TECHNISCHE UNIVERSITÄT ILMENAU

FAKULTÄT FÜR MATHEMATIK UND NATURWISSENSCHAFTEN

INSTITUT FÜR MATHEMATIK

On Cycles and Independence in Graphs

Dissertation

zur Erlangung des akademischen Grades
Dr. rer. nat.

eingereicht von

Dipl.-Math. Friedrich Regen

Institut für Mathematik
TU Ilmenau

Juli 2010

Betreuender Hochschullehrer: Univ.-Prof. Dr. rer. nat. habil. Dieter Rautenbach
(Technische Universität Ilmenau)

urn:nbn:de:gbv:ilm1-2010000384

Meinen lieben Eltern

Contents

Zusammenfassung

Die vorliegende Arbeit behandelt Fragestellungen im Zusammenhang mit Kreisen und unabhängigen Mengen in Graphen.

Kapitel 2 handelt von unabhängigen Kreisen: Die Parameter ν_e bzw. ν_v geben die maximale Größe kanten- bzw. eckendisjunkter Kreispackungen an, d.h. die maximale Anzahl von Kreisen in einem Graph, die paarweise keine gemeinsame Kante bzw. Ecke haben. Da die Berechnung dieser Parameter bekanntermaßen schon für subkubische Graphen schwer ist, geht es im ersten Abschnitt um die Komplexität eines einfacheren Problems, des Packens von Kreisen einer festen Länge ℓ in Graphen mit Maximalgrad Δ. Für $\ell = 3$ und beliebiges Δ wurde diese Komplexität bereits von Caprara and Rizzi in [12] bestimmt, und wir verallgemeinern ihre Ergebnisse auf alle größeren Kreislängen ℓ. Im zweiten Abschnitt von Kapitel 2 untersuchen wir die Struktur von Graphen, für die $\mu(G) - \nu_e(G)$ bzw. $\mu(G) - \nu_v(G)$ einen vorgegebenen Wert haben. Die 2-zusammenhängenden derartigen Graphen können erzeugt werden, indem eine einfache Erweiterungsregel auf eine endliche Menge von Graphen angewandt wird. Aus diesem Strukturergebnis können wir folgern, daß die Parameter $\nu_e(G)$ und $\nu_v(G)$ „fixed parameter tractable“ bezüglich ihrer Differenz zur zyklomatischen Zahl sind.

In Kapitel 3 bestimmen wir die Größenordnung der minimalen Anzahl von Kreislängen in einem Hamiltongraph mit q Sehnen. Wir geben eine Familie von Beispielen an, in denen nur $\sqrt{q+1}$ Kreislängen auftreten, zeigen aber, daß jeder Hamiltongraph mit q Sehnen mindestens $\sqrt{\frac{4}{7}q}$ Kreislängen enthält. Der Beweis beruht auf einem Lemma von Faudree et al. in [23], demzufolge der Graph, der aus einem Weg mit Endecken x und y und q gleichlangen Sehnen besteht, x-y-Wege von mindestens $\frac{q}{3}$ verschiedenen Längen enthält. Im ersten Abschnitt korrigieren wir den ursprünglich fehlerhaften Beweis und leiten zusätzliche Schranken her. Im zweiten Abschnitt folgern wir daraus die Unterschranke für die Anzahl der Kreislängen.

Im letzten Kapitel untersuchen wir Unterschranken für den Unabhängigkeitsquotienten, d.h. den Bruch $\frac{\alpha(G)}{n(G)}$, für Graphen gegebener Dichte. Wir stellen fest, daß bestmögliche Schranken für die Klasse aller Graphen sowie für große zusammenhängende Graphen bereits bekannt sind. Deshalb verändern wir die Fragestellung, indem wir Graphenklassen betrachten, die durch das Verbot kleiner ungerader Kreise definiert sind. Das Hauptergebnis des ersten Abschnitts ist eine Verallgemeinerung eines Ergebnisses von Heckman und Thomas, das die bestmögliche Schranke für zusammenhängende dreiecksfreie Graphen mit Durchschnittsgrad bis zu $\frac{10}{3}$ impliziert und die extremalen Graphen charakterisiert. Der Rest der ersten beiden Abschnitte enthält Vermutungen ähnlichen Typs für zusammenhängende dreiecksfreie Graphen mit Durchschnittsgrad im Intervall $\left[\frac{10}{3}, \frac{54}{13}\right]$ und für zusammenhängende Graphen mit ungerader Taillenweite 7 mit Durchschnittsgrad bis zu $\frac{14}{5}$. Der letzte Abschnitt enthält analoge Beobachtungen zum Bipartitionsquotienten. Möglicherweise lassen sich viele Unterschranken für den Unabhängigkeitsquotienten auf den Bipartitionsquotienten übertragen, indem man sie einfach verdoppelt. Diese neuen Schranken sind stärker, und die zugehörigen Klassen extremaler Graphen oft viel reichhaltiger. Am Ende dieser Arbeit diskutieren wir Vermutungen dieser Art.

Dank

Meinem Doktorvater Professor Dieter Rautenbach möchte ich herzlich dafür danken, daß er mir ermöglicht hat, in die graphentheoretische Forschung einzusteigen und diese Arbeit zu schreiben. Er hat die Arbeit mit großem Einsatz betreut, und ich habe ihn nicht nur als Mathematiker, sondern auch als Menschen sehr zu schätzen gelernt. Auch die Zusammenarbeit mit Dr. Stephan Brandt, Dr. Christian Löwenstein, Janina Müttel und Anders Sune Pedersen hat mir viel Freude gemacht und war für mich eine wertvolle Erfahrung.

Die Herzlichkeit der Ilmenauer Graphentheoriegruppe und der Kammerchor der TU Ilmenau trugen wesentlich dazu bei, daß mir mein Aufenthalt in Ilmenau in guter Erinnerung bleiben wird. Meinen Eltern und Geschwistern bin ich für ihre Unterstützung und viele gute Gespräche sehr dankbar.

1 Introduction

1.1 Summary

This thesis discusses several problems related to cycles and the independence number in graphs.

In Chapter 2, we discuss independent sets of cycles. The parameters ν_e resp. ν_v denote the maximum cardinality of edge-disjoint resp. vertex-disjoint cycle packings, i.e. the maximum number of cycles in a graph that can be arranged such that no two of them share an edge resp. a vertex. Since the computation of these parameters is known to be hard even for subcubic graphs, the first section discusses the complexity of a simpler problem, packing cycles of fixed length ℓ in graphs of maximum degree Δ. For $\ell = 3$ and arbitrary Δ, the complexity has been determined by Caprara and Rizzi in [12], and we extend their results to all greater values of ℓ. In the second section of Chapter 2, we discuss the structure of graphs for which $\mu(G) - \nu_e(G)$ resp. $\mu(G) - \nu_v(G)$ equals some given integer. The 2-connected graphs of this kind can be obtained by a simple extension rules applied to a finite set of graphs, which yields a fixed-parameter-tractability result for $\nu_e(G)$ and $\nu_v(G)$.

In Chapter 3, we approximate the minimum number of cycle lengths in a Hamiltonian graph with q chords. We give a family of examples that contain only $\sqrt{q+1}$ cycle lengths, but show that $\sqrt{\frac{4}{7}q}$ cycle lengths can be guaranteed. The proof relies on a lemma by Faudree et al. in [23], which states that the graph that contains a path with endvertices x and y and q chords of equal length contains paths between x and y of at least $\frac{q}{3}$ different lengths. In the first section, we correct the originally faulty proof and derive additional bounds. The second section we use these bounds to derive the lower bound on the size of the cycle spectrum.

In the last chapter, we study lower bounds on the independence ratio, i.e. the fraction $\frac{\alpha(G)}{n(G)}$, for graphs of given density. We observe that best-possible bounds are already known both for arbitrary graphs and for large connected graphs. Therefore, we modify the question by considering classes of graphs defined by forbidding small odd cycles as subgraphs. The main result of the first section is a generalisation of a result of Heckman and Thomas that determines the best possible lower bound for connected triangle-free graphs with average degree at most $\frac{10}{3}$ and characterises the extremal graphs. The rest of the first two sections contains conjectures with similar statements on connected triangle-free graphs of average degree in $\left[\frac{10}{3}, \frac{54}{13}\right]$ and on connected graphs of odd girth 7 with average degree up to $\frac{14}{5}$. The last section collects analogous observations for the bipartite ratio. It seems possible to translate many lower bounds on the independence ratio to bounds on the bipartite ratio by just doubling them. Those new bounds are stronger and the corresponding classes of extremal graphs usually much richer. The thesis ends with some conjectures for statements of such generalisations.

