Communauté UNIVERSITÉ Grenoble Alpes

THÈSE

pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Informatique et Mathématiques Appliquées**

Arrêté ministériel : 25 mai 2016

Présentée par
Andrea Munaro

Thèse dirigée par **András Sebő** et **Matěj Stehlík**

préparée au sein du
Laboratoire G-SCOP
dans **l'école doctorale Mathématiques, Sciences et Technologies de l'Information, Informatique (MSTII)**

Sur quelques invariants classiques et nouveaux des hypergraphes

Thèse soutenue publiquement le **1er décembre 2016**,
devant le jury composé de:

Dieter Rautenbach
Professeur des Universités, Universität Ulm, Rapporteur
Nicolas Trotignon
Directeur de Recherche CNRS, ENS de Lyon, Rapporteur
András Gyárfás
Senior Researcher, Alfréd Rényi Institute of Mathematics, Examinateur
Frédéric Maffray
Directeur de Recherche CNRS, Univ. Grenoble Alpes, Laboratoire G-SCOP, Examinateur et Président du jury
András Sebő
Directeur de Recherche CNRS, Univ. Grenoble Alpes, Laboratoire G-SCOP, Directeur de thèse
Matěj Stehlík
Maître de Conférences, Univ. Grenoble Alpes, CNRS, Laboratoire G-SCOP, Co-encadrant de thèse

Acknowledgements

First of all, I am grateful to my supervisors András Sebő and Matěj Stehlík. Thank you for your support, encouragement, patience and for all the helpful suggestions. I also appreciated a lot the freedom you gave me during these intense three years.

Then, I would like to thank the other members of my thesis committee: András Gyárfás, Frédéric Maffray, Dieter Rautenbach and Nicolas Trotignon. Thank you for accepting to be part of it and for travelling to Grenoble the day of my defense.

I would also like to thank all the members of the "Optimisation Combinatoire" group, my fellow Ph.D. students, in particular Lucas, Quentin and Yohann, and my officemates Kaveh, Saleh and Van-Thao. A special thanks goes to Javi, Miguel and Teri for the nice reunions we had.

I am also grateful to Fadila Messaoud-Djebara, Marie-Josephe Perruet, Christine Rouzier and Myriam Torlini for their support in administrative matters.

Finally, I want to express my gratitude to my family, my friends and to you, Marie.

Contents

List of Figures

Introduction

In general, the structure of a (hyper)graph can be complicated, both from a combinatorial and algorithmic point of view. On the other hand, it is often the case that restricting the range makes possible to obtain certain properties. In this thesis, we consider several (hyper)graph parameters and study whether restrictions to subclasses of hypergraphs allow to obtain desirable combinatorial or algorithmic properties. Most of the parameters we consider are special instances of packings and transversals of hypergraphs and we examine essentially three kinds of "desirable" properties:

- Existence of a large packing;
- Existence of an upper bound for the transversal number in terms of the packing number;
- Polynomial-time decidability of the packing and transversal numbers and of some (hyper)graph properties.

Along the way, we also consider a prominent measure of the "complexity" of a hypergraph: its VC-dimension.

Let us now properly define the notions introduced above. A packing of a hypergraph $\mathcal{H} = (V, E)$ is a set of pairwise disjoint edges of $\mathcal{H}$. A related notion is that of a transversal (also known as hitting set or covering) of $\mathcal{H}$, which is a subset $X \subseteq V$ intersecting each edge of $\mathcal{H}$. Clearly, every hypergraph has a packing (the empty set) and so it is natural to look for maximum packings, namely packings with as many edges as possible. Similarly, every hypergraph has a transversal (the vertex set) and we are interested in minimum transversals, namely transversals with as few vertices as possible. The packing number $\nu(\mathcal{H})$ is the number of edges in a packing of $\mathcal{H}$ of maximum size (a maximum packing) and the transversal number $\tau(\mathcal{H})$ is the number of vertices in a transversal of $\mathcal{H}$ of minimum size (a minimum transversal).

In Chapter 2, we study the packing number of two hypergraphs arising from graphs. The first is obtained by considering the graph itself, and so a packing is what is usually known as a matching. The other hypergraph is the dual of the clique hypergraph of a graph G, where the clique hypergraph is the hypergraph having as vertices the vertices of G and as edges the maximal cliques of G. It is easy to see that a packing of this hypergraph is nothing but an independent set of G. A natural relation between the notions of matching and independent set arises when considering the class of line graphs[1]: Indeed, there is a bijection between the matchings of a graph and the independent sets of its line graph. Line graphs constitute a rich and ubiquitous class of graphs, introduced by Whitney [183]. In Chapter 2, we concentrate on a subclass of line graphs: line graphs of subcubic triangle-free graphs. In particular, we provide several characterizations of this class and we study the independence number of its graphs. The famous Brooks' Theorem asserts that every connected graph G which is neither a complete graph nor an odd cycle must be $\Delta(G)$-colourable and so $\alpha(G) \geq |V(G)|/\Delta(G)$. Following this result, several authors considered the problem of finding tight lower bounds for the independence number of graphs having bounded maximum degree and not containing cliques on 3 or 4 vertices [66, 67, 93, 129, 176]. In particular, Kang et al. [104] showed that if G is a connected (K_4, claw)-free 4-regular graph on n vertices then, apart from three

[1] The line graph of a graph G is the graph having as vertices the edges of G, two vertices being adjacent if the corresponding edges intersect.

exceptions, $\alpha(G) \geq (8n-3)/27$. We show that if G is a $(K_4, \text{claw}, \text{diamond})$-free graph on n vertices, then the tight bound $\alpha(G) \geq 3n/10$ holds.

The question of whether a graph admits a perfect matching has been deeply investigated but not much is known about general lower bounds for the matching number. Biedl et al. [24] showed that any subcubic graph G has a matching of size $(|V(G)|-1)/3$ and that any cubic graph G has a matching of size $(4|V(G)|-1)/9$ and Henning et al. [95] showed that any connected cubic triangle-free graph G has a matching of size $(11|V(G)|-2)/24$. Given the correspondence between matchings of a graph and independent sets of its line graph, our previous lower bound for the independence number implies the following tight lower bound for the matching number: if G is a subcubic triangle-free graph with n_i vertices of degree i, then $\alpha'(G) \geq 3n_1/20 + 3n_2/10 + 9n_3/20$.

The problem of deciding, given a hypergraph $\mathcal{H}$ and an integer k, whether $\tau(\mathcal{H}) \leq k$ is NP-complete in general. In Section 2.2.3, we consider the special case where $\mathcal{H}$ is the cycle hypergraph of G, i.e. the hypergraph whose vertices are the vertices of G and whose edges are the vertex sets of cycles of G. This problem is known as FEEDBACK VERTEX SET. Speckenmeyer [174, 175] showed that it remains NP-hard even for planar graphs with maximum degree 4 and Ueno et al. [180] showed that it becomes solvable in polynomial time for subcubic graphs. In Section 2.2.3, we strengthen the results in [174, 175] by showing the NP-hardness for line graphs of planar cubic bipartite graphs and we provide an inapproximability result for line graphs of subcubic triangle-free graphs. Finally, in Section 2.2.4, we address two other well-known NP-complete graph problems: HAMILTONIAN CYCLE and HAMILTONIAN PATH. In particular, we show that they remain NP-hard for some subclasses of line graphs of planar cubic bipartite graphs, thus strengthening a result by Lai and Wei [119].

One possible generalization of line graphs is given by the following construction: For an integer $k \geq 2$, the k-line graph $L_k(G)$ of a graph G is the graph having as vertices the cliques of G of size k, two vertices being adjacent if the corresponding cliques intersect in a clique of size $k-1$. This notion has been introduced independently and with different motivations by several authors [43–45]. Clearly, 2-line graphs are the usual line graphs, whereas 3-line graphs are also known as triangle graphs. Unlike line graphs, the class of k-line graphs with $k \geq 3$ is not hereditary. Nevertheless, motivated by Tuza's Conjecture[2], we provide in Section 2.3 a partial list of forbidden induced subgraphs for this class.

Let us now consider again packings and coverings of a hypergraph $\mathcal{H}$ and see how they interact. Since no vertex covers two edges of a packing, we have $\tau(\mathcal{H}) \geq \nu(\mathcal{H})$. Therefore, a large packing can be considered as an obstruction to a small transversal. A family of hypergraphs satisfies the Min-Max Property if $\nu(\mathcal{H}) = \tau(\mathcal{H})$, for each member $\mathcal{H}$ of the family. There are several families of hypergraphs satisfying the Min-Max Property and the most prominent example is probably given by the well-known König-Egerváry Theorem[3]. The Min-Max Property allows a good characterization of the packing and transversal numbers. Indeed, to show that $\nu(\mathcal{H}) \leq k$ or $\tau(\mathcal{H}) \geq k$, it is enough to exhibit a transversal or a packing of size k, respectively. Unfortunately, most families of hypergraphs do not satisfy the Min-Max Property, but it is still of interest to find an upper bound for τ in terms of ν, if any. A family of hypergraphs satisfies the Erdős-Pósa Property if there exists a function f such that $\tau(\mathcal{H}) \leq f(\nu(\mathcal{H}))$, for each member $\mathcal{H}$ of the family. This implies that one parameter is characterized by its obstructing analogue, or dual: either $\mathcal{H}$ contains a packing of size k or it contains a transversal of size $f(k)$. The family of r-uniform hypergraphs satisfies the Erdős-Pósa Property. Indeed, consider

[2] This conjecture will be covered in Chapter 3. [3] This theorem asserts that the family of bipartite graphs satisfies the Min-Max Property.

an r-uniform hypergraph $\mathcal{H}$ and a maximal packing of $\mathcal{H}$. Since the union of the edges in this packing intersects all the edges of $\mathcal{H}$, we have $\tau(\mathcal{H}) \leq r\nu(\mathcal{H})$ and it is not difficult to see that this inequality is tight.

In Chapter 3, we consider three families of hypergraphs satisfying the Erdős-Pósa Property and we seek to determine the optimal bounding functions. The first family consists of duals of clique hypergraphs. It is easy to see that the transversals of such a hypergraph correspond to the clique covers of the underlying graph and we have already mentioned that packings correspond to independent sets. For historical reasons, it is convenient to stick to the underlying graph and we say that a class of graphs $\mathcal{G}$ is θ-bounded if there exists a function $f : \mathbb{N} \to \mathbb{R}$ such that for all $G \in \mathcal{G}$ and all induced subgraphs H of G, we have $\theta(H) \leq f(\alpha(H))$. Such a function f is a θ-bounding function for $\mathcal{G}$. Gyárfás [81] introduced the concept of θ-bounded class in order to provide a natural extension of the class of perfect graphs: indeed, this class is exactly the class of graphs θ-bounded by the identity function. In [81], he also formulated the following meta-question: given a class $\mathcal{G}$, what is the smallest θ-bounding function for $\mathcal{G}$, if any? In Section 3.2, we consider this question for classes of graphs having bounded maximum degree. It is easy to see that $\theta(G) \leq k\alpha(G)$, for any graph G with maximum degree at most k. On the other hand, for $k = 3$, we show that this bound is far from optimal: $f(x) = \lfloor \frac{3}{2}x \rfloor$ is a θ-bounding function for the class of subcubic graphs and it is best possible. Moreover, we give some insight for the case of graphs with maximum degree four. The study of these two cases is also motivated by a result of Henning et al. [95] showing that $\theta(G) \leq \frac{3}{2}\alpha(G)$, for any subcubic triangle-free graph G, and by a result of Joos [101] showing that $\theta(G) \leq \frac{7}{4}\alpha(G)$, for any triangle-free graph G with maximum degree four. In Section 3.2, we also treat some algorithmic aspects related to clique covering and we show, in particular, that the problem of finding a minimum-size clique cover of a planar graph admits a polynomial-time approximation scheme.

In Section 3.3, we consider the triangle hypergraph of a graph G, which is the hypergraph having as vertices the edges of G and whose edges are the subsets spanning triangles of G. Since the triangle hypergraph is 3-uniform, its transversal number is at most three times the packing number. In other words, the minimum number of edges of G whose deletion results in a triangle-free graph is at most three times the maximum number of edge-disjoint triangles of G. In fact, Tuza [178] conjectured that this can be improved: the transversal number of the triangle hypergraph is at most twice the packing number. If true, this conjecture would be tight, as shown by the complete graph on 4 vertices. Several partial results on Tuza's Conjecture have been obtained and in Section 3.3 we concentrate on some subclasses of K_4-free graphs. The classes we consider are essentially of two kinds: graphs with edges in few triangles (at most four) and graphs obtained by forbidding certain odd-wheels. We show that, in these cases, it is in fact possible to considerably reduce the constant 2 in Tuza's Conjecture. The results proved in Section 3.2 play a big role in our reasoning.

Finally, in Section 3.4, we consider the cycle hypergraph. Contrary to the triangle hypergraph, the cycle hypergraph need not be uniform. Nevertheless, a fundamental result by Erdős and Pósa [59] asserts that the transversal number of cycle hypergraphs is bounded by a function of the packing number. Kloks et al. [107] conjectured that cycle hypergraphs of planar graphs admit the constant function 2 as a bounding function. In other words, the minimum number of vertices of a planar graph whose deletion makes it acyclic is at most twice the maximum number of vertex-disjoint cycles. This is known as Jones' Conjecture and, if true, it would be best possible, as shown by wheel graphs. Not much is known about Jones' Conjecture and, in Section 3.4, we show it holds for claw-free graphs with maximum degree

4 and we provide some properties a minimum subcubic counterxample must have, if any.

In Chapter 4, we consider a prominent measure of the "complexity" of a set system[4]: its VC-dimension. Given a set system $\mathcal{H}$ on X and a subset $Y \subseteq X$, we say that Y is shattered by $\mathcal{H}$ if $\{E \cap Y : E \in \mathcal{H}\} = 2^Y$ and the VC-dimension of $\mathcal{H}$ is defined as the maximum size of a set shattered by $\mathcal{H}$, or as ∞ if arbitrarily large subsets can be shattered. The notion of VC-dimension was introduced by Vapnik and Chervonenkis [181] and, in Chapter 4, we are interested in the VC-dimension of set systems arising from graphs. For every family $\mathcal{P}$ of subgraphs of a given graph G, we can naturally define the VC-dimension $\mathsf{VC}_{\mathcal{P}}(G)$ of G with respect to the family $\mathcal{P}$ as the VC-dimension of the set system induced by $\mathcal{P}$. In this way we obtain several different notions of VC-dimension, each one related to a special family of subgraphs. The VC-dimension with respect to some of these families is equal to well-known parameters: if $\mathcal{P}$ is the family of complete subgraphs then $\mathsf{VC}_{\mathcal{P}}$ is the clique number, while if $\mathcal{P}$ is the family of subgraphs induced by independent sets then $\mathsf{VC}_{\mathcal{P}}$ is the independence number.

Kranakis et al. [115] initiated a systematic study of the VC-dimensions of graphs with respect to families of subgraphs. In particular, they showed that the VC-dimension $\mathsf{VC}_{\mathsf{con}}(G)$ of a graph G with respect to connected subgraphs differs by at most 1 from the connected domination number of G. In Chapter 4, we continue their systematic study and we concentrate on the VC-dimension with respect to k-connected subgraphs. Given a graph G, this quantity can be thought as the maximum size of a subset $A \subseteq V(G)$ such that, no matter how many vertices of A are deleted from G, there is a k-connected subgraph of G containing the remaining vertices of A. We extend the results in [115] by providing tight upper and lower bounds for the VC-dimension with respect to k-connected subgraphs, for $k \geq 2$.

Papadimitriou and Yannakakis [155] considered the problem of deciding the VC-dimension of a general set system: Given a set system $\mathcal{H}$ (by its incidence matrix) and an integer s, does $\mathcal{H}$ have VC-dimension at least s? They introduced the complexity class LOGNP and showed that the problem in question is complete for it. In this context, it is natural to investigate GRAPH $\mathsf{VC}_{\mathcal{P}}$ DIMENSION: the problem of deciding, given a graph G and an integer s, whether $\mathsf{VC}_{\mathcal{P}}(G) \geq s$ holds. Kranakis et al. [115] showed that GRAPH $\mathsf{VC}_{\mathsf{con}}$ DIMENSION is NP-complete. In Section 4.3, we extend this result by showing that GRAPH $\mathsf{VC}_{k\text{-}\mathsf{con}}$ DIMENSION is NP-complete even for split graphs, for any k. On the positive side, we show it can be decided in linear time for graphs of bounded clique-width and in polynomial time for the subclass of split graphs having Dilworth number at most 2. Finally, we prove that GRAPH $\mathsf{VC}_{k\text{-}\mathsf{con}}$ DIMENSION remains NP-hard for some subclasses of planar bipartite graphs in the cases $k = 1$ and $k = 2$.

In Section 4.4, we provide complexity dichotomies for GRAPH $\mathsf{VC}_{\mathsf{con}}$ DIMENSION and CONNECTED DOMINATING SET when restricted to classes of graphs obtained by forbidding a single induced subgraph (monogenic classes). The first dichotomy in monogenic classes was obtained by Korobitsin [108] for DOMINATING SET and only few other dichotomies in monogenic classes are known [1, 79, 102, 114]. Our results show that the complexities of GRAPH $\mathsf{VC}_{\mathsf{con}}$ DIMENSION, CONNECTED DOMINATING SET and DOMINATING SET all agree in monogenic classes.

In Chapter 5, we study in more details some of the algorithmic graph problems mentioned above (and some others). We have seen that most of them are NP-hard even for restricted classes of graphs, while they might become solvable in polynomial time for some subclasses. Therefore, assuming that P $\neq$ NP, it is natural to ask when a certain "hard" graph problem

[4] This notion coincides with that of hypergraph.

becomes "easy": Is there any "boundary" separating "easy" and "hard" instances? In Chapter 5, we consider this question for hereditary graph classes. For a problem Π, we say that a hereditary class of graphs X is Π-hard if Π is NP-hard for X, and Π-easy if Π is solvable in polynomial time for graphs in X. In a first attempt to answer the meta-question posed above, one might be tempted to consider maximal Π-easy classes and minimal Π-hard classes. In fact, the first approach immediately turns out to be meaningless: there are no maximal Π-easy classes. Moreover, minimal Π-hard classes might not exist at all, as there might exist infinite descending chains of Π-hard classes. This suggests that the "limit" of an infinite "decreasing" sequence of Π-hard classes should play a role in the search of a "boundary" between easy and hard classes. Alekseev [5] formalized this intuition by introducing the notions of limit class and boundary class for INDEPENDENT SET. In fact, these concepts are completely general and can be stated as follows [8]. Given an NP-hard graph problem Π and a Π-hard class of graphs X, a class Y is a limit class for Π with respect to X ((Π, X)-limit in short) if there exists a sequence $Y_1 \supseteq Y_2 \supseteq \dots$ of Π-hard subclasses of X such that $\bigcap_{n \geq 1} Y_n = Y$. The class Y is a limit class for Π (Π-limit in short) if there exists a Π-hard class X such that Y is (Π, X)-limit. Finally, an inclusion-wise minimal (Π, X)-limit class is a boundary class for Π with respect to X ((Π, X)-boundary in short) and a class Y is a boundary class for Π (Π-boundary) if there exists a Π-hard class X such that Y is (Π, X)-boundary.

Alekseev [5] and Alekseev et al. [8] showed that a boundary class with respect to a Π-hard class represents indeed a meaningful notion of "boundary" between Π-hard and Π-easy subclasses: A class X is Π-hard if and only if it contains a (Π, X)-boundary class. Moreover, Π-boundary classes can be used to characterize the finitely defined graph classes[5] which are Π-hard: A finitely defined class is Π-hard if and only if it contains a Π-boundary class [8].

Alekseev [5] studied INDEPENDENT SET and revealed the first boundary class for this problem. Other problems have been studied in the context of boundary classes. For example, Alekseev et al. [7] revealed three boundary classes for DOMINATING SET, Malyshev [135] found a fourth boundary class for this problem and Korpelainen et al. [110] revealed two boundary classes for HAMILTONIAN CYCLE. So far, the complete description of boundary classes has been obtained only for a single problem, the so-called LIST EDGE-RANKING [136]. In Chapter 5, we continue the study of boundary classes for NP-hard problems. In Section 5.2, we provide the first boundary class for the closely related HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE and HAMILTONIAN PATH. In Section 5.3, we reveal the first boundary class for FEEDBACK VERTEX SET. Finally, in Sections 5.4 and 5.5 we make some progress towards the determination of some boundary classes for two other problems involving non-local properties: CONNECTED DOMINATING SET and CONNECTED VERTEX COVER.

Chapter 2 is partially based on [144], Chapter 3 on [143, 144], Chapter 4 on [145] and Chapter 5 on [142].

[5] Recall that a hereditary class is finitely defined if the set of its minimal forbidden induced subgraphs is finite.

CHAPTER 1
Preliminaries

In this chapter we introduce most terminology and basic results used in the thesis. Further definitions are presented later, in special cases where more explanations are necessary.

A reader familiar with the basic topics in graph theory may skip single sections or even the whole chapter. We refer to [28, 52, 182] for complete introductions to graph theory and for the missing proofs.

Graphs and subgraphs. A *graph* is a pair $G = (V, E)$ consisting of a set V of *vertices* and a set E (disjoint from V) of *edges* together with a function that associates to each edge e a pair of (not necessarily distinct) vertices, called the *endpoints* of e. If the edge e has endpoints u and v, we usually write $e = uv$ in place of $e = \{u, v\}$. In this thesis we consider only *finite* graphs, namely graphs with a finite number of vertices. Given a graph G, we usually denote its vertex set by $V(G)$ and its edge set by $E(G)$. The *order* of G is the number of vertices of G, sometimes denoted by $n(G)$, and the size of G is the number of edges of G, also denoted by $m(G)$. A *loop* is an edge whose endpoints are equal and *multiple edges* are edges having the same pair of endpoints. For the most part of this thesis, we consider only *simple* graphs, namely graphs with no loops or multiple edges. However, in some parts of Sections 3.4.1 and 5.3 we allow loops and multiple edges and this will be explicitly mentioned whenever it is the case.

A graph $G' = (V', E')$ is a *subgraph* of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$, and we write $G' \subseteq G$. Moreover, G' is a *spanning subgraph* of G if G' is a subgraph of G and $V' = V$. Finally, if G' is a subgraph of G and G' contains all the edges of G with both endpoints in V', then G' is an *induced subgraph* of G and we we write $G' = G[V']$.

Two (simple) graphs $G = (V, E)$ and $G' = (V', E')$ are *isomorphic* if there exists a bijection $\varphi\colon V \to V'$ such that $xy \in E$ if and only if $\varphi(x)\varphi(y) \in E'$, for each pair of vertices x and y.

Neighbourhoods and degrees. A vertex v of a graph G is *incident* to an edge $e \in E(G)$ if it is an endpoint of e. Two vertices of G are *adjacent*, or *neighbours*, if they are endpoints of an edge of G. Two edges are *incident* if they share an endpoint. For a vertex $v \in V(G)$, the *neighbourhood* $N_G(v)$ is the set of vertices adjacent to v in G and the *closed neighbourhood* $N_G[v]$ is the set $N(v) \cup \{v\}$. Moreover, if $S \subseteq V(G)$, then $N_G(S) = \bigcup_{v \in S} N_G(v)$ and $N_G[S] = \bigcup_{v \in S} N_G[v]$. Note that we usually drop the subscripts when the context is clear.

Given two subsets S and T of $V(G)$, the set of edges having one endpoint in S and the other in T is denoted by $[S, T]$. An *edge cut* is a set of edges of the form $[S, \overline{S}]$, where S is a non-empty proper subset of $V(G)$ and $\overline{S} = V(G) \setminus S$.

The *degree* $d_G(v)$ of a vertex $v \in V(G)$ is the number of edges incident to v in G, with the exception that each loop counts as two edges. A graph is *even* if all its vertices have even degrees. A *k-vertex* is a vertex of degree k and we refer to a 3-vertex as a *cubic vertex*. A 0-vertex is an *isolated* vertex. We denote by $d_k(G)$ the set of k-vertices of G and we usually write $n_k(G)$ for $|d_k(G)|$. The *maximum degree* $\Delta(G)$ of G is the number $\max\{d_G(v) : v \in V\}$ and G is *subcubic* if $\Delta(G) \leq 3$. Similarly, the *minimum degree* $\delta(G)$ of G is the quantity $\min\{d_G(v) : v \in V\}$. If all the vertices of G have the same degree k, then G is *k-regular* and a 3-regular graph is usually called *cubic*. A *k-factor* of a graph is a spanning k-regular subgraph.

Paths and cycles. A *path* is a non-empty graph $P = (V, E)$ with $V = \{x_0, x_1, \ldots, x_k\}$ and $E = \{x_0x_1, x_1x_2, \ldots, x_{k-1}x_k\}$, and where the x_i are all distinct. The vertices x_0 and x_k are *linked* by P and they are called the *ends* of P. The vertices $x_1, \ldots, x_{k-1}$ are the *inner* vertices of P. The size of a path is its *length* and the path of order n is denoted by P_n. We refer to a path P by a natural sequence of its vertices: $P = x_0x_1 \cdots x_k$. Such a path P is a path *between* x_0 and x_k, or a x_0, x_k-path. More generally, given a graph $G = (V, E)$ and two subsets X and Y of V, an *X, Y-path* is a path which has one end in X, the other end in Y, and whose inner vertices belong to neither X nor Y. Two or more paths are *independent* if none of them contains an inner vertex of another. A family of independent x, Y-paths with distinct ends in Y is an *x, Y-fan*.

The *distance* $d_G(u, v)$ from a vertex u to a vertex v in a graph G is the minimum length of a path between u and v. If u and v are not linked by any path in G, we set $d_G(u, v) = \infty$. The greatest distance between any two vertices of G is the *diameter* of G, denoted by $\operatorname{diam}(G)$. The radius of G is the quantity $\min_{x\in V(G)} \max_{y\in V(G)} d_G(x, y)$.

If $P = x_0 \cdots x_k$ is a path and $k \geq 2$, then the graph $P + x_kx_0$ with vertex set $V(P)$ and edge set $E(P) \cup \{x_kx_0\}$ is called a *cycle*. The cycle on n vertices is denoted by C_n. The *girth* of a graph containing a cycle is the length of a shortest cycle and a graph with no cycle has infinite girth.

A *Hamiltonian path* of a graph G is a path of G which is spanning. A *Hamiltonian cycle* of G is a spanning cycle of G and a graph is *Hamiltonian* if it contains a Hamiltonian cycle.

Graph operations. Let $G = (V, E)$ be a graph and $V' \subseteq V$ and $E' \subseteq E$. The operation of *deleting the set of vertices V'* from G results in the graph $G - V' = G[V \setminus V']$. The operation of *deleting the set of edges E'* from G results in the graph $G - E' = (V, E \setminus E')$. The *complement* of a simple graph G is the graph $\overline{G}$ with vertex set $V(G)$ and such that $uv \in E(\overline{G})$ if and only if $uv \notin E(G)$. The *union* of simple graphs G and H is the graph $G \cup H$ with vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H)$. If G and H are disjoint, their union is denoted by $G + H$ and the union of k (disjoint) copies of G is denoted by kG. The *join* of simple graphs G and H is the graph $G \vee H$ obtained from $G + H$ by adding the edges $\{xy : x \in V(G), y \in V(H)\}$. A *$k$-subdivision* of G is the graph obtained from G by adding k new vertices for each edge of G, i.e. each edge is replaced by a path of length $k + 1$. The *cartesian product* $G \square H$ of G and H is the graph with vertex set $V(G) \times V(H)$ and such that (u, v) is adjacent to (u', v') if and only if either $u = u'$ and $vv' \in E(H)$ or $v = v'$ and $uu' \in E(G)$.

Graph classes and special graphs. If a graph does not contain induced subgraphs isomorphic to graphs in a set Z, then it is *Z-free* and the set of all Z-free graphs is denoted by *Free*(Z). A class of graphs is *hereditary* if it is closed under deletions of vertices. It is well-known and easy to see that a class of graphs X is hereditary if and only if it can be defined by a set of forbidden induced subgraphs, i.e. $X =$ *Free*(Z) for some set of graphs Z. The minimal set Z with this property is unique and it is denoted by *Forb*(X). If the set of minimal forbidden induced subgraphs for a hereditary class X is finite, then X is *finitely defined*. If $X \subseteq Y$ and *Forb*$(X) \setminus$ *Forb*(Y) is a finite set, then X is *defined by finitely many forbidden induced subgraphs with respect to* Y.

A *complete graph* is a graph whose vertices are pairwise adjacent and the complete graph on n vertices is denoted by K_n. A graph G is r*-partite*, for $r \geq 2$, if its vertex set admits a partition into r classes such that every edge has its endpoints in different classes. An r-partite graph in which every two vertices from different partition classes are adjacent is called *complete*. 2-partite graphs are usually called *bipartite* and the complete bipartite graph with partition classes of size n and m is denoted by $K_{n,m}$. A graph of the form $K_{1,n}$ is usually called a *star*.

The unique complete $(r-1)$-partite graph on $n \geq r-1$ vertices and whose partition classes differ in size by at most 1 is denoted by $T^{r-1}(n)$ and called a *Turán graph*. It is easy to see that $T^{r-1}(n)$ has at most $\frac{1}{2}n^2\frac{r-2}{r-1}$ edges. Turán graphs are extremal in the following sense:

Theorem 1.0.1 (Turán's Theorem). *For all integers r and n with $r > 1$, every graph of order n with no K_r subgraph and with the largest possible size is the Turán graph $T^{r-1}(n)$.*

A *triangle* is the graph K_3, a *claw* is the graph $K_{1,3}$ and a *diamond* is the graph obtained from K_4 by deleting an edge. For $n \geq 3$, an n*-wheel* W_n is the graph $C_n \vee K_1$. An *odd-wheel* is a graph W_n with n odd. Figure 1.1 depicts other recurrent graphs appearing in the thesis. Note that the prefix "co-" denotes the complement of a certain graph: for example, the co-banner is the complement of the banner.

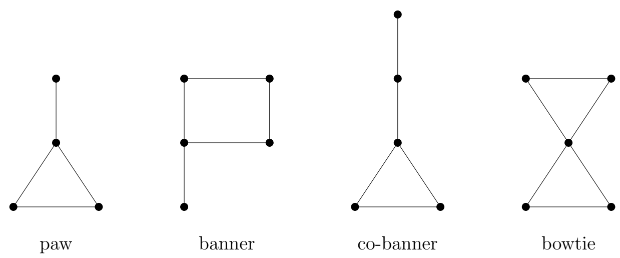


Figure 1.1: Some special graphs.

A *hole* in a graph G is an induced subgraph isomorphic to C_n, for $n \geq 4$. An *antihole* in G is an induced subgraph isomorphic to the complement of C_n, for $n \geq 4$. An *odd-hole* is a hole isomorphic to C_n, with n odd. Similarly, an *odd-antihole* is an antihole isomorphic to $\overline{C_n}$, with n odd.

A *split graph* is a graph in which the vertex set can be partitioned into a clique and an independent set (see below for the definitions of clique and independent set). A *cograph* (or *complement reducible* graph) is defined recursively as follows: K_1 is a cograph, the disjoint union of cographs is a cograph, the complement of a cograph is a cograph. In fact, the class of cographs coincides with that of P_4-free graphs.

Connectivity. A non-empty graph $G = (V, E)$ is *connected* if any two of its vertices are linked by a path in G. A *component* of a graph is a maximal connected subgraph. A *separating set* or *vertex cut* of G is a set $S \subseteq V(G)$ such that $G - S$ has more than one component. Given two vertices x and y of G, an x, y*-cut* is a set $S \subseteq V(G) \setminus \{x, y\}$ such that $G - S$ has no x, y-path. A *cut-vertex* of a graph is a vertex whose deletion increases the number of components.

A *block* of G is a maximal connected subgraph with no cut-vertex. The *connectivity* of G is the minimum size of a vertex set S such that $G - S$ is not connected or has only one vertex. A graph G is k-connected if its connectivity is at least k. By definition, a graph different from K_1 is connected if and only if it is 1-connected. Moreover, we consider K_1 as 1-connected.

The following easy but useful property of k-connectedness will be employed in Chapter 4:

Lemma 1.0.2 (Expansion Lemma). *If G is a k-connected graph and G' is obtained from G by adding a new vertex with at least k neighbours in G, then G' is k-connected.*

The following fundamental min-max theorem will also be used in Chapter 4 (see also Section 3.1):

Theorem 1.0.3 (Menger's Theorem). *Let G be a graph and x and y two of its vertices. The minimum size of an x,y-cut is equal to the maximum number of independent x,y-paths. Moreover, G is k-connected if and only if it contains k independent paths between any two vertices.*

An easy corollary of Theorem 1.0.3 is given by the following:

Lemma 1.0.4 (Fan Lemma). *For $k \geq 2$, a graph G is k-connected if and only if it has at least $k+1$ vertices and, for every choice of $x \in V(G)$ and $U \subseteq V(G)$ with $|U| \geq k$, it has an x,U-fan of size k.*

A *cut-edge* (or *bridge*) of a graph is an edge whose deletion increases the number of components. A *bridgeless graph* is a graph without cut-edges. A *disconnecting set* of edges is a set $F \subseteq E(G)$ such that $G - F$ has more than one component and a graph is *k-edge-connected* if every disconnecting set has at least k edges.

A graph G is *cyclically k-edge-connected* if the deletion of fewer than k edges from G does not create two components both containing at least one cycle.

Trees. A *tree* is a connected graph not containing any cycle as a subgraph. The vertices of degree 1 are its *leaves*. The following assertions can all be easily verified. Every non-trivial tree has at least two leaves. A graph T is a tree if and only if any two vertices of T are linked by a unique path in T. A connected graph with n vertices is a tree if and only if it has $n-1$ edges. Every connected graph contains a spanning tree.

If we consider one vertex of a tree as special, we call it the *root* of the tree. A tree with a fixed root is a *rooted tree*. Let T be a tree rooted at r and, for each $v \in V(T)$, let $P(v)$ denote the unique v,r-path in T. The *descendants* of a vertex $v \in V(T)$ are the vertices u such that $P(u)$ contains v.

A spanning tree is a *maximum leaf spanning tree* (MLST) if there is no spanning tree with a larger number of leaves. Given a graph G, we denote by $\ell(G)$ the number of leaves in a MLST of G.

Graph parameters. In this thesis we often consider the following parameters of a graph G. An *independent set* of a graph is a set of pairwise non-adjacent vertices. The maximum size of an independent set of G is the *independence number* $\alpha(G)$. A *clique* of a graph is a set of pairwise adjacent vertices. The *clique number* $\omega(G)$ is the maximum size of a clique of G.

Clearly, we have $\alpha(G) = \omega(\overline{G})$.

A *colouring* of a graph G is a partition of $V(G)$ into independent sets and the minimum number of partition classes is the *chromatic number* $\chi(G)$. The graph G is k*-colourable* if $\chi(G) \leq k$. A *clique cover* of a graph is a set of cliques such that each vertex of the graph belongs to at least one of them. The minimum size of a clique cover of G is denoted by $\theta(G)$. Clearly, we have $\chi(G) = \theta(\overline{G})$.

A *matching* of a graph is a set of pairwise non-incident edges and the *matching number* $\alpha'(G)$ is the maximum size of a matching of G.

A *vertex cover* of a graph is a subset of vertices containing at least one endpoint of every edge. The minimum size of a vertex cover of G is denoted by $\beta(G)$. Clearly, $S \subseteq V(G)$ is a vertex cover of G if and only if $V(G) \setminus S$ is an independent set of G. A *connected vertex cover* of G is a vertex cover S of G such that $G[S]$ is connected and we denote by $\beta_c(G)$ the minimum size of a connected vertex cover of G.

An *edge cover* of a graph G is a subset $S \subseteq E(G)$ of edges such that every vertex of G is incident to an edge in S. We denote by $\beta'(G)$ the minimum size of an edge cover of G. A result known as one of the Gallai's identities asserts that $\alpha'(G) + \beta'(G) = |V(G)|$, for any graph G without isolated vertices (see below for an extension and a proof).

A *dominating set* of G is a subset $D \subseteq V(G)$ such that each vertex in $V(G) \setminus D$ is adjacent to a vertex in D. The minimum size of a dominating set of G is denoted by $\gamma(G)$. A *connected dominating set* of G is a dominating set D of G such that $G[D]$ is connected and we denote by $\gamma_c(G)$ the minimum size of a connected dominating set of G, also known as the *connected domination number* of G.

A *feedback vertex set* of G is a subset $T \subseteq V(G)$ such that $G - T$ is acyclic, i.e. it contains no cycle. We denote by $\tau_c(G)$ the minimum size of a feedback vertex set of G.

A *vertex triangle-transversal* of G is a subset $T \subseteq V(G)$ such that $G - T$ is triangle-free. We denote by $\tau_\Delta(G)$ the minimum size of a vertex triangle-transversal of G. Similarly, an *edge triangle-transversal* of G is a subset $T \subseteq E(G)$ such that $G - T$ is triangle-free and we denote by $\tau'_\Delta(G)$ the minimum size of an edge triangle-transversal of G.

A *nonseparating independent set* of G is an independent set $I \subseteq V(G)$ such that there is no $X \subseteq I$ for which $G - X$ has more components than G. Therefore, for a connected graph G, the complement of a nonseparating independent set of G is a connected vertex cover of G. We denote by $z(G)$ the maximum size of a nonseparating independent set of G.

Graphs on surfaces. For the contents of this section, we refer the reader to [140].

Let X be a topological space. A *curve* in X (or *arc*) is the image of a continuous function $f\colon [0,1] \to X$. The curve $f([0,1])$ is said to *join* (or *connect*) its *endpoints* $f(0)$ and $f(1)$. The curve is *simple* if f is injective. It is *closed* if $f(0) = f(1)$. A topological space X is *arcwise connected* if any two elements of X are connected by a simple arc in X. The existence of a simple arc between two points of X determines an equivalence relation whose equivalence classes are the *arcwise connected components* of X, or *regions*.

A graph G is *embedded* in a topological space X if the vertices of G are distinct elements of X and every edge of G is a simple arc connecting in X its two ends and such that the interior is disjoint from other edges and vertices. The *faces* of a graph G embedded in X are the regions of $X \setminus G$. An *embedding* of G in X is an isomorphism of G with a graph G' embedded in X and the graph G' is usually called a *drawing* of G in X. An embedding of G in X is *cellular* if every face of G' is homeomorphic to an open disc, where G' is the drawing of

G under the embedding. If there is an embedding of G in X, we say that G *can be embedded* in X.

Graphs which can be embedded in the Euclidean plane $\mathbb{R}^2$ are called *planar*. A graph which is embedded in the plane is a *plane* graph. The *outer face* of a plane graph is the unbounded face.

Theorem 1.0.5 (Jordan Curve Theorem). *For every simple closed curve C in the plane, $\mathbb{R}^2 \setminus C$ consists of exactly two arcwise connected components. Precisely one of them is bounded and C is the boundary of each region.*

Each face of a 2-connected plane graph G is bounded by a cycle of G. Each 3-connected planar graph has a unique planar drawing, in the sense that the facial cycles are uniquely determined. The *(geometric) dual* G^* of a 2-connected plane graph G is the plane graph having one vertex in each face of G and such that if e is an edge of G, then G^* has an edge e^* crossing e and joining the vertices of G^* in the two faces of G that contains e on the boundary.

Given a planar graph G and a drawing Γ of G in $\mathbb{R}^2$, we define L_1 to be the set of vertices incident to the outer face and, for $i > 1$, L_i is defined recursively as the set of vertices on the outer face of the planar drawing obtained by deleting the vertices in $\bigcup_{j=1}^{i-1} L_j$. We call L_i the *i-th layer* of the drawing Γ. A graph is *k-outerplanar* if it has a planar drawing with at most k layers. This notion was introduced by Baker [18]. 1-outerplanar graphs are simply known as *outerplanar* graphs.

Algorithms and complexity. We refer the reader to [15, 55] for introductions to computational complexity.

A *decision problem* Π is a pair $(\mathcal{I}, S)$, where $\mathcal{I}$ is the set of *instances* of Π and $S\colon \mathcal{I} \to \{\text{“yes”}, \text{“no”}\}$ is a function. *Solving* (or *deciding*) Π for an instance $I \in \mathcal{I}$ means deciding whether $S(I) =$ “yes” or $S(I) =$ “no”.

The *(worst-case) time-complexity* of an algorithm is the (worst-case) running time of the algorithm as a function of the input size, i.e. the maximum running time over all inputs of the same size. In order to estimate the running time of an algorithm, it is necessary to specify the computational model in which the algorithm is implemented and it turns out that a reasonable choice is the so-called Turing machine. Indeed, it is widely believed that every physically realizable computation device can be simulated by a Turing machine (this is the *Church–Turing thesis*). The complexity class P is then defined as the class of all decision problems that are solvable in polynomial time by a (deterministic) Turing machine.

A decision problem $\Pi = (\mathcal{I}_\Pi, S_\Pi)$ is *polynomial-time reducible* to a decision problem $\Lambda = (\mathcal{I}_\Lambda, S_\Lambda)$ (denoted by $\Pi \leq_p \Lambda$) if there exists a polynomial-time computable function $f\colon \mathcal{I}_\Pi \to \mathcal{I}_\Lambda$ (called a *reduction function* from Π to Λ) such that $S_\Pi(I) =$ “yes” if and only if $S_\Lambda(f(I)) =$ “yes”. It is easy to see that the relation $\leq_p$ is transitive and that if $\Pi \leq_p \Lambda$ and $\Lambda \in \mathsf{P}$, then $\Pi \in \mathsf{P}$.

The complexity class NP is formally defined as the class of all decision problems that are solvable in polynomial time by a nondeterministic Turing machine. Loosely speaking, a decision problem $\Pi = (\mathcal{I}, S)$ is in NP if there exists a polynomial-time algorithm $V(\cdot,\cdot)$ (a *verifier*) such that, for every instance $I \in \mathcal{I}$, $S(I) =$ “yes” if and only if there exists a *certificate* $C(I)$ of size polynomial in that of I such that $V(I, C(I))$ returns “yes”. Therefore, the problems in NP can be thought as those having “efficiently verifiable solutions”.

A (decision) problem Π is NP-hard if $\Lambda \leq_p \Pi$, for every $\Lambda \in$ NP and it is NP-complete if in addition $\Pi \in$ NP. Clearly, if Π is NP-hard and $\Pi \in$ P then P $=$ NP and so an NP-hard problem can be considered at least as hard as any other problem in NP. In Table 1.1, we summarize most of the NP-complete problems which will be considered in the thesis.

Let $\Sigma_1^{\mathsf{P}} = \mathsf{NP}$. For $i \geq 2$, the complexity class Σ_i^{P} is defined recursively as the class of decision problems solvable in polynomial time by a nondeterministic Turing machine having access to an oracle which decides problems in $\Sigma_{i-1}^{\mathsf{P}}$. The *polynomial hierarchy* is the set $\bigcup_i \Sigma_i^{\mathsf{P}}$.

An *optimization problem* Π is a quadruple $(\mathcal{I}, S, c, \mathrm{opt})$, where $\mathcal{I}$ is the set of *instances* of Π, $S(I)$ is the set of *feasible solutions* of an instance $I \in \mathcal{I}$, the function $c \colon \mathcal{I} \times S \to \mathbb{N}$ is the *objective function* and $\mathrm{opt} \in \{\max, \min\}$. *Solving* Π for an instance $I \in \mathcal{I}$ means finding a solution $s \in S(I)$ which maximizes or minimizes $c(I,s)$, according to whether $\mathrm{opt} = \max$ or $\mathrm{opt} = \min$. We denote by $\mathrm{opt}(I)$ the value $\mathrm{opt}\{c(I,s) : s \in S(I)\}$ and an *optimum solution* of an instance I is a feasible solution $s \in S(I)$ with $c(I,s) = \mathrm{opt}(I)$.

For $k \geq 1$, a *k-factor approximation algorithm* for an optimization problem $\Pi = (\mathcal{I}, S, c, \mathrm{opt})$ is a polynomial-time algorithm that computes for each instance $I \in \mathcal{I}$ a solution $s \in S(I)$ with

$$\max\left\{\frac{c(I,s)}{\mathrm{opt}(I)}, \frac{\mathrm{opt}(I)}{c(I,s)}\right\} \leq k.$$

A *polynomial-time approximation scheme* (PTAS, in short) for an optimization problem $\Pi = (\mathcal{I}, S, c, \mathrm{opt})$ is a polynomial-time algorithm accepting as input an instance $I \in \mathcal{I}$ and a constant $\varepsilon > 0$ and such that, for each fixed ε, it is a $(1+\varepsilon)$-approximation algorithm for Π.

Let $0 < \alpha < \beta$. A minimization problem $\Pi = (\mathcal{I}, S, c, \min)$ has an NP-*hard gap* $[\alpha, \beta]$ if there exist an NP-complete decision problem $\Lambda = (\mathcal{I}_\Lambda, S_\Lambda)$ and a polynomial-time reduction f from Λ to Π such that, for every $I \in \mathcal{I}_\Lambda$, the following holds:

- If $S_\Lambda(I) =$ "yes", then $\min(f(I)) \leq \alpha$;
- If $S_\Lambda(I) =$ "no", then $\min(f(I)) > \beta$.

We have an obvious analogue definition for a maximization problem.

Lemma 1.0.6. *If Π is an optimization problem with an* NP-*hard gap $[\alpha, \beta]$, for some $0 < \alpha < \beta$, then there is no $\frac{\beta}{\alpha}$-approximation algorithm for Π, unless* P $=$ NP.

A *gap-preserving reduction* from a maximization problem $\Pi = (\mathcal{I}_\Pi, S_\Pi, c_\Pi, \max)$ to a minimization problem $\Lambda = (\mathcal{I}_\Lambda, S_\Lambda, c_\Lambda, \min)$ is a function f mapping every instance of Π to an instance of Λ in polynomial time, together with constants $\alpha_\Pi \leq 1$ and $\alpha_\Lambda \geq 1$ and functions g_Π and g_Λ such that:

- If $\max(I) \geq g_\Pi(I)$, then $\min(f(I)) \leq g_\Lambda(f(I))$;
- If $\max(I) < \alpha_\Pi g_\Pi(I)$, then $\min(f(I)) > \alpha_\Lambda g_\Lambda(f(I))$.

The definition above can be easily adapted to the other three possible cases of a reduction between two optimization problems. Similarly to Lemma 1.0.6, the following holds:

Lemma 1.0.7. *Suppose there exists a gap-preserving reduction from a maximization problem $\Pi = (\mathcal{I}_\Pi, S_\Pi, c_\Pi, \max)$ to a minimization problem $\Lambda = (\mathcal{I}_\Lambda, S_\Lambda, c_\Lambda, \min)$. If it is* NP-*hard to distinguish between those $I \in \mathcal{I}_\Pi$ for which $\max(I) \geq g_\Pi(I)$ and those for which $\max(I) <$*

$\alpha_\Pi g_\Pi(I)$, then it is NP*-hard to distinguish between those $f(I) \in \mathcal{I}_\Lambda$ for which $\min(f(I)) \leq g_\Lambda(f(I))$ and those for which $\min(f(I)) > \alpha_\Lambda g_\Lambda(f(I))$. In particular, Λ is not approximable within α_Λ, unless* P = NP.

We use the standard notation for considerations on the asymptotic behaviour. Given two real-valued functions $f(n)$ and $g(n)$ depending on n, we write $f(n) = O(g(n))$ if there exists a constant $c > 0$ such that $|f(n)| \leq c \cdot |g(n)|$, for all sufficiently large n.

Hypergraphs. A *hypergraph* (or *set system*) $\mathcal{H}$ is a pair $\mathcal{H} = (X, \mathcal{F})$, where X is a set (the *vertex set*) and $\mathcal{F}$ is a family of subsets of X (the *hyperedges*). We refer to a hypergraph with vertex set X as a hypergraph on X. The vertex set of $\mathcal{H}$ is denoted by $V(\mathcal{H})$ and the family of hyperedges by $E(\mathcal{H})$. A hypergraph is *k-uniform* if all its hyperedges have size k and *complete* if it contains all possible hyperedges. For a hypergraph $\mathcal{H} = (X, \mathcal{F})$, the *dual hypergraph* $\mathcal{H}^* = (Y, \mathcal{G})$ is defined as follows: $Y = \{y_S : S \in \mathcal{F}\}$, where the y_S are pairwise distinct vertices and for each $x \in X$, we have that $\{y_S : S \in \mathcal{F},\ x \in S\}$ is a set in the family $\mathcal{G}$.

A *hitting set* (or *transversal*) of a set system $\mathcal{H} = (X, \mathcal{F})$ is a subset $T \subseteq X$ which intersects all the sets in $\mathcal{F}$. Given a set system $\mathcal{H}$, the *hitting set problem* consists in finding a minimum-size hitting set of $\mathcal{H}$. The dual problem is the *set cover problem*: given a set system $\mathcal{H} = (X, \mathcal{F})$, the goal is to find a minimum-size subfamily $\mathcal{F}' \subseteq \mathcal{F}$ such that $\bigcup_{F' \in \mathcal{F}'} F' = X$. Both problems are known to be NP-hard (see, e.g., [75]).

Polymatroids. We refer the reader to [112, 132] for introductions to matroids and 2-polymatroids, respectively.

A 2*-polymatroid* is a pair $P = (S, f)$, where S is a finite set and f is a function $f\colon 2^S \to \mathbb{Z}$ satisfying the following properties:

(P1) $f(\varnothing) = 0$;

(P2) $f(X) \leq f(Y)$, for any $X \subseteq Y \subseteq S$;

(P3) $f(X \cup Y) + f(X \cap Y) \leq f(X) + f(Y)$, for any $X, Y \subseteq S$;

(P4) $f(\{x\}) \leq 2$, for any $x \in S$.

If in addition $f(\{x\}) \leq 1$, for any $x \in S$, then P is a *matroid*. A subset $X \subseteq S$ is a *matching* of P if $f(X) = 2|X|$ and it is a *spanning set* of P if $f(X) = f(S)$. The maximum size of a matching of P is denoted by $\nu(P)$ while the minimum size of a spanning set of P is denoted by $\rho(P)$.

It is easy to see that, given a (simple) graph $G = (V, E)$, the pair $P = (E, f)$ with $f(X) = |\bigcup_{e \in X} e|$ is a 2-polymatroid. Moreover, a matching of P is a matching of G and a spanning set of P is an edge cover of G. We have already mentioned that the following holds:

Lemma 1.0.8 (Gallai's identity). *If G is a graph without isolated vertices, then $\alpha'(G) + \beta'(G) = |V(G)|$.*

In fact, Lovász provided the following generalization (see [132]):

Theorem 1.0.9. *For any 2-polymatroid $P = (S, f)$, we have $\rho(P) + \nu(P) = f(S)$.*

Proof. Let M be a matching of P of maximum size $\nu(P)$. It is easy to see that the function $f_M : 2^{S\setminus M} \to \mathbb{Z}$ defined by $f_M(X) = f(M \cup X) - f(M)$ satisfies (P1) to (P3). Moreover, by maximality, we have $f_M(\{x\}) = f(M \cup \{x\}) - f(M) \leq 1$, for any $x \in S \setminus M$. We claim that $f_M(X) = \max\{|Y| : Y \subseteq X,\ f_M(Y) = |Y|\}$, for every $X \subseteq S \setminus M$. Indeed, let $X \subseteq S \setminus M$ and $Y \subseteq X$ of maximum size such that $f_M(Y) = |Y|$. By maximality, we have $f_M(Y \cup \{x\}) < |Y| + 1$, for every $x \in X \setminus Y$. But then (P2) implies that $f_M(Y \cup \{x\}) = |Y|$ and repeated applications of (P3) give $f_M(X) = |Y|$.

Consider now a subset $B \subseteq S \setminus M$ of maximum size such that $f_M(B) = |B|$. By the previous claim, we have that $f_M(B) = f_M(S \setminus M)$ and so $f(M \cup B) = f(M) + f_M(B) = f(M) + f_M(S \setminus M) = f(S)$. Therefore, $M \cup B$ is a spanning set of P and we have

$$\rho(P) \leq |M \cup B| = |M| + f_M(S \setminus M) = |M| + f(S) - f(M) = f(S) - |M| = f(S) - \nu(P).$$

Conversely, let T be a spanning set of P of minimum size and consider a subset $M \subseteq T$ of maximum size such that $f(M) = 2|M|$. By maximality, the function $f_M : 2^{T\setminus M} \to \mathbb{Z}$ defined by $f_M(X) = f(M \cup X) - f(M)$ satisfies (P1) to (P3) and $f_M(\{x\}) \leq 1$, for any $x \in T \setminus M$. Therefore, we have that

$$f(S) = f(T) = f(M \cup (T \setminus M)) = f_M(T \setminus M) + f(M) \leq |T \setminus M| + 2|M| \leq |T| + |M|.$$

But then $\rho(P) = |T| \geq f(S) - |M| \geq f(S) - \nu(P)$. $\square$

In Section 3.4.1, we will see that also feedback vertex sets and nonseparating independent sets can be interpreted as spanning sets and matchings of a certain 2-polymatroid, respectively.

A 2-polymatroid (S, f) is *linearly representable* (over a field $\mathbb{F}$) if there exists a matrix $A = (A_e)_{e\in S} \in \mathbb{F}^{d\times 2S}$ obtained by concatenating $d \times 2$ matrices $A_e \in \mathbb{F}^{d\times 2}$ and such that $f(X) = \operatorname{rank} A(X)$, for any $X \subseteq S$, where d is a positive integer and $A(X) = (A_e)_{e\in X}$ denotes the submatrix of A obtained by selecting the corresponding columns.

Tree-width and Clique-width. Graphs of bounded tree-width are particularly interesting from an algorithmic point of view: many NP-complete problems can be solved in linear time for them. The notion of tree-width was introduced by Robertson and Seymour [162] in their seminal work on graph minors:

A *tree decomposition* of a graph $G = (V, E)$ is a pair (X, T), where $T = (I, F)$ is a tree and $X = \{X_i : i \in I\}$ is a family of subsets of V such that:

- $\bigcup_{i\in I} X_i = V$;
- for all edges $vw \in E$, there is an $i \in I$ such that $\{v, w\} \subseteq X_i$;
- for all vertices $v \in V$, the set $\{i \in I : v \in X_i\}$ forms a connected subtree of T.

The *width* of the tree decomposition (X, T) is $\max_{i\in I} |X_i| - 1$ and the *tree-width* of a graph G is the minimum width among all tree decompositions of G. It is easy to see that forests have tree-width at most 1 and the tree-width measures, loosely speaking, how far a given graph is from a tree. In fact, the graphs having tree-width at most k are exactly the so-called partial k-trees (see, e.g., [26] for a proof and other characterizations).

The rough idea is that, for certain problems, once a tree decomposition of the input graph with small width is found, it can be used in a dynamic programming algorithm to solve the original problem (see, e.g., [64] for some examples). For a fixed k, it is in fact possible to test in linear time whether a graph has tree-width at most k and, if so, to find a tree-decomposition with width at most k [25]. A celebrated algorithmic meta-theorem of Courcelle [46] provides a way to quickly establish that a certain problem is decidable in linear time on graphs of bounded tree-width: all (graph) problems expressible in monadic second-order logic with edge-set quantification are decidable in linear time on graphs of bounded tree-width, assuming a tree decomposition is given (see also [14]). Let us briefly recall that monadic second-order logic is an extension of first-order logic by quantification over sets. The language of *monadic second-order logic of graphs* (MSO_1 in short) contains the expressions built from the following elements:

- Variables $x, y, \ldots$ for vertices and $X, Y, \ldots$ for sets of vertices;
- Predicates $x \in X$ and $\mathit{adj}(x, y)$;
- Equality for variables, standard Boolean connectives and the quantifiers $\forall$ and $\exists$.

By considering edges and sets of edges as other sorts of variables and the incidence predicate $\mathit{inc}(v, e)$, we obtain *monadic second-order logic of graphs with edge-set quantification* (MSO_2 in short).

A notion related to tree-width is that of clique-width, introduced by Courcelle et al. [48]. The *clique-width* of a graph G is the minimum number of labels needed to construct G using the following operations:

- Creation of a new vertex v with label i;
- Disjoint union of two labelled graphs G and H;
- Joining by an edge each vertex with label i to each vertex with label j;
- Renaming label i to j.

Every graph can be defined by an algebraic expression using these four operations and such an expression is a k-expression if it uses k different labels. As shown by Courcelle and Olariu [47], every graph of bounded tree-width has bounded clique-width but there are graphs of bounded clique-width having unbounded tree-width (for example, complete graphs). Therefore, clique-width can be viewed as a more general concept than tree-width. An important class of graphs having bounded clique-width is that of cographs: it directly follows from the definition that cographs have clique-width at most 2. We refer to [103] for other examples of graph classes of bounded clique-width.

Similarly to tree-width, having bounded clique-width has interesting algorithmic implications. If a graph property is expressible in the more restricted MSO_1, then Courcelle et al. [49] showed that it is decidable in linear time even for graphs of bounded clique-width, assuming a k-expression of the graph is explicitly given. On the other hand, for fixed k, Oum and Seymour [152] provided a polynomial-time algorithm that given a graph G either decides G has clique-width at least $k+1$ or outputs a $2^{3k+2}-1$-expression. Therefore, a graph property expressible in MSO_1 is decidable in polynomial time for graphs of bounded clique-width. We will see several applications of this result in Chapter 4.

Discrete geometry. The usual scalar product in $\mathbb{R}^d$ is denoted by $\langle \cdot, \cdot \rangle$. A *closed half-space* in $\mathbb{R}^d$ is a set of the form $\{x \in \mathbb{R}^d : \langle a, x \rangle \geq b\}$, for some $a \in \mathbb{R}^d \setminus \{0\}$ and $b \in \mathbb{R}$. An *axis-parallel box* in $\mathbb{R}^d$ is a set of the form $[a_1, b_1] \times \cdots \times [a_d, b_d]$, with $\{a_1, \ldots, a_d, b_1, \ldots, b_d\} \subseteq \mathbb{R}^d$.

A set $C \subseteq \mathbb{R}^d$ is *convex* if for every two points x and y of C and for every $t \in [0, 1]$, the point $tx + (1 - t)y$ belongs to C. The convex hull $\mathrm{conv}(X)$ of a set $X \subseteq \mathbb{R}^d$ is the intersection of all the convex sets in $\mathbb{R}^d$ containing X. The following is a basic property of convexity in $\mathbb{R}^d$:

Theorem 1.0.10 (Radon's Theorem). *If X is a set of $d+2$ points in $\mathbb{R}^d$, there exists a partition of X into sets X_1 and X_2 such that $\mathrm{conv}(X_1) \cap \mathrm{conv}(X_2) \neq \emptyset$.*

FEEDBACK VERTEX SET

Instance: A graph G and a positive integer k.
Question: Does G contain a feedback vertex set of size at most k?

HAMILTONIAN PATH

Instance: A graph G.
Question: Does G contain a Hamiltonian path?

HAMILTONIAN CYCLE

Instance: A graph G.
Question: Does G contain a Hamiltonian cycle?

HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE

Instance: A graph $G = (V, E)$ and $e \in E$.
Question: Does G contain a Hamiltonian cycle through e?

CLIQUE COVER

Instance: A graph G and a positive integer k.
Question: Does G contain a clique cover of size at most k?

DOMINATING SET

Instance: A graph G and a positive integer k.
Question: Does G contain a dominating set of size at most k?

CONNECTED DOMINATING SET

Instance: A graph G and a positive integer k.
Question: Does G contain a connected dominating set of size at most k?

VERTEX COVER

Instance: A graph G and a positive integer k.
Question: Does G contain a vertex cover of size at most k?

CONNECTED VERTEX COVER

Instance: A graph G and a positive integer k.
Question: Does G contain a connected vertex cover of size at most k?

Table 1.1 – The decision problems considered in the thesis.

