C}$ satisfies (i).

Now suppose that for any choice of J and K, there is always a vertex of K_n, say w, such that $w \notin V(J \cup K)$. Then $r_w = |V(J)| = |V(K)|$. As J and K were arbitrary, every clique in $\mathcal{C}$ must have the same size, say $q+1$. As there are n cliques in $\mathcal{C}$, an edge count gives,

$$\frac{n(n-1)}{2} = n\frac{(q+1)q}{2},$$

implying, $n = q^2 + q + 1$. Hence, $\mathcal{C}$ satisfies (ii).

Thus it follows that $\mathcal{C}$ satisfies either (i) or (ii). □

Corollary 2.4.2 *If G is a graph on n vertices, and is neither the complete graph nor the empty graph then,*

$$\mathrm{cp}(G) + \mathrm{cp}(\bar{G}) \geq n.$$

Proof. Suppose G is a graph on n vertices, and is neither the complete graph nor the empty graph. If $\mathrm{cp}(G) + \mathrm{cp}(\bar{G}) < n$, then we would get a clique partition of K_n into fewer than n cliques, contradicting Theorem 2.4.1. □

The de Bruijn-Erdős Theorem can be used to construct clique partitions of the complement of small graphs, as done in Sections $2.5 - 2.8$. The Fano plane in Section 2.3 gives an example of Theorem 2.4.1 (ii).

2.5 Complement of a clique

We wish to know how many cliques are required to partition the edge set of $K_n \backslash K_m$, for $m < n$. For $n \geq 3$, Orlin [22] first noted that Theorem 2.4.1 gives $\mathrm{cp}(K_n \backslash K_2) = n-1$ and

$$\mathrm{cp}(K_{q^2+q+1} \backslash K_{q+1}) = q^2 + q$$

whenever a projective plane of order q exists. As for a lower bound, we get

$$\mathrm{cp}(K_n \backslash K_m) \geq n-1$$

for $1 < m < n$, by Corollary 2.4.2.

In the next section we will see a result due to Wallis [26], who shows if $m > 1$ is fixed or $m \leq \sqrt{n}$, then

$$\mathrm{cp}(K_n \backslash K_m) \sim n.$$

Using the edge chromatic number of a graph, Pullman and Donald [23] show that

$$\mathrm{cp}(K_n \backslash K_m) = \frac{1}{2}(n-m)(3m-n+1)$$

for $n > m \geq \frac{1}{2}(n-x)$, (where $x = 0$ for $n-m$ odd, and $x = 1$ otherwise).

To try and fill in the gap, Erdős, Faudree and Ordman [13] use affine planes to prove that for $m = o(n)$,

$$\mathrm{cp}(K_n \backslash K_m) \sim \max\{n, m^2\}.$$

2.5.1 The Pullman-Donald Theorem

Let $\chi'(G)$ denote the edge chromatic number of G, that is, the minimum number of colours required to colour the edges of G, so that no pair of adjacent edges have the same colour.

We define the *join* $H_1 \vee H_2$ of two vertex disjoint graphs H_1 and H_2, to be the graph with vertex set $V(H_1) \cup V(H_2)$, and edge set consisting of $E(H_1) \cup E(H_2)$ and edges ij, where $i \in V(H_1)$ and $j \in V(H_2)$. If H is a graph, then Pullman and Donald [23] compute the clique partition number of the join of H and $\bar{K}_q$, for $q \geq \chi'(H)$. Choosing $H = K_n$ will then provide us with the clique partition number of $K_n \backslash K_m$, for particular values of m.

Theorem 2.5.1 *Let H be a graph with h vertices and e edges. If $q \geq \chi'(H)$ then $\mathrm{cp}(H \vee \bar{K}_q) = hq - e$. Further, any minimal clique partition consists of edges and triangles only.*

Proof. Let H be a graph with h vertices and e edges and fix $q \geq \chi'(H)$. We first construct a clique partition of $H \vee \bar{K}_q$ into $hq - e$ edges and triangles. Suppose E_i is the set of edges of colour i in a minimal edge colouring of H, for $1 \leq i \leq \chi'(H)$. For convenience, we denote the vertices of $\bar{K}_q$ by $\{v_1, v_2, \ldots, v_q\}$.

Note that E_i is a matching. Let T_i be the family of $|E_i|$ triangles with one vertex at v_i and the opposite edge in E_i. This gives a clique partition of $H \vee \bar{K}_q$ into edges and triangles. Namely, we use the triangles from $T_1, T_2, \ldots, T_{\chi'(H)}$ and all of the remaining edges. As every edge of H belongs to a triangle using a vertex from $\bar{K}_q$, there are e triangles. Removing these triangles from $H \vee \bar{K}_q$ gives a triangle-free graph with $hq - 2e$ edges. Thus,

$$\mathrm{cp}(H \vee \bar{K}_q) \leq e + (hq - 2e) = hq - e.$$

Now suppose $\mathcal{C}$ is a clique partition of $H \vee \bar{K}_q$ having r of its members $C_1, C_2, \ldots, C_r$ with edges in H. For convenience, let $c_j = |V(C_j) \cap V(H)|$. For each C_j, the clique on vertices $V(C_j) \cap V(H)$ has at least $c_j - 1$ edges, with equality if $|V(C_j) \cap V(H)| = 2$. Thus, $e + r \geq \sum_{j=1}^{r} c_j$. If $e + r = \sum_{j=1}^{r} c_j$ then no member of $\mathcal{C}$ has size more than three. Let s be the number of edges outside of H that are covered by $C_1, C_2, \ldots, C_r$. Then $s \leq \sum_{j=1}^{r} c_j$. Note that $|\mathcal{C}| = r + (hq - s)$, as removing the $C_1, C_2, \ldots, C_r$ from $\mathcal{C}$ leaves a triangle-free graph with $hq - s$ edges. Then,

$$|\mathcal{C}| = r + (hq - s) \geq r + hq - \sum_{j=1}^{r} c_j \geq hq - e.$$

If we have equality, then no member of $\mathcal{C}$ may have more than three vertices. Hence, $\mathrm{cp}(H \vee \bar{K}_q) = hq - e$, and any minimal clique partition uses only triangles and edges. □

If $q < \chi'(H)$ then Theorem 2.5.1 may be false. For example, if $n \geq 10$, $H = K_{n-3}$ and $q = 3$, then Theorem 2.5.1 gives $\mathrm{cp}(H \vee \bar{K}_q) = (n-3)3 - \binom{n-3}{2} \leq 0$, but $\mathrm{cp}(H \vee \bar{K}_q) > 0$ as $H \vee \bar{K}_q$ is not the empty graph.

To get a result on the complement of clique, we note that $\chi'(K_{2k}) = \chi'(K_{2k-1}) = 2k - 1$. Thus when $m \geq \chi'(K_{n-m})$ we get $\mathrm{cp}(K_n \backslash K_m) = (n-m)m - \binom{n-m}{2}$ by Theorem 2.5.1.

Theorem 2.5.2 *For $n > m \geq \frac{1}{2}(n - x)$, (where $x = 0$ for $n - m$ odd, and $x = 1$ otherwise),*

$$\mathrm{cp}(K_n \backslash K_m) = \frac{1}{2}(n-m)(3m - n + 1).$$

Also, any minimal clique partition consists of edges and triangles only.

2.5.2 The Erdős-Faudree-Ordman Theorem

The proof of Theorem 2.5.1 relies on the fact that there are hq edges connecting H to $\bar{K}_q$; two lie in the same clique only if that clique contains at least one edge selected from the e edges of H. Erdős, Faudree and Ordman [13] make use of this fact to extend several existing lower bounds on the clique partition number of a graph.

We first need a numerical lemma.

Lemma 2.5.3 *Let $\sum_{i=1}^{q} e_i = c$ and $\sum_{i=1}^{q} e_i^2 \leq d$. Then $q \geq c^2/d$.*

Proof. Note that the e_i's may be distinct. If we substitute $e_i = c/q$, for $i = 1, 2, \ldots, q$, then the first sum is preserved. This substitution can only decrease $\sum_{i=1}^{q} e_i^2$, as the sum of the e_i^2 is minimized when $e_1 = e_2 = \cdots = e_q = c/q$. This is because $\sum_i x_i^2$ is a convex function. Thus, $\sum_{i=1}^{q}(c/q)^2 \leq d$ implying that $q(c/q)^2 \leq d$. □

Let G be a graph with n vertices, and partition $V(G)$ into two sets A and B (which we call *sides*) such that $A \cap B = \emptyset$. The edges of the graph G now fall into three classes, which we call *A edges*, *B edges*, and *connecting edges*, depending as to whether their ends lie both in A, both in B, or one in each. If a clique in G contains more than one of the connecting edges of G, then it must contain some A edges or B edges (or both). If the number of connecting edges is large, then there will not be enough A edges or B edges of G available to combine the connecting edges into a few number of cliques. Erdős, Faudree and Ordman [13] provide a lower bound on $\mathrm{cp}(G)$ if one of the sides has no edges.

Lemma 2.5.4 *Suppose G is a graph with n vertices, and $A, B \subseteq V(G)$ such that $A \cap B = \emptyset$. If G has k edges in side A, no edges in side B, and c connecting edges, then*

$$\mathrm{cp}(G) \geq \frac{c^2}{(2k + c)}.$$

Proof. If G is partitioned by q cliques and clique i has e_i connecting edges, then clique i has $e_i(e_i - 1)/2$ edges in side A. Then $\sum_{i=1}^{q} e_i = c$ and

$$k \geq \sum_{i=1}^{q} \frac{e_i(e_i - 1)}{2} = \frac{1}{2}\left(\sum_{i=1}^{q} e_i^2 - \sum_{i=1}^{q} e_i\right) = \frac{1}{2}\left(\sum_{i=1}^{q} e_i^2 - c\right).$$

Hence, $\sum_{i=1}^{q} e_i^2 \leq 2k + c$, and the result follows by Lemma 2.5.3. □

Thus we obtain a lower bound for the clique partition number of the complement of a clique.

Theorem 2.5.5 *For $1 < m < n$,*

$$\mathrm{cp}(K_n \backslash K_m) \geq \frac{(n-m)m^2}{(n-1)}.$$

Proof. Substituting $c = m(n-m)$ and $k = \binom{n-m}{2}$ into Lemma 2.5.4 gives,

$$\mathrm{cp}(K_n \backslash K_m) \geq \frac{(n-m)^2 m^2}{(n-m)(n-m-1) + m(n-m)} = \frac{(n-m)m^2}{(n-1)}.$$

□

To get an upper bound for the complement of a clique, Erdős, Faudree and Ordman [13] use a modification of the strategy that Wallis [26] used, that is, affine planes. One problem is our lack of knowledge of affine planes when q is not a prime power. To get around this problem we use the following result.

Lemma 2.5.6 *There exists a constant $\alpha \in [1/2, 1)$, such that if $p < q$ are consecutive primes, then,*

$$q - p = O(p^{\alpha}).$$

Note that $\alpha = 21/40$ is possible, as shown in [3] by Baker, Harman and Pintz.

By constructing a clique partition using an affine plane, Erdős, Faudree and Ordman [13] showed that if $m = f(n)$ and for n sufficiently large, $\sqrt{n} < m < n$, we have $\mathrm{cp}(K_n \backslash K_m) < m^2 + o(m^2)$.