Acknowledgement

I would like to thank Prof. Dieter Rautenbach for giving me the opportunity to enter graph theoretical research and to write this dissertation. He was a most competent, friendly and committed advisor. I have also enjoyed the collaboration with Dr. Stephan Brandt, Dr. Christian Löwenstein, Janina Müttel and Anders Sune Pedersen in joint projects very much.

The Ilmenau graph theory group and the Kammerchor der TU Ilmenau made me feel at home in Ilmenau. I am very grateful to my family for their support and inspiration.

1.2 Notation

In this section, we briefly define the graph theoretical concepts used in this thesis. It is included merely as a reference and should be skipped both by the graph theorist and the newcomer, who will find well motivated and accessible introductions in the books by Diestel [18] and by Korte and Vygen [41]. The former covers more areas of “purely mathematical” interest while the latter emphasises algorithmic questions.

1.2.1 Graph Theory

$V(G)$, $E(G)$

$n(G)$, $m(G)$

A *graph* G is a pair $\big(V(G),\ E(G)\big)$, where $V(G)$ is an arbitrary set called the *vertex set* of G, and $E(G) \subseteq \big\{\{v,w\}\colon v,w \in V(G), v \neq w\big\}$ is called the *edge set* of G. We may also refer to an edge $\{v,w\}$ by the shorthand notations vw or wv. Throughout this thesis, we only consider *finite* graphs, i.e. graphs whose vertex set is finite. The elements of $V(G)$ are called the *vertices*, the elements of $E(G)$ the *edges* of G. The cardinality of $V(G)$ is the *order* $n(G) := |V(G)|$ of G, the cardinality of $E(G)$ the *size* $m(G) := |E(G)|$ of G.

$N_G(v)$, $N_G[v]$

$d_G(v)$

$\delta(G)$, $\Delta(G)$, $d(G)$

Neighbourhoods and degrees A vertex v is *incident* with an edge e, if $v \in e$. The (open) *neighbourhood* of a vertex v is defined by $N_G(v) = \{w \in V(G)\colon \{v,w\} \in E(G)\}$. w is called adjacent to v, if $w \in N_G(v)$. The closed neighbourhood of v is defined by $N_G[v] := N_G(v) \cup \{v\}$. The *degree* of v is $d_G(v) := |N_G(v)|$. The *minimum degree* of G is $\delta(G) := \min\{d_G(v) \mid v \in V(G)\}$, the *maximum degree* of G is $\Delta(G) := \max\{d_G(v) \mid v \in V(G)\}$, and the *average degree* of G is $d(G) := \frac{2m(G)}{n(G)}$, since each edge contributes to the degree of two vertices. Vertices of degree 0 are called *isolated*. A graph all of whose vertices have degree r is called *r-regular*. In particular, 3-regular graphs are called *cubic*, 4-regular graphs are called *quartic*, and a graph is called *subcubic*, if its maximum degree is at most 3. Two edges in G are called *adjacent*, if they share a common vertex.

Operations on graphs A *subgraph* of G is a graph G' with $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$. For each subset $T \subseteq V(G)$, we define $G[T]$ to be the subgraph with vertex set T and the maximal edge set, i.e. $E(G[T]) = \big\{\{v,w\} \in E(G)\colon v,w \in T\big\}$. Those subgraphs are called *induced subgraphs*, while subgraphs that contain all vertices of G are called *spanning*. The edge sets of subgraphs of maximum degree at most one are called *matchings*. For $X \subseteq V(G)$, let $G - X$ be the induced subgraph $G[V(G) \setminus X]$, and for $v \in V(G)$ we define $G - v := G - \{v\}$. For $Y \subseteq E(G)$, let $G - Y$ be the spanning subgraph of G with edge set $E(G) \setminus Y$, and for $e \in E(G)$ we define $G - e := G - \{e\}$. To avoid an ambiguity, the expression $G - \{u, v\}$ with $u, v \in V(G)$ always denotes a deletion of two vertices instead of one edge, since the edge deletion can be expressed by $G - uv$.

$G - \cdots$ $G - \{u, v\}$

The *complement* $\overline{G}$ is given by $V(\overline{G}) = V(G)$, and $E(\overline{G}) = \binom{V(G)}{2} \setminus E(G)$. For a set $Y \subseteq E(\overline{G})$, let $G + Y$ be the graph with vertex set $V(G)$ and edge set $E(G) \cup Y$, and for a single edge $e \in E(\overline{G})$, we define $G + e := G + \{e\}$. If G_1 and G_2 are two graphs, then the *cartesian product* $G_1 \square G_2$ of G_1 and G_2 is the graph G with vertex set $V(G_1) \times V(G_2)$ and edge set

$\overline{G}$ $G + \cdots$ $G_1 \square G_2$

$$E(G) = \Big\{ \{(a, b), (c, d)\} \subseteq V(G) \colon \big((a = c) \wedge (b, d) \in G_2\big) \vee \big((a, c) \in G_1 \wedge (b = d)\big) \Big\} .$$

If X is a nonempty set of vertices in a graph G, then the *vertex identification* with respect to X yields the graph G' with vertex set $V(G') = (V(G) \setminus X) \cup \{\xi\}$ and edge set

$$E(G') = E(G - X) \cup \{\xi w \colon \exists x \in X, w \notin X \colon xw \in E(G)\}.$$

If $X = \{u, v\}$ induces a connected subgraph in G, then the vertex identification of X is called a *contraction*.

Morphisms A *homomorphism* φ from $G \to G'$ between two graphs is a map $\varphi \colon V(G) \to V(G')$ with $\forall e \in E(G) \colon \varphi(e) \in E(G')$. An *isomorphism* φ between G and G' is a bijective map $\varphi \colon V(G) \to V(G')$ such that φ and φ^{-1} are homomorphisms.

In many situations, it is common to talk about specific graphs instead of isomorphism classes: For example, the statement "The graph G does not contain $K_{3,3}$ as a subgraph" usually only means that G contains no subgraph which is *isomorphic* to $K_{3,3}$. Similarly, lists of graphs with special properties should be understood as lists as isomorphism classes of graphs. We adopt this simplified notation although it is imprecise.

caveat

Special graphs The following special graph classes are used throughout the following text.

For a set $L = \{l_1, \ldots, l_k\} \subseteq \mathbb{Z}/n\mathbb{Z}$, we define the *circulant graph* $Ci_n[l_1, \ldots, l_k]$ to be the graph with vertex set $\{v_i\}_{i \in \mathbb{Z}/n\mathbb{Z}}$ and edge set $\big\{v_i v_j \colon i - j \in L\}\big\}$. Graphs that are isomorphic to $K_n := Ci_n[1, \ldots, n]$ are called *complete* graphs on n vertices. Graphs that are isomorphic to $C_n := Ci_n[1]$ are called *cycles*, if $n \geq 3$. Graphs that are isomorphic to P_n with $P_1 = K_1$, $P_2 = K_2$ and $P_n = C_n - v_1 v_n$ for $n \geq 3$ are called *paths* with *endvertices* v_1 and v_n. An expression of the form "$P = x_1 x_2 \ldots x_n$" defines a path P on the vertex set $\{x_1, \ldots, x_n\}$ in which two vertices are adjacent if and only if they are consecutive entries of the sequence. Similarly, a cycle can be given by a sequence $x_1 x_2 \ldots x_n x_1$ of vertices. The *length* of a path or a cycle is its size. A cycle of length k is called a *k-cycle*.

$Ci_n[\ldots]$ K_n, C_n, P_n

$\gamma(G)$

Parameters A vertex set $X \subseteq V(G)$ is a *dominating set* of a graph G, if $V(G) = \bigcup_{v \in X} N_G[v]$. The minimum cardinality of a dominating set is the *dominating number* $\gamma(G)$ of G.