CHAPTER 2

Generalized Line Graphs

Contents

Line graphs constitute a rich and well-studied class of graphs. In this chapter, we focus on line graphs of subcubic triangle-free graphs and show that any such graph G has an independent set of size at least $\frac{3}{10}|V(G)|$, the bound being sharp. As an immediate consequence, any subcubic triangle-free graph G with n_i vertices of degree i has a matching of size at least $\frac{3}{20}n_1 + \frac{3}{10}n_2 + \frac{9}{20}n_3$.
Then we study the complexity of FEEDBACK VERTEX SET, HAMILTONIAN CYCLE and HAMILTONIAN PATH for subclasses of line graphs of subcubic triangle-free graphs: we show that these problems remain NP-hard and we provide an inapproximability result for FEEDBACK VERTEX SET.
Finally, we consider the class of k-line graphs, a generalization of line graphs, and make some observations which will be used in Chapter 3.

2.1 Introduction

The *line graph* $L(G)$ of a graph G is the graph having as vertices the edges of G, two vertices being adjacent if the corresponding edges intersect. Line graphs constitute a rich and ubiquitous class of graphs. They were introduced by Whitney [183] who showed that, with the exception of K_3 and $K_{1,3}$, any two connected graphs having isomorphic line graphs are isomorphic. Krausz [116] characterized line graphs as those graphs admitting a partition of the edges into complete subgraphs such that every vertex belongs to at most two of them. What is arguably a cornerstone in the whole theory of graph classes is Beineke's characterization in terms of forbidden induced subgraphs:

Theorem 2.1.1 (Beineke [19]). *G is a line graph if and only if it does not contain any of the graphs depicted in Figure 2.1 as an induced subgraph.*

Theorem 2.1.1 immediately implies that line graphs can be recognized in polynomial time. The trivial algorithm was improved by Roussopoulos [165] and Lehot [126], who

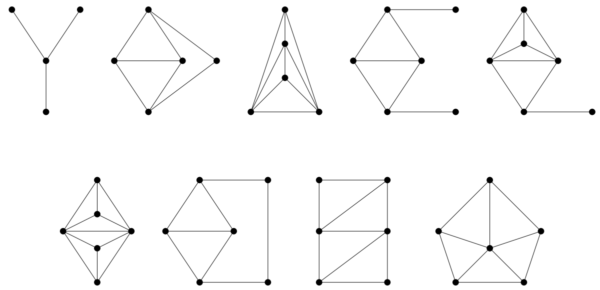


Figure 2.1: The minimal forbidden induced subgraphs for the class of line graphs.

showed that recognition is possible in linear time. To further emphasize the importance of having a forbidden subgraph characterization, let us remark the following anecdote reported in [110]. In 1969, the "Journal of Combinatorial Theory" published a paper entitled "An interval graph is a comparability graph" while one year later the same journal published another paper entitled "An interval graph is not a comparability graph". Having an induced subgraph characterization at hand would have prevented this situation since it is easy to see that *Free*$(M_1) \subseteq$ *Free*(M_2) if and only if for every graph $G \in M_2$ there is a graph $H \in M_1$ such that H is an induced subgraph of G.

In Section 2.2, we concentrate on line graphs of subcubic triangle-free graphs. Specifically, in Section 2.2.1 we provide several characterizations of this class and we observe it coincides with the class of $(K_4, \text{claw}, \text{diamond})$-free graphs. Moreover, we show that the line graphs of cubic triangle-free graphs are exactly those 4-regular graphs for which every edge belongs to exactly one K_3.

In Section 2.2.2, we consider the independence number. The famous Brooks' Theorem asserts that every connected graph G which is neither a complete graph nor an odd cycle must be $\Delta(G)$-colourable and so $\alpha(G) \geq |V(G)|/\Delta(G)$. Following this result, several authors considered the problem of finding tight lower bounds for the independence number of graphs having bounded maximum degree and not containing cliques on 3 or 4 vertices [66, 67, 93, 129, 176]. Kang et al. [104] showed that if G is a connected (K_4, claw)-free 4-regular graph on n vertices then, apart from three exceptions, $\alpha(G) \geq (8n-3)/27$. Motivated by this result, we show that if G is a $(K_4, \text{claw}, \text{diamond})$-free graph on n vertices, then $\alpha(G) \geq 3n/10$. This gives a tight bound, as can be seen by considering the following:

Example 2.1.2. Let G be the graph depicted in Figure 2.2. Since $|E(G)| = 10$ and $\alpha'(G) = 3$, we have that $L(G)$ has 10 vertices and $\alpha(L(G)) = \alpha'(G) = 3$.

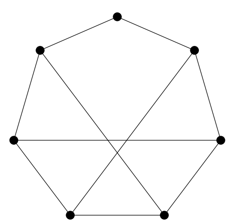


Figure 2.2: A subcubic triangle-free graph G with 10 edges and $\alpha'(G) = 3$.

The well-known Petersen's Theorem asserts that every 3-regular bridgeless graph has a perfect matching and the question of whether a graph admits a perfect matching has been deeply investigated (see [3] for a survey). On the other hand, not much is known about general lower bounds for the matching number. Biedl et al. [24] showed that any subcubic graph G has a matching of size $(|V(G)|-1)/3$ and that any cubic graph G has a matching of size $(4|V(G)|-1)/9$. Henning et al. [95] investigated lower bounds in the case of cubic graphs with odd girth. In particular, they showed that any connected cubic triangle-free graph G has a matching of size $(11|V(G)|-2)/24$. By recalling that the matchings of a graph G are in bijection with the independent sets of its line graph $L(G)$, our result on the independence number of $(K_4, \text{claw}, \text{diamond})$-free graphs directly translates into a tight lower bound for the matching number. Indeed, we show that if G is a subcubic triangle-free graph with n_i vertices of degree i, then $\alpha'(G) \geq 3n_1/20 + 3n_2/10 + 9n_3/20$.

Consider now FEEDBACK VERTEX SET: the problem of deciding, given a graph G and an integer k, whether $\tau_c(G) \leq k$. Ueno et al. [180] showed that FEEDBACK VERTEX SET can be solved in polynomial time for subcubic graphs by a reduction to a matroid parity problem (see Sections 3.4.1 and 5.3 for a proof). On the other hand, FEEDBACK VERTEX SET becomes NP-hard for graphs with maximum degree 4, even if restricted to be planar, as shown by Speckenmeyer [174, 175]. In Section 2.2.3, we strengthen this result by showing the NP-hardness for line graphs of planar cubic bipartite graphs. This is done in two steps. We first show that if G is the line graph of a cubic triangle-free graph H, then $\tau_c(G) \geq |V(G)|/3 + 1$, with equality if and only if H contains a Hamiltonian path. We then show that the well-known HAMILTONIAN PATH remains NP-hard even for planar cubic bipartite graphs. This matches the fact that even HAMILTONIAN CYCLE remains NP-hard for that class [4] and may be of independent interest. We conclude the section with an inapproximability result for FEEDBACK VERTEX SET restricted to line graphs of subcubic triangle-free graphs.

Despite the fact that Hamiltonicity in line graphs has been widely investigated, beginning with the works of Chartrand [34, 35] and Harary and Nash-Williams [86], to the best of our knowledge no result is known on HAMILTONIAN CYCLE restricted to line graphs. Concerning HAMILTONIAN PATH, Bertossi [23] showed that the problem is NP-complete for line graphs. Lai and Wei [119] strengthened this result by showing that it remains NP-hard even when restricted to line graphs of bipartite graphs. In Section 2.2.4, we prove that HAMILTONIAN CYCLE remains NP-hard for line graphs of 1-subdivisions of planar cubic bipartite graphs and for line graphs of planar cubic bipartite graphs. Finally, we show that HAMILTONIAN PATH remains NP-hard for line graphs of 1-subdivisions of planar cubic bipartite graphs, thus strengthening the result by Lai and Wei [119].

As a side remark, note that line graphs of subcubic triangle-free graphs are not necessarily 3-colourable, as Example 2.1.2 shows. In fact, 3-COLOURABILITY remains NP-hard even when restricted to line graphs of cubic triangle-free graphs [114]: if G is the line graph of a cubic triangle-free graph H, then G is 3-colourable if and only if H is of class 1.

One possible generalization of line graphs is given by the following construction: For an integer $k \geq 2$, the *k-line graph* $L_k(G)$ of a graph G is the graph having as vertices the cliques of G of size k, two vertices being adjacent if the corresponding cliques intersect in a clique of size $k-1$. This notion has been introduced independently and with different motivations by several authors [43–45]. Clearly, 2-line graphs are the usual line graphs, whereas 3-line graphs are also known as *triangle graphs*. Unlike line graphs, the class of k-line graphs with $k \geq 3$ is not hereditary, as will become evident in the next paragraph. Nevertheless, it is still of interest to find forbidden induced subgraphs for this class. In particular, it follows directly from the definition that every k-line graph is $K_{1,k+1}$-free and in Section 2.3 we will expand this list. Our main motivation for studying k-line graphs is the relation between triangle graphs and Tuza's Conjecture (Conjecture 3.3.1), which asserts that the minimum number of edges of a graph whose deletion results in a triangle-free graph is at most two times the maximum number of edge-disjoint triangles. In the case of K_4-free graphs, triangle-transversals and triangle-packings correspond to clique covers and independent sets of the triangle graph, respectively. The list of forbidden induced subgraphs obtained in Section 2.3 will then be used in Section 3.3 in the context of Tuza's Conjecture.

Two interesting classes of spanning subgraphs of k-line graphs are defined as follows. The *k-Gallai graph* $\Gamma_k(G)$ of a graph G is the graph having as vertices the cliques of G of size k, two vertices being adjacent if the corresponding cliques intersect in a clique of size $k-1$ but their union is not a clique of size $k+1$. Conversely, the *anti-k-Gallai graph* $\Delta_k(G)$ of G is the graph having as vertices the cliques of G of size k, two vertices being adjacent if the union of the corresponding cliques is a clique of size $k+1$. Clearly, $\Delta_k(G)$ is the complement of $\Gamma_k(G)$ in $L_k(G)$. 2-Gallai graphs are simply known as *Gallai graphs* and were introduced by Gallai [73] in his work on comparability graphs. Anti-2-Gallai graphs are also known as *anti-Gallai graphs* or *triangular line graphs* [9]. The classes of k-Gallai graphs and anti-k-Gallai graphs are not hereditary: for every graph G, the k-Gallai graph $\Gamma_k(\overline{G} \vee K_{k-1})$ has a component isomorphic to G and the anti-k-Gallai graph $\Delta_k(G \vee K_{k-1})$ contains G as an induced subgraph. On the other hand, Le [124] gave a Krausz-type characterization of these two classes. Anand et al. [9] showed that recognizing anti-Gallai graphs is NP-complete. In fact, they showed that even deciding whether a connected graph is the anti-Gallai graph of some K_4-free graph is NP-complete. This was recently used by Lakshmanan et al. [120] in order to show that, for every fixed $k \geq 3$, deciding whether a given graph is the k-line graph of a K_{k+1}-free graph is NP-complete. Quite surprisingly, this is in sharp contrast with the case of line graphs mentioned above. Moreover, they completed the picture about generalized anti-Gallai graphs by showing that, for every fixed $k \geq 3$, recognizing anti-k-Gallai graphs is NP-complete. On the other hand, the recognition of k-Gallai graphs remains a major open problem.

2.2 Line graphs of subcubic triangle-free graphs

2.2.1 Characterizations

The purpose of this section is to characterize line graphs of subcubic triangle-free graphs. Most of the stated results are not original and our goal is to put them under a unified framework.

Let $\mathcal{F}$ be a family of non-empty sets. The *intersection graph* $G_{\mathcal{F}}$ of $\mathcal{F}$ is the graph having as vertices the sets in $\mathcal{F}$, two vertices being adjacent if the corresponding sets intersect. The set $\bigcup_{F \in \mathcal{F}} F$ is the *ground set* of $G_{\mathcal{F}}$. If a graph G is the intersection graph of a family $\mathcal{F}$, then $\mathcal{F}$ is a *realization* of G.

Clearly, the line graph of G is the intersection graph of the family $E(G)$. An equivalent definition of line graphs is that of 2-intersection graphs: a graph is a 2-*intersection* graph if it is the intersection graph of a family of subsets of positive integers, each of size 2.

We now consider a geometric realization. A graph G of order n is a 2-*interval* graph if it is the intersection graph of a set of n unions of two disjoint intervals on the real line, i.e. each vertex corresponds to a union of two disjoint intervals $I^k = I^k_\ell \cup I^k_r$ and there is an edge between I^j and I^k if and only if $I^j \cap I^k \neq \varnothing$. Note that the two intervals corresponding to a vertex are naturally ordered into a left interval and a right interval. A 2-interval graph is (x, x)-*interval* if it has a realization in which the intervals of the ground set have length x, integral endpoints and they are open[1]. Note that the class of $(1, 1)$-interval graphs coincides with that of 2-intersection graphs (and so with line graphs). This can be seen by associating to each vertex $\{x, y\} \subseteq \mathbb{N}$ of a 2-intersection graph the union $(x-1, x) \cup (y-1, y)$ and, conversely, after an appropriate translation of the intervals, by associating to each vertex $(x, x+1) \cup (y, y+1)$ the set $\{x+1, y+1\}$ of the right endpoints of the two intervals.

Line graphs of bipartite graphs have another interesting geometric characterization: they are equivalent to gridline graphs (see, e.g., [156]). A *gridline* graph is a graph whose vertices correspond to distinct points in $\mathbb{R}^2$ and such that (x, y) and (x', y') are adjacent if and only if $x = x'$ or $y = y'$, i.e. two vertices are adjacent if and only if they are on a common horizontal or vertical line. Note that the vertices of a gridline graph may be assumed to lie in $\mathbb{N}^2$.

Before stating our characterization, we need some more definitions. If $\mathcal{F}$ is the realization of a 2-intersection graph G, an *incidence matrix* $M = [m_{ij}]$ of $\mathcal{F}$ is a $(0, 1)$-matrix whose rows correspond to the vertices of G (i.e. two-element subsets $S_i \subset \mathbb{N}$), whose columns correspond to the elements of the ground set of G (i.e. integers $j \in \bigcup S_i$) and such that $m_{ij} = 1$ if and only if $j \in S_i$. Let now A, B and C be matrices. B is a *submatrix* of A if it can be obtained by deleting some rows and columns of A. The matrix A is C-*free* if C is not a submatrix of A. Similarly, for a set S of matrices, A is S-*free* if A is M-free, for any $M \in S$. Moreover, let F be defined as follows:

$$F = \left\{ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Note that the 3×3 matrices in F are the *cycle matrices* of order 3, i.e. the edge-vertex incidence matrices of cycles of length 3.

Theorem 2.2.1. *The following are equivalent, for any graph G:*

(a) G is a $(K_4, \text{claw}, \text{diamond})$-free graph.

(b) G is the line graph of a subcubic triangle-free graph.

(c) G is a $(1, 1)$-interval graph such that there do not exist four vertices sharing the same interval and there do not exist three vertices such that any two of them share a different interval.

(d) G is a 2-intersection graph such that any incidence matrix of the realization of G is F-free.

Proof. $(a) \Rightarrow (b)$: By Theorem 2.1.1, we have $G = L(H)$, for some graph H. Let H' be the graph obtained from H by replacing each component isomorphic to K_3 with a claw. Clearly, $G = L(H) = L(H')$. Since G is K_4-free, H' is subcubic. Suppose now H' contains a triangle T. By construction, there exists a vertex $v \in V(H') \setminus V(T)$ which is adjacent to a vertex of T and so there exists an induced diamond in $L(H')$, a contradiction. Therefore, G is the line graph of a subcubic triangle-free graph.

[1] We refer the reader to [74] for a survey on 2-interval graphs.

$(b) \Rightarrow (c)$: Let $G = L(H)$, for a subcubic triangle-free graph H with $V(H) = \{v_1, \dots, v_n\}$. By associating to each vertex $e_{ij} = v_i v_j$ of $L(H)$, with $i < j$, the union $I_{ij} = (2i, 2i+1) \cup (2j, 2j+1)$, it is easy to see that G is $(1,1)$-interval. Since G is subcubic, there do not exist four vertices sharing the same interval. Moreover, since G is triangle-free, there do not exist three vertices such that any two of them share a different interval.

$(c) \Rightarrow (d)$: By appropriately translating the intervals and associating to each vertex $(i, i+1) \cup (j, j+1)$ of G the set $\{i+1, j+1\}$ of the right endpoints of the two intervals, it is easy to see that G is 2-intersection. Since there do not exist four vertices sharing the same interval and there do not exist three vertices such that any two of them share a different interval, any incidence matrix of the realization of G is F-free.

$(d) \Rightarrow (a)$: Let M be an F-free incidence matrix of the realization of G and, for any $v \in V(G)$, let S_v be the two-element set corresponding to v. It is easy to see that no 2-intersection graph contains an induced claw. Suppose now G contains a copy H of K_4. For $v \in V(H)$, the sets S_v are pairwise intersecting and $|S_v| = 2$. But then $\bigcap_{v \in V(H)} S_v \neq \varnothing$, contradicting the fact that M is F-free. Finally, if G contains an induced diamond H, there exists a triangle $T \subseteq H$ such that $\bigcap_{v \in V(T)} S_v = \varnothing$. This means that a cycle matrix of order 3 is a submatrix of M, a contradiction. □

The observation that a graph is $(\text{claw}, \text{diamond})$-free if and only if it is the line graph of a triangle-free graph is probably due to Harary and Holzmann [85]. Let us now recall the following well-known characterization of diamond-free graphs:

Theorem 2.2.2 (Folklore). *A graph G is diamond-free if and only if it can be obtained as the edge-disjoint union of complete subgraphs of order at least 3 plus possibly some edges which are not in any triangle and isolated vertices such that each triangle of G is contained in one of the complete subgraphs.*

Proof. The "if" part is trivial. Conversely, we claim that if G' is an inclusion-wise maximal complete subgraph of the diamond-free graph G, then each triangle of G is either contained in G' or edge-disjoint from it. Indeed, if G is triangle-free, the statement is vacuously true. Therefore, suppose there exists a triangle uvw having exactly one edge in $E(G')$, say $uv \in E(G')$ and $w \notin V(G')$. By maximality, $|V(G')| \geq 3$. Moreover, every $w' \in V(G') \setminus \{u, v\}$ is adjacent to both u and v and so, since G is diamond-free, every $w' \in V(G') \setminus \{u, v\}$ is adjacent to w as well. Therefore, $V(G') \cup \{w\}$ is a clique, contradicting the maximality of G'.

We now claim that the graph G'' obtained from G by deleting the edges of G' is diamond-free. Indeed, since G is diamond-free, a diamond $K_4 - e$ can arise in G'' only if $e \in E(G')$ and the two remaining vertices of the diamond are not in $V(G')$. This means there exists a triangle with exactly one edge in G', contradicting the paragraph above.

Since G'' is diamond-free, an easy induction shows that G can be obtained as the edge-disjoint union of complete subgraphs of order at least 3 plus possibly some edges which are not in any triangle and isolated vertices such that each triangle of G is contained in one of the complete subgraphs. □

Corollary 2.2.3. *Let G be a $(K_4, \text{claw}, \text{diamond})$-free graph. For $v \in V(G)$, the possible subgraphs induced by $N[v]$ are exactly those depicted in Figure 2.3.*

Proof. By Theorem 2.2.2, G can be obtained as the edge-disjoint union of triangles plus possibly some edges which are not in any triangle. Moreover, each $v \in V(G)$ belongs to at most two edge-disjoint triangles and $d(v) \leq 4$. If $v \in V(G)$ belongs to no triangle, then $d(v) \leq 2$. If it belongs to exactly one triangle, then $G[N(v)]$ is isomorphic to either K_2 or $K_2 + K_1$. Finally, if it belongs to exactly two triangles, then $G[N(v)] = 2K_2$. □

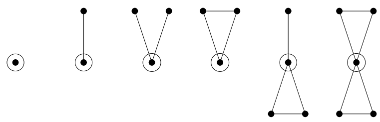


Figure 2.3: The possible subgraphs induced by the closed neighbourhood of a vertex of a $(K_4, \text{claw}, \text{diamond})$-free graph.

A graph G is *locally linear* if the subgraph induced by $N_G(v)$ is 1-regular, for every $v \in V(G)$. Clearly, a locally linear graph has no odd degree vertices. Fronček [68] observed that a graph is locally linear if and only if every edge belongs to exactly one triangle. Using this, he showed that there exists a k-regular locally linear graph, for any even k. This can be seen by considering the following recursive construction. K_3 is clearly 2-regular and locally linear. A $(k + 2)$-regular locally linear graph is then obtained by taking the cartesian product of a k-regular locally linear graph with K_3. Combining Theorem 2.2.1 and Corollary 2.2.3, we immediately get the following:

Theorem 2.2.4. *The following are equivalent, for any 4-regular graph G:*

(a) G is a $(K_4, \text{claw}, \text{diamond})$-free graph.

(b) G is the line graph of a cubic triangle-free graph.

(c) G is a locally linear graph.

(d) G is such that each $e \in E(G)$ belongs to exactly one triangle.

2.2.2 Independence number

For a graph G, we denote by $i(G)$ the *independence ratio* $\alpha(G)/|V(G)|$ of G. The well-known Brooks' Theorem asserts that every connected graph G which is neither a complete graph nor an odd cycle must be $\Delta(G)$-colourable and so $i(G) \geq 1/\Delta(G)$. If G is K_k-free, Brooks' bound can be strengthen to $i(G) \geq 2/(\Delta(G) + k)$, as shown by Fajtlowicz [62], who also provided some cases for which equality holds [61]. Staton [176] improved both Brooks' and Fajtlowicz's results in the case of subcubic triangle-free graphs by showing that $i(G) \geq 5/14$, for any such graph G. Moreover, Heckman [92] showed that there are exactly two connected subcubic triangle-free graphs with independence ratio equal to $5/14$. Fraughnaugh and Locke [66] proved that every connected subcubic triangle-free graph G on n vertices has $\alpha(G) \geq 11n/30 - 2/15$. Fraughnaugh Jones [67] improved Fajtlowicz's result in the case of triangle-free graphs with maximum degree 4 by showing that $i(G) \geq 4/13$, for any such graph G. Locke and Lou [129] showed that if G is a connected K_4-free 4-regular graph with n vertices, then $\alpha(G) \geq (7n - 4)/26$. Kang et al. [104] studied the independence number of 4-regular (K_4, claw)-free graphs. In particular, they proved the following:

Theorem 2.2.5 (Kang et al. [104]). *If G is a connected $(K_4, claw)$-free 4-regular graph on n vertices then, apart from three exceptions, $\alpha(G) \geq (8n-3)/27$. Moreover, equality holds only if G is the line graph of a cubic graph.*

If we further assume the graph G in Theorem 2.2.5 to be 2-connected, then $\alpha(G) = \lfloor n/3 \rfloor$ [105]. Motivated by these results, we seek a tight lower bound for the independence number of $(K_4, \text{claw}, \text{diamond})$-free graphs. It appears that forbidding diamonds is stronger than relaxing the regularity assumption, in the sense that the independence ratio of $(K_4, \text{claw}, \text{diamond})$-free graphs is strictly bigger than the independence ratio of (K_4, claw)-free 4-regular graphs:

Theorem 2.2.6. *If G is a $(K_4, claw, diamond)$-free graph on n vertices, then $\alpha(G) \geq \frac{3}{10}n$. Moreover, the bound is tight, as shown by the graph in Example 2.1.2.*

Recall that there is a bijection between the matchings of a graph G and the independent sets of its line graph $L(G)$. Biedl et al. [24] showed that every subcubic graph G has a matching of size $(|V(G)|-1)/3$ and that every cubic graph G has a matching of size $(4|V(G)|-1)/9$. Moreover, O and West [149] characterized the cubic graphs G with $\alpha'(G) = (4|V(G)|-1)/9$. Henning et al. [95] showed that $\alpha'(G) \geq (11|V(G)|-2)/24$, for any connected cubic triangle-free graph G, and characterized the graphs attaining equality. If we consider subcubic triangle-free graphs then, as a direct consequence of Theorem 2.2.6, we obtain the following tight lower bound:

Corollary 2.2.7. *If G is a subcubic triangle-free graph with n_i vertices of degree i, then $\alpha'(G) \geq \frac{3}{20}n_1 + \frac{3}{10}n_2 + \frac{9}{20}n_3$.*

Note that Corollary 2.2.7 is similar in nature to a series of results by Haxell and Scott [90]: they completely characterized the set of 3-tuples of real coefficients (α, β, γ) for which there exists a constant K such that $\alpha'(G) \geq \alpha n_1 + \beta n_2 + \gamma n_3 - K$ for every connected subcubic graph G.

As a side remark, note that INDEPENDENT SET is polynomial for claw-free graphs: Minty [139] reduced this problem to a matching problem in an auxiliary graph and this can be solved by Edmonds' algorithm (see, e.g., [28, 182]). Another proof was given, independently, by Sbihi [167].

By further exploiting the relation between matchings of G and independent sets of $L(G)$, we can obtain the following simple lemma, already implicitly stated in [105], and which will be used in the proof of Theorem 2.2.6:

Lemma 2.2.8. *If G is a 2-connected 4-regular $(K_4, claw, diamond)$-free graph, then $\alpha(G) = \frac{|V(G)|}{3}$.*

Proof. By Theorem 2.2.4, G is the line graph of a cubic triangle-free graph H. Moreover, since G is 2-connected, H is bridgeless. Therefore, by Petersen's Theorem, H has a perfect matching and $\alpha(G) = \alpha'(H) = \frac{|V(H)|}{2} = \frac{|V(G)|}{3}$. □

We can finally proceed to the proof of Theorem 2.2.6:

Proof of Theorem 2.2.6. Without loss of generality, we may assume G to be connected. Recall that, for $v \in V(G)$, the possible subgraphs induced by $N[v]$ are exactly those depicted in Figure 2.3. We proceed by induction on the number of vertices and we will repeatedly make use of the following simple claim:

Claim 1. *If there exists an independent set S of G such that $|S| \geq \frac{3}{10}|N_G[S]|$, then $\alpha(G) \geq \frac{3}{10}n$.*

Indeed, suppose such an S exists and let $G' = G - N_G[S]$. By the induction hypothesis, there exists an independent set I' of G' of size at least $\frac{3}{10}(n - |N_G[S]|)$. But then $I' \cup S$ is an independent set of G of size at least $\frac{3}{10}(n - |N_G[S]|) + |S| \geq \frac{3}{10}n$. ■

By Brooks' Theorem, if $\Delta(G) \leq 3$, then G is 3-colourable and so $\alpha(G) \geq \frac{n}{3}$. Therefore, we may assume there exists a 4-vertex $v \in V(G)$ and let $N_G(v) = \{v_1, v_2, v_3, v_4\}$. We have that $G[N(v)] = 2K_2$ and, without loss of generality, $E(G[N(v)]) = \{v_1v_2, v_3v_4\}$. We distinguish the following cases:

Case 1: There exists a vertex v_i, with $i \in \{1, 2\}$, such that $d_G(v_i) = 2$ (see Figure 2.4(a)). Then $S = \{v_i\}$ is clearly an independent set with $|N_G[S]| = 3$.

Case 2: $d_G(v_1) = 3$ and $d_G(v_2) = 3$ (see Figure 2.4(b)).
Let $N_G(v_1) \setminus \{v, v_2\} = \{u_1\}$ and $N_G(v_2) \setminus \{v, v_1\} = \{u_2\}$. Clearly, $u_1 \neq u_2$. Let G' be the graph obtained from G by deleting the set of vertices $\{v, v_1, v_2\}$ and by adding, if necessary (i.e. without creating multiple edges), the edge u_1u_2. Suppose first such a graph is $(K_4, \text{claw}, \text{diamond})$-free. By the induction hypothesis, there exists an independent set I of G' of size at least $\frac{3}{10}(n-3)$. Clearly, at most one of u_1 and u_2 belongs to I. But then, if u_1 or u_2 is in I, we have that $I \cup \{v_2\}$ or $I \cup \{v_1\}$ is an independent set of G, respectively, of size at least $\frac{3}{10}(n-3) + 1 > \frac{3}{10}n$.

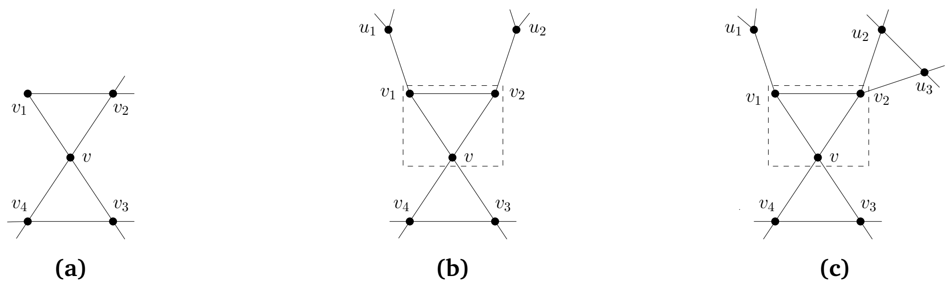


Figure 2.4: Three cases in the proof of Theorem 2.2.6.

It remains to consider the case of G' not being $(K_4, \text{claw}, \text{diamond})$-free. Note that, in this case, $u_1u_2 \notin E(G)$. Since $G - \{v, v_1, v_2\}$ is diamond-free, then G' is K_4-free. Moreover, since $G[N(u_i)]$ is isomorphic to either K_1, $2K_1$ or $K_2 + K_1$, then G' is claw-free. Therefore, only diamonds containing u_1u_2 may arise in G'. In order for this to happen, it must be that one of $G[N(u_1)]$ and $G[N(u_2)]$ is isomorphic to $K_2 + K_1$. Without loss of generality, suppose that w_1 and w_1' are the remaining neighbours of u_1. But then u_2 must be adjacent to either w_1 or w_1'. Therefore, since $S = \{v, u_1, u_2\}$ is an independent set of G and $|N_G[S]| \leq 10$, we conclude by Claim 1.

Case 3: $d_G(v_1) = 3$ and $d_G(v_2) = 4$ (see Figure 2.4(c)).
Let $N_G(v_1) \setminus \{v, v_2\} = \{u_1\}$ and $N_G(v_2) \setminus \{v, v_1\} = \{u_2, u_3\}$. Note that $\{u_1\} \cap \{u_2, u_3\} = \varnothing$ and

$u_2u_3 \in E(G)$. The reasoning is similar to the one adopted in the previous case. Let G_1 be the graph obtained from G by deleting the set of vertices $\{v, v_1, v_2\}$ and by adding, if necessary (i.e. without creating multiple edges), the edges u_1u_2 and u_1u_3. Suppose first such a graph is $(K_4, \text{claw}, \text{diamond})$-free. By the induction hypothesis, there exists an independent set I of G_1 of size at least $\frac{3}{10}(n-3)$. Clearly, at most one of u_1, u_2 and u_3 belongs to I. But then, if u_1 or a vertex in $\{u_2, u_3\}$ belongs to I, we have that $I \cup \{v_2\}$ or $I \cup \{v_1\}$ is an independent set of G, respectively, of size at least $\frac{3}{10}(n-3) + 1 > \frac{3}{10}n$.

It remains to consider the case of G_1 not being $(K_4, \text{claw}, \text{diamond})$-free. Since both G and $G - \{v, v_1, v_2\}$ are diamond-free, then G_1 is K_4-free. Moreover, since $G[N(u_1)]$ is isomorphic to either K_1, $2K_1$ or $K_2 + K_1$ and $G[N(u_2)]$ and $G[N(u_3)]$ are isomorphic to either K_2, $K_2 + K_1$ or $2K_2$, then G_1 is claw-free. Therefore, only diamonds containing newly added edges may arise in G_1. It is easy to see that this implies there exists $w \in V(G) \setminus \{u_1, u_2, u_3\}$ such that $u_1w \in E(G)$ and either $u_2w \in E(G)$ or $u_3w \in E(G)$. By symmetry, we may assume $u_2w \in E(G)$. Suppose $w \in \{v_3, v_4\}$, say without loss of generality $w = v_3$. Then $u_1u_2 \in E(G)$ and $S = \{u_1, v_2, v_4\}$ is an independent set of G with $|N_G[S]| \leq 10$.

Therefore, we may assume $w \notin \{v_3, v_4\}$. Consider now the 4-vertex v_2. We have that v_1 is a 3-vertex and v is a 4-vertex. Let G_2 be the graph obtained from G by deleting the set of vertices $\{v, v_1, v_2\}$ and by adding, if necessary (i.e. without creating multiple edges), the edges u_1v_3 and u_1v_4. By the reasoning above, if G_2 is $(K_4, \text{claw}, \text{diamond})$-free we are done. Otherwise, as we have already seen, the only possibility is that a diamond arises in G_2. But this means there exists $w' \in V(G) \setminus \{u_1, v_3, v_4\}$ such that $u_1w' \in E(G)$ and either $v_3w' \in E(G)$ or $v_4w' \in E(G)$, say without loss of generality $v_4w' \in E(G)$. Since G is $(K_4, \text{claw}, \text{diamond})$-free, it is easy to see that $w' \notin \{u_2, u_3\}$. Therefore, suppose first $w = w'$. Then either $u_1v_4 \in E(G)$, $u_1u_2 \in E(G)$ or $u_2v_4 \in E(G)$. If $u_1v_4 \in E(G)$, then $S = \{v, u_1, u_2\}$ is an independent set of G with $|N_G[S]| \leq 10$. If $u_1u_2 \in E(G)$, then $S = \{u_2, v_1, v_4\}$ is an independent set of G with $|N_G[S]| \leq 10$. Finally, if $u_2v_4 \in E(G)$, then $S = \{u_1, v_2, v_4\}$ is an independent set of G with $|N_G[S]| \leq 10$.

Therefore, we may further assume $w \neq w'$. But then $ww' \in E(G)$. We now claim we may assume that u_2 and v_4 have both degree 4. Indeed, if $d_G(u_2) = 3$, then $S = \{v, u_1, u_2\}$ is an independent set of G with $|N_G[S]| \leq 10$. Moreover, if $d_G(v_4) = 3$, then $S = \{u_1, v_2, v_4\}$ is an independent set of G with $|N_G[S]| \leq 10$. This means that w' and v_4 have a common neighbour and the same holds for w and u_2. Suppose now the common neighbour of w' and v_4 is u_3. Then $S = \{u_1, v_2, v_4\}$ is an independent set of G with $|N_G[S]| \leq 10$. Similarly, if the common neighbour of w and u_2 is v_3, then $S = \{v, u_1, u_2\}$ is an independent set of G with $|N_G[S]| \leq 10$. Therefore, we may assume that the common neighbour a of w' and v_4 and the common neighbour b of w and u_2 are such that $\{a, b\} \cap \{v, v_1, v_2, v_3, v_4, u_1, u_2, u_3, w, w'\} = \varnothing$. Moreover, since G is $(K_4, \text{claw}, \text{diamond})$-free and $ww' \in E(G)$, we have $a \neq b$.

Consider now the graph G_3 obtained from G by deleting the set $\{v, v_1, v_2, v_3, v_4, u_1, u_2, u_3, w, w'\}$ and by adding, if necessary (i.e. without creating multiple edges), the edge ab. Suppose first such a graph is $(K_4, \text{claw}, \text{diamond})$-free. By the induction hypothesis, there exists an independent set I of G_3 of size at least $\frac{3}{10}(n-10)$. Clearly, at most one of a and b belongs to I. But then, if a or b is in I, we have that $I \cup \{u_1, u_2, v\}$ or $I \cup \{v_2, v_4, u_1\}$ is an independent set of G, respectively, of size at least $\frac{3}{10}(n-10) + 3 \geq \frac{3}{10}n$.

It remains to consider the case of G_3 not being $(K_4, \text{claw}, \text{diamond})$-free. Note that, in this case, $ab \notin E(G)$. Clearly, G_3 is K_4-free. Moreover, since $G[N(a)]$ and $G[N(b)]$ are isomorphic to either K_2, $K_2 + K_1$ or $2K_2$, then G_3 is claw-free. Therefore, only diamonds containing ab may arise in G_3. In order for this to happen, it must be that there exist two distinct vertices

$a_1 \in V(G_3)$ and $a_2 \in V(G_3)$, with $\{a_1, a_2\} \cap \{a, b\} = \varnothing$, such that $\{a_1 a, a_1 b\} \subseteq E(G)$ and a_2 is a common neighbour of either $\{a_1, a\}$ or $\{a_1, b\}$. By symmetry, we may assume that a_2 is a common neighbour of $\{a_1, a\}$.

Now we repeat once again the reasoning of the previous paragraphs. Let G_4 be the graph obtained from G by deleting the set of vertices $\{v, v_1, v_2, v_4, u_1, u_2, w, w', a, b\}$ and by adding, if necessary (i.e. without creating multiple edges), the edge $u_3 v_3$. Suppose first such a graph is $(K_4, \text{claw}, \text{diamond})$-free. By the induction hypothesis, there exists an independent set I of G_4 of size at least $\frac{3}{10}(n-10)$. Clearly, at most one of u_3 and v_3 belongs to I. But then, if u_3 or v_3 is in I, we have that $I \cup \{v_1, v_4, w\}$ or $I \cup \{v_1, u_2, w'\}$ is an independent set of G, respectively, of size at least $\frac{3}{10}(n-10) + 3 \geq \frac{3}{10}n$.

It remains to consider the case of G_4 not being $(K_4, \text{claw}, \text{diamond})$-free. Note that, in this case, $u_3 v_3 \notin E(G)$. Moreover, we have that G_4 is K_4-free and claw-free. Therefore, only diamonds containing $u_3 v_3$ may arise in G_4. In order for this to happen, it must be that there exist $v_5 \in V(G_4)$ and $v_6 \in V(G_4)$ such that $\{v_5 v_3, v_5 u_3\} \subseteq E(G)$ and v_6 is a common neighbour of either $\{v_3, v_5\}$ or $\{u_3, v_5\}$. In the following, we assume that v_6 is a common neighbour of $\{v_3, v_5\}$. The remaining case can be treated similarly and we leave its verification to the interested reader. Clearly, $v_5 \neq a_2$ and $\{v_5, v_6\} \cap \{a_1\} = \varnothing$. Moreover, if $v_6 = a_2$, then $S = \{v, u_1, u_2, a, v_5\}$ is an independent set of G with $|N_G[S]| \leq 16$ and we conclude by Claim 1. Therefore, we may assume $\{v_5, v_6\} \cap \{a_1, a_2\} = \varnothing$. We now claim that v_5 and a_1 both have degree 4. Indeed, if $d_G(v_5) = 3$, then $S = \{v, u_1, u_2, a, v_5\}$ is an independent set of G with $|N_G[S]| \leq 16$. Moreover, if $d_G(a_1) = 3$, then $S = \{a_1, v_1, v_3, u_2, w'\}$ is an independent set of G with $|N_G[S]| \leq 16$.

Therefore, v_5 and a_1 both have degree 4. This means there exist vertices x and y such that x is the common neighbour of $\{a_1, b\}$ and y is the common neighbour of $\{v_5, u_3\}$. If $x = v_6$, then $S = \{a_1, v_1, v_3, u_2, w'\}$ is an independent set of G with $|N_G[S]| \leq 16$. Moreover, if $y = a_2$, then $S = \{a_2, v_1, v_3, u_2, w'\}$ is an independent set of G with $|N_G[S]| \leq 16$. Therefore, $\{x, y\} \cap \{v, v_1, v_2, v_3, v_4, v_5, v_6, u_1, u_2, u_3, w, w', a, b, a_1, a_2\} = \varnothing$. If $x = y$, then $S = \{x, a_2, v_1, u_2, v_3, w'\}$ is an independent set of G with $|N_G[S]| \leq 19$. Finally, if $x \neq y$, then $S = \{u_2, v_1, v_3, w', a_1, y\}$ is an independent set of G with $|N_G[S]| \leq 20$.

Case 4: $d_G(v_1) = 4$ and $d_G(v_2) = 4$.
By the previous cases, we may assume that each neighbour of a 4-vertex is a 4-vertex. Therefore, by connectedness, we have that G is 4-regular. If G is also 2-connected, we conclude by Lemma 2.2.8.

It remains to consider the case of G having a cut-vertex v. Let G_1 and G_2 be two non-trivial induced subgraphs of G such that $G = G_1 \cup G_2$ and $V(G_1) \cap V(G_2) = \{v\}$. Suppose that $|V(G_1)| = 10k_1 + a$ and $|V(G_2)| = 10k_2 + b$, with $1 \leq a, b \leq 10$. Therefore, $\frac{3}{10}|V(G)| = 3k_1 + 3k_2 + \frac{3}{10}(a + b - 1)$. By the induction hypothesis, there exist an independent set I_1 of $G_1 - v$ of size at least $\lceil \frac{3}{10}(|V(G_1)| - 1) \rceil = \lceil 3k_1 + \frac{3}{10}(a-1) \rceil$ and an independent set I_2 of $G_2 - v$ of size at least $\lceil \frac{3}{10}(|V(G_2)| - 1) \rceil = \lceil 3k_2 + \frac{3}{10}(b-1) \rceil$. Clearly, $I_1 \cup I_2$ is an independent set of G. It is enough to distinguish the following cases:

Subcase 4.1: $1 \leq a - 1 \leq 3$.
We have $\lceil 3k_1 + \frac{3}{10}(a-1) \rceil = 3k_1 + 1$. If $1 \leq b - 1 \leq 3$, then

$$|I_1 \cup I_2| \geq \left\lceil 3k_1 + \frac{3}{10}(a-1) \right\rceil + \left\lceil 3k_2 + \frac{3}{10}(b-1) \right\rceil = 3k_1 + 3k_2 + 2 \geq 3k_1 + 3k_2 + \frac{3}{10}(a+b-1),$$

except when $a - 1 = b - 1 = 3$. In that case, let I_1' be an independent set of G_1 of size at least

$\lceil \frac{3}{10}|V(G_1)|\rceil = 3k_1 + \lceil \frac{3}{10}a\rceil$ and I_2' be an independent set of G_2 of size at least $\lceil \frac{3}{10}|V(G_2)|\rceil = 3k_2 + \lceil \frac{3}{10}b\rceil$. By eventually removing v from $I_1' \cup I_2'$, we get an independent set of G of size at least

$$\left(3k_1 + \left\lceil \frac{3}{10}a \right\rceil + 3k_2 + \left\lceil \frac{3}{10}b \right\rceil\right) - 1 \geq 3k_1 + 3k_2 + 3 \geq 3k_1 + 3k_2 + \frac{3}{10}(a+b-1).$$

If $4 \leq b-1 \leq 6$, then

$$|I_1 \cup I_2| \geq \left\lceil 3k_1 + \frac{3}{10}(a-1) \right\rceil + \left\lceil 3k_2 + \frac{3}{10}(b-1) \right\rceil = 3k_1 + 3k_2 + 3 \geq 3k_1 + 3k_2 + \frac{3}{10}(a+b-1).$$

If $7 \leq b-1 \leq 9$, then

$$|I_1 \cup I_2| \geq \left\lceil 3k_1 + \frac{3}{10}(a-1) \right\rceil + \left\lceil 3k_2 + \frac{3}{10}(b-1) \right\rceil = 3k_1 + 3k_2 + 4 \geq 3k_1 + 3k_2 + \frac{3}{10}(a+b-1).$$

Subcase 4.2: $4 \leq a-1 \leq 6$.
We have $\lceil 3k_1 + \frac{3}{10}(a-1)\rceil = 3k_1 + 2$. If $4 \leq b-1 \leq 6$, then

$$|I_1 \cup I_2| \geq \left\lceil 3k_1 + \frac{3}{10}(a-1) \right\rceil + \left\lceil 3k_2 + \frac{3}{10}(b-1) \right\rceil = 3k_1 + 3k_2 + 4 \geq 3k_1 + 3k_2 + \frac{3}{10}(a+b-1).$$

If $7 \leq b-1 \leq 9$, then

$$|I_1 \cup I_2| \geq \left\lceil 3k_1 + \frac{3}{10}(a-1) \right\rceil + \left\lceil 3k_2 + \frac{3}{10}(b-1) \right\rceil = 3k_1 + 3k_2 + 5 \geq 3k_1 + 3k_2 + \frac{3}{10}(a+b-1).$$

Subcase 4.3: $7 \leq a-1 \leq 9$.
We have $\lceil 3k_1 + \frac{3}{10}(a-1)\rceil = 3k_1 + 3$. If $7 \leq b-1 \leq 9$, then

$$|I_1 \cup I_2| \geq \left\lceil 3k_1 + \frac{3}{10}(a-1) \right\rceil + \left\lceil 3k_2 + \frac{3}{10}(b-1) \right\rceil = 3k_1 + 3k_2 + 6 \geq 3k_1 + 3k_2 + \frac{3}{10}(a+b-1).$$

Subcase 4.4: $a-1 = 0$.
Let I_1' be an independent set of G_1 of size at least $\lceil \frac{3}{10}|V(G_1)|\rceil = 3k_1 + 1$ and I_2' be an independent set of G_2 of size at least $\lceil \frac{3}{10}|V(G_2)|\rceil = 3k_2 + \lceil \frac{3}{10}b\rceil$. By eventually removing v from $I_1' \cup I_2'$, we get an independent set of G of size at least

$$\left(3k_1 + 1 + 3k_2 + \left\lceil \frac{3}{10}b \right\rceil\right) - 1 \geq 3k_1 + 3k_2 + \frac{3}{10}b = \frac{3}{10}|V(G)|.$$

This concludes the proof. □

We leave as an open problem the characterization of graphs attaining equality in Theorem 2.2.6:

Problem 2.2.9. Characterize the $(K_4, \text{claw}, \text{diamond})$-free graphs G such that $\alpha(G) = \frac{3}{10}|V(G)|$.

2.2.3 Complexity of Feedback Vertex Set

Ueno et al. [180] showed that FEEDBACK VERTEX SET can be solved in polynomial time for subcubic graphs. In this section, strengthening a result by Speckenmeyer [175], we show the NP-hardness for line graphs of planar cubic bipartite graphs, a subclass of 4-regular planar graphs.

We begin by considering (vertex) triangle-transversals of $(K_4, \text{claw}, \text{diamond})$-free graphs or, equivalently, line graphs of subcubic triangle-free graphs (Theorem 2.2.1). Given a subset of vertices S of a graph G, an S-*cover* of G is a subset of $E(G)$ covering each vertex in S. Clearly, $V(G)$-covers are the usual edge covers of G. Recall that $d_k(G) = \{v \in V(G) : d_G(v) = k\}$.

Lemma 2.2.10. *For a subcubic triangle-free graph H, there is a bijection between the $d_3(H)$-covers of H and the triangle-transversals of $L(H)$.*

Proof. Using the bijection between edges of H and vertices of $L(H)$, the assertion follows from the fact that the triangles of $L(H)$ are in bijection with the cubic vertices of H. □

Given a graph G, let us now introduce the graph $T'(G)$ defined as follows: the vertices of $T'(G)$ are the triangles of G, two vertices being adjacent if the corresponding triangles share a vertex (in Section 2.3, we will study the "edge analogue" of $T'(G)$). Note that if G is the line graph of a subcubic triangle-free graph H, then $T'(G)$ is nothing but the subgraph of H induced by the cubic vertices. Lemma 2.2.10 immediately tells us that a triangle-transversal of a $(K_4, \text{claw}, \text{diamond})$-free graph G essentially corresponds to an edge cover of $T'(G)$:

Corollary 2.2.11. *If G is a $(K_4, \text{claw}, \text{diamond})$-free graph, then $\tau_\Delta(G) = \beta'(T'(G) - S) + |S|$, where S is the set of isolated vertices of $T'(G)$.*

Another consequence is that finding a minimum-size triangle-transversal of a $(K_4, \text{claw}, \text{diamond})$-free graph is polynomial:

Corollary 2.2.12. *It is possible to find a minimum-size triangle-transversal of a $(K_4, \text{claw}, \text{diamond})$-free graph with n vertices in $O(n^3)$ time.*

Proof. Let G be a $(K_4, \text{claw}, \text{diamond})$-free graph of order n and size m. By Corollary 2.2.11, it is enough to find a minimum-size edge cover of $T'(G) - S$, where S denotes the set of isolated vertices of $T'(G)$. This can be done by Edmonds' maximum matching algorithm in $O(n^3)$ time. □

The following lower bound on the size of a feedback vertex set will be crucial for our NP-hardness proof of FEEDBACK VERTEX SET restricted to line graphs.

Lemma 2.2.13. *Let H be a subcubic triangle-free graph with $\delta(H) \geq 2$ and let $G = L(H)$ be its line graph. We have $\tau_c(G) \geq \frac{|d_3(H)|}{2} + 1$, with equality if and only if H contains a Hamiltonian path.*

Proof. By Lemma 2.2.10, there exists a bijection between the triangle-transversals of G and the $d_3(H)$-covers of H. Therefore, we have $\tau_c(G) \geq \tau_\Delta(G) \geq \frac{|d_3(H)|}{2}$ and the last inequality is an equality if and only if the induced subgraph $H[d_3(H)]$ contains a 1-factor.

Suppose, to the contrary, that $\tau_c(G) = \frac{|d_3(H)|}{2}$. This means there exists a minimum triangle-transversal T of G (of size $\frac{|d_3(H)|}{2}$) which is also a feedback vertex set of G. Moreover, T corresponds to a 1-factor H' of $H[d_3(H)]$ and so, since $\delta(H) \geq 2$, we have that $H - E(H')$ is 2-regular. But then there exists a cycle in $G - T$, a contradiction. This implies that $\tau_c(G) \geq \frac{|d_3(H)|}{2} + 1$.

Suppose now equality holds, i.e. $\tau_c(G) = \frac{|d_3(H)|}{2} + 1$, and let T be a minimum feedback vertex set of G. Moreover, let T_Δ be a triangle-transversal of G having minimum size among those contained in T. Clearly, we have $\frac{|d_3(H)|}{2} \leq |T_\Delta| \leq \frac{|d_3(H)|}{2} + 1$. If $|T_\Delta| = \frac{|d_3(H)|}{2}$, then T_Δ corresponds to a 1-factor H' of $H[d_3(H)]$ and so $H - E(H')$ is a 2-factor F with $p \geq 1$ components. Since the components of F give rise to p vertex-disjoint cycles in $G - T_\Delta$, then $\tau_c(G) = \frac{|d_3(H)|}{2} + 1$ implies that $p = 1$ and so $F \subseteq H$ is a Hamiltonian cycle.

Suppose now that T_Δ has size $\frac{|d_3(H)|}{2} + 1$ (in particular, $T = T_\Delta$), and let T' be the corresponding $d_3(H)$-cover of H. We have that T' contains at most two edges with an endpoint not in $d_3(H)$. If T' contains only edges of $H[d_3(H)]$ then, by the minimality of T_Δ, we have that T' consists of a maximum matching M of size $\frac{|d_3(H)|}{2} - 1$ together with two edges, each one covering exactly one vertex uncovered by M. If T' contains exactly one edge e with an endpoint not in $d_3(H)$, then $T' \setminus \{e\}$ consists of a maximum matching M of $H[d_3(H)]$ of size $\frac{|d_3(H)|}{2} - 1$ and an edge of $H[d_3(H)]$ covering the vertex uncovered by $M \cup \{e\}$. Finally, if T' contains exactly two edges e_1 and e_2 with an endpoint not in $d_3(H)$, then e_1 and e_2 cover distinct cubic vertices. Moreover, $T' \setminus \{e_1, e_2\}$ consists of a maximum matching of $H[d_3(H)]$ of size $\frac{|d_3(H)|}{2} - 1$. It is easy to see that, in all the three cases above, the graph $H - T'$ either contains a single isolated vertex and all the remaining vertices have degree 2, or it contains exactly two 1-degree vertices with all the remaining ones having degree 2. On the other hand, $H - T'$ is a forest, or else there would be a cycle in $G - T$. This implies that all the vertices of $H - T'$ have degree 2, except two of them having degree 1, and that $H - T'$ is a path. Therefore, H contains a Hamiltonian path.

Conversely, suppose that H contains a Hamiltonian path P. The number of edges in $E(H) \setminus E(P)$ is $|d_2(H)| + \frac{3}{2}|d_3(H)| - (|d_2(H)| + |d_3(H)| - 1) = \frac{|d_3(H)|}{2} + 1$ and these edges constitute a $d_3(H)$-cover of H. If T is the corresponding triangle-transversal of G of size $\frac{|d_3(H)|}{2} + 1$, we have that $G - T \subseteq L(P)$ and so T is in fact a feedback vertex set. □

The strategy becomes evident: we would like to reduce from HAMILTONIAN PATH restricted to planar cubic bipartite graphs. Therefore, we now deal with the hardness of this problem. In order to do so, we first need an auxiliary result related to the following:

HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE	
Instance:	A graph $G = (V, E)$ and $e \in E$.
Question:	Does G contain a Hamiltonian cycle through e?

HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE was shown to be NP-complete even when restricted to planar cubic bipartite graphs [118]. We find useful to present the proof, since it introduces an operation which will be used in Section 2.2.4 as well. Given a graph G and a 3-vertex u, a *hexagon implant* is the operation replacing u by the gadget depicted in Figure 2.5.

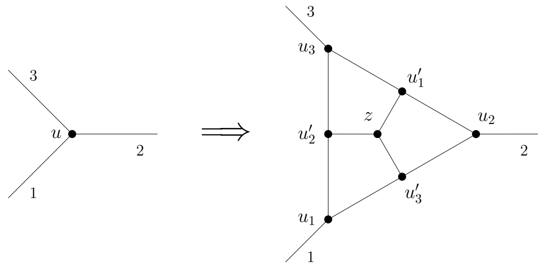


Figure 2.5: Hexagon implant: the cubic vertex u is replaced by a gadget containing 7 vertices.

Theorem 2.2.14 (Labarre [118]). HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE *is* NP-*complete even for planar cubic bipartite graphs.*

Proof. We reduce from HAMILTONIAN CYCLE, known to be NP-complete even for planar cubic bipartite graphs [4]. Given a planar cubic bipartite graph G, we apply a hexagon implant to a vertex $u \in V(G)$ (see Figure 2.5) and we set $e = u'_2 z$. Clearly, the resulting graph G' is planar, cubic and bipartite. Moreover, it is easy to see that G' contains a Hamiltonian cycle through e if and only if G contains a Hamiltonian cycle. □

We can now prove that HAMILTONIAN PATH remains NP-hard for planar cubic bipartite graphs. This comes as no surprise, since the related HAMILTONIAN CYCLE is NP-hard for the same class [4]. It is worth noticing, and an easy exercise, that a connected cubic bipartite graph is always 2-connected.

Theorem 2.2.15. HAMILTONIAN PATH *is* NP-*complete even for planar cubic bipartite graphs.*

Proof. We reduce from HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE restricted to planar cubic bipartite graphs, which is NP-complete by Theorem 2.2.14. Given an instance of this problem, i.e. a graph G as above and $uv \in E(G)$, we build a graph G' by substituting the edge uv with the gadget depicted in Figure 2.6. It is easy to see that G' is planar, cubic and bipartite. We claim that G contains a Hamiltonian cycle through uv if and only if G' contains a Hamiltonian path.

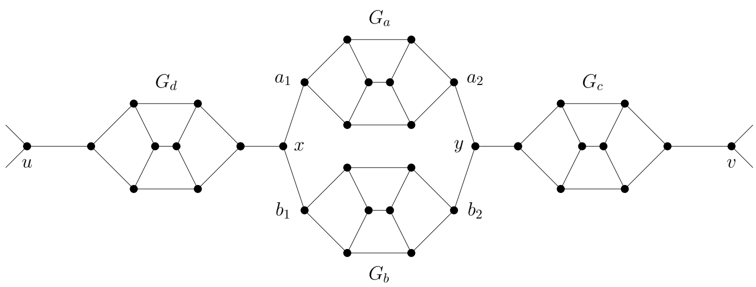


Figure 2.6: The gadget in G' replacing the edge uv.

Suppose first G contains a Hamiltonian cycle C through uv. It is easy to see that $C - uv$ can be extended to a Hamiltonian a_1, b_2-path in G'.

Conversely, suppose G' contains a Hamiltonian path P. It is easy to see that $|E(P) \cap \{a_1x, a_2y, b_1x, b_2y\}|$ is either 2 or 3. If it is equal to 2, then $E(P)$ contains exactly one edge incident to the subgraph G_a and exactly one edge incident to the subgraph G_b. This implies that both G_a and G_b contain a vertex of degree 1 in P. Suppose now that $|E(P) \cap \{a_1x, a_2y, b_1x, b_2y\}| = 3$, say without loss of generality $E(P)$ contains $\{a_1x, a_2y, b_2y\}$. By the reasoning above, both G_b and G_c contain a vertex of degree 1 in P. This means that, in either case, the vertices of degree 1 in P belong to the gadget in G' replacing uv and so there exists a Hamiltonian u, v-path in G. The conclusion immediately follows. $\square$

We finally have all the machinery to address FEEDBACK VERTEX SET:

Theorem 2.2.16. FEEDBACK VERTEX SET *remains* NP*-complete even for line graphs of planar cubic bipartite graphs.*

Proof. We reduce from HAMILTONIAN PATH, which is NP-complete even when restricted to planar cubic bipartite graphs (Theorem 2.2.15). Given such a graph H, consider its line graph $G = L(H)$. By Lemma 2.2.13, we know that $\tau_c(G) \leq \frac{|V(G)|}{3} + 1$ if and only if H contains a Hamiltonian path. This concludes the proof. $\square$

Given the NP-hardness of FEEDBACK VERTEX SET restricted to $(K_4, \text{claw}, \text{diamond})$-free graphs, it is natural to ask whether the problem admits a PTAS. We now show that this is not the case, unless P $=$ NP. To this end, let us recall the following problem:

E3-OCC-MAX-E2-SAT	
Instance:	A formula Φ with variable set X and clause set C, such that each variable has exactly three literals (in three different clauses) and each clause is the disjunction of exactly two literals (of two different variables).
Task:	Find a truth assignment maximizing the number of satisfied clauses.

Berman and Karpinski [21] showed that, for every sufficiently small $\varepsilon > 0$, it is NP-hard to distinguish between those instances of E3-OCC-MAX-E2-SAT for which there is a truth assignment satisfying at least $(\frac{788}{792} - \varepsilon)|C|$ clauses and those for which every truth assignment satisfies at most $(\frac{787}{792} + \varepsilon)|C|$ clauses.

Let Φ be an instance of E3-OCC-MAX-E2-SAT. Let $t(f)$ be the number of clauses of Φ satisfied by a truth assignment f and let $t(\Phi)$ be the maximum value of $t(f)$, taken over all truth assignments f of Φ. Clearly, we may assume that the three literals of each variable are neither all positive nor all negative. Moreover, by eventually replacing each variable x appearing twice negated and once unnegated by $\overline{x}$, we obtain a new instance Φ' with $t(\Phi) = t(\Phi')$ and such that each variable appears twice unnegated and once negated.

Theorem 2.2.17. FEEDBACK VERTEX SET *is not approximable within* $\frac{2117}{2116}$*, unless* P $=$ NP*, even when restricted to* $(K_4, \text{claw}, \text{diamond})$*-free graphs.*

Proof. We construct a gap-preserving reduction from E3-OCC-MAX-E2-SAT. Given a formula Φ with variable set X and clause set C, we build a graph G as follows. First, for any variable $x \in X$, we introduce the gadget G_x depicted in Figure 2.7: the *literal vertices* x_1 and x_2 correspond to the unnegated occurrences of x, while the *literal vertex* $\overline{x}_1$ corresponds to the negated one. Finally, for any clause $c = x \vee y$, we create a triangle having as vertices the literal vertices x and y and a new vertex v_c. It is easy to see that the resulting graph G is $(K_4, \text{claw}, \text{diamond})$-free. We claim that $\tau_c(G) = 4|X| + |C| - t(\Phi)$.

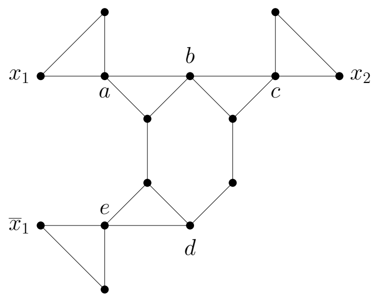


Figure 2.7: The variable gadget G_x.