Theorem 2.5.7 *Let $m = f(n)$ and $\sqrt{n} < m < n$. Then for n sufficiently large,*

$$\mathrm{cp}(K_n \backslash K_m) < m^2 + o(m^2).$$

Proof. Let t be the smallest integer such that $t \geq m$ and there is an affine plane of order t. By Lemma 2.5.6, we may assume for sufficiently large m,

$$m \leq t \leq m + cm^{\alpha},$$

for some constant c and $\alpha \in [1/2, 1)$. Thus, $t^2 = m^2 + o(m^2)$. In an affine plane of order t, choose a line L, and delete all but m points from L. In the other lines, delete a total of $(t^2 - n) - (t - m)$ points. This leaves a total of n points, with m of them on a selected line. Use this design to construct a clique partition of $K_n \backslash K_m$ into at most $t^2 + t - 1$ cliques. Thus,

$$\mathrm{cp}(K_n \backslash K_m) \leq t^2 + t - 1 < m^2 + o(m^2),$$

as required. □

Theorem 2.5.8 *If $m = o(n)$ and $m \geq \sqrt{n}$, then*

$$\mathrm{cp}(K_n \backslash K_m) \sim m^2.$$

Proof. If $m = o(n)$ and $m \geq \sqrt{n}$ then for a lower bound, we note that for n large enough, Theorem 2.5.5 gives,

$$\mathrm{cp}(K_n \backslash K_m) \geq \frac{(n-m)m^2}{n-1} \sim m^2.$$

For an upper bound, we note that for n large enough, Theorem 2.5.7 gives,

$$\mathrm{cp}(K_n \backslash K_m) \leq m^2 + o(m^2) \sim m^2.$$

□

If $m \leq \sqrt{n}$, we will see in the next section that

$$\mathrm{cp}(K_n \backslash K_m) \sim n.$$

For comparison, we note that Theorem 2.5.2 gives,

$$\mathrm{cp}(K_n \backslash K_m) \sim \frac{(1-c)(3c-1)}{2c^2} m^2$$

when $m = cn$, for some constant $1/2 \leq c \leq 1$. If $c = 1/2$ this gives $\frac{1}{8}n^2$, and if $c = 3/4$ this gives $\frac{5}{32}n^2$. Various authors have analyzed the case $c < 1/2$. See Monson, Pullman, and Rees [21] for these results.

2.6 Complement of small graphs

In the previous section, we analyzed the clique partition number of the graph $K_n \backslash H$, where H is K_m. What can be said about the clique partition number of $K_n \backslash H$, if H is not a clique?

As in Theorem 2.5.7, we will exploit the second part of Theorem 2.4.1 to construct clique partitions of complements of small graphs. Wallis [26, 28] examined the asymptotic behaviour of the clique partition number of complements of graphs with few vertices, thus proving (i) and (ii) of the following theorem. Using the same argument, we can prove if H has a subgraph H' which has $O(\sqrt{n})$ components, with each component having at most $\sqrt{n}$ vertices, and further if H has $O(\sqrt{n})$ edges between components of H', then $\mathrm{cp}(K_n \backslash H) = O(n^{3/2})$.

Theorem 2.6.1 *Let H be a graph.*

(i) *If H has $\sqrt{n}$ or less vertices, then,*

$$\mathrm{cp}(K_n \backslash H) = O(n).$$

(ii) *If H has $o(\sqrt{n})$ vertices then,*

$$\mathrm{cp}(K_n \backslash H) \sim n.$$

(iii) *If H has a subgraph H' which has $O(\sqrt{n})$ components, with each component having at most $\sqrt{n}$ vertices, and further if H has $O(\sqrt{n})$ edges between components of H', then*

$$\mathrm{cp}(K_n \backslash H) = O(n^{3/2}).$$

Proof. Let H be a graph satisfying one of the three conditions in the theorem and suppose H has h vertices. Let t be the smallest integer such that $t \geq \sqrt{n}$, and such that there is a projective plane of order t. Then by Lemma 2.5.6, $t^2 = n + o(n)$. Since $K_n \backslash H$ is an induced subgraph of $K_{t^2+h} \backslash H$, by Lemma 2.1.2,

$$\mathrm{cp}(K_n \backslash H) \leq \mathrm{cp}(K_{t^2+h} \backslash H)$$

Suppose H satisfies (i) or (ii). Then take a finite projective plane of order t, and construct a copy of K_{t^2+t+1} with a distinguished subgraph H. Do this by identifying points of the plane with vertices in such a way that the vertices of H are identified with the points of one block, say B. This can be done since, $h \leq t$. Then we delete from K_{t^2+t+1}, all the points of B except the ones belonging to H. This results in a K_{t^2+h} partitioned into t^2+t+1 cliques, with one clique of size h, and the remaining cliques of size t or $t+1$. We replace the clique of size h by a clique partition of $K_h \backslash H$ using $\mathrm{cp}(K_h \backslash H)$ cliques. This gives a clique partition of $K_{t^2+h} \backslash H$ into $t^2+t+\mathrm{cp}(K_h \backslash H)$ cliques. Hence,

$$\mathrm{cp}(K_n \backslash H) \leq \mathrm{cp}(K_{t^2+h} \backslash H) \leq t^2+t+\mathrm{cp}(K_h \backslash H).$$

If H satisfies (i) then, $\mathrm{cp}(K_h \backslash H) \leq \frac{1}{2}h(h-1) < t^2$, giving that

$$\mathrm{cp}(K_n \backslash H) = O(t^2) = O(n).$$

If H satisfies (ii), then $h = o(\sqrt{n})$, implying $\frac{1}{2}h(h-1) = o(n)$. Thus, $\mathrm{cp}(K_n \backslash H) \leq n + o(n)$ But $\mathrm{cp}(K_n \backslash H) \geq n - \sqrt{n}$ by Theorem 2.4.2, thus,

$$\mathrm{cp}(K_n \backslash H) \sim n.$$

Finally, suppose H satisfies (iii). Thus, the vertices of H can be identified with points in a projective plane of order t such that the edges of H' appear in $r = O(\sqrt{n})$ blocks, say, $B_1, B_2, \ldots, B_r$. Suppose there are $e = O(\sqrt{n})$ edges in H not in H', then, $K_{t^2+t+1} \backslash H$ can be partitioned into at most

$$t^2+t+1-(r+e)+(r+e)(n/2)+ \leq n+o(n)+O(n^{3/2}) = O(n^{3/2})$$

cliques, as the edges of H not in H' are contained in at most e blocks and each block which contains edges of H can be partitioned into less than $n/2$ edges. Hence, we get a clique partition of $K_n \backslash H$ into at most $O(n^{3/2})$ cliques, by deleting points not identified with a vertex of H. □

Note that if H is the graph K_m, for $m \leq \sqrt{n}$, then we get

$$\mathrm{cp}(K_n \backslash K_m) \sim n,$$

as $\mathrm{cp}(K_h \backslash H)$ would be zero in the proof of Theorem 2.6.1 (i). Thus, along with Theorem 2.5.8 we have established the following theorem.

Theorem 2.6.2 *If $m = o(n)$, then*

$$\mathrm{cp}(K_n \backslash K_m) \sim \max\{n, m^2\}.$$

More generally, the proof of Theorem 2.6.1 (iii) implies that if H has a subgraph H' which has $O(\sqrt{n})$ components, with each component having at most $\sqrt{n}$ vertices, and there are $o(n)$ edges connecting these components, then

$$\mathrm{cp}(K_n \backslash H) = o(n^2).$$

2.7 Complement of paths, cycles, and perfect matchings

If we fix H in Theorem 2.6.1(iii) to be a particular graph, we can provide better bounds on the clique partition number of the complement of H. In particular, consider the clique partition numbers of the complement of a path P_n, a cycle C_n, and a perfect matching M_n each on n vertices.

In this essay, $\bar{M}_n$ denotes the complement of a perfect matching on n vertices, where n must be even. For convenience, if n is odd, we use the notation $\bar{M}_n$ to denote the complement graph $K_n \backslash M_{n-1}$. It was Orlin [22] who first asked about the asymptotics of the clique partition number of the complement of a perfect matching. Wallis [27] showed that for any fixed $\epsilon > 0$, $\mathrm{cp}(\bar{M}_n) = o(n^{1+\epsilon})$. Later, Gregory, McGuinness and Wallis [15] proved for n sufficiently large, we have

$$n \leq \mathrm{cp}(\bar{M}_n) \leq (1 + o(1))n \log_2 \log_2 n$$

and conjecture that $\mathrm{cp}(\bar{M}_n) \sim n$.

The first result regarding the clique partition number of the complement of a cycle is by Wallis [29], who shows $\mathrm{cc}(\bar{C}_n) < \frac{3}{16}n^2$. Using the same argument as for the complement of a perfect matching, Wallis [28] proves that for the path, $\mathrm{cp}(\bar{P}_n) \leq (1 + o(1))n \log_2 \log_2 n$, for n sufficiently large. Wallis [28] also notes (but does not prove), that this technique can be extended to the complement of H, where every vertex of H is of degree one or two, namely C_n. Since de Caen and Gregory [10] prove that for $n \geq 11$, $\mathrm{cp}(\bar{C}_n) \geq n$, for the complement of a cycle we have,

$$n \leq \mathrm{cp}(\bar{C}_n) \leq (1 + o(1))n \log_2 \log_2 n.$$

Lemma 2.1.2 gives $\mathrm{cp}(\bar{P}_n) \geq \mathrm{cp}(\bar{C}_{n-1})$. Thus, for the complement of a path we have,

$$n - 1 \leq \mathrm{cp}(\bar{P}_n) \leq (1 + o(1))n \log_2 \log_2 n.$$

In this essay, we will prove the result $\mathrm{cp}(\bar{P}_n) \leq (1+o(1))n \log_2 \log_2 n$ by Wallis [28]. Note that this result can be used to show that $\mathrm{cp}(\bar{C}_n)$ and $\mathrm{cp}(\bar{M}_n)$ are $O(n \log_2 \log_2 n)$, by Lemma 2.1.2.

Theorem 2.7.1 *If P_n is the path on n vertices then,*

$$\mathrm{cp}(\bar{P}_n) \leq (1 + o(1))n \log_2 \log_2 n,$$

for n sufficiently large.

Proof. Note that by Lemma 2.1.2, $\mathrm{cp}(\bar{P}_n) \leq \mathrm{cp}(\bar{P}_r)$, for $n \leq r$. Let $d = \lfloor \sqrt{n} \rfloor$ and choose e to be the smallest integer such that $de \geq n$.

Let t be the smallest prime power such that $t \geq \sqrt{n}$. We embed a path P_n in a copy of an affine plane of order t. One parallel class, say P_1, is chosen and all but d of its lines are deleted. Now select another parallel class P_2 and delete $t - e$ of its lines, leaving e lines. Denote the point of intersection of line i of P_1 with line j of P_2 by a_{ij}. Replace line j of P_2 by a copy of $\overline{P_d}$, in particular, the complement of path $(a_{1j}, a_{2j}, \ldots, a_{dj})$, for $j = 1, 2, \ldots, e$. Now replace line 1 of P_1 by the complement of path $(a_{11}, a_{12}, \ldots, a_{1e})$ and line d of P_1 by the complement of path $(a_{d1}, a_{d2}, \ldots, a_{de})$. What we have done is selected $e+2$ blocks and replaced them with copies of complements of paths, while the remaining blocks are left untouched. By construction, we have an embedded path P_{ed} along with $e - 1$ additional edges in an affine plane of order t. Thus,

$$\begin{aligned} \mathrm{cp}(\bar{P}_n) &\leq \mathrm{cp}(\bar{P}_{ed}) \leq t^2 + t - (e + 2) + e\,\mathrm{cp}(\bar{P}_d) + 2\,\mathrm{cp}(\bar{P}_e) + e - 1, \\ &\leq t^2 + t - 3 + e\,\mathrm{cp}(\bar{P}_d) + 2[\mathrm{cp}(\bar{P}_d) + 3d - 1], \end{aligned}$$

since $\bar{P}_e$ can be partitioned into at most $\mathrm{cp}(\bar{P}_d) + 3d - 1$ cliques, as $e \leq d + 3$. So,

$$\begin{aligned} \mathrm{cp}(\bar{P}_n) &\leq t^2 + d\,\mathrm{cp}(\bar{P}_d) + 5\,\mathrm{cp}(\bar{P}_d) + t + 6d - 5, \\ &\leq t^2 + d\,\mathrm{cp}(\bar{P}_d) + o(d^2). \end{aligned}$$

by Theorem 2.6.1 (iii), as the path satisfies the condition of the theorem. We write $c(x) = \mathrm{cp}(\bar{P}_{\lfloor x \rfloor})$, so that,

$$c(x) \leq t^2 + \sqrt{x}c(\sqrt{x}) + o(x),$$

where t is the closest integer bigger than $\sqrt{x}$. Hence,

$$c(x) \leq x + \sqrt{x}c(\sqrt{x}) + o(x).$$

By Lemma 5.2.1,

$$c(x) \leq (1 + o(1))x \log_2 \log_2 x,$$

for x sufficiently large. □

2.8 Complement of forests

In this section, we use projective planes to prove that if F_n is a forest on n vertices, then the edge set of $\bar{F}_n$ can be partitioned into $O(n(\log_2 n)^{\log_2 3})$ cliques. We use the techniques of the previous section to set up a recursion and solve it.