$\alpha(G)$, $\omega(G)$, $\chi(G)$

A vertex set is called an *independent set* of a graph G, if it does not contain a pair of adjacent vertices. It is called a *clique*, if its vertices are pairwise adjacent. The maximum cardinality of an independent set is the *independence number* $\alpha(G)$. The *clique number* $\omega(G) = \alpha(\overline{G})$ denotes the maximum cardinality of a clique. The *chromatic number* $\chi(G)$ denotes the minimum value of k such that G admits a homomorphism $G \to K_k$. Since the fibres of homomorphisms are independent sets, this means that $\chi(G)$ is the minimum cardinality of a partition of $V(G)$ into independent sets. Obviously, $\chi(G) \geq \omega(G)$ for every graph G. Graphs are called *perfect*, if all of their induced subgraphs G' satisfy $\chi(G') = \omega(G')$.

$\nu_v(G)$, $\nu_e(G)$

Cycle subgraphs A set of cycles in a graph no two of which contain a common vertex is called a *vertex-disjoint cycle packing*. A set of cycles in a graph no two of which share an edge is called an *edge-disjoint cycle packing*. The parameters $\nu_v(G)$ and $\nu_e(G)$ denote the largest cardinalities of vertex-disjoint resp. edge-disjoint cycle packings of G.

A cycle or a path in G of order $n(G)$ is called *Hamiltonian*. A graph is called *Hamiltonian*, if it contains a Hamiltonian cycle. The *cycle spectrum* of a graph is the set of the lengths of all cycles. If a graph G has the maximum possible cycle spectrum $\{3, \dots, n(G)\}$, then G is called *pancyclic*.

$g(G)$

The *girth* $g(G)$ of a graph is the length of a shortest cycle, if G contains a cycle, and ∞ if no such cycle exists. The *odd girth* $g_{\text{odd}}(G)$ of a graph G is the minimum length of a cycle of odd length in G, and ∞ G contains no odd cycle.

Definitions by forbidden cycles A *forest* is a graph that does not contain a cycle. A *tree* is a connected forest. Vertices of degree one in a forest are called *leaves*. A *star* is a tree in which at most one vertex is no leaf.

$K_{r,s}$

A *cactus* is a graph for which every edge is contained in at most one cycle. A *bipartite graph* is a graph that does not contain a cycle of odd length. For each bipartite graph G, there exist disjoint independent sets A and B with $V(G) = A \cup B$; this decomposition is called a *bipartition* of G. For $r, s \in \mathbb{N}$, the *complete bipartite* $K_{r,s}$ is the graph whose vertex set is the disjoint union of two A, B with $|A| = r$, $|B| = s$ and which contains all edges with exactly one endvertex in A.

$\mathrm{dist}_G(v, w)$, G^k

Connectivity In a graph G, the *distance* $\mathrm{dist}_G(v, w)$ between two vertices is defined to be the length of the shortest path containing both vertices and ∞ if no such path exists. The maximum distance between two vertices of G is the *diameter* $\mathrm{diam}(G) := \sup_{v,w \in V(G)} \mathrm{dist}_G(v, w)$. The *k-th power* of a graph G is the graph G^k with $V(G^k) = V(G)$ and $E(G^k) = \left\{ \{v, w\} \in \binom{V(G)}{2} : \mathrm{dist}_G(v, w) \leq k \right\}$.

$\kappa(G), \mu(G)$

A graph is *connected*, if $\mathrm{diam}(G) < \infty$. The equivalence classes of the equivalence relation $\sim \subseteq V(G) \times V(G)$ with $v \sim w \;:\Leftrightarrow\; \mathrm{dist}_G(v, w) < \infty$ are called the *(connected) components* of G. The number of components of a graph G is denoted by $\kappa(G)$. The *cyclomatic number* $\mu(G)$ of a graph G with $\kappa(G)$ components is given by $\mu(G) = m(G) - n(G) + \kappa(G)$ and counts, for any spanning forest of G, the edges not contained in this forest.

For $k \in \mathbb{N}$, a graph is called *k-connected*, if for every $X \subseteq V(G)$ with $|X| < k$, $G - X$ is connected, and it is called *k-edge-connected*, if for every $X \subseteq E(G)$ with $|X| < k$, $G - E$ is connected. A *cutvertex* of a connected graph G is a vertex whose removal disconnects G. A *block* of a graph is a maximal subgraph without a cutvertex, i.e. a maximal 2-connected subgraph. An *endblock* of a connected graph G is a block that contains at most one cutvertex of G. A *bridge* is an edge whose removal increases the number of components.

The *cut* induced by some vertex set $X \subseteq V(G)$ is the set of edges for which exactly one endpoint is contained in X.

Ear decompositions A *proper ear* of G is a path in G of length at least 1 such that all its internal vertices have degree 2 in G. An ear of G is *maximal*, if it is not properly contained in another ear of G. If P is an ear of G and I is the set of internal vertices of P, then we say that G arises from $G' := G - I$ by *adding the ear P* and that G' arises from G by *removing the ear P*. Whitney [66, 79] has proved that a graph of order at least 3 is 2-connected if and only if it has an *proper ear decomposition*, i.e. it arises from a cycle by iteratively adding ears.

Multigraphs

In Section 2.2, we use multigraphs instead of graphs. In multigraphs, multiple edges between vertices are allowed, and an edge may connect a vertex with itself. Since we allow multiple edges, the edge set cannot be defined as a subset of $\binom{V(G)}{2}$. Instead, a finite *multigraph* $(V(G), E(G), \varphi_G)$ is a triple of a finite vertex set $V(G)$, an arbitrary edge set $E(G)$, and an *incidence map* $\varphi_G\colon E(G) \to \{\{v,w\} \subseteq V(G)\}$ that assigns to each edge its endvertices. An edge with only one endvertex would be called a *loop*, but in this thesis we shall only consider loopless multigraphs. Any edges that are incident to the same set of endvertices are called *parallel*. Again, an edge with endvertices v and w is denoted by vw, although this may not be a unique description, if $\{v, w\}$ has more than one preimage under φ_G. $\varphi(G)$

An edge e is *incident* with a vertex v, if $v \in \varphi_G(e)$. The *degree* of a vertex of a loopless multigraph is the number of incident edges. A *cycle* in G is a connected 2-regular subgraph of G. In particular, multigraphs may contain cycles of length 2.

1.2.2 Complexity Theory

When evaluating the efficiency of algorithms, we use standard terminology of Complexity Theory as introduced e.g. in chapters 15 and 16 of [41]. Although most fundamental concepts such as polynomial time solvability or NP-completeness do not rely on specific definitions of machine models and graph representations, the claim that an algorithm can be implemented in linear time does. For such claims, we assume that the algorithms run on a random access machine (RAM) and that the input graph is given by an adjacency matrix. This implies in particular that we can verify or modify membership of a constant in a set of edges or vertices in constant time. These assumptions allow to estimate the asymptotic running time of the implementation of a typical contemporary computer, provided that the input size does not exceed some constant fraction of the address space.

2 Cycle Packings

The problems to determine the cycle packing numbers ν_v and ν_e are algorithmically hard: Extending a method of Chuzhoy and Khanna in [15], Friggstad and Salavatipour showed in [26] that, even if restricted to graphs of maximum degree at most 3, both parameters are hard to approximate within ratio $O(\sqrt[2+\varepsilon]{\log n})$ for any $\varepsilon > 0$ provided that $NP \not\subseteq ZPTIME(n^{polylog(n)})$. On the other hand, Krivelevich et al. have shown in [42] that a slightly enhanced greedy algorithm approximates ν_e with approximation ratio $O(\sqrt{\log n})$. For the parameter ν_v, the best approximation algorithm known so far is due to Salivatipour and Verstraëte [62] and has approximation ratio $O(\log n)$.

In this chapter, we study two simplifications of the general cycle packing problem.

The first section deals with packing only cycles of a given length $\ell \in \mathbb{N}$. For graphs of a fixed maximum degree Δ, the maximum cardinality of ℓ-cycle packings can be approximated in linear time: The cycles of length ℓ can be enumerated in $O(n(G) \cdot \Delta^\ell)$, and at least $\frac{1}{\ell\Delta^\ell}$ of them can be included in an ℓ-cycle packing because each cycle shares a vertex with less than $\ell\Delta^\ell$ other cycles.