Given a truth assignment f of Φ such that $t(f) = t(\Phi)$, we build a feedback vertex set T of G as follows. For any variable x, we add to T either $\{x_1, x_2, b, e\}$, if x evaluates to true, or $\{\overline{x}_1, a, c, d\}$, otherwise. Clearly, it only remains to check whether T intersects the cycles corresponding to clauses. If a clause c is satisfied, the corresponding triangle already contains a vertex in T. Otherwise, we simply add to T one literal vertex belonging to c, thus obtaining a feedback vertex set of G. Therefore, we have $\tau_c(G) \leq 4|X| + |C| - t(\Phi)$.

Conversely, let T be a feedback vertex set of G such that $|T| = \tau_c(G)$. Note that $|T \cap V(G_x)| \geq 4$, for any variable x. We define a truth assignment f of Φ as follows: we set x to true if $T \cap \{x_1, x_2\} \neq \varnothing$, and to false otherwise. Consider now a clause c not satisfied by f. If $v_c \notin T$ then, since the literal vertex corresponding to an unnegated literal (which evaluates to false under f) does not belong to T, it must be that c contains a negated literal $\overline{x}_1$ such that the corresponding literal vertex belongs to T. Moreover, since x is set to true, $T \cap \{x_1, x_2\} \neq \varnothing$. Therefore, T contains both an unnegated and a negated literal vertex of G_x and so it is easy to see that $|T \cap V(G_x)| \geq 5$. Summarizing, we have that for each unsatisfied clause c, either $v_c \in T$ or c contains a negated literal $\overline{x}_1$ and $|T \cap V(G_x)| \geq 5$. But then, denoting by p the number of unsatisfied clauses c such that $v_c \in T$, we have

$$\begin{aligned}
\tau_c(G) &= |T| \\
&\geq p + 5(|C| - t(f) - p) + 4(|X| - (|C| - t(f) - p)) \\
&= 4|X| + |C| - t(f) \\
&\geq 4|X| + |C| - t(\Phi).
\end{aligned}$$

Now let $m = |C|$ and $n = |X|$. Clearly, $3n = 2m$. Recall that it is NP-hard to distinguish between those instances Φ of E3-OCC-MAX-E2-SAT for which $t(\Phi) \geq (\frac{788}{792} - \varepsilon)m$ and those for which $t(\Phi) \leq (\frac{787}{792} + \varepsilon)m$. On the other hand, if $t(\Phi) \geq (\frac{788}{792} - \varepsilon)m$, then

$$\tau_c(G) \leq 4 \cdot \frac{2}{3}m + m - \left(\frac{788}{792} - \varepsilon\right)m = \left(\frac{2116}{792} + \varepsilon\right)m$$

and if $t(\Phi) \leq (\frac{787}{792} + \varepsilon)m$, then

$$\tau_c(G) \geq 4 \cdot \frac{2}{3}m + m - \left(\frac{787}{792} + \varepsilon\right)m = \left(\frac{2117}{792} - \varepsilon\right)m.$$

Therefore, there is no $\frac{2117}{2116}$-approximation algorithm for FEEDBACK VERTEX SET restricted to $(K_4, \text{claw}, \text{diamond})$-free graphs, unless P = NP. $\square$

Note that we made no effort in optimizing the constant in Theorem 2.2.17. In fact, it could be improved by using the improved inapproximability result for E3-OCC-MAX-E2-SAT obtained in [22]. An even stronger improvement could be possibly obtained by a direct gap-preserving reduction from the problem E3-OCC-E2-LIN-2 (see [22]).

A planarity constraint cannot be added to Theorem 2.2.17, since Demaine and Hajiaghayi [50] showed that FEEDBACK VERTEX SET admits a PTAS when restricted to planar graphs. Moreover, we leave as an open problem to determine the approximation hardness of FEEDBACK VERTEX SET restricted to line graphs of cubic bipartite graphs.

Bafna et al. [17] showed that FEEDBACK VERTEX SET has a 2-approximation for general graphs. Given the fact that we can find a minimum-size triangle-transversal of a $(K_4, \text{claw}, \text{diamond})$-free graph in polynomial time, if we want to approximate FEEDBACK VERTEX SET for that same class, it is natural to consider the following algorithm:

Algorithm 1

Require: A $(K_4, \text{claw}, \text{diamond})$-free graph G.
Ensure: A feedback vertex set of G.
1: Find a minimum-size triangle-transversal T_Δ of G.
2: Find a minimum-size feedback vertex set T of $G - T_\Delta$.
3: **return** $T_G = T \cup T_\Delta$.

Unfortunately, Algorithm 1 does not improve on the factor 2 algorithm for general graphs. Indeed, there exists an infinite sequence of graphs for which the approximation factor r of Algorithm 1 gets arbitrarily close to 2: just consider the graph depicted in Figure 2.8 and containing $2k$ triangles. Algorithm 1 returns a feedback vertex set of size $k + \frac{2k-2}{2}$, whereas an optimum solution has size $\frac{2k-2}{2} + 2$.

Figure 2.8: A graph with r arbitrarily close to 2.

On the other hand, if we consider 4-regular graphs, we obtain a $\frac{3}{2}$-approximation algorithm:

Theorem 2.2.18. *Algorithm 1 is a $\frac{3}{2}$-approximation algorithm for 4-regular $(K_4, \text{claw}, \text{diamond})$-free graphs. It runs in $O(n^3)$ time, where n is the number of vertices of the input graph.*

Proof. Algorithm 1 clearly returns a feedback vertex set of G. Consider now the approximation factor r. Let $\tau_c(G) = \tau_\Delta(G) + a$, for some $a \geq 0$, and let $T_G = T \cup T_\Delta$ be the feedback vertex set found by Algorithm 1, where T_Δ is a minimum-size triangle-transversal of G and T is a minimum-size feedback vertex set of $G - T_\Delta$. We have that

$$r = \frac{|T_G|}{\tau_c(G)} = \frac{\tau_\Delta(G) + |T|}{\tau_\Delta(G) + a} \leq 1 + \frac{|T|}{\tau_\Delta(G)}.$$

Recall now that $G = L(H)$, for a cubic triangle-free graph H. By Corollary 2.2.11 and Gallai's identity, we have $\tau_\Delta(G) = |V(H)| - \alpha'(H) \geq \frac{|V(H)|}{2}$. Moreover, $|T| \leq \frac{|V(G)|-\tau_\Delta(G)}{4}$ and so

$$\frac{|T_G|}{\tau_c(G)} \leq 1 + \frac{|T|}{\tau_\Delta(G)} \leq 1 + \frac{1}{4}\left(\frac{|V(G)|}{\tau_\Delta(G)} - 1\right) \leq 1 + \frac{1}{4}\left(\frac{3}{2} \cdot \frac{|V(H)|}{\tau_\Delta(G)} - 1\right) \leq \frac{3}{2}.$$

As for the running time, since $G - T_\Delta$ has maximum degree 2, we can find in linear time a minimum-size feedback vertex set of $G - T_\Delta$. Therefore, by Corollary 2.2.12, Algorithm 1 runs in $O(n^3)$ time. □

2.2.4 Complexity of Hamiltonian Cycle and Hamiltonian Path

In this section, we investigate the computational complexity of HAMILTONIAN CYCLE and HAMILTONIAN PATH when restricted to line graphs. The topic of Hamiltonicity in line graphs has a rich history, dating back to the works of Chartrand [34, 35] and Harary and Nash-Williams [86]. At its core lies the concept of sequential ordering. A *sequential ordering* of the m edges of a graph G is an ordering $e_0, e_1, \ldots, e_{m-1}$ such that e_i and e_{i+1} are incident, for any $i \in \{0, \ldots, m-1\}$ (indices are taken modulo m). A graph admitting such an ordering is called *sequential*. It is easy to see that every Hamiltonian graph is sequential and the following result reveals the relevance of sequential graphs.

Theorem 2.2.19 (Chartrand [34]). *Given a graph G, its line graph $L(G)$ is Hamiltonian if and only if G is sequential.*

In particular, Theorem 2.2.19 implies that if G is Hamiltonian then $L(G)$ is Hamiltonian too. The converse is in general not true, as can be seen by considering $K_{2,3}$, which is bipartite, or the Petersen graph, which is cubic. It is not the case even if G is cubic and bipartite, as can be seen by considering [57, Figure 4]. On the other hand, passing to a 1-subdivision, we obtain the following:

Lemma 2.2.20. *Let G be a subcubic graph and let G' be a 1-subdivision of G. If $L(G')$ is Hamiltonian, then G is Hamiltonian too.*

Proof. If $L(G')$ is Hamiltonian then, by Theorem 2.2.19, G' is sequential and let $e_0, e_1, \ldots, e_{m-1}$ be a sequential ordering of G'. Clearly, $\delta(G) \geq 2$ and consider a 3-vertex $u \in V(G') \cap V(G)$, i.e. u is not the result of an edge subdivision. Let a_1, a_2 and a_3 be the neighbours of u in G'. It is not difficult to see that the edges ua_1, ua_2 and ua_3 are consecutive (modulo m) in the sequential ordering, say $e_i = ua_1$, $e_{i+1} = ua_2$ and $e_{i+2} = ua_3$. But then $e_{i-1} = a_1v$ and $e_{i+3} = a_3w$, for some $\{v, w\} \subseteq V(G') \cap V(G)$ with $\{uv, uw\} \subseteq E(G)$. Therefore, for each 3-vertex $u \in V(G)$, we select the edges in $E(G)$ which correspond to the "leftmost" and "rightmost" edge incident to u in the sequential ordering of G'. A similar reasoning applies to 2-vertices. The resulting subgraph is 2-regular and connected. □

Note that Lemma 2.2.20 does not hold if G contains vertices of degree 4, as can be seen by considering a 1-subdivision of the bowtie (see Figure 1.1).

It is easy to see that if a graph G is Hamiltonian, then a 1-subdivision of G is sequential. Therefore, Theorem 2.2.19 and Lemma 2.2.20 imply the following:

Corollary 2.2.21. *Let G be a subcubic graph and let G' be a 1-subdivision of G. We have that $L(G')$ is Hamiltonian if and only if G is.*

We can now state our first result, which implies that HAMILTONIAN CYCLE remains NP-hard for planar cubic line graphs.

Theorem 2.2.22. HAMILTONIAN CYCLE *is* NP*-complete even for line graphs of* 1*-subdivisions of planar cubic bipartite graphs.*

Proof. It is known that HAMILTONIAN CYCLE is NP-complete even for planar cubic bipartite graphs [4]. Given an instance G of this problem, consider the line graph $L(G')$ of a 1-subdivision G' of G. The statement follows by Corollary 2.2.21. □

Note that the construction in the proof of Theorem 2.2.22 can be rephrased in terms of the following operation. Given a cubic graph G, a *Y-extension* of G consists in replacing each $u \in V(G)$ by a triangle T_u, where each $x \in V(T_u)$ corresponds to an edge incident to u, and connecting the vertices of the triangles which correspond to the same edge.

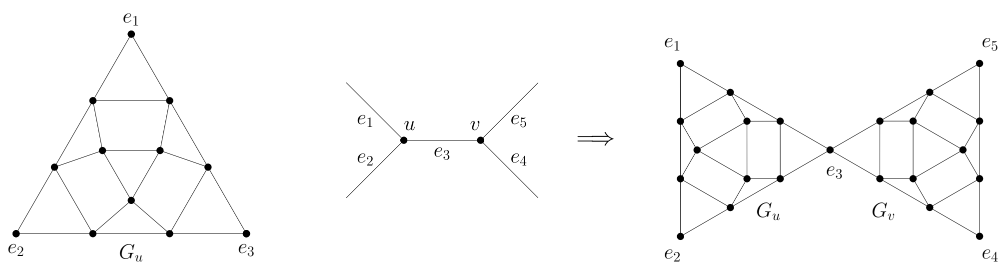
We now proceed with the case of line graphs of planar cubic bipartite graphs, for which the hexagon implant operation defined in Section 2.2.3 comes in handy (see Figure 2.5).

Theorem 2.2.23. HAMILTONIAN CYCLE *is* NP*-complete even for line graphs of planar cubic bipartite graphs.*

Proof. Once again, we reduce from HAMILTONIAN CYCLE restricted to planar cubic bipartite graphs [4]. Given an instance G of this problem, we build a graph G' as follows: $G' = L(H)$, where H is the graph obtained from G by applying a hexagon implant to each vertex of G. Since G is planar, cubic and bipartite, the same holds for H. We claim that G is Hamiltonian if and only if G' is.

Suppose first G is Hamiltonian. It is easy to see that H is Hamiltonian too (in fact, also the converse holds) and so, by Theorem 2.2.19, $G' = L(H)$ is Hamiltonian.

Conversely, suppose G' has a Hamiltonian cycle C'. At this point, it is useful to consider an equivalent construction of G'. Starting from G, we first replace each $u \in V(G)$ by the gadget G_u depicted in Figure 2.9(a). It has three specified vertices, called the *gates*, each one corresponding to a different edge incident to u. Finally, we identify corresponding gates. It is easy to see that the resulting graph is isomorphic to G'.



(a) **(b)**

Figure 2.9: (a) The gadget G_u corresponding to $u \in V(G)$: the gate $e_i \in V(G_u)$ corresponds to the edge $e_i \in E(G)$ incident to u. (b) The construction of G'.

We say that the Hamiltonian cycle C' *crosses* the gate e_i if the two edges of C' incident to e_i belong to different gadgets. It is not difficult to see that, for each gadget, C' crosses exactly two of its gates. But then, by selecting the edges in G corresponding to the crossed gates in

G', we obtain a Hamiltonian cycle in G. $\square$

As an immediate consequence of Theorem 2.2.23, we can state the following:

Corollary 2.2.24. *Let G be a cubic graph and let G' be the graph obtained from G by applying a hexagon implant to each vertex of G. We have that $L(G')$ is Hamiltonian if and only if G is.*

We can finally show an analogue of Theorem 2.2.22 for HAMILTONIAN PATH which strengthens the result in [119].

Theorem 2.2.25. HAMILTONIAN PATH *is* NP*-complete even for line graphs of* 1*-subdivisions of planar cubic bipartite graphs.*

Proof. We reduce from HAMILTONIAN PATH restricted to planar cubic bipartite graphs, shown to be NP-complete in Theorem 2.2.15. Given an instance G of this problem, the graph G' is obtained by a Y-extension of G. Note that there is a bijection between the edges of G and the edges of G' which do not belong to any triangle and a bijection between the vertices of G and the triangles of G'. We denote by T_u the triangle of G' corresponding to $u \in V(G)$ and let $V(T_u) = \{u_1, u_2, u_3\}$. Clearly, G' is isomorphic to the line graph of a 1-subdivision of G. We claim that G has a Hamiltonian path if and only if G' has.

Suppose first G has a Hamiltonian path P between a and b. By Corollary 2.2.21, we may assume $a \neq b$. We start by selecting the edges of G' corresponding to those of $E(P)$. For each $u \in V(G) \setminus \{a, b\}$, the selected edges are incident to exactly two vertices of T_u, say without loss of generality u_1 and u_2. For each such triangle, we now select the edges u_3u_1 and u_3u_2. In this way, we obtain a path P' containing all but four vertices (two vertices of T_a and two vertices of T_b). But then it is easy to extend P' to a Hamiltonian path in G'.

Conversely, suppose G' has a Hamiltonian a_1, b_1-path P (recall this means $a_1 \in V(T_a)$ and $b_1 \in V(T_b)$). By Corollary 2.2.21, we may assume $a_1 \neq b_1$ and $a_1b_1 \notin E(G')$. In particular, $a \neq b$. For $u \in V(G) \setminus \{a, b\}$, we have that $E(P)$ contains exactly two edges incident to T_u and $P[V(T_u)]$ is connected. Suppose first that, for each $u \in \{a, b\}$, the neighbour of u_1 in P belongs to T_u. Then $E(P)$ contains exactly one edge incident to T_u and $P[V(T_u)]$ is connected. Therefore, by contracting each triangle to a single vertex, we obtain a Hamiltonian a, b-path in G. Suppose now there exists $u \in \{a, b\}$ such that the neighbour of u_1 in P does not belong to T_u. Without loss of generality, we may assume $u = a$ and let the neighbour of a_1 in P belong to T_v. If the vertices of T_v do not occur consecutively in P, then $b_1 \in V(T_v)$ (and so $v = b$). In this case, it is easy to see that, by deleting the edge of P incident to a_1 and contracting each triangle to a single vertex, we obtain a Hamiltonian cycle in G. On the other hand, if the vertices of T_v occur consecutively in P, then $v \neq b$ and consider T_b. By the same reasoning as above, we may assume that either the neighbour of b_1 in P belongs to T_b and $P[V(T_b)]$ is connected, or the neighbour of b_1 in P belongs to T_w, with $w \neq b$, and the vertices of T_w occur consecutively in P. In the former case, we delete the edge of P incident to a_1 and we contract each triangle to a single vertex. In the latter case, we delete the edges of P incident to a_1 and to b_1 and we contract each triangle to a single vertex. In either case, it is easy to see that we obtain a Hamiltonian path in G. $\square$

2.3 k-Line graphs

Recall that, for an integer $k \geq 2$, the *k-line graph* $L_k(G)$ of a graph G is the graph having as vertices the cliques of G of size k, two vertices being adjacent if the corresponding cliques intersect in a clique of size $k-1$. A 2-line graph is the usual line graph and we refer to a 3-line graph as a triangle graph, denoting it by $T(G)$.

In this section, we show some basic properties of k-line graphs. Despite the fact that this class is not hereditary, we provide a partial list of forbidden induced subgraphs. As remarked in Section 2.1, our main motivation for studying k-line graphs is that Tuza's Conjecture (Conjecture 3.3.1), in the case of K_4-free graphs, can be expressed in terms of the triangle graph. Let us briefly explain this (a more detailed discussion is postponed to Chapter 3).

In Section 2.2.3, we have introduced vertex triangle-transversals. Similarly, we define an *edge triangle-transversal* of a graph G as a subset of $E(G)$ whose deletion results in a triangle-free graph. We denote by $\tau'_\Delta(G)$ the minimum size of an edge triangle-transversal of G and by $\nu'_\Delta(G)$ the maximum number of edge-disjoint triangles of G. Clearly, the following holds:

Fact 2.3.1. *For any graph G, we have $\nu'_\Delta(G) = \alpha(T(G))$.*

This is the "higher dimensional" analogue of the bijection between the matchings of a graph and the independent sets of its line graph. But what about $\tau'_\Delta(G)$? Continuing with the analogy, the case $k = 2$ gives us the following well-known fact:

Fact 2.3.2. *If G is a triangle-free graph, then $\beta(G) = \theta(L(G))$.*

The proof of Fact 2.3.2 follows by noticing that there are two types of cliques in a line graph G. Moreover, if G is the line graph of a triangle-free graph H, then a clique of G corresponds to the edges incident to a fixed vertex of H. The same situation occurs for $k \geq 3$:

Observation 2.3.3 (Lakshmanan et al. [120]). *Every n-clique of a k-line graph $L_k(G)$ either corresponds to n k-cliques of G sharing a fixed $(k-1)$-clique or to n k-cliques contained in a common $(k+1)$-clique.*

Proof. Let $c_1, \ldots, c_n$ be the vertices of a n-clique of $L_k(G)$ and let $C_1, \ldots, C_n$ be the corresponding k-cliques of G. Moreover, let $C_1 = \{v_1, \ldots, v_k\}$. Since c_1 and c_2 are adjacent in $L_k(G)$, we have $|C_1 \cap C_2| = k - 1$ and so we may assume $C_2 = \{u, v_2, \ldots, v_k\}$, for some $u \notin C_1$. Suppose now there exists a C_i with $i > 2$ such that $\{v_2, \ldots, v_k\} \nsubseteq C_i$, say without loss of generality $v_k \notin C_i$. Since c_i is adjacent to both c_1 and c_2, we have $C_i = \{u, v_1, \ldots, v_{k-1}\}$. But then, any other C_j must be a subset of the clique $\{u, v_1, v_2, \ldots, v_k\}$. □

Observation 2.3.3 tells us that if G is a K_4-free graph, then a clique of $T(G)$ corresponds to a set of triangles of G sharing a common edge and so a clique cover of $T(G)$ corresponds to an edge triangle-transversal of G:

Fact 2.3.4 (Lakshmanan et al. [122]). *If G is a K_4-free graph, then $\tau'_\Delta(G) = \theta(T(G))$.*

Fact 2.3.1 and Fact 2.3.4 will be repeatedly used in Section 3.3 in order to reduce Tuza's Conjecture to a more manageable statement about bounded clique covers (see Section 3.2). In this context, it is useful to see which graphs cannot appear as induced subgraphs in a triangle graph. The same can be asked, more generally, for k-line graphs. In fact, the definition immediately implies the following:

Observation 2.3.5. *Every k-line graph is $K_{1,k+1}$-free.*

Clearly, any two triangles of a graph share at most one edge and so Observation 2.3.3 implies the following:

Observation 2.3.6. *If G is a K_4-free graph, then two maximal cliques of $T(G)$ cannot have more than one vertex in common. In particular, $T(G)$ is diamond-free.*

Observation 2.3.6 shows once again that the class of triangle graphs is not hereditary: a diamond is not a triangle graph, whereas it is an induced subgraph of $T(K_2 \vee \overline{P_3})$.

Observation 2.3.7 (Le and Prisner [125]). *If c_1 and c_2 are two non-adjacent vertices of a k-line graph $G = L_k(H)$, then $G[N(c_1) \cap N(c_2)]$ is an induced subgraph of C_4.*

Proof. Let C_1 and C_2 be the k-cliques of H corresponding to c_1 and c_2 and suppose that $N(c_1) \cap N(c_2) \neq \varnothing$. It is easy to see that $|C_1 \cap C_2| = k - 2$ and so we may assume $C_1 = \{u_1, u_2, a_1, \ldots, a_{k-2}\}$ and $C_2 = \{v_1, v_2, a_1, \ldots, a_{k-2}\}$. On the other hand, there are exactly four sets of size k having $k - 1$ elements in common with both C_1 and C_2, namely the sets of the form $\{u_i, v_j, a_1, \ldots, a_{k-2}\}$ for $\{i, j\} \subseteq \{1, 2\}$, and the conclusion follows. $\square$

Corollary 2.3.8. *Every k-line graph is $K_{2,3}$-free and $\overline{K_2} \vee \overline{P_3}$-free.*

We now show that the graph depicted in Figure 2.10 does not appear as an induced subgraph.

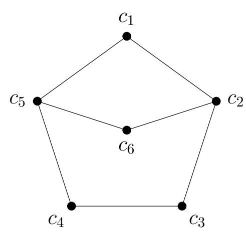


Figure 2.10: Twin-C_5.

Observation 2.3.9. *Every k-line graph is twin-C_5-free.*

Proof. Suppose $G = L_k(H)$ is a k-line graph containing an induced twin-C_5 with vertex set $\{c_1, \ldots, c_6\}$, as depicted in Figure 2.10. For $1 \leq i \leq 6$, let C_i be the k-clique of H corresponding to c_i. Since c_1 and c_6 are both adjacent to c_5 and c_2 but $c_1c_6 \notin E(G)$ and $c_2c_5 \notin E(G)$, it is easy to see that $C_1 = \{u_1, u_2, a_1, \ldots, a_{k-2}\}$, $C_6 = \{v_1, v_2, a_1, \ldots, a_{k-2}\}$, $C_5 = \{u_1, v_1, a_1, \ldots, a_{k-2}\}$ and $C_2 = \{u_2, v_2, a_1, \ldots, a_{k-2}\}$, where $\{u_1, u_2\} \cap \{v_1, v_2\} = \varnothing$.

Suppose now the clique C_4 does not contain some a_i, say a_{k-2}. Since $|C_4 \cap C_5| = k-1$ and $|C_4 \cap C_2| = k-2$, we have either $C_4 = \{u_1, v_1, u_2, a_1, \ldots, a_{k-3}\}$ or $C_4 = \{u_1, v_1, v_2, a_1, \ldots, a_{k-3}\}$. In either case we obtain a contradiction to the fact that $c_4c_1 \notin E(G)$ and $c_4c_6 \notin E(G)$. Therefore, we have $\{a_1, \ldots, a_{k-2}\} \subseteq C_4$. On the other hand, since $c_4c_1 \notin E(G)$ and $c_4c_6 \notin E(G)$, then $C_4 \cap \{u_1, u_2, v_1, v_2\} = \varnothing$, a contradiction to the fact that $c_4c_5 \in E(G)$. □

As noticed in the proof of Observation 2.3.9, an induced C_4 in a triangle graph $T(G)$ corresponds to a W_4 in G. Let us now consider an induced C_5:

Lemma 2.3.10. *An induced C_5 in the triangle graph of a K_4-free graph G corresponds to an induced W_5 in G.*

Proof. Let $uvwxy$ be the 5-cycle in $T(G)$. It is easy to see that the 4-path $uvwx$ corresponds either to the subgraph depicted in Figure 2.11(a) or to the one in Figure 2.11(b) (note that these subgraphs are not necessarily induced).

Figure 2.11: The two not necessarily induced subgraphs of G corresponding to an induced P_4 in $T(G)$.

Suppose first the latter situation occurs. The triangle y shares edges with u and x but not with v and w and so it has to contain two edges in the set $\{ab, ac, df, ef\}$, which is clearly impossible. Therefore, $uvwx$ corresponds to the subgraph in Figure 2.11(a). Again, the triangle y shares edges with u and x but not with v and w and so it contains two edges in the set $\{ab, bd, df, ef\}$. This can happen only if $\{b, d, f\}$ induces a triangle, therefore giving rise to a W_5. □

Note that if G is not K_4-free, then an induced C_5 in $T(G)$ may also correspond to a K_5 in G. For general C_n, the situation becomes even more complicated. Nevertheless, Lakshmanan et al. [121] provided a forbidden subgraph characterization of graphs with C_n-free triangle graphs, for any specified $n \geq 3$.

Let us now consider anti-Gallai graphs. Recall that the anti-Gallai graph $\Delta(G)$ of G is the graph having as vertices the edges of G, two vertices being adjacent if the corresponding edges are incident and span a triangle in G. It directly follows from the definition that every edge of an anti-Gallai graph belongs to at least one triangle. Moreover, if G is K_4-free, it is easy to see that every edge of $\Delta(G)$ belongs to at most one triangle. Recall that a graph G is locally linear if each edge of G belongs to exactly one triangle or, equivalently, if $G[N(v)]$ is 1-regular for every $v \in V(G)$ (see Section 2.2.1). Therefore, the following holds:

Observation 2.3.11. *If G is a K_4-free graph, then $\Delta(G)$ is locally linear.*

If G is K_4-free, it is not difficult to interpret edge triangle-transversals and edge-disjoint triangles of G in terms of the anti-Gallai graph of G. Indeed, by Observation 2.3.11, the map f which sends a triangle of G with edge set $\{e_1, e_2, e_3\}$ to the triangle of $\Delta(G)$ with vertex set $\{e_1, e_2, e_3\}$ is a bijection between the triangles of G and those of $\Delta(G)$. This implies that

$\nu'_\Delta(G) = \nu_\Delta(\Delta(G))$ and $\tau'_\Delta(G) = \tau_\Delta(\Delta(G))$.

Recall now that the *clique graph* $K(G)$ of G is the graph having as vertices the maximal cliques of G, two vertices being adjacent if the corresponding cliques share at least one vertex. If G is K_4-free, Observation 2.3.11 implies that $\Delta(G)$ is K_4-free as well and, if every edge of G belongs to a triangle, the triangles of $\Delta(G)$ are exactly its maximal cliques. Therefore, the map f introduced above gives a bijection between the vertices of $T(G)$ (i.e. the triangles of G) and the vertices of $K(\Delta(G))$ which clearly preserves adjacency. This implies that the following holds:

Observation 2.3.12. *If G is a K_4-free graph such that each edge belongs to a triangle, then $T(G) \cong K(\Delta(G))$.*

In fact, Lakshmanan et al. [120] showed that the converse holds as well: a connected graph F is the anti-Gallai graph of a K_4-free graph G such that every edge of G belongs to a triangle if and only if $K(F) \cong T(G)$. This was used in order to reduce the recognition of anti-Gallai graphs of K_4-free graphs to that of triangle graphs. As mentioned in Section 2.1, the former is NP-complete [9] and so they showed that recognizing triangle graphs is NP-complete as well.

We conclude this chapter with the following observation, which will be used in Section 3.3.1:

Observation 2.3.13 (Lakshmanan et al. [120]). *If G is the k-line graph of a K_{k+1}-free graph, then it is also the k'-line graph of a $K_{k'+1}$-free graph, for any $k' > k$.*

Proof. Let $G = L_k(H)$, for a K_{k+1}-free graph H, and consider the join $H' = H \vee K_{k'-k}$. Since H is K_{k+1}-free, we have that H' is $K_{k'+1}$-free and every k'-clique of H' corresponds to a k-clique of H. Moreover, two k'-cliques of H' intersect in a clique of size $k'-1$ if and only if the corresponding k-cliques of H intersect in a clique of size $k-1$. Therefore, we have $L_{k'}(H') = L_k(H) = G$. $\square$

CHAPTER 3

Approximate Min-Max Theorems

Contents

In this chapter, we consider three hypergraphs having the Erdős-Pósa Property and we seek to determine the optimal bounding functions. First, extending a result by Henning et al. [95], we show that $f(x) = \lfloor \frac{3}{2}x \rfloor$ is a θ-bounding function for the class of subcubic graphs and that it is best possible. Moreover, we provide a θ-bounding function for the class of graphs with maximum degree at most 4. Finally, we study CLIQUE COVER and show it admits a PTAS for planar graphs.
Then we consider Tuza's Conjecture, which asserts that the minimum number of edges of a graph whose deletion results in a triangle-free graph is at most twice the maximum number of edge-disjoint triangles. We show that the constant 2 can be improved for some K_4-free graphs whose edges are contained in at most four triangles and graphs obtained by forbidding certain odd-wheels.
Finally, we consider Jones' Conjecture: the minimum number of vertices of a planar graph whose deletion makes it acyclic is at most twice the maximum number of vertex-disjoint cycles. We show it holds for claw-free graphs with maximum degree at most 4 and we make some observations for subcubic graphs.

3.1 Introduction

The chromatic number $\chi(G)$ of a graph G is the minimum number of independent sets needed to cover the vertices of G, while the clique number $\omega(G)$ is the maximum size of a clique of G. Since any independent set contains at most one vertex of a clique, we have $\chi(G) \geq \omega(G)$. A class of graphs $\mathcal{G}$ is χ-*bounded* if there exists a function $f\colon \mathbb{N} \to \mathbb{R}$ such that for all $G \in \mathcal{G}$ and all induced subgraphs H of G, we have $\chi(H) \leq f(\omega(H))$. Such a function f is a χ-*bounding* function for $\mathcal{G}$. Note that not all graphs are χ-bounded: the so-called Mycielski's construction provides triangle-free graphs with arbitrarily large chromatic number (see, e.g.,

[182]). Gyárfás [81] introduced the concept of χ-bounded class in order to provide a natural extension of the class of perfect graphs: indeed, this class is exactly the class of graphs χ-bounded by the identity function. The notion of χ-bounded class has been extensively studied, especially in the context of hereditary classes (see, e.g., [81, 161]), and many conjectures on the χ-boundedness of certain classes have been formulated [81].

By substituting χ with θ and ω with α, we obtain the notion of θ-boundedness and the two are complementary, in the sense that G is χ-bounded if and only if $\overline{G}$ is θ-bounded. In [81], Gyárfás formulated the following meta-question: given a class $\mathcal{G}$, what is the smallest θ-bounding function for $\mathcal{G}$, if any? In Section 3.2, we consider this question for classes of graphs having bounded maximum degree. It is easy to see that $\theta(G) \leq k\alpha(G)$, for any graph G with maximum degree at most k. Indeed, given a maximal independent set I of G, we have that the edges incident to I constitute a clique cover of G and their number is at most $k\alpha(G)$. On the other hand, for $k = 3$, we show that this bound is far from optimal: $f(x) = \lfloor \frac{3}{2}x \rfloor$ is a θ-bounding function for the class of subcubic graphs and it is best possible. Moreover, we give some insight for the case of graphs with maximum degree four. The study of these two cases is also motivated by a result of Henning et al. [95] showing that $\theta(G) \leq \frac{3}{2}\alpha(G)$, for any subcubic triangle-free graph G, and by a result of Joos [101] showing that $\theta(G) \leq \frac{7}{4}\alpha(G)$, for any triangle-free graph G with maximum degree four. In Section 3.2, we also treat some algorithmic aspects related to clique covering: in particular, answering a question by Cerioli et al. [31], we provide a PTAS for Clique Cover when restricted to planar graphs.

Let us now give an equivalent formulation of the concept of θ-boundedness (and, similarly, of χ-boundedness) as a "packing and covering relation". In fact, all the problems we treat in this chapter can be expressed in this way and in order to give the global picture, we first recall some definitions. A packing of a hypergraph $\mathcal{H} = (V, E)$ is a set of pairwise disjoint edges of $\mathcal{H}$ and a transversal (also known as hitting set or covering) of $\mathcal{H}$ is a subset $X \subseteq V$ intersecting each edge of $\mathcal{H}$. The packing number $\nu(\mathcal{H})$ of $\mathcal{H}$ is the number of edges in a packing of $\mathcal{H}$ of maximum size (a maximum packing) and the transversal number $\tau(\mathcal{H})$ of $\mathcal{H}$ is the number of vertices in a transversal of $\mathcal{H}$ of minimum size (a minimum transversal).

We can now reformulate the notion of θ-boundedness as follows. Given a graph G, the *clique hypergraph* $\mathcal{H}(G)$ of G is the hypergraph having as vertices the vertices of G and as edges the maximal cliques of G. Consider its dual $\mathcal{H}(G)^*$: we have that $\theta(G) = \tau(\mathcal{H}(G)^*)$ and $\alpha(G) = \nu(\mathcal{H}(G)^*)$. Therefore, a class of graphs $\mathcal{G}$ is θ-bounded if and only there exists a function f such that for all $G \in \mathcal{G}$ and all induced subgraphs H of G, we have $\tau(\mathcal{H}(H)^*) \leq f(\nu(\mathcal{H}(H)^*)$. A similar relation can be obtained for χ-boundedness by considering the hypergraphs having as edges the maximal independent sets.

Let us now come back to the general packing and covering setting. Since no vertex covers two edges of a packing, we have $\tau(\mathcal{H}) \geq \nu(\mathcal{H})$. This implies that obtaining a packing and a transversal of the same size proves each of them to be optimal. A family of hypergraphs satisfies the *Min-Max Property* if $\nu(\mathcal{H}) = \tau(\mathcal{H})$, for each member $\mathcal{H}$ of the family. There are several families of hypergraphs satisfying the Min-Max Property and the most prominent example is probably given by the following well-known result (see, e.g., [182]):

Theorem 3.1.1 (König-Egerváry Theorem). *The family of bipartite graphs satisfies the Min-Max Property.*

Another example is given by Menger's Theorem (Theorem 1.0.3). Indeed, given a graph G and vertices x and y of G, consider the hypergraph on the vertex set of G having as hyperedges the vertex sets of x, y-paths. Clearly, the transversals of this hypergraph are exactly the x, y-cuts of G and the packings are exactly the families of independent x, y-paths in G.

The Min-Max Property allows a good characterization of the packing and transversal numbers. Indeed, to show that $\nu(\mathcal{H}) \leq k$ or $\tau(\mathcal{H}) \geq k$, it is enough to exhibit a transversal or a packing of size k, respectively. Unfortunately, most families of hypergraphs do not satisfy the Min-Max Property, but it is still of interest to find an upper bound for τ in terms of ν, if any. If we have this sort of approximate min-max relation, we usually say that the family of hypergraphs satisfies the Erdős-Pósa Property (this naming will appear consistent in the following paragraphs): a family of hypergraphs satisfies the *Erdős-Pósa Property* if there exists a function f such that $\tau(\mathcal{H}) \leq f(\nu(\mathcal{H}))$, for each member $\mathcal{H}$ of the family. This implies that one parameter is characterized by its obstructing analogue, or dual: either $\mathcal{H}$ contains a packing of size k or it contains a transversal of size $f(k)$.

The family of r-uniform hypergraphs satisfies the Erdős-Pósa Property. Indeed, consider an r-uniform hypergraph $\mathcal{H}$ and a maximal packing of $\mathcal{H}$. Since the union of the edges in this packing intersects all the edges of $\mathcal{H}$, we have

$$\tau(\mathcal{H}) \leq r\nu(\mathcal{H}). \tag{3.1}$$

Note that (3.1) is tight. A first example is given by the hypergraph $\mathcal{P}_r$ consisting of the lines of some projective plane of order $r-1$: indeed, $\nu(\mathcal{P}_r) = 1$ and $\tau(\mathcal{P}_r) = r$. Since projective planes do not exist for every value of r, it is worth considering also the following construction. Let $\mathcal{H}$ be the hypergraph consisting of all the subsets of size r of a set of size $kr-1$. Clearly, we have $\nu(\mathcal{H}) = k-1$ and $\tau(\mathcal{H}) = r(k-1)$.

A long-standing open problem known as Ryser's Conjecture and formulated in [94] asserts that (3.1) can be improved if the hypergraph is r-partite. Recall that a hypergraph is *r-partite* if the vertex set can be partitioned into r classes such that each edge of the hypergraph contains at most one vertex for each class.

Conjecture 3.1.2 (Ryser's Conjecture). *If $\mathcal{H}$ is an r-uniform r-partite hypergraph, then $\tau(\mathcal{H}) \leq (r-1)\nu(\mathcal{H})$.*

Note that Theorem 3.1.1 is the case $r = 2$ of Conjecture 3.1.2. Besides this, not much is known and the only other solved case is $r = 3$, settled by Aharoni [2] through a nice topological argument.

Theorem 3.1.3 (Aharoni [2]). *If $\mathcal{H}$ is a 3-uniform 3-partite hypergraph, then $\tau(\mathcal{H}) \leq 2\nu(\mathcal{H})$.*

In Section 3.3, we are interested in another family of hypergraphs constructed from graphs and which, contrary to clique hypergraphs, turns out to be uniform. Given a graph G, the *triangle hypergraph* $\mathcal{H}(G)$ of G is the hypergraph whose vertices are the edges of G and whose edges are the subsets spanning triangles in G. Since $\mathcal{H}(G)$ is 3-uniform, (3.1) implies that $\tau(\mathcal{H}(G)) \leq 3\nu(\mathcal{H}(G))$. This observation can be rephrased as follows: the minimum number of edges of a graph G whose deletion results in a triangle-free graph is at most three times the maximum number of edge-disjoint triangles of G. Considering the complete graphs on 4 and 5 vertices, it is clear that $\tau(\mathcal{H}(G))$ may be as large as $2\nu(\mathcal{H}(G))$ and Tuza conjectured that these cases are extremal:

Conjecture 3.1.4 (Tuza's Conjecture [178]). *For any graph G, we have $\tau(\mathcal{H}(G)) \leq 2\nu(\mathcal{H}(G))$.*

Note that, although similar in nature, Conjecture 3.1.4 is essentially different from Theorem 3.1.3, since a triangle hypergraph is in general not 3-partite. However, we will see in Section 3.3 that Theorem 3.1.3 has indeed some interesting consequences for Tuza's Conjecture.

Several partial results on Tuza's Conjecture have been obtained: for example, it is known to be true for graphs on n vertices with at least $\frac{7}{16}n^2$ edges [178], for 4-colourable graphs [122] and for graphs with maximum average degree less than 7 [159]. All these classes contain the complete graph K_4 and so the constant 2 in Conjecture 3.1.4 is optimal for them. But a natural question arises: What happens if we forbid K_4? Haxell et al. [91] showed that the constant 2 cannot essentially be improved: for every $\varepsilon > 0$, there exists a K_4-free graph G such that $\tau(\mathcal{H}(G)) > (2-\varepsilon)\nu(\mathcal{H}(G))$.

In Section 3.3, we consider subclasses of K_4-free graphs and we see how the results proved in Section 3.2 turn out to be useful in the context of packing and covering triangles. The main idea comes from Section 2.3, where we showed that if G is a K_4-free graph, then the size of a minimum triangle transversal of G is equal to the size of a minimum clique cover of the triangle graph $T(G)$. Moreover, for any graph G, the size of a maximum triangle packing of G is equal to the size of a maximum independent set of $T(G)$. Therefore, Tuza's Conjecture restricted to K_4-free graphs translates into the following equivalent assertion: $\theta(T(G)) \leq 2\alpha(T(G))$, for any K_4-free graph G. The classes we consider in Section 3.3 are essentially of two kinds: graphs with edges in few triangles (at most four) and graphs obtained by forbidding certain odd-wheels. We show that, in these cases, it is in fact possible to considerably reduce the constant 2 in Conjecture 3.1.4.

As a side remark, Krivelevich [117] proved that the two fractional relaxations of Tuza's Conjecture obtained by replacing τ with τ^* and ν with ν^* indeed hold. Moreover, Chapuy et al. [33] showed that these fractional versions hold even for graphs with multiple edges.

Let us now consider another family of hypergraphs arising from graphs. Given a graph G, the *cycle hypergraph* $\mathcal{H}(G)$ of G is the hypergraph whose vertices are the vertices of G and whose edges are the vertex sets of cycles of G. Note that a transversal of $\mathcal{H}(G)$ is nothing but a feedback vertex set of G. Contrary to the triangle hypergraph, the cycle hypergraph need not be uniform. Nevertheless, a fundamental result by Erdős and Pósa [59] shows that cycle hypergraphs indeed admit an approximate Min-Max Property:

Theorem 3.1.5 (Erdős and Pósa [59]). *For any graph G, we have $\tau(\mathcal{H}(G)) = O(\nu(\mathcal{H}(G)) \log \nu(\mathcal{H}(G)))$. Moreover, the bound is sharp.*

Kloks et al. [107] conjectured that cycle hypergraphs of planar graphs admit a considerably smaller bounding function[1]:

Conjecture 3.1.6 (Jones' Conjecture [107]). *If G is a planar graph, then $\tau(\mathcal{H}(G)) \leq 2\nu(\mathcal{H}(G))$.*

If true, Conjecture 3.1.6 would be sharp, as can be seen by considering wheel graphs. Kloks et al. [107] showed that it holds for outerplanar graphs and, in general, they proved the weaker $\tau(\mathcal{H}(G)) \leq 5\nu(\mathcal{H}(G))$. The factor 5 was later improved to 3 by a series of authors [32, 36, 134] and it currently gives the best general bound. To the best of our knowledge, these are the only works related to Conjecture 3.1.6.

In Section 3.4, we show that Jones' Conjecture indeed holds for claw-free graphs with

[1] In fact, this conjecture is known as Jones' Conjecture, where Jones is an alternate name for C. M. Lee (see [32]).

maximum degree 4. Moreover, we consider the case of subcubic graphs and we provide some properties a minimum counterxample must have, if any.

3.2 Bounded clique cover

As we have seen in Section 3.1, a class of graphs $\mathcal{G}$ is θ-bounded if there exists a function $f\colon \mathbb{N} \to \mathbb{R}$ such that for all $G \in \mathcal{G}$ and all induced subgraphs H of G, we have $\theta(H) \leq f(\alpha(H))$. Such a function f is a θ-bounding function for $\mathcal{G}$. The notion of θ-boundedness and its complementary χ-boundedness were introduced by Gyárfás [81] in order to provide a natural extension of the class of perfect graphs. In this section, we consider the following question formulated in [81]: given a class $\mathcal{G}$, what is the smallest θ-bounding function for $\mathcal{G}$, if any? In particular, we give an answer for the class of subcubic graphs:

Theorem 3.2.1. *If G is a subcubic graph, then $\theta(G) \leq \frac{3}{2}\alpha(G)$. Moreover, $f(x) = \lfloor \frac{3}{2}x \rfloor$ is the smallest θ-bounding function for the class of subcubic graphs.*

Elaborating on a result by Choudum et al. [39], Pedersen conjectured that $\theta(G) \leq \frac{3}{2}\alpha(G)$, for any subcubic triangle-free graph G (see [95]). Recall that, if G is a triangle-free graph and $\alpha'(G)$ denotes the maximum size of a matching in G, then $\theta(G) = \alpha'(G) + (|V(G)| - 2\alpha'(G)) = |V(G)| - \alpha'(G)$. Pedersen's conjecture was confirmed by Henning et al. [95], who actually proved the following generalization:

Theorem 3.2.2 (Henning et al. [95]). *If G is a subcubic graph, then*

$$\frac{3}{2}\alpha(G) + \alpha'(G) + \frac{1}{2}t(G) \geq |V(G)|,$$

where $t(G)$ denotes the maximum number of vertex-disjoint triangles of G. Moreover, equality holds if and only if every component of G is in $\{K_3, K_4, C_5, G_{11}\}$ (see Figure 3.1).

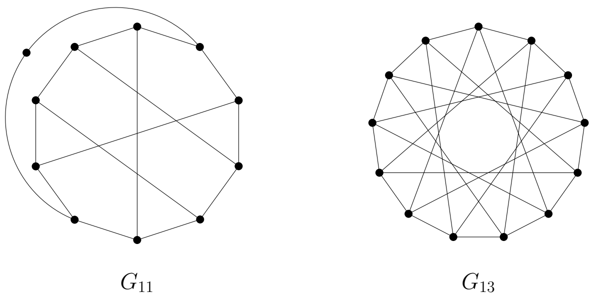


Figure 3.1: The graphs G_{11} and G_{13}.

Theorem 3.2.2 implies that $f(x) = \lfloor \frac{3}{2}x \rfloor$ is the smallest θ-bounding function for the class of subcubic triangle-free graphs. Consider now the class $\mathcal{C}$ containing those graphs G such that $\alpha(H) \geq \frac{|V(H)|}{3}$, for every induced subgraph H of G. Gyárfás et al. [82] showed that $f(x) = \lfloor \frac{8}{5}x \rfloor$ is the smallest θ-bounding function for the class $\mathcal{C}$. In particular, they proved the following:

Theorem 3.2.3 (Gyárfás et al. [82]). *If $G \in \mathcal{C}$, then $\theta(G) \leq \frac{8}{5}\alpha(G)$.*

By Brooks' Theorem, every subcubic graph different from K_4 belongs to $\mathcal{C}$ and so $f(x) = \lfloor \frac{8}{5}x \rfloor$ is a θ-bounding function for the class of subcubic graphs as well. On the other hand, Gyárfás et al. [83] provided evidence for the following meta-statement: the graphs for which the difference $\theta - \alpha$ is large are triangle-free. It would therefore be natural to expect that the ratio $\frac{\theta}{\alpha}$ is maximum for triangle-free graphs and Theorem 3.2.1 partially confirms this intuition.

Our proof of Theorem 3.2.1 in Section 3.2.1 is inspired by that of Theorem 3.2.3. The main idea is rather simple and it is based on the notion of θ-criticality, a graph G being *θ-critical* if $\theta(G - v) < \theta(G)$, for every $v \in V(G)$. First, we show that a minimum counterexample is connected and θ-critical. We then rely on the following result by Gallai (see [177] for a short proof and an extension):

Theorem 3.2.4 (Gallai [72]). *If v is any vertex of a connected θ-critical graph G, then G has a minimum-size clique cover in which v is the only isolated vertex. In particular, $\theta(G) \leq \frac{|V(G)|+1}{2}$.*

The final contradiction is then reached by using an appropriate lower bound for the independence number of a subcubic graph. Note that the following statement implies Theorem 3.2.1: for every subcubic graph G, either $\alpha(G) > \frac{|V(G)|}{3}$ or $\theta(G) < \frac{|V(G)|+1}{2}$. Moreover, Brooks' Theorem and Theorem 3.2.4 imply that Theorem 3.2.1 is equivalent to the fact that there is no subcubic connected θ-critical graph G with $\alpha(G) = \frac{|V(G)|}{3}$ and $\theta(G) = \frac{|V(G)|+1}{2}$.

Let us now consider the class of graphs with maximum degree at most 4. Joos [101] relaxed the degree condition in Theorem 3.2.2 and showed that $\theta(G) \leq \frac{7}{4}\alpha(G)$, for any triangle-free graph G with $\Delta(G) \leq 4$.

Theorem 3.2.5 (Joos [101]). *If G is a triangle-free graph with $\Delta(G) \leq 4$, then $\frac{7}{4}\alpha(G) + \alpha'(G) \geq |V(G)|$. Moreover, equality holds if and only if every component C of G has order 13, $\alpha(C) = 4$ and $\alpha'(C) = 6$.*

It would be tempting to extend Theorem 3.2.5 to the class of graphs with maximum degree 4 in the same way we extend Theorem 3.2.2 to the class of subcubic graphs. Unfortunately, the method adopted in the proof of Theorem 3.2.1 does not seem to be powerful enough for this purpose and the price we have to pay is a bigger θ-bounding function (see Theorem 3.2.8), likely to be far from the optimal.

We close Section 3.2.1 with an observation related to Theorems 3.2.1 and 3.2.2. We have seen in Theorem 3.2.2 that there are exactly two connected subcubic triangle-free graphs for which $\theta \leq \frac{3}{2}\alpha$ holds with equality: C_5 and G_{11}. Both of them clearly contain an induced C_5 and so it is natural to ask what happens by forbidding C_5. We show that if G is a subcubic (K_3, C_5)-free graph, then $\theta(G) \leq \frac{10}{7}\alpha(G)$.

In Section 3.2.2, we consider the problem of finding a clique cover of minimum size for graphs with bounded maximum degree and for planar graphs. The decision version of this well-known NP-complete problem is formulated as follows:

CLIQUE COVER	
Instance:	A graph G and a positive integer k.
Question:	Does $\theta(G) \leq k$ hold?

Since any subset of a clique is again a clique, CLIQUE COVER is equivalent to the following problem:

CLIQUE PARTITION	
Instance:	A graph G and a positive integer k.
Question:	Does there exist a partition of $V(G)$ into k disjoint cliques?

Moreover, CLIQUE PARTITION is clearly equivalent to the well-known COLOURING problem on the complement graph.

Cerioli et al. [31] studied CLIQUE COVER on planar graphs and on subclasses of subcubic graphs. In particular, they showed that CLIQUE COVER is NP-complete even for planar cubic graphs and that the optimization version is MAX SNP-hard for cubic graphs. Moreover, they asked whether the problem admits a PTAS for planar cubic graphs and conjectured that it has a polynomial-time approximation algorithm with a fixed ratio for graphs with bounded maximum degree. In Section 3.2.2, we answer both questions in the affirmative. We also provide some hardness results for subclasses of planar graphs and subcubic graphs.

3.2.1 θ-Bounding functions

We begin this section with a proof of Theorem 3.2.1. As mentioned in Section 3.2, our proof makes use of an appropriate lower bound for the independence number given by Harant et al. [84]. In order to state their result, we need the following definitions. A block of a graph is *difficult* if it is isomorphic to one of the four graphs depicted in Figure 3.2. Moreover, a connected graph is *bad* if its blocks are either difficult or are edges between difficult blocks. For a graph G, the number of bad components of G is denoted by $\lambda(G)$ and the maximum number of vertex-disjoint triangles of G is denoted by $t(G)$.

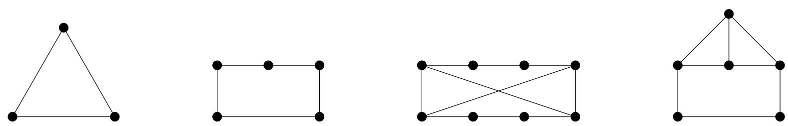


Figure 3.2: The difficult blocks.

Theorem 3.2.6 (Harant et al. [84]). *Every subcubic K_4-free graph G has an independent set of size at least $\frac{1}{7}(4n(G) - m(G) - \lambda(G) - t(G))$.*

We also require the notion of distance between sets of vertices of a graph. Given two subsets X and Y of $V(G)$, the *distance* from X to Y is the quantity $d(X,Y) = \min_{x\in X, y\in Y} d(x,y)$, i.e. it is the minimum length of a path between a vertex in X and a vertex in Y. With a slight abuse of notation, if T is a triangle, we write $d(T,Y)$ instead of $d(V(T),Y)$.

We can finally proceed to the proof of Theorem 3.2.1:

Proof of Theorem 3.2.1. Let us begin by showing that $\theta(G) \leq \frac{3}{2}\alpha(G)$, for any subcubic graph G. Suppose, by contradiction, that G is a counterexample with the minimum number of vertices. In the following, we deduce some structural properties of G and we show how they lead to a contradiction. Each claim is followed by a short proof.

Claim 2. *G is connected.*

Otherwise, G is the disjoint union of two non-empty graphs G_1 and G_2. By minimality, we have

$$\theta(G) = \theta(G_1) + \theta(G_2) \leq \frac{3}{2}(\alpha(G_1) + \alpha(G_2)) = \frac{3}{2}\alpha(G),$$

a contradiction. ■

Claim 3. *G is θ-critical.*

Indeed, suppose there exists a vertex $v \in V(G)$ such that $\theta(G) = \theta(G - v)$. By minimality, we have

$$\theta(G) = \theta(G - v) \leq \frac{3}{2}\alpha(G - v) \leq \frac{3}{2}\alpha(G),$$

a contradiction. ■

Claim 4. *G has minimum degree at least 2.*

Suppose there exists a 1-vertex u of G. By minimality, we have

$$\theta(G) \leq \theta(G - N[u]) + 1 \leq \frac{3}{2}\alpha(G - N[u]) + 1 \leq \frac{3}{2}(\alpha(G) - 1) + 1 < \frac{3}{2}\alpha(G),$$

a contradiction. ■

Claim 5. *G is 2-connected.*

Since every connected subcubic bridgeless graph is 2-connected, it is enough to show that G has no cut-edges. Therefore, suppose $e = u_1u_2$ is a cut-edge and let G_1 be the component of $G - e$ containing u_1 and $G_2 = G - V(G_1)$ (therefore, $u_2 \in V(G_2)$). Clearly, $\theta(G) \leq \theta(G_1) + \theta(G_2)$. If there exists $i \in \{1, 2\}$ such that a maximum independent set of G_i avoids u_i, then $\alpha(G) \geq \alpha(G_1) + \alpha(G_2)$ and so, by minimality,

$$\theta(G) \leq \theta(G_1) + \theta(G_2) \leq \frac{3}{2}\alpha(G_1) + \frac{3}{2}\alpha(G_2) \leq \frac{3}{2}\alpha(G),$$

a contradiction. Therefore, for each $i \in \{1, 2\}$, every maximum independent set of G_i contains u_i. This means that $\alpha(G_i - u_i) = \alpha(G_i) - 1$, for each $i \in \{1, 2\}$. Moreover, denoting by I_i a maximum independent set of G_i, we have that $I_1 \cup (I_2 \setminus \{u_2\})$ is an independent set of G and so $\alpha(G) \geq \alpha(G_1) + \alpha(G_2) - 1$. But then, by minimality,

$$\begin{aligned}
\theta(G) &\leq \theta(G_1 - u_1) + \theta(G_2 - u_2) + 1 \\
&\leq \frac{3}{2}\alpha(G_1 - u_1) + \frac{3}{2}\alpha(G_2 - u_2) + 1 \\
&= \frac{3}{2}(\alpha(G_1) - 1) + \frac{3}{2}(\alpha(G_2) - 1) + 1 \\
&= \frac{3}{2}(\alpha(G_1) + \alpha(G_2)) - 2 \\
&< \frac{3}{2}\alpha(G),
\end{aligned}$$

a contradiction. ■

Claim 6. *G does not contain a diamond.*

Suppose G contains a diamond and let u and v be its 2-vertices. Since G is connected and $\theta(K_4) = \alpha(K_4)$, we have $uv \notin E(G)$. Therefore, by minimality,

$$\theta(G) \leq \theta(G - N[u] - N[v]) + 3 \leq \frac{3}{2}\alpha(G - N[u] - N[v]) + 3 \leq \frac{3}{2}(\alpha(G) - 2) + 3 \leq \frac{3}{2}\alpha(G),$$

a contradiction. ■

Claim 7. *$d(u, T) \geq 4$, for any 2-vertex $u \in V(G)$ and any triangle $T \subseteq G$.*

Suppose first a triangle T contains a 2-vertex. By minimality, we have

$$\theta(G) \leq \theta(G - V(T)) + 1 \leq \frac{3}{2}\alpha(G - V(T)) + 1 \leq \frac{3}{2}(\alpha(G) - 1) + 1 < \frac{3}{2}\alpha(G),$$

a contradiction. Therefore, we have $d(u, T) \geq 1$, for any 2-vertex $u \in V(G)$ and any triangle $T \subseteq G$. Suppose now $d(u, T) = 1$, for a triangle $T \subseteq G$ and a 2-vertex $u \in V(G) \setminus V(T)$. This means T contains a vertex v such that $uv \in E(G)$ and let $v' \in V(T) \setminus \{v\}$. By Claim 6, we have $uv' \notin E(G)$ and so, by minimality,

$$\theta(G) \leq \theta(G - N[u] - N[v']) + 3 \leq \frac{3}{2}\alpha(G - N[u] - N[v']) + 3 \leq \frac{3}{2}(\alpha(G) - 2) + 3 \leq \frac{3}{2}\alpha(G),$$

a contradiction. Using a similar argument, it is an easy but tedious task to check the two remaining cases and we leave it to the interested reader. ■

Claim 8. *$d(u, v) \geq 3$, for any two distinct 2-vertices u and v of G.*

Suppose first there exist two adjacent 2-vertices u and v and let $u' \in N(u) \setminus \{v\}$ and $v' \in N(v) \setminus \{v\}$. By Claim 7, we have $u' \neq v'$. If there exists a vertex $w \in V(G)$ adjacent to both u' and v' then, by minimality,

$$\theta(G) \leq \theta(G - N[u] - N[w]) + 3 \leq \frac{3}{2}\alpha(G - N[u] - N[w]) + 3 \leq \frac{3}{2}(\alpha(G) - 2) + 3 \leq \frac{3}{2}\alpha(G),$$

a contradiction. Therefore, no vertex of G is adjacent to both u' and v'.

Consider now the graph G' obtained from G by deleting $\{u, v\}$ and by adding, if necessary, the edge $u'v'$. The graph G' is clearly simple and subcubic. Since a maximum independent set I' of G' is also an independent set of $G - \{u, v\}$ and I' contains at most one of the vertices u' and v', we have $\alpha(G) \geq \alpha(G') + 1$. Moreover, we claim that $\theta(G) \leq \theta(G') + 1$. Indeed, consider a minimum clique cover C' of G'. If no clique in C' contains $\{u', v'\}$, then $C' \cup \{u, v\}$ is a clique cover of G of size $\theta(G') + 1$. On the other hand, by the paragraph above, if a clique in C' contains $\{u', v'\}$, then it must be of size 2. Therefore, $(C' \setminus \{u', v'\}) \cup \{u', u\} \cup \{v', v\}$ is a clique cover of G of size $\theta(G') + 1$ and we have established our claim. But then, again by minimality, we have

$$\theta(G) \leq \theta(G') + 1 \leq \frac{3}{2}\alpha(G') + 1 \leq \frac{3}{2}(\alpha(G) - 1) + 1 < \frac{3}{2}\alpha(G),$$

a contradiction.