But first, we need to know that we can partition F_n into components of size at most $\sqrt{n}$ such that two components intersect in at most one vertex. If the components are small, then this corresponds to using lots of blocks of our projective plane, thus increasing the number of cliques required to partition $\bar{F}_n$. Hence, we will prove a lower bound on the size of the components.

We assume that we are working with trees, as a clique partition for $\bar{F}_n$ can be extended to a clique partition of the complement of some tree T_n with $E(F_n) \subseteq E(T_n)$, by adding at most $n-1$ edges (namely edges $E(T_n) \backslash E(F_n)$) to our clique partition of $\bar{F}_n$.

Definition 2.8.1 A *tree partition* of a tree T_n is a collection of subtrees

$$\{T^1, T^2, \ldots, T^r\},$$

such that every edge of T_n is in exactly one subtree, $T_n = \cup_{i=1}^r T^i$ and

$$|V(T^i) \cap V(T^j)| \leq 1$$

for all $i \neq j$.

For any positive integers k and b with $2 \leq k \leq n$ and $b > 1$, we say a tree partition is (k, b)-*smooth* if $k/b \leq |T^i| \leq k$, for $i = 1, 2, \ldots, r$. We will prove that a $(k, 3)$-smooth tree partition of T_n always exists.

Lemma 2.8.2 *Let T_n be a tree on n vertices, and $2 \leq k \leq n$ be a positive integer. Then there exists a $(k, 3)$-smooth tree partition of T_n.*

Proof. Let T_n be a tree on n vertices, and $2 \leq k \leq n$ be a positive integer. Note that there is always a tree partition of T_n into two subtrees. It is well known that there exists a tree partition $\{T^1, T^2\}$ of T_n such that $n/3 \leq |T^i| \leq 2n/3$, for $i = 1, 2$. To see this, take an arbitrary tree partition $\{T^1, T^2\}$ of T_n so that $||V(T^1)| - |V(T^2)||$ is minimized, and assume T^1 and T^2 share vertex v. If $n/3 \leq |T^i| \leq 2n/3$, for $i = 1, 2$, then we are done. Without loss of generality, suppose that $|T^1| < n/3$. As $||V(T^1)| - |V(T^2)||$ is minimized, v is adjacent to at least two vertices of T^2. Form a tree partition $\{J^1, J^2\}$ of T^2, such that J^1 and J^2 share vertex v. Then,

$$\frac{2n}{3} + 1 < \{|T^1| + |J^1|\} + \{|T^1| + |J^2|\} < \frac{4n}{3} + 1,$$

as $2n/3 + 1 < 2|T^1| + |T^2| + 1 < 4n/3 + 1$. Hence, $n/3 < |V(T^1 \cup J^1)| < 2n/3$ or $n/3 < |V(T^1 \cup J^2)| < 2n/3$. Thus, we constructed a tree partition of T_n consisting of two subtrees that share a vertex, with sizes between $n/3$ and $2n/3$.

We will construct a $(k, 3)$-smooth tree partition $\mathcal{S}$ of T_n as follows. Repeatedly split trees using the above argument that have size more than k. Then, for all subtrees T^i of $\mathcal{S}$, we have that $k/3 \leq |T^i| \leq k$, as required. We get a $(k, 3)$-smooth tree partition of T_n, as each time we split a subtree, the two resulting trees share exactly one vertex. □

For convenience of the next proof, we define the following functions $g(n)$ and $h(x)$, where $n \in \mathbb{N}$ and $x \in \mathbb{R}^+$,

$$g(n) = \max\{\mathrm{cp}(\bar{T}_n) : T_n \text{ is a tree on } n \text{ vertices}\},$$

$$h(x) = \max\{\mathrm{cp}(K_{9\lfloor x \rfloor} \backslash T_{\lfloor x \rfloor}) : T_{\lfloor x \rfloor} \text{ is a tree on } \lfloor x \rfloor \text{ vertices}\}.$$

Note that $g(n) \leq h(n)$, and $\mathrm{cp}(\bar{T}_n) \leq g(n)$, for any tree T_n.

Theorem 2.8.3 *If F_n is a forest on n vertices then,*

$$\mathrm{cp}(\bar{F}_n) = O(n(\log_2 n)^{\log_2 3}),$$

for n sufficiently large.

Proof. We assume that we are working with trees, as a clique partition for $\bar{F}_n$ can be extended to a clique partition of $\bar{T}_n$, for some tree T_n with $E(F_n) \subseteq E(T_n)$, by adding $O(n)$ edges to our clique partition of $\bar{F}_n$. Since $\mathrm{cp}(\bar{T}_n) \leq g(n)$, for any tree T_n, it suffices to show that $g(n) = O(n(\log_2 n)^{\log_2 3})$. Let T_n be a tree on n vertices such that $h(n) = \mathrm{cp}(K_{9n} \backslash T_n)$. Let t be the smallest integer such that $t \geq 4\sqrt{n} + 11$ and there is a projective plane of order t. Lemma 2.5.6 gives that $t = 4\sqrt{n} + o(\sqrt{n})$ which gives $t^2 = 16n + o(n)$ upon squaring. Let $\mathcal{S} = \{T^1, T^2, \ldots, T^r\}$ be a $(\lfloor\sqrt{n}\rfloor, 3)$-smooth tree partition of T_n, which exists by Lemma 2.8.2. Then, we have for $i = 1, 2, \ldots, r$,

$$\frac{\lfloor\sqrt{n}\rfloor}{3} \leq |T^i| \leq \lfloor\sqrt{n}\rfloor.$$

Without loss of generality, suppose that

$$V(T^i) \cap \left(\bigcup_{j=1}^{i-1} V(T^j)\right) = \{v_i\}$$

for $i = 2, 3, \ldots, r$. Note that we need $|T^1| + (|T^2| - 1) + (|T^3| - 1) + \cdots + (|T^r| - 1) = n$, which implies that

$$r \leq \frac{3(n-1)}{\sqrt{n} - 4}$$

as $\sqrt{n} \leq \lfloor\sqrt{n}\rfloor + 1$. Then we get $r \leq 3\sqrt{n} + 13$, if n is large enough.

Choose a projective plane of order t. We claim that we can find blocks

$$\{B^1, B^2, \ldots, B^r\}$$

such that B^i contains T^i, no two vertices of T_n are identified with the same point of the projective plane and the blocks B^i have the same tree like structure as the subtrees T^i. First identify the vertices of T^1 with points from an arbitrary block, say B^1, of the projective plane, where the vertex v_2 is identified with some point w_2, and all other vertices of T^1 are identified arbitrarily with points from $B^1 \backslash \{w_2\}$. We will show how to find blocks $B^2, B^3, \ldots, B^r$ satisfying the requirements. Suppose for some

$2 \leq i \leq r$, we have already identified the vertices of T^{i-1} with the points of B^{i-1}, such that vertex v_i is identified with point w_i of some block B^j, where $j \leq i-1$. Pick a block B^i (different from $B^1, B^2, \ldots, B^{i-1}$) that contains the point w_i. There exists such a block as there are $t+1 \geq 4\sqrt{n}+11$ blocks containing the point w_i (and at most $r \leq 3\sqrt{n}+13$ blocks have been used). Identify the vertices of T^i with points from B^i such that v_i is identified with w_i, and all other vertices of T^i are identified arbitrarily with points from $B^i \backslash W$, where W is the set of points from $B^1 \cup B^2 \cup \cdots \cup B^{i-1}$ that intersect with B^i. Note that $|W| < r$, as B^i intersects every other block in at most one point. This identification can be done, as each block has $t+1 \geq 4\sqrt{n}+12$ points, each tree has at most $\lfloor \sqrt{n} \rfloor \leq \sqrt{n}$ vertices, and removing at most $r-1 \leq 3\sqrt{n}+12$ points of block B^i leaves at least $\sqrt{n}$ points which can be identified with T^i. This gives a collection of blocks $\{B^1, B^2, \ldots, B^r\}$ that contain the edges of $\bar{T}_n$.

Then by Lemma 2.1.2,

$$g(n) \leq h(n) = \mathrm{cp}(K_{9n} \backslash T_n) \leq \mathrm{cp}(K_{t^2+t+1} \backslash T_n).$$

As r is the number of blocks B^i that contain edges of T_n, we have

$$\begin{aligned} h(n) &\leq t^2+t+1-r+\sum_{i=1}^{r} \mathrm{cp}(K_{|B^i|} \backslash T^i) \\ &\leq O(n)+\sum_{i=1}^{r} \mathrm{cp}(K_{|B^i|} \backslash T^i) \\ &\leq O(n)+\sum_{i=1}^{r} \mathrm{cp}(K_{9\lfloor\sqrt{n}\rfloor} \backslash T^i). \end{aligned}$$

The last inequality follows as, Lemma 2.5.6 implies $t+1 \leq 4\sqrt{n}+o(\sqrt{n}) \leq 9\lfloor\sqrt{n}\rfloor$, for n sufficiently large. But, if T_p is an induced subgraph of T_q, then $\mathrm{cp}(\bar{T}_p) \leq \mathrm{cp}(\bar{T}_q)+q-p$, for $p \leq q$, by Lemma 2.1.2. Hence, we can extend the trees T^i to have $\lfloor\sqrt{n}\rfloor$ vertices. Then we have,

$$\begin{aligned} h(n) &\leq O(n)+r \cdot h(\sqrt{n}) \\ &\leq O(n)+(3\sqrt{n}+13)h(\sqrt{n}) \\ &= O(n)+3\sqrt{n}h(\sqrt{n}) \end{aligned}$$

since $h(\sqrt{n}) = O(n)$. Hence,

$$h(x) \leq O(x)+3\sqrt{x}h(\sqrt{x}).$$

Dividing through by x and setting $z = \log_2 \log_2 x$, and $r(z) = h(x)/x$ gives,

$$r(z) \leq O(1) + 3 \cdot r(z-1).$$

So,

$$r(z) = O(3^z)$$

for x (and hence z) arbitrarily large. Hence,

$$h(x) < O(x3^z) = O(x(\log_2 x)^{\log_2 3}),$$

implying that

$$g(n) = O(n(\log_2 n)^{\log_2 3}).$$

□

It would be interesting to know if there is a forest F_n such that

$$\lim_{n\to\infty} \frac{\mathrm{cp}(\bar{F}_n)}{n} = \infty.$$

Regardless, we conjecture that a bound of $O(n \log_2 \log_2 n)$ is possible.