In [12], Caprara and Rizzi have considered the case $\ell = 3$, i.e. the problem of packing edge-disjoint and vertex-disjoint triangles. For both problems, they determined the weakest condition on the maximum degree of a graph that allows to determine the maximum number of disjoint 3-cycles in polynomial time. Furthermore, they proved that for all weaker conditions this optimisation problem is *APX-hard*, i.e. there does not exist an approximation algorithm with arbitrary good approximation factor unless $P = NP$. The main result of this section is a generalisation of this result to cycles of arbitrary, but still fixed length. APX-hard

In the second section, we characterise the structure of graphs in which one of the packing numbers differs from the cyclomatic number only by a given constant k. Since the edge-disjoint cycle packing number of a graph is the sum of the cycle packing numbers of its blocks, it suffices to consider two-connected graphs. We show that all blocks with $\mu(G) - \nu_e(G) = k$ can be constructed from a finite list of graphs by a simple easily reversible operation in which edges are replaced by what we call *cycle paths*. This implies that for given k, there is a linear-time algorithm which decides if $\mu(G) - \nu_e(G) = k$ for a given graph G.

We obtain a similar result for vertex-disjoint cycle packings. Although the vertex-disjoint cycle packing number of a graph does not need to be the sum of the cycle packing numbers of its blocks, we can still give a linear time algorithm for diciding whether $\mu(G) - \nu_v(G) = k$.

2.1 Cycles of a given length

$\nu_{v,\ell}(G)$, $\nu_{e,\ell}(G)$

Definition 1. *For a graph G and an integer $\ell \in \mathbb{N}$, the parameter $\nu_{v,\ell}(G)$ resp. $\nu_{e,\ell}(G)$ is the cardinality of a largest set of vertex-disjoint resp. edge-disjoint cycles of length ℓ. We denote the optimisation problems to determine $\nu_{v,\ell}$ and $\nu_{e,\ell}$ by ℓ-VCP and ℓ-ECP.*

Note that for a graph G with $\Delta(G) \leq 3$, a pair of cycles that shares a vertex also shares an edge, so subcubic graphs satisfy $\nu_e(G) = \nu_v(G)$ and $\nu_{e,\ell}(G) = \nu_{v,\ell}(G)$ for every $\ell \geq 3$.

The case $\ell = 3$ has been thoroughly studied in the literature: In [27], Garey and Johnson show that the problem to find the maximum number of vertex-disjoint triangles in a graph is NP-hard. In [36], Holyer proves that the problem to find the maximum number of edge-disjoint triangles in a graph is NP-hard. Both sources actually consider arbitrarily large cliques instead of triangles. Finally, in [12], Caprara and Rizzi study the approximability of both triangle packing problems. On one hand, they give a polynomial time algorithm for the restriction of 3-ECP to graphs with maximum degree 4 and a polynomial time algorithm for the restriction of 3-VCP to graphs with maximum degree 3. On the other hand, they show that both problems are APX-hard under all weaker maximum degree assumptions. In this section, we generalise their results to packing cycles of arbitrary, but fixed length.

Theorem 1 (Caprara and Rizzi [12])**.** *The restrictions of 3-ECP to graphs with maximum degree 5 and of 3-VCP to graphs with maximum degree 4 are APX-hard.*

For given $k \in \mathbb{N}$, this result immediately implies that the restriction of $3k$-ECP to graphs with maximum degree 5 and of $3k$-VCP to graphs with maximum degree 4 are APX-hard: This follows by considering the classes of graphs which arise from graphs of maximum degree 5 resp. 4 by subdividing[1] each edge $k-1$ times.

In this section, we are going to show that the restriction of ℓ-ECP to subcubic graphs is APX-hard for any $\ell \geq 6$. Due to the maximum degree condition, the same result applies to ℓ-VCP. For $\ell \in \{4, 5\}$, we give polynomial time algorithms for the restriction of ℓ-ECP to subcubic graphs — and thus for the restriction of ℓ-VCP to subcubic graphs — but we show that both ℓ-ECP and ℓ-VCP are APX-hard under all weaker maximum degree restrictions.

The results of this section are based on [59].

2.1.1 Exact Algorithms

As in [12], we assign to each graph G two *auxiliary graphs* which reflect the intersection relation of its cycles.

$EC(G, \ell)$, $VC(G, \ell)$

Definition 2. *For a graph G and a positive integer ℓ, let $\mathcal{C}(G, \ell)$ be the set of cycles of length ℓ in G. We define*

[1]Subdividing an edge uv x times means removing this edge, adding x new vertices $w_1, w_2, \ldots, w_x$ and adding the edges of the path $uw_1w_2 \ldots w_xv$.

$$EC(G,\ \ell) := \Big(\mathcal{C}(G,\ell),\ \big\{\{C,D\} \subseteq \mathcal{C}(G,\ell)\colon C \neq D,\ E(C)\cap E(D) \neq \emptyset\big\}\Big)$$
$$\textit{and}\quad VC(G,\ell) := \Big(\mathcal{C}(G,\ell),\ \big\{\{C,D\} \subseteq \mathcal{C}(G,\ell)\colon C \neq D,\ V(C)\cap V(D) \neq \emptyset\big\}\Big).$$

Clearly, $\nu_{e,\ell}(G) = \alpha(EC(G,\ \ell))$ and $\nu_{v,\ell}(G) = \alpha(VC(G,\ell))$. Since all pairs of cycles of length g can be trivially enumerated in $O(n^{2\ell})$, the auxiliary graphs can be constructed in polynomial time for fixed ℓ, so the problems ℓ-ECP and ℓ-VCP can be polynomially reduced to determining the independence number of auxiliary graphs. For instances with a maximum degree restriction, the auxiliary graphs can even be constructed in linear time because all pairs of intersecting cycles that contain a specified vertex can be found in constant time. Note that the restrictions of the general cycle packing problem do not allow this reduction, since e.g. any cubic subgraph of $Ci_{4k}[1,2]$ contains at least 2^k cycles of length $3k$.

Theorem 2. *The restrictions of* 4*-ECP and* 5*-ECP to subcubic graphs can be solved in polynomial time.*

Proof. It is sufficient to show that $EC(G,\ \ell)$ is claw-free because Minty [50] and Sbihi [63] have designed polynomial time algorithms for determining the independence number of a *claw-free* graph, i.e. of a graph that does not contain the star $K_{1,3}$ as an induced subgraph. Indeed, if $EC(G,\ \ell)$ vertex v with three neighbours v_1, v_2 and v_3, then G contains a cycle C of length at most five that intersects three other cycles C_1, C_2 and C_3. Each of these cycles uses two of the at most five edges of the cut induced by $V(C)$, so they cannot all be edge-disjoint, i.e. $\{v, v_1, v_2, v_3\}$ does not induce a claw. □

Note that these problems coincide with the restrictions of 4-VCP and 5-VCP to subcubic graphs.

The same argument shows that auxiliary graphs of graphs with girth 4 are *quasi-line graphs*, i.e. the neighbourhood of each of their vertices can be partitioned into two cliques. Indeed, only few auxiliary graphs of girth 4 do not allow trivial reductions, so this case can be solved by a simple algorithm that does not rely on the deep results by Minty and Sbihi.