Suppose now there exist two 2-vertices u and v such that $d(u,v) = 2$. Since $uv \notin E(G)$ then, by minimality, we have

$$\theta(G) \le \theta(G - N[u] - N[v]) + 3 \le \frac{3}{2}\alpha(G - N[u] - N[v]) + 3 \le \frac{3}{2}(\alpha(G) - 2) + 3 \le \frac{3}{2}\alpha(G),$$

a contradiction. ■

Claim 9. *$d(T_1, T_2) \ge 3$, for any two distinct triangles T_1 and T_2 of G.*

By Claim 6 and since G is subcubic, we have $d(T_1, T_2) \ge 1$. Suppose first there exist two triangles T_1 and T_2 at distance 1 and let $u_1 \in V(T_1)$ and $u_2 \in V(T_2)$ be such that $d(u_1, u_2) = 1$. Moreover, consider a vertex $u_2' \in V(T_2) \setminus \{u_2\}$. By minimality,

$$\theta(G) \le \theta(G - N[u_1] - N[u_2']) + 3 \le \frac{3}{2}\alpha(G - N[u_1] - N[u_2']) + 3 \le \frac{3}{2}(\alpha(G) - 2) + 3 \le \frac{3}{2}\alpha(G),$$

a contradiction. The remaining case can be treated similarly, leading again to a contradiction. ■

Claim 10. *Each cycle of G on four vertices contains only cubic vertices.*

Indeed, suppose there exists a cycle $C \subseteq G$ on four vertices containing a 2-vertex u of G. Let $v \in V(C)$ be such that $d(u,v) = 2$. By minimality,

$$\theta(G) \le \theta(G - N[u] - N[v]) + 3 \le \frac{3}{2}\alpha(G - N[u] - N[v]) + 3 \le \frac{3}{2}(\alpha(G) - 2) + 3 \le \frac{3}{2}\alpha(G),$$

a contradiction. ■

Claim 11. *$n_2(G) > 0$.*

Suppose this is not the case. By Claim 4 and Claim 5, G is a cubic bridgeless graph and the well-known Petersen's Theorem implies it has a perfect matching. But then $\theta(G) \le \frac{n(G)}{2} \le \frac{3}{2}\alpha(G)$, a contradiction. ■

Claim 12. *$6t(G) \le n_3(G) - 6$.*

Let u be a 2-vertex of G. We first show that the set $S = \{v \in V(G) : d(v,u) = 2\}$ of vertices at distance 2 from u has size 4. Indeed, by Claim 8, the neighbours u' and u'' of u are cubic vertices and, by Claim 7, $u'u'' \notin E(G)$. Moreover, by Claim 10, no neighbour of u' different from u is also a neighbour of u'' and so $|S| = 4$. Note that, by Claim 8, each $v \in S$ is a cubic vertex.

Consider now the triangles of G. By Claim 7 and Claim 6, each vertex of a triangle T has a neighbour not in T and any two such neighbours are distinct. Moreover, by Claim 9, the set of neighbours of T_1 does not intersect the set of neighbours of T_2, for any two (vertex-disjoint) triangles T_1 and T_2 of G. Finally, by Claim 7, no vertex in $S \cup \{u', u''\}$ belongs to a triangle or is a neighbour of a triangle and each neighbour of a triangle is a cubic vertex. Therefore, we have $6t(G) \le n_3(G) - 6$. ■

We are finally in a position to conclude our proof. By Claim 5, G is 2-connected and since no graph in Figure 3.2 is a counterexample, we have $\lambda(G) = 0$. Therefore, by Theorem 3.2.6 and recalling that $n(G) = n_3(G) + n_2(G)$ and $6t(G) \leq n_3(G) - 6$, we get

$$\begin{aligned}
\alpha(G) &\geq \left\lceil \frac{1}{7}\Big(4n(G) - m(G) - t(G)\Big)\right\rceil \\
&= \left\lceil \frac{1}{7}\Big(4n_3(G) + 4n_2(G)\Big) - \frac{1}{7}\Big(\frac{3}{2}n_3(G) + n_2(G)\Big) - \frac{1}{7}t(G)\right\rceil \\
&\geq \left\lceil \frac{1}{7}\Big(4n_3(G) + 4n_2(G)\Big) - \frac{1}{7}\Big(\frac{3}{2}n_3(G) + n_2(G)\Big) - \frac{1}{42}(n_3(G) - 6)\right\rceil \\
&= \left\lceil \frac{1}{3}n_3(G) + \frac{3}{7}n_2(G) + \frac{1}{7}\right\rceil. \qquad (3.2)
\end{aligned}$$

We now claim that

$$\left\lceil \frac{1}{3}n_3(G) + \frac{3}{7}n_2(G) + \frac{1}{7}\right\rceil \geq \frac{1}{3}(n_3(G) + n_2(G) + 1). \qquad (3.3)$$

This can be easily seen if $n_2(G) \geq 2$. Therefore, suppose $n_2(G) = 1$ and let $n_3(G) = 3k + a$, for some integer $0 \leq a \leq 2$. Inequality (3.3) is then equivalent to

$$\left\lceil \frac{7a + 12}{21}\right\rceil \geq \frac{a+2}{3},$$

which clearly holds for $0 \leq a \leq 2$.

On the other hand, by Claim 2 and Claim 3, G is a connected θ-critical graph and so, by Theorem 3.2.4, we have $\theta(G) \leq \frac{n(G)+1}{2}$. Therefore, combining this with (3.2) and (3.3), we have

$$\frac{3}{2}\alpha(G) \geq \frac{3}{2}\left\lceil \frac{1}{3}n_3(G) + \frac{3}{7}n_2(G) + \frac{1}{7}\right\rceil \geq \frac{1}{2}(n_3(G) + n_2(G) + 1) = \frac{1}{2}(n(G) + 1) \geq \theta(G),$$

a contradiction. This concludes the proof of the first statement in Theorem 3.2.1.

As for the second statement, we need to show that, for each integer $x \geq 0$, there exists a subcubic graph G such that $\alpha(G) = x$ and $\theta(G) = \lfloor \frac{3}{2}\alpha(G)\rfloor$. If x is even, we construct G as the disjoint union of $\frac{x}{2}$ copies of C_5. On the other hand, if x is odd, it is enough to construct G as the disjoint union of $\lfloor \frac{x}{2}\rfloor$ copies of C_5 together with an isolated vertex. □

As mentioned in Section 3.2, the ratio $\frac{\theta}{\alpha}$ should be maximum for triangle-free graphs and so we propose the following:

Conjecture 3.2.7. *If G is a subcubic graph, then $\theta(G) = \frac{3}{2}\alpha(G)$ if and only if every component of G is either C_5 or G_{11}.*

Let us now consider the class of graphs with maximum degree at most 4. Joos [101] relaxed the degree condition in Theorem 3.2.2 and showed that $\theta(G) \leq \frac{7}{4}\alpha(G)$, for any triangle-free graph G with $\Delta(G) \leq 4$. Following the intuition expressed in Section 3.2, it would be natural to expect that the previous inequality holds for any graph with maximum degree at most 4. Unfortunately, the method adopted in the proof of Theorem 3.2.1 gives an upper bound for the ratio $\frac{\theta(G)}{\alpha(G)}$ which gets larger as the size of a maximum clique of G increases. Since we believe the maximum is attained by triangle-free graphs, we present only the bound for K_4-free graphs, which is already substantially larger than the expected $\frac{7}{4}$.

Theorem 3.2.8. *If G is a K_4-free graph with maximum degree at most 4, then $\theta(G) \le \frac{193}{98}\alpha(G)$.*

As mentioned above, our proof of Theorem 3.2.8 resembles that of Theorem 3.2.1: a counterexample G with minimum order must be connected and θ-critical; Theorem 3.2.4 then guarantees the existence of a clique cover of size at most $\frac{n(G)+1}{2}$ and by using an appropriate lower bound on the independence number, we derive a contradiction, assuming the value of $\alpha(G)$ is large enough. On the other hand, if the value of $\alpha(G)$ is small, then $n(G)$ is small and it is useful to consider several results on the *covering gap* (the difference between the minimum size of a clique cover and the maximum size of an independent set) of small graphs, as stated in the following theorem:

Theorem 3.2.9 (Gyárfás et al. [83]). *The graph G_{13} in Figure 3.1 is the only graph on 13 vertices with covering gap 3. Moreover, for any graph G, the following hold:*

- *If $n(G) \le 22$, then $\theta(G) - \alpha(G) \le 5$;*
- *If $n(G) \le 19$, then $\theta(G) - \alpha(G) \le 4$;*
- *If $n(G) \le 16$, then $\theta(G) - \alpha(G) \le 3$;*
- *If $n(G) \le 12$, then $\theta(G) - \alpha(G) \le 2$;*
- *If $n(G) \le 9$, then $\theta(G) - \alpha(G) \le 1$.*

We can finally proceed to the proof of Theorem 3.2.8.

Proof of Theorem 3.2.8. Suppose, by contradiction, that G is a counterexample with the minimum number of vertices. As in the proof of Theorem 3.2.1, it is easy to see that G is connected and θ-critical. On the other hand, Locke and Lou [129] showed that, for any connected K_4-free graph G with $\Delta(G) \le 4$, we have $\alpha(G) \ge \frac{7n(G)-4}{26}$. Combining this with Theorem 3.2.4, and if $\alpha(G) \ge 7$, we get

$$\theta(G) \le \frac{n(G)+1}{2} \le \frac{26\alpha(G)+11}{14} \le \frac{193}{98}\alpha(G).$$

For the small values of $\alpha(G)$, we rely on Theorem 3.2.9. If $\alpha(G) = 6$, then $n(G) \le 22$ and so $\theta(G) \le 11$. If $\alpha(G) = 5$, then $n(G) \le 19$ and so $\theta(G) \le 9$. If $\alpha(G) = 4$, then $n(G) \le 15$ and so $\theta(G) \le 7$. If $\alpha(G) = 3$, then $n(G) \le 11$ and so $\theta(G) \le 5$. If $\alpha(G) = 2$, then $n(G) \le 8$ and so $\theta(G) \le 3$. Finally, if $\alpha(G) \le 1$, then G is complete and $\theta(G) = \alpha(G)$. It immediately follows that, for all the values of $\alpha(G)$, we have $\theta(G) \le \frac{193}{98}\alpha(G)$, a contradiction. □

As remarked before, we believe $\frac{7}{4}$ should be the optimal constant:

Conjecture 3.2.10. *If G is a graph with maximum degree at most 4, then $\theta(G) \le \frac{7}{4}\alpha(G)$, with equality if and only if every component of G is G_{13}.*

It is worth noticing that a slight modification of the proof of Theorem 3.2.8 gives a short proof of Theorem 3.2.5:

Proof of Theorem 3.2.5. Suppose, by contradiction, that G is a counterexample with the minimum number of vertices. As we have seen above, G is a connected θ-critical graph. On the other hand, Fraughnaugh Jones [67] showed that $\alpha(G) \geq \frac{4}{13}n(G)$, for any triangle-free graph G with $\Delta(G) \leq 4$. Combining this with Theorem 3.2.4, and if $\alpha(G) > 4$, we get

$$\theta(G) \leq \frac{n(G)+1}{2} \leq \frac{13\alpha(G)+4}{8} < \frac{7}{4}\alpha(G).$$

For the remaining values of $\alpha(G)$, we use again Theorem 3.2.9. If $\alpha(G) = 4$, then $n(G) \leq 13$ and so $\theta(G) \leq 7$. If $\alpha(G) = 3$, then $n(G) \leq 9$ and so $\theta(G) \leq 4$. If $\alpha(G) = 2$, then $n(G) \leq 6$ and so $\theta(G) \leq 3$. Finally, if $\alpha(G) \leq 1$, then G is complete and $\theta(G) = \alpha(G)$. It immediately follows that $\theta(G) \leq \frac{7}{4}\alpha(G)$, a contradiction.

Note that, if G is connected and θ-critical, then equality holds only if $\alpha(G) = 4$ and $n(G) = 13$. But the graph G_{13} in Figure 3.1 (also known as the $(3,5)$-Ramsey graph) is the only graph G such that $\omega(G) = 2$, $\omega(\overline{G}) = \alpha(G) = 4$ and $n(G) = 13$ (see [160]). On the other hand, it is easy to see that equality cannot hold if G is not θ-critical. □

We conclude this section with two observations on the subcubic case we treated above. Theorem 3.2.1 (or, alternatively, Theorem 3.2.2) implies that $f(x) = \lfloor \frac{3}{2}x \rfloor$ is a θ-bounding function for the class of subcubic triangle-free graphs. This class is equivalent to *Free*$(K_{1,4}, K_3)$ and the same function is θ-bounding for the superclass *Free*$(K_{1,4}, \text{paw})$:

Lemma 3.2.11. *If G is a $(K_{1,4}, paw)$-free graph, then $\theta(G) \leq \frac{3}{2}\alpha(G)$. Equality holds if and only if every component of G is either C_5 or G_{11}.*

Proof. Clearly, it is enough to show the assertion for a connected G. A well-known result by Olariu [150] states that a connected graph is paw-free if and only if it is either triangle-free or complete multipartite. Suppose first G is triangle-free. Since G is $K_{1,4}$-free as well, it must be subcubic and so, by Theorem 3.2.2, we have that $\theta(G) \leq \frac{3}{2}\alpha(G)$. Moreover, we know equality holds if and only if G is either C_5 or G_{11}.

Finally, if G is a complete multipartite graph, then $\overline{G}$ is a disjoint union of cliques and so $\theta(G) = \chi(\overline{G}) = \omega(\overline{G}) = \alpha(G)$. □

We have seen in Theorem 3.2.2 that there are exactly two connected subcubic triangle-free graphs for which $\theta \leq \frac{3}{2}\alpha$ holds with equality: the 5-cycle C_5 and the graph G_{11}. Both of them clearly contain an induced C_5 and so it is natural to ask whether forbidding this graph leads to a smaller bounding function. It appears this is indeed the case, as shown by the following:

Theorem 3.2.12. *If G is a subcubic (K_3, C_5)-free graph, then $\theta(G) \leq \frac{10}{7}\alpha(G)$.*

As we have seen throughout the section, our approach for proving such statements ultimately relies on a good lower bound for the independence number. For a triangle-free graph G, Theorem 3.2.6 implies Staton's bound $\alpha(G) \geq \frac{5}{14}n(G)$ [176] and Heckman [92] showed that there are exactly two graphs attaining equality. They both have 14 vertices and contain an induced C_5 (one such graph is the so-called generalized Petersen graph $P(7,2)$). Fraughnaugh and Locke [66] showed that, for graphs with more than 14 vertices, a better lower bound is possible: $\alpha(G) \geq \frac{11}{30}n(G) - \frac{2}{15}$, for any connected subcubic triangle-free graph. Moreover, they showed that if G is not cubic and does not belong to a certain family, the previous bound can be further improved. Let us now define this special family.

We denote by $\mathcal{F}_{11}$ the class of graphs obtained by the following construction. Given any tree T with maximum degree at most four, we replace each vertex of T of degree at least two by a copy of G_8 (see Figure 3.3) and each vertex of degree one by a copy of either G_8 or G_{11} (see Figure 3.1). For each vertex v of T replaced by $G \in \{G_8, G_{11}\}$, the edges of T incident to v become incident to vertices of degree two in G, at most one such edge to each vertex of degree two in G. Clearly, every graph $G \in \mathcal{F}_{11}$ thus obtained is subcubic and triangle-free and it is not difficult to see that $m(G) - 7n(G) + 15\alpha(G) = -1$.

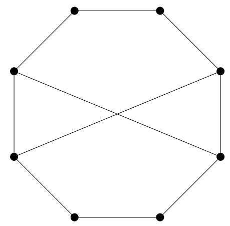


Figure 3.3: The graph G_8.

Denoting by $\gamma(G)$ the quantity $m(G) - 7n(G) + 15\alpha(G)$, Fraughnaugh and Locke [66] showed the following:

Theorem 3.2.13 (Fraughnaugh and Locke [66]). *Let G be a connected subcubic triangle-free graph. If G is cubic, then $\gamma(G) \geq -2$. If $G \in \mathcal{F}_{11}$, then $\gamma(G) = -1$. Finally, if G is not cubic and $G \notin \mathcal{F}_{11}$, then $\gamma(G) \geq 0$.*

The following corollary is immediate:

Corollary 3.2.14. *Let G be a connected subcubic triangle-free graph. If G is cubic, then $\alpha(G) \geq \frac{11n(G)-4}{30}$, while if G is not cubic and $G \notin \mathcal{F}_{11}$, then $\alpha(G) \geq \frac{11n(G)+1}{30}$.*

Note that if G is C_5-free, then $G \notin \mathcal{F}_{11}$, as both G_8 and G_{11} contain induced copies of C_5. We can finally proceed to the proof of Theorem 3.2.12.

Proof of Theorem 3.2.12. Suppose, by contradiction, that G is a counterexample with the minimum number of vertices. As we have repeatedly seen in the previous paragraphs, G is connected and θ-critical. We now show it is in fact 2-connected. The proof is almost identical to the one of Claim 5 in Theorem 3.2.1.

Claim 13. *G is 2-connected.*

Since every connected subcubic bridgeless graph is 2-connected, it is enough to show that G has no cut-edges. Therefore, suppose $e = u_1u_2$ is a cut-edge and let G_1 be the component of $G - e$ containing u_1 and $G_2 = G - V(G_1)$ (therefore, $u_2 \in V(G_2)$). Since G is triangle-free, we have that $G - e$ is triangle-free and C_5-free as well. Moreover, $\theta(G) \leq \theta(G_1) + \theta(G_2)$. If there exists $i \in \{1,2\}$ such that a maximum independent set of G_i avoids u_i, then $\alpha(G) \geq \alpha(G_1) + \alpha(G_2)$ and so, by minimality,

$$\theta(G) \leq \theta(G_1) + \theta(G_2) \leq \frac{10}{7}\alpha(G_1) + \frac{10}{7}\alpha(G_2) \leq \frac{10}{7}\alpha(G),$$

a contradiction. Therefore, for each $i \in \{1,2\}$, every maximum independent set of G_i contains u_i. This means that $\alpha(G_i - u_i) = \alpha(G_i) - 1$, for each $i \in \{1,2\}$. Moreover, denoting by I_i a maximum independent set of G_i, we have that $I_1 \cup (I_2 \setminus \{u_2\})$ is an independent set of G and so $\alpha(G) \geq \alpha(G_1) + \alpha(G_2) - 1$. But then, by minimality,

$$\begin{aligned}
\theta(G) &\leq \theta(G_1 - u_1) + \theta(G_2 - u_2) + 1 \\
&\leq \frac{10}{7}\alpha(G_1 - u_1) + \frac{10}{7}\alpha(G_2 - u_2) + 1 \\
&= \frac{10}{7}(\alpha(G_1) - 1) + \frac{10}{7}(\alpha(G_2) - 1) + 1 \\
&= \frac{10}{7}(\alpha(G_1) + \alpha(G_2)) - \frac{13}{7} \\
&< \frac{10}{7}\alpha(G),
\end{aligned}$$

a contradiction. ■

Claim 14. *G is cubic.*

Suppose it is not. Since G is C_5-free, we have that $G \notin \mathcal{F}_{11}$ and so, by Corollary 3.2.14, we have $\alpha(G) \geq \frac{11n(G)+1}{30}$. Combining this with Theorem 3.2.4 and if $\alpha(G) \geq 7$, we obtain

$$\theta(G) \leq \frac{n(G)+1}{2} \leq \frac{30\alpha(G)+10}{22} \leq \frac{10}{7}\alpha(G).$$

For the remaining values of $\alpha(G)$, we rely on Theorem 3.2.9:

- If $\alpha(G) = 6$, then $n(G) \leq 16$ and $\theta(G) \leq 9$. Suppose now that $\theta(G) = 9$. This means that $\theta(G) = \frac{3}{2}\alpha(G)$ and so, by Theorem 3.2.2, we have $G \in \{C_5, G_{11}\}$, a contradiction to the fact that G is C_5-free. Therefore, $\theta(G) \leq 8 < \frac{10}{7}\alpha(G)$.

- If $\alpha(G) = 5$, then $n(G) \leq 13$ and $\theta(G) \leq 8$. Since G_{13} is the only graph on 13 vertices with covering gap 3 and since it contains an induced C_5, we have that $\theta(G) \leq 7 < \frac{10}{7}\alpha(G)$.

- If $\alpha(G) = 4$, then $n(G) \leq 10$ and $\theta(G) \leq 6$. By the same argument as in the case $\alpha(G) = 6$, we have that $\theta(G) \leq 5 < \frac{10}{7}\alpha(G)$.

- If $\alpha(G) = 3$, then $n(G) \leq 8$ and so $\theta(G) \leq 4 < \frac{10}{7}\alpha(G)$.

- If $\alpha(G) = 2$, then $n(G) \leq 5$ and $\theta(G) \leq 3$. As above, by Theorem 3.2.2, we have that $\theta(G) = 2$.

- Finally, if $\alpha(G) \leq 1$, then G is complete and again $\theta(G) = \alpha(G)$.

In all the cases we have $\theta(G) \leq \frac{10}{7}\alpha(G)$, a contradiction. ■

We are now in a position to conclude our proof. Since G is cubic and 2-connected, Petersen's Theorem implies it has a perfect matching. Moreover, by Corollary 3.2.14, we have that $\alpha(G) \geq \frac{11n(G)-4}{30}$. Therefore, if $\alpha(G) \geq 3$, we obtain

$$\theta(G) = \frac{n(G)}{2} \leq \frac{30\alpha(G)+4}{22} < \frac{10}{7}\alpha(G).$$

If $\alpha(G) = 2$, then $n(G) \leq 5$ and since $G \not\cong C_5$, we have $\theta(G) = 2$. Finally, if $\alpha(G) \leq 1$, then G is complete and $\theta(G) = \alpha(G)$. In all the cases we have $\theta(G) < \frac{10}{7}\alpha(G)$, a contradiction. This concludes the proof. □

We do not have any example of a graph attaining equality in Theorem 3.2.12. In fact, considering C_7, we believe the optimal constant is $\frac{4}{3}$:

Conjecture 3.2.15. *If G is a subcubic (K_3, C_5)-free graph, then $\theta(G) \leq \frac{4}{3}\alpha(G)$.*

3.2.2 Algorithmic aspects of Clique Cover

CLIQUE COVER is a well-known NP-complete problem polynomially equivalent to COLOURING. Cerioli et al. [31] showed that the optimization version of CLIQUE COVER is MAX SNP-hard for cubic graphs. We begin this section with an inapproximability gap result for CLIQUE COVER restricted to subcubic line graphs:

Theorem 3.2.16. CLIQUE COVER *is not approximable within $\frac{391}{390}$, unless* P = NP, *even when restricted to line graphs of 2-subdivisions of cubic triangle-free graphs.*

Proof. Chlebík and Chlebíková [38] showed that it is NP-hard to approximate VERTEX COVER within $\frac{391}{390}$, even for 2-subdivisions of cubic graphs: they construct a gap-preserving reduction from VERTEX COVER restricted to cubic graphs (for which they provided an NP-hard gap in [37]) just by a 2-subdivision of the input graph. Since their NP-hard gap result for VERTEX COVER in [37] holds for cubic triangle-free graphs as well, it follows that it is NP-hard to approximate VERTEX COVER within $\frac{391}{390}$, even for 2-subdivisions of cubic triangle-free graphs. Therefore, given a 2-subdivision of a cubic triangle-free graph G, we simply construct its line graph $L(G)$. Since $\theta(L(G)) = \beta(G)$ (Fact 2.3.2), the conclusion immediately follows. □

We now turn to the decision version of CLIQUE COVER when restricted to planar line graphs with maximum degree at most 4.

Theorem 3.2.17. CLIQUE COVER *is* NP-*complete even for line graphs of 2-subdivisions of planar cubic triangle-free graphs.*

Proof. Note that, given a graph G and a 2-subdivision G' of G, we have $\beta(G') = \beta(G) + |E(G)|$ (see [38] for the short proof). Since VERTEX COVER is NP-hard for planar cubic triangle-free graphs [179], it is NP-hard for 2-subdivisions of planar cubic triangle-free graphs and, by Fact 2.3.2, we can easily obtain the claimed NP-hardness of CLIQUE COVER. □

Cerioli et al. [31] showed that CLIQUE COVER is NP-hard for planar cubic graphs. Not surprisingly, it remains NP-hard even for planar 4-regular graphs, as implied by the following theorem. The proof immediately follows by the result in [179] mentioned above.

Theorem 3.2.18. CLIQUE COVER *remains* NP-*complete even when restricted to line graphs of planar cubic triangle-free graphs.*

It is therefore natural to look for approximation algorithms for CLIQUE COVER when restricted to graphs having bounded maximum degree or to planar graphs. Cerioli et al. [31] showed that CLIQUE COVER admits a polynomial-time $\frac{5}{4}$-approximation algorithm for subcubic graphs and they conjectured it has a polynomial-time approximation algorithm with a fixed ratio for graphs with bounded maximum degree. This can be easily verified once we notice the close relation between θ-boundedness of a certain class of graphs and approximation algorithms for CLIQUE COVER for that class. Indeed, since $\alpha(G) \leq \theta(G)$, if we could show "algorithmically" that a class of graphs is θ-bounded by a linear function, we would obtain a constant-factor approximation algorithm for CLIQUE COVER. This algorithmic feature of θ-boundedness was first observed by Gyárfás [81]. Note that it is not clear whether the proofs of Theorems 3.2.1 and 3.2.8 can be turned into "algorithmic" ones. Nevertheless, the following holds:

Theorem 3.2.19. CLIQUE COVER *admits a linear-time k-approximation algorithm for graphs with maximum degree at most k.*

Proof. Consider the following greedy algorithm: first, find a maximal independent set I of the input graph G and set $C = \varnothing$; then, for each $v \in I$, add the edges incident to v to the set C and return C. Clearly, the algorithm runs in linear time and it returns a clique cover of G. Moreover, we have

$$|C| \leq k|I| \leq k\alpha(G) \leq k\theta(G),$$

and so this greedy algorithm is indeed a k-approximation algorithm for graphs with maximum degree at most k. □

Theorem 3.2.16 shows that CLIQUE COVER admits no PTAS, even for subcubic graphs. Cerioli et al. [31] asked whether this could be possible in the special case of planar cubic graphs. In the rest of this section, using Baker's well-known technique [18], we show that CLIQUE COVER indeed admits a PTAS even for planar graphs. The idea of Baker's technique is the following: partition the planar graph into k-outerplanar graphs, solve the problem optimally for each k-outerplanar graph and finally show that the union of these solutions is in fact a "near-optimal solution" for the original graph. As it stands, k-outerplanar graphs play a key role in the reasoning and so let us recall their definition. Given a planar graph G and a fixed planar drawing Γ of G, we define L_1 to be the set of vertices incident to the outer face and, for $i > 1$, L_i is defined recursively as the set of vertices on the outer face of the planar drawing obtained by deleting the vertices in $\bigcup_{j=1}^{i-1} L_j$. We call L_i the i-th layer of the drawing Γ and a graph is k-outerplanar if it has a planar drawing with at most k layers.

Bodlaender [26, 27] showed that k-outerplanar graphs have tree-width at most $3k - 1$. Moreover, using dynamic programming, it is possible to determine $\theta(G)$ in polynomial time for any graph G of bounded tree-width (see, e.g., [80]). In fact, many other problems are solvable in polynomial time for graphs of bounded tree-width (see Chapter 1) and this led

Baker [18] to design the mentioned technique, which allows a PTAS for some of them (for example, INDEPENDENT SET and VERTEX COVER), when restricted to planar graphs. From the sketch of the technique we gave above, it should be clear that the problems which can be treated are those for which the local solutions can be combined into a global solution. We now show that CLIQUE COVER is indeed one of them:

Theorem 3.2.20. CLIQUE COVER *admits a PTAS for planar graphs.*

Proof. Given a planar drawing Γ of the input graph G, we construct the layers L_i as defined above. Note that the neighbours of a vertex $v_i \in L_i$ must be in $L_{i-1} \cup L_i \cup L_{i+1}$. Finally, given $\varepsilon > 0$, we set $k = \lceil \frac{2}{\varepsilon} \rceil$.

A *slice* G_{ij} is an induced subgraph defined as follows. For $1 \le i \le k$, we denote by G_{i0} the subgraph of G induced by the vertices which belong to the consecutive layers between the first and the i-th. Moreover, for a fixed $1 \le i \le k$ and a $j \ge 1$, we denote by G_{ij} the subgraph of G induced by the vertices which belong to the k consecutive layers whose indices range between $(j-1)(k-1)+i$ and $j(k-1)+i$ (note that j runs until each vertex of G belongs to at least one G_{ij}). By definition, each slice is k-outerplanar and so we can determine in polynomial time a minimum-size clique cover C_{ij} of G_{ij}, for each $1 \le i \le k$ and $j \ge 0$. Finally, for each i, we set $C_i = \bigcup_{j \ge 0} C_{ij}$. By construction, $\bigcup_{j\ge 0} V(G_{ij}) = V(G)$ and so each C_i is a clique cover of G. We return the one with minimum size.

Let now Q denote a minimum-size clique cover of G and, for $0 \le i \le k-1$, denote by Q_i the set of cliques in Q which contain at least one vertex in $\bigcup_{j \equiv i \pmod{k}} L_j$. Clearly, $\bigcup Q_i = Q$ and each clique in Q belongs to at most two distinct Q_i's. But then there exists an index ℓ such that

$$|Q_\ell| \le \frac{2}{k}|Q| \le \varepsilon |Q|.$$

Let $Q_{\ell j}$ denotes the set of cliques in Q containing at least one vertex in $V(G_{\ell j})$ (if $\ell = 0$, we set $G_{\ell j} = G_{kj}$). Since $C_{\ell j}$ is a minimum-size clique cover of $G_{\ell j}$, then $|C_{\ell j}| \le |Q_{\ell j}|$, for each $j \ge 0$. Consider now the sum $\sum_j |Q_{\ell j}|$. Each clique is counted exactly once, except those which contain vertices from layers L_j with $j \equiv \ell \pmod{k}$ (i.e. those in Q_ℓ), which are counted exactly twice. But then $\sum_j |Q_{\ell j}| = |Q| + |Q_\ell|$. Summarizing, we have

$$|C_\ell| \le \sum_j |C_{\ell j}| \le \sum_j |Q_{\ell j}| = |Q| + |Q_\ell| \le |Q| + \varepsilon|Q| = (1+\varepsilon)|Q|.$$

Therefore, the algorithm above is a polynomial-time approximation scheme. □

3.3 On Tuza's Conjecture

What is the minimum number of edges of a graph whose deletion results in a triangle-free graph? An obvious obstruction to a small set of edges meeting all the triangles is the presence of a large family of edge-disjoint triangles. On the other hand, deleting the edge set of a family of edge-disjoint triangles of maximum size results in a triangle-free graph. In Section 3.1, we have phrased this approximate Min-Max Property in terms of the triangle hypergraphs. For our purposes, it is convenient to stick to the underlying graph and we recall that, for a graph G, we denote by $\tau'_\Delta(G)$ the transversal number of the triangle hypergraph of G and by $\nu'_\Delta(G)$ its packing number. Therefore, Tuza's Conjecture can be rewritten as follows:

Conjecture 3.3.1 (Tuza's Conjecture [178]). *For any graph G, we have $\tau'_\Delta(G) \leq 2\nu'_\Delta(G)$.*

Note that, in this section, a *triangle-transversal* of G is intended to be a transversal of the triangle hypergraph of G, i.e. a subset of $E(G)$ whose deletion results in a triangle-free graph.

Despite having received considerable attention, Conjecture 3.3.1 is still open. To date, the best non-trivial bound is $\tau'_\Delta(G) \leq (3 - \frac{3}{23})\nu'_\Delta(G)$, as shown by Haxell [89]. Moreover, several graph classes for which it holds are known. Tuza [178] proved it for planar graphs and for "dense" graphs, specifically for graphs on n vertices and with at least $\frac{7}{16}n^2$ edges. It is well-known that every graph G has a bipartite subgraph with at least $\frac{|E(G)|}{2}$ edges (see, e.g., [182]). Since the complement of this edge set is clearly a triangle-transversal of G, we have that Tuza's Conjecture holds if G has many edge-disjoint triangles, more precisely at least $\frac{|E(G)|}{4}$. We now present a proof of Tuza's Conjecture for complete graphs, due to Sebő [172], and which uses exactly this reasoning.

Theorem 3.3.2. *For any integer n, we have $\tau'_\Delta(K_n) \leq 2\nu'_\Delta(K_n)$.*

Proof. Let $K_n = (V, E)$ and $t = \lfloor \frac{n}{2} \rfloor$. Moreover, let $A = \{a_i : i \in \{0, 1, \ldots, t-1\}\}$ and $B = \{b_i : i \in \{0, 1, \ldots, t-1\}\}$. In order to highlight the fact that for some values of n the constant 2 can be improved [2], we distinguish several cases according to the parity of n. Note that, in the following, addition of indices is considered modulo t.

Case 1: n is even.
We have that $n = 2t$ and we may assume $V = A \cup B$. Clearly, the triangles in the family

$$T = \{a_i a_j b_{i+j} : i \neq j \text{ and } \{i, j\} \subseteq \{0, \ldots, t-1\}\}$$

are edge-disjoint and $|T| = \frac{t(t-1)}{2}$. Moreover, the set of edges $E \setminus [A, B]$ meets all the triangles of K_n and it has size $t(t-1)$. Therefore, $\tau'_\Delta(K_n) \leq 2\nu'_\Delta(K_n)$.

Case 2: $n \equiv 3 \pmod 4$, i.e. t is odd.
We have that $n = 2t + 1$ and we may assume $V = A \cup B \cup \{v\}$, for some $v \notin A \cup B$. It is easy to see that

$$T_1 = \{a_i a_j b_{i+j} : i \neq j \text{ and } \{i, j\} \subseteq \{0, \ldots, t-1\}\} \quad \text{and} \quad T_2 = \{v a_i b_{2i} : i \in \{0, \ldots, t-1\}\}$$

are families of edge-disjoint triangles and that their union is again a family of edge-disjoint triangles. Clearly, $|T_1 \cup T_2| = \frac{t(t+1)}{2}$. Moreover, the set of edges $E \setminus [A \cup \{v\}, B]$ meets all the triangles of K_n and it has size t^2. Therefore, we have $\tau'_\Delta(K_n) \leq t^2 < 2\frac{t(t+1)}{2} \leq 2\nu'_\Delta(K_n)$.

Case 3: $n \equiv 1 \pmod 4$, i.e. t is even.
As above, we have that $n = 2t + 1$ and $V = A \cup B \cup \{v\}$.

$$T_1 = \{a_i a_j b_{i+j} : i \neq j \text{ and } \{i, j\} \subseteq \{0, \ldots, t-1\}\} \quad \text{and} \quad T_2 = \{v a_i b_{2i} : i \in \{0, \ldots, \frac{t}{2} - 1\}\}$$

are families of edge-disjoint triangles and their union is again a family of edge-disjoint triangles. Clearly, $|T_1 \cup T_2| = \frac{t^2}{2}$. Moreover, $E \setminus [A \cup \{v\}, B]$ meets all the triangles of K_n and its size is t^2. Therefore, we have $\tau'_\Delta(K_n) \leq 2\nu'_\Delta(K_n)$.

This concludes the proof. □

[2] In fact, Tuza [178] showed that $\lim_{n\to\infty} \tau'_\Delta(K_n)/\nu'_\Delta(K_n) = \frac{3}{2}$.

Corollary 3.3.3. *If G is the line graph of a triangle-free graph, then $\tau'_\triangle(G) \leq 2\nu'_\triangle(G)$.*

Proof. Let $G = L(H)$, for a triangle-free graph H. Each star in H corresponds to a clique in G and these cliques partition the edges of G. The conclusion follows by Theorem 3.3.2. □

Lakshmanan et al. [122] showed that Tuza's Conjecture holds for the class of triangle-3-colourable graphs, where a graph G is *triangle-3-colourable* if its edges can be coloured with three colours so that the edges of each triangle receive three distinct colours. This is a direct consequence of the case $r = 3$ of Ryser's Conjecture proved by Aharoni [2] (Theorem 3.1.3): indeed, if G is triangle-3-colourable, then the triangle hypergraph of G is clearly 3-uniform and 3-partite. Since the class of triangle-3-colourable graphs contains that of 4-colourable graphs [122], this is a generalization of the planar case mentioned above. Another generalization of the planar case was given by Krivelevich [117], who showed Tuza's Conjecture holds for graphs with no $K_{3,3}$-subdivision. Recently, this was further generalized by Puleo [159] who showed, using the discharging method, that it holds for graphs having maximum average degree less than 7, where the *maximum average degree* of a graph G is defined as $\max\{2|E(H)|/|V(H)| : H \subseteq G\}$.

For all the classes mentioned so far, Conjecture 3.3.1 is tight, since they all contain the complete graph K_4. Therefore, a natural question arises: What happens if we forbid K_4? Haxell et al. [91] showed that the constant 2 cannot essentially be improved: for every $\varepsilon > 0$, there exists a K_4-free graph G such that $\tau'_\triangle(G) > (2-\varepsilon)\nu'_\triangle(G)$. Note that K_4 can be viewed as the 3-wheel. For a reason which will become apparent in the next paragraphs, we now show that the situation does not change even if we further forbid the 4-wheel and the 5-wheel. We follow the reasoning by Haxell et al. [91].

Lemma 3.3.4. *For each $\varepsilon > 0$, there exists a (W_3, W_4, W_5)-free graph G such that $\tau'_\triangle(G) > (2 - \varepsilon)\nu'_\triangle(G)$.*

Proof. Erdős [58] showed that, for any fixed k and sufficiently large n, there exists a graph on n vertices with girth greater than k and independence number smaller than $n^{\frac{2k}{2k+1}}$. For $k = 5$, let G_n be such a graph. We construct G by adding a universal vertex v_0 to G_n, i.e. a new vertex adjacent to all the vertices of G_n. Clearly, since G_n has girth at least 6, the graph G is (W_3, W_4, W_5)-free. Moreover, every triangle of G contains v_0 and so $\nu'_\triangle(G) = \alpha'(G_n) \leq \frac{n}{2}$. We now claim there exists a minimum-size triangle-transversal of G containing only edges incident to v_0. Indeed, if T is a triangle-transversal containing $uv \in E(G_n)$, we have that $(T \setminus \{uv\}) \cup \{v_0v\}$ is a triangle-transversal of G having size at most $|T|$. On the other hand, it is easy to see that a subset $F \subseteq V(G_n)$ is a vertex cover of G_n if and only if the set $\{v_0v : v \in F\}$ is a triangle-transversal of G. Therefore, we have $\tau'_\triangle(G) = n - \alpha(G_n) > n - n^{\frac{10}{11}}$. But then, for each $\varepsilon > 0$, we can find an n such that $n - n^{\frac{10}{11}} > (2 - \varepsilon)\frac{n}{2}$. □

There are two quantities which is natural to consider when dealing with packing and covering of triangles in a graph G: the number of triangles containing a certain edge and the number of edges shared by a certain triangle with other triangles. These quantities are in some sense dual to each other: the former corresponds to the degree of a vertex in the triangle hypergraph $\mathcal{H}(G)$ while the latter corresponds to the degree of a vertex in the subhypergraph of $\mathcal{H}(G)^*$ consisting of those hyperedges of size greater than 1.

Let us begin by considering graphs with edges in few triangles. Suppose we are given a graph G such that each of its edges belongs to at most one triangle, i.e. G is the edge-disjoint union of triangles plus possibly some edges that do not belong to any triangle. Clearly, we have $\tau'_\Delta(G) = \nu'_\Delta(G)$. It is then natural to consider the "next case": What happens if each edge of G belongs to at most two triangles? Following [88], we refer to graphs having this property as *flat* graphs. Note that, if G is flat, each edge of a K_4 subgraph of G is in no triangle other than those contained in K_4 and so we may restrict ourselves to consider K_4-free flat graphs. In this case, Haxell et al. [88] showed that the constant 2 can be dropped to $3/2$:

Theorem 3.3.5 (Haxell et al. [88]). *If G is a K_4-free flat graph, then $\tau'_\Delta(G) \leq \frac{3}{2}\nu'_\Delta(G)$. Equality holds if and only if G is the edge-disjoint union of 5-wheels plus possibly edges which are not in any triangle.*

Using Theorem 3.3.5, Haxell et al. [88] showed that the same bound holds for K_4-free planar graphs: If G is a K_4-free planar graph, then $\tau'_\Delta(G) \leq \frac{3}{2}\nu'_\Delta(G)$.

Note that flatness can be easily expressed in terms of the triangle graph[3] and passing to triangle graphs will be a recurrent theme of this section. By Observation 2.3.3, a K_4-free graph G is flat if and only if $T(G)$ is triangle-free.

The bound in Theorem 3.3.5 actually holds for the superclass of K_4-free graphs such that each triangle shares its edges with at most three other triangles. We can express this property succintly by invoking the triangle graph: each triangle of G shares its edges with at most three other triangles if and only if $T(G)$ is subcubic. In Section 3.2.1, we showed that $f(x) = \frac{3}{2}x$ is a θ-bounding function for the class of subcubic graphs and so we can immediately state the following generalization of Theorem 3.3.5:

Theorem 3.3.6. *If G is a K_4-free graph such that $T(G)$ is subcubic, then $\tau'_\Delta(G) \leq \frac{3}{2}\nu'_\Delta(G)$.*

Indeed, we have seen in Section 2.3 that if G is a K_4-free graph, we can translate an approximate min-max relation for τ'_Δ and ν'_Δ into an equivalent one for θ and α in the triangle graph $T(G)$. Since this reasoning will be extensively used throughout the section, it is convenient to restate it properly: By Fact 2.3.1 and Fact 2.3.4, if G is K_4-free, then

$$\nu'_\Delta(G) = \alpha(T(G)) = \omega(\overline{T(G)}) \text{ and } \tau'_\Delta(G) = \theta(T(G)) = \chi(\overline{T(G)}). \tag{3.4}$$

Note that if G is a graph having a subcubic triangle graph, then each triangle in a K_4 subgraph of G shares its edges only with the other triangles in K_4 and so a K_4 in G gives rise to a component of $T(G)$.

Let us now consider the second quantity mentioned above: the number of edges shared by a triangle with other triangles. Suppose we are given a graph G and a maximal family of edge-disjoint triangles. If every triangle of G shares (at most) one edge with other triangles, it is enough to take one edge for each triangle in the family in order to obtain a triangle-transversal of G. Therefore, in this case, we have $\tau'_\Delta(G) = \nu'_\Delta(G)$. Moreover, if every triangle of G shares at most two edges with other triangles, the reasoning above gives us $\tau'_\Delta(G) \leq 2\nu'_\Delta(G)$. Quite surprisingly, this trivial observation guarantees an essentially tight bound for the class in question. Indeed, it is easy to see that each triangle of the graph constructed in Lemma 3.3.4 shares at most two edges with other triangles. Nevertheless, as we will see in Section 3.3.1, if in addition each edge is in at most four triangles, the constant 2 can be dropped to $3/2$:

[3] Recall that the triangle graph $T(G)$ of G is the graph having as vertices the triangles of G, two vertices being adjacent if the corresponding triangles share an edge.

Theorem 3.3.7. *If G is a (K_4-free) graph such that each triangle shares at most two of its edges with other triangles and each edge belongs to at most four triangles (or, equivalently, $T(G)$ is (K_5, claw)-free), then $\tau'_\Delta(G) \leq \frac{3}{2}\nu'_\Delta(G)$. Equality holds if and only if each component of $T(G)$ is either C_5 or $L(G_{13})$.*

For the classes we consider, the constant 2 in Conjecture 3.3.1 can be improved and we try to provide the optimal one, in the spirit of the following problem:

Problem 3.3.8 (Lakshmanan et al. [122]). Given a class of graphs $\mathcal{G}$ in which at least one $G \in \mathcal{G}$ is not triangle-free, determine the infimum of constants c such that $\tau'_\Delta(G) \leq c\nu'_\Delta(G)$ holds for every $G \in \mathcal{G}$.

For the classes of graphs in Theorems 3.3.5 to 3.3.7, we have $c = 3/2$. In general, for the class of all graphs, Tuza's Conjecture claims the answer is $c = 2$ and we certainly have $c \geq 2$ for every class containing K_4. Moreover, Lemma 3.3.4 tells us that even for $\mathcal{G} = \mathit{Free}(W_3, W_4, W_5)$ we cannot expect $c < 2$. On the other hand, forbidding all odd-wheels could lead to a $c < 2$. Lakshmanan et al. [122] showed that Tuza's Conjecture holds for the class of odd-wheel-free graphs and they noticed that the odd-wheel-free graph $\overline{C_7}$ implies $c \geq 4/3$. The nice feature of this class is that a triangle-transversal and a triangle-packing witnessing $\tau'_\Delta \leq 2\nu'_\Delta$ can be easily constructed from a "local optimum", as we show in the following:

Theorem 3.3.9 (Lakshmanan et al. [122]). *If G is an odd-wheel-free graph, then $\tau'_\Delta(G) \leq 2\nu'_\Delta(G)$.*

Proof. Note first that every subgraph (not necessarily induced) of an odd-wheel-free graph is odd-wheel-free. Suppose now, to the contrary, that G is a counterexample with the minimum number of edges.

Consider a non-isolated vertex $v \in V(G)$ and the subgraph G_v of G induced by $N(v)$. Since G is odd-wheel-free, we have that G_v is bipartite and so the König-Egerváry Theorem (Theorem 3.1.1) implies that $\beta(G_v) = \alpha'(G_v)$. Consider now the graph G' obtained from G by deleting the edges incident to v and the edges in a maximum matching M of G_v. Clearly, we have $\nu'_\Delta(G) \geq \nu'_\Delta(G') + \alpha'(G_v)$. On the other hand, as we have seen in the proof of Lemma 3.3.4, a vertex cover of G_v corresponds to a set of edges which meets all the triangles having v as a vertex. Therefore, if we additionally add this edge set to the edges in M and to those in a triangle-transversal of G', we obtain a triangle-transversal of G. This implies that $\tau'_\Delta(G) \leq \tau'_\Delta(G') + \beta(G_v) + \alpha'(G_v) = \tau'_\Delta(G') + 2\alpha'(G_v)$. But then, by minimality, we have

$$\tau'_\Delta(G) \leq \tau'_\Delta(G') + 2\alpha'(G_v) \leq 2\nu'_\Delta(G') + 2\alpha'(G_v) \leq 2\nu'_\Delta(G),$$

a contradiction. □

Theorem 3.3.9 was generalized by Puleo [159], who showed the following:

Theorem 3.3.10 (Puleo [159]). *If G is a graph with no subgraph isomorphic to any odd-wheel W_n, for $n \geq 5$, then $\tau'_\Delta(G) \leq 2\nu'_\Delta(G)$.*

The main idea (used also in his proof of the fact that Tuza's Conjecture holds for graphs with maximum average degree less than 7) revolves around the concept of reducible set, which is in some sense a "local optimum". Let us briefly sketch some details. Given a set S of triangles of a graph G, an *S-edge* is an edge of some triangle in S. A non-empty set $V_0 \subseteq V(G)$ is *reducible* if there exist a set S of edge-disjoint triangles of G and a set $X \subseteq E(G)$ such that the following conditions hold:

- $|X| \leq 2|S|$;
- $G - X$ has no triangle containing a vertex of V_0;
- X contains every S-edge whose endpoints are both outside V_0.

When V_0, S and X satisfy the conditions above, V_0 is said to be *reducible* using S and X. Note that in the proof of Theorem 3.3.9 we showed, thanks to the König-Egerváry Theorem, that any non-isolated vertex v is reducible using the following S and X: the set S is formed by the triangles having v as one vertex and having one edge in a maximum matching M of G_v; the set X is the union of M with the set $\{vw : w$ is in a minimum vertex cover of $G_v\}$. We then saw how the existence of such a reducible set leads to a contradiction. In a similar fashion, one can show that a minimal counterexample to Tuza's Conjecture has no reducible set:

Lemma 3.3.11 (Puleo [159]). *Let G be a graph and let $V_0 \subseteq V(G)$ be reducible using S and X. Moreover, let $G' = (G - X) - V_0$. If $\tau'_\Delta(G') \leq 2\nu'_\Delta(G')$, then $\tau'_\Delta(G) \leq 2\nu'_\Delta(G)$.*

Following the reasoning in the proof of Theorem 3.3.9, it is easy to see that a sufficient condition for a vertex $v \in V(G)$ (belonging to at least one triangle) to be reducible is that $\beta(G[N(v)]) = \alpha'(G[N(v)])$. A graph G such that $\beta(G) = \alpha'(G)$ is usually called *König-Egerváry* [51]. In other words, using the terminology introduced in Section 3.1, the class of König-Egerváry graphs is exactly the class of graphs satisfying the Min-Max Property. Puleo [159] introduced a generalization of König-Egerváry graphs: a graph G is *weak König-Egerváry* if G has a matching M and a vertex set $Q \subseteq V(G)$ such that $|Q| \leq |M|$ and Q is a vertex cover of $G - M$. Following the proof of Theorem 3.3.9, it is not difficult to see that this relaxed notion is still sufficient to guarantee reducibility: for any graph G and $v \in V(G)$, if $G[N(v)]$ is weak König-Egerváry then $\{v\}$ is reducible [159]. He then showed that every graph with no odd cycle of length greater than 3 is weak König-Egerváry. This fact, together with the previous remarks, immediately implies Theorem 3.3.10.

In Section 3.3.2, motivated by the previous discussion, we are interested in graphs without the odd-wheels W_3 and W_5. As a partial result towards a proof of Tuza's Conjecture for this class, we obtain the following:

Theorem 3.3.12. *If G is a (W_3, W_5)-free graph such that $T(G)$ is co-banner-free, then $\tau'_\Delta(G) \leq \frac{10}{7}\nu'_\Delta(G)$.*

Consider now the class of K_4-free graphs with odd-hole-free triangle graph. It is easy to see that this is a subclass of odd-wheel-free graphs. In this case, the Strong Perfect Graph Theorem [41] implies that the equality $\tau'_\Delta = \nu'_\Delta$ holds:

Theorem 3.3.13 (Lakshmanan et al. [122]). *If G is a K_4-free graph such that $T(G)$ is C_{2k+1}-free for all $k \geq 2$, then $\tau'_\Delta(G) = \nu'_\Delta(G)$.*

Indeed, we know that if G is K_4-free, then $T(G)$ is diamond-free (Observation 2.3.6) and so $\overline{C_{2k+1}}$-free, for any $k \geq 3$. Therefore, recalling (3.4), we have $\tau'_\Delta(G) = \theta(T(G)) = \alpha(T(G)) = \nu'_\Delta(G)$.

Theorem 3.3.13 implies in particular that the equality $\tau'_\Delta = \nu'_\Delta$ holds for the K_4-free graphs with a P_4-free triangle graph. In Section 3.3.2, we show the following tight bound for the "next case":

Theorem 3.3.14. *If G is a K_4-free graph such that $T(G)$ is P_5-free, then $\tau'_\Delta(G) \leq \frac{3}{2}\nu'_\Delta(G)$.*

Most of the results on Tuza's Conjecture in Sections 3.3.1 and 3.3.2 are in fact obtained by providing θ-bounding functions for classes related to triangle graphs. In this context, we will see how the (partial) list of forbidden induced subgraphs for triangle graphs given in Section 2.3 ($\{K_{1,4}, K_{2,3}, \text{diamond}, \text{twin-}C_5\}$) comes in handy.

We conclude this section with a table summarizing the results we are going to prove. All the classes of K_4-free graphs which we study in relation to Problem 3.3.8 can actually be characterized in terms of the triangle graph.

G	$T(G)$	Upper bound for $\tau'_\Delta(G)/\nu'_\Delta(G)$	Reference
K_4-free	subcubic	$3/2$	Theorem 3.3.6
K_4-free	K_4-free, maximum degree 4	$193/98$	Theorem 3.3.16
K_4-free	(K_5, claw)-free	$3/2$	Theorem 3.3.7
K_4-free	$(C_5, \text{co-banner})$-free	$10/7$	Theorem 3.3.12
K_4-free	P_5-free	$3/2$	Theorem 3.3.14

It is easy to see that the classes above are mutually incomparable with respect to set inclusion.

3.3.1 Graphs with edges in few triangles

In this section, we consider K_4-free graphs having edges in at most four triangles and we address Problem 3.3.8. Our results are mostly simple consequences of those obtained in Section 3.2.1. The equality $\tau'_\Delta(G) = \nu'_\Delta(G)$ trivially holds if every edge of G is in at most one triangle and Haxell et al. [88] showed that $\tau'_\Delta(G) \leq \frac{3}{2}\nu'_\Delta(G)$, for any K_4-free graph such that each edge is in at most two triangles. This was generalized by Theorem 3.3.6, which we now restate:

Theorem 3.3.6. *If G is a K_4-free graph such that $T(G)$ is subcubic, then $\tau'_\Delta(G) \leq \frac{3}{2}\nu'_\Delta(G)$.*

Proof. By (3.4) and Theorem 3.2.1, we have $\tau'_\Delta(G) = \theta(T(G)) \leq \frac{3}{2}\alpha(T(G)) = \frac{3}{2}\nu'_\Delta(G)$. □

An edge-disjoint union of 5-wheels (plus possibly edges not in triangles) shows that the constant $3/2$ is optimal. Note that, as remarked in Section 3.2.1, we do not have a characterization of the subcubic graphs G such that $\theta(G) = \frac{3}{2}\alpha(G)$ and we conjectured that every component of such an extremal G is either C_5 or G_{11} (see Figure 3.1). On the other hand, it is easy to see that the graph G_{11} contains an induced copy of a twin-C_5 and so, by Observation 2.3.9, it cannot appear as an induced subgraph of a triangle graph. Therefore, it is natural to conjecture the following:

Conjecture 3.3.15. *If G is a K_4-free graph such that $T(G)$ is subcubic, then $\tau'_\Delta(G) = \frac{3}{2}\nu'_\Delta(G)$ if and only if G is the edge-disjoint union of 5-wheels plus possibly edges which are not in any triangle.*

What about if each edge belongs to at most three triangles? In this case, each triangle shares its edges with at most six other triangles. If we allow each triangle to share its edges only with at most four other triangles, we can state the following:

Theorem 3.3.16. *If G is a K_4-free graph such that each edge belongs to at most three triangles and each triangle shares its edges with at most four other triangles (or, equivalently, $T(G)$ is K_4-free and has maximum degree 4), then $\tau'_\Delta(G) \leq \frac{193}{98}\nu'_\Delta(G)$.*

Proof. The equivalence of the two classes in the statement is immediate. By (3.4) and Theorem 3.2.8, we have $\tau'_\Delta(G) = \theta(T(G)) \leq \frac{193}{98}\alpha(T(G)) = \frac{193}{98}\nu'_\Delta(G)$. □

We believe the constant in Theorem 3.3.16 is far from optimal but it seems difficult even to formulate a conjecture on the optimal value. In fact, we do not have any example for which the ratio is greater than $3/2$. In Section 3.2.1, we conjectured that if G is a graph with maximum degree 4, then $\theta(G) \leq \frac{7}{4}\alpha(G)$, with equality if and only if every component of G is G_{13} (see Figure 3.1). On the other hand, since G_{13} contains an induced $K_{1,4}$, it cannot be a triangle graph.

Let us now consider graphs such that each triangle shares at most two of its edges with other triangles. Clearly, any such graph G is K_4-free and we have already seen that the essentially tight bound $\tau'_\Delta(G) \leq 2\nu'_\Delta(G)$ holds. We now show that, if in addition each edge is in at most four triangles, the constant 2 can be dropped to $3/2$:

Theorem 3.3.7. *If G is a (K_4-free) graph such that each triangle shares at most two of its edges with other triangles and each edge belongs to at most four triangles (or, equivalently, $T(G)$ is (K_5, claw)-free), then $\tau'_\Delta(G) \leq \frac{3}{2}\nu'_\Delta(G)$. Equality holds if and only if each component of $T(G)$ is either C_5 or $L(G_{13})$.*

Note that, by Observation 2.3.13, $L(G_{13})$ is indeed a triangle graph of a K_4-free graph: for example, we have $L(G_{13}) = T(K_1 \vee G_{13})$.

The crucial observation for the proof of Theorem 3.3.7 is that the triangle graph is in fact a line graph. Indeed, by Observation 2.3.6, we know that the triangle graph of a K_4-free graph is diamond-free and the following lemma tells us that (K_5, claw, diamond)-free graphs are exactly line graphs of triangle-free graphs with maximum degree 4. The proof is similar to that of Theorem 2.2.1:

Lemma 3.3.17. *A graph G is the line graph of a triangle-free graph with maximum degree 4 if and only if it is (K_5, claw, diamond)-free.*

Proof. If G is the line graph of a triangle-free graph with maximum degree 4, then it is clearly (K_5, claw, diamond)-free.

Suppose now G is (K_5, claw, diamond)-free. By Theorem 2.1.1, we have that $G = L(H)$, for some graph H. Consider the graph H' obtained from H by replacing each component isomorphic to K_3 with a claw. Clearly, $G = L(H) = L(H')$. Since G is K_5-free, H' has maximum degree 4. Moreover, H' is triangle-free. Indeed, if H' contains a triangle T, then there exists a vertex $v \notin V(T)$ adjacent to a vertex of T and so there exists an induced diamond in $L(H')$, a contradiction. □

By the discussions above, we know it would be enough to show that $\theta(G) \leq \frac{3}{2}\alpha(G)$, for any $(K_5,$ claw, diamond)-free graph G. But this is an easy corollary of the following result by Joos [101], which is in some sense complementary to Theorem 3.2.5:

Theorem 3.3.18 (Joos [101]). *If G is a triangle-free graph with $\Delta(G) \leq 4$, then $\alpha(G) + \frac{3}{2}\alpha'(G) \geq |V(G)|$. Equality holds if and only if every component C of G is either in $\{K_1, C_5\}$ or has order 13, $\alpha(C) = 4$ and $\alpha'(C) = 6$.*

Corollary 3.3.19. *If G is a $(K_5,$claw, diamond$)$-free graph, then $\theta(G) \leq \frac{3}{2}\alpha(G)$. Equality holds if and only if every component of G is either C_5 or $L(G_{13})$.*

Proof. By Lemma 3.3.17, G is the line graph of a triangle-free graph H with maximum degree at most 4. Clearly, $\alpha(G) = \alpha'(H)$ and, by Fact 2.3.2, we have that $\theta(G) = \beta(H)$. Therefore, by Theorem 3.3.18 and since the complement of a vertex cover is an independent set, we have

$$\theta(G) = \beta(H) = |V(H)| - \alpha(H) \leq \frac{3}{2}\alpha'(H) = \frac{3}{2}\alpha(G).$$

Let us now consider the cases of equality. As remarked in Section 3.2.1, the $(3,5)$-Ramsey graph G_{13} is the only graph on 13 vertices with parameters $\omega = 2$ and $\alpha = 4$. If equality holds then, by Theorem 3.3.18, each component of H is in $\{K_1, C_5, G_{13}\}$. This implies that each component of $G = L(H)$ is either C_5 or $L(G_{13})$. The converse clearly holds. □

We can finally proceed to the proof of Theorem 3.3.7:

Proof of Theorem 3.3.7. The equivalence of the two classes in the statement is immediate. Moreover, since $T(G)$ is diamond-free then, by (3.4) and Corollary 3.3.19, we have $\tau'_\Delta(G) = \theta(T(G)) \leq \frac{3}{2}\alpha(T(G)) = \frac{3}{2}\nu'_\Delta(G)$. The characterization of equality follows again by Corollary 3.3.19. □

3.3.2 θ-Bounding functions for classes related to triangle graphs

In this section, we prove Theorems 3.3.12 and 3.3.14. As we are dealing with K_4-free graphs, we use our standard approach and proceed by showing θ-bounding functions for the classes *Free*(co-banner, odd-antihole, $K_{1,4}$) and *Free*(P_5, diamond, $K_{2,3}$) which are related to Theorem 3.3.12 and Theorem 3.3.14, respectively. These results might be of independent interest.

Let us begin by restating Theorem 3.3.12:

Theorem 3.3.12. *If G is a (W_3, W_5)-free graph such that $T(G)$ is co-banner-free, then $\tau'_\Delta(G) \leq \frac{10}{7}\nu'_\Delta(G)$.*

We know that an induced C_5 in $T(G)$ corresponds to an induced W_5 in G (Lemma 2.3.10). Moreover, we have that $T(G)$ is diamond-free. Therefore, $T(G)$ is in fact odd-antihole-free. For our purposes it will be convenient to work with the complement graph and our proof of

Theorem 3.3.12 relies on a characterization of (banner, odd-hole)-free graphs given by Hoàng [96]. Before stating his result, let us recall some definitions.

Let A and B be disjoint subsets of $V(G)$ and let $b \in V(G) \setminus A$. The vertex b is *complete* to A if b is adjacent to every vertex of A and b is *anticomplete* to A if b is non-adjacent to every vertex of A. If every vertex of A is complete to B, then A is *complete* to B. Similarly, if every vertex of A is anticomplete to B, then A is *anticomplete* to B. A *module* in G is a subset $M \subseteq V(G)$ such that every vertex in $V(G) \setminus M$ is either complete or anticomplete to M. A *homogeneous set* in G is a module in G containing at least two vertices and different from $V(G)$.