Conjecture 2.8.4 *If F_n is a forest on n vertices then,*

$$\mathrm{cp}(\bar{F}_n) = O(n \log_2 \log_2 n),$$

for n sufficiently large.

We believe that F_n can be partitioned into $(1 + o(1))\sqrt{n}$ components each with size $(1 + o(1))\sqrt{n}$, such that any two components intersect in at most one vertex. If this is true, then Theorem 2.8.3 can be adapted to give $O(n \log_2 \log_2 n)$.

2.9 Complement of graphs with bounded maximum degree

The Probabilistic Method has recently been developed intensively. We use the terminology from Alon [2] and follow Beth, Jungnickel and Lenz [5] regarding the existence of Steiner systems. Necessary existence conditions for the existence of an $\mathcal{S}(n, k)$ are

$$n \equiv 1 \mod k-1,$$

$$n(n-1) \equiv 0 \mod k(k-1).$$

Wilson's theorem [5] says that the necessary conditions above for the existence of an $\mathcal{S}(n,k)$ are sufficient for almost all $n \in \mathbb{N}$. However, the proofs presented by Wilson do not give an explicit constant $n_0(k)$ such that an $\mathcal{S}(n,k)$ exists for all $n \geq n_0(k)$ satisfying the necessary conditions. Recently, Chang showed that $n_0(k) \leq \exp(\exp(k^{k^2}))$ (see page 800 in [5]). In the next theorem, we will show that if a graph G has bounded maximum degree, then the existence of a Steiner system gives rise to a clique partition of $\bar{G}$.

Theorem 2.9.1 *Let G be a graph on g vertices with maximum degree $\Delta = o(n/\log^4 n)$, where $n = O(g)$. Let $k = \lceil \sqrt{n}/\sqrt{2\Delta} \rceil$. If n is large enough, and if a Steiner system $\mathcal{S}(n,k)$ exists then,*

$$\mathrm{cp}(\overline{G}) = O(n^{3/2}\sqrt{\Delta}(\log_2 n)^2).$$

Proof. Suppose G, n, k satisfy the conditions of the theorem. Let $\mathcal{S} = (X, \mathcal{B})$ be a Steiner system with blocks of size k on n points. For a random permutation of the points, the probability that a fixed set of k points is a fixed block in $\mathcal{B}$ is exactly $1/\binom{n}{k}$. Take G to be a fixed graph on the same set of n points, with maximum degree Δ. Let G_B denote the subgraph of G spanned by the edges contained in a block $B \in \mathcal{B}$.

Consider the event $|E(G_B)| \geq r$, for some integer r. Pick a subgraph H_B of G_B with exactly r edges. If the maximum size of a matching in H_B is i for some positive integer $i \leq r$, and if there are s vertices of H_B which are unsaturated by a maximum matching, then

$$\max\left\{\frac{r - \binom{2i}{2}}{2i}, 0\right\} \leq s \leq r - i.$$

For convenience, let $s_i = (r - \binom{2i}{2})/2i$. Let $A_B(i,s)$ denote the event that the largest matching in H_B has size i and H_B has $s + 2i$ vertices. Fixing a matching M of size i in G, there are at most $(2i\Delta)^s$ ways to choose a set S of s vertices so that $V(M) \cup S = H_B$. Then there are $\binom{n-2i-s}{k-2i-s}$ ways to choose the vertices of $B \backslash V(H_B)$ from G. Therefore

$$\mathbb{P}[A_B(i,s)] \leq \frac{1}{\binom{n}{k}} \binom{\Delta n}{i} (2i\Delta)^s \binom{n-2i-s}{k-2i-s}.$$

Since $\{|E(G_B)| \geq r\} \subset \bigcup_{i,s} A_B(i,s)$, it follows that

$$\mathbb{P}[|E(G_B)| \geq r] \leq \frac{1}{\binom{n}{k}} \sum_{i=1}^{r} \sum_{\substack{s \geq s_i \\ s \geq 0}}^{r-i} \binom{\Delta n}{i} (2i\Delta)^s \binom{n-2i-s}{k-2i-s}.$$

To estimate the sums on the right, we use the inequality

$$\frac{\binom{a-t}{b-t}}{\binom{a}{b}} < \frac{b^t}{a^t}.$$

Let j be the largest integer such that

$$r - \binom{2j}{2} \geq 0$$

so that definitely $\sqrt{r/2} - 1 \leq j \leq \sqrt{r/2} + 1$. Put $k = \lceil \sqrt{n}/\sqrt{2\Delta} \rceil$. Then

$$\begin{aligned}
\mathbb{P}[|E(G_B)| \geq r] &\leq \sum_{i=1}^{r} \sum_{\substack{s \geq s_i \\ s \geq 0}}^{r-i} \binom{\Delta n}{i} (2i\Delta)^s \frac{\binom{n-2i-s}{k-2i-s}}{\binom{n}{k}} \\
&< \sum_{i=1}^{r} \sum_{\substack{s \geq s_i \\ s \geq 0}}^{r-i} (\Delta n)^i (2r\Delta)^s \frac{k^{2i+s}}{n^{2i+s}} \\
&= \sum_{i=1}^{r} \left(\frac{\Delta k^2}{n}\right)^i \sum_{\substack{s \geq s_i \\ s \geq 0}}^{r-i} \left(\frac{2rk\Delta}{n}\right)^s \\
&\leq \sum_{i=1}^{j} \left(\frac{1}{2}\right)^i \sum_{s=s_i}^{r-i} \left(\frac{1}{2}\right)^s + \sum_{i=j+1}^{r} \left(\frac{1}{2}\right)^i \sum_{s=0}^{r-i} \left(\frac{1}{2}\right)^s \qquad (*) \\
&\leq 2 \sum_{i=1}^{j} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{s_i} + 4 \left(\frac{1}{2}\right)^{j+1} \\
&= \sqrt{2} \sum_{i=1}^{j} \left(\frac{1}{2}\right)^{r/(2i)} + \left(\frac{1}{2}\right)^{j-1} \\
&\leq \sqrt{2} j \left(\frac{1}{2}\right)^{r/(2j)} + \left(\frac{1}{2}\right)^{j-1}.
\end{aligned}$$

Note that $(2rk\Delta)/n \le 1/2$ is true if we set $r = 2(\log_2(2n^2\Delta))^2$ and suppose

$$\frac{\Delta(\log_2 n)^4}{n} \to 0,$$

for large n. Also, $\Delta k^2/n \le 1/2$ by our choice of k. Hence, $(*)$ follows. We want the above expression less than $1/|\mathcal{B}|$. That is,

$$\left[\sqrt{2}j\left(\frac{1}{2}\right)^{\frac{r}{2j}} + \left(\frac{1}{2}\right)^{j-1}\right]\frac{n(n-1)}{k(k-1)} < 1.$$

But, the left hand side (LHS) of this is

$$\begin{aligned}
LHS &= O\left(\frac{n^2}{k^2}j\left(\frac{1}{2}\right)^{\frac{r}{2j}}\right)\\
&= O\left(\Delta n \log_2(2n^2\Delta)\left(\frac{1}{2}\right)^{\frac{r}{2\log_2(2n^2\Delta)+1}}\right)\\
&= O\left(\Delta n \log_2(2n^2\Delta)\left(\frac{1}{2}\right)^{\log_2(2n^2\Delta)}\right)\\
&= O\left(\frac{\log_2(2n^2\Delta)}{n}\right)\\
&= O(\log_2 n/n).
\end{aligned}$$

So with positive probability $|E(G_B)| < r$ for all $B \in \mathcal{B}$, and

$$\mathrm{cp}(\overline{G}) < rk|\mathcal{B}| = O(n^{3/2}\sqrt{\Delta}\log^2 n),$$

if n is sufficiently large. This completes the proof. □

Provided that $\Delta = o(n/\log^4 n)$, the expression in the theorem is $o(n^2)$. We conjecture the following stronger statement:

Conjecture 2.9.2 *Let G_n be a graph on n vertices with $o(n^2)$ edges. Then the clique partition number of $\bar{G}_n$ is $o(n^2)$.*

Part III

Clique Coverings

3.1 Intersection graphs and Kneser graphs

Let G be a graph. A *clique covering* of G is a set of cliques of G, which together contain each edge of G at least once. The smallest cardinality of any clique covering of G is called the *clique covering number* of G, and is denoted by $\mathrm{cc}(G)$. This number exists as the edge set of G forms a clique covering for G. A clique covering of G with size $\mathrm{cc}(G)$ is referred to as a *minimum clique covering* of G. The results on clique partitions in Sections $2.1 - 2.2$ also hold for clique coverings. Namely, Lemma 2.1.1, Lemma 2.1.2 and Theorem 2.2.1. Brigham and Dutton [7] list other results that the effect vertex and edge deletion have on the clique covering number of a graph.

Before Erdős, Goodman and Pósa [14] proved Theorem 2.2.1 for clique partitions, Hall [18] proved that the edge set of any graph G on n vertices can be covered using at most $\lfloor n^2/4 \rfloor$ cliques, none of which need to be larger than a triangle. We note that Hall's result follows trivially from Theorem 2.2.1. This is because a clique partition of a graph G is also a clique covering of G, so we have the inequality $\mathrm{cc}(G) \leq \mathrm{cp}(G)$. Thus, by Theorem 2.2.1, we have $\mathrm{cc}(G) \leq \mathrm{cp}(G) \leq \lfloor n^2/4 \rfloor$, with equality everywhere if and only if $G \cong K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

Let $\mathcal{S} = \{S_1, S_2, \ldots, S_m\}$ be a family of subsets of $[n]$. The *intersection graph* $\Omega(\mathcal{S})$ is the graph whose vertex set is S, with two vertices being adjacent if their sets intersect. That is, S_i is adjacent to S_j if and only if $S_i \cap S_j \neq \emptyset$, for $i \neq j$, $S_i, S_j \in S$. The Kneser graph $K_{n:k}$ is the complement of the intersection graph of all distinct k-subsets of an n-set. We have the following relationship between intersection graphs and clique coverings which is essentially due to Spilrajn-Marczewski [24] and Erdős, Goodman and Pósa [14].

Theorem 3.1.1 *Given a graph G,*

$$G \cong \Omega(\mathcal{S}) \iff cc(G) \leq n$$

where $\mathcal{S} = \{S_1, S_2, \ldots, S_m\}$ is a family of subsets of $[n]$.

Proof. We first prove that $G \cong \Omega(\mathcal{S})$ implies $\mathrm{cc}(G) \leq n$. Let $G \cong \Omega(\mathcal{S})$ with $\mathcal{S} = \{S_1, S_2, \ldots, S_m\}$, $S_i \subseteq [n]$. Let $C_j = \{S_i : j \in S_i, i \in [m]\}$, for each $j \in [n]$. Then C_j induces a clique whenever $|C_j| \geq 2$. Then $\{C_j : |C_j| \geq 2\}$ is a clique covering of G of size at most n, implying $\mathrm{cc}(G) \leq n$.