Input: A subcubic graph G
Output: $\nu_{e,4}(G)$

1 **begin**
2     $H \longleftarrow EC(G,\ 4)$;
3     $k \longleftarrow 0$;
4   **while** *H contains a vertex v of degree at most 1* **do**
5         $k \longleftarrow k + 1$;
6         $H \longleftarrow H - N_H[v]$;
7         $G \longleftarrow G - E(v)$;
8   **end**
9   Let $V_1, \ldots, V_t$ be the connected components of H;
10   **for** $i \longleftarrow 1$ **to** t **do**
11   **if** $\Delta(H[V_i]) = 2$ **then**
12             $k \longleftarrow k + \left\lfloor \frac{|V_i|}{2} \right\rfloor$;
13   **else**
14             $k \longleftarrow k + \left\lfloor \frac{\left|\bigcup_{C \in V_i} V(C)\right|}{4} \right\rfloor$;
15   **end**
16   **end**
17   **return** k;
18 **end**

Algorithm 1: SubcubicFourCyclePacking

Theorem 3. *Algorithm 1 solves the restriction of* 4*-ECP to subcubic graphs in linear time.*

Proof. The construction of $EC(G,\ 4)$ in line 2 can be performed in linear time because the maximum degree of G is bounded. The running time for the reduction step (line 4) is linearly bounded in the number of deleted vertices of H and edges of H. Note that $E(v)$ is the edge set of the cycle $v \in V(H)$. Since the size of $EC(G,\ 4)$ is linear in the size of G by the above argument, this step can also be performed in linear time. Counting and labelling of the connected components in line 9 can be implemented in $O(|E(H)|+|V(H)|)$ by repeatedly performing breadth-first-search at an unlabelled vertex. It remains to prove that $k = \nu_{e,4}(G)$ at the end of the algorithm.

In each reduction step in line 4, G and H are modified such that $\alpha(H)$ drops by one, k increases by one and the property $H = EC(G,\ 4)$ continues to hold. Therefore, it suffices to show that in each step of the **for**-loop, k is increased by $\alpha(H[V_i])$. If $\Delta(H[V_i]) = 2$, this holds because $H[V_i]$ is a cycle. Otherwise, $H[V_i]$ contains a vertex of degree at least 3.

Let $G' := G\left[\bigcup_{C \in V_i} V(C)\right]$. Since $H[V_i]$ is connected, so is G'. We may assume that every edge of is contained in a 4-cycle, since a removal of edges in G' which are not contained in a 4-cycle does not affect the auxiliary graph of G' and the final *for*-loop only depends on the auxiliary graph $H[V_i]$ of G'. Now it suffices to show that G' is one of the graphs in Figure 2.1, since each of those contains exactly $\left\lfloor \frac{|V(G')|}{4} \right\rfloor$ edge-disjoint 4-cycles.

Case 1: G' contains a triangle T.

If G' does not contain the diamond $K_4 - e$ as a subgraph, then each edge of T lies in a 4-cycle with two vertices outside T. Since every vertex of T has at most one neighbour outside of T, the union of these three 4-cycles is the *prism graph* of the triangle, i.e. the cubic graph that can be constructed by two disjoint copies of a triangle by connecting the three pairs of corresponding vertices. Since G' is subcubic and connected, it does not contain any further vertices. If G' contains the diamond $D = K_4 - e$ as a subgraph, let v and w be the vertices of degree 3 in the diamond. Since the edge $\{v, w\}$ is contained in a 4-cycle, the two other vertices in D are connected with each other, so $G' = K_4$. Note that in both subcases, $H[V_i]$ is a triangle, so $\Delta(H[V_i]) = 2$.

Case 2: G' contains no triangle, but a $K_{2,3}$ subgraph induced by the union of the independent sets $\{v_1, v_2\}$ and $\{w_1, w_2, w_3\}$.

Since this subgraph contains only three 4-cycles, we may assume that w_1 has another neighbour x. In order for the edge w_1x to be contained in a 4-cycle, x must be adjacent to one of the vertices w_2 and w_3, w.l.o.g. it is adjacent to w_2. Therefore, G' contains $K_{3,3} - e$ as a subgraph. Since the vertices x and w_3 have distance 3 in the $K_{3,3} - e$ subgraph and every edge of G is contained in a 4-cycle, any further edge that contains one of them must contain them both. Therefore, in this subcase either $G' = K_{3,3} - e$ or $G' = K_{3,3}$.

Case 3: G' contains neither a triangle nor a $K_{2,3}$ subgraph.

In this case, each pair of different 4-cycles in G' shares at most one edge. Since $H[V_i]$ is claw-free but contains a vertex of degree at least 3, G' contains three 4-cycles each pair of which shares exactly an edge. These three 4-cycles form a subgraph that consists of an induced 6-cycle $v_1w_1v_2w_2v_3w_3v_1$ and a vertex y with $N_{G'}(y) = \{v_1, v_2, v_3\}$. As in the above case, we may assume that G' contains a path w_1xw_2 for a new vertex x, and either G' contains no further edge or another edge xw_3. In the first case, $G' = K_2 \square P_4 - e$, and in the second case, $G' = K_2 \square C_4$. $\square$

The proof implies that nonempty subcubic graphs which allow no trivial reductions (i.e. they are connected, each edge is contained in a 4-cycle, and each 4-cycle intersects at least two other 4-cycles) either belong to one of the graphs in Figure 2.1 or realise a cycle graph, in which case they belong to one of the following two families:

- Möbius ladders $M_k = Ci_{2k}[1, k]$ for $k \geq 2$ $\quad (M_2 = K_4,\ EC(M_k,\ 4) \cong C_k)$
- cycle prisms $K_2 \square C_k$ for $k \geq 3$ $\quad (EC(C_k \square P_2,\ 4) \cong C_k$ for $k \neq 4)$

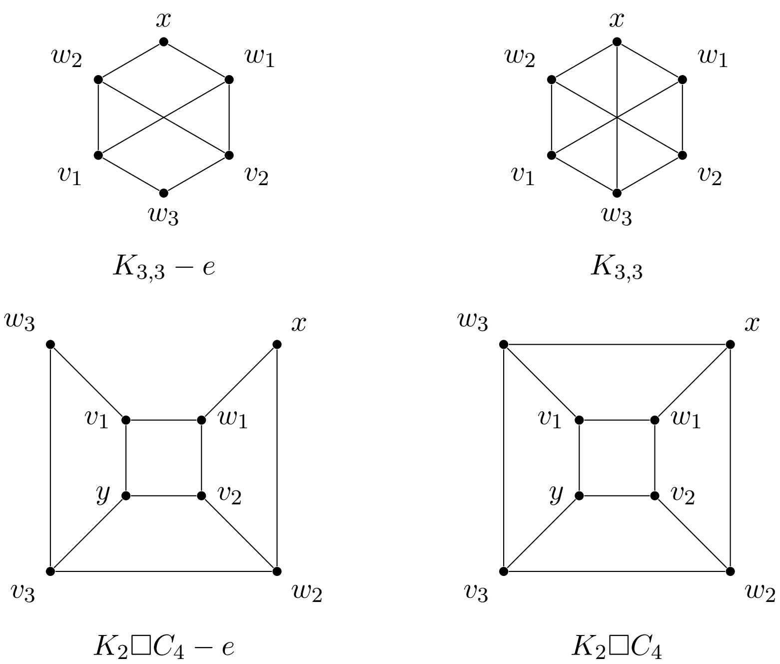


Figure 2.1: Minimal subcubic realisations of connected auxiliary graphs H with $\delta(H) \geq 2$ and $\Delta(H) > 2$

2.1.2 Hardness of Approximation

In order to show the hardness of packing ℓ-cycles in graphs of a given maximum degree, we are going to use hardness results for finding the independence number in a graph class with the property that each member of the class is isomorphic to some graph $EC(G,\ \ell)$ resp. $VC(G,\ \ell)$.

A MAX-SAT instance consists of a set X of some s Boolean variables $x_1, \ldots, x_s$ and of a set $\mathcal{Z}$ of some t clauses, which are subsets of the set L of literals, where L is the disjoint union of X and the set $\{\overline{x_1}, \ldots, \overline{x_s}\}$ of negations of the variables. We say that a truth assignment $X \to \{true, false\}$ satisfies a clause C, if C contains a Boolean variable set to true or the negation of a Boolean variable set to false. The maximum satisfiability problem asks for the maximum number of clauses that can be satisfied by a truth assignment.

Our proofs of the hardness of ℓ-cycle packings rely on a result of Berman and Karpinski on the 3-OCC-MAX 2SAT problem. This problem is the restriction of the maximum satisfiability problem to instances for which each clause contains at most two variables and each variable x occurs in at most three clauses, i.e. at most three clauses contain one of the literals x and $\overline{x}$.