Theorem 3.3.20 (Hoàng [96]). *If G is a (banner, odd-hole)-free graph, then either G is perfect, or $\alpha(G) \leq 2$, or every odd-antihole of G belongs to a homogeneous set M in G such that $G[M]$ is co-triangle-free.*

Some remarks are in place. Recently, Scott and Seymour [171] solved a conjecture by Gyárfás [81] on χ-bounded classes: they showed that the class of odd-hole-free graphs is χ-bounded by the function $f(x) = 2^{2^{x+2}}$. It is likely that this exponential function can be improved, although they provided a series of examples showing that a linear bounding function is not possible. Using Theorem 3.3.20, Hoàng [96] showed that (banner, odd-hole)-free graphs are 2-divisible. Recall that a graph G is *k-divisible* if the vertex set of each induced subgraph H of G with at least one edge can be partitioned into k sets none of which contains a clique of size $\omega(H)$. An easy induction shows that $\chi(G) \leq k^{\omega(G)-1}$, for any k-divisible graph G. Therefore, the bound $\chi \leq 2^{2^{\omega+2}}$ can be improved to $\chi \leq 2^{\omega-1}$ for the subclass *Free*(banner, odd-hole). We now show that by further forbidding $\overline{K_{1,4}}$, we can obtain a linear χ-bounding function:

Lemma 3.3.21. *If G is a (banner, odd-hole, $\overline{K_{1,4}}$)-free graph, then $\chi(G) \leq \frac{10}{7}\omega(G)$.*

Proof. We proceed by induction on the number of vertices. If G is perfect, then $\chi(G) = \omega(G)$. If $\alpha(G) \leq 2$, then $\overline{G}$ is triangle-free and so, since it is $K_{1,4}$-free as well, it must be subcubic. Moreover, $\overline{G}$ is by assumption C_5-free and so Theorem 3.2.12 implies that

$$\chi(G) = \theta(\overline{G}) \leq \frac{10}{7}\alpha(\overline{G}) = \frac{10}{7}\omega(G).$$

Therefore, in view of Theorem 3.3.20, we may assume G contains an odd-antihole $\overline{C_{2k+1}}$, with $k \geq 3$, which belongs to a homogeneous set M. In particular, M contains a triangle. Now let $A \cup B$ be a partition of $V(G) \setminus M$ such that A is complete to M and B is anticomplete to M. If $A = \varnothing$, then G is the disjoint union of $G[M]$ and $G[B]$, and we immediately conclude by the induction hypothesis. Moreover, if $B = \varnothing$, then G is the join of $G[A]$ and $G[M]$ and by the induction hypothesis we have

$$\chi(G) = \chi(G[A]) + \chi(G[M]) \leq \frac{10}{7}\omega(G[A]) + \frac{10}{7}\omega(G[M]) = \frac{10}{7}\omega(G).$$

Therefore, we may assume both A and B are non-empty. Since M contains a triangle and G is $\overline{K_{1,4}}$-free, it is easy to see that A must be complete to B. But then G is the join of $G[A]$ and $G[M \cup B]$ and we conclude as above. $\square$

We do not have any example of a graph attaining equality in Lemma 3.3.21 and we suspect that the optimal constant is $4/3$, as given by the graph $\overline{C_7}$. As a side remark, note that Chudnovsky et al. [40] showed that (odd-hole, K_4)-free graphs are 4-colourable, with $\overline{C_7}$ being 4-chromatic.

The proof of Theorem 3.3.12 is now an immediate consequence of Lemma 3.3.21:

Proof of Theorem 3.3.12. Since G is K_4-free, we have that $\tau'_\Delta(G) = \chi(\overline{T(G)})$ and $\nu'_\Delta(G) = \omega(\overline{T(G)})$ (see (3.4)). By Lemma 2.3.10, $T(G)$ is C_5-free (the triangle graph of a K_4-free graph is C_5-free if and only if the original graph is W_5-free). Moreover, since $T(G)$ is diamond-free as well, we have that $\overline{T(G)}$ is odd-hole-free. Finally, since $T(G)$ is $K_{1,4}$-free, Lemma 3.3.21 implies that $\tau'_\Delta(G) \leq \frac{10}{7}\nu'_\Delta(G)$. □

Suppose now G is a K_4-free graph such that $T(G)$ is P_5-free. Clearly, 5-wheels show that the ratio τ'_Δ/ν'_Δ is at least $3/2$ and Theorem 3.3.14 tells us that this is an extremal case:

Theorem 3.3.14. *If G is a K_4-free graph such that $T(G)$ is P_5-free, then $\tau'_\Delta(G) \leq \frac{3}{2}\nu'_\Delta(G)$.*

In order to prove Theorem 3.3.14, we use a structural characterization of $(P_5, \text{diamond})$-free graphs by Brandstädt [29]. Before stating his result, we need to introduce the following classes of graphs which constitute the basic graphs in the characterization.

A graph is a:

- *thin spider* if it is partitionable into a clique C and an independent set I, with $|C| = |I|$ or $|C| = |I| + 1$, such that the edges between C and I form a matching and at most one vertex in C is not covered by the matching;
- *matched co-bipartite graph* if it is partitionable into two cliques C_1 and C_2, with $|C_1| = |C_2|$ or $|C_1| = |C_2| + 1$, such that the edges between C_1 and C_2 form a matching and at most one vertex in C_1 and C_2 is not covered by the matching;
- *bipartite chain graph* if it is bipartite, with bipartition $X_1 \cup X_2$, and each X_i forms a chain, i.e. for $1 \leq i \leq 2$, we have that $\{N(x_i) : x_i \in X_i\}$ is linearly ordered with respect to set inclusion;
- *co-bipartite chain graph* if it is the complement of a bipartite chain graph;
- *enhanced co-bipartite chain graph* if it is partitionable into a co-bipartite chain graph with cliques C_1 and C_2 and three additional vertices a, b and c (a and c optional) such that $N(a) = C_1 \cup C_2$, $N(b) = C_1$ and $N(c) = C_2$;
- *enhanced bipartite chain graph* if it is the complement of an enhanced co-bipartite chain graph.

In the following, a graph G is *co-connected* if its complement $\overline{G}$ is connected.

Theorem 3.3.22 (Brandstädt [29]). *If G is a connected and co-connected $(P_5, \text{diamond})$-free graph, then either G contains a homogeneous set (inducing a P_3-free subgraph) or one of the following holds: G is a matched co-bipartite graph or a thin spider or an enhanced bipartite chain graph or it has at most 9 vertices.*

The strategy is clear: we reduce the statement in Theorem 3.3.14 into an equivalent one for the triangle graph and we use the fact that this graph is diamond-free and $K_{2,3}$-free (see Section 2.3).

Theorem 3.3.23. *If G is a $(P_5, diamond, K_{2,3})$-free graph, then $\theta(G) \leq \frac{3}{2}\alpha(G)$.*

Proof. We proceed by induction on the number of vertices. If G is not connected, then it is the disjoint union of two non-empty graphs G_1 and G_2. By the induction hypothesis,

$$\theta(G) = \theta(G_1) + \theta(G_2) \leq \frac{3}{2}\alpha(G_1) + \frac{3}{2}\alpha(G_2) = \frac{3}{2}\alpha(G).$$

Similarly, if $\overline{G}$ is not connected, then it is the disjoint union of two non-empty graphs G_1 and G_2 and, by the induction hypothesis,

$$\theta(G) = \chi(\overline{G}) = \max\{\chi(G_1), \chi(G_2)\} \leq \frac{3}{2}\max\{\omega(G_1), \omega(G_2)\} = \frac{3}{2}\omega(\overline{G}) = \frac{3}{2}\alpha(G).$$

Therefore, we may assume G to be connected and co-connected.

Suppose now G contains a homogeneous set M. By definition, $|M| \geq 2$ and $M \neq V(G)$. Moreover, M induces a P_3-free subgraph, i.e. it is a disjoint union of cliques. Let $A \cup B$ be a partition of $V(G) \setminus M$, where A is complete to M and B is anticomplete to M. Suppose first $|A| \leq 1$. This implies that M contains a simplicial vertex v, i.e. a vertex whose neighbourhood is a clique. If $N[v] = V(G)$, then $\theta(G) = \alpha(G)$. Otherwise, by the induction hypothesis,

$$\theta(G) \leq \theta(G - N[v]) + 1 \leq \frac{3}{2}\alpha(G - N[v]) + 1 \leq \frac{3}{2}(\alpha(G) - 1) + 1 < \frac{3}{2}\alpha(G).$$

Therefore, we may assume $|A| \geq 2$. Since G is diamond-free, we have that either A and M are both independent sets or they are both cliques. But if the latter holds, then M contains a simplicial vertex and we conclude as in the previous paragraph. Therefore, we may further assume that A and M are both independent sets and since G is $K_{2,3}$-free, we have $|A| = |M| = 2$. This implies that $\theta(G) \leq \theta(G - A - M) + 2$. On the other hand, $\alpha(G) \geq \alpha(G - A - M) + 2$ and so, by the induction hypothesis, we have

$$\theta(G) \leq \theta(G - A - M) + 2 \leq \frac{3}{2}\alpha(G - A - M) + 2 < \frac{3}{2}\alpha(G).$$

Therefore, we may assume G does not contain a homogeneous set. By Theorem 3.3.22, G is either a matched co-bipartite graph or a thin spider or an enhanced bipartite chain graph or it has at most 9 vertices. It is easy to see that in all the first three cases $\theta(G) = \alpha(G)$. Moreover, if G has at most 9 vertices then, by Theorem 3.2.9, we have $\theta(G) - \alpha(G) \leq 1$. If $\alpha(G) \leq 1$, then G is complete and $\theta(G) = \alpha(G)$. Otherwise, $\theta(G) \leq \alpha(G) + 1 \leq \frac{3}{2}\alpha(G)$. □

Proof of Theorem 3.3.14. We have seen that triangle graphs are $K_{2,3}$-free (Corollary 2.3.8). Moreover, since G is K_4-free, $T(G)$ is diamond-free. Therefore, Theorem 3.3.23 implies that $\tau'_\Delta(G) = \theta(T(G)) \leq \frac{3}{2}\alpha(T(G)) = \frac{3}{2}\nu'_\Delta(G)$. □

3.4 On Jones' Conjecture

Given a graph G, we have seen that the cycle hypergraph $\mathcal{H}(G)$ of G is the hypergraph whose vertices are the vertices of G and whose edges are the vertex sets of cycles of G. Clearly, a transversal of $\mathcal{H}(G)$ is nothing but a feedback vertex set of G and so $\tau(\mathcal{H}(G)) = \tau_c(G)$. Similarly, we denote the size of a maximum packing of $\mathcal{H}(G)$ (i.e. the maximum number of vertex-disjoint cycles of G) by $\nu_c(G)$.

Erdős and Pósa [59] showed that $\tau_c(G)$ can in fact be upper bounded in terms of $\nu_c(G)$: for any graph G, we have $\tau_c(G) = O(\nu_c(G) \log \nu_c(G))$, the bound being sharp. Kloks et al. [107] conjectured that this can be considerably improved in the case of planar graphs:

Conjecture 3.4.1 (Jones' Conjecture [107]). *If G is a planar graph, then $\tau_c(G) \leq 2\nu_c(G)$.*

If true, Conjecture 3.4.1 would be sharp, as can be seen by considering wheel graphs. Kloks et al. [107] showed it holds for outerplanar graphs and, in general, they proved the weaker $\tau_c(G) \leq 5\nu_c(G)$. The factor 5 was subsequently improved to 3 by a series of authors [32, 36, 134]. Chappell et al. [32] actually showed a stronger statement: $\tau_c(G) \leq 3\nu_c(G)$, for any graph G that embeds in a closed surface of non-negative Euler characteristic. The proofs in [32, 36, 134] all essentially rely on a refinement of the idea adopted in [107]. Using the discharging method, they showed that every 2-edge-connected triangle-free plane graph with minimum degree at least 3 either has a 4-face containing at least one cubic vertex or a 5-face containing at least four cubic vertices.

To the best of our knowledge, these are the only works related to Jones' Conjecture, which seems to be non-trivial even in the case of subcubic graphs. In this section, we prove Conjecture 3.4.1 for claw-free graphs with maximum degree at most 4:

Theorem 3.4.2. *If G is a planar claw-free graph with maximum degree at most 4, then $\tau_c(G) \leq 2\nu_c(G)$.*

We remark that Ma et al. [134] showed that a minimum counterexample to Jones' Conjecture is 3-connected and Plummer [158] showed that a planar 3-connected claw-free graph has maximum degree at most 6.

In order to prove Theorem 3.4.2, we first show that the bound $\tau_c \leq 2\nu_c$ holds for line graphs of subcubic triangle-free graphs, a class we have studied in Section 2.2 and which contains non-planar graphs.

In Section 3.4.1, we finally consider the case of subcubic graphs and provide a list of properties a minimum subcubic counterexample must have, if any.

Let us remark that in this section a triangle-transversal of G is intended to be a subset of $V(G)$ meeting all the triangles of G.

Theorem 3.4.3. *If G is a $(K_4, \text{claw}, \text{diamond})$-free graph, then $\tau_c(G) \leq 2\nu_c(G)$.*

Proof. We proceed by induction on the number of vertices of G. Without loss of generality, we may assume G to be connected. Recall that, for $v \in V(G)$, the possible subgraphs induced by $N[v]$ are those depicted in Figure 2.3.

Suppose first G is triangle-free. This implies that $\Delta(G) \leq 2$ and G is either a path or a cycle, from which $\tau_c(G) = \nu_c(G)$.

Therefore, we may assume G contains a triangle T. Suppose there exists $v \in V(T)$ such that $d_G(v) \leq 3$. Since any cycle containing v passes through one vertex in $V(T) \setminus \{v\}$, a feedback vertex set of G can be obtained from a feedback vertex set of $G - V(T)$ by adding the two vertices in $V(T) \setminus \{v\}$. Moreover, by the induction hypothesis, we have $\tau_c(G - V(T)) \leq$

$2\nu_c(G - V(T))$ and so

$$\tau_c(G) \leq \tau_c(G - V(T)) + 2 \leq 2\nu_c(G - V(T)) + 2 \leq 2\nu_c(G).$$

Therefore, we may assume that every triangle of G contains only 4-vertices. Connectedness then implies that G is 4-regular and so we have that $G = L(H)$, for a cubic triangle-free graph H.

Suppose now G is 3-connected. This implies that H is 3-edge-connected. Jackson and Yoshimoto [100] showed that every 3-edge-connected graph with n vertices has a spanning even subgraph in which each component has at least $\min\{n, 5\}$ vertices. Therefore, let F be the 2-factor of H whose existence is guaranteed by the previous result. Then F has at most $\frac{|V(H)|}{5}$ components. Moreover, the edges of $H - E(F)$ constitute a perfect matching of H and the vertices of G corresponding to this matching form a triangle-transversal T of G (see Lemma 2.2.10). Consider now the set obtained by taking exactly one edge for each component of F and let T' be the corresponding set of vertices of G. It is easy to see that $T \cup T'$ is a feedback vertex set of G of size at most $\frac{|V(H)|}{2} + \frac{|V(H)|}{5}$. Now it remains to properly lower bound $\nu_c(G)$. Denoting by $\nu_\Delta(G)$ the maximum number of vertex-disjoint triangles of G, we have $\nu_c(G) \geq \nu_\Delta(G) = \alpha(H)$. On the other hand, Staton [176] showed that $\alpha(H) \geq \frac{5}{14}|V(H)|$, for every subcubic triangle-free graph H. Therefore, combining everything, we get

$$\tau_c(G) \leq \frac{|V(H)|}{2} + \frac{|V(H)|}{5} < 2 \cdot \frac{5}{14}|V(H)| \leq 2\nu_c(G).$$

By the paragraph above, we may assume G is not 3-connected. Suppose first G has a cut-vertex v. Let G_1 and G_2 be two non-trivial induced subgraphs of G such that $G = G_1 \cup G_2$ and $V(G_1) \cap V(G_2) = \{v\}$. If $\nu_c(G) \geq \nu_c(G_1) + \nu_c(G_2)$ then, by the induction hypothesis,

$$\begin{aligned}
\tau_c(G) &\leq \tau_c(G_1 - N_{G_1}[v]) + \tau_c(G_2 - N_{G_2}[v]) + 4 \\
&\leq 2\nu_c(G_1 - N_{G_1}[v]) + 2\nu_c(G_2 - N_{G_2}[v]) + 4 \\
&\leq 2(\nu_c(G_1) - 1) + 2(\nu_c(G_2) - 1) + 4 \\
&\leq 2\nu_c(G_1) + 2\nu_c(G_2) \\
&\leq 2\nu_c(G).
\end{aligned}$$

Finally, if $\nu_c(G) = \nu_c(G_1) + \nu_c(G_2) - 1$ then, for each $i \in \{1, 2\}$, every maximum size cycle packing of G_i contains a cycle through v. This means that, for each $i \in \{1, 2\}$, we have $\nu_c(G_i - v) \leq \nu_c(G_i) - 1$ and, by the induction hypothesis,

$$\begin{aligned}
\tau_c(G) &\leq \tau_c(G_1 - v) + \tau_c(G_2 - v) + 1 \\
&\leq 2\nu_c(G_1 - v) + 2\nu_c(G_2 - v) + 1 \\
&\leq 2(\nu_c(G_1) - 1) + 2(\nu_c(G_2) - 1) + 1 \\
&\leq 2\nu_c(G_1) + 2\nu_c(G_2) - 3 \\
&< 2\nu_c(G).
\end{aligned}$$

Therefore, we may assume G is 2-connected but not 3-connected. In particular, G has a 2-cut $\{u, v\}$. Let G_1 and G_2 be two non-trivial induced subgraphs of G such that $G = G_1 \cup G_2$ and $V(G_1) \cap V(G_2) = \{u, v\}$. If $\nu_c(G - \{u, v\}) \leq \nu_c(G) - 1$ then, by the induction hypothesis,

$$\tau_c(G) \leq \tau_c(G - \{u, v\}) + 2 \leq 2\nu_c(G - \{u, v\}) + 2 \leq 2\nu_c(G).$$

Therefore, we may assume $\nu_c(G - \{u, v\}) = \nu_c(G)$. But then $\nu_c(G) = \nu_c(G - \{u, v\}) = \nu_c(G_1 - \{u, v\}) + \nu_c(G_2 - \{u, v\})$, from which $\nu_c(G) \leq \nu_c(G_1 - \{u, v\}) + \nu_c(G_2)$. On the other hand, we have that $\nu_c(G) \geq \nu_c(G_1 - \{u, v\}) + \nu_c(G_2)$ and so $\nu_c(G) = \nu_c(G_1 - \{u, v\}) + \nu_c(G_2)$. Therefore, $\nu_c(G_2 - \{u, v\}) = \nu_c(G_2)$ and, similarly, $\nu_c(G_1 - \{u, v\}) = \nu_c(G_1)$. Combining everything, we get $\nu_c(G) = \nu_c(G_1) + \nu_c(G_2)$. Since G is 2-connected, u has neighbours both in G_1 and G_2 and so, again by the induction hypothesis, we have

$$\begin{aligned}
\tau_c(G) &\leq \tau_c(G_1 - N_{G_1}[u]) + \tau_c(G_2 - N_{G_2}[u]) + 4 \\
&\leq 2\nu_c(G_1 - N_{G_1}[u]) + 2\nu_c(G_2 - N_{G_2}[u]) + 4 \\
&\leq 2(\nu_c(G_1) - 1) + 2(\nu_c(G_2) - 1) + 4 \\
&\leq 2\nu_c(G_1) + 2\nu_c(G_2) \\
&= 2\nu_c(G).
\end{aligned}$$

□

Note that the factor 2 in Theorem 3.4.3 is best possible. This can be seen by considering the line graph of the graph obtained from the 6-cycle by adding an edge between two vertices at distance 3.

Using Theorem 3.4.3, we can finally show that Jones' Conjecture holds for claw-free graphs with maximum degree at most 4.

Proof of Theorem 3.4.2. We proceed by contradiction. Therefore, let G be a counterexample with the minimum number of vertices and consider a fixed planar embedding of G. In the following, we show that G is $(K_4,\text{diamond})$-free, thus reaching a contradiction to Theorem 3.4.3. Each claim is followed by a proof.

Claim 15. *Let C be a triangle of G. For each $v \in V(C)$, there exists a cycle of G passing through v and avoiding $V(C) \setminus \{v\}$.*

Indeed, suppose there exists $v \in V(C)$ such that every cycle through v contains at least one vertex in $V(C) \setminus \{v\}$. If T is a minimum feedback vertex set of $G - V(C)$, we have that $T \cup (V(C) \setminus \{v\})$ is a feedback vertex set of G. Therefore, by minimality, $2\nu_c(G) \geq 2\nu_c(G - V(C)) + 2 \geq \tau_c(G - V(C)) + 2 \geq \tau_c(G)$, a contradiction. ■

Claim 16. *G is K_4-free.*

Indeed, suppose G contains a copy of K_4 and let $\{v_1, v_2, v_3, v_4\}$ be its vertex set. The cycle $C = v_1v_2v_3$ is a simple closed curve in the plane and, without loss of generality, v_4 belongs to the interior of C. By Claim 15, for each $v \in V(C)$, there exists a cycle in G passing through v but avoiding $V(C) \setminus \{v\}$. In particular, each vertex of C has degree 4 in G. Let $v_1' \in N(v_1) \setminus \{v_2, v_3, v_4\}$. Since $\Delta(G) \leq 4$, a cycle C' through v_1 which avoids v_2 and v_3 must contain the edges v_1v_4 and v_1v_1'. Moreover, C' contains a vertex $v_4' \in N(v_4) \setminus \{v_1, v_2, v_3\}$. By the Jordan Curve Theorem (Theorem 1.0.5), v_1' and v_4' both belong to the interior of either $v_1v_2v_4$ or $v_1v_3v_4$. Say, without loss of generality, $\{v_1', v_4'\} \subseteq \mathrm{int}(v_1v_2v_4)$. But then, again by the Jordan Curve Theorem and the fact that $\Delta(G) \leq 4$, every cycle through v_3 contains either v_1 or v_2, a contradiction. ■

Claim 17. *G is diamond-free.*

Indeed, suppose G contains an induced diamond with vertex set $\{v_1, v_2, v_3, v_4\}$, where v_2 and v_4 are the vertices of degree 3. By Claim 15, v_2 has a neighbour $v_2' \in V(G) \setminus \{v_1, v_3, v_4\}$ and v_4 has a neighbour $v_4' \in V(G) \setminus \{v_1, v_2, v_3\}$. Since G is claw-free, either $v_2'v_1 \in E(G)$ or $v_2'v_3 \in E(G)$, say without loss of generality the former holds. Moreover, since G is K_4-free, we have $v_2' \neq v_4'$. But then either $v_4'v_1 \in E(G)$ or $v_4'v_3 \in E(G)$. Suppose first the latter holds. Clearly, $v_1v_2v_2'$ and $v_3v_4v_4'$ are two vertex-disjoint triangles and $T \cup \{v_1, v_2', v_3, v_4'\}$ is a feedback vertex set of G, for any feedback vertex set T of $G - \{v_1, v_2, v_2', v_3, v_4, v_4'\}$. But then, by the minimality of G, we have

$$2\nu_c(G) \geq 2\nu_c(G - \{v_1, v_2, v_2', v_3, v_4, v_4'\}) + 4 \geq \tau_c(G - \{v_1, v_2, v_2', v_3, v_4, v_4'\}) + 4 \geq \tau_c(G),$$

a contradiction. Therefore, $v_4'v_3 \notin E(G)$ and $v_4'v_1 \in E(G)$. By Claim 15, there exists a cycle of G passing through v_4' and avoiding v_1 and v_4. In particular, v_4' is adjacent to a and b, where $\{a, b\} \cap \{v_1, v_2, v_3, v_4\} = \varnothing$. Moreover, since G is claw-free and $\Delta(G) \leq 4$, we have $ab \in E(G)$. But then $v_1v_2v_4$ and $v_4'ab$ are two vertex-disjoint triangles and $T \cup \{v_2, v_4, a, b\}$ is a feedback vertex set of G, for any feedback vertex set T of $G - \{v_1, v_2, v_4, v_4', a, b\}$. Once again, by the minimality of G, we have

$$2\nu_c(G) \geq 2\nu_c(G - \{v_1, v_2, v_4, v_4', a, b\}) + 4 \geq \tau_c(G - \{v_1, v_2, v_4, v_4', a, b\}) + 4 \geq \tau_c(G),$$

a contradiction. ■

By Claim 16 and Claim 17, we have that G is $(K_4, \text{claw}, \text{diamond})$-free. Theorem 3.4.3 then implies that $\tau_c(G) \leq 2\nu_c(G)$, a contradiction. This concludes the proof. □

3.4.1 Subcubic graphs

In this section, we consider Jones' Conjecture for subcubic graphs and we provide a list of properties a minimum counterexample (if any) must have. In order to do so, we first present some results related to feedback vertex sets of (sub)cubic graphs, a topic which has been extensively studied (see, e.g., [77, 128, 174]). In particular, we will see that the minimum size of a feedback vertex set of a cubic graph can be expressed in terms of two "topological" parameters: the Betti number and the maximum genus.

Note that in this section we allow graphs to contain loops and multiple edges, unless otherwise stated.

Let us begin by recalling some definitions. A nonseparating independent set of a graph G is an independent set $I \subseteq V(G)$ such that there is no $X \subseteq I$ for which $G - X$ has more components than G. The maximum size of a nonseparating independent set of G is denoted by $z(G)$. The *Betti number* $\mu(G)$ of G (also known as the *cyclomatic number* or the *circuit rank*) is the minimum number of edges that must be deleted from G in order to make it acyclic. It is easy to see that if G has c components, then $\mu(G) = |E(G)| - |V(G)| + c$.

Speckenmeyer [174] showed that $\tau_c(G) + z(G) = \mu(G)$, for any connected cubic simple graph G. Ueno et al. [180] showed that the previous relation actually holds for every cubic graph (thus allowing loops and multiple edges):

Theorem 3.4.4 (Ueno et al. [180]). *If G is a cubic graph, then $\tau_c(G) + z(G) = \mu(G)$.*

In the following, we present their proof. The idea is that τ_c and z can be interpreted as the size of a minimum spanning set and the size of a maximum matching, respectively, of a suitably defined 2-polymatroid. Let us recall the definitions:

A 2-polymatroid is a pair $P = (S, f)$, where S is a finite set and f is a function $f : 2^S \to \mathbb{Z}$ satisfying the following properties:

(P1) $f(\varnothing) = 0$;

(P2) $f(X) \leq f(Y)$, for any $X \subseteq Y \subseteq S$;

(P3) $f(X \cup Y) + f(X \cap Y) \leq f(X) + f(Y)$, for any $X, Y \subseteq S$;

(P4) $f(\{x\}) \leq 2$, for any $x \in S$.

A subset $X \subseteq S$ is a matching of P if $f(X) = 2|X|$ and it is a spanning set of P if $f(X) = f(S)$.

Given a graph G, we now define a function $f : 2^{V(G)} \to \mathbb{Z}$ by $f(X) = \mu(G) - \mu(G - X)$. It is easy to see that such an f satisfies (P1) to (P3) and that the following holds:

Observation 3.4.5. *For any graph G and $v \in V(G)$, we have that $\mu(G) - \mu(G - v) \leq d_G(v) - 1$. Moreover, the inequality is strict if and only if v is a cut-vertex or the endpoint of a loop.*

Therefore, if G is a cubic graph, $P(G) = (V(G), f)$ is indeed a 2-polymatroid and we can show the following crucial result:

Theorem 3.4.6 (Ueno et al. [180]). *Let G be a cubic graph. A subset $T \subseteq V(G)$ is a feedback vertex set of G if and only if it is a spanning set of the 2-polymatroid $P(G)$. Moreover, $I \subseteq V(G)$ is a nonseparating independent set of G if and only if it is a matching of $P(G)$.*

Proof. The first assertion directly follows by the definitions. Therefore, let us consider the second one and suppose I is a nonseparating independent set of G. We proceed by induction on $|I|$. If $|I| = 1$, say $I = \{v\}$, then v is neither a cut-vertex nor the endpoint of a loop and so, by Observation 3.4.5, we have $f(I) = 2 = 2|I|$. Consider now a nonseparating independent set I with $|I| > 1$. For any $v \in I$, we have that $I \setminus \{v\}$ is a nonseparating independent set and so, by the induction hypothesis, $f(I \setminus \{v\}) = 2|I \setminus \{v\}|$. By definition, v is neither a cut-vertex of $G - (I \setminus \{v\})$ nor the endpoint of a loop and $d_{G-(I\setminus\{v\})}(v) = 3$. Therefore, by Observation 3.4.5, we have $\mu(G - (I \setminus \{v\})) - \mu(G - I) = d_{G-(I\setminus\{v\})}(v) - 1 = 2$ and so $f(I) = f(I \setminus \{v\}) + 2 = 2|I|$.

Suppose now I is not a nonseparating independent set of G. We claim that $f(I) < 2|I|$. By (P3), it is enough to consider such an I which is minimal, i.e. $I \setminus \{v\}$ is a nonseparating independent set for any $v \in I$. If I consists of one vertex v, then v is either a cut-vertex or the endpoint of a loop and so, by Observation 3.4.5, we have $f(I) < 2 = 2|I|$. If $|I| > 1$, then $I \setminus \{v\}$ is a nonseparating independent set, for any $v \in I$. On the other hand, since I is not a nonseparating independent set, either v is adjacent to some vertex in $I \setminus \{v\}$ or there exists a separating set $S \subseteq I$ containing v. Therefore, either $d_{G-(I\setminus\{v\})}(v) \leq 2$ or v is a cut-vertex of $G - (I \setminus \{v\})$. But then, by Observation 3.4.5, we have $\mu(G - (I \setminus \{v\})) - \mu(G - I) < 2$ and so $f(I) < f(I \setminus \{v\}) + 2 \leq 2|I|$. □

At this point, Theorem 3.4.4 is an immediate consequence of the "generalized Gallai's identity" stated in Theorem 1.0.9. The algorithmic implications of Theorem 3.4.6 will be discussed in Chapter 5.

Let us now introduce the seemingly unrelated notion of maximum genus [4]. The *maximum genus* $\gamma_M(G)$ of a connected graph G is the maximum integer k such that there exists a cellular

[4] We refer the reader to [20, 140] for introductions to topological graph theory.

embedding of G in the orientable surface S_k of genus k. Note that the notion of maximum genus makes sense only for connected graphs and in the following we will implicitly assume that.

Huang and Liu [97] showed a remarkable connection between nonseparating independent sets and the topological notion of maximum genus:

Theorem 3.4.7 (Huang and Liu [97]). *If G is a connected cubic graph, then $z(G) = \gamma_M(G)$.*

By Theorem 3.4.4, we can express the minimum size of a feedback vertex set in terms of the Betti number and the maximum genus:

Corollary 3.4.8. *If G is a connected cubic graph, then $\tau_c(G) = \mu(G) - \gamma_M(G)$.*

Euler's polyhedron formula tells us that if a graph G has a cellular embedding with v vertices, e edges and f faces in an orientable surface of genus g, then $v - e + f = 2 - 2g$. Therefore, for a connected graph G, introducing the Betti number $\mu(G) = e - v + 1$, we obtain $2g = \mu(G) + 1 - f$. Since every embedding has at least one face, we have $\gamma_M(G) \leq \lfloor \frac{\mu(G)}{2} \rfloor$ and a graph G is *upper-embeddable* if $\gamma_M(G) = \lfloor \frac{\mu(G)}{2} \rfloor$. In other words, a graph is upper-embeddable if and only if it admits a cellular embedding in an orientable surface with at most two faces. For example, trees and cycles are clearly upper-embeddable. Moreover, the complete graph K_4 can be embedded in the torus and the graph $K_{3,3}$ can be embedded in the double torus and so they are upper-embeddable as well.

The notion of maximum genus was introduced by Nordhaus et al. [148] and several combinatorial characterizations were subsequently found (see, e.g., [146, 147, 184]). Moreover, the maximum genus can be computed in polynomial time: Furst et al. [70] reduced this problem to a matroid matching problem and this should come as no surprise thanks to the previous discussions.

In case G is upper-embeddable, Corollary 3.4.8 translates into the following:

Corollary 3.4.9. *If G is a connected cubic upper-embeddable graph on n vertices, then $\tau_c(G) = \lfloor \frac{n}{4} \rfloor + 1$.*

Proof. Since $\tau_c(G) = \mu(G) - \gamma_M(G)$ and $\gamma_M(G) = \lfloor \frac{\mu(G)}{2} \rfloor$, we have $\tau_c(G) = \mu(G) - \lfloor \frac{\mu(G)}{2} \rfloor$. The conclusion immediately follows by substituting $\mu(G) = \frac{n}{2} + 1$. □

Let us now come back to Jones' Conjecture and recall that a counterexample with the minimum number of vertices must be 3-connected [134]. In the case of subcubic graphs, it has to further satisfy the following:

Theorem 3.4.10. *If G is a (simple) subcubic graph which is a counterexample to Jones' Conjecture and which has the minimum number of vertices, then G has girth 5, it is cubic and it is not upper-embeddable. In particular, the last property implies that G is not cyclically 4-edge-connected.*

Proof. Suppose first G contains a triangle C with $V(C) = \{v_1, v_2, v_3\}$. If T is a minimum feedback vertex set of $G - V(C)$, then $T \cup \{v_1, v_2\}$ is a feedback vertex set of G. By minimality, we have $2\nu_c(G) \geq 2\nu_c(G - V(C)) + 2 \geq \tau_c(G - V(C)) + 2 \geq \tau_c(G)$, a contradiction. Suppose now G contains a cycle C of length 4 and let v_1 and v_2 be two vertices of C at distance 2. If T is a minimum feedback vertex set of $G - V(C)$, we have that $T \cup \{v_1, v_2\}$ is a feedback vertex set of G and we obtain a contradiction by minimality as above.

Suppose now G contains a vertex v with $d(v) < 3$. If $d(v) = 1$ then, by minimality, $2\nu_c(G) = 2\nu_c(G-v) \geq \tau_c(G-v) = \tau_c(G)$. If v is a 2-vertex with neighbours u_1 and u_2, let G' be the graph obtained by deleting v and adding the edge u_1u_2. Clearly, G' is subcubic and, by the paragraph above, it is simple. Therefore, by minimality, $2\nu_c(G) \geq 2\nu_c(G') \geq \tau_c(G') \geq \tau_c(G)$, a contradiction.

An easy consequence of Euler's formula is that a connected planar graph with n vertices and girth g has at most $\frac{g}{g-2}(n-2)$ edges (see, e.g., [52]). Since G is cubic, this immediately implies that G has girth 5.

Suppose now G is upper-embeddable and let $n = |V(G)|$. By Corollary 3.4.9, we have $\tau_c(G) = \lfloor \frac{n}{4} \rfloor + 1$. Consider now the dual graph G^* of G. Since G is cubic and 3-connected, an independent set of G^* corresponds to a set of vertex-disjoint cycles of G. By Euler's formula, we have that $|V(G^*)| = \frac{n}{2} + 2$ and so the Four-Colour Theorem [164] implies that $\alpha(G^*) \geq \frac{|V(G^*)|}{4} = \frac{n+4}{8}$. Therefore, we have

$$2\nu_c(G) \geq 2\alpha(G^*) \geq \frac{n}{4} + 1 \geq \left\lfloor \frac{n}{4} \right\rfloor + 1 = \tau_c(G),$$

a contradiction.

As for the last assertion in the statement, recall that a graph G is cyclically 4-edge-connected if the deletion of fewer than 4 edges from G does not create two components both containing at least one cycle. Xuong [185] observed that cyclically 4-edge-connected cubic graphs are upper-embeddable. □

We now introduce a class of upper-embeddable graphs. A *fullerene graph* is a 3-connected cubic planar graph with all faces of size 5 or 6 (see [10] for a survey on fullerene graphs). Došlić [54] showed that fullerene graphs are cyclically 5-edge-connected and so, by [185], upper-embeddable. Therefore, by the proof of Theorem 3.4.10, we have that $\tau_c(G) \leq 2\nu_c(G)$, for any such graph G. The smallest fullerene graph is the dodecahedron graph, depicted in Figure 3.4. This graph shows that the previous bound is tight even for the class of fullerene graphs: indeed, we have $\tau_c = 6$ (for example, by Corollary 3.4.9) and $\nu_c = 3$. This gives another "non-artificial" graph (other than wheels) for which Jones' Conjecture would be tight.

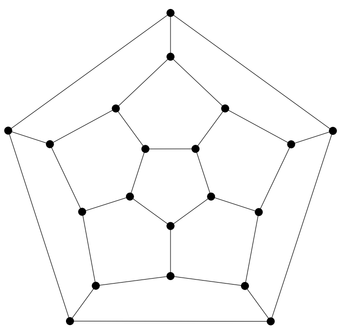


Figure 3.4: The dodecahedron graph.

Note that, in relation to Theorem 3.4.10, there exist planar cubic 3-connected graphs with girth 5 which are not upper-embeddable. The following example is due to Liu [127]. Consider a planar cubic 3-connected graph G on n vertices, with $n \geq 6$, and let H be a graph obtained from the dodecahedron by a 1-subdivision of three edges of a facial 5-cycle bounding the outerface. Clearly, H has 23 vertices and $\tau_c(H) = 6$. Let now G' be a graph obtained by replacing each vertex of G with a copy of H and by adding edges so that the contraction of each copy of H to a single vertex results in G. It is easy to see that G' is planar, cubic, 3-connected and of girth 5. Moreover, $\tau_c(G') \geq 6n > \lfloor \frac{23n}{4} \rfloor + 1$ and so, by Corollary 3.4.9, G' is not upper-embeddable.

We conclude the section with an observation. Corollary 3.4.8 implies that, for a cubic connected graph G, Jones' Conjecture is equivalent to the statement $2\nu_c(G) \geq \mu(G) - \gamma_M(G)$. An apparently similar statement was proved by Kotrbčík [113], who showed that $\nu_c(G) \geq \mu(G) - 2\gamma_M(G)$, for any connected graph G. We give the short proof to illustrate a combinatorial characterization of the maximum genus. Recall that, given a spanning tree T of a connected graph G and an edge e of G such that $e \notin E(T)$, the *fundamental cycle* of e with respect to T is the unique cycle in $T + e$. We define a graph $G \sharp T$ having as vertices the edges of $G - E(T)$, two vertices being adjacent if the fundamental cycles of the corresponding edges are not vertex-disjoint. Nebeský [147] showed that, for every connected graph G, we have $\gamma_M(G) = \min_T \alpha'(G \sharp T)$, where the minimum is taken over all spanning trees of G. Therefore, let T be a spanning tree of G attaining the minimum. The graph $G \sharp T$ has $\mu(G)$ vertices and so $\mu(G) - 2\gamma_M(G)$ of them are not covered by a maximum matching. Clearly, these vertices are pairwise non-adjacent and so the corresponding fundamental cycles are vertex-disjoint.

CHAPTER 4

VC-Dimension

Contents

In this chapter, we study the VC-dimension of certain set systems arising from graphs. In particular, we consider the set system on the vertex set of some graph which is induced by the family of its k-connected subgraphs. Kranakis et al. [115] showed that, in the case $k = 1$, the VC-dimension of this set system differs by at most 1 from the connected domination number of the underlying graph. We extend this result to each $k \geq 1$, by providing tight upper and lower bounds for the VC-dimension.

Moreover, we show that computing the VC-dimension of the set system above is NP-complete and it remains NP-complete for split graphs, for any $k \geq 1$, and for some subclasses of planar bipartite graphs in the cases $k = 1$ and $k = 2$. On the positive side, we observe it can be decided in linear time for graphs of bounded clique-width. In the final part of the chapter, we completely determine the computational complexity of this problem (in the case $k = 1$) when restricted to monogenic classes, i.e. classes of graphs obtained by forbidding a single induced subgraph: for each such class, the problem is either proved to be NP-hard or polynomial-time solvable. The same dichotomy holds for CONNECTED DOMINATING SET.

4.1 Introduction

Let $\mathcal{H}$ be a set system on X. For a subset $Y \subseteq X$, the *trace* of $\mathcal{H}$ on Y is the set $\{E \cap Y : E \in \mathcal{H}\}$ of all the possible intersections of Y with a set in $\mathcal{H}$. A subset $Y \subseteq X$ is *shattered by* $\mathcal{H}$ if $\{E \cap Y : E \in \mathcal{H}\} = 2^Y$ and the *VC-dimension* of $\mathcal{H}$ is defined as the maximum size of a set shattered by $\mathcal{H}$, or as ∞ if arbitrarily large subsets can be shattered. The VC-dimension of a set system was introduced by Vapnik and Chervonenkis [181]. The initial interest was in the contexts of empirical process theory and learning theory, where it proved to be a fundamental concept. It represents a prominent measure of the "complexity" of a set system. A result proved independently by Shelah [173], Sauer [166] and Vapnik and Chervonenkis [181] (known in the literature as Sauer's Lemma) asserts that a hypergraph on n vertices with bounded VC-dimension has at most a polynomial (in n) number of edges. More precisely, if

$\mathcal{H}$ has VC-dimension at most k, then $|E(\mathcal{H})| \leq \sum_{i=0}^{k} \binom{n}{i}$[1]. The class of hypergraphs with VC-dimension less than k can be viewed as the class of hypergraphs with a certain forbidden trace, namely the complete hypergraph on k vertices, and extremal problems on traces are a popular topic in hypergraph theory (see [11, 69] for some surveys).

The notion of VC-dimension has a particular relevance for the NP-hard problem HITTING SET [2]: Brönnimann and Goodrich [30] provided an $O(\log \mathrm{opt})$-approximation algorithm for hypergraphs with bounded VC-dimension. In fact, the link between VC-dimension and HITTING SET was first noticed in a seminal paper by Haussler and Welzl [87], who showed that hypergraphs with small VC-dimension admit hitting sets of small size for the "heavy" hyperedges, i.e. the hyperedges whose size is at least an ε-fraction of the vertex set (such hitting sets are called ε-nets).

It turns out that many hypergraphs of "geometric nature" have bounded VC-dimension and in the following we give some examples to increase the familiarity with this notion. Our examples are of the form $(X, \mathcal{R})$, where $X \subseteq \mathbb{R}^d$ is a set of points in the Euclidean space and $\mathcal{R} = \{X \cap r : r \in \mathcal{R}'\}$, for some family $\mathcal{R}'$ of geometric objects (see Chapter 1). Consider first the hypergraph of points and axis-parallel boxes in $\mathbb{R}^d$. We claim it has VC-dimension $2d$. Indeed, it is easy to see that the following set of $2d$ points in $\mathbb{R}^d$ can be shattered:

$$\{(1, 0, \ldots, 0), (-1, 0, \ldots, 0), \ldots, (0, 0, \ldots, 1), (0, 0, \ldots, -1)\} .$$

On the other hand, let X be an arbitrary subset in $\mathbb{R}^d$ of size $2d + 1$ and construct a set S as follows: for each $1 \leq i \leq d$, choose a point in X with the smallest i-th coordinate and a point with the largest i-th coordinate. Every axis-parallel box containing S contains X as well, and so no set of $2d + 1$ points in $\mathbb{R}^d$ can be shattered. Consider now the hypergraph of points and closed half-spaces in $\mathbb{R}^d$. By Radon's Theorem (Theorem 1.0.10) no set of $d + 2$ points can be shattered. On the other hand, it is not difficult to see that any $d + 1$ affinely independent points can be shattered and so the VC-dimension of this hypergraph is $d + 1$.

In this chapter, we are interested in the VC-dimension of set systems arising from graphs. We have already seen several examples of such set systems in Chapter 3 and a first example in the context of VC-dimension was given by Haussler and Welzl [87], who considered the set system induced by the closed neighbourhoods of the vertices of a graph. Since a graph of order n has n closed neighbourhoods, its VC-dimension is at most $\lfloor \log_2 n \rfloor$. On the other hand, if the graph is planar, it is not difficult to see that the VC-dimension is at most 4 (see below for a proof).

More generally, we can consider set systems induced by a certain family of subgraphs. In this way we obtain several different notions of VC-dimension, each one related to a special family of subgraphs. Kranakis et al. [115] initiated a systematic study of these notions and adapted the definition of VC-dimension to the graph theoretic setting as follows:

Definition 4.1.1. Let $G = (V, E)$ be a graph and let $\mathcal{P}$ be a family of subgraphs of G. A subset $A \subseteq V$ is *$\mathcal{P}$-shattered* if every subset of A can be obtained as the intersection of $V(H)$ with A, for $H \in \mathcal{P}$. The *VC-dimension* of G with respect to $\mathcal{P}$ is the maximum size of a $\mathcal{P}$-shattered subset and it is denoted by $\mathrm{VC}_{\mathcal{P}}(G)$.

According to Definition 4.1.1, we denote by $\mathrm{VC}_{\mathrm{tree}}$, $\mathrm{VC}_{\mathrm{con}}$, $\mathrm{VC}_{k\text{-con}}$, $\mathrm{VC}_{\mathrm{nbd}}$, $\mathrm{VC}_{\mathrm{path}}$, $\mathrm{VC}_{\mathrm{cycle}}$

[1] A beautiful proof of Sauer's Lemma using the so-called linear algebra method was given by Frankl and Pach [65] (see also [16]). [2] The problem of deciding, given a hypergraph $\mathcal{H}$ and an integer k, whether $\tau(\mathcal{H}) \leq k$ holds (see Chapter 1).

and VC_{star} the VC-dimensions with respect to families of subgraphs which are trees, connected, k-connected, induced by the closed neighbourhoods of the vertices, paths, cycles and stars, respectively. Note that the VC-dimension with respect to some families of subgraphs is equal to well-known parameters in graph theory: if $\mathcal{P}$ is the family of complete subgraphs then $VC_{\mathcal{P}}$ is the clique number, while if $\mathcal{P}$ is the family of subgraphs induced by independent sets then $VC_{\mathcal{P}}$ is the independence number.

We have already noticed that $VC_{\text{nbd}}(G) \leq \lfloor \log_2 |V(G)| \rfloor$ and it is not difficult to see that this bound is tight [12]. Indeed, consider the graph H built as follows. Take a set S of $\lfloor \log_2 n \rfloor$ independent vertices and, for each non-singleton subset $R \subseteq S$, add a vertex v_R adjacent to precisely the vertices of R. The resulting graph H has at most n vertices and $VC_{\text{nbd}}(H) = \lfloor \log_2 n \rfloor$.

Anthony et al. [12] showed that if G contains no subdivision of K_{r+1}, then $VC_{\text{nbd}}(G) \leq r$. Let us briefly sketch the proof. Suppose, to the contrary, that there exists a shattered set S of size $r+1$ and let x and y be two non-adjacent vertices in S. Since S is shattered, there exists a vertex w such that $\{x, y\} = N[w] \cap S$. Since x and y are non-adjacent, we have $w \neq x$ and $w \neq y$. Therefore, w is a vertex in $V(G) \setminus S$ such that the only vertices in S adjacent to w are x and y. But then the subgraph induced by S and by the vertices w as above contains a subdivision of K_{r+1}, a contradiction.

Kranakis et al. [115] observed that if G is a graph with maximum degree Δ, then $\Delta \leq VC_{\text{star}}(G) \leq \Delta + 1$. Moreover, they showed that the VC-dimension with respect to trees is the same as the VC-dimension with respect to connected subgraphs: this is an immediate consequence of the fact that a connected graph contains a spanning tree. In addition, $VC_{\text{con}}(G)$ differs by at most 1 from the number of leaves $\ell(G)$ in a maximum leaf spanning tree of G:

Theorem 4.1.2 (Kranakis et al. [115]). *$\ell(G) \leq VC_{\text{con}}(G) \leq \ell(G) + 1$, for any connected graph G.*

Note that the VC-dimension of a graph with respect to connected subgraphs is the maximum of the VC-dimensions of its components.

In this chapter, we continue the systematic study of the VC-dimension of set systems defined by graphs which was initiated in [115]. In particular, we focus on the VC-dimension with respect to k-connected subgraphs. Given a graph G, this quantity can be thought as the maximum size of a subset $A \subseteq V(G)$ such that, no matter how many vertices of A are deleted from G, there is a k-connected subgraph of G containing the remaining vertices of A. In Section 4.2, we extend Theorem 4.1.2 by giving tight upper and lower bounds on the VC-dimension with respect to k-connected subgraphs, for $k \geq 2$. These are given, similarly to Theorem 4.1.2, in terms of the number of leaves in a maximum leaf spanning tree. By Definition 4.1.1, it follows that if $\mathcal{P} \subseteq \mathcal{P}'$, then $VC_{\mathcal{P}}(G) \leq VC_{\mathcal{P}'}(G)$. Therefore, for any connected graph G and $k \geq 2$, Theorem 4.1.2 implies that $VC_{k\text{-con}}(G) \leq \ell(G) + 1$ and we show that this bound can be improved to the optimal $VC_{k\text{-con}}(G) \leq \ell(G) - k + 1$.

Papadimitriou and Yannakakis [155] considered the problem of deciding the VC-dimension: Given a set system $\mathcal{H}$ on X (by its incidence matrix) and an integer s, does $\mathcal{H}$ have VC-dimension at least s? Since the VC-dimension is at most $\log_2 |\mathcal{H}|$, it can clearly be computed by brute force in $O(|X|^{\log_2 |\mathcal{H}|})$ time and so the problem looks unlikely to be NP-complete. In fact, they introduced the complexity class LOGNP and showed that the problem in question is complete for it (see [155] for all the details). On the other hand, if the set system is represented by a circuit, Schaefer [168] showed that deciding the VC-dimension is complete for the third level of the polynomial hierarchy (see Chapter 1).

In this context, it is natural to investigate the computational complexity of computing $\mathrm{VC}_{\mathcal{P}}(G)$, for a given graph G and a family of its subgraphs $\mathcal{P}$. The decision problem is formulated as follows:

GRAPH $\mathrm{VC}_{\mathcal{P}}$ DIMENSION

Instance: A graph G and a number $s \geq 1$.
Question: Does $\mathrm{VC}_{\mathcal{P}}(G) \geq s$ hold?

Kranakis et al. [115] showed that GRAPH $\mathrm{VC}_{\mathrm{con}}$ DIMENSION is NP-complete. In Section 4.3, we extend this result by showing that GRAPH $\mathrm{VC}_{k\text{-con}}$ DIMENSION is NP-complete even for split graphs, for any $k \geq 1$. On the positive side, we show it can be decided in linear time for graphs of bounded clique-width and in polynomial time for the subclass of split graphs having Dilworth number at most 2. Finally, we prove that GRAPH $\mathrm{VC}_{\mathrm{con}}$ DIMENSION and GRAPH $\mathrm{VC}_{2\text{-con}}$ DIMENSION remain NP-hard for some subclasses of planar bipartite graphs with maximum degree at most 4. The following table summarizes the known results on the computational complexity of GRAPH $\mathrm{VC}_{\mathcal{P}}$ DIMENSION:

Family $\mathcal{P}$	Graph G	Comp. Compl.	Reference
star		P	Kranakis et al. [115]
neighbourhood		LOGNP-complete	Kranakis et al. [115]
path		Σ_3^{P}-complete	Schaefer [169]
cycle		Σ_3^{P}-complete	Schaefer [169]
k-connected	split	NP-complete	Theorem 4.3.1
k-connected	bounded clique-width	P	Corollary 4.3.3
k-connected	split, Dilworth number ≤ 2	P	Theorem 4.3.4
connected	planar, bipartite, $\Delta(G) = 3$	NP-complete	Theorem 4.3.5
2-connected	planar, bipartite, $\Delta(G) = 4$	NP-complete	Theorem 4.3.6

It is interesting to notice how GRAPH $\mathrm{VC}_{\mathcal{P}}$ DIMENSION is a source of natural problems with such diverse complexities. Note also that GRAPH $\mathrm{VC}_{\mathrm{path}}$ DIMENSION and GRAPH $\mathrm{VC}_{\mathrm{cycle}}$ DIMENSION are some of the very few known problems which are complete for the third level of the polynomial hierarchy (see [170]).

A class of graphs is *monogenic* if it is defined by a single forbidden induced subgraph. Moreover, we say that a (decision) graph problem admits a dichotomy in monogenic classes if, for each monogenic class, the problem is either NP-complete or decidable in polynomial time. In Section 4.4, we provide complexity dichotomies in monogenic classes for GRAPH $\mathrm{VC}_{\mathrm{con}}$ DIMENSION and CONNECTED DOMINATING SET. The first dichotomy in monogenic classes was obtained by Korobitsin [108], who showed that DOMINATING SET is decidable in polynomial time for $\mathit{Free}(H)$ if H is an induced subgraph of $P_4 + tK_1$, for $t \geq 0$, and NP-complete otherwise. Besides this, only few other dichotomies in monogenic classes are known, most notably for COLOURING [114] and ℓ-LIST COLOURING [79]. Theorem 4.1.2 hints at the fact that GRAPH $\mathrm{VC}_{\mathrm{con}}$ DIMENSION and CONNECTED DOMINATING SET are two related problems (recall that $\ell(G) = |V(G)| - \gamma_c(G)$, where $\gamma_c(G)$ is the minimum size of a connected dominating set of G). In fact, we show that the complexities of GRAPH $\mathrm{VC}_{\mathrm{con}}$ DIMENSION, CONNECTED DOMINATING SET and DOMINATING SET all agree in monogenic classes.

4.2 Bounds on the VC-dimension

The main result of this section is an extension of Theorem 4.1.2 by considering families of k-connected subgraphs, for $k \geq 2$:

Theorem 4.2.1. $\mathrm{VC}_{k\text{-con}}(G) \leq \ell(G) - k + 1$*, for any connected graph* G *and* $k \geq 2$*.*

The proof of Theorem 4.2.1 consists essentially of two parts. First, we show that $\mathrm{VC}_{k\text{-con}}(G) \leq \ell(G) - k + 2$, for any connected graph G and $k \geq 1$ (the case $k = 1$ thus gives a proof for the upper bound in Theorem 4.1.2). The idea is to construct a spanning tree T with at least $\mathrm{VC}_{k\text{-con}}(G) + k - 2$ leaves. We fix a shattered set A of maximum cardinality and choose an appropriate vertex $r \in A$ as the root. Then we consider some k neighbours of r, say $u_1, \ldots, u_k$, and we try to "attach" the remaining vertices of A to the graph $(\{r, u_1, \ldots, u_k\}, \{ru_1, \ldots, ru_k\})$ via appropriate paths. A crucial step is the distinction between two types of vertices of A that we are going to add as leaves of T: lower leaves and upper leaves (see below for definitions). In the second part, we provide the actual proof of Theorem 4.2.1 by contradiction. In particular, we suppose that $\ell(G) \leq \mathrm{VC}_{k\text{-con}}(G) + k - 2$ and we show how to modify the tree T constructed in the first part in order to obtain a contradiction.

Proof of Theorem 4.2.1. We begin by showing that $\mathrm{VC}_{k\text{-con}}(G) \leq \ell(G) - k + 2$, for any connected graph G and $k \geq 1$. Let A be a shattered set of maximum cardinality. Since our aim is to construct a spanning tree with at least $|A| + k - 2$ leaves, note that it is enough to construct a tree $T \subseteq G$ with as many leaves. For $w \in A$, we denote by G_w a fixed k-connected subgraph such that $V(G_w) \cap A = \{w\}$. Similarly, $G_{ww'}$ denotes a fixed k-connected subgraph such that $V(G_{ww'}) \cap A = \{w, w'\}$. We choose a vertex $r \in A$ having the minimum number of neighbours in $V(G) \setminus A$ as the root of T. Clearly, $d_{G_r}(r) \geq k$ and let $U = \{u_1, \ldots, u_k\}$ be a set of k arbitrary vertices in $N_{G_r}(r)$[3]. We select the edges $u_1 r, \ldots, u_k r$. By Menger's Theorem (Theorem 1.0.3), there exist k independent w, r-paths in G_{wr}. We call $w \in A \setminus \{r\}$ a *lower leaf* for T if there exist k independent w, r-paths $P_1, \ldots, P_k$ in G with no inner vertex in A and such that each of them contains (exactly) one edge in $\{u_1 r, \ldots, u_k r\}$ (in other words, w is a lower leaf if there exists a w, U-fan in $G - (A \setminus \{w\})$ of size k). In particular, no path P_i contains two vertices in U. Otherwise, we call w an *upper leaf* for T.

We set $L = \{u_1, \ldots, u_k\}$ and we view L as the set of potential leaves for T. For any $w \in A \setminus \{r\}$, we do the following (see Figure 4.1):

- If w is an upper leaf, select a w, r-path $P \subseteq G_{wr}$ such that $V(P) \cap U = \varnothing$. Such a path exists: by Menger's Theorem (Theorem 1.0.3), there exist k independent w, r-paths in G_{wr} and, if each of them contained a (different) vertex in U, we would obtain a w, U-fan in $G - (A \setminus \{w\})$ of size k. Finally, add w to L and remove cycles and appropriate edges so that the selected subgraph is a tree.

- If w is a lower leaf, select a w, r-path $P \subseteq G$ as in the definition of lower leaf and such that $E(P) \cap \{u_1 r, \ldots, u_k r\} = \{u_1 r\}$, add w to L and remove u_1 from L. Finally, delete edges from the newly added paths so that the selected subgraph is a tree.

In this way, we get a tree T, rooted at r and in which the elements of L are leaves. Moreover, $|L| \geq |A| + k - 2$. The construction works for any k, the case $k = 1$ giving the upper bound in Theorem 4.1.2.

[3] If $k = 1$, we simply let u_1 be any vertex in $V(G) \setminus A$.

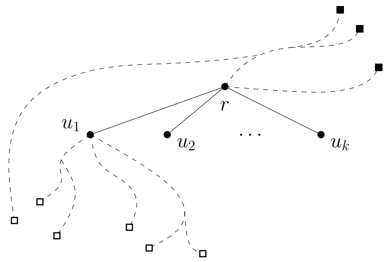


Figure 4.1: The black squared vertices are the upper leaves, while the white squared vertices are the lower leaves. The selected paths are dashed.

From now on, we assume $k \geq 2$ and we prove Theorem 4.2.1 by contradiction. Therefore, let G be a counterexample and let A be a shattered set of maximum cardinality. Using the procedure described in the previous part, we can obtain a tree $T \subseteq G$, rooted at r and with at least $|A| + k - 2$ leaves. Recall that r is chosen as a vertex of A having the minimum number of neighbours in $V(G) \setminus A$. In the following, we deduce some structural properties of G and show how they lead to a contradiction. Each claim is followed by a short proof.

Claim 18. *Each leaf of T is adjacent to at most one vertex not in T.*

Indeed, if there exists a leaf of T adjacent to at least two vertices not in T, we immediately get a tree with at least $|A| + k - 1$ leaves. ■

Claim 19. *T contains at least one upper leaf.*

Indeed, suppose T has no upper leaves. For any $2 \leq i \leq k$, select a u_i, u_1-path in G_r with no inner vertex in $(U \setminus \{u_1, u_i\}) \cup \{r\}$. Clearly, such a path has no inner vertex in $A \cup U$. But then we can obtain a tree, rooted at u_1, with at least $|A| + k - 1$ leaves (r becomes an additional leaf). ■

Claim 20. *T contains an upper leaf w with $d_T(w, r) \geq 2$.*

Indeed, suppose this is not the case. By Claim 19, the set W of upper leaves for T is non-empty and each of them is adjacent to r. Suppose now every $w \in W$ has one neighbour (in G) which is contained in $T_{u_1} - A$, where T_{u_1} denotes the subtree induced by u_1 and its descendants. Then we select the edges joining each upper leaf to $T_{u_1} - A$ and, for any $2 \leq i \leq k$, we select a u_i, u_1-path in G_r with no inner vertex in $A \cup U$. In this way, we obtain a tree rooted at u_1 and with at least $|A| + k - 1$ leaves, a contradiction. Therefore, there exists $w \in W$ such that $N_G(w) \cap (V(T_{u_1}) \setminus A) = \varnothing$. By Claim 18, w has at most one neighbour in $V(G) \setminus V(T)$ and so $N_{G_w}(w) \subseteq N_G(w) \setminus A \subseteq \{u_2, \ldots, u_k\} \cup \{x\}$, for some $x \in V(G) \setminus V(T)$. But since w has at least k neighbours in G_w, we have $N_{G_w}(w) = N_G(w) \setminus A = \{u_2, \ldots, u_k\} \cup \{x\}$.