We now show that $\mathrm{cc}(G) \leq n$ implies that $G \cong \Omega(\mathcal{S})$, for some family $\mathcal{S}$ of subsets of $[n]$. Let $\{C_1, C_2, \ldots, C_n\}$ be a clique covering of G of size n. Let $S_v = \{i \in [n] : v \in V(C_i)\}$. Let $\mathcal{S} = \{S_v : v \in V(G)\}$. Then G is the intersection graph of $\mathcal{S}$. □

Notice that $\mathrm{cc}(G)$ does not change if isolated vertices are removed from G. We give another operation on G which does not affect $\mathrm{cc}(G)$. We call vertices $x, y \in V(G)$ *equivalent* if $xy \in E(G)$ and $\Gamma(x)\backslash\{y\} = \Gamma(y)\backslash\{x\} \neq \emptyset$. Then if x is equivalent to another vertex of G, $\mathrm{cc}(G) = \mathrm{cc}(G - x)$. Gyárfás [17] gives a lower bound for a graph with no isolated vertices and no equivalent vertices.

Theorem 3.1.2 *Let G be a graph on n vertices that contains neither isolated vertices nor equivalent vertices. Then*

$$\mathrm{cc}(G) \geq \log_2(n+1).$$

Proof. Let $\mathcal{C}$ be a clique covering of G of minimum cardinality k. Index the vertices of G by $v_1, v_2, \ldots, v_n$. Let F_i consist of those members of $\mathcal{C}$ having v_i as a vertex and define $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$. Then $\Omega(\mathcal{F})$ is isomorphic to G. Since G contains no isolated vertices, $F_i \neq \emptyset$, for $i \in [n]$. We now show that $F_i \neq F_j$, for $i \neq j$. Suppose that $i \neq j$. If $v_i v_j \notin E(G)$, then $F_i \neq F_j$ as they are disjoint sets. If $v_i v_j \in E(G)$, then by assumption they are not equivalent. So there is a vertex v_r, for some $r \in [n]\backslash\{i, j\}$, such that v_r is adjacent to exactly one of v_i and v_j. Without loss of generality, suppose $v_i v_r \in E(G)$ and $v_j v_r \notin E(G)$. Let $K \in \mathcal{C}$ be the clique that covers edge $v_i v_r$. Then $K \in F_i$ and $K \notin F_j$ showing that $F_i \neq F_j$. As the sets F_i are distinct non-empty subsets of $[k]$, we have that

$$n = |\mathcal{F}| \leq 2^k - 1.$$

Hence, $k \geq \log_2(n+1)$, as required. □

Gyárfás [17] showed the inequality in Theorem 3.1.2 is tight for infinitely many n. However, if we know more about the intersection properties of the sets F_i in the proof of Theorem 3.1.2, then we may obtain a better upper bound on $|\mathcal{F}|$ than $2^k - 1$, thus improving the lower bound on the clique covering number.

3.2 Clique covers of complete multipartite graphs

In Section 2.5 we looked at the clique partition number of the complement of a clique. The corresponding problem for clique coverings is an easy one. Namely, $\mathrm{cc}(K_n \backslash K_m) = m$, for $m > 2$, because we require a unique clique in our clique covering for each vertex of K_m.

A more interesting problem is determining the clique covering number of the complement of the union of complete graphs. Gregory and Pullman [16] compute the clique covering number of the complement of a perfect matching. We generalize this result of Gregory and Pullman [16] to give bounds on the clique covering number of the complement of the union of complete graphs. Note that the complement of the union of complete graphs is the s-partite complete graph $K_{t_1,t_2,\ldots,t_s}$, whose parts are of size $t_1, t_2, \ldots, t_s$ respectively. If each part has the same size, $t_1 = t_2 = \cdots = t_s = t > 1$, then we denote the graph by $K_s(t)$. In the next section we will prove the following theorem.

Theorem 3.2.1 *If $0 < \epsilon < 1$ and $t > 1$ are fixed, then for s sufficiently large,*

$$\mathrm{cc}(K_s(t)) \geq \log_b(st) + \frac{1-\epsilon}{2} \log_b \log_b(st)$$

where

$$b = \frac{t}{(t-1)^{(t-1)/t}}.$$

3.2.1 The Erdős-Ko-Rado Theorem and clique coverings

We will use a variant of the Erdős-Ko-Rado Theorem [6] to give a lower bound on the clique covering number of $K_s(t)$. We use the usual definitions from set theory found in Bollobás [6]. An *antichain* is a family $\mathcal{F}$ of sets such that $F \not\subset G$, for all $F \neq G$ in $\mathcal{F}$. In 1928, Sperner [6] proved the following theorem on the maximum size of an antichain:

Theorem 3.2.2 *Let $\mathcal{F}$ be an antichain on $[n]$, then*

$$|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

In 1961, Erdős, Ko, and Rado [6] gave an upper bound for the maximum size of an intersecting family of k-element sets in $[n]$, for $n \geq 2k$. This is known as the Erdős-Ko-Rado Theorem.

Theorem 3.2.3 *If $\mathcal{F}$ is an intersecting family of k-element sets in $[n]$, where $n \geq 2k$, then $|\mathcal{F}| \leq \binom{n-1}{k-1}$. Also, if $n > 2k$, then we get equality if and only if the sets in $\mathcal{F}$ all contain a common element.*

We say $\mathcal{F}$ is an *intersecting antichain* if it is an antichain of sets F_i such that $F_i \cap F_j \neq \emptyset$, for all F_i, F_j in $\mathcal{F}$. If $\mathcal{F}$ is an intersecting antichain whose sets are of size at most k, then a variant of the Erdős-Ko-Rado Theorem holds. The proof uses a technique known as the Katona Circle Method [6].

Theorem 3.2.4 *Let $n \geq 2k$. If $\mathcal{F}$ is an antichain and an intersecting family of sets of size at most k in $[n]$, then $|\mathcal{F}| \leq \binom{n-1}{k-1}$.*

Proof. Let $\mathcal{F}$ be an intersecting family of subsets of $[n]$, such that $\mathcal{F}$ is an antichain. Note that there are exactly $\frac{1}{2}(n-1)!$ circular permutations on n elements. We will count the ordered pairs (π, F), where π is a circular permutation of $[n]$, and F is a set in $\mathcal{F}$ whose image under π is an interval. Fix any particular circular permutation π. At most k members of $\mathcal{F}$ have their image under π as an interval, as each member of $\mathcal{F}$ must intersect the others (and have size at most k). Thus, the number of ordered pairs (π, F) is at most $k \cdot \frac{1}{2}(n-1)!$.

Now, fix any particular member F of $\mathcal{F}$. Then F is the image of an interval for $\frac{1}{2} \cdot |F|! \cdot (n - |F|)!$ circular permutations, as there are $|F|!$ ways to permute the elements of F, and $(n-|F|)!$ ways to permute the remaining elements for the circular permutation. Hence, the number of ordered pairs (π, F) is

$$\sum_{F \in \mathcal{F}} \frac{1}{2} \cdot |F|! \cdot (n - |F|)! \leq k \cdot \frac{1}{2}(n-1)!$$

Thus, it suffices to show

$$|\mathcal{F}| \cdot k! \cdot (n-k)! \leq \sum_{F \in \mathcal{F}} |F|! \cdot (n - |F|)!$$

But,

$$\sum_{F\in\mathcal{F}} \frac{|F|!\cdot(n-|F|)!}{k!\cdot(n-k)!} = \sum_{F\in\mathcal{F}} \left[\binom{n}{k}\Big/\binom{n}{|F|}\right] \geq \sum_{F\in\mathcal{F}} 1 = |\mathcal{F}|$$

since $\binom{n}{k} \geq \binom{n}{|F|}$, for each $F \in \mathcal{F}$, as $|F| \leq k \leq n/2$. □

3.2.2 t-Balanced families

Definition 3.2.5 A family $\mathcal{F}$ is *t-balanced* if it is the intersection graph of the complete multipartite graph on $|\mathcal{F}|$ vertices, whose components each have size t.

It is implied that $|\mathcal{F}|$ is divisible by t and $\mathcal{F}$ is a set of subsets in $[n]$, for some $n \in \mathbb{N}$. It is easy to see that a t-balanced family is an antichain. Suppose on the contrary that $\mathcal{F}$ is a t-balanced family which is not an antichain. Then there are sets $A, B \in \mathcal{F}$ with $A \subset B$. Consider $D \in \mathcal{F}$ such that $B \cap D = \emptyset$. But then $A \cap D \neq \emptyset$ by the definition of t-balanced, a contradiction. The Erdős-Ko-Rado Theorem can be used to provide an upper bound on the size of a t-balanced family.

Theorem 3.2.6 *Let $\mathcal{F}$ be a t-balanced collection of sets in $[n]$. Then*

$$|\mathcal{F}| \leq t\binom{n-1}{\lceil n/t\rceil}.$$

Further, there exists a t-balanced family $\mathcal{F}$ that gives equality when $t = 2$.

Proof. Take $\mathcal{F}$ to be a t-balanced collection of sets in $[n]$. Then for some s, $\Omega(\mathcal{F}) \cong K_s(t)$, and the vertices of $K_s(t)$ correspond to sets in $[n]$. For each part of $K_s(t)$, choose a single vertex whose set in $[n]$ is smallest among all other vertices of that part, and let B be the set of these vertices. Then B consists of s sets of size at most $\lfloor n/t\rfloor$. Since $\mathcal{F}$ is t-balanced, $\mathcal{F}$ is an antichain implying that B is an antichain. Also, B is intersecting as it contains one vertex from each part. Thus, by Theorem 3.2.4,

$$|\mathcal{F}| = t|B| \leq t\binom{n-1}{\lfloor n/t\rfloor - 1} = t\binom{n-1}{\lceil n/t\rceil}.$$

Note that we may obtain equality when $t = 2$. Let $r = \lfloor n/2\rfloor - 1 < n/2$, and B be the set of r-subsets of $[n]$ containing a point x. Then $\mathcal{F} = B \cup \bar{B}$ gives equality. □

This theorem provides us with a lower bound for $\mathrm{cc}(K_s(t))$. Throughout the rest of this essay let

$$\sigma_t(s) = \min\left\{ n : s \leq \binom{n-1}{\lceil n/t \rceil} \right\}.$$

Corollary 3.2.7 *For all $s, t > 1$, $\mathrm{cc}(K_s(t)) \geq \sigma_t(s)$, where $\sigma_t(s)$ is defined above.*

Proof. Let $\mathcal{C}$ be a clique covering of $K_s(t)$ of minimum cardinality k. Index the vertices of $K_s(t)$ by $v_1, v_2, \ldots, v_{ts}$. Let F_i consist of those members of $\mathcal{C}$ having v_i as a vertex and define $\mathcal{F} = \{F_1, F_2, \ldots, F_{ts}\}$. Since $\Omega(\mathcal{F})$ is isomorphic to $K_s(t)$, the family $\mathcal{F}$ is t-balanced. Thus by Theorem 3.2.6,

$$|\mathcal{F}| = ts \leq t\binom{k-1}{\lceil k/t \rceil}$$

and hence, $\mathrm{cc}(K_s(t)) = k \geq \sigma_t(s)$ by definition of $\sigma_t(s)$. □

For $t = 2$, we may obtain equality. In general, suppose equality holds in Theorem 3.2.6. Take $k = \sigma_t(s)$ and choose a t-balanced family $\mathcal{F}$ in $S = \{1, 2, \ldots, k\}$ with cardinality

$$|\mathcal{F}| = ts \leq t\binom{k-1}{\lceil k/t \rceil}.$$

Then the intersection graph of $\mathcal{F}$ is isomorphic to $K_s(t)$ and $\mathrm{cc}(\Omega(\mathcal{F})) \leq k$, by Theorem 3.1.1. Thus, for values of t where equality holds in Theorem 3.2.6, $\mathrm{cc}(K_s(t)) = \sigma_t(s)$. We conjecture that $\mathrm{cc}(K_s(t)) = \sigma_t(s)$, for $t > 1$. Theorem 3.2.1 now follows as it is straight forward to compute a lower bound on $\sigma_t(s)$. See Lemma 5.1.1 in the Appendix for details. Theorem 3.1.2 gives

$$\mathrm{cc}(K_s(t)) \geq \log_2(st+1).$$

Theorem 3.2.1 is an improvement on this result since $b \to 1$ as $t \to \infty$. This gives a lower bound for complete multipartite graphs whose parts have different size. If $K_{t_1,t_2,\ldots,t_s}$ has r parts of size bigger than one and if

$$t = \min\{t_i : t_i > 1,\ i = 1, 2, \ldots, s\}$$

then

$$\mathrm{cc}(K_{t_1,t_2,\ldots,t_s}) \geq \sigma_t(r) \geq \log_b r$$

where

$$b = \frac{t}{(t-1)^{(t-1)/t}}.$$

This follows from Lemma 2.1.2 for clique covers, by deleting vertices in parts of size one, and deleting vertices in parts of size bigger than t, until every part has size t.