Theorem 4 (Berman and Karpinski [7])**.** *For every $\varepsilon > 0$, it is NP-hard to approximate 3-OCC-MAX 2SAT within a factor of $\frac{2012}{2011} - \varepsilon$.*

Definition 3. *We call a 3-OCC-MAX 2SAT instance* reduced*, if for any two different literals $l_1, l_2 \in L$,* reduced

1. *none of its clauses is of the form $\{l_1, \overline{l_1}\}$,*
2. *at least two clauses contain one of the literals l_1 and $\overline{l_1}$,*
3. *the instance does not contain both the clauses $\{l_1, l_2\}$ and $\{\overline{l_1}, \overline{l_2}\}$,*
4. *the instance does not contain two clauses $\{l_1, \overline{l_2}\}$ and $\{l_1, l_2\}$ and a third clause that contains the literal $\overline{l_1}$.*

Lemma 1. *For every $\varepsilon > 0$, it is NP-hard to approximate the restriction of 3-OCC-MAX 2SAT to reduced instances within a factor of $\frac{2012}{2011} - \varepsilon$.*

Proof. By Theorem 4, it is sufficient to show that for every unreduced instance I of the 3-OCC-MAX 2SAT problem, we can compute an integer d and a smaller instance I' with $OPT(I) = OPT(I') + d$ in polynomial time.

If the first condition on reduced instances is violated, we construct I' from I by removing the clause $\{l_1, \overline{l_1}\}$ and setting $d := 1$.

If the second condition is violated, let C be a clause such that neither l_1 nor $\overline{l_1}$ occurs outside of C. Then we construct I' from I by removing the clause C and set $d := 1$.

If the third condition is violated, we have two clauses $\{l_1, l_2\}$ and $\{\overline{l_1}, \overline{l_2}\}$. Let x_1 and x_2 be the two variables corresponding to the literals l_1 and l_2. If there exists a partial truth assignment $\{x_1, x_2\} \to \{true, false\}$ that satisfies all clauses in which the literals

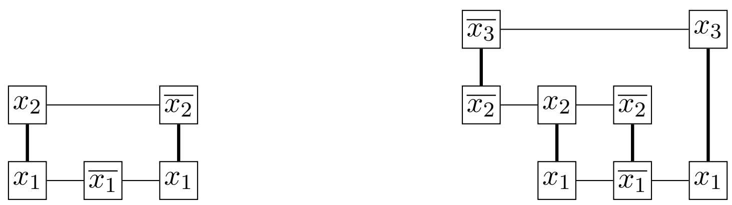
x_1, $\overline{x_1}$, x_2 and $\overline{x_2}$ occur, we set d to be the number of these clauses and construct I' from I by removing them. Otherwise, there are w.l.o.g. two clauses $\{l_1, l_3\}$ and $\{l_2, l_4\}$ for literals l_3, l_4 corresponding to two further variables. Any truth assignment that assigns the same value to l_1 and l_2 satisfies *at most* three of the four clauses in which l_1 and l_2 occur, while a truth assignment that assigns different values to these literals satisfies *at least* three of the four clauses. Therefore, we can set $d := 2$ and construct I' from I by removing the clauses $\{l_1, l_2\}$ and $\{\overline{l_1}, \overline{l_2}\}$ and replacing the literal l_1 by the literal $\overline{l_2}$.

If the fourth condition is violated, at most one of the literals l_2 and $\overline{l_2}$ occurs in a fourth clause; w.l.o.g. there is no further occurrence of the literal $\overline{l_2}$. Then any optimal truth setting remains optimal after the value of the variable corresponding to l_2 is adjusted such that l_2 is *true*. Therefore, we can set d to be the number of occurrences of the literal l_2 in I and construct I' from I by removing the clauses in which l_2 occurs and replacing the clause $\{l_1, \overline{l_2}\}$ by $\{l_1\}$. □

We can associate graphs to MAX-SAT instances via a construction from Karp's proof of the NP-completeness of STABLE SET [39].

SAT graph

Definition 4. *For a given MAX-SAT instance the vertices of the* SAT graph *correspond to the pairs* $(l, C) \in L \times \mathcal{Z}$ *with* $l \in C$. *Its edge set is a union* $E_C \cup E_V$ *of the set* E_C *of* clause edges *between each pair of vertices* (l_1, C) *and* (l_2, C) *that belong to the same clause, and of the set* E_V *of* variable edges *between each pair* (x, C_1) *and* $(\overline{x}, C_2)$ *of vertices.*



$(x_1 \vee x_2), (\overline{x_1}), (x_1 \vee \overline{x_2})$ $(\overline{x_2} \vee \overline{x_3}), (x_1 \vee x_2), (\overline{x_1} \vee \overline{x_2}), (x_1 \vee x_3)$

Figure 2.2: SAT graphs to two *unreduced* 3-OCC MAX 2-SAT instances

Obviously, the solution of the MAX-SAT problem corresponds the size of a maximum independent set in its SAT graph. For reduced 3-OCC MAX 2-SAT instances the SAT graphs have some properties summarised in the following lemma.

Lemma 2. *The SAT graph corresponding to a reduced 3-OCC-MAX 2SAT instance is a simple graph* G *with* $\delta(G) \leq 2 \leq 3 \leq \Delta(G)$ *and* $g(G) \geq 6$, *whose vertices of degree 3 induce a graph of maximum degree* 1.

Proof. Let H be the SAT graph for a reduced 3-OCC-MAX 2SAT instance. Parallel edges in SAT graphs arise only if a clause contains two literals corresponding to the same variable, and this case is excluded for reduced graphs. The degree of any vertex in H cannot be larger than three because it is incident to at most one clause edge and at most two variable edges. As each variable occurs in at least two clauses, there is no vertex of degree one.

Let us assume that H contains a 3-cycle T. Since the clause edges are a matching, T contains at least two variable edges. Therefore, all three vertices in T correspond to the same variable, so none of the three edges is a clause edge. This is impossible because (V, E_V) is a disjoint union of paths of length one and two.

Let us assume that H contains a 4-cycle Q. Since the clause edges are a matching, and every path with three edges in H contains at least one clause edge, clause edges and variable edges alternate on Q. Therefore, the vertices of C correspond to two clauses of the form $\{l_1, l_2\}$ and $\{\overline{l_1}, \overline{l_2}\}$, but this is the third excluded case for reduced instances.

Let us assume that H contains a 5-cycle P. As the clause edges are a matching, P contains at most two of them, so the vertices of P correspond to only two variables, and the 5-cycle corresponds to clauses $\{l_1, \overline{l_2}\}$, $\{l_1, l_2\}$ and $\{\overline{l_1}\}$, where the third clause may contain another literal. This is the fourth excluded case for reduced instances.

Finally, since the clause edges are a matching, every vertex v with $d_G(v) = 3$ is adjacent to one clause edge and two variable edges. The two vertices that are connected to v by variable edges are not incident to a clause edge, so their degree is at most two, which implies that each vertex of degree three has at most one neighbour of degree three. □

To prove the hardness results, we use a construction which, under suitable conditions, provides graphs with given girth and given auxiliary graph.

Definition 5. *For any triangle-free graph H of maximum degree Δ and any integer $\ell \geq \max\{3, \Delta\}$, we call a graph G a $C(\ell,\ H)$*-graph*, if it is obtained from the disjoint union G' of ℓ-cycles C_v for each vertex $v \in V(H)$ by the following identification process: For each $v \in V(H)$, we select $d_H(v)$ different* glueing edges *$\{e_{v,w}\}_{w \in N_H(v)} \subseteq E(C_v)$ such that for each pair $w \neq w'$ of neighbours of v, the distance between the vertex sets $e_{v,w}$ and $e_{v,w'}$ is at least $\left\lfloor \frac{\ell - d_H(v)}{d_H(v)} \right\rfloor = \left\lfloor \frac{\ell}{d_H(v)} - 1 \right\rfloor$. For each edge $\{v, w\} \in E(H)$, we define two* identification sets *$\{a_1, a_2\}$ and $\{b_1, b_2\}$ with $e_{v,w} = \{a_1, b_1\}$ and $e_{w,v} = \{a_2, b_2\}$. We obtain G from G' by identifying the vertices that are contained in a common identification set and removing double edges, i.e. one edge from each pair of glueing edges.* $C(\ell,\ H)$

For fixed ℓ, this construction can be performed in polynomial time. It does not yield a *unique* $C(\ell,\ H)$-graph, as it depends upon the choice of glueing edges and identification sets. Each $C(\ell,\ H)$-graph contains the ℓ-cycles C_v for each $v \in V(H)$ as induced subgraphs, and for each edge $\{v, w\} \in H$, it contains an edge $e_{\{v,w\}}$ representing the pair of glueing edges $e_{v,w}$ and $e_{w,v}$. Figure 2.3 shows a $C(5,\ K_4 - e)$-graph that, besides the 5-cycles C_a, C_b, C_c, and C_d, contains two 3-cycles and an additional 4-cycle, so its $EC(\cdot,\ 4)$-graph is *not* $K_4 - e$. However the following definition and lemma describe a condition on H that guarantees that all additional cycles in the constructed graph have length greater than ℓ.