Consider now $w' \in A$ such that $ww' \notin E(G)$. There exists a w, w'-path P in $G_{ww'}$ with no inner vertex in $U \setminus \{u_1\}$. Clearly, $wx \in E(P)$. Moreover, P does not contain any vertex in $T_{u_1} - A$, or else x could become an additional leaf of T. By selecting these paths, together with edges connecting vertices in A to w, it is easy to see we can get a new tree T' rooted at w and with at least $|A| + k - 2$ leaves. Moreover, if an upper leaf for T has at least two neighbours in the subtree $T_{u_1} - A$, then we can get an additional leaf for T'. Therefore, each upper leaf for T has at most one neighbour in $T_{u_1} - A$.

Now let $w' \neq w$ be an upper leaf for T. We claim that w' is adjacent to u_2. Indeed, suppose $w'u_2 \notin E(G)$. By Claim 18, w' has at most one neighbour in $G - V(T)$ and by the paragraph above it has at most one neighbour in $T_{u_1} - A$. Since $d_{G_{w'}}(w') \geq k$, we have that $N_{G_{w'}}(w') = (U \setminus \{u_1, u_2\}) \cup \{y, z\}$, for some $y \in V(G) \setminus V(T)$ and $z \in V(T_{u_1}) \setminus A$. Moreover, $u_2 \notin V(G_{w'})$. Otherwise, there exists a w', u_2-path P in $G_{w'}$ with no inner vertex in $(U \setminus \{u_1, u_2\}) \cup \{z\}$. Clearly, $w'y \in E(P)$ and P does not contain vertices in $T_{u_1} - A$, or else y would become an additional leaf of T. But then there exists a w', U-fan in $G - (A \setminus \{w'\})$ of size k, contradicting the fact that w' is an upper leaf. Therefore, $u_2 \notin V(G_{w'})$ and so there exists a y, z-path in $G_{w'}$ with no inner vertex in $A \cup (U \setminus \{u_1\})$. Again, by adding this path to the initial tree T, we get a new tree with at least $|A| + k - 1$ leaves, a contradiction. This means that $w'u_2 \in E(G)$, for any upper leaf w'. On the other hand, for any lower leaf w'', there exists a w'', u_2-path with no inner vertex in $A \cup U$. Finally, for any $u_i \in U \setminus \{u_2\}$, there exists a u_i, u_2-path in G_r with no inner vertex in $A \cup U$. But then it is easy to construct a new tree rooted at u_2 and with at least $|A| + k - 1$ leaves. ■

Claim 21. *The vertex r has at most one neighbour in $V(T) \setminus (A \cup U)$.*

Indeed, suppose r has at least two neighbours in $V(T) \setminus (A \cup U)$. The subtree $T - (\{u_2, \ldots, u_k\} \cup (A \setminus \{r\}))$ contains a leaf q and let $Q = N_T(q) \cap (A \setminus \{r\})$. Moreover, let $w \in Q$ and consider G_w. Clearly, w has at least $k - 1 \geq 1$ neighbours in $V(G) \setminus (A \cup \{q\})$. If for any $w \in Q$, one of these neighbours is contained in $V(T) \setminus (\{u_2, \ldots, u_k\} \cup \{q\})$, then we get a tree rooted at r and in which q is an additional leaf. Therefore, there exists $w \in Q$ with no neighbours in $V(T) \setminus (\{u_2, \ldots, u_k\} \cup \{q\} \cup A)$. By Claim 18, w has at most one neighbour in $V(G) \setminus V(T)$. But then w has at most $k + 1$ neighbours in $V(G) \setminus A$, contradicting the minimality of r. ■

By Claim 21 and Claim 20, r has exactly $k + 1$ neighbours in $V(T) \setminus A$. Therefore, let $N_T(r) \setminus (A \cup U) = \{z\}$ and let T_z be the subtree induced by z and its descendants.

Claim 22. $V(T_{u_1}) \cap V(T_z) = \varnothing$.

Indeed, suppose there exists $x \in V(T_{u_1}) \cap V(T_z)$ and let $P(x)$ be the unique x, r-path in T. By definition, $P(x)$ contains both z and u_1 and so a cycle arises in T. ■

Claim 23. *No leaf of T_z is a lower leaf for T.*

Indeed, suppose there exists a lower leaf w for T which is a leaf of T_z. This means that the w, u_1-subpath P in the definition of w intersects T_z. Then, select the u_1, x-subpath of P, where x is the first intersection of P with T_z when traversed from u_1. Moreover, for any $u_i \in U \setminus \{u_1\}$, select a u_i, u_1-path in G_r with no inner vertex in $A \cup U$ and delete the edge set $\{u_2 r, \ldots, u_k r, rz\}$. After removing cycles and appropriate edges from the selected subgraph, we get a tree T' rooted at u_1 and with at least $|A| + k - 2$ leaves. Finally, by Claim 18 and the minimality of r, any upper leaf for T adjacent to r has a neighbour in $V(T') \setminus (U \setminus \{u_1\} \cup A)$ and so r could become an additional leaf of T'. ■

Clearly, the tree $T_z - A$ contains a leaf q and let $Q = N_T(q) \cap A$. By Claim 23, Q is a set of upper leaves. By an argument similar to the one in the proof of Claim 21, there exists $w' \in Q$ such that $N_{G_{w'}}(w') \subseteq N_G(w') \setminus A \subseteq (U \setminus \{u_1\}) \cup \{q, x\}$, for some $x \in V(G) \setminus V(T)$, and w' is adjacent to at least $k-2$ vertices in $U \setminus \{u_1\}$ (see Figure 4.2). Moreover, $U \setminus \{u_1\} \subseteq V(G_{w'})$, or else $\{q, x\} \subseteq V(G_{w'})$ and there would exist an x, q-path in $G_{w'}$ with no inner vertex in $A \cup (U \setminus \{u_1\})$, thus turning x into an additional leaf of T.

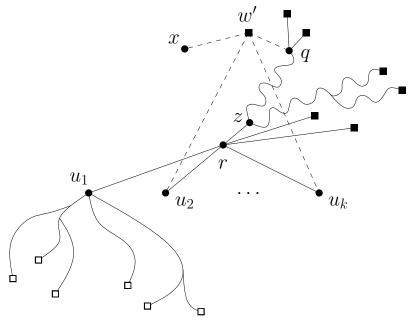


Figure 4.2: Neighbourhood of w'.

Claim 24. *There is no path P (in G) between w' and $v \in V(T_{u_1}) \setminus A$ with no inner vertex in $A \cup (U \setminus \{u_1\})$.*

Suppose such a path P exists. Then $w'x \notin E(P)$, or else x becomes an additional leaf. Therefore, $w'q \in E(P)$. By the paragraph above, w' is adjacent to at least $k-2$ vertices in $U \setminus \{u_1\}$, say $\{w'u_3, w'u_4, \ldots, w'u_k\} \subseteq E(G)$, and there exists a w', u_2-path $P' \subseteq G_{w'}$ with no inner vertex in $(U \setminus \{u_1, u_2\}) \cup \{q\}$. If $w'u_2 \in E(G)$, then there exists a w', U-fan in $G - (A \setminus \{w'\})$ of size k, contradicting the fact that w' is an upper leaf. Therefore, $w'u_2 \notin E(G)$ and $w'x \in E(P')$. But then $P + P(v)$ (where $P(v)$ is the unique v, u_1-path in T) and P' do not intersect, or else x becomes an additional leaf and once again we obtain a w', U-fan in $G - (A \setminus \{w'\})$ of size k, a contradiction. ■

Claim 25. *Each $w \in A$ has at most one neighbour in $T_{u_1} - A$.*

Consider $w \in A \setminus \{w'\}$ such that $w'w \notin E(G)$. There exists a w', w-path P in $G_{w'w}$ with no inner vertex in $A \cup (U \setminus \{u_1\})$. Moreover, by Claim 24, P does not contain any vertex of $T_{u_1} - A$. For any $w \in A \setminus \{w'\}$ such that $w'w \notin E(G)$, we select these paths. For any $w \in A \setminus \{w'\}$ such that $w'w \in E(G)$, we select the corresponding edges. Moreover, recall that w' is adjacent to at least $k-2$ vertices in $U \setminus \{u_1\}$, say $\{u_3, \ldots, u_k\}$, and $U \setminus \{u_1\} \subseteq V(G_{w'})$. But then there exists a w', u_2-path in $G_{w'}$ with no inner vertex in $A \cup U$ and so, by Claim 24, with no inner vertex in $T_{u_1} - A$ as well. Therefore, we can obtain a new tree T' rooted at w' and with at least $|A| + k - 2$ leaves. But then each $w \in A$ has at most one neighbour in $T_{u_1} - A$, or else we could get an additional leaf of T'. ■

Claim 26. *If w is an upper leaf, then $((U \setminus \{u_1\}) \cup \{v\}) \nsubseteq V(G_w)$, for any $v \in V(T_{u_1})$.*

Indeed, suppose this is not the case and let $v \in V(T_{u_1})$ be a vertex with minimum $d_T(v, u_1)$ among those satisfying $((U \setminus \{u_1\}) \cup \{v\}) \subseteq V(G_w)$. Since the set $U' = (U \setminus \{u_1\}) \cup \{v\}$ is contained in $V(G_w)$, there exists a w, U'-fan in G_w of size k (Lemma 1.0.4). By minimality, no path in the fan intersects the unique v, u_1-path in T in a vertex different from v and so we can obtain a w, U-fan in $G - (A \setminus \{w\})$ of size k, contradicting the fact that w is an upper leaf. ■

Claim 27. *$V(G_w) \cap V(T_{u_1}) = \varnothing$, for any $w \in Q$.*

Indeed, suppose there exists $w \in Q$ and a vertex s such that $s \in V(G_w) \cap V(T_{u_1})$. By Claim 26, we have $U \setminus \{u_1\} \nsubseteq V(G_w)$. By Claim 25, w has at most one neighbour in $T_{u_1} - A$ and, by Claim 18, w has at most one neighbour in $V(G) \setminus V(T)$. But then there exists $v \in V(G_w) \cap V(T_z)$ with $v \neq w$, or else $N_{G_w}(w) = (U \setminus \{u_1, u_i\}) \cup \{x, y\}$, for some $x \in V(G) \setminus V(T)$, $y \in V(T_{u_1}) \setminus A$ and $2 \leq i \leq k$, and we could find an x, y-path in G_w with no inner vertex in $A \cup (U \setminus \{u_1\})$. Therefore, there exists a v, s-path $P \subseteq G_w$ with no inner vertex in $A \cup (U \setminus \{u_1\})$, contradicting Claim 24. ■

We now show how to build a tree rooted at u_2 and with at least $|A| + k - 1$ leaves. This will provide a contradiction, thus concluding the proof.

Consider $w \in Q$. By Claim 27 and Claim 18, if $(V(G_w) \setminus \{w\}) \cap (V(T_z) \setminus \{q\}) = \varnothing$, then $N_{G_w}(w) \subseteq (U \setminus \{u_1\}) \cup \{q, x\}$, for some $x \in V(G) \setminus V(T)$. Moreover, as we have seen before, we have that $U \setminus \{u_1\} \subseteq V(G_w)$, or else x could become an additional leaf of T.

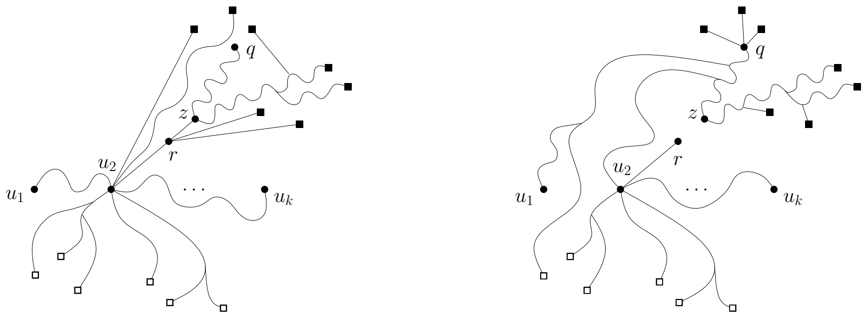


Figure 4.3: The different constructions of a tree T with at least $|A| + k - 1$ leaves.

We start by deleting $V(T_{u_1}) \cup U$ from T. Suppose now that, for every lower leaf w'', the w'', u_2-subpath P in the definition of w'' contains no vertices of T_z and, for every $u_i \in U \setminus \{u_2\}$, each u_i, u_2-path P' in G_r with no inner vertex in $A \cup U$ contains no vertices of T_z. In particular, these paths do not contain q and we select them (see Figure 4.3(a)). Moreover, for any $w \in Q$, we do the following. If there exists $v \in (V(G_w) \setminus \{w\}) \cap (V(T_z) \setminus \{q\})$, then we select a w, v-path in G_w with no inner vertex in $A \cup (U \setminus \{u_2\}) \cup \{q\}$ (such a path exists by Claim 27). Otherwise, by the paragraph above, we select a w, u_2-path in G_w with no inner vertex in $A \cup U \cup \{q\}$ (again, the existence follows from Claim 27). After removing cycles and appropriate edges we get a new tree T, rooted at u_2 and with at least $|A| + k - 1$ leaves (q becomes an additional leaf), a contradiction.

Therefore, there exists either a lower leaf $\overline{w}$ such that the $\overline{w}, u_2$-subpath P in the definition of $\overline{w}$ contains vertices of T_z, or a $u_i \in U \setminus \{u_2\}$ such that a u_i, u_2-path P' in G_r with no inner vertex in $A \cup U$ contains vertices of T_z. Then we select either the $\overline{w}, x$-subpath and the y, u_2-subpath of P, where x and y are, respectively, the first and the last intersection of P with T_z when traversed from $\overline{w}$, or the u_i, x'-subpath and the y', u_2-subpath of P', where x' and y' are, respectively, the first and the last intersection of P' with T_z when traversed from u_i (see

Figure 4.3(b)). In this way, we get a new tree T' and we grow it as follows. Let w'' be a lower leaf and $P_{w''}$ the w'', u_2-subpath in the definition of w''. For any lower leaf w'', we select the $w'', x_{w''}$-subpath of $P_{w''}$, where $x_{w''}$ is the first intersection of $P_{w''}$ with the tree constructed so far and in which w'' becomes a leaf. Similarly, we add the remaining vertices of $U \setminus \{u_2\}$ as leaves. Finally, consider an upper leaf w'' for the original T such that $d_T(w'', r) = 1$. By the minimality of r and by an argument similar to Claim 18, w'' is adjacent to a vertex in $V(T') \setminus (A \cup U \setminus \{u_2\})$. But then we can add these edges to the tree T' rooted at u_2 in order to obtain a new tree with at least $|A| + k - 1$ leaves (r becomes an additional leaf), a contradiction. This concludes the proof. □

Our bound is tight in the sense of the following:

Proposition 4.2.2. *For any $k \geq 2$ and $x \geq 2k$, there exists a graph G with $\ell(G) = x$ and $\mathrm{VC}_{k\text{-con}}(G) = \ell(G) - k + 1$.*

For the proof we need to recall the Expansion Lemma (Lemma 1.0.2), which will be used in the upcoming sections as well: If G is a k-connected graph and G' is obtained from G by adding a new vertex with at least k neighbours in G, then G' is k-connected.

Proof of Proposition 4.2.2. For a fixed $k \geq 2$ and $x = 2k$, consider the graph G_k constructed as follows. Start with a clique H_k of size $k+1$. For each subset $S \subset H_k$ of size k, add a vertex adjacent to precisely the vertices of S, and let A be the set of the added vertices. Clearly, $|V(G_k)| = 2k+2$ and $\ell(G_k) = 2k$. Moreover, by the Expansion Lemma, A is shattered. For $x \geq 2k$, let G be the graph obtained from G_k by adding $x - 2k$ vertices adjacent to exactly k vertices of H_k. It is easy to see that $\ell(G) = x$ and $\mathrm{VC}_{k\text{-con}}(G) = \ell(G) - k + 1$. □

Let us now consider a lower bound for the VC-dimension. Theorem 4.1.2 asserts that $\mathrm{VC}_{\mathrm{con}}(G) \geq \ell(G)$, for any connected graph G. Indeed, consider a spanning tree T of G and its set of leaves L. For any subset $A \subseteq L$, there exists a subtree of T whose set of leaves is A and so the set L can be shattered by connected subgraphs. Therefore, in the case $k = 1$, we have that $\mathrm{VC}_{\mathrm{con}}(G)$ is either $\ell(G)$ or $\ell(G) + 1$, for any connected graph G. The situation changes for $k \geq 2$, as the difference $\ell(G) - \mathrm{VC}_{k\text{-con}}$ can be arbitrarily large. A simple example in the case $k = 2$ is given by the wheel W_n: we have that $\mathrm{VC}_{2\text{-con}}(W_n) = 3$ and $\ell(W_n) = n$.

Nevertheless, we can lower bound $\mathrm{VC}_{k\text{-con}}(G)$ in terms of $\ell(G)$, $|V(G)|$ and $|E(G)|$. The obvious idea is that having a sufficiently large complete subgraph guarantees shattering by k-connected subgraphs.

Theorem 4.2.3. *Let G be a connected graph of order n, size m, and maximum degree Δ. For $k \geq 2$,*

$$\mathrm{VC}_{k\text{-con}}(G) \geq \ell(G) - k + 1 - \left(n + 2 - \left\lceil \frac{n-2}{\Delta - 1} \right\rceil - \frac{n^2}{n^2 - 2m}\right).$$

Proof. By Turán's theorem (Theorem 1.0.1), if $m > \left(1 - \frac{1}{r}\right)\frac{n^2}{2}$, then G contains K_{r+1} as a subgraph. Therefore, by the Expansion Lemma, a set of size $r + 1 - (k+1)$ can be shattered by k-connected subgraphs. The condition above is equivalent to $r < \frac{n^2}{n^2 - 2m}$ and so, taking

$r = \left\lceil \frac{n^2}{n^2-2m} - 1 \right\rceil$, we get

$$\mathrm{VC}_{k\text{-con}}(G) \geq \left\lceil \frac{n^2}{n^2-2m} - 1 \right\rceil + 1 - (k+1).$$

Let now T be a spanning tree of G and $d_i = |\{v \in V(T) : d_T(v) = i\}|$. We want to find an upper bound for $\ell(G)$. We have

$$\sum_{i=1}^{\Delta} d_i = n \quad \text{and} \quad 2(n-1) = \sum_{v \in V(T)} d_T(v) = \sum_{i=1}^{\Delta} i d_i.$$

Using the two relations above, it is easy to see that

$$n - d_1 = \sum_{i=2}^{\Delta} d_i \geq \sum_{i=2}^{\Delta} \frac{i-1}{\Delta-1} d_i = \frac{1}{\Delta-1} \sum_{i=2}^{\Delta} (i-1) d_i = \frac{n-2}{\Delta-1}.$$

Since $\ell(G)$ is the maximum of d_1 taken over all spanning trees of G, we have

$$\ell(G) \leq n - \left\lceil \frac{n-2}{\Delta-1} \right\rceil.$$

Summarizing, we get

$$\begin{aligned}
\mathrm{VC}_{k\text{-con}}(G) &\geq \left\lceil \frac{n^2}{n^2-2m} - 1 \right\rceil - k \\
&\geq \frac{n^2}{n^2-2m} - 1 - k + \left(\ell(G) - n + \left\lceil \frac{n-2}{\Delta-1} \right\rceil \right) \\
&\geq \ell(G) - k + 1 - \left(n + 2 - \left\lceil \frac{n-2}{\Delta-1} \right\rceil - \frac{n^2}{n^2-2m} \right).
\end{aligned}$$ $\square$

Complete graphs show that the bound in Theorem 4.2.3 is tight.

4.3 The decision problem

In this section we show that Graph $\mathrm{VC}_{k\text{-con}}$ Dimension is NP-complete, for any k. The case $k = 1$ was addressed by Kranakis et al. [115] and our proof is in fact inspired by theirs.

Consider the following decision problem, usually called Set Multicover:

Set Multicover	
Instance:	A set $S = \{a_1, \ldots, a_n\}$, a collection of subsets $S_1, \ldots, S_m \subseteq S$, and integers k and t.
Question:	Is there an index set $I \subseteq \{1, \ldots, m\}$ such that $\bigcup_{i \in I} S_i = S$, each a_i is covered by at least k distinct subsets, and $\lvert I \rvert \leq t$?

Being a generalization of the well-known Set Cover (also known as Minimum Cover [75]), it is NP-complete and we use it for our reduction. Recall that a split graph is a graph in which the vertex set can be partitioned into a clique and an independent set.

Theorem 4.3.1. Graph VC$_{k\text{-con}}$ Dimension *is* NP*-complete even for split graphs.*

Proof. First, we show that the problem is in NP. Our proof is based on the following elementary result:

Claim 28. *If G and G' are two k-connected graphs such that $|V(G) \cap V(G')| \geq k$, then $G \cup G'$ is k-connected as well.*

Indeed, let $S \subset V(G \cup G')$ be a subset such that $|S| < k$ and let v and w be two distinct vertices in $(G \cup G') - S$. If both v and w are in G or in G', then there is a v,w-path in $(G \cup G') - S$ by assumption. Otherwise, since $|V(G) \cap V(G')| \geq k$, there exists $u \in V(G) \cap V(G') \cap V((G \cup G') - S)$. Moreover, since $G - S$ and $G' - S$ are connected, there exist a v,u-path in $G - S$ and a u,w-path in $G' - S$. But then there exists a v,w-path in $(G \cup G') - S$. ■

Let $G = (V, E)$ and $s \geq 1$ be an instance of Graph VC$_{k\text{-con}}$ Dimension. We claim we can check in polynomial time whether a subset $V' \subseteq V$ with $|V'| \geq s$ is shattered. By Claim 28, it is enough to check all the $O(|V|^{k+1})$ subsets of V' of size at most $k+1$. Recall now that finding a minimum separating set of a graph G is polynomial in the order of G. Moreover, if $S \subseteq V(G)$ is a minimum separating set and $A \cup B$ is a partition of $V(G - S)$ into two non-empty sets such that any path between a vertex in A and a vertex in B contains a vertex in S then, for $k > |S|$, the vertices of every k-connected subgraph of G are entirely contained in either $A \cup S$ or $B \cup S$. These observations, as shown in [106, Theorem 1], allow to test whether G has a k-connected subgraph in polynomial time. Therefore, for any $B \subseteq V'$ of size at most $k+1$, we can check in polynomial time if there exists a k-connected subgraph contained in $G - (V' \setminus B)$ and containing B.

Now we prove the NP-hardness by a reduction from Set Multicover. Given an instance of this problem, we construct a graph $G = (V, E)$ as follows (see Figure 4.4). The set of vertices V is formed by four pairwise disjoint sets A, B, C and D. A is an independent set of $n \cdot (t + k + 1)$ vertices arranged in n columns of $t + k + 1$ vertices each (every element in the j-th column corresponds to a copy of a_j), $B = \{v_1, \ldots, v_m\}$ is a clique (v_i corresponds to the set S_i), C is a clique of size k and D is an independent set of $t+m+1$ vertices. Each vertex in C is connected to all vertices in B (therefore, $B \cup C$ is a clique of size $m + k$) and D. Finally, $v_i \in B$ is connected to each copy of $a_j \in A$ if and only if $a_j \in S_i$.

Since $B \cup C$ is a clique and $A \cup D$ is an independent set, G is a split graph. We claim that there is an index set $I \subseteq \{1, \ldots, m\}$ such that $\bigcup_{i \in I} S_i = S$, each a_i is covered by at least k distinct subsets and $|I| \leq t$ if and only if VC$_{k\text{-con}}(G) \geq |V| - (t + k)$.

Suppose first such an index set I exists. We claim that the set

$$V' = A \cup D \cup \{v_i \in B : i \notin I\}$$

is shattered. Indeed, the subgraph $G' = G[C \cup \{v_i \in B : i \in I\}]$ is a clique of size at least $k + 1$ and each vertex in V' has at least k neighbours in G'. Therefore, V' is shattered by the Expansion Lemma (Lemma 1.0.2). Finally, $|I| \leq t$ implies that $|V'| \geq |V| - t - k$.

Conversely, let V' be a shattered set of maximum cardinality. Then $|V'| \geq |V| - (t + k)$. Suppose there exists $c \in V' \cap C$. This implies that no vertex in D belongs to V' and so $|V'| \leq |V| - (t + m + 1) < |V| - (t + k)$, a contradiction. Therefore, no vertex of C is in V' and $D \subseteq V'$. Moreover, at least one vertex $v \in A$ for each column is in V' and so v has at

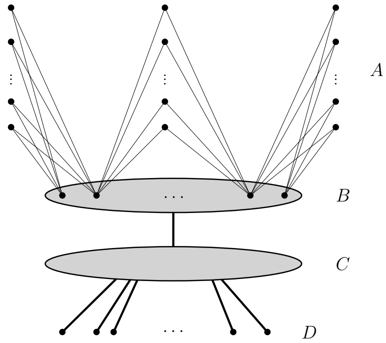


Figure 4.4: The graph G for the reduction. The grey ovals are cliques. A thick edge joining a vertex $v \in D$ to C means that v is adjacent to all the vertices of C. Similarly, the thick edge between the ovals means that $B \cup C$ is a clique.

least k neighbours in the clique $B \setminus V'$. By the Expansion Lemma (Lemma 1.0.2) and since all the vertices in the column of v have identical neighbourhoods, each vertex in the column of v belongs to V' and so $A \subseteq V'$. Therefore, the number of vertices in B which are in V' is at least $|V| - (t + k) - |A| - |D| = m - t$. We claim that $I = \{i : v_i \in B \setminus V'\}$ is a "yes"-instance of SET MULTICOVER. Indeed, any vertex of A has at least k neighbours in $B \setminus V'$. In other words, each $a_j \in S$ is contained in at least k of the subsets S_i with $i \in I$. Moreover, $|I| \leq m - (m - t) = t$. □

4.3.1 Graphs of bounded clique-width

Graphs of bounded clique-width are particularly interesting from an algorithmic point of view: many NP-complete problems can be solved in linear time for them. In fact, all graph properties which are expressible in monadic second-order logic are decidable in linear time on graphs of bounded clique-width (see Chapter 1). Let us recall that monadic second-order logic is an extension of first-order logic by quantification over sets. The language of *monadic second-order logic of graphs* (MSO_1 in short) contains the expressions built from the following elements:

- Variables $x, y, \ldots$ for vertices and $X, Y, \ldots$ for sets of vertices;
- Predicates $x \in X$ and $\mathit{adj}(x, y)$;
- Equality for variables, standard Boolean connectives and the quantifiers $\forall$ and $\exists$.

By considering edges and sets of edges as other sorts of variables and the incidence predicate $\mathit{inc}(v, e)$, we obtain *monadic second-order logic of graphs with edge-set quantification* (MSO_2 in short). If a graph property is expressible in the more restricted MSO_1, then Courcelle et al. [49] showed that it is decidable in linear time for graphs of bounded clique-width, assuming a k-expression of the input graph is explicitly given.

We now show that, for any graph, being shattered by its k-connected subgraphs is a property that can be expressed in monadic second-order logic of graphs:

Lemma 4.3.2. *Being shattered by k-connected subgraphs is expressible in* MSO_1.

Proof. Let $G = (V, E)$ be a graph. The following MSO_2 sentence says that the subgraph induced by $X \subseteq V$ is connected:

$$\textbf{conn}(X) = \forall_{Y \subseteq V}[(\exists_{u \in X}\, u \in Y \wedge \exists_{v \in X}\, v \notin Y) \to (\exists_{e \in E}\exists_{u \in X}\exists_{v \in X}\, \textit{inc}(u,e) \wedge \textit{inc}(v,e) \wedge u \in Y \wedge v \notin Y)].$$

It is easy to see that the quantification over single edges can be expressed by a MSO_1 sentence as follows:

$$\exists_{a \in V}\exists_{b \in V}\exists_{u \in X}\exists_{v \in X}\, \textit{adj}(a,b) \wedge (u = a \vee u = b) \wedge (v = a \vee v = b) \wedge u \in Y \wedge v \notin Y.$$

The following MSO_1 sentence says that the subgraph induced by $X \subseteq V$ is k-connected:

$$\textbf{k-conn}(X) = \exists_{v_1 \in X} \ldots \exists_{v_{k+1} \in X}(\forall_{u_1 \in V} \ldots \forall_{u_{k-1} \in V}\, \textbf{conn}(X \setminus \{u_1, \ldots, u_{k-1}\})).$$

Finally, the following MSO_1 sentence says that the set $A \subseteq V$ is shattered by k-connected subgraphs:

$$\textbf{shatt}(A) = \forall_{B \subseteq A}\exists_{X \subseteq V}\, \textbf{k-conn}(X) \wedge X \cap A = B.$$ □

Therefore, as an immediate consequence of the meta-theorem stated above, we have the following:

Corollary 4.3.3. GRAPH VC$_{k\text{-con}}$ DIMENSION *is decidable in linear time for graphs of bounded clique-width, assuming a k-expression of the input graph is explicitly given.*

We have seen that GRAPH VC$_{k\text{-con}}$ DIMENSION is NP-hard even for split graphs. Are there some (non-trivial) subclasses of split graphs for which the problem becomes easy? Recall that the *Dilworth number* of a graph G is the size of a largest antichain (or, equivalently, the size of a minimum chain partition) with respect to the quasi-order $\preceq$ defined on the vertices of G as follows: $x \preceq y$ if and only if $N_G(x) \subseteq N_G[y]$. Graphs with Dilworth number 1 are precisely the well-known *threshold* graphs, which are the P_4-free split graphs [42]. Therefore, they have clique-width at most 2 (see Chapter 1) and we have seen we can decide the VC-dimension in linear time. On the other hand, already a small jump for the Dilworth number of a split graph from 1 to 2 changes the clique-width from bounded to unbounded [111]. Nevertheless, deciding the VC-dimension remains easy:

Theorem 4.3.4. GRAPH VC$_{k\text{-con}}$ DIMENSION *is decidable in polynomial time for split graphs with Dilworth number at most* 2.

Proof. Let $G = (V, E)$ be the input graph. We assume the unique partition of V into a clique of size $\omega(G)$ and an independent set I of size $\alpha(G)$ is given. It is well-known that the problem of finding a minimum chain partition of a poset can be translated into a maximum bipartite matching problem (see, e.g., [132]) and the same holds for a quasi-order. Therefore, we can find in polynomial time a chain partition $I_1 \cup I_2$ of I. For $j \in \{1, 2\}$, let $I_{j,\geq k} = \{u \in I_j : d(u) \geq k\}$ and $I_{j,\succeq u} = \{v \in I_j : v \succeq u\}$. Note that, if $\omega(G) \leq k$, then G contains no k-connected subgraph. Therefore, we may assume $\omega(G) > k$. But then a maximum-size shattered set containing vertices from at most one of I_1 and I_2 has size

$$\max\{\omega(G) - (k+1) + |I_{1,\geq k}|,\ \omega(G) - (k+1) + |I_{2,\geq k}|\}.$$

On the other hand, it is not difficult to see that, for any pair $x \in I_{1,\geq k}$ and $y \in I_{2,\geq k}$, a maximum-size shattered set containing x as the minimal element from $I_{1,\geq k}$ and y as the minimal element from $I_{2,\geq k}$ has size

$$\omega(G) - \max\{k+1,\ 2k - |N(x) \cap N(y)|\} + |I_{1,\succeq x}| + |I_{2,\succeq y}|. \qquad \square$$

It is not difficult to see that outerplanar graphs have tree-width at most 2, and so bounded clique-width as well. In the next section, we consider GRAPH VC$_{k\text{-con}}$ DIMENSION for planar graphs.

4.3.2 Subclasses of planar graphs

Since a planar graph has a vertex of degree at most 5, its connectivity is at most 5. Therefore, GRAPH VC$_{k\text{-con}}$ DIMENSION restricted to planar graphs is non-trivial for at most five values of k and in this section we study its computational complexity in the cases $k = 1$ and $k = 2$.

Let us begin with $k = 1$. Note that planar graphs, in general, do not have bounded clique-width: the class of planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3 has unbounded clique-width [133]. In fact, we show that GRAPH VC$_{\text{con}}$ DIMENSION is NP-complete for this class. Given the connection between VC$_{\text{con}}$ and the connected domination number (Theorem 4.1.2), it is no surprise we construct a reduction from CONNECTED DOMINATING SET. Douglas [53] showed that, given a planar subcubic graph G, it is NP-complete to decide whether $\gamma_c(G) \leq \frac{|V(G)|}{2} - 1$ holds. Note that, for a subcubic graph, the quantity $\frac{|V(G)|}{2} - 1$ is the minimum possible size of a connected dominating set[4]. In Section 5.4, we strengthen Douglas' result by showing that the following problem is NP-complete for any $\ell \geq 2$ (Corollary 5.4.2):

CONNECTED DOMINATING SET	
Instance:	A planar bipartite $(C_4, \ldots, C_{2\ell})$-free graph $G = (V, E)$ with maximum degree 3.
Question:	Does there exist $D \subseteq V$ such that $G[D]$ is connected, every vertex in $V \setminus D$ is adjacent to at least one vertex in D and $|D| \leq \frac{|V|}{2} - 1$?

As mentioned, we postpone the proof of this result to Section 5.4. Instead, we immediately use it to show the NP-hardness of GRAPH VC$_{\text{con}}$ DIMENSION for the class of subcubic planar bipartite graphs. The idea is as follows. It is easy to see that the complement of every connected dominating set can be shattered by connected subgraphs. Moreover, if G is a subcubic graph with sufficiently large order and VC-dimension, we show that the converse holds as well: more precisely, $\gamma_c(G) \leq \frac{|V(G)|}{2} - 1$ if and only if VC$_{\text{con}}(G) \geq \frac{|V(G)|}{2} + 1$.

Theorem 4.3.5. *For any $\ell \geq 2$, GRAPH VC$_{\text{con}}$ DIMENSION is NP-complete even for planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3.*

Proof. We prove NP-hardness by a reduction from the variant of CONNECTED DOMINATING SET introduced above. Let $G = (V, E)$ be an instance of CONNECTED DOMINATING SET where G is a planar bipartite $(C_4, \ldots, C_{2\ell})$-free graph, $\Delta(G) = 3$ and $|V| = n$. Clearly, we may

[4] This follows easily by recalling $\ell(G) = |V(G)| - \gamma_c(G)$.

assume $n \geq 46$. We claim G has a connected dominating set D with $|D| \leq \frac{n}{2} - 1$ if and only if $\mathrm{VC}_{\mathrm{con}}(G) \geq \frac{n}{2} + 1$.

Suppose first G has a connected dominating set D with $|D| \leq \frac{n}{2} - 1$. Then each vertex in $V \setminus D$ can be joined to $G[D]$ independently of one another, and so $V \setminus D$ can be shattered by connected subgraphs. Therefore, $\mathrm{VC}_{\mathrm{con}}(G) \geq |V \setminus D| \geq \frac{n}{2} + 1$.

Conversely, suppose $\mathrm{VC}_{\mathrm{con}}(G) \geq \frac{n}{2} + 1$ and let A be a shattered set of maximum size. Recall that, by assumption, we have $|A| \geq 24$. The idea is to show there exists a component C of $G - A$ such that each vertex of A is joined to C. Claim 29 would then imply that $|C| \geq |A| - 2 \geq \frac{n}{2} - 1$ and so C would be a connected dominating set of size $\frac{n}{2} - 1$. We now elaborate on this idea in a series of claims, each followed by a short proof.

Claim 29. *Each component C of $G - A$ has at most $|C| + 2$ neighbours in A.*

Indeed, if C contains at most two 1-vertices, then the claim clearly holds. Therefore, let $\{u_1, \dots, u_k\}$, with $k \geq 3$, be the set of 1-vertices in C. Since C is connected, there exists a u_1, u_2-path P in C. Moreover, any u_3, u_1-path intersects P in a 3-vertex. Applying this reasoning again to each u_i with $i \geq 3$, we have that $d_1(C) - 2 \leq d_3(C)$, where $d_i(C) = |\{v \in C : d_C(v) = i\}|$. But then the number of neighbours of C in A is at most $2d_1(C) + d_2(C) = |C| + d_1(C) - d_3(C) \leq |C| + 2$. ■

Since $\Delta(G) = 3$ and $|A| \geq 24$, each vertex $u \in A$ has at least one neighbour in $G - A$, otherwise it would not be possible to shatter u and a vertex in $A \setminus N(u)$. Let now $C_1, \dots, C_k$ be the components of $G - A$.

Claim 30. *There exists a vertex in A joined to less than three components of $G - A$.*

Indeed, suppose each vertex in A is joined to exactly three components. By Claim 29 and double counting the size of the edge cut $[A, \overline{A}]$, we have $3|A| \leq \sum(|C_i| + 2)$. Therefore,

$$k \geq \frac{3|A| - \sum |C_i|}{2} = \frac{3|A| - (n - |A|)}{2} \geq \frac{n}{2} + 2,$$

contradicting the fact that $|V \setminus A| \leq \frac{n}{2} - 1$. ■

Claim 31. *No vertex in A is joined to exactly two components of $G - A$.*

Suppose, to the contrary, there exists $u \in A$ joined to exactly two components, say C_1 and C_2. Then, every vertex in $A \setminus N(u)$ is joined to either C_1 or C_2. By Claim 29, we have $|C_1| + |C_2| + 3 \geq |A| - 1$, whence $|C_1| + |C_2| \geq \frac{n}{2} - 3$ and so $\sum_{i=3}^{k} |C_i| \leq 2$.

We now claim that the (at least $|A| - 1$) vertices in $A \setminus N(u)$ are all joined to C_1 or all joined to C_2. Suppose, to the contrary, there exist v and w in $A \setminus N[u]$ such that v is joined to C_1, w is joined to C_2 but none of them is joined to both C_1 and C_2. Let $B \subseteq A$ be the subset of vertices which are joined to both C_1 and C_2. Since $|C_1| + 2 \geq |B|$ and $|C_2| + 2 \geq |B|$, we have $|B| \leq \frac{|C_1| + |C_2|}{2} + 2 \leq \frac{n+6}{4}$. Therefore, at least $\frac{n}{2} + 1 - \frac{n+6}{4} = \frac{n-2}{4}$ vertices of A are not joined to both C_1 and C_2. But then the set B' of vertices of $A \setminus (N[v] \cup N[w])$ which are not joined to both C_1 and C_2 has size at least $\frac{n-2}{4} - 6 \geq 5$. Since $\{x, v\}$ and $\{x, w\}$ are shattered for any $x \in B'$, each vertex in B' is joined to some component different from C_1 and C_2, contradicting the fact that the remaining components can be joined to at most four vertices in A.

Therefore, the (at least $|A|-1$) vertices in $A \setminus N(u)$ are all joined to the same component, say C_1. Suppose now there exists $u' \in N(u)$ not joined to C_1. By Claim 29, we have $|C_1| \geq \frac{n}{2} - 2$. Moreover, the set $\{u', w\}$ is shattered, for any $w \in A \setminus \{u, u'\}$, contradicting the fact that the remaining component has at most 3 neighbours in A. Therefore, each $v \in A$ is joined to C_1. Since $|C_1| \geq |A| - 2 \geq \frac{n}{2} - 1$, it follows that $C_2 = \varnothing$, a contradiction. ■

By Claim 30 and Claim 31, there exists $v \in A$ joined to exactly one component C of $G - A$. Clearly, v has at most two neighbours in A. Moreover, each of the (at least $|A| - 3$) non-neighbours of v in A is joined to C, otherwise it would not be possible to shatter the set $\{v, w\}$, for some $w \in A \setminus N[v]$. Suppose now there exists $v' \in N(v) \cap A$ not joined to C. By Claim 29, we have $|C| + 2 \geq |A| - 2$, hence $|C| \geq \frac{n}{2} - 3$ and so the remaining components contain at most two vertices. Since the set $\{v', w\}$ is shattered, for any $w \in A \setminus N[v']$, and since there are at least $|A| - 3 \geq 21$ such non-neighbours, we get a contradiction to the fact that a component in $G - A - C$ can have at most four neighbours in A. Therefore, each $v \in A$ is joined to C and $|C| \geq |A| - 2 \geq \frac{n}{2} - 1$. But then $|C| = \frac{n}{2} - 1$ and C is a connected dominating set for G. □

Clearly, Graph $\mathrm{VC}_{\mathrm{con}}$ Dimension becomes easy for graphs G with $\Delta(G) \leq 2$. Indeed, $\mathrm{VC}_{\mathrm{con}}(P_n) = 2$, for $n \geq 3$, and $\mathrm{VC}_{\mathrm{con}}(C_3) = 2$ and $\mathrm{VC}_{\mathrm{con}}(C_n) = 3$, for $n \geq 4$. Note that, if $\Delta(G) \leq 2$, then G has tree-width at most 2. Therefore, the conclusion follows from Corollary 4.3.3 as well.

We now consider Graph $\mathrm{VC}_{2\text{-con}}$ Dimension:

Theorem 4.3.6. Graph $\mathrm{VC}_{2\text{-con}}$ Dimension *is* NP-*complete even for planar bipartite graphs with maximum degree* 4.

Proof. Our reduction is from Hamiltonian Cycle, which remains NP-complete even for planar cubic bipartite graphs [4]. Given an instance $G = (V, E)$ of this problem, where G is a planar cubic bipartite graph with $|V| = n$, we construct a graph G' by replacing each vertex u of G with the gadget depicted in Figure 4.5.

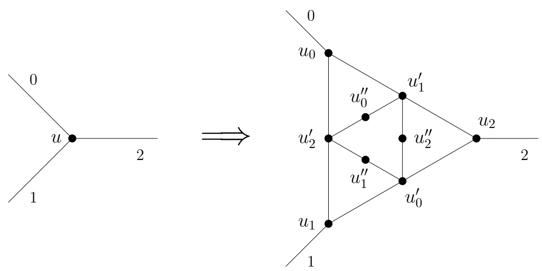


Figure 4.5: Construction of the graph G': The vertex u is replaced by a gadget containing 9 vertices.

For $0 \leq i \leq 2$, the vertices u_i are the *gates* and the vertices u'_i are the *connectors*. Finally, each pair $u'_i u'_{i+1}$ (indices modulo 3) of connectors in the gadget is connected by a path of length 2 with inner vertex u''_{i+2}, called a *crossing vertex*. Clearly, the construction can be done in polynomial time and the resulting graph $G' = (V', E')$ is planar, bipartite and $\Delta(G') = 4$. We claim that $\mathrm{VC}_{2\text{-con}}(G') \geq |V'| - 5n$ if and only if G contains a Hamiltonian cycle.

Suppose first G contains a Hamiltonian cycle C. Without loss of generality, $u \in V$ is incident to the edges 1 and 2 in C. Then we augment the subgraph induced by $E(C)$ in G'

with the path $u_1u_0'u_2$. Repeating this procedure for every vertex of G, we get a cycle in G' containing three vertices from every gadget. The vertices in $\{u_0, u_0'', u_1'', u_2''\}$ can be joined to this cycle, independently of one another, via paths through the connectors u_0', u_1' and u_2'. In all the cases, the resulting subgraph is clearly 2-connected. Repeating this process for every gadget, we have that a set of size $|V'| - 5n$ can be shattered.

Suppose now $\mathrm{VC}_{2\text{-con}}(G') \geq |V'| - 5n$ and let A be a shattered set of maximum cardinality.

Claim 32. *For any gadget $H \subseteq G'$, exactly one gate and the three crossing vertices are in A.*

We show first that at most four vertices of H are in A. Suppose, to the contrary, H contains at least five vertices of A. Then at least one crossing vertex is in A, otherwise at least two gates would be in A, contradicting the fact that the set consisting of a connector in H and a vertex not in H is shattered. Therefore, at least one crossing vertex is in A, say without loss of generality u_1''. This implies that the connectors u_0' and u_2' are not in A. If another crossing vertex is in A, then $u_1' \notin A$ and at least two gates are in A, a contradiction. Therefore, u_1'' is the only crossing vertex in A. But then all the gates are in A, a contradiction again.

Since $\mathrm{VC}_{2\text{-con}}(G') \geq |V'| - 5n$, exactly four vertices per gadget are in A and we have seen that at most one gate per gadget is in A. Moreover, exactly one gate per gadget belongs to A, otherwise a crossing vertex and one of its neighbouring connectors would both be in A. Let u_0 be the gate of the gadget H belonging to A. If one of its neighbouring connectors is in A (clearly, there exists at most one such connector), then a crossing vertex and at least one of its neighbouring connectors are both in A, a contradiction. Therefore, both u_1' and u_2' are not in A and it is easy to see that it must be $A \cap V(H) = \{u_0, u_0'', u_1'', u_2''\}$. ■

By Claim 32, there exists a 2-connected subgraph of G' containing crossing vertices in every gadget and avoiding exactly one gate per gadget. Therefore, for any gadget, this subgraph contains exactly two of the edges incident to its gates and originally in G. This means that, contracting each gadget to a single vertex, we obtain a 2-regular connected spanning subgraph. Therefore, G contains a Hamiltonian cycle. □

We conclude this section with some open problems. The first is to study the complexity of Graph $\mathrm{VC}_{2\text{-con}}$ Dimension for planar bipartite graphs with maximum degree 3. We believe this problem to be NP-hard. More generally, it would be interesting to determine the complexity of Graph $\mathrm{VC}_{k\text{-con}}$ Dimension for planar graphs in the remaining cases $3 \leq k \leq 5$.

4.4 Complexity dichotomies in monogenic classes

A class of graphs $\mathcal{G}$ is *monogenic* if it is defined by a single forbidden induced subgraph, i.e. $\mathcal{G} = \mathit{Free}(H)$, for some graph H. We say that a (decision) graph problem admits a dichotomy in monogenic classes if, for each monogenic class, the problem is either NP-complete or decidable in polynomial time. The first result in this direction was obtained by Korobitsin [108], who showed that Dominating Set is decidable in polynomial time if H is an induced subgraph of $P_4 + tK_1$, for $t \geq 0$, and NP-complete otherwise. Král' et al. [114] showed that Colouring is decidable in polynomial time if H is an induced subgraph of P_4 or of $P_3 + K_1$ and NP-complete otherwise. Kamiński [102] showed that Simple Max-Cut is decidable in polynomial time if H is an induced subgraph of P_4 and NP-complete otherwise. Other dichotomies were obtained by Golovach et al. [79] for Precolouring Extension and ℓ-List

COLOURING and by AbouEisha et al. [1] for UPPER DOMINATING SET. Other problems, like k-COLOURING and INDEPENDENT SET, appear more stubborn and no complete dichotomy is available: we refer the reader to [78] for a survey on the status of k-COLOURING and we just mention that a major question related to INDEPENDENT SET is whether it is decidable in polynomial time for P_k-free graphs with $k \geq 6$.

In the following, we enlarge the list above by providing dichotomies for the closely related GRAPH VC$_{\text{con}}$ DIMENSION and CONNECTED DOMINATING SET. Note that, in Chapter 5, the latter problem will be further studied in the context of boundary classes.

Let us begin with GRAPH VC$_{\text{con}}$ DIMENSION. Definition 4.1.1 has an analogue formulation for edge sets:

Definition 4.4.1 (Kranakis et al. [115]). Let $G = (V, E)$ be a graph and let $\mathcal{P}$ be a family of sets of edges of G. A subset $A \subseteq E$ is *$\mathcal{P}$-edge-shattered* if every subset of A can be obtained as the intersection of a $C \in \mathcal{P}$ with A. The *edge VC-dimension* of G with respect to $\mathcal{P}$ is defined as the maximum size of a $\mathcal{P}$-edge-shattered subset and it is denoted by EVC$_{\mathcal{P}}(G)$.

We denote by EVC$_{\text{con}}$ the edge VC-dimension with respect to the family of connected edge sets. Since a graph G is connected if and only if $L(G)$ is, it is easy to see that EVC$_{\text{con}}(G) =$ VC$_{\text{con}}(L(G))$ [115]. Moreover, given a graph G and a number $s \geq 1$, it is NP-complete to decide whether EVC$_{\text{con}}(G) \geq s$ holds [115]. It immediately follows that GRAPH VC$_{\text{con}}$ DIMENSION is NP-complete even for line graphs.

We now have all the machinery to prove our first dichotomy:

Theorem 4.4.2. GRAPH VC$_{\text{con}}$ DIMENSION *restricted to H-free graphs is decidable in polynomial time if H is an induced subgraph of $P_4 + tK_1$ and* NP*-complete otherwise.*

Proof. Suppose first H contains an induced cycle C_k. If k is odd, then the problem is NP-complete since it is NP-complete when restricted to bipartite graphs (Theorem 4.3.5). If k is even, then the problem is again NP-complete since it is NP-complete when restricted to split graphs (Theorem 4.3.1).

Suppose now H is a forest with a vertex of degree at least 3. Then H contains an induced claw and the problem is NP-complete since it is NP-complete when restricted to line graphs.

Finally, suppose H is the disjoint union of paths. If H contains at least two paths on at least 2 vertices, then H contains $2K_2$ and the problem is NP-complete since it is NP-complete when restricted to split graphs. The same conclusion holds if H contains a path on at least 5 vertices. It remains to consider the case of H being of the form $P_k + tK_1$, for some $k \leq 4$ and $t \geq 0$. Therefore, let G be such a $P_k + tK_1$-free graph. If G is in addition P_k-free, then it has bounded clique-width (see Chapter 1) and the problem is decidable in linear time (Corollary 4.3.3). On the other hand, if G contains an induced copy G' of P_k, then there are at most $t - 1$ pairwise non-adjacent vertices of G none of which is adjacent to a vertex of G'. Denoting this set by S, we have that $V(G') \cup S$ is a dominating set of size at most $t + 3$. Moreover, denoting by $\gamma(G)$ the domination number of G, Duchet and Meyniel [56] showed that $\gamma_c(G) \leq 3\gamma(G) - 2$ and so we have $\gamma_c(G) \leq 3t + 7$. If G has n vertices, Theorem 4.1.2 implies that $n - \gamma_c(G) \leq$ VC$_{\text{con}}(G) \leq n - \gamma_c(G) + 1$. But then we simply check all the subsets of $V(G)$ of size $n - \gamma_c(G)$ and $n - \gamma_c(G) + 1$: their number is $\binom{n}{\gamma_c(G)} + \binom{n}{\gamma_c(G)-1} = O(n^{3t+7})$ and so, by the proof of Theorem 4.3.1, we can compute VC$_{\text{con}}(G)$ in polynomial time. $\square$

Let us now consider CONNECTED DOMINATING SET. In order to show that this problem is NP-hard even for line graphs, it is useful to introduce the following notion. A *connected edge dominating set* of a graph $G = (V, E)$ is a subset $D \subseteq E$ such that $G[D]$ is connected and every edge in $E \setminus D$ is incident to at least one edge in D. We denote by $\gamma'_c(G)$ the minimum size of a connected edge dominating set of G and consider the following natural problem:

CONNECTED EDGE DOMINATING SET	
Instance:	A graph G and a positive integer s.
Question:	Does $\gamma'_c(G) \leq s$ hold?

Lemma 4.4.3 (Folklore). CONNECTED EDGE DOMINATING SET *is* NP-*complete.*

Proof. Recall that a connected vertex cover of a graph G is a vertex cover S such that $G[S]$ is connected and we denote by $\beta_c(G)$ the minimum size of a connected vertex cover of G. We first claim that $\gamma'_c(G) = \beta_c(G) - 1$, for any graph G. Indeed, consider a minimum connected edge dominating set D of G. It is easy to see that $G[D]$ is a tree. But then the vertex set of $G[D]$ is a connected vertex cover of G of size $\gamma'_c(G) + 1$.

Conversely, let D be a minimum connected vertex cover of G and let T be a spanning tree of $G[D]$. We have that $E(T)$ is a connected edge dominating set of G of size $\beta_c(G) - 1$.

This observation implies that CONNECTED EDGE DOMINATING SET and CONNECTED VERTEX COVER are polynomially equivalent and since the latter is NP-complete [76] (see also Section 5.5), the conclusion follows. □

Since there is an obvious bijection between the connected edge dominating sets of a graph G and the connected dominating sets of $L(G)$, we have that CONNECTED EDGE DOMINATING SET polynomially reduces to CONNECTED DOMINATING SET for line graphs. Therefore, Lemma 4.4.3 implies the following:

Lemma 4.4.4. CONNECTED DOMINATING SET *is* NP-*complete even for line graphs.*

We have seen in Chapter 1 that if a graph property is expressible in MSO_1, then it is decidable in linear time for graphs of bounded clique-width. We now show that being a connected dominating set is an example of such a property:

Lemma 4.4.5. *Being a connected dominating set is expressible in* MSO_1.

Proof. Let $G = (V, E)$ be a graph. As we have seen in Lemma 4.3.2, the following MSO_2 sentence says that the subgraph induced by $X \subseteq V$ is connected:

$$\mathbf{conn}(X) = \forall_{Y \subseteq V}[(\exists_{u \in X}\, u \in Y \wedge \exists_{v \in X}\, v \notin Y) \rightarrow (\exists_{e \in E}\exists_{u \in X}\exists_{v \in X}\, \mathit{inc}(u,e) \wedge \mathit{inc}(v,e) \wedge u \in Y \wedge v \notin Y)].$$

Moreover, the quantification over single edges can be expressed by a MSO_1 sentence as follows:

$$\exists_{a \in V}\exists_{b \in V}\exists_{u \in X}\exists_{v \in X}\, \mathit{adj}(a,b) \wedge (u = a \vee u = b) \wedge (v = a \vee v = b) \wedge u \in Y \wedge v \notin Y.$$

Finally, the following MSO_1 sentence says that $D \subseteq V$ is a connected dominating set:

$$\mathbf{cds}(D) = \mathbf{conn}(D) \wedge \forall_{v \in V \setminus D}\exists_{u \in D}\, \mathit{adj}(u,v).$$

□

Corollary 4.4.6. CONNECTED DOMINATING SET *is decidable in linear time for graphs of bounded clique-width, assuming a k-expression of the input graph is explicitly given. In particular, it is decidable in linear time for P_4-free graphs.*

We can finally prove our second dichotomy. The proof is similar to that of Theorem 4.4.2.

Theorem 4.4.7. CONNECTED DOMINATING SET *restricted to H-free graphs is decidable in polynomial time if H is an induced subgraph of $P_4 + tK_1$ and* NP*-complete otherwise.*

Proof. Suppose first H contains an induced cycle C_k. If k is odd, then the problem is NP-complete since it is NP-complete when restricted to bipartite graphs [157]. If k is even, then the problem is again NP-complete since it is NP-complete when restricted to split graphs [123].

Suppose now H is a forest with a vertex of degree at least 3. Then H contains an induced claw and the problem is NP-complete since, by Lemma 4.4.4, it is NP-complete when restricted to line graphs.

Finally, suppose H is the disjoint union of paths. If H contains at least two paths on at least 2 vertices, then H contains $2K_2$ and the problem is NP-complete since it is NP-complete when restricted to split graphs. The same conclusion holds if H contains a path on at least 5 vertices. Therefore, it remains to consider the case of H being of the form $P_k + tK_1$, for some $k \leq 4$ and $t \geq 0$. Let G be such a $P_k + tK_1$-free graph. If G is in addition P_k-free, then the problem is decidable in linear time by Corollary 4.4.6. On the other hand, if G contains an induced copy of P_k, we have seen in the proof of Theorem 4.4.2 that $\gamma_c(G) \leq 3t + 7$. Therefore, it suffices to check all the subsets of $V(G)$ of size at most $3t + 7$, and this can be clearly done in polynomial time. □

CHAPTER 5

Boundary Classes For NP-Hard Graph Problems

Contents

In this chapter, we consider the following meta-questions: when does a certain "hard" graph problem become "easy"?; Is there any "boundary" separating "easy" and "hard" instances? In order to answer these questions in the case of hereditary classes, Alekseev [5] introduced the notion of a boundary class for an NP-hard problem. The fundamental feature of this notion is that a problem Π is NP-hard for a finitely defined hereditary class X if and only if X contains a boundary class for Π [5, 8].

In the context of determining the boundary classes of an NP-hard problem, Alekseev [5] studied VERTEX COVER, Alekseev et al. [7] and Malyshev [135] DOMINATING SET and Korpelainen et al. [110] HAMILTONIAN CYCLE. In this chapter, we continue the search of boundary classes for several other problems involving non-local properties. In Section 5.2, we provide the first boundary class for the closely related HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE and HAMILTONIAN PATH. In Section 5.3, we reveal the first boundary class for FEEDBACK VERTEX SET. Finally, in Sections 5.4 and 5.5 we make some progress towards the determination of some boundary classes for CONNECTED DOMINATING SET and CONNECTED VERTEX COVER.

5.1 Introduction

In the previous chapters we have encountered several graph problems: most of them are "hard" even for restricted classes of graphs, while they become "easy" for some subclasses. For example, Theorem 4.4.7 tells us that CONNECTED DOMINATING SET is NP-complete for P_5-free graphs, while it is decidable in polynomial time for the subclass of P_4-free graphs. It is therefore natural to ask when a certain "hard" graph problem becomes "easy": Is there any "boundary" separating "easy" and "hard" instances? In this chapter, we consider hereditary graph classes[1] and, in order to make the previous concepts more precise, we introduce the following vocabulary. Given a graph problem Π, we say that a hereditary class of graphs X is Π-*hard* if Π is NP-hard for X, and Π-*easy* if Π is solvable in polynomial time for graphs in X. Note that throughout this chapter we assume P $\neq$ NP, or else the notion of "boundary" becomes vacuous.

In a first attempt to answer the meta-question posed above, one might be tempted to consider maximal Π-easy classes and minimal Π-hard classes. In fact, the first approach immediately turns out to be meaningless: there are no maximal Π-easy classes. Indeed, every Π-easy class X can be extended to another Π-easy class simply by adding to X a graph $G \notin X$ together with all its induced subgraphs. Even the approach through minimal Π-hard classes is not completely satisfactory: it works well for minor-closed graph classes but it generally fails for hereditary graph classes. Recall that a class of graphs is minor-closed if it is closed under vertex deletion, edge deletion and edge contraction (in particular, a minor-closed graph class is hereditary). Robertson and Seymour [163] showed that the tree-width of graphs in a minor-closed class X is bounded if and only if X excludes at least one planar graph. In other words, in the family of minor-closed graph classes there is a unique minimal class of unbounded tree-width. As we have seen in Chapter 1, many graph problems are solvable in polynomial time for graphs of bounded tree-width and so the class of planar graphs can be effectively considered as a boundary in the family of minor-closed graph classes.

On the other hand, for general hereditary classes, minimal Π-hard classes might not exist at all. We have seen a first example of this behaviour in Chapter 4, where we showed that, for any $\ell \geq 2$, GRAPH VC$_{\text{con}}$ DIMENSION is NP-hard for planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3. This clearly gives an infinite descending chain of Π-hard classes. Many other examples of this kind are known for other problems like INDEPENDENT SET, DOMINATING SET, HAMILTONIAN CYCLE [5, 7, 110]. In other situations, minimal Π-hard classes indeed exist, as the following example due to Malyshev and Pardalos [138] shows. Consider the TRAVELING SALESMAN PROBLEM: given a graph G, a weight function $w\colon E(G) \to \mathbb{R}$ and a number s, does there exist a Hamiltonian cycle C of G such that $\sum_{e \in E(C)} w(e) \leq s$? A simple reduction from HAMILTONIAN CYCLE shows that the problem is NP-hard for the class of complete graphs. On the other hand, every proper hereditary subclass of the class of complete graphs is finite and so the problem can be clearly solved in polynomial time for such a subclass. This means that the class of complete graphs is a minimal hard class for the TRAVELING SALESMAN PROBLEM.