3.3 Complement of paths, cycles, and perfect matchings

In Section 2.7 we analyzed the clique partition number of the complement of paths, cycles and perfect matchings. It would be nice to know how the clique covering number behaves for these graphs. It was Orlin [22] who first asked about the asymptotics of the clique covering number of the complement of a perfect matching. Gregory and Pullman [16] answer this question and show that asymptotically, $\mathrm{cc}(\bar{M}_m) \sim \log_2 m$. Regarding the complement of a cycle and path, de Caen, Gregory, and Pullman [11] have computed exact values for $\mathrm{cc}(\bar{C}_m)$ and $\mathrm{cc}(\bar{P}_m)$, for small values of m.

m	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\mathrm{cc}(\bar{C}_m)$	2	5	5	7	6	7	6	8	7	7	7	8	7	8	7	8	8
$\mathrm{cc}(\bar{P}_m)$	3	4	5	5	6	6	6	6	7	7	7	7	7	7	7	7	8

They show that for m large enough,

$$\log_2 m \leq \mathrm{cc}(\bar{C}_m) - 2 \leq \mathrm{cc}(\bar{P}_m) \leq 2 \log_2 m$$

and conjecture that asymptotically, $\mathrm{cc}(\bar{C}_m) \sim \log_2 m$ and $\mathrm{cc}(\bar{P}_m) \sim \log_2 m$. Alles and Poljak [1] improve the upper bound to $\mathrm{cc}(\bar{P}_m) \leq 1.695 \log_2 m$, for m sufficiently large. Kohayakawa [19] again improves the upper bound to $\mathrm{cc}(\bar{P}_m) \leq 1.459 \log_2 m$, for m sufficiently large, by finding a long induced path in the Kneser graph.

First, note that if we have a clique covering for $\bar{P}_{m-1}$ then this gives rise to a clique covering of $\bar{C}_m$ using at most $\mathrm{cc}(\bar{P}_{m-1}) + 2$ cliques. Using Lemma 2.1.1 for clique coverings, we get that $\mathrm{cc}(\bar{C}_m) \leq \mathrm{cc}(\bar{P}_{m-1}) + 2 \leq \mathrm{cc}(\bar{P}_m) + 2$.

Remark 3.3.1 $\mathrm{cc}(\bar{C}_m) \leq \mathrm{cc}(\bar{P}_m) + 2$.

We use the results from the previous section to provide a lower bound on the clique covering number of the complement of a cycle, and hence, the complement of a path. We note that if we have a clique covering for $\bar{C}_m$, then by adding at most three cliques, we get a clique covering for $\bar{M}_m$.

Remark 3.3.2 $\mathrm{cc}(\bar{M}_m) \leq \mathrm{cc}(\bar{C}_m) + 3$.

Hence, for m sufficiently large and fixed $\epsilon > 0$, Corollary 3.2.1 gives

$$\log_2 m + \frac{1-\epsilon}{2} \log_2 \log_2 m - 3 \leq \mathrm{cc}(\bar{C}_m) \leq \mathrm{cc}(\bar{P}_m) + 2.$$

Thus,

$$\log_2 m \leq \mathrm{cc}(\bar{C}_m) - 2 \leq \mathrm{cc}(\bar{P}_m),$$

for m sufficiently large.

3.3.1 Induced cycles and paths in Kneser graphs

In this section, we will look at how to get an upper bound for the clique covering number of the complement of a path, and hence, the complement of a cycle. We will follow Vander Meulen [25] for most of the results in this section. Theorem 3.1.1 says that finding an induced graph H in the Kneser graph $K_{n:k}$ gives rise to a clique covering of $\bar{H}$. Thus, we look at the problem of finding the order of the longest cycle (or path) in the Kneser graph $K_{n:k}$. We define $p(n,k)$ and $c(n,k)$ to be the maximum order (number of vertices) of an induced path and cycle respectively in $K_{n:k}$. Upper bounds on $p(n,k)$ and $c(n,k)$ have been given by Alles and Poljak [1] and Kohayakawa [19]. A simple argument relates these two numbers. If C_m is an induced cycle in a graph, then removing a vertex gives an induced path of order $m-1$.

Proposition 3.3.3 *Fix $k \geq 1$, then $p(n,k) \geq c(n,k) - 1$, for $n \geq 2k+1$.*

Further, if H is an induced subgraph of the Kneser graph $K_{n:k}$, the intersection graph of $V(H)$ is the complement of H. Hence, $\mathrm{cc}(\bar{H}) \leq n$. Also, if there is an induced path of order m in $K_{n:k}$, then $\mathrm{cc}(\bar{P}_m) \leq n$. In particular, if H is an induced path or cycle, $\mathrm{cc}(\bar{P}_m) \leq \mathrm{cc}(\bar{P}_{m+1})$ implies the following:

Lemma 3.3.4 *Given $k \geq 1$, $n > 2k$, we have $\mathrm{cc}(\bar{C}_{c(n,k)}) \leq n$ and $\mathrm{cc}(\bar{P}_s) \leq n$, for all positive integers $s \leq p(n,k)$.*

To provide an upper bound on the clique covering number of the complement of a path, we must obtain a lower bound on $p(2k+1,k)$. Kohoyakawa [19] bounds $p(2k+1,k)$ from below by using a bipartite graph and giving a recursive construction for an induced path in $K_{2k+1:k}$. Let $s \in \mathbb{N}$ and define G_s to be a bipartite graph with vertex classes $[2s]^{(s)}$ and $[2s]^{(s-1)}$. Two vertices in different classes are adjacent if and only if they are disjoint. (Note that we are using G_s to represent this bipartite graph, and not a graph on s vertices).

We define $w(s)$ to be the maximum number of vertices in $[2s]^{(s)}$ in an induced path in G_s. Note that $w(s)$ is roughly half the order of a maximum induced path in G_s. The proof of the following theorem gives a construction of long induced paths in the Kneser graph $K_{2k+1:k}$. For the next proof, we use $K_{a_1,a_2,\ldots,a_l:k}$ to denote the graph whose vertex set consists of the k-subsets of $\{a_1,a_2,\ldots,a_l\}$, and where two vertices are adjacent if and only if they are disjoint. Note that $K_{a_1,a_2,\ldots,a_l:k}$ is isomorphic to $K_{l:k}$.

Theorem 3.3.5 *For $k \geq 2, s \geq 1$,*

$$p(2(k+s)+1,k+s) \geq \begin{cases} w(s)\cdot(p(2k+1,k)+1)-1, & \text{if } p(2k+1,k) \text{ is odd} \\ w(s)\cdot p(2k+1,k)-1, & \text{if } p(2k+1,k) \text{ is even} \end{cases}$$

Proof. Let $A_1B_1A_2B_2\ldots B_{m-1}A_m$ be an induced path in G_s, where $A_i \in [2s]^{(s)}$, for $1 \leq i \leq m$, and $B_j \in [2s]^{(s-1)}$, for $1 \leq j \leq m-1$. Let $V_1V_2,\ldots V_n$ be an induced path of order n (with n odd) in

$$K_{2s+1,2s+2,\ldots,2(k+s)+1:k}$$

which is isomorphic to $K_{2k+1:k}$. Consider the paths,

$$P_i = A_i \cup V_1, \bar{A}_i \cup V_2, A_i \cup V_3, \ldots, \bar{A}_i \cup V_{n-1}, A_i \cup V_n$$

where $\bar{A}_i$ is the complement of A_i in $[2s]^{(s)}$, for $1 \leq i \leq m$. The paths end in vertex $A_i \cup V_n$, as n was chosen to be odd.

Then the set of paths $P = \cup_i P_i$ is a collection of m vertex disjoint induced paths in $K_{2(k+s)+1:k+s}$. Note that we could have $A_i = \bar{A}_j$, for some $i \neq j$, as $A_i \in [2s]^{(s)}$. We join these paths together to create a path of order $m(n+1)-1$ in $K_{2(k+s)+1:k+s}$. Take $b \in V_3 \backslash V_1$ and $a \in V_{n-2} \backslash V_n$. Then, for $i \in \{1,2,\ldots,m-1\}$, define

$$C_i = \begin{cases} V_{n-1} \cup B_i \cup \{a\}, & \text{if } i \text{ is odd} \\ V_2 \cup B_i \cup \{b\}, & \text{if } i \text{ is even} \end{cases}$$

Then, as vertices, the C_i are non-adjacent as $V_{n-1} \cap V_2 \subseteq C_i \cap C_j$ and $V_{n-1} \cap V_2 \neq \emptyset$ (for $n \geq 5$ and $1 \leq i, j \leq m-1$). Also, each C_i is connected to exactly two vertices in P, namely $V_n \cup A_i$ and $V_n \cup A_{i+1}$ for odd i, or $V_1 \cup A_i$ and $V_1 \cup A_{i+1}$ for even i.

Thus, we get a path of order $m(n+1)-1$ in $K_{2(k+s)+1:k+s}$. Since the longest induced path in G_s might not have initial and end vertices in $[2s]^{(s)}$, then we simply delete these vertices to get a path as above. This can be done as $w(s)$ counts the vertices in $[2s]^{(s)}$. Taking $m = w(s)$ and

$$n = \begin{cases} p(2k+1,k)-1, & \text{if } p(2k+1,k) \text{ is even} \\ p(2k+1,k), & \text{if } p(2k+1,k) \text{ is odd} \end{cases}$$

gives the result. □

Theorem 3.3.5 gives a lower bound on $p(2k+1,k)$.

Corollary 3.3.6 *Fix $s \geq 1$. Then $p(2k+1,k) \geq w(s)^{\lfloor (k-1)/s \rfloor}$, for all $k \geq 1$.*

Proof. Note that the theorem is true when $s = 1$, as $w(s) = 1$. Fix $s \geq 2$ and recursively use $p(2k+1,k+1) \geq w(s) \cdot p(2(k-s)+1, k-s) - 1$ to get

$$p(2k+1,k+1) \geq w(s)^t \cdot p(2(k-ts)+1, k-ts) - \sum_{i=0}^{t-1} w(s)^i$$

for any integer $1 \leq t \leq \frac{k-1}{s}$. As $w(s) \geq 2$ for $s \geq 1$, we get

$$p(2k+1,k+1) \geq w(s)^t \cdot [p(2(k-ts)+1, k-ts) - 1].$$

As $p(3,1) = 2$, taking $t = \lfloor (k-1)/s \rfloor$ gives

$$p(2k+1,k+1) \geq w(s)^{\lfloor (k-1)/s \rfloor}[p(3,1)-1] = w(s)^{\lfloor (k-1)/s \rfloor}$$

as required to prove. □

By using a computer search, Kohayakawa [19] found a long induced path in G_6.