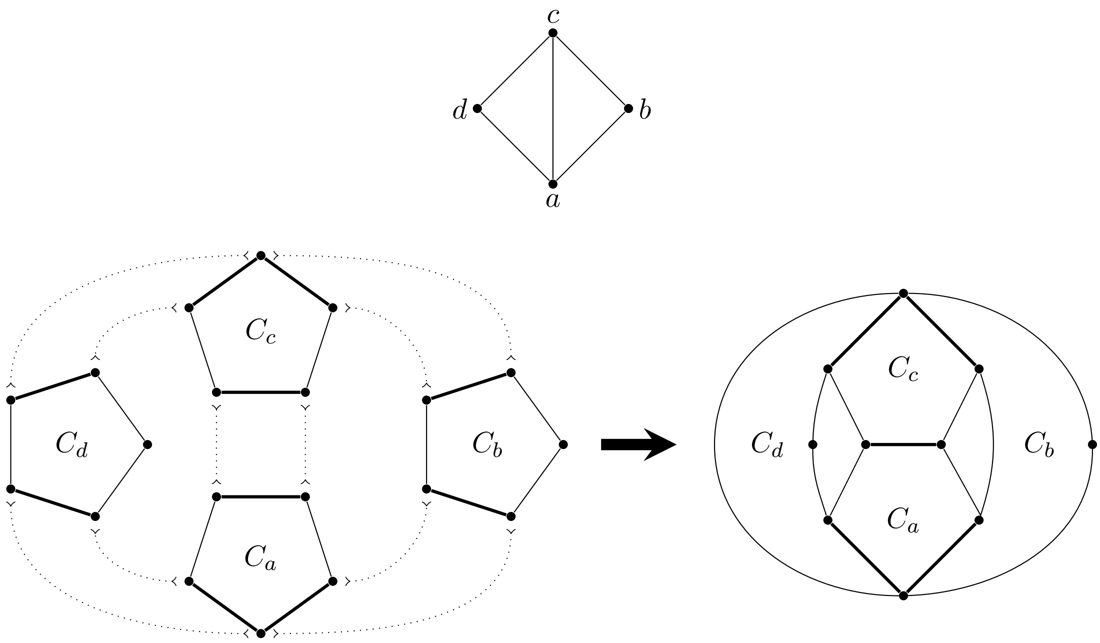


Figure 2.3: Construction of a $C(\ell,\ H)$-graph for $\ell = 5$ and $H = K_4 - e$

Definition 6. *The ℓ-**weight** of a cycle C_H in a graph H is*

$$w_\ell(C_H) := \sum_{v \in V(C_H)} \left\lfloor \frac{\ell}{d_H(v)} - 1 \right\rfloor .$$

Lemma 3. *If a graph H contains no cycle C_H of ℓ-weight $w_\ell(C_H) \leq \ell$, then the girth of any $C(\ell,\ H)$-graph G is ℓ, and $EC(C(\ell,\ H), \ell) \cong H$.*

Proof. It suffices to show that G contains no cycle $C \notin \{C_v \colon v \in H\}$ of length less than or equal to ℓ. Let us assume that C is such a cycle. Then we can split the sequence of the edges of C into a sequence $P_1, P_2, \ldots, P_l$ of consecutive paths such that for $1 \leq i \leq l$ the path P_i is entirely contained in some ℓ-cycle C_{v_i}. Furthermore, allowing paths of length 0, it is possible to choose these paths such that $\forall i \in \{1, \ldots, l\} \colon v_{i-1}v_i \in E(H)$ with $v_0 := v_l$. The length of P_i is at least $\left\lfloor \frac{\ell}{d_H(v_i)} - 1 \right\rfloor$. Since the length of C is at most ℓ, we obtain that $\ell \geq \sum_{i=1}^{l} \left\lfloor \frac{\ell}{d_H(v_i)} - 1 \right\rfloor$ and that $\forall i \in \{0, \ldots, l-2\} \colon v_i \neq v_{i+2}$. As $l > 1$, this implies that the sequence $v_0, v_1, v_2, \ldots, v_l$ contains a cycle C_H of H with $w_\ell(C_H) \leq \ell$. $\square$

Theorem 5. *For every $\varepsilon > 0$ and every $\ell \geq 6$, it is NP-hard to approximate the restriction of ℓ-ECP to graphs with maximum degree at most 3 within a factor of $\frac{2012}{2011} - \varepsilon$.*

Proof. Let H be the SAT graph for a given reduced instance of the 3-OCC-MAX 2SAT problem. By Lemma 1, the theorem can be proved by a polynomial time construction of a graph G with maximum degree 3 such that $EC(G,\ \ell) \cong H$.

We choose G as an arbitrary $C(\ell,\ H)$-graph. Since $\ell \geq 2\Delta(H)$, the glueing edges in any $C(\ell,\ H)$-graph G are disjoint, so $\Delta(G) \leq 3$. By Lemma 3, G has girth ℓ, and $EC(G,\ \ell) \cong H$, provided that the ℓ-weight of each cycle C_H in H is greater than ℓ.

As the clause edges on C_H are disjoint, each vertex of C_H is contained in a path of length one or two that contains only variable edges. Now each path of variable edges contains at most one vertex of degree three, so at least half of the vertices of C_H have degree two. Therefore, $w_\ell(C_H) = \sum_{v\in V(C_H)} \left\lfloor \frac{\ell}{d_H(v)} - 1 \right\rfloor \geq 3 \cdot \left(\lfloor \frac{\ell}{2} \rfloor + \lfloor \frac{\ell}{3} \rfloor - 2\right)$. Since $\ell \geq 6$, this implies that the ℓ-weight of C_H is strictly greater than ℓ. □

Theorem 6. *For every $\varepsilon > 0$ and each $\ell \in \{4,5\}$, it is NP-hard to approximate the restriction of ℓ-ECP to graphs with maximum degree at most 4 within a factor of $\frac{6036}{6035} - \varepsilon$.*

Proof. Let H' be the SAT graph to an arbitrary reduced instance of the 3-OCC-MAX 2SAT problem. By Lemma 2, the vertices of degree 3 induce a subgraph of H' of maximum degree 1.

In polynomial time, we can determine a set A of vertices of degree 2 of H' such that every cycle of H' contains a vertex of A and subject to this property the set A is minimal with respect to inclusion. Since A is independent, $|A| \leq \alpha(H')$.

Let H be the graph that we obtain by each vertex $v \in A$ with neighbours a and b by adding five vertices $v_1, \ldots, v_5$ of degree 2, such that $av_1v_2v_3v_4v_5b$ is a path in H. It is easy to see that $\alpha(H) = \alpha(H') + 2|A|$ and that every independent set I of H efficiently yields an independent set I' of H' with $|I'| \geq |I| - 2|A|$. This implies that every independent set I of H with $\frac{\alpha(H)}{|I|} \leq 1+\delta$ would efficiently yield an independent set I' of H' for which $\frac{\alpha(H')}{|I'|} \leq \frac{1+\delta}{1-2\delta}$.

Since each cycle in H' contains at least two vertices of degree 2, each cycle C_H in H contains at least six vertices of degree 2, so $w_\ell(C_H) \geq 6 > \ell$, and by Lemma 3, any $C(\ell,\ H)$-graph G has girth ℓ and satisfies $EC(G,\ \ell) \cong H$.