The previous discussion suggests that the "limit" of an infinite "decreasing" sequence of Π-hard classes should play a role in the search of a "boundary" between easy and hard classes. Alekseev [5] formalized this intuition by introducing the notions of limit class and boundary class for INDEPENDENT SET. In fact, these concepts are completely general and the following definition is due to Alekseev et al. [8] (we remark that all the graph classes considered in this chapter are hereditary):

Definition 5.1.1 (Alekseev et al. [8]). Let Π be an NP-hard graph problem and X a Π-hard class of graphs. A class of graphs Y is a *limit class for* Π *with respect to* X ((Π, X)-limit, in short) if there exists a sequence $Y_1 \supseteq Y_2 \supseteq \ldots$ of Π-hard subclasses of X such that $\bigcap_{n \geq 1} Y_n = Y$. The class Y is a *limit class for* Π (Π-limit) if there exists a Π-hard class X such

[1] Recall that a class of graphs X is hereditary if it is closed under deletions of vertices. This is equivalent to the fact that X can be defined by a set of forbidden induced subgraphs, i.e. $X = \mathit{Free}(Z)$ for some set of graphs Z. The minimal set Z with this property is unique and it is denoted by $\mathit{Forb}(X)$.

that Y is (Π, X)-limit.

An inclusion-wise minimal (Π, X)-limit class is a *boundary class for* Π *with respect to* X ((Π, X)-boundary, in short). The class Y is a *boundary class for* Π (Π-boundary) if there exists a Π-hard class X such that Y is (Π, X)-boundary.

Note that in the definition of a limit class for Π with respect to X, the Π-hard subclasses of X in a sequence $Y_1 \supseteq Y_2 \supseteq \ldots$ need not be distinct. In particular, every Π-hard subclass of X is (Π, X)-limit. On the other hand, a Π-limit class need not be Π-hard. Indeed, consider again GRAPH VC$_{\text{con}}$ DIMENSION. Denoting by Y_ℓ the class of bipartite $(C_4, \ldots, C_{2\ell})$-free graphs, we have that GRAPH VC$_{\text{con}}$ DIMENSION is NP-hard for Y_ℓ, for any $\ell \geq 2$ (Theorem 4.3.5). Moreover, $Y_\ell \supseteq Y_{\ell+1}$ and $\bigcap_{n \geq 1} Y_n$ is clearly the class of forests, for which the problem is known to be solvable in linear time (Corollary 4.3.3).

Another important remark is the following:

Remark 5.1.2. A Π-limit subclass of a Π-hard class X is not necessarily (Π, X)-limit. Indeed, let Π be HAMILTONIAN CYCLE. This problem is NP-hard for graphs with arbitrarily large girth [13, 110] and so the class of forests is a limit class for Π. Moreover, Müller [141] showed that the class $X = \mathit{Free}(C_3, C_5, C_6, \ldots)$[2] is Π-hard. The class of forests is clearly contained in X but it is easy to see that it is not (Π, X)-limit. Indeed, suppose there exists a sequence $Y_1 \supseteq Y_2 \supseteq \ldots$ of Π-hard subclasses of X such that $\bigcap_{n \geq 1} Y_n$ coincides with the class of forests. Clearly, there exists a class Y_i not containing C_4. But then Y_i is the class of forests, which is Π-easy, a contradiction.

The existence of a boundary class with respect to every Π-hard class is guaranteed by the following theorem:

Theorem 5.1.3 (Alekseev et al. [8]). *A class X is Π-hard if and only if it contains a (Π, X)-boundary class.*

Note that the "if" direction in Theorem 5.1.3 is trivial: if $Y \subseteq X$ is a (Π, X)-boundary class, there exists a sequence $Y_1 \supseteq Y_2 \supseteq \ldots$ of Π-hard subclasses of X and so Π is NP-hard for X as well.

Theorem 5.1.3 shows that a boundary class with respect to a Π-hard class represents indeed a meaningful notion of "boundary" between Π-hard and Π-easy subclasses. Moreover, Π-boundary classes can be used to characterize the finitely defined graph classes[3] which are Π-hard:

Theorem 5.1.4 (Alekseev et al. [8]). *A finitely defined class is Π-hard if and only if it contains a Π-boundary class.*

Alekseev [5] studied INDEPENDENT SET and revealed the first boundary class for this problem: the class $\mathcal{T}$ of forests whose components have at most three leaves. This shows that Theorem 5.1.4 is not true in general: the class of forests contains $\mathcal{T}$ but it is easy for the problem (see Chapter 1). So far, this is the only known boundary class for INDEPENDENT SET and in fact he conjectured no other boundary class exists. It is easy to see that this conjecture is equivalent to the following statement: for each $G \in \mathcal{T}$, INDEPENDENT SET is not NP-hard

[2] This is the class of *chordal bipartite* graphs.

[3] Recall that a hereditary class is finitely defined if the set of its minimal forbidden induced subgraphs is finite.

for *Free*(G). The conjecture seems to be very challenging since it is already a major open problem to determine whether INDEPENDENT SET is NP-hard for P_k-free graphs with $k > 6$ (see [130]).

Other problems have been studied in the context of boundary classes. For example, Alekseev et al. [7] revealed three boundary classes for DOMINATING SET, one of them being $\mathcal{T}$, and Malyshev [135] found a fourth boundary class. Alekseev et al. [8] further emphasized the speacial role played by $\mathcal{T}$ and showed that, among others, it is boundary for INDEPENDENT DOMINATING SET, INDUCED MATCHING and EDGE DOMINATING SET (see [8] for definitions). On the other hand, $\mathcal{T}$ is not boundary for HAMILTONIAN CYCLE and Korpelainen et al. [110] revealed two boundary classes for this problem. So far, the complete description of boundary classes has been obtained only for a single problem, the so-called LIST EDGE-RANKING: Malyshev [136] showed it admits exactly ten boundary classes. Note that some problems may admit infinitely many boundary classes and it is known that there is a continuum set of boundary classes for VERTEX k-COLOURING [110, 137].

In this chapter, we continue the study of boundary classes for NP-hard problems. In Section 5.2, we provide the first boundary class for the closely related HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE and HAMILTONIAN PATH. This class was in fact shown to be boundary for HAMILTONIAN CYCLE [110]. In Section 5.3, we reveal the first boundary class for FEEDBACK VERTEX SET: the class of forests whose components have at most four leaves and no two vertices of degree three. Finally, in Sections 5.4 and 5.5 we make some progress towards the determination of some boundary classes for two other problems involving non-local properties: CONNECTED DOMINATING SET and CONNECTED VERTEX COVER.

We conclude this section with the proofs of Theorems 5.1.3 and 5.1.4 and some additional observations. We also invite the reader to notice that the notions of limit class and boundary class make sense for every partially ordered set (see [109]).

Lemma 5.1.5 (Alekseev et al. [8]). *If Y is a (Π, X)-limit class and $Y \subseteq Z \subseteq X$, then Z is a (Π, X)-limit class as well.*

Proof. Let $Y_1 \supseteq Y_2 \supseteq \dots$ be a sequence of Π-hard subclasses of X such that $\bigcap_{n\geq 1} Y_n = Y$. Clearly, the class $Z_n = Y_n \cup Z$ is Π-hard for every n. Moreover, we have $X \supseteq Z_1 \supseteq Z_2 \supseteq \dots$ and $\bigcap_{n\geq 1} Z_n = Z$. □

We have seen that, in general, a (Π, X)-limit class is not Π-hard. On the other hand, we now show that this is the case if the limit class is defined by finitely many forbidden induced subgraphs with respect to X[4]:

Lemma 5.1.6 (Alekseev et al. [8]). *If Y is a (Π, X)-limit class which is defined by finitely many forbidden induced subgraphs with respect to X, then it is Π-hard.*

Proof. Suppose that *Forb*$(Y) \setminus$ *Forb*$(X) = \{G_1, \dots, G_k\}$. Since Y is (Π, X)-limit, there exists a sequence $Y_1 \supseteq Y_2 \supseteq \dots$ of Π-hard subclasses of X such that $\bigcap_{n\geq 1} Y_n = Y$. Moreover, there exists an index n such that Y_n is $(G_1, \dots, G_k)$-free. But then $Y_i = Y$, for each $i \geq n$, and so Y is Π-hard. □

[4] Recall that a subclass $Y \subseteq X$ is defined by finitely many forbidden induced subgraphs with respect to X if *Forb*$(Y) \setminus$ *Forb*(X) is a finite set.

Lemma 5.1.7 (Alekseev et al. [8]). *If $Y_1 \supseteq Y_2 \supseteq \dots$ is a sequence of (Π, X)-limit classes, then $Y = \bigcap_{n \geq 1} Y_n$ is a (Π, X)-limit class.*

Proof. Let $Forb(Y) = \{G_1, G_2, \dots\}$ and, for each k, let $Y^{(k)}$ be the class $Free(G_1, \dots, G_k)$. Clearly, for each k, there exists an index n such that Y_n is $(G_1, \dots, G_k)$-free and so $Y_n \subseteq Y^{(k)}$. By Lemma 5.1.5, we have that $Y^{(k)}$ is a (Π, X)-limit class and so, by Lemma 5.1.6, it is Π-hard. On the other hand, $Y^{(k)} \subseteq Y^{(k+1)}$ and $\bigcap_k Y^{(k)} = Y$. Therefore, Y is a (Π, X)-limit class. □

We can finally conclude the proof Theorem 5.1.3. It remains to show that if Π is NP-hard for the class X, then X contains a (Π, X)-boundary class.

Proof of Theorem 5.1.3. We may assume the graphs in X are indexed as $X = \{G_1, G_2, \dots\}$ and we recursively define a sequence $Y_1 \supseteq Y_2 \supseteq \dots$ of subclasses of X as follows. We set first $Y_1 = X$. Then, if there exists an index j such that $G_j \in Y_n$ and $Y_n \cap Free(G_j)$ is a (Π, X)-limit class, we take a minimum such j and define $Y_{n+1} = Y_n \cap Free(G_j)$. Otherwise, we set $Y_{n+1} = Y_n$.

We claim that the class $Y = \bigcap_{n \geq 1} Y_n$ is a (Π, X)-boundary class. By Lemma 5.1.7, we have that Y is (Π, X)-limit. Suppose now, to the contrary, that there exists a (Π, X)-limit class Z which is properly contained in Y. This means there exists a graph $G_i \in X$ which belongs to Y but not to Z and so $Z \subseteq Y \cap Free(G_i) \subseteq Y_n \cap Free(G_i)$, for each n. Therefore, by Lemma 5.1.5, we have that $Y_n \cap Free(G_i)$ is a (Π, X)-limit class, for each n. Moreover, there exists an n such that the index i is the minimum with the property that $Y_n \cap Free(G_i)$ is a (Π, X)-limit class. Therefore, $Y_{n+1} = Y_n \cap Free(G_i)$ and G_i does not belong to any class Y_k with $k > n$, contradicting the fact that $G_i \in Y$. □

We can actually prove the following stronger version of Theorem 5.1.3 which clearly implies Theorem 5.1.4:

Theorem 5.1.8 (Alekseev et al. [8]). *A subclass $Y \subseteq X$ defined by finitely many forbidden induced subgraphs with respect to X is Π-hard if and only if Y contains a (Π, X)-boundary class.*

Proof. If Y contains a (Π, X)-boundary class, then Y is (Π, X)-limit (Lemma 5.1.5) and so it is Π-hard (Lemma 5.1.6). Conversely, if Y is Π-hard, then it contains a (Π, Y)-boundary class Z (Theorem 5.1.3), which is (Π, X)-limit. But then Z contains a (Π, X)-boundary class. □

Proving the minimality of a certain limit class is in general not an easy task and currently there is no unified framework to address such a problem. Nevertheless, the following sufficient condition turns out to be useful and it will be employed in all our proofs:

Lemma 5.1.9 (Alekseev and Malyshev [6]). *A (Π, X)-limit class $Y = Free(M)$ is (Π, X)-boundary if for every $G \in Y$ there exists a finite set of graphs $A \subseteq M$ such that $Free(A \cup \{G\})$ is a Π-easy class.*

Proof. Suppose Y is not (Π, X)-boundary. This means there exists a (Π, X)-limit class $Z \subsetneq Y$ and let G be a graph in $Y \setminus Z$. By assumption, there exists a finite set $A \subseteq M$ such that $Z' = Free(A \cup \{G\})$ is Π-easy. Moreover, since Z is (Π, X)-limit, there exists a sequence $Z_1 \supseteq Z_2 \supseteq \dots$ of Π-hard subclasses of X such that $\bigcap_{n \geq 1} Z_n = Z$. But then, for each n, we have that $Z'_n = Z_n \cup Z'$ is Π-hard. Moreover, for each k, we have that $Z'_k \supseteq Z'_{k+1}$ and $\bigcap_{n\geq 1} Z'_n = Z'$. In other words, Z' is a Π-limit class as well. Since the set of forbidden induced subgraphs for $Z' = Free(A \cup \{G\})$ is finite, there exists a class Z'_n which is $(A \cup \{G\})$-free. Therefore, Z'_n is Π-easy, a contradiction. $\square$

5.2 Hamiltonian Cycle Through Specified Edge and Hamiltonian Path

In this section, we provide the first boundary class for HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE and HAMILTONIAN PATH. We have seen that these problems are NP-hard for subcubic graphs (Theorems 2.2.14 and 2.2.15) and in fact we determine a boundary class with respect to the class of subcubic graphs. To this end, we have to introduce some notation. For positive integers i, j and k, let $Y_{i,j,k}$ be the graph depicted in Figure 5.1, called a *tribranch*. Moreover, let $\mathcal{Y}_p = \{Y_{i,j,k} : i, j, k \leq p\}$ and $\mathcal{C}_p = \{C_k : k \leq p\}$. Finally, denote by $\mathcal{Q}_p$ the class of subcubic $\mathcal{Y}_p \cup \mathcal{C}_p$-free graphs such that each cubic vertex has a non-cubic neighbour.

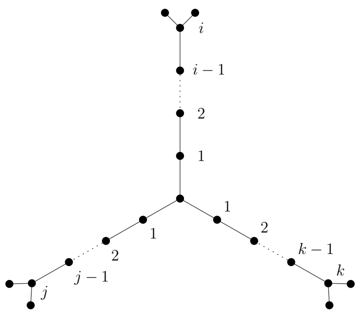


Figure 5.1: A tribranch $Y_{i,j,k}$.

Korpelainen et al. [110] showed that the class $\mathcal{Q} = \bigcap_{p\geq 1} \mathcal{Q}_p$ is a boundary class for HAMILTONIAN CYCLE:

Theorem 5.2.1 (Korpelainen et al. [110]). *For any $p \geq 1$,* HAMILTONIAN CYCLE *is* NP-*complete for graphs in $\mathcal{Q}_p$. Moreover, $\mathcal{Q}$ is a boundary class for* HAMILTONIAN CYCLE.

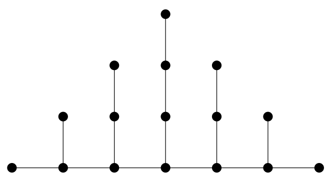


Figure 5.2: A caterpillar with hairs of arbitrary length.

The class $\mathcal{Q}$ is clearly contained in the class of forests and in fact it coincides with the class of graphs whose components are caterpillars with hairs of arbitrary length, where a *caterpillar with hairs of arbitrary length* is a subcubic tree in which all cubic vertices belong to a single path (see Figure 5.2):

Lemma 5.2.2 (Korpelainen et al. [110]). *A graph G belongs to $\mathcal{Q}$ if and only if each component of G is a caterpillar with hairs of arbitrary length.*

In the following, we adapt the reasoning of Korpelainen et al. [110] in order to show that $\mathcal{Q}$ is a boundary class also for HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE and HAMILTONIAN PATH. Let us begin with the former:

HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE	
Instance:	A graph $G = (V, E)$ and $e \in E$.
Question:	Does G contain a Hamiltonian cycle through e?

We first show that $\mathcal{Q}$ is a limit class:

Lemma 5.2.3. *For any $p \geq 1$,* HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE *is* NP*-complete for graphs in $\mathcal{Q}_p$.*

Proof. We reduce from HAMILTONIAN CYCLE for graphs in $\mathcal{Q}_{2p+3}$, which is NP-complete by Theorem 5.2.1. Given a graph $G = (V, E)$ in $\mathcal{Q}_{2p+3}$, we construct a graph G' as follows. Clearly, we may assume G contains a cubic vertex v. But then v has a non-cubic neighbour v' and we subdivide vv' with a new vertex u. Finally, we set $e = uv$.

We claim that the resulting graph G' belongs to $\mathcal{Q}_p$. The only not completely trivial verification to make is that G' is $\mathcal{Y}_p$-free. Note that G does not contain a tribranch in $\mathcal{Y}_p$, even as a subgraph. Suppose now G' contains an induced $Y_{i,j,k}$ with $i, j, k \leq p$. By the previous remark, u belongs to $Y_{i,j,k}$. But then, by contracting an edge of $Y_{i,j,k}$ incident to u, we obtain a tribranch in G which belongs to $\mathcal{Y}_p$, a contradiction.

Finally, G has a Hamiltonian cycle if and only if G' has a Hamiltonian cycle through e. □

Remark 5.2.4. The proof of Lemma 5.2.3 shows that HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE remains NP-hard for graphs in $\mathcal{Q}_p$ even when $e = uv$ is such that $d(u) = 3$ and $d(v) = 2$. This fact will be used in the proof that HAMILTONIAN PATH is NP-hard for graphs in $\mathcal{Q}_p$ (Lemma 5.2.6).

We now show that $\mathcal{Q}$ is a minimal limit class. The idea is to use Lemma 5.1.9. More precisely, we show that for every $G \in \mathcal{Q}$ there exists a constant p such that *Free*$(M \cup \{G\})$ is an easy class for HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE, where M is the set with $\mathcal{Q}_p = $ *Forb*(M). The applicability of Lemma 5.1.9 follows from the fact that $\mathcal{Q}_p$ is finitely defined. Indeed, the set of forbidden induced subgraphs for $\mathcal{Q}_p$ contains finitely many cycles and tribranches. Similarly, the conditions that every graph in $\mathcal{Q}_p$ is subcubic and that every cubic vertex has a non-cubic neighbour can be expressed by finitely many forbidden induced subgraphs.

Theorem 5.2.5. *$\mathcal{Q}$ is a boundary class for* HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE.

Proof. As remarked above, it is enough to show that for every $G \in \mathcal{Q}$ there exists a constant p such that HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE is solvable in polynomial time for G-free graphs in $\mathcal{Q}_p$.

We say that an edge of a graph G is *good* if it belongs to all Hamiltonian cycles of G (if any), whereas it is *bad* if it does not belong to any Hamiltonian cycle of G. Korpelainen et al.

[110] showed that, for each $G \in \mathcal{Q}$, there exists a constant p' such that the following holds: given a G-free graph $G' \in \mathcal{Q}_{p'}$, for any cubic vertex $v \in V(G')$, there is a polynomial-time algorithm that labels at least two edges incident to v as good, or it returns as output that the graph has no Hamiltonian cycle.

We claim it is enough to take the constant p'. In other words, we show that for every $G \in \mathcal{Q}$, HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE is solvable in polynomial time for G-free graphs in $\mathcal{Q}_{p'}$. Therefore, consider an input of this problem consisting of a G-free graph G' in $\mathcal{Q}_{p'}$ and an edge $e \in E(G')$. Clearly, we may assume G' has no vertices of degree 1, or else G' has no Hamiltonian cycle. But then every vertex of G' has degree 2 or 3 and we can clearly label all the edges incident to vertices of degree 2 as good. Now, for each cubic vertex, we simply run the algorithm mentioned above, thus obtaining a labelling of the edges of G'. If e is labelled bad, then G' has no Hamiltonian cycle through e. Therefore, suppose e is labelled good. If there exists $v \in V(G')$ incident to three good edges, then G' has no Hamiltonian cycle. Otherwise, each vertex of G' is incident to exactly two good edges and these edges induce a collection of disjoint cycles. If the collection contains exactly one cycle, then there exists a Hamiltonian cycle in G' through e. On the other hand, if the collection contains more than one cycle, then G' contains no Hamiltonian cycle at all. □

We now consider HAMILTONIAN PATH and show that $\mathcal{Q}$ is a boundary class for this problem as well. Denote by $\mathcal{R}_p$ the subclass of $\mathcal{Q}_p$ consisting of the graphs G with $\delta(G) \geq 2$. The following lemma implies that $\mathcal{Q}$ is a limit class for HAMILTONIAN PATH:

Lemma 5.2.6. *For any $p \geq 1$,* HAMILTONIAN PATH *is* NP*-complete for graphs in $\mathcal{R}_p$.*

Note that the reason for the apparently useless restriction to $\mathcal{R}_p$ will appear in Lemma 5.3.6.

Proof. We reduce from HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE for graphs in $\mathcal{Q}_{2p+3}$, which is NP-complete by Lemma 5.2.3. Let $G = (V, E)$ and $uv \in E$ be an instance of this problem, where $G \in \mathcal{Q}_{2p+3}$. Clearly, we may assume $\delta(G) \geq 2$ and, by Remark 5.2.4, we may further assume that $d(u) = 3$ and $d(v) = 2$. We construct a graph $G' = (V', E')$ as follows (see Figure 5.3). Set first $V' = V \cup \{a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4\}$, where each a_i and b_i is a new vertex, and $E' = (E \setminus \{uv\}) \cup \{a_1u, b_1v, a_1a_2, a_1a_4, a_2a_3, a_2a_4, a_3a_4, b_1b_2, b_1b_4, b_2b_3, b_2b_4, b_3b_4\}$. Finally, subdivide each edge in $\{a_1u, b_1v, a_1a_4, a_2a_3, a_3a_4, b_1b_4, b_2b_3, b_3b_4\}$ with p new vertices (thus belonging to V'). Note that the cubic vertices in $V' \setminus V$ are exactly the vertices in $\{a_1, a_2, a_4, b_1, b_2, b_4\}$.

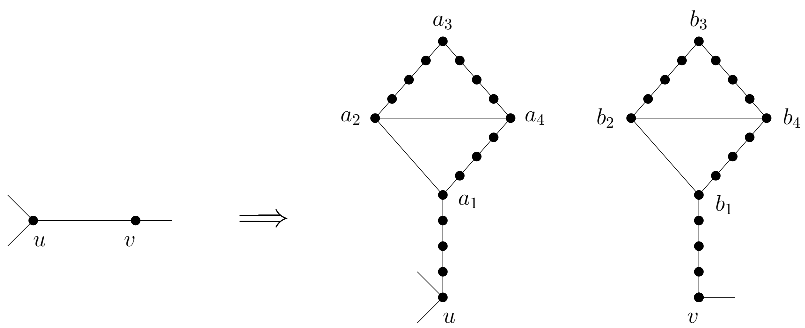


Figure 5.3: Construction of the graph G'.

Clearly, G' is a subcubic $\mathcal{C}_p$-free graph with $\delta(G') \geq 2$. Moreover, it is easy to see that each cubic vertex of G' has a non-cubic neighbour.

Suppose now G' contains an induced tribranch $Y_{i,j,k}$ with $i, j, k \leq p$. Since G does not contain a tribranch in $\mathcal{Y}_p$ (even as a subgraph), we have that $Y_{i,j,k}$ is not contained in $G - uv$ and so it must contain at least one vertex in $V' \setminus V$. On the other hand, it is easy to see that no vertex in $V' \setminus V$ can be a cubic vertex of $Y_{i,j,k}$ and so it must be that the unique vertex x in $N_{G'}(u) \setminus V$ is a 1-vertex of $Y_{i,j,k}$ and $V(Y_{i,j,k}) \cap (V' \setminus V) = \{x\}$. Moreover, if v does not belong to $Y_{i,j,k}$, then G contains a subgraph in $\mathcal{Y}_p$ (just substitute x by v), a contradiction. Therefore, v belongs to $Y_{i,j,k}$. But then the distance between u and v in G' is at most $2p+1$ and so, since $uv \in E$, we have that G contains a cycle of length at most $2p+2$, a contradiction. This implies that $G' \in \mathcal{R}_p$.

Finally, we claim that G has a Hamiltonian cycle through uv if and only if G' has a Hamiltonian path. Suppose first G' has a Hamiltonian path P. It is easy to see that a_1 is a 2-vertex of P and so a vertex in the gadget attached to u is a 1-vertex of P. Similarly, a vertex in the gadget attached to v is a 1-vertex of P. Therefore, there exists a Hamiltonian path in G between u and v and so a Hamiltonian cycle through uv. Conversely, it is easy to see that if G has a Hamiltonian cycle through uv, then G' has a Hamiltonian path between a_2 and b_2. □

We now show the minimality of $\mathcal{Q}$. Our proof is based again on Lemma 5.1.9 and it is inspired by the proof in [110] that $\mathcal{Q}$ is boundary for HAMILTONIAN CYCLE.

It is useful to consider the following special graphs in $\mathcal{Q}$: for $d \geq 2$, the graph T_d is the caterpillar consisting of a path of length $2d$ and $2d-1$ consecutive hairs of lengths $1, 2, \ldots, d-1, d, d-1, \ldots, 2, 1$ (the caterpillar in Figure 5.2 is in fact T_3). Obviously, every graph in $\mathcal{Q}$ is an induced subgraph of some T_d:

Observation 5.2.7 (Korpelainen et al. [110]). *Every graph in $\mathcal{Q}$ is an induced subgraph of T_d, for some $d \geq 2$.*

The idea is that for every $G \in \mathcal{Q}$, we can carefully choose a constant p such that a G-free graph in $\mathcal{Q}_p$ is "locally" a graph in $\mathcal{Q}$. This allows to implement a labelling procedure as in the proof of Theorem 5.2.5. To determine the local structure, we make use of the following elementary result whose proof can be found in [52]:

Lemma 5.2.8. *If G is a graph of radius at most r and maximum degree at most $k \geq 3$, then $|V(G)| < \frac{k}{k-2}(k-1)^r$.*

Theorem 5.2.9. *$\mathcal{Q}$ is a boundary class for* HAMILTONIAN PATH.

Proof. As remarked above, we show that for every $G \in \mathcal{Q}$ there exists a constant p such that HAMILTONIAN PATH is solvable in polynomial time for G-free graphs in $\mathcal{Q}_p$. By Lemma 5.1.9, this would conclude the proof.

Consider a graph $G \in \mathcal{Q}$. By Observation 5.2.7, we have that G is an induced subgraph of T_d, for some $d \geq 2$, and we define $p = 3 \cdot 2^d$. We claim we can decide in polynomial time whether a T_d-free graph in $\mathcal{Q}_p$ contains a Hamiltonian Path. This would clearly imply the assertion in the paragraph above, thus concluding the proof. Therefore, let G' be a T_d-free graph in $\mathcal{Q}_p$. For each pair of vertices u and v of G' and edges uu' and vv', we show that it is possible to check in polynomial time whether there exists a Hamiltonian u, v-path containing

uu' and vv'. The conclusion would then follow by repeating this procedure $O(|V(G')|^2)$ times.

Clearly, we may assume G' contains a cubic vertex, or else the problem is trivial. Moreover, it is enough to prove our claim for graphs with no 1-vertex. Indeed, suppose G' contains a 1-vertex w and let P be a shortest path linking w to a cubic vertex w'. Clearly, any Hamiltonian path of G' contains P as a subpath. Therefore, it is enough to check whether there exists a Hamiltonian path in $G' - (V(P) \setminus \{w'\})$ having w' as one end. By the proof of Theorem 5.2.5, we may also assume that u and v are non-adjacent, or else a Hamiltonian u, v-path is equivalent to a Hamiltonian cycle through uv and the constant p we take is the same as the one in [110] and in Theorem 5.2.5 (note that we make this assumption just in order to shorten the proof).

We say that an edge of G' is *good* if it belongs to all Hamiltonian u, v-paths of G' containing uu' and vv' (if any), whereas it is *bad* if it does not belong to any Hamiltonian u, v-path of G' through uu' and vv' (clearly, uu' and vv' are good). We provide a polynomial-time algorithm that either labels at least two edges incident to w as good, for each vertex $w \in V(G') \setminus \{u, v\}$, or returns as output that the graph has no Hamiltonian u, v-path through uu' and vv'. More precisely, we address the vertices sequentially and, if during the labelling process some edges are relabelled (i.e. a good edge becomes bad or vice versa), we have that G' does not contain any Hamiltonian u, v-path through uu' and vv'. Suppose now we have obtained such a labelling. If there exists $w \in V(G') \setminus \{u, v\}$ incident to three good edges, then G' has no Hamiltonian u, v-path through uu' and vv'. Otherwise, each vertex in $V(G') \setminus \{u, v\}$ is incident to exactly two good edges and these edges induce a path and possibly some cycles. If there are no cycles, we have found a desired Hamiltonian path. Otherwise, no such path exists.

Let us finally proceed with the description of the labelling algorithm. Clearly, uu' and vv' are good and the same holds for the edges incident to a 2-vertex $w \in V(G') \setminus \{u, v\}$. Moreover, the other edges incident to u and v are bad. Therefore, it remains to consider the cubic vertices in $V(G') \setminus \{u, v\}$. Let w be such a vertex and let H_w be the subgraph of G' induced by the set of vertices at distance at most d from w. Since H_w belongs to $\mathcal{Q}_p$, it is a subcubic $\mathcal{Y}_p \cup \mathcal{C}_p$-free graph such that each cubic vertex has a non-cubic neighbour. On the other hand, since H_w is subcubic, we have that $|V(H_w)| < 3 \cdot 2^d = p$ (Lemma 5.2.8) and so H_w is $\mathcal{Y}_p \cup \mathcal{C}_p$-free for any $k > p$. Therefore, H_w belongs to $\mathcal{Q}$ and, being connected, it must be a caterpillar with hairs of arbitrary length. Moreover, each leaf of H_w is at distance exactly d from w, or else it would be a 1-vertex of G'. Consider now a path P of H_w connecting two leaves and containing all cubic vertices of H_w. Since G' is T_d-free, we have that P contains 2-vertices, and let w_i be a 2-vertex of P having shortest distance from w. Moreover, let $w_0 w_1 \cdots w_i$ be the subpath between w_i and $w_0 = w$. Each vertex w_j with $j \neq i$ has a neighbour w'_j not on the path and which is a 2-vertex different from w_i. We denote by W the set $\{w_0, \ldots, w_{i-1}, w_i, w'_0, \ldots, w'_{i-1}\}$ and we distinguish several cases according to the size of the intersection $W \cap \{u, v\}$. In each case, we are going to argue that there exists a subpath $w_0 w_1 \cdots w_j$ whose edges are labelled alternately good and bad. This suffices to label two edges incident to $w = w_0$ as good. Indeed, if $w_0 w_1$ is bad, the other two edges incident to w_0 are good. On the other hand, suppose $w_0 w_1$ is good and let $w''_0 \in N(w_0) \setminus \{w_1, w'_0\}$. If w'_0 is either u or v (say without loss of generality $w'_0 = u$) and $u' \neq w_0$, we have that $w_0 w''_0$ is good. Otherwise, the remaining good edge must be $w_0 w'_0$.

Suppose first $W \cap \{u, v\} = \varnothing$. This means that the edges incident to a 2-vertex in W are both labelled good. Since $w_i w_{i-1}$ and $w_{i-1} w'_{i-1}$ are both good, $w_{i-1} w_{i-2}$ is bad. Moreover, since each cubic vertex in W has at least two good incident edges, $w_{i-2} w_{i-3}$ is good.

Therefore, the edges on the subpath $w_0w_1 \cdots w_i$ are labelled alternately good and bad (while traversing the path from w_i to w_0).

Suppose now W contains exactly one vertex from $\{u, v\}$, say without loss of generality $u \in W$. If u is a cubic vertex then, by assumption, we have that $u = w_j$ with $j > 0$ and $u' = w'_j$ (otherwise $v = w'_j$). Moreover, for each 2-vertex in W, its incident edges are labelled good. But then w_jw_{j-1} is bad and, similarly to the paragraph above, the edges on the subpath $w_0w_1 \cdots w_j$ are labelled alternately bad and good. Suppose now u is a 2-vertex in W. If $u = w_i$ and u' is the neighbour of u different from w_{i-1}, the edges on the subpath $w_0w_1 \cdots w_i$ are labelled alternately bad and good. If $u = w_i$ and $u' = w_{i-1}$, the edges on the subpath $w_0w_1 \cdots w_i$ are labelled alternately good and bad. If $u = w'_j$ and $u' = w_j$, for some $0 \leq j \leq i-1$, we have again that the edges on the subpath $w_0w_1 \cdots w_i$ are labelled alternately good and bad. Finally, if $u = w'_j$ and u' is the neighbour of u different from w_j, for some $0 \leq j \leq i-1$, the edges on the subpath $w_0w_1 \cdots w_j$ are labelled alternately good and bad.

Finally, suppose that $\{u, v\} \subseteq W$. Consider the smallest index $j < i$ such that either w_j or w'_j is a vertex in $\{u, v\}$. Note that, since $uv \notin E(G')$, it cannot be that $\{u, v\} \subseteq \{w_j, w'_j\}$ and we assume, without loss of generality, that $u \in \{w_j, w'_j\}$. If $u = w_j$, then $u' = w'_j$ and, by minimality, each 2-vertex w'_k with $k < j$ is incident to two good edges. Therefore, w_jw_{j-1} is bad and the edges on the subpath $w_0w_1 \cdots w_j$ are labelled alternately bad and good. If $u = w'_j$ and u' is the neighbour of w'_j different from w_j, then $w_j \neq v$ and the edges w_jw_{j-1} and w_jw_{j+1} are both good. Moreover, by minimality, each 2-vertex w'_k with $k < j$ is incident to two good edges and so the edges on the subpath $w_0w_1 \cdots w_j$ are labelled alternately good and bad. It remains to consider the case $u = w'_j$ and $u' = w_j$. If $v = w_k$ and $v' = w'_k$, for some $j < k < i$, it is easy to see that the edges on the subpath $w_0w_1 \cdots w_k$ are labelled alternately bad and good. If $v = w'_k$ and $v' = w_k$, for some $j < k < i$, the edges on the subpath $w_0w_1 \cdots w_i$ are labelled alternately good and bad. If $v = w'_k$ and v' is the neighbour of w'_k different from w_k, for $j < k < i$, the edges on the subpath $w_0w_1 \cdots w_k$ are labelled alternately good and bad. Finally, in the case $v = w_i$, the edges on the subpath $w_0w_1 \cdots w_i$ are labelled alternately good and bad, if $v' = w_{i-1}$, or bad and good otherwise.

This concludes the proof. □

We conclude this section with some observations on Hamiltonian cycles and paths in planar bipartite graphs. It is well-known that HAMILTONIAN CYCLE and HAMILTONIAN PATH are NP-hard even when restricted to planar graphs [75] and so they admit a boundary class with respect to planar graphs. In fact, Arkin et al. [13] showed that HAMILTONIAN CYCLE remains NP-hard even for subcubic planar graphs with arbitrarily large girth. We now strengthen this result by adding a bipartiteness constraint:

Theorem 5.2.10. *For any $\ell \geq 2$,* HAMILTONIAN CYCLE *is* NP-*complete even for planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3.*

It should be noticed that in the reduction used by Korpelainen et al. [110] to prove Theorem 5.2.1, planarity is in general not preserved. For the proof of Theorem 5.2.10, we need to introduce a useful operation defined on graphs containing cubic vertices. For an integer $k \geq 0$, the operation T^k is defined as follows: the graph $T^k(G)$ is obtained by replacing an arbitrary cubic vertex u of G with the gadget depicted in Figure 5.4. Note that, for $0 \leq i \leq 2$, there are k unmarked 2-vertices between u_i and u'_{i+1} (indices modulo 3).

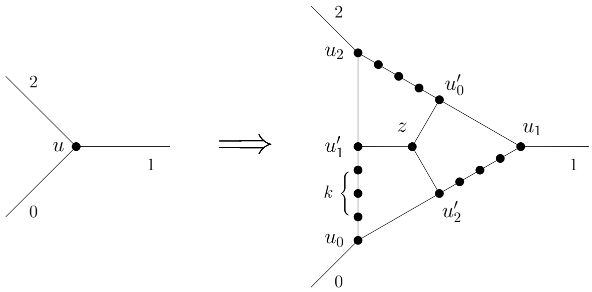


Figure 5.4: Application of T^k for $k = 3$: the 3-vertex u is replaced by a gadget containing $3k + 7$ vertices.

Lemma 5.2.11. *If G is a graph containing a 3-vertex, then G is Hamiltonian if and only if $T^k(G)$ is. Moreover, if G is subcubic, planar and bipartite then, for k even, $T^k(G)$ is subcubic, planar and bipartite as well.*

Proof. Any Hamiltonian cycle of G can be clearly extended to a Hamiltonian cycle of $T^k(G)$. Conversely, suppose $T^k(G)$ contains a Hamiltonian cycle C. It is easy to see that if C contains the edges u'_0z and u'_1z, then it contains the edges 1 and 2 but not the edge 0. The other cases are symmetric. Therefore, by contracting the gadget to a single vertex, we obtain a Hamiltonian cycle of G.

The second statement is immediate. □

Proof of Theorem 5.2.10. We reduce from Hamiltonian Cycle, known to be NP-complete even for planar cubic bipartite graphs [4]. Given a planar cubic bipartite graph $G = (V, E)$, we construct a graph G' as follows. We first apply $T^{2\ell}$ to each $v \in V$. Note that, for $k \leq 2\ell$, no C_k is created inside any gadget and the length of a cycle passing through $v \in V$ increases by at least 4. Therefore, applying $T^{2\ell}$ sufficiently many times, we get rid of the induced cycles C_i with $i \leq 2\ell$ and so, by Lemma 5.2.11, the resulting G' is a planar bipartite $(C_4, \ldots, C_{2\ell})$-free graph with maximum degree 3. Finally, again by Lemma 5.2.11, G contains a Hamiltonian cycle if and only if G' does. □

Theorem 5.2.10 shows that the class of subcubic forests is limit for Hamiltonian Cycle with respect to planar bipartite graphs. On the other hand, we suspect this class is not minimal and a natural candidate for minimality would be the class $\mathcal{Q}$[5]. We leave as an open problem to determine if this is the case.

In the following, we show that a result similar to Theorem 5.2.10 holds for Hamiltonian Cycle Through Specified Edge and Hamiltonian Path.

Theorem 5.2.12. *For any $\ell \geq 2$,* Hamiltonian Cycle Through Specified Edge *is* NP-*complete even for planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3.*

Proof. We reduce from Hamiltonian Cycle for planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3, which is NP-complete by Theorem 5.2.10. Let G be an instance of this problem. Clearly, we may assume G has a vertex u of degree 3. By applying $T^{2\ell}$ to u, we obtain a planar bipartite $(C_4, \ldots, C_{2\ell})$-free graph G' with maximum degree 3. Let v be the neighbour of u_1 in the gadget introduced by $T^{2\ell}$ and such that $d(v) = 2$. We set $e = u_1v$.

[5] Note that $\mathcal{Q}$ is not a priori a limit class with respect to planar bipartite graphs (cfr. Remark 5.1.2).

It is easy to see that G' contains a Hamiltonian cycle through e if and only if G contains a Hamiltonian cycle. □

Theorem 5.2.13. *For any $\ell \geq 2$,* HAMILTONIAN PATH *is* NP*-complete even for planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3 and exactly two vertices of degree 1.*

Proof. Let $G = (V, E)$ and $uv \in E$ be an instance of HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE, where G is a planar bipartite $(C_4, \ldots, C_{2\ell})$-free graph with maximum degree 3. Clearly, we may assume G has no vertices of degree 1. Our reduction constructs a graph $G' = (V', E')$ as follows. Let $V' = V \cup \{a, b\}$, where a, b are new vertices, and $E' = (E \setminus \{uv\}) \cup \{au, bv\}$. Clearly, G' is a planar bipartite $(C_4, \ldots, C_{2\ell})$-free graph with maximum degree 3 and where a and b are the only vertices of degree 1.

It is easy to see that G has a Hamiltonian cycle through uv if and only if G' has a Hamiltonian path (between a and b). □

Theorem 5.2.13 will be used in Section 5.4 in the search of a boundary class for CONNECTED DOMINATING SET.

5.3 Feedback Vertex Set

In this section, we provide the first boundary class for FEEDBACK VERTEX SET. Ueno et al. [180] showed that FEEDBACK VERTEX SET and the related CONNECTED VERTEX COVER can be solved in polynomial time for subcubic graphs by a reduction to a matroid matching problem. On the other hand, FEEDBACK VERTEX SET is NP-hard for planar graphs with maximum degree at most 4, as first shown by Speckenmeyer [175] (see Theorem 2.2.16 for a strengthening), and so it admits a boundary class with respect to the class of planar graphs with maximum degree at most 4. We begin by showing that the class of forests whose components have at most four leaves and no two vertices of degree three is in fact a limit class.

For $k \geq 1$, we denote by $\mathcal{S}_k$ the class of planar bipartite $(C_4, \ldots, C_{2k}, H_1, \ldots, H_k)$-free graphs with maximum degree at most 4 (see Figure 5.5).

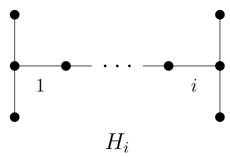


Figure 5.5: The graph H_i.

Theorem 5.3.1. *For any $k \geq 1$,* FEEDBACK VERTEX SET *is* NP*-complete for graphs in $\mathcal{S}_k$.*

Proof. We reduce from FEEDBACK VERTEX SET for planar graphs with maximum degree at most 4, which is NP-complete by Theorem 2.2.16. Given a planar graph G with maximum degree at most 4, we construct a graph G' by subdividing each edge of G with $2k + 1$ new vertices. It is easy to see that $G' \in \mathcal{S}_k$ and $\tau_c(G) = \tau_c(G')$. □

We denote by $\mathcal{S}$ the class of forests whose components have at most four leaves and no two vertices of degree three. Each graph in $\mathcal{S}$ has components of the form $S_{i,j,k,\ell}$, for some non-negative integers i, j, k, ℓ (see Figure 5.6). It is not difficult to see that $\bigcap_{k\geq 1} \mathcal{S}_k = \mathcal{S}$ and so $\mathcal{S}$ is a limit class for FEEDBACK VERTEX SET with respect to the class of planar bipartite graphs with maximum degree at most 4.

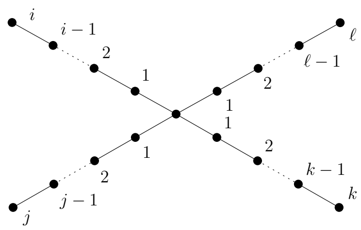


Figure 5.6: The graph $S_{i,j,k,\ell}$.

In order to show the minimality of $\mathcal{S}$, we first prove that FEEDBACK VERTEX SET can be solved in polynomial time for graphs with maximum degree at most 4 and bounded number of 4-vertices. This result follows from the fact that even the weighted version of FEEDBACK VERTEX SET can be solved in polynomial time for cubic graphs which, in turn, is an easy corollary of a deep result on the weighted matroid matching problem recently obtained by Iwata [99] and Pap [154]. Therefore, let us begin by recalling the matroid machinery introduced in Section 3.4.1. Note that in the following we allow graphs to contain loops and multiple edges, unless otherwise stated.

A 2-polymatroid is a pair $P = (S, f)$, where S is a finite set and f is a function $f\colon 2^S \to \mathbb{Z}$ satisfying the following properties:

(P1) $f(\varnothing) = 0$;

(P2) $f(X) \leq f(Y)$, for any $X \subseteq Y \subseteq S$;

(P3) $f(X \cup Y) + f(X \cap Y) \leq f(X) + f(Y)$, for any $X, Y \subseteq S$;

(P4) $f(\{x\}) \leq 2$, for any $x \in S$.

In Section 3.4.1, we showed that if G is a cubic graph, the function $f\colon 2^{V(G)} \to \mathbb{Z}$ defined by $f(X) = \mu(G) - \mu(G - X)$ gives a 2-polymatroid $P(G) = (V(G), f)$. The connection between polymatroids and feedback vertex sets was then revealed in the following:

Theorem 3.4.6 (Ueno et al. [180]). *Let G be a cubic graph. A subset $T \subseteq V(G)$ is a feedback vertex set of G if and only if it is a spanning set of the 2-polymatroid $P(G)$. Moreover, $I \subseteq V(G)$ is a nonseparating independent set of G if and only if it is a matching of $P(G)$.*

Recall that a subset $X \subseteq S$ is a matching of $P = (S, f)$ if $f(X) = 2|X|$ and it is a spanning set if $f(X) = f(S)$. Lovász [131] provided a polynomial-time algorithm that finds a maximum matching of a special class of 2-polymatroids, the so-called linearly represented 2-polymatroids (see also [71, 132, 151]). A 2-polymatroid (S, f) is linearly representable (over a field $\mathbb{F}$) if there exists a matrix $A = (A_e)_{e\in S} \in \mathbb{F}^{d\times 2S}$ obtained by concatenating $|S|$ matrices $A_e \in \mathbb{F}^{d\times 2}$ and such that $f(X) = \text{rank } A(X)$, for any $X \subseteq S$, where d is a positive integer and $A(X) = (A_e)_{e\in X}$ denotes the submatrix of A obtained by selecting the corresponding columns. In view of Theorem 1.0.9, Lovász's result implies we can find in polynomial time a minimum spanning set of a linearly represented 2-polymatroid. Moreover, Ueno et al. [180] showed that the 2-polymatroid $P(G)$ is linearly representable and that a linear representation can be obtained in polynomial time. Therefore, the mentioned results and Theorem 3.4.6 imply we can find in polynomial time a maximum nonseparating independent set and a minimum feedback vertex set of any cubic graph. Note that the same conclusion holds for a minimum connected vertex cover, since the complement of a nonseparating independent set of a connected graph is a connected vertex cover. Moreover, applying standard cleaning procedures (see Theorem 5.3.4 for an example), it is easy to see that these results hold for subcubic graphs.

In a recent breakthrough, Iwata [98, 99] and Pap [154] showed, independently, that even the following weighted version of the matroid matching problem can be solved in polynomial time:

WEIGHTED MATROID MATCHING	
Instance:	A linearly represented 2-polymatroid on S and a function $w\colon S \to \mathbb{R}$.
Task:	Find a matching $X \subseteq S$ with maximum $w(X) = \sum_{x \in X} w(x)$.

We now highlight the interesting consequences that this result has for the weighted versions of FEEDBACK VERTEX SET and CONNECTED VERTEX COVER. Indeed, consider a 2-polymatroid (S, f) and its 2-*dual* (S, f^*), where $f^*\colon 2^S \to \mathbb{Z}$ is defined by $f^*(X) = 2|X| + f(S \setminus X) - f(S)$ (see, e.g., [153]). Clearly, X is a matching of (S, f^*) if and only if $S \setminus X$ is a spanning set of (S, f) and so the problem of finding a minimum weight spanning set for (S, f) is equivalent to WEIGHTED MATROID MATCHING for the 2-dual (S, f^*). Moreover, given a linear representation of (S, f), it is not difficult to find a linear representation of (S, f^*) in polynomial time. Therefore, by Theorem 3.4.6, the following holds:

Theorem 5.3.2. *The problems of finding a minimum-weight feedback vertex and a minimum-weight connected vertex cover of a cubic graph can be solved in polynomial time.*

Using Theorem 5.3.2, we can show that FEEDBACK VERTEX SET is solvable in polynomial time for graphs with maximum degree at most 4 and bounded number of 4-vertices. Since we want to reduce this problem to the cubic case, we first introduce the following operation.

Let $x \in V(G)$ be a 4-vertex of a graph G (as already remarked, G may contain loops and multiple edges). A *vertex stretching* with respect to x is the operation replacing x with two new vertices x_1 and x_2 as depicted in Figure 5.7. Note that there is some freedom in determining the neighbours of x_1 and x_2. A vertex stretching with respect to x has the following fundamental properties:

- It does not alter the cycles of G, except possibly increasing the length of a cycle through x by 1;
- It does not modify the degrees of the vertices adjacent to x and it gives $d(x_1) = d(x_2) = 3$.

The following is an immediate but useful observation:

Lemma 5.3.3. *Let $x \in V(G)$ be a 4-vertex of G and let G' be the graph obtained from G by a vertex stretching with respect to x. We have that $\tau_c(G - x) = \tau_c(G' - \{x_1, x_2\})$. Moreover, G has a feedback vertex set avoiding x if and only if G' has a feedback vertex set avoiding $\{x_1, x_2\}$*

and the size of a minimum feedback vertex set of G avoiding x is the same as that of a minimum feedback vertex set of G' avoiding $\{x_1, x_2\}$.

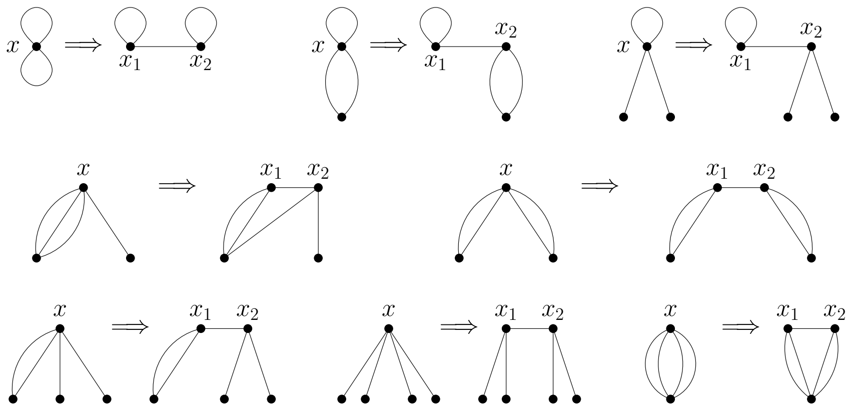


Figure 5.7: Vertex stretching.

We can finally prove the following:

Theorem 5.3.4. FEEDBACK VERTEX SET *can be solved in polynomial time for graphs with maximum degree at most* 4 *and bounded number of* 4*-vertices.*

Proof. Let G be a graph with $\Delta(G) \leq 4$ and $|d_4(G)| \leq c$, for some constant c. The following observations are immediate: if v is a 1-vertex of G, then $\tau_c(G) = \tau_c(G - v)$; if v is a 2-vertex of G which is the endpoint of a loop, then $\tau_c(G) = \tau_c(G - v) + 1$; if v is a 2-vertex of G adjacent to u and u' (note that we might have $u = u'$) then, denoting by H the graph obtained from G by deleting v and adding the edge uu', we have $\tau_c(G) = \tau_c(H)$. Therefore, we can reduce FEEDBACK VERTEX SET for G to the same problem for the graph G' obtained by the following cleaning procedure: if v is a 1-vertex or a 2-vertex which is the endpoint of a loop, we delete v; if v is a 2-vertex adjacent to u and u' (possibly $u = u'$), we delete v from G and add the edge uu'. It is easy to see that the graph G' obtained by applying these operations as long as possible either contains only vertices of degree 3 or 4, or it is empty.

We now proceed by a brute force argument: for each subset $S \subseteq d_4(G')$, we find a minimum feedback vertex set T of G' subject to $T \cap d_4(G') = S$ (if any). This is done as follows. We fix $S \subseteq d_4(G')$ and we apply a vertex stretching to each 4-vertex of G' in order to obtain a cubic graph G''. Then we define a weight function $w \colon V(G'') \to \mathbb{R}$ as follows: $w(x) = 0$, if x is the result of a vertex stretching with respect to a vertex in S; $w(x) = |V(G'')| + 1$, if x is the result of a vertex stretching with respect to a vertex in $d_4(G') \setminus S$; $w(x) = 1$ otherwise.

By Theorem 5.3.2, we can find in polynomial time a minimum-weight feedback vertex set T_S of (G'', w). Note that, without loss of generality, we may assume that every zero-weight vertex belongs to T_S. If $w(T_S) > |V(G'')|$ then, by Lemma 5.3.3, there exists no feedback vertex set T of G' such that $T \cap d_4(G') \subseteq S$. Otherwise, we remove from T_S the vertices which are the result of a vertex stretching with respect to a 4-vertex v (which must belong to S) and we add v to T_S, in order to obtain a set T'_S. Clearly, T'_S is a minimum feedback vertex set of G' subject to $T'_S \cap d_4(G') = S$. We thus build a vector indexed by the at most 2^c subsets S of $d_4(G')$ and we return the minimum value of $|T'_S|$. Correctness and polynomiality are evident. $\square$

We now have all the machinery to show that $\mathcal{S}$ is a boundary class for FEEDBACK VERTEX SET. Our proof relies once again on Lemma 5.1.9. Note that, in the following, all considered graphs are simple, i.e. loops and multiple edges are not allowed anymore.

Theorem 5.3.5. *$\mathcal{S}$ is a boundary class for* FEEDBACK VERTEX SET.

Proof. Let $\mathcal{S}'_p$ denote the class of $(C_3, \ldots, C_p, H_1, \ldots, H_p)$-free graphs with maximum degree 4. In view of Lemma 5.1.9, it is enough to show that, for each $H \in \mathcal{S}$, there exists a constant p such that FEEDBACK VERTEX SET is solvable in polynomial time for H-free graphs in $\mathcal{S}'_p$. Therefore, consider $H \in \mathcal{S}$ and let c be the number of its components. Since each component of H is of the form $S_{i,j,k,\ell}$, for some non-negative integers i, j, k and ℓ, there exists a non-negative integer d such that H is an induced subgraph of $cS_{d,d,d,d}$. We define $p = 2 \cdot 3^d$ and we claim that FEEDBACK VERTEX SET can be solved in polynomial time for $cS_{d,d,d,d}$-free graphs in $\mathcal{S}'_p$. Lemma 5.1.9 would then imply that $\mathcal{S}$ is a boundary class.

Let G be a $cS_{d,d,d,d}$-free graph in $\mathcal{S}'_p$ and consider a 4-vertex $v \in V(G)$ (if G is subcubic, we know the problem is solvable in polynomial time). Let G_v be the subgraph of G induced by the set of vertices at distance at most d from v. Since G_v belongs to $\mathcal{S}'_p$, it is a $(C_3, \ldots, C_p, H_1, \ldots, H_p)$-free graph with maximum degree 4. Moreover, since each vertex of G_v has degree at most 4, we have that $|V(G_v)| < 2 \cdot 3^d = p$ (Lemma 5.2.8) and so G_v is (C_k, H_k)-free for any $k > p$. Therefore, G_v belongs to $\mathcal{S}$ and, being connected, it must be of the form S_{i_1,i_2,i_3,i_4}, for some non-negative integers $i_1, i_2, i_3, i_4 \leq d$. Suppose now some i_j is strictly less than d and let v_{i_j} be the leaf of S_{i_1,i_2,i_3,i_4} at distance i_j from v. Clearly, no cycle of G contains a vertex belonging to the unique v, v_{i_j}-path P in S_{i_1,i_2,i_3,i_4} and different from v. Therefore, we may delete $V(P) \setminus \{v\}$ from G. Repeating this operation for each 4-vertex $v \in V(G)$, we obtain a graph G' with maximum degree at most 4. This graph is such that $\tau_c(G) = \tau_c(G')$ and, for each 4-vertex $v \in V(G')$, the induced subgraph G_v is isomorphic to $S_{d,d,d,d}$.

We now claim that G' has a bounded number of 4-vertices. In view of Theorem 5.3.4, this would conclude the proof. Let $F \subseteq d_4(G')$ be a subset of maximum size such that the corresponding induced copies of $S_{d,d,d,d}$ (for each $v \in F$, the induced subgraph G_v is isomorphic to $S_{d,d,d,d}$) are pairwise vertex-disjoint and denote by F' this corresponding set. Note that no two subgraphs in F' are connected by an edge, or else a copy of H_{2d+1} would arise in $G' \in \mathcal{S}'_{2 \cdot 3^d}$. Therefore, since G' is $cS_{d,d,d,d}$-free, we have $|F| < c$. Moreover, we claim that $|d_4(G')| \leq 17|F|$. By definition, for each $v \in d_4(G') \setminus F$, we have that G_v intersects a graph in F'. It is therefore enough to show that each branch of $S_{d,d,d,d}$ can intersect at most four other copies of $S_{d,d,d,d}$, where a branch is the unique path between a leaf and the 4-vertex. To this end, let S be a copy of $S_{d,d,d,d}$ with 4-vertex v and let P be a v, v_i-branch. The following are easy observations. If v belongs to another copy S' of $S_{d,d,d,d}$, then $V(P) \subseteq V(S')$ and v_i is the 4-vertex of S', and so P intersects at most four other copies of $S_{d,d,d,d}$. If v does not belong to another copy of $S_{d,d,d,d}$ but an inner vertex of P does, then P intersects at most one other copy of $S_{d,d,d,d}$. Finally, if v_i is the only vertex of P which belongs to another copy of $S_{d,d,d,d}$, then P intersects at most three other copies of $S_{d,d,d,d}$. □

We have seen in Theorem 2.2.16 that FEEDBACK VERTEX SET is NP-hard for line graphs of subcubic triangle-free graphs and so it admits a boundary class with respect to this subclass of line graphs. In the rest of this section, we make some progress towards the determination of such a boundary class.

Recall that $\mathcal{Q}_p$ is the class of subcubic $\mathcal{Y}_p \cup \mathcal{C}_p$-free graphs such that each cubic vertex has a non-cubic neighbour and $\mathcal{R}_p$ is the subclass of $\mathcal{Q}_p$ consisting of the graphs G with $\delta(G) \geq 2$ (see Section 5.2). By Lemma 5.2.6, HAMILTONIAN PATH is NP-complete for graphs in $\mathcal{R}_p$, for any $p \geq 1$, and so it is clearly NP-complete for triangle-free graphs in $\mathcal{R}_p$.

Lemma 5.3.6. *For any $p \geq 1$,* FEEDBACK VERTEX SET *is* NP*-complete for line graphs of triangle-free graphs in $\mathcal{R}_p$.*

Proof. We reduce from HAMILTONIAN PATH for triangle-free graphs in $\mathcal{R}_p$. Let $G = (V, E)$ be an instance of this problem. In particular, G is a subcubic triangle-free graph with $\delta(G) \geq 2$. Consider now its line graph $G' = L(G)$. By Lemma 2.2.13, we have that $\tau_c(G') \leq \frac{|d_3(G)|}{2} + 1$ if and only if G contains a Hamiltonian path. The conclusion immediately follows. □

Denoting by $L(\mathcal{Q})$ the class $\{L(G) : G \in \mathcal{Q}\}$, Lemma 5.3.6 implies the following:

Corollary 5.3.7. *$L(\mathcal{Q})$ is a limit class for* FEEDBACK VERTEX SET.

We suspect that $L(\mathcal{Q})$ is indeed a minimal limit class and we leave this verification as an open problem.

5.4 Connected Dominating Set

In this section, we consider CONNECTED DOMINATING SET and we show that the class of subcubic forests is a limit class for this problem. Alekseev et al. [7] showed that the class $\mathcal{T}$ of forests whose components have at most three leaves is boundary for the related DOMINATING SET. Clearly, $\mathcal{T}$ is contained in the class of subcubic forests but we provide some evidence for the fact that it may not be boundary for CONNECTED DOMINATING SET.

Douglas [53] showed that the following variant of CONNECTED DOMINATING SET is NP-hard even for subcubic planar graphs:

$(\frac{\lvert V\rvert}{2} - 1)$-CONNECTED DOMINATING SET	
Instance:	A graph $G = (V, E)$.
Question:	Does $\gamma_c(G) \leq \frac{\lvert V\rvert}{2} - 1$ hold?

In particular, CONNECTED DOMINATING SET is NP-hard for subcubic planar graphs and in the following we show that the same holds for the class of subcubic planar bipartite graphs with arbitrarily large girth. This would clearly imply that the class of subcubic forests is a limit class for CONNECTED DOMINATING SET.

Recall that, for any graph G, we have $\gamma_c(G) = |V(G)| - \ell(G)$[6]. Moreover, it is easy to see that for any subcubic graph G, we have $\ell(G) \leq \frac{|V(G)|}{2} + 1$, with equality if and only if G contains a $\{1, 3\}$-spanning tree, i.e. a spanning tree with no vertices of degree 2. Therefore, $(\frac{|V|}{2} - 1)$-CONNECTED DOMINATING SET restricted to subcubic graphs is polynomially equivalent to the following problem:

[6] $\ell(G)$ denotes the maximum number of leaves in a spanning tree of G.

$\{1,3\}$-SPANNING TREE	
Instance:	A graph $G = (V, E)$.
Question:	Does there exist a spanning tree T of G such that $d_T(v)$ is either 1 or 3, for any $v \in V$?