Remark 3.3.7 *There exists an induced path in G_6 with* 300 *vertices in* $[12]^{(6)}$ *(that is, $w(6) \geq 300$).*

Thus by Corollary 3.3.6, $p(2k+1,k) \geq 300^{\lfloor (k-1)/6 \rfloor} \geq (2.587)^{k-1}$, for all $k \geq 1$. By Lemma 3.3.4, $\mathrm{cc}(\bar{P}_m) \leq 2k+1$, for $m \leq p(2k+1,k)$. Take $m = \lfloor (2.587)^{k-1} \rfloor + 1$. Then we have,

$$k \leq \frac{\log_2 m}{\log_2 2.587} + 1.$$

This implies for m large enough,

$$\mathrm{cc}(\bar{P}_m) \leq 2k+1 \leq 1.459 \log_2 m.$$

Kohayakawa [19] conjectured that $\sup w(s)^{1/s} = 4$. If this is correct, then we will be able to improve the 1.459 in the upper bound to $1 + o(1)$. Note that de Caen and Gregory [11] also conjectured that $\mathrm{cc}(\bar{P}_m) \sim \log_2 m$. Thus we have established the following theorem.

Theorem 3.3.8 *For m large enough,*

$$\log_2 m \leq \mathrm{cc}(\bar{C}_m) - 2 \leq \mathrm{cc}(\bar{P}_m) \leq 1.459 \log_2 m.$$

3.4 Complement of graphs with maximum degree two

In this section we will obtain bounds on the clique covering number of the complement of graphs whose maximum degree is two. Let G_n be a graph on n vertices with maximum degree two. Then G_n is a graph whose components are paths and cycles. We will assume that the paths and cycles are nontrivial. de Caen, Gregory and Pullman [11] prove that for n sufficiently large

$$\log_2 n \leq \mathrm{cc}(\bar{G}_n) \leq 5.8 \log_2 n.$$

To provide a lower bound on $\mathrm{cc}(\bar{G}_n)$, where G_n is a graph whose components are paths and cycles, we use the complement of a perfect matching. Notice that by removing at most $\lceil 3n/5 \rceil$ vertices from G_n, we obtain a perfect matching on at least $\lfloor 2n/5 \rfloor$ vertices. Thus by Lemma 2.1.2 and Lemma 5.1.1, if n is large enough,

$$\mathrm{cc}(\bar{G}_n) \geq \mathrm{cc}(\bar{M}_{\lfloor 2n/5 \rfloor}) \geq \log_2 n.$$

For an upper bound on $\mathrm{cc}(\bar{G}_n)$, we first need the following lemma which may be found in de Caen, Gregory and Pullman [11].

Lemma 3.4.1 *Let G be a graph with components G^i, for $i = 1, 2, \ldots, s$. Then*

$$\mathrm{cc}(\bar{G}) \leq \max\{\mathrm{cc}(\bar{G}^i) : i = 1, 2, \ldots, s\} + \mathrm{cc}(K_{t_1,t_2,\ldots,t_s})$$

where t_i is the chromatic number of G^i, for $i = 1, 2, \ldots, s$.

Proof. We will first cover the edges of the form vw such that $v \in G^i \implies w \notin G^i$. Since t_i is the chromatic number of G^i, we can select t_i cliques in $\bar{G}^i$ which partition the vertices of $\bar{G}^i$. Call the cliques $K^i(1), K^i(2), \ldots, K^i(t_i)$. Note that $K^i(j)$ could be a single vertex if $t_j = 1$. Recall the join and union of graphs [12]. Consider the graph

$$H = \bigvee_{i=1}^{s} (K^i(1) \cup K^i(2) \cup \cdots \cup K^i(t_i)).$$

A clique covering of H covers the edges of the form vw such that $v \in G^i \implies w \notin G^i$. Note that if $x, y \in V(K^i(j))$, then x, y are equivalent. Hence, $\mathrm{cc}(H) = \mathrm{cc}(K_{t_1,t_2,\ldots,t_s})$.

If K^i is a clique in $\bar{G}^i$, for each $i = 1, 2, \ldots, s$, their join $K_1 \vee K_2 \vee \cdots K_s$ is a clique in $\bar{G}$. By joining cliques from minimum clique coverings of the $\bar{G}^i$, for $i = 1, 2, \ldots, s$, the remaining edges within the graphs $\bar{G}^i$ can be covered by $\max\{\mathrm{cc}(G^i) : i = 1, 2, \ldots, s\}$ cliques in G. □

de Caen, Gregory and Pullman [11] use this result to obtain an upper bound on $\mathrm{cc}(\bar{G}_n)$.

Theorem 3.4.2 *If G_n has components consisting of nontrivial cycles and paths then*

$$\mathrm{cc}(\bar{G}_n) = O(\log_2 n)$$

for sufficiently large n.

Proof. Suppose G_n has components G^i for $i = 1, 2, \ldots, s$, where each G^i is a cycle or a path on $n_i > 1$ vertices. As the chromatic number of each G^i is at most three, Lemma 3.4.1 gives

$$\mathrm{cc}(\bar{G}_n) \leq \max\{\mathrm{cc}(\bar{G}^i) : i = 1, 2 \ldots, s\} + \mathrm{cc}(K_s(3))$$

Note that as Theorem 3.3.8 is for n_i sufficiently large, $\mathrm{cc}(\bar{G}^i) \leq 1.459 \log_2 n_i + O(1)$, for $n_i > 1$. For an upper bound on $\mathrm{cc}(K_s(3))$, we note that Lemma 3.4.1 gives

$$\mathrm{cc}(K_{3s}(3)) \leq \mathrm{cc}(K_s(3)) + \mathrm{cc}(K_3(3)) = \mathrm{cc}(K_s(3)) + 9.$$

Thus,

$$\mathrm{cc}(K_s(3)) \leq 9 \log_3 s + O(1).$$

Note that de Caen, Gregory and Pullman [11] provide an upper bound of $6 \log_3 s + O(1)$ for $\mathrm{cc}(K_s(3))$ by using Latin squares, however, $9 \log_3 s + O(1)$ is sufficient for our purposes. This implies

$$\mathrm{cc}(\bar{G}_n) \leq 1.459 \log_2 n + 9 \log_3 s = O(\log_2 n)$$

for n sufficiently large. □

de Caen, Gregory and Pullman [11] ask whether or not for n sufficiently large, we can get

$$\mathrm{cc}(\bar{G}_n) < (1 + o(1)) \log_2 n$$

in Theorem 3.4.2. We give an example of a graph, which is the complement of paths and cycles, where this fails to be true. Consider G_n to be the graph where each component is a K_3, and n is divisible by three. Corollary 3.2.1 gives,

$$\mathrm{cc}(\bar{G}_n) \geq 1.088 \log_2 n.$$

So $\bar{G}_n$ is a counterexample.

3.5 Complement of forests

In this section, we use the techniques of the previous section to obtain bounds on the clique covering number of the complement of a forest, $\bar{F}_n$, were F_n denotes a forest on n vertices. For the first time, we show

$$\log_2 k \leq \mathrm{cc}(\bar{F}_n) \leq 10.3 \log_2 n$$

where k is the length of the longest path in F_n.

To provide a lower bound on $\mathrm{cc}(\bar{F}_n)$ we use the complement of a path. Let P_k be an induced path in F_n with k vertices. Then by Lemma 2.1.2 and Theorem 3.3.8, if n is large enough,

$$\mathrm{cc}(\bar{F}_n) \geq \mathrm{cc}(\bar{P}_k) \geq \log_2 k.$$

Note that this bound is achieved by the star graph. Also note that this lower bound holds if k is replaced the size of the largest induced matching of F_n.

We now obtain an upper bound on $\mathrm{cc}(\bar{F}_n)$. For the remainder of this section, let g be the function defined by

$$g(n) = \max\{\mathrm{cc}(\bar{F}_n) : F_n \text{ is a forest on } n \text{ vertices}\}.$$

It is easy to see that $g(n) \leq g(n+1) \leq g(n) + 2$.

Lemma 3.5.1

$$g(n) \leq g(\lfloor 2n/3 \rfloor) + 6.$$

Proof. Let F_n be a forest such that $g(n) = \mathrm{cc}(\bar{F}_n)$. By Lemma 2.8.2, we may split F_n into forests F^1 and F^2 such that $|F^i| \leq \lfloor 2n/3 \rfloor$, for $i = 1, 2$, and $V(F^1) \cap V(F^2) = \{v\}$, for some vertex v of F_n. We will construct a clique covering of $\bar{F}_n$. By Lemma 3.4.1, $F_n - v$ can be covered using at most

$$\max\{\mathrm{cc}(\overline{F^1 - v}), \mathrm{cc}(\overline{F^2 - v})\} + \mathrm{cc}(K_2(2))$$

cliques. But $\mathrm{cc}(K_2(2)) = 4$. Also, the edges adjacent to v in $\bar{F}_n$ can be covered by at most two cliques. The result now follows, as

$$\max\{\mathrm{cc}(\overline{F^1 - v}), \mathrm{cc}(\overline{F^2 - v})\} \leq g(\lfloor 2n/3 \rfloor).$$

□

We will show if n is large enough, then $\mathrm{cc}(\bar{F}_n) = O(\log_2 n)$.

Theorem 3.5.2 *Let F_n be a forest on n vertices. Then*

$$\mathrm{cc}(\bar{F}_n) \leq 10.3 \log_2 n$$

for n sufficiently large.

Proof. Let F_n be a forest on n vertices. Then $\mathrm{cc}(\bar{F}_n) \leq g(n)$, so it suffices to prove that $g(n) \leq 10.3 \log_2 n$. Let $r = \log_{3/2} n$, so that $n \leq \lfloor (3/2)^r \rfloor + 1$. Then by repeatedly using Lemma 3.5.1 and the fact that $g(n) \leq g(n+1) \leq g(n) + 2$ gives,

$$g(n) \leq g(\lfloor (3/2)^r \rfloor + 1) \leq 6r + O(1) \leq 10.3 \log_2 n$$

for n sufficiently large. □

It would be interesting to know whether or not

$$\mathrm{cc}(\bar{F}_n) \leq (1 + o(1)) \log_2 n$$

for n sufficiently large. If in fact $\mathrm{cc}(\bar{P}_n) \sim \log_2 n$, then this would seem like a reasonable guess. Perhaps the upper bound is $O(\log_2 k)$, where k is the size of a largest matching in F_n.

Part IV

Conclusion

4.1 Open Problems

In Chapter II, we looked that the clique partition number of graphs. In Section 2.7, we gave a result due to Gregory, McGuinness and Wallis [15], that is, for n sufficiently large,

$$n \leq \mathrm{cp}(\bar{M}_n) \leq (1 + o(1))n \log_2 \log_2 n.$$

They conjecture that $\mathrm{cp}(\bar{M}_n) \sim n$. Similar results hold for $\bar{P}_n$ and $\bar{C}_n$, although it does not appear that projective planes can be used to get a linear bound for the clique partition number of $\bar{M}_n$, $\bar{P}_n$, and $\bar{C}_n$.

Conjecture 4.1.1 *If G_n is M_n, P_n or C_n, then $\mathrm{cp}(\bar{G}_n) \sim n$.*

In Section 2.8, the clique partition number of the complement of a forest was looked at. We believe that F_n can be partitioned into $(1 + o(1))\sqrt{n}$ components each with size $(1+o(1))\sqrt{n}$, such that any two components intersect in at most one vertex.

Conjecture 4.1.2 *If F_n is a forest on n vertices then,*

$$\mathrm{cp}(\bar{F}_n) = O(n \log_2 \log_2 n),$$

for n sufficiently large.

A probabilistic argument could give the existence of a forest whose clique partition number is not linear.