In fact, Douglas [53] first showed that $\{1,3\}$-SPANNING TREE is NP-hard for subcubic planar graphs and then used the equivalence above to deduce that the same holds for $(\frac{|V|}{2} - 1)$-CONNECTED DOMINATING SET. In the following, we take the same path and show that $\{1,3\}$-SPANNING TREE is NP-hard even for the class of subcubic planar bipartite graphs with arbitrarily large girth:

Theorem 5.4.1. *For any $\ell \geq 2$, $\{1,3\}$-SPANNING TREE is NP-complete even for planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3.*

Proof. We reduce from HAMILTONIAN PATH for planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3 and exactly two vertices of degree 1, which is NP-complete by Theorem 5.2.13. Given such a graph $G = (V, E)$, our reduction constructs a graph $G' = (V', E')$ as follows. We replace each 3-vertex u of G by the gadget depicted in Figure 5.8 with $k = 2\ell + 1$. We proceed similarly for any 2-vertex u of G (in this case, u_3 will be a 2-vertex in G'). Note that there is some freedom on the way the edges incident to u are attached to the gadget. It is easy to see that G' is a planar bipartite $(C_4, \ldots, C_{2\ell})$-free graph with maximum degree 3.

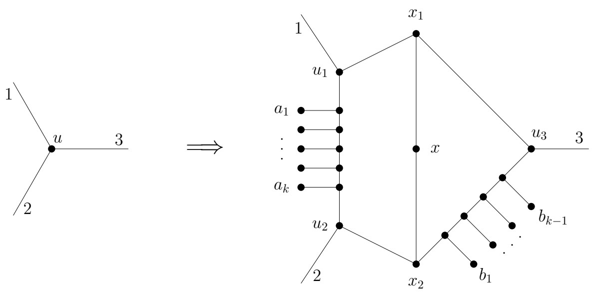
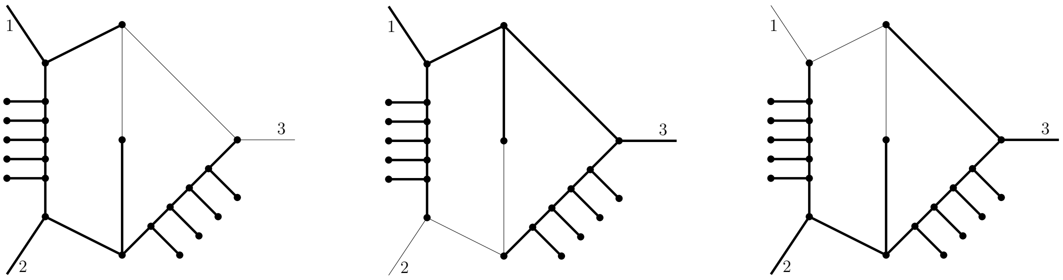


Figure 5.8: Construction of the graph G': The vertex u is replaced by a gadget containing $2(2\ell + 1) + 2 \cdot 2\ell + 6$ vertices.

Let v_1 and v_2 be the vertices of degree 1 in G (and so in G'). We claim that G has a Hamiltonian path (between v_1 and v_2) if and only if there exists a spanning tree T of G' such that $d_T(v)$ is either 1 or 3, for any $v \in V'$.

Suppose first G contains a Hamiltonian path P between v_1 and v_2. Each $u \in V \setminus \{v_1, v_2\}$ is incident to exactly two edges in $E(P)$. We select the corresponding edges in G' and, for each gadget, we select the bold edges as described in Figure 5.9 (if the gadget substitutes a 2-vertex of G, the construction is similar). In this way, we obtain a spanning tree T of G' such that $d_T(v)$ is either 1 or 3, for any $v \in V'$.

Conversely, suppose there exists a spanning tree T of G' such that $d_T(v)$ is either 1 or 3, for any $v \in V'$, and consider the gadget replacing a 3-vertex $u \in V$ (see Figure 5.8). Clearly, the edges incident to the 1-vertices a_i's and b_i's are all in $E(T)$. Moreover, by the degree constraint and the connectedness of T, all the edges incident to the 3-neighbours of the a_i's and b_i's are in $E(T)$. We now claim that $|E(T) \cap \{1,2,3\}| = 2$. Since T is a spanning tree, $|E(T) \cap \{1,2,3\}| \geq 1$. Suppose first $\{1,2,3\} \subseteq E(T)$. By the degree constraint, we have



(a) **(b)** **(c)**

Figure 5.9: Construction of the spanning tree T. If P passes through the edges 1 and 2, we select the bold edges in (a). Similarly for the other two cases illustrated in (b) and (c).

that $\{u_1x_1, u_2x_2, u_3x_1\} \subseteq E(T)$ and so T contains a cycle, a contradiction. Suppose now $|E(T) \cap \{1,2,3\}| = 1$ and, without loss of generality, $E(T) \cap \{1,2,3\} = \{1\}$ (the other two cases are treated similarly). By the degree constraint, we have $u_1x_1 \in E(T)$, $u_2x_2 \notin E(T)$ and $u_3x_1 \notin E(T)$, a contradiction to the fact that at least one of xx_1 and xx_2 is in $E(T)$. Therefore, we have $|E(T) \cap \{1,2,3\}| = 2$. If $xx_1 \in E(T)$, then it must be $E(T) \cap \{1,2,3\} = \{1,3\}$ (see Figure 5.9(b)). Otherwise, i.e. if $xx_2 \in E(T)$, we either have $E(T) \cap \{1,2,3\} = \{1,2\}$ or $E(T) \cap \{1,2,3\} = \{2,3\}$ (see Figures 5.9(a) and 5.9(c)). A similar reasoning applies to the gadget replacing a 2-vertex. But then, by contracting each gadget in G' to a single vertex, we obtain a connected spanning subgraph P of G such that, for each vertex of degree at least 2, exactly two of its incident edges are in $E(P)$. This implies that P is a Hamiltonian v_1, v_2-path in G. □

We have seen that a subcubic graph G contains a $\{1,3\}$-spanning tree if and only if $\gamma_c(G) \leq \frac{|V(G)|}{2} - 1$. Therefore, Theorem 5.4.1 has the following two immediate consequences:

Corollary 5.4.2. *For any $\ell \geq 2$, $(\frac{|V|}{2} - 1)$-*Connected Dominating Set *is* NP-*complete even for planar bipartite $(C_4, \ldots, C_{2\ell})$-free graphs with maximum degree 3.*

Corollary 5.4.3. *The class of subcubic forests is a limit class for $(\frac{|V|}{2} - 1)$-*Connected Dominating Set *and* Connected Dominating Set.

We now show that the class $\mathcal{T}$ of forests whose components have at most three leaves (which is boundary for Dominating Set) is not boundary for $(\frac{|V|}{2} - 1)$-Connected Dominating Set. Clearly, $\mathcal{T}$ does not contain the graph H_1 depicted in Figure 5.5.

Lemma 5.4.4. *If X is a class of subcubic forests which is boundary for $(\frac{|V|}{2} - 1)$-*Connected Dominating Set*, then $H_1 \in X$.*

Proof. Since $(\frac{|V|}{2} - 1)$-Connected Dominating Set is polynomially equivalent to $\{1,3\}$-Spanning Tree when restricted to subcubic graphs, we have that X is a boundary class for $\{1,3\}$-Spanning Tree with respect to the class Y of subcubic (C_3, C_4)-free graphs.

Suppose now $H_1 \notin X$. Then the class $Y \cap Free(H_1)$ contains X and it is defined by finitely many forbidden induced subgraphs with respect to Y. Moreover, by the previous paragraph and Theorem 5.1.8, we have that $\{1,3\}$-Spanning Tree is NP-hard for $Y \cap Free(H_1)$. On the other hand, consider a graph $G \in Y \cap \mathit{Free}(H_1)$. If G has maximum degree 2, then $\{1,3\}$-Spanning Tree is trivial for G. Moreover, if G contains a cubic vertex v, then each neighbour of v has degree at most 2. This implies that v has degree 3 in a $\{1,3\}$-spanning tree of G and so such a tree exists if and only if $G = K_{1,3}$. Therefore, $\{1,3\}$-Spanning Tree is trivial for the class $Y \cap \mathit{Free}(H_1)$, a contradiction. □

Remark 5.4.5. Similarly to Lemma 5.4.4, it is not difficult to show that if X is a class of subcubic forests which is boundary for $(\frac{|V|}{2} - 1)$-Connected Dominating Set, then X contains a tribranch. Therefore, the class $\mathcal{Q}$ is not boundary for this problem either.

Unfortunately, we do not know whether Lemma 5.4.4 holds for Connected Dominating Set and the major open problem is to find a boundary class with respect to subcubic planar bipartite graphs. Another interesting problem (related to Lemma 5.4.4) is to determine the computational complexity of Connected Dominating Set for (C_3, C_4, H_1)-free subcubic graphs. We have seen that $(\frac{|V|}{2} - 1)$-Connected Dominating Set is trivial for this class and we know that Dominating Set is NP-hard for it [7]. Moreover, in Section 4.4 we showed that Dominating Set and Connected Dominating Set belong to the same complexity class when restricted to $\mathit{Free}(H)$, for any graph H. If Connected Dominating Set is polynomial for (C_3, C_4, H_1)-free subcubic graphs, we would obtain a first example[7] of a non-trivial hereditary class for which Dominating Set and Connected Dominating Set belong to different complexity classes.

5.5 Connected Vertex Cover

In this short section, we consider the connected variant of the vertex cover problem which asks for a minimum-size vertex cover inducing a connected graph:

Connected Vertex Cover	
Instance:	A graph $G = (V, E)$ and a positive integer k.
Question:	Does $\beta_c(G) \leq k$ hold?

Connected Vertex Cover was introduced by Garey and Johnson [76], who showed it is NP-complete for planar graphs with maximum degree 4. Fernau and Manlove [63] strengthened this result by showing that it remains NP-hard even for planar bipartite graphs with maximum degree 4 (see also [60]). On the other hand, we have seen in Section 5.3 that it is solvable in polynomial time for subcubic graphs. In the following, we make some observations towards determining the first boundary classes for this problem.

Alekseev [5] showed that the class of forests whose components have at most three leaves is boundary for Vertex Cover[8] and conjectured no other boundary class exists. For Connected Vertex Cover, we show there are at least two boundary classes. One of them is a subclass of line graphs of bipartite graphs:

Lemma 5.5.1. Connected Vertex Cover *is* NP-*complete even for line graphs of planar cubic bipartite graphs.*

[7] To the best of our knowledge. [8] He actually stated this result for Independent Set but the two problems are polynomially equivalent.

Note that, on the contrary, VERTEX COVER restricted to line graphs can be solved in polynomial time by a reduction to a matching problem [139, 167].

Our proof of Lemma 5.5.1 is based on the following lemma:

Lemma 5.5.2. *If G is the line graph of a cubic triangle-free graph H, then $\beta_c(G) \geq \frac{2}{3}|V(G)|$. Equality holds if and only if H contains a Hamiltonian cycle.*

Proof. Clearly, there is a bijection between the vertices of H and the triangles of G. Moreover, any vertex cover of G contains at least two vertices for every triangle. Since there are $|V(H)| = \frac{2}{3}|V(G)|$ triangles in G and any two of them share at most one vertex, we have that $\beta_c(G) \geq \beta(G) \geq \frac{2}{3}|V(G)|$.

Suppose now that $\beta_c(G) = \frac{2}{3}|V(G)|$. This means there exists a connected vertex cover S of G containing exactly two vertices for each triangle of G. Consider now the set of edges $S' \subseteq E(H)$ corresponding to S. Since $G[S]$ is connected, we have that $H[S']$ is connected as well. Therefore, $S' \subseteq E(H)$ is a set of edges such that each vertex of H is incident to exactly two of them and $H[S']$ is connected and so S' induces a Hamiltonian cycle in H.

Conversely, suppose H contains a Hamiltonian cycle C and let S be the set of vertices of G corresponding to $E(C)$. Since every edge of H is incident to an edge in $E(C)$, we have that S is a vertex cover of G. Moreover, $G[S]$ is connected and so $\beta_c(G) \leq |V(H)| = \frac{2}{3}|V(G)|$. □

Proof of Lemma 5.5.1. We reduce from HAMILTONIAN CYCLE for planar cubic bipartite graphs, which is known to be NP-complete [4]. Let G be an instance of this problem and consider its line graph $G' = L(G)$. By Lemma 5.5.2, we have that $\beta_c(G') \leq \frac{2}{3}|V(G')|$ if and only if G contains a Hamiltonian cycle. □

The results in [60, 63] mentioned above show that the class of planar bipartite graphs with maximum degree 4 is limit. On the other hand, we now show it is not minimal. The following operation proves to be helpful: Given a graph $G = (V, E)$, an edge $uv \in E$ and an integer $p \geq 1$, the graph $A_p(G)$ is obtained from G by replacing uv with the gadget depicted in Figure 5.10.

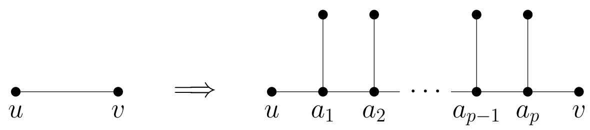


Figure 5.10: The operation A_p.

The fundamental property of A_p, which is left as an easy exercise, is that $\beta_c(A_p(G)) = \beta_c(G) + p$. We can now provide the other limit class (see Figure 5.11):

Lemma 5.5.3. *For any $k \geq 1$,* CONNECTED VERTEX COVER *is* NP-*complete for planar bipartite $(C_4, \ldots, C_{2k}, H'_1, \ldots, H'_k)$-free graphs with maximum degree at most 4.*

Proof. We reduce from CONNECTED VERTEX COVER for planar graphs with maximum degree 4, which is NP-complete by Lemma 5.5.1. Given an instance G of this problem we construct a graph G' by applying the operation A_p, with $p = 2k + 1$, to each edge of G. The statement follows from the fact that $\beta_c(A_p(G)) = \beta_c(G) + p$. □

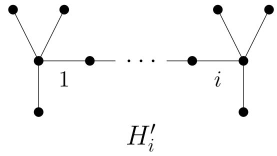


Figure 5.11: The graph H'_i.

Note that the two boundary classes whose existence is guaranteed by Lemmas 5.5.1 and 5.5.3 are distinct. Indeed, by Lemma 5.5.1, there exists a boundary class $\mathcal{C}_1$ which is a subclass of line graphs and, by Lemma 5.5.3, there exists a boundary class $\mathcal{C}_2$ which is a subclass of forests. On the other hand, $\mathcal{C}_1$ must contain K_3 or else, by Theorem 5.1.8, CONNECTED VERTEX COVER would be NP-hard for triangle-free line graphs and so for graphs with maximum degree 2.

We finally conclude this section with the following problem:

Problem 5.5.4. Find boundary classes for CONNECTED VERTEX COVER with respect to the classes described in Lemmas 5.5.1 and 5.5.3.

The classes $L(\mathcal{Q})$ and $\mathcal{S}$ would be two natural candidates, respectively.

Bibliography

[1] H. AbouEisha, S. Hussain, V. Lozin, J. Monnot, and B. Ries. “A Dichotomy for Upper Domination in Monogenic Classes.” In: *Combinatorial Optimization and Applications*. Ed. by Z. Zhang, L. Wu, W. Xu, and D.-Z. Du. Vol. 8881. Lecture Notes in Computer Science. 2014, pp. 258–267 (cit. on pp. 4, 101).

[2] R. Aharoni. “Ryser’s Conjecture for Tripartite 3-Graphs.” In: *Combinatorica* 21(1) (2001), pp. 1–4 (cit. on pp. 47, 64).

[3] J. Akiyama and M. Kano. *Factors and Factorizations of Graphs*. Lecture Notes in Mathematics. Springer, 2011 (cit. on p. 21).

[4] T. Akiyama, T. Nishizeki, and N. Saito. “NP-Completeness of the Hamiltonian Cycle Problem for Bipartite Graphs.” In: *Journal of Information Processing* 3(2) (1980), pp. 73–76 (cit. on pp. 21, 33, 38, 99, 116, 126).

[5] V. E. Alekseev. “On easy and hard hereditary classes of graphs with respect to the independent set problem.” In: *Discrete Applied Mathematics* 132(1–3) (2003), pp. 17–26 (cit. on pp. 5, 105–107, 125, 143, 144).

[6] V. E. Alekseev and D. S. Malyshev. “A criterion for a class of graphs to be a boundary class and applications.” In: *Diskretn. Anal. Issled. Oper.* 15(6) (2008). (In Russian), pp. 3–10 (cit. on p. 109).

[7] V. E. Alekseev, D. V. Korobitsyn, and V. V. Lozin. “Boundary classes of graphs for the dominating set problem.” In: *Discrete Mathematics* 285(1–3) (2004), pp. 1–6 (cit. on pp. 5, 105, 106, 108, 122, 125).

[8] V. E. Alekseev, R. Boliac, D. V. Korobitsyn, and V. V. Lozin. “NP-hard graph problems and boundary classes of graphs.” In: *Theoretical Computer Science* 389(1–2) (2007), pp. 219–236 (cit. on pp. 5, 105–109).

[9] P. Anand, H. Escuadro, R. Gera, S. G. Hartke, and D. Stolee. “On the hardness of recognizing triangular line graphs.” In: *Discrete Mathematics* 312(17) (2012), pp. 2627–2638 (cit. on pp. 22, 43).

[10] V. Andova, F. Kardoš, and R. Škrekovski. “Mathematical aspects of fullerenes.” In: *Ars Mathematica Contemporanea* 11(2) (2016), pp. 353–379 (cit. on p. 80).

[11] R. P. Anstee. “A Survey of Forbidden Configuration Results.” In: *Electronic Journal of Combinatorics* DS20 (2013) (cit. on p. 84).

[12] M. Anthony, G. Brightwell, and C. Cooper. “The Vapnik-Chervonenkis Dimension of a Random Graph.” In: *Discrete Mathematics* 138(1) (1995), pp. 43–56 (cit. on p. 85).

[13] E. M. Arkin, S. P. Fekete, K. Islam, H. Meijer, J. S. B. Mitchell, Y. Núñez-Rodríguez, V. Polishchuk, D. Rappaport, and H. Xiao. “Not being (super)thin or solid is hard: A study of grid Hamiltonicity.” In: *Computational Geometry* 42(6) (2009), pp. 582–605 (cit. on pp. 107, 115).

[14] S. Arnborg, J. Lagergren, and D. Seese. “Easy problems for tree-decomposable graphs.” In: *Journal of Algorithms* 12(2) (1991), pp. 308–340 (cit. on p. 16).

[15] S. Arora and B. Barak. *Computational Complexity - A Modern Approach*. Cambridge University Press, 2009 (cit. on p. 12).

[16] L. Babai and P. Frankl. *Linear Algebra Method in Combinatorics*. Preliminary Version 2. University of Chicago, 1992 (cit. on p. 84).

[17] V. Bafna, P. Berman, and T. Fujito. “A 2-Approximation Algorithm for the Undirected Feedback Vertex Set Problem.” In: *SIAM Journal on Discrete Mathematics* 12(3) (1999), pp. 289–297 (cit. on p. 36).

[18] B. S. Baker. “Approximation Algorithms for NP-complete Problems on Planar Graphs.” In: *J. ACM* 41(1) (1994), pp. 153–180 (cit. on pp. 12, 61, 62).

[19] L. W. Beineke. “Characterizations of Derived Graphs.” In: *Journal of Combinatorial Theory* 9(2) (1970), pp. 129–135 (cit. on p. 19).

[20] L. W. Beineke and R. J. Wilson, eds. *Topics in Topological Graph Theory*. Cambridge University Press, 2009 (cit. on p. 78).

[21] P. Berman and M. Karpinski. *Efficient Amplifiers and Bounded Degree Optimization*. Tech. rep. TR01-053. Electronic Colloquium on Computational Complexity, 2003 (cit. on p. 34).

[22] P. Berman and M. Karpinski. *Improved Approximation Lower Bounds on Small Occurrence Optimization*. Tech. rep. TR03-008. Electronic Colloquium on Computational Complexity, 2003 (cit. on p. 36).

[23] A. A. Bertossi. “The edge Hamiltonian path problem is NP-complete.” In: *Information Processing Letters* 13(4–5) (1981), pp. 157–159 (cit. on p. 21).

[24] T. Biedl, E. D. Demaine, C. A. Duncan, R. Fleischer, and S. G. Kobourov. “Tight bounds on maximal and maximum matchings.” In: *Discrete Mathematics* 285(1–3) (2004), pp. 7–15 (cit. on pp. 2, 21, 26).

[25] H. L. Bodlaender. “A Linear-Time Algorithm for Finding Tree-Decompositions of Small Treewidth.” In: *SIAM Journal on Computing* 25(6) (1996), pp. 1305–1317 (cit. on p. 16).

[26] H. L. Bodlaender. “A partial k-arboretum of graphs with bounded treewidth.” In: *Theoretical Computer Science* 209(1) (1998), pp. 1–45 (cit. on pp. 16, 61).

[27] H. L. Bodlaender. *Planar Graphs With Bounded Treewidth*. Tech. rep. RUU-CS-88-14. Rijksuniversiteit Utrecht, 1988 (cit. on p. 61).

[28] J. A. Bondy and U. S. R. Murty. *Graph Theory*. Graduate Texts in Mathematics. Springer, 2008 (cit. on pp. 7, 26).

[29] A. Brandstädt. “(P_5,diamond)-free graphs revisited: structure and linear time optimization.” In: *Discrete Applied Mathematics* 138(1–2) (2004), pp. 13 –27 (cit. on p. 72).

[30] H. Brönnimann and M. T. Goodrich. “Almost Optimal Set Covers in Finite VC-Dimension.” In: *Discrete & Computational Geometry* 14(4) (1995), pp. 463–479 (cit. on p. 84).

[31] M. R. Cerioli, L. Faria, T. O. Ferreira, C. A. J. Martinhon, F. Protti, and B. Reed. “Partition into cliques for cubic graphs: Planar case, complexity and approximation.” In: *Discrete Applied Mathematics* 156(12) (2008), pp. 2270–2278 (cit. on pp. 46, 51, 60, 61, 143, 144).

[32] G. Chappell, J. Gimbel, and C. Hartman. “On cycle packings and feedback vertex sets.” In: *Contributions to Discrete Mathematics* 9(2) (2014) (cit. on pp. 48, 74).

[33] G. Chapuy, M. DeVos, J. McDonald, B. Mohar, and D. Scheide. “Packing Triangles in Weighted Graphs.” In: *SIAM Journal on Discrete Mathematics* 28(1) (2014), pp. 226–239 (cit. on p. 48).

[34] G. Chartrand. “On Hamiltonian line-graphs.” In: *Trans. Amer. Math. Soc.* 134 (1968), pp. 559–566 (cit. on pp. 21, 37).

[35] G. Chartrand. “The existence of complete cycles in repeated line-graphs.” In: *Bull. Amer. Math. Soc.* 71(4) (1965), pp. 668–670 (cit. on pp. 21, 37).

[36] H. B. Chen, H. L. Fu, and C. H. Shih. “Feedback vertex set on planar graphs.” In: *Taiwanese Journal of Mathematics* 16(6) (2012), pp. 2077–2082 (cit. on pp. 48, 74).

[37] M. Chlebík and J. Chlebíková. “Complexity of approximating bounded variants of optimization problems.” In: *Theoretical Computer Science* 354(3) (2006), pp. 320–338 (cit. on p. 60).

[38] M. Chlebík and J. Chlebíková. “The Complexity of Combinatorial Optimization Problems on d-Dimensional Boxes.” In: *SIAM Journal on Discrete Mathematics* 21(1) (2007), pp. 158–169 (cit. on p. 60).

[39] S. A. Choudum, T. Karthick, and M. A. Shalu. “Linear Chromatic Bounds for a Subfamily of $3K_1$-free Graphs.” In: *Graphs and Combinatorics* 24(5) (2008), pp. 413–428 (cit. on p. 49).

[40] M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas. “K_4-free graphs with no odd holes.” In: *Journal of Combinatorial Theory, Series B* 100(3) (2010), pp. 313–331 (cit. on p. 72).

[41] M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas. “The strong perfect graph theorem.” In: *Annals of Mathematics* 164(1) (2006), pp. 51–229 (cit. on p. 67).

[42] V. Chvátal and P. L. Hammer. “Aggregation of Inequalities in Integer Programming.” In: *Studies in Integer Programming*. Ed. by P. L. Hammer, E. L. Johnso, B. H. Korte, and G. L. Nemhauser. Vol. 1. Annals of Discrete Mathematics. 1977, pp. 145–162 (cit. on p. 96).

[43] M. Conforti, D. G. Corneil, and A. R. Mahjoub. “K_i-covers I: Complexity and polytopes.” In: *Discrete Mathematics* 58(2) (1986), pp. 121–142 (cit. on pp. 2, 21).

[44] M. Conforti, D. C. Corneil, and A. R. Mahjoub. “K_i-covers. II. K_i-perfect graphs.” In: *Journal of Graph Theory* 11(4) (1987), pp. 569–584 (cit. on pp. 2, 21).

[45] C. R. Cook. “Two characterizations of interchange graphs of complete m-partite graphs.” In: *Discrete Mathematics* 8(4) (1974), pp. 305–311 (cit. on pp. 2, 21).

[46] B. Courcelle. “The Monadic Second-Order Logic of Graphs. I. Recognizable Sets of Finite Graphs.” In: *Information and Computation* 85(1) (1990), pp. 12–75 (cit. on p. 16).

[47] B. Courcelle and S. Olariu. “Upper bounds to the clique width of graphs.” In: *Discrete Applied Mathematics* 101(1–3) (2000), pp. 77–114 (cit. on p. 16).

[48] B. Courcelle, J. Engelfriet, and G. Rozenberg. "Handle-Rewriting Hypergraph Grammars." In: *Journal of Computer and System Sciences* 46(2) (1993), pp. 218–270 (cit. on p. 16).

[49] B. Courcelle, J. A. Makowsky, and U. Rotics. "Linear Time Solvable Optimization Problems on Graphs of Bounded Clique-Width." In: *Theory of Computing Systems* 33(2) (2000), pp. 125–150 (cit. on pp. 16, 95).

[50] E. D. Demaine and M. T. Hajiaghayi. "The Bidimensionality Theory and Its Algorithmic Applications." In: *The Computer Journal* (2007) (cit. on p. 36).

[51] R. W. Deming. "Independence numbers of graphs-an extension of the Koenig-Egervary theorem." In: *Discrete Mathematics* 27(1) (1979), pp. 23–33 (cit. on p. 67).

[52] R. Diestel. *Graph Theory*. Graduate Texts in Mathematics. Springer, 2005 (cit. on pp. 7, 80, 113).

[53] R. J. Douglas. "NP-completeness and degree restricted spanning trees." In: *Discrete Mathematics* 105(1–3) (1992), pp. 41–47 (cit. on pp. 97, 122, 123).

[54] T. Došlić. "Cyclical Edge-Connectivity of Fullerene Graphs and $(k, 6)$-Cages." In: *Journal of Mathematical Chemistry* 33(2) (2003), pp. 103–112 (cit. on p. 80).

[55] Ding-Zhu Du, Ker-I Ko, and Xiaodong Hu. *Design and Analysis of Approximation Algorithms*. Springer Optimization and Its Applications. Springer, 2012 (cit. on p. 12).

[56] P. Duchet and H. Meyniel. "On Hadwiger's number and the stability number." In: *Ann. Discr. Math.* 13 (1982), pp. 71–74 (cit. on p. 101).

[57] M. N. Ellingham and J. D. Horton. "Non-hamiltonian 3-Connected Cubic Bipartite Graphs." In: *Journal of Combinatorial Theory, Series B* 34(3) (1983), pp. 350–353 (cit. on p. 37).

[58] P. Erdős. "Graph theory and probability." In: *Canad. J. Math.* 11 (1959), pp. 34–38 (cit. on p. 64).

[59] P. Erdős and L. Pósa. "On independent circuits contained in a graph." In: *Canad. J. Math.* 17 (1965), pp. 347–352 (cit. on pp. 3, 48, 74).

[60] B. Escoffier, L. Gourvès, and J. Monnot. "Complexity and approximation results for the connected vertex cover problem in graphs and hypergraphs." In: *Journal of Discrete Algorithms* 8(1) (2010), pp. 36–49 (cit. on pp. 125, 126).

[61] S. Fajtlowicz. "Independence, clique size and maximum degree." In: *Combinatorica* 4(1) (1984), pp. 35–38 (cit. on p. 25).

[62] S. Fajtlowicz. "On the size of independent sets in graphs." In: *Congressus Numerantium* 21 (1978), pp. 269–274 (cit. on p. 25).

[63] H. Fernau and D. F. Manlove. "Vertex and edge covers with clustering properties: Complexity and algorithms." In: *Journal of Discrete Algorithms* 7(2) (2009), pp. 149–167 (cit. on pp. 125, 126).

[64] F. V. Fomin and D. Kratsch. *Exact Exponential Algorithms*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2010 (cit. on p. 16).

[65] P. Frankl and J. Pach. "On the number of sets in a null t-design." In: *European Journal of Combinatorics* 4 (1983), pp. 21–23 (cit. on p. 84).

[66] K. Fraughnaugh and S. C. Locke. “11/30 (Finding Large Independent Sets in Connected Triangle-Free 3-Regular Graphs).” In: *Journal of Combinatorial Theory, Series B* 65(1) (1995), pp. 51–72 (cit. on pp. 1, 20, 25, 58).

[67] K. Fraughnaugh Jones. “Independence in Graphs with Maximum Degree Four.” In: *Journal of Combinatorial Theory, Series B* 37(3) (1984), pp. 254–269 (cit. on pp. 1, 20, 25, 57).

[68] D. Fronček. “Locally linear graphs.” In: *Mathematica Slovaca* 39(1) (1989), pp. 3–6 (cit. on p. 25).

[69] Z. Füredi and J. Pach. “Traces of finite sets: extremal problems and geometric applications.” In: *Extremal Problems for Finite Sets*. Ed. by P. Frankl et al. János Bolyai Math. Soc., 1994, pp. 251–282 (cit. on p. 84).

[70] M. L. Furst, J. L. Gross, and L. A. McGeoch. “Finding a Maximum-Genus Graph Imbedding.” In: *J. ACM* 35(3) (1988), pp. 523–534 (cit. on p. 79).

[71] H. N. Gabow and M. Stallmann. “An augmenting path algorithm for linear matroid parity.” In: *Combinatorica* 6(2) (1986), pp. 123–150 (cit. on p. 118).

[72] T. Gallai. “Kritische Graphen II.” In: *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 8 (1963), pp. 373–395 (cit. on p. 50).

[73] T. Gallai. “Transitiv Orientierbare Graphen.” In: *Acta Mathematica Academiae Scientiarum Hungarica* 18(1) (1967), pp. 25–66 (cit. on p. 22).

[74] P. Gambette and S. Vialette. “On Restrictions of Balanced 2-Interval Graphs.” In: *Graph-Theoretic Concepts in Computer Science*. Ed. by A. Brandstädt, D. Kratsch, and H. Müller. Vol. 4769. Lecture Notes in Computer Science. 2007, pp. 55–65 (cit. on p. 23).

[75] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979 (cit. on pp. 14, 93, 115).

[76] M. R. Garey and D. S. Johnson. “The Rectilinear Steiner Tree Problem is NP-Complete.” In: *SIAM Journal on Applied Mathematics* 32(4) (1977), pp. 826–834 (cit. on pp. 102, 125).

[77] M. Gentner and D. Rautenbach. “Feedback vertex sets in cubic multigraphs.” In: *Discrete Mathematics* 338(12) (2015), pp. 2179–2185 (cit. on p. 77).

[78] P. A. Golovach, M. Johnson, D. Paulusma, and J. Song. “A Survey on the Computational Complexity of Coloring Graphs with Forbidden Subgraphs.” In: *Journal of Graph Theory* (2016), n/a–n/a. URL: `http://dx.doi.org/10.1002/jgt.22028` (cit. on p. 101).

[79] P. A. Golovach, D. Paulusma, and J. Song. “Closing complexity gaps for coloring problems on H-free graphs.” In: *Information and Computation* 237 (2014), pp. 204–214 (cit. on pp. 4, 86, 100).

[80] F. Gurski. “A comparison of two approaches for polynomial time algorithms computing basic graph parameters.” In: *CoRR* abs/0806.4073 (2008). URL: `http://arxiv.org/abs/0806.4073` (cit. on p. 61).

[81] A. Gyárfás. “Problems from the world surrounding perfect graphs.” In: *Zastosowania Matematyki* 19(3-4) (1987), pp. 413–441 (cit. on pp. 3, 46, 49, 61, 71).

[82] A. Gyárfás, Z. Li, R. Machado, A. Sebő, S. Thomassé, and N. Trotignon. “Complements of nearly perfect graphs.” In: *Journal of Combinatorics* 4(3) (2013), pp. 299–310 (cit. on p. 49).

[83] A. Gyárfás, A. Sebő, and N. Trotignon. “The chromatic gap and its extremes.” In: *Journal of Combinatorial Theory, Series B* 102(5) (2012), pp. 1155–1178 (cit. on pp. 50, 56).

[84] J. Harant, M. A. Henning, D. Rautenbach, and I. Schiermeyer. “The independence number in graphs of maximum degree three.” In: *Discrete Mathematics* 308(23) (2008), pp. 5829–5833 (cit. on p. 51).

[85] F. Harary and C. Holzmann. “Line graphs of bipartite graphs.” In: *Revista de la Sociedad Matematica de Chile* 1 (1974), pp. 19–22 (cit. on p. 24).

[86] F. Harary and C. St. J. A. Nash-Williams. “On eulerian and hamiltonian graphs and line graphs.” In: *Canad. Math. Bull.* 8 (1965), pp. 701–709 (cit. on pp. 21, 37).

[87] D. Haussler and E. Welzl. “ε-Nets and Simplex Range Queries.” In: *Discrete & Computational Geometry* 2 (1987), pp. 127–151 (cit. on p. 84).

[88] P. Haxell, A. Kostochka, and S. Thomassé. “Packing and Covering Triangles in K_4-free Planar Graphs.” In: *Graphs and Combinatorics* 28(5) (2012), pp. 653–662 (cit. on pp. 65, 68).

[89] P. E. Haxell. “Packing and covering triangles in graphs.” In: *Discrete Mathematics* 195(1-3) (1999), pp. 251–254 (cit. on p. 63).

[90] P. E. Haxell and A. D. Scott. “On Lower Bounds for the Matching Number of Subcubic Graphs.” In: *Journal of Graph Theory* (2016), n/a–n/a. URL: `http://dx.doi.org/10.1002/jgt.22063` (cit. on p. 26).

[91] P. E. Haxell, A. V. Kostochka, and S. Thomassé. “A stability theorem on fractional covering of triangles by edges.” In: *Eur. J. Comb.* 33(5) (2012), pp. 799–806 (cit. on pp. 48, 64).

[92] C. C. Heckman. “On the tightness of the $\frac{5}{14}$ independence ratio.” In: *Discrete Mathematics* 308(15) (2008), pp. 3169–3179 (cit. on pp. 25, 57).

[93] C. C. Heckman and R. Thomas. “Independent sets in triangle-free cubic planar graphs.” In: *Journal of Combinatorial Theory, Series B* 96(2) (2006), pp. 253–275 (cit. on pp. 1, 20).

[94] J. R. Henderson. “Permutation decomposition of $(0,1)$-matrices and decomposition transversals.” PhD thesis. California Institute of Technology, 1971 (cit. on p. 47).

[95] M. A. Henning, C. Löwenstein, and D. Rautenbach. “Independent sets and matchings in subcubic graphs.” In: *Discrete Mathematics* 312(11) (2012), pp. 1900–1910 (cit. on pp. 2, 3, 21, 26, 45, 46, 49).

[96] C. T. Hoàng. “On the structure of (banner, odd hole)-free graphs.” In: *CoRR* abs/1510.02324 (2015). URL: `http://arxiv.org/abs/1510.02324` (cit. on p. 71).

[97] Y. Huang and Y. Liu. “Maximum genus and maximum nonseparating independent set of a 3-regular graph.” In: *Discrete Mathematics* 176(1–3) (1997), pp. 149–158 (cit. on p. 79).

[98] S. Iwata. Personal communication. 2016 (cit. on p. 119).

[99] S. Iwata. "A weighted linear matroid parity algorithm." In: *Proceedings of the 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications*. 2013, pp. 251–259 (cit. on pp. 118, 119).

[100] B. Jackson and K. Yoshimoto. "Spanning Even Subgraphs of 3-Edge-Connected Graphs." In: *Journal of Graph Theory* 62(1) (2009), pp. 37–47 (cit. on p. 75).

[101] F. Joos. "Independence and matching number in graphs with maximum degree 4." In: *Discrete Mathematics* 323 (2014), pp. 1–6 (cit. on pp. 3, 46, 50, 55, 70).

[102] M. Kamiński. "MAX-CUT and containment relations in graphs." In: *Theoretical Computer Science* 438 (2012), pp. 89–95 (cit. on pp. 4, 100).

[103] M. Kamiński, V. V. Lozin, and M. Milanič. "Recent developments on graphs of bounded clique-width." In: *Discrete Applied Mathematics* 157(12) (2009), pp. 2747–2761 (cit. on p. 16).

[104] L. Kang, D. Wang, and E. Shan. "Independent sets in $\{$claw, $K_4\}$-free 4-regular graphs." In: *Discrete Mathematics* 332 (2014), pp. 40–44 (cit. on pp. 1, 20, 25, 26).

[105] L. Kang, D. Wang, and E. Shan. "The independence number of connected $($claw, $K_4)$-free 4-regular graphs." In: *Taiwanese Journal of Mathematics* 17(1) (2013), pp. 275–285 (cit. on p. 26).

[106] L. M. Kirousis, M. J. Serna, and P. G. Spirakis. "Parallel Complexity of the Connected Subgraph Problem." In: *SIAM J. Comput.* 22(3) (1993), pp. 573–586 (cit. on p. 94).

[107] T. Kloks, C. M. Lee, and J. Liu. "New Algorithms for k-Face Cover, k-Feedback Vertex Set, and k-Disjoint Cycles on Plane and Planar Graphs." In: *Graph-Theoretic Concepts in Computer Science*. Ed. by Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, and Luděk Kučera. Vol. 2573. Lecture Notes in Computer Science. 2002, pp. 282–295 (cit. on pp. 3, 48, 74).

[108] D. V. Korobitsin. "On the complexity of domination number determination in monogenic classes of graphs." In: *Discrete Mathematics and Applications* 2(2) (1992), pp. 191–200 (cit. on pp. 4, 86, 100).

[109] N. Korpelainen. "Boundary Properties of Graphs." PhD thesis. University of Warwick, 2012 (cit. on p. 108).

[110] N. Korpelainen, V. V. Lozin, D. S. Malyshev, and A. Tiskin. "Boundary properties of graphs for algorithmic graph problems." In: *Theoretical Computer Science* 412(29) (2011), pp. 3545–3554 (cit. on pp. 5, 20, 105–108, 110, 111, 113–115).

[111] N. Korpelainen, V. V. Lozin, and C. Mayhill. "Split Permutation Graphs." In: *Graphs and Combinatorics* 30(3) (2014), pp. 633–646 (cit. on p. 96).

[112] B. Korte and J. Vygen. *Combinatorial Optimization*. 4th. Algorithms and Combinatorics. Springer, 2007 (cit. on p. 14).

[113] M. Kotrbčík. "A note on disjoint cycles." In: *Information Processing Letters* 112(4) (2012), pp. 135–137 (cit. on p. 81).

[114] D. Král', J. Kratochvíl, Z. Tuza, and G. J. Woeginger. "Complexity of Coloring Graphs without Forbidden Induced Subgraphs." In: *Graph-Theoretic Concepts in Computer Science*. Ed. by A. Brandstädt and V. B. Le. Vol. 2204. Lecture Notes in Computer Science. 2001, pp. 254–262 (cit. on pp. 4, 21, 86, 100).

[115] E. Kranakis, D. Krizanc, B. Ruf, J. Urrutia, and G. J. Woeginger. “The VC-dimension of set systems defined by graphs.” In: *Discrete Applied Mathematics* 77(3) (1997), pp. 237–257 (cit. on pp. 4, 83–86, 93, 101, 143, 144).

[116] J. Krausz. “Démonstration nouvelle d’un théorème de Whitney sur les réseaux.” In: *Mat. Fiz. Lapok* 50 (1943), pp. 75–85 (cit. on p. 19).

[117] M. Krivelevich. “On a conjecture of Tuza about packing and covering of triangles.” In: *Discrete Mathematics* 142(1-3) (1995), pp. 281–286 (cit. on pp. 48, 64).

[118] A. Labarre. *Comment on “Complexity of finding 2 vertex-disjoint $(|V|/2)$-cycles in cubic graphs?”* 2011. URL: `http://cstheory.stackexchange.com/questions/6107/complexity-of-finding-2-vertex-disjoint-v-2-cycles-in-cubic-graphs` (cit. on pp. 32, 33).

[119] T. H. Lai and S. S. Wei. “The Edge Hamiltonian Path problem is NP-complete for bipartite graphs.” In: *Information Processing Letters* 46(1) (1993), pp. 21–26 (cit. on pp. 2, 21, 39).

[120] S. A. Lakshmanan, C. Bujtás, and Z. Tuza. “Generalized line graphs: Cartesian products and complexity of recognition.” In: *Electr. J. Comb.* 22(3) (2015), P3.33 (cit. on pp. 22, 40, 43).

[121] S. A. Lakshmanan, C. Bujtás, and Z. Tuza. “Induced cycles in triangle graphs.” In: *Discrete Applied Mathematics* 209 (2016), pp. 264–275 (cit. on p. 42).

[122] S. A. Lakshmanan, C. Bujtás, and Z. Tuza. “Small Edge Sets Meeting all Triangles of a Graph.” In: *Graphs and Combinatorics* 28(3) (2012), pp. 381–392 (cit. on pp. 40, 48, 64, 66, 67).

[123] R. Laskar and J. Pfaff. *Domination and irredundance in split graphs*. Tech. rep. 430. Dept. Mathematical Sciences, Clemson Univ., 1983 (cit. on p. 103).

[124] V. B. Le. “Gallai graphs and anti-Gallai graphs.” In: *Discrete Mathematics* 159(1–3) (1996), pp. 179–189 (cit. on p. 22).

[125] V. B. Le and E. Prisner. “Iterated k-Line Graphs.” In: *Graphs and Combinatorics* 10(2) (1994), pp. 193–203 (cit. on p. 41).

[126] P. G. H. Lehot. “An Optimal Algorithm to Detect a Line Graph and Output Its Root Graph.” In: *J. ACM* 21(4) (1974), pp. 569–575 (cit. on p. 19).

[127] C. H. Liu. Personal communication. 2016 (cit. on p. 80).

[128] J. Liu and C. Zhao. “A new bound on the feedback vertex sets in cubic graphs.” In: *Discrete Mathematics* 148(1) (1996), pp. 119–131 (cit. on p. 77).

[129] S. C. Locke and F. Lou. “Finding Independent Sets in K_4-Free 4-Regular Connected Graphs.” In: *Journal of Combinatorial Theory, Series B* 71(1) (1997), pp. 85–110 (cit. on pp. 1, 20, 25, 56).

[130] D. Lokshtanov, M. Pilipczuk, and E. J. van Leeuwen. “Independence and Efficient Domination on P_6-free Graphs.” In: *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA ’16. 2016, pp. 1784–1803 (cit. on p. 108).

[131] L. Lovász. “Matroid Matching and Some Applications.” In: *Journal of Combinatorial Theory, Series B* 28(2) (1980), pp. 208–236 (cit. on p. 118).

[132] L. Lovász and M. D. Plummer. *Matching Theory*. North-Holland, 1986 (cit. on pp. 14, 96, 118).

[133] V. V. Lozin and D. Rautenbach. “The tree- and clique-width of bipartite graphs in special classes.” In: *The Australasian Journal of Combinatorics* 34 (2006), pp. 57–67 (cit. on p. 97).

[134] J. Ma, X. Yu, and W. Zang. “Approximate min-max relations on plane graphs.” In: *Journal of Combinatorial Optimization* 26(1) (2013), pp. 127–134 (cit. on pp. 48, 74, 79).

[135] D. S. Malyshev. “A complexity dichotomy and a new boundary class for the dominating set problem.” In: *Journal of Combinatorial Optimization* 32(1) (2016), pp. 226–243 (cit. on pp. 5, 105, 108).

[136] D. S. Malyshev. “Classes of graphs critical for the edge list-ranking problem.” In: *Journal of Applied and Industrial Mathematics* 8(2) (2014), pp. 245–255 (cit. on pp. 5, 108).

[137] D. S. Malyshev. “Continued sets of boundary classes of graphs for colorability problems.” In: *Diskretn. Anal. Issled. Oper.* 16(5) (2009). (In Russian), pp. 41–51 (cit. on p. 108).

[138] D. S. Malyshev and P. M. Pardalos. “Critical hereditary graph classes: a survey.” In: *Optimization Letters* (2015), pp. 1–20 (cit. on p. 106).

[139] G. J. Minty. “On Maximal Independent Sets of Vertices in Claw-Free Graphs.” In: *Journal of Combinatorial Theory, Series B* 28(3) (1980), pp. 284–304 (cit. on pp. 26, 126).

[140] B. Mohar and C. Thomassen. *Graphs on Surfaces*. The Johns Hopkins University Press, 2001 (cit. on pp. 11, 78).

[141] H. Müller. “Hamiltonian circuits in chordal bipartite graphs.” In: *Discrete Mathematics* 156(1) (1996), pp. 291–298 (cit. on p. 107).

[142] A. Munaro. “Boundary classes for graph problems involving non-local properties.” In: (). Submitted (cit. on p. 5).

[143] A. Munaro. “Bounded clique cover of some sparse graphs.” In: *Discrete Mathematics* (). To appear (cit. on p. 5).

[144] A. Munaro. “On line graphs of subcubic triangle-free graphs.” In: *Discrete Mathematics* (). To appear (cit. on p. 5).

[145] A. Munaro. “The VC-dimension of graphs with respect to k-connected subgraphs.” In: *Discrete Applied Mathematics* 211 (2016), pp. 163–174 (cit. on p. 5).

[146] L. Nebeský. “A new characterization of the maximum genus of a graph.” In: *Czechoslovak Mathematical Journal* 31(4) (1981), pp. 604–613 (cit. on p. 79).

[147] L. Nebeský. “Characterizing the maximum genus of a connected graph.” In: *Czechoslovak Mathematical Journal* 43(1) (1993), pp. 177–185 (cit. on pp. 79, 81).

[148] E. A. Nordhaus, B. M. Stewart, and A. T. White. “On the Maximum Genus of a Graph.” In: *Journal of Combinatorial Theory, Series B* 11(3) (1971), pp. 258–267 (cit. on p. 79).

[149] S. O and D. B. West. "Balloons, Cut-Edges, Matchings, and Total Domination in Regular Graphs of Odd Degree." In: *Journal of Graph Theory* 64(2) (2010), pp. 116–131 (cit. on p. 26).

[150] S. Olariu. "Paw-free graphs." In: *Information Processing Letters* 28(1) (1988), pp. 53–54 (cit. on p. 57).

[151] J. B. Orlin. "A Fast, Simpler Algorithm for the Matroid Parity Problem." In: *Integer Programming and Combinatorial Optimization: 13th International Conference, IPCO 2008*. Ed. by A. Lodi, A. Panconesi, and G. Rinaldi. 2008, pp. 240–258 (cit. on p. 118).

[152] S. Oum and P. Seymour. "Approximating clique-width and branch-width." In: *Journal of Combinatorial Theory, Series B* 96(4) (2006), pp. 514–528 (cit. on p. 16).

[153] J. Oxley and G. Whittle. "A Characterization of Tutte Invariants of 2-Polymatroids." In: *Journal of Combinatorial Theory, Series B* 59(2) (1993), pp. 210–244 (cit. on p. 119).

[154] G. Pap. "Weighted linear matroid matching." In: *Proceedings of the 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications*. 2013, pp. 411–413 (cit. on pp. 118, 119).

[155] C. H. Papadimitriou and M. Yannakakis. "On Limited Nondeterminism and the Complexity of the V-C Dimension." In: *Journal of Computer and System Sciences* 53(2) (1996), pp. 161–170 (cit. on pp. 4, 85).

[156] D. Peterson. "Gridline graphs: a review in two dimensions and an extension to higher dimensions." In: *Discrete Applied Mathematics* 126(2–3) (2003), pp. 223–239 (cit. on p. 23).

[157] J. Pfaff, R. Laskar, and S. T. Hedetniemi. *NP-completeness of total and connected domination, and irredundance for bipartite graphs*. Tech. rep. 428. Dept. Mathematical Sciences, Clemson Univ., 1983 (cit. on p. 103).

[158] M. D. Plummer. "Claw-free maximal planar graphs." In: *Congressus Numerantium* 72 (1990), pp. 9–23 (cit. on p. 74).

[159] G. J. Puleo. "Tuza's Conjecture for graphs with maximum average degree less than 7." In: *European Journal of Combinatorics* 49 (2015), pp. 134–152 (cit. on pp. 48, 64, 66, 67).

[160] S. P. Radziszowski. "Small Ramsey Numbers." In: *Electronic Journal of Combinatorics* DS1 (2014) (cit. on p. 57).

[161] B. Randerath and I. Schiermeyer. "Vertex Colouring and Forbidden Subgraphs – A Survey." In: *Graphs and Combinatorics* 20(1) (2004), pp. 1–40 (cit. on p. 46).

[162] N. Robertson and P. D. Seymour. "Graph Minors. II. Algorithmic Aspects of Tree-width." In: *Journal of Algorithms* 7(3) (1986), pp. 309–322 (cit. on p. 15).

[163] N. Robertson and P. D. Seymour. "Graph Minors. V. Excluding a Planar Graph." In: *Journal of Combinatorial Theory, Series B* 41(1) (1986), pp. 92–114 (cit. on p. 106).

[164] N. Robertson, D. Sanders, P. Seymour, and R. Thomas. "The Four-Colour Theorem." In: *Journal of Combinatorial Theory, Series B* 70(1) (1997), pp. 2–44 (cit. on p. 80).

[165] N. D. Roussopoulos. "A max $\{m, n\}$ algorithm for determining the graph H from its line graph G." In: *Information Processing Letters* 2(4) (1973), pp. 108–112 (cit. on p. 19).

[166] N. Sauer. "On the density of families of sets." In: *Journal of Combinatorial Theory, Series A* 13(1) (1972), pp. 145–147 (cit. on p. 83).

[167] N. Sbihi. "Algorithme de recherche d'un stable de cardinalité maximum dans un graphe sans étoile." In: *Discrete Mathematics* 29(1) (1980), pp. 53–76 (cit. on pp. 26, 126).

[168] M. Schaefer. "Deciding the Vapnik–Červonenkis Dimension is Σ_3^p-complete." In: *Journal of Computer and System Sciences* 58(1) (1999), pp. 177–182 (cit. on p. 85).

[169] M. Schaefer. *Deciding the VC-dimension is Σ_3^p-complete, II*. Tech. rep. TR00-006. DePaul University, 2000 (cit. on p. 86).

[170] M. Schaefer and C. Umans. "Completeness in the Polynomial-Time Hierarchy - A Compendium." In: *SIGACT News* 33(3) (2002). URL: `http://ovid.cs.depaul.edu/documents/phcom.pdf` (cit. on p. 86).

[171] A. Scott and P. Seymour. "Induced subgraphs of graphs with large chromatic number. I. Odd holes." In: *Journal of Combinatorial Theory, Series B* (2015) (cit. on p. 71).

[172] A. Sebő. Personal communication. 2014 (cit. on p. 63).

[173] S. Shelah. "A combinatorial problem; stability and order for models and theories in infinitary languages." In: *Pac. J. Math* 41 (1972), pp. 247–261 (cit. on p. 83).

[174] E. Speckenmeyer. "On feedback vertex sets and nonseparating independent sets in cubic graphs." In: *Journal of Graph Theory* 12(3) (1988), pp. 405–412 (cit. on pp. 2, 21, 77).

[175] E. Speckenmeyer. "Untersuchungen zum Feedback Vertex Set Problem in ungerichteten Graphen." PhD thesis. Paderborn, 1983 (cit. on pp. 2, 21, 31, 117).

[176] W. Staton. "Some Ramsey-type numbers and the independence ratio." In: *Trans. Amer. Math. Soc.* 256 (1979), pp. 353–370 (cit. on pp. 1, 20, 25, 57, 75).

[177] M. Stehlík. "Critical graphs with connected complements." In: *Journal of Combinatorial Theory, Series B* 89(2) (2003), pp. 189–194 (cit. on p. 50).

[178] Z. Tuza. "A Conjecture on Triangles of Graphs." In: *Graphs and Combinatorics* 6(4) (1990), pp. 373–380 (cit. on pp. 3, 47, 48, 63).

[179] R. Uehara. NP-*complete problems on a* 3-*connected cubic planar graph and their applications*. Tech. rep. 4. Tokyo Woman's Christian University, 1996 (cit. on pp. 60, 61).

[180] S. Ueno, Y. Kajitani, and S. Gotoh. "On the nonseparating independent set problem and feedback set problem for graphs with no vertex degree exceeding three." In: *Discrete Mathematics* 72(1) (1988), pp. 355–360 (cit. on pp. 2, 21, 31, 77, 78, 117–119).

[181] V. N. Vapnik and A. Y. Chervonenkis. "On the uniform convergence of relative frequencies of events to their probabilities." In: *Theory of Probability and Its Applications* 16(2) (1971), pp. 264–280 (cit. on pp. 4, 83).

[182] D. B. West. *Introduction to Graph Theory*. 2nd ed. Prentice Hall, 2001 (cit. on pp. 7, 26, 46, 63).

[183] H. Whitney. "Congruent Graphs and the Connectivity of Graphs." In: *American Journal of Mathematics* 54(1) (1932), pp. 150–168 (cit. on pp. 1, 19).

[184] N. H. Xuong. "How to Determine the Maximum Genus of a Graph." In: *Journal of Combinatorial Theory, Series B* 26(2) (1979), pp. 217–225 (cit. on p. 79).

[185] N. H. Xuong. "Upper-embeddable Graphs and Related Topics." In: *Journal of Combinatorial Theory, Series B* 26(2) (1979), pp. 226–232 (cit. on p. 80).

Index Of Notions Not Defined In Chapter 1

Résumé — Dans cette thèse, nous considérons plusieurs paramètres des hypergraphes et nous étudions si les restrictions aux sous-classes des hypergraphes permettent d'obtenir des propriétés combinatoires et algorithmiques souhaitables. La plupart des paramètres que nous prenons en compte sont des instances spéciales des *packings* et *transversals* des hypergraphes.

Dans la première partie, nous allons nous concentrer sur les *line graphs* des graphes subcubiques sans triangle et nous allons démontrer que pour tous ces graphes il y a un *independent set* de taille au moins $\frac{3}{10}|V(G)|$ et cette borne est optimale. Conséquence immédiate: nous obtenons une borne inférieure optimale pour la taille d'un couplage maximum dans les graphes subcubiques sans triangle. De plus, nous montrons plusieurs résultats algorithmiques liés au FEEDBACK VERTEX SET, HAMILTONIAN CYCLE et HAMILTONIAN PATH quand restreints aux *line graphs* des graphes subcubiques sans triangle.

Puis nous examinons trois hypergraphes ayant la propriété d'Erdős-Pósa et nous cherchons à déterminer les fonctions limites optimales. Tout d'abord, nous apportons une fonction *θ-bounding* pour la classe des graphes subcubiques et nous étudions CLIQUE COVER: en répondant à une question de Cerioli et al. [31], nous montrons qu'il admet un PTAS pour les graphes planaires. Par la suite, nous nous intéressons à la Conjecture de Tuza et nous montrons que la constante 2 peut être améliorée pour certains graphes sans K_4 et avec arêtes contenues dans au maximum quatre triangles et pour les graphes sans certains *odd-wheels*. Enfin, nous nous concentrons sur la Conjecture de Jones: nous la démontrons dans le cas des graphes sans griffes avec degré maximal 4 et nous faisons quelques observations dans le cas des graphes subcubiques.

Nous étudions ensuite la *VC-dimension* de certains hypergraphes résultants des graphes. En particulier, nous considérons l'hypergraphe sur l'ensemble des sommets d'un certain graphe qui est induit par la famille de ses sous-graphes k-connexes. En généralisant les résultats de Kranakis et al. [115], nous fournissons des bornes supérieures et inférieures optimales pour la *VC-dimension* et nous montrons que son calcul est NP-complet, pour chacun $k \geq 1$. Enfin, nous démontrons que ce problème (dans le cas $k = 1$) et le problème étroitement lié CONNECTED DOMINATING SET sont soit solvables en temps polynomial ou NP-complet, quand restreints aux classes de graphes obtenues en interdisant un seul sous-graphe induit.

Dans la partie finale de cette thèse, nous nous attaquons aux meta-questions suivantes: Quand est-ce qu'un certain problème "difficile" de graphe devient "facile"?; Existe-t-il des frontières séparant des instances "faciles" et "difficiles"? Afin de répondre à ces questions, dans le cas des classes héréditaires, Alekseev [5] a introduit la notion de *boundary class* pour un problème NP-difficile et a montré qu'un problème Π est NP-difficile pour une classe héréditaire X finiment défini si et seulement si X contient un *boundary class* pour Π. Nous continuons la recherche des *boundary classes* pour les problèmes suivants: HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE, HAMILTONIAN PATH, FEEDBACK VERTEX SET, CONNECTED DOMINATING SET and CONNECTED VERTEX COVER.

Mots clés : Line graphs, paramètres des hypergraphes, θ-bounding functions, VC-dimension, NP-complétude, boundary classes

Abstract — In this thesis, we consider several hypergraph parameters and study whether restrictions to subclasses of hypergraphs allow to obtain desirable combinatorial or algorithmic properties. Most of the parameters we consider are special instances of packings and transversals of hypergraphs.

In the first part, we focus on line graphs of subcubic triangle-free graphs and show that any such graph G has an independent set of size at least $\frac{3}{10}|V(G)|$, the bound being sharp. As an immediate consequence, we obtain a tight lower bound for the matching number of subcubic triangle-free graphs. Moreover, we prove several algorithmic results related to FEEDBACK VERTEX SET, HAMILTONIAN CYCLE and HAMILTONIAN PATH when restricted to line graphs of subcubic triangle-free graphs.

Then we consider three hypergraphs having the Erdős-Pósa Property and we seek to determine the optimal bounding functions. First, we provide an optimal θ-bounding function for the class of subcubic graphs and we study CLIQUE COVER: answering a question by Cerioli et al. [31], we show it admits a PTAS for planar graphs. Then we focus on Tuza's Conjecture and show that the constant 2 in the statement can be improved for some K_4-free graphs whose edges are contained in at most four triangles and graphs obtained by forbidding certain odd-wheels. Finally, we concentrate on Jones' Conjecture: we prove it in the case of claw-free graphs with maximum degree at most 4 and we make some observations in the case of subcubic graphs.

Then we study the VC-dimension of certain set systems arising from graphs. In particular, we consider the set system on the vertex set of some graph which is induced by the family of its k-connected subgraphs. Generalizing results by Kranakis et al. [115], we provide tight upper and lower bounds for the VC-dimension and we show that its computation is NP-complete, for each $k \geq 1$. Finally, we show that this problem (in the case $k = 1$) and the closely related CONNECTED DOMINATING SET are either NP-complete or polynomial-time solvable when restricted to classes of graphs obtained by forbidding a single induced subgraph.

In the final part of the thesis, we consider the following meta-questions: when does a certain "hard" graph problem become "easy"?; Is there any "boundary" separating "easy" and "hard" instances? In order to answer these questions in the case of hereditary classes, Alekseev [5] introduced the notion of a boundary class for an NP-hard problem and showed that a problem Π is NP-hard for a finitely defined (hereditary) class X if and only if X contains a boundary class for Π. We continue the search of boundary classes for the following problems: HAMILTONIAN CYCLE THROUGH SPECIFIED EDGE, HAMILTONIAN PATH, FEEDBACK VERTEX SET, CONNECTED DOMINATING SET and CONNECTED VERTEX COVER.

Keywords: Line graphs, hypergraph parameters, θ-bounding functions, VC-dimension, NP-completeness, boundary classes

Univ. Grenoble Alpes, Laboratoire G-SCOP