Conjecture 4.1.3 *There exists a forest such that*

$$\lim_{n\to\infty} \frac{\mathrm{cp}(\bar{F}_n)}{n} = \infty.$$

In Section 2.9 we used a probabilistic argument to give an upper bound on the clique partition number of the complement of a graph with bounded maximum degree. If the maximum degree is not bounded, a similar argument might exist.

Conjecture 4.1.4 *If* $|E(G_n)| = o(n^2)$ *then* $\mathrm{cp}(\bar{G}_n) = o(n^2)$.

In Chapter II, we analyzed the clique covering number of graphs. As $\mathrm{cc}(K_s(2)) = \sigma_2(s)$, we suspect that $\mathrm{cc}(K_s(t)) = \sigma_t(s)$, for all $t > 1$.

In Section 3.3, $\mathrm{cc}(\bar{P}_m)$ and $\mathrm{cc}(\bar{C}_m)$ were analyzed. de Caen, Gregory, and Pullman [11], as well as Kohayakawa [19], conjecture that $\mathrm{cc}(\bar{C}_m) \sim \log_2 m$ and $\mathrm{cc}(\bar{P}_m) \sim \log_2 m$. It would be interesting to know if this is the case and whether or not $\sup w(s)^{1/s}$ is indeed equal to four.

Conjecture 4.1.5 $\mathrm{cc}(\bar{C}_m) \sim \log_2 m$ *and* $\mathrm{cc}(\bar{P}_m) \sim \log_2 m$.

In Section 3.5 we looked at the clique covering number of the complement of a forest. We suspect that the upper bound is $\log_2 n$.

Conjecture 4.1.6

$$\mathrm{cc}(\bar{F}_n) \leq (1 + o(1)) \log_2 n$$

for n *sufficiently large.*

Part V

Appendix

5.1 Lower bound for $\sigma_t(s)$

Lemma 5.1.1 *Let $t > 1$ be an integer, $\sigma_t(s) = \min\{n : s \leq \binom{n-1}{\lceil n/t \rceil}\}$ and fix $0 < \epsilon < 1$. Then for s sufficiently large,*

$$\sigma_t(s) \geq \log_b(st) + \frac{1-\epsilon}{2} \log_b \log_b(st)$$

where

$$b = \frac{t}{(t-1)^{(t-1)/t}}.$$

Proof. We first note Stirling's formula,

$$\sqrt{2\pi n}(n/e)^n e^{1/(12n+1)} \leq n! \leq \sqrt{2\pi n}(n/e)^n e^{1/(12n)}.$$

The following inequality can be derived from Stirling's formula,

$$\binom{tn}{n} \leq \sqrt{\frac{t}{2\pi(t-1)}} \frac{1}{\sqrt{n}} \left(\frac{t^t}{(t-1)^{t-1}} \right)^n.$$

Then, as $\sigma_t(s) = \min\{n : s \leq \binom{n-1}{\lceil n/t \rceil}\}$, we get,

$$s \leq \binom{n-1}{\lceil n/t \rceil} = \frac{n - \lceil n/t \rceil}{n} \binom{n}{\lceil n/t \rceil} \leq \frac{t-1}{t} \binom{n}{\lceil n/t \rceil} \leq \frac{t-1}{t} \binom{t\lceil n/t \rceil}{\lceil n/t \rceil}.$$

Thus,

$$s \leq \frac{t-1}{t} \sqrt{\frac{t}{2\pi(t-1)}} \frac{1}{\sqrt{n/t}} \left(\frac{t^t}{(t-1)^{t-1}} \right)^{\lceil n/t \rceil}.$$

For convenience, we let

$$b = \frac{t}{(t-1)^{(t-1)/t}} \qquad \text{and} \qquad c = \frac{t^2 b^2 (t-1)}{2\pi}.$$

Then as $b > 1$ for $t > 1$,

$$st \leq \sqrt{\frac{t-1}{2\pi}} \frac{1}{\sqrt{n}} \cdot b^{n+1} t = b^n \sqrt{\frac{c}{n}}.$$

Then,

$$\log_b(st) \le n - \frac{1}{2}\log_b n + \frac{1}{2}\log_b c$$
$$n \ge \log_b(st) + \frac{1}{2}\log_b n - \frac{1}{2}\log_b c.$$

Take s sufficiently large so that

$$\log_b(st) \ge c\left(\log_b(st)\right)^{1-\epsilon} + \frac{1}{2}\log_b c.$$

This can be done as t, b, c, ϵ are all constant. This implies

$$\begin{aligned} n &\ge \log_b(st) + \frac{1}{2}\log_b\left(c\left(\log_b(st)\right)^{1-\epsilon}\right) - \frac{1}{2}\log_b c \\ &= \log_b(st) + \frac{1-\epsilon}{2}\log_b\log_b(st). \end{aligned}$$

As $\sigma_t(s)$ is the smallest such n, we are done. □

Note that $b = 2$ when $t = 2$. Also, as $t \to \infty$, $b \to 1$. A similar argument can be used to give a lower bound on $\sigma_t(s)$. Note that we actually have that

$$\binom{tn}{n} \sim \sqrt{\frac{t}{2\pi(t-1)}}\frac{1}{\sqrt{n}}\left(\frac{t^t}{(t-1)^{t-1}}\right)^n$$

For convenience, denote $\sigma = \sigma_t(s)$. Thus, for a lower bound on s we can use

$$s \ge \binom{\sigma-2}{\lceil\frac{\sigma-1}{t}\rceil} \ge d_1(t)\binom{t\lceil\frac{\sigma-1}{t}\rceil}{\lceil\frac{\sigma-1}{t}\rceil}$$

where $d_1(t)$ is some function of t. Then we have,

$$s \ge d_2(t)\frac{b^\sigma}{\sqrt{\sigma}}$$

for some function $d_2(t)$. There is also a function $d_3(t)$ so that we have,

$$d_2(t)\frac{b^\sigma}{\sqrt{\sigma}} \le s \le d_3(t)\frac{b^\sigma}{\sqrt{\sigma}}.$$

Taking logarithms base b, dividing by σ and taking the limit as $s \to \infty$ (and hence $\sigma \to \infty$), we have

$$1 \le \lim_{s\to\infty}\frac{\log_b s}{\sigma} \le 1$$

Thus, $\sigma_t(s) \sim \log_b s$, where

$$b = \frac{t}{(t-1)^{(t-1)/t}}.$$

5.2 A recurrence relation

Lemma 5.2.1 *Let x be a positive real number and $\epsilon > 0$. If*

$$\frac{c(x)}{x} \leq 1 + \epsilon + \frac{c(\sqrt{x})}{\sqrt{x}},$$

then

$$c(x) \leq (1+\epsilon)x \log_2 \log_2 x + Mx$$

for some constant M.

Proof. Let $z = \log_2 \log_2 x$ and define $h(z) = \frac{c(x)}{x}$. Then

$$\begin{aligned} \log_2 \log_2 \sqrt{x} &= \log_2 \frac{1}{2} \log_2 x \\ &= \log_2 \log_2 x + \log_2 \frac{1}{2} \\ &= z - 1. \end{aligned}$$

Thus, $h(z-1) = \frac{c(\sqrt{x})}{\sqrt{x}}$. By our assumption in the lemma,

$$h(z) \leq 1 + \epsilon + h(z-1).$$

By repeatedly using this inequality, we get

$$h(z) \leq (1+\epsilon)\lfloor z \rfloor + h(z - \lfloor z \rfloor).$$

Let

$$M = \sup_{z \in [0,1]} h(z).$$

Then

$$h(z) \leq (1+\epsilon)z + M$$

implying

$$c(x) \leq (1+\epsilon)x \log_2 \log_2 x + Mx.$$

□

List of notation

References

[1] P. Alles and S. Poljak Long induced paths and cycles in Kneser graphs *Graphs and Combinatorics* **5** (1989) 303-306.

[2] N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley, New York, (1992). (Second Edition, 2000).

[3] R.C. Baker, G. Harman, and J. Pintz, The difference between consecutive primes, *Proc. London Math. Soc.* **83**(3) (2001) 532-562.

[4] L.M. Batten and A. Beutelspacher, *The theory of finite linear spaces*, Cambridge Univ. Press, New York, (1993).

[5] T. Beth, D. Jungnickel, and H. Lenz, *Design Theory, Volume* 2, *Second Edition*, Cambridge Univ. Press, New York, (1999).

[6] B. Bollobás *Combinatorics: set systems, hypergraphs, families of vectors, and combinatorial probability*, Univ. of Cambridge, New York, (1994).

[7] R.C. Brigham and R.D. Dutton, Changing and unchanging invariants: The edge clique covering number, *Congr. Numer.* **70** (1990) 145-152.

[8] N.G. de Bruijn and P. Erdős, On a combinatorial problem, *Indag. Math.* **10** (1948) 421-423.

[9] P.J. Cameron and J.H. Van Lint, *Designs, graphs, codes and their links*, London Mathematical Society Student Texts 22, Cambridge Univ. Press, New York, (1991).

[10] D. de Caen and D.A. Gregory, Partitions of the edge-set of a multigraph by complete subgraphs, *Congressus Numerantium* **47** (1985) 255-263.

[11] D. de Caen, D.A. Gregory, and N.J. Pullman, Clique coverings of complements of paths and cycles, *Annals of Discrete Math.* **27** (1985) 25-268.

[12] J.A. Bondy and U.S.R. Murty, *Graph theory with applications*, MacMillan, London, (1972).

[13] P. Erdős, R. Faudree, and E.T. Ordman, Clique partitions and clique coverings, *Discrete Math.* **72** (1988) 93-101.

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

[15] D.A. Gregory, S. McGuinness, and W. Wallis, Clique partitions of the cocktail party graph, *Discrete Math.* **59** (1986) 267-273.

[16] D.A. Gregory and N.J. Pullman, On a clique covering problem of Orlin, *Discrete Math.* **41** (1982) 97-99.

[17] A. Gyárfás, A simple lower bound on edge coverings by cliques, *Discrete Math.* **85** (1990) 103-104.

[18] M. Hall, Jr., A problem in partitions, *Bull. Amer. Math. Soc.*, **47** (1941) 801-807.

[19] Y. Kohayakawa, A note on induced cycles in Kneser graphs, *Combinatorica*, **11** (1991) 245-251.

[20] S.D. Monson, The effects of vertex deletion and edge deletion on the clique partition number, *Ars Combin.* **42** (1996), 89-96.

[21] S.D. Monson, N.J. Pullman, and R. Rees, A survey of clique and biclique coverings and factorizations of $(0,1)$-matrices, *Bull. Inst. Combin. Appl.* **14** (1995), 17-86.

[22] J. Orlin, Contentment in Graph Theory, *Indag. Math.* **39** (1977), 406-424.

[23] N.J. Pullman and A. Donald, Clique coverings of graphs II - Compliments of cliques, *Utilitas Mathematica* **19** (1981) 207-213.

[24] E. Spilrajn-Marczewski, Sur des propriétés des classes d'ensembles, *Fund. Math.* **33** (1945), 303-307.

[25] K. Vander Meulen, Induced subgraphs of Kneser graphs, M.Sc. Thesis, Queen's University, Kingston, ON (1991).

[26] W.D. Wallis, Asymptotic values of clique partition numbers, *Combinatorica*, **2**(1) (1982) 99-101.

[27] W.D. Wallis, Clique partitions of the complement of a one-factor, *Congressus Numerantium*, **46** (1985) 317-319.

[28] W.D. Wallis, Finite planes and clique partitions, *Contemporary Mathematics*, **111** (1990) 279-285.

[29] W.D. Wallis, The clique partition number of the complement of a cycle, *Annals of Discrete Mathematics*, **27** (1985) 335-344.