The vertices of degree 3 in H induce a subgraph of maximum degree 1 as they do in H', i.e. a collection of isolated vertices and disjoint edges. We are now going to show that this allows us to use the freedom of choosing the glueing edges and the identification sets in such a way that we obtain a $C(\ell,\ H)$-graph G of maximum degree at most 4. For vertices v of H of degree 3 all neighbours of which are of degree 2, the identification processes involving the edges of C_v cannot create vertices of degree more than 4. If vw is an edge of H between two vertices of degree 3, then we choose the glueing edge $e_{v,w} = xy$ in C_v such that y is not contained in another glueing edge of C_v and the glueing edge $e_{w,v} = x'y'$ in C_w such that x' is not contained in another glueing edge of C_w. Furthermore, for the edge vw we choose the identification sets $\{x, x'\}$ and $\{y, y'\}$. By these choices, the identification processes involving the edges of C_v and C_w do not create vertices of degree more than 4. Hence G has maximum degree at most 4.

By Lemma 1, we can finish the proof by showing that the modified graph H is still the SAT graph of a reduced instance of the 3-OCC-MAX 2SAT problem. It suffices to show that for each vertex v of degree 2 in the SAT graph of a reduced instance I of the

3-OCC-MAX 2SAT problem, we can construct an instance I' whose SAT graph is the graph obtained by replacing v with a path $v_1v_2v_3v_4v_5$ as above.

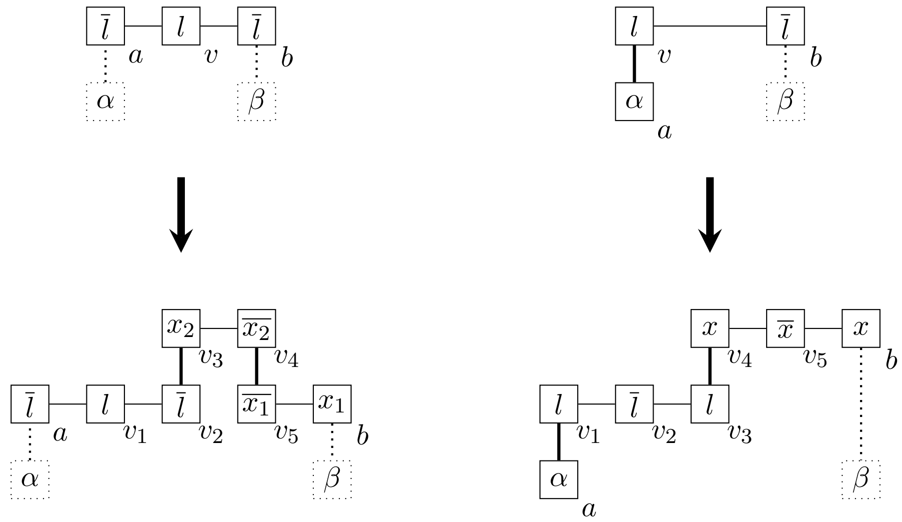


Figure 2.4: Replacement of a vertex v of degree two in Theorem 6

If v corresponds to the only literal l in a clause $\{l\}$, then there exist two further clauses C_1 and C_2 containing the literal $\bar{l}$. In this case, we construct I from I' by replacing the occurrence of $\bar{l}$ in C_2 with a new variable x_1 and adding two new clauses $\{\bar{l}, x_2\}$ and $\{\overline{x_2}, \overline{x_1}\}$, where x_2 is a second new variable. Otherwise, v corresponds to a literal l in a clause $\{l, \alpha\}$ that contains another literal α, and the literal $\bar{l}$ occurs in precisely one clause. In this case, we construct I from I' by replacing the occurrence of $\bar{l}$ by a new variable x and adding three new clauses $\{\bar{l}\}$, $\{l, x\}$, and $\{\overline{x}\}$. $\square$

Finally, we show the APX-hardness of ℓ-VCP for $\ell < 6$ by a similar construction in which each two ℓ-cycles intersect in at most one vertex.

Theorem 7. *For every $\varepsilon > 0$ and every $\ell \in \{3, 4, 5\}$, it is NP-hard to approximate the restriction of ℓ-VCP to graphs with maximum degree at most 4 within a factor of $\frac{2012}{2011} - \varepsilon$.*

Proof. Let H be the SAT graph for a given reduced instance of the 3-OCC-MAX 2SAT problem. We are going give a polynomial time construction of a graph G of girth ℓ and maximum degree at most 4 whose shortest cycles are a set $\{C_v\}_{v \in V(H)}$, such that two cycles C_v and C_w are vertex-disjoint if and only if $\{v, w\} \notin E(H)$. Since vertex-disjoint packings of shortest cycles in G correspond to stable sets in H, Lemma 1 then implies the statement.

Let G' be the disjoint union of $|H|$ ℓ-cycles C_v. Since the maximum degree of H is at most 3, we can select vertices $x_{v,w} \in C_v$ and $x_{w,v} \in C_w$ for each edge $\{v, w\} \in E(H)$ such that all $2|E(H)|$ selected vertices are pairwise different. We construct G by identifying $x_{v,w}$ with $x_{w,v}$ for each $\{v, w\} \in E(H)$. Then each vertex of G is contained in at most two ℓ-cycles C_v, so the maximum degree of G is at most 4. It remains to show that the length of any cycle $C_G \notin \{C_v\}_{v\in H}$ is greater than g.

Indeed, the edge sequence C_G can be uniquely decomposed into maximal non-empty subpaths $P_1, P_2, \ldots, P_l$, such that $l > 2$, and for each $1 \leq i \leq l$ the edges of the path P_i are contained in some cycle C_{v_i} for $v_i \in H$. Then the edges v_1v_l and v_iv_{i+1} for $i \in \{1, \ldots, l-1\}$ are contained in H, and since $\forall i \in \{1, \ldots, l-2\}\colon v_i \neq v_{i+2}$, the vertex sequence contains a cycle of H. Since the girth of H is at least 6 by Lemma 2, we have $|E(C_G)| \geq l \geq 6 > \ell$. □

2.2 Cyclomatic Number

In this section, graphs are considered to be multigraphs without loops as defined in 1.2.1. Clearly, the cyclomatic number $\mu(G)$ is an upper bound for the maximum cardinality $\nu_e(G)$ of an edge-disjoint cycle packing, since if G would contain a collection of more than $\mu(G)$ edge-disjoint cycles, removing one edge from each cycle would not create any new components and thus yield a graph with a negative cyclomatic number, a contradiction.

In Section 2.2.1, we prove that for every fixed nonnegative integer k, there is a finite set $\mathcal{P}_e(k)$ of isomorphism types of graphs such that every 2-connected graph G with $n(G) \geq 2$ and $\mu(G) - \nu_e(G) = k$ arises by applying a simple extension rule to one of the graphs in $\mathcal{P}_e(k)$, i.e. there are essentially only finitely many configurations which cause $\mu(G)$ and $\nu_e(G)$ to deviate by k. Furthermore, we determine $\mathcal{P}_e(k)$ for $k \leq 2$. The results of this section are based on [32]. $\mathcal{P}_e(k)$

A related problem is to find the minimum value p such that all graphs G with $m(G) - n(G) \geq p$ contain k edge-disjoint cycles. This parameter is defined as $p(k)$, and there are several classical results: $p(k)$

$$p(k) = \begin{cases} 0 & , k = 1 \\ 4 & , k = 2 \quad [21] \\ 10 & , k = 3 \quad [53] \\ 18 & , k = 4 \quad [8, 73] \\ \Theta\left(k \log k\right) & \qquad\quad [21, 69, 74, 73]. \end{cases}$$

In Section 2.2.2, we obtain results analogous to those in Section 2.2.1, applied to vertex-disjoint cycle packings. We obtain a similar characterisation of the two-connected graphs with $\mu(G) - \nu_v(G) = k$ and apply this result to construct a linear-time algorithm that determines in linear time whether an arbitrary graph satisfies $\mu(G) - \nu_v(G) = k$. The results of this section are based on [60].

2.2.1 Graphs G with $\mu(G) - \nu_e(G) = k$

The connected graphs G with $\mu(G) - \nu_e(G) = 0$ are exactly the *cactus graphs*, i.e. their blocks are either edges or arise by possibly subdividing the edges of a cycle of length 2: These are precisely the graphs in which all cycles are edge-disjoint. If all cycles in G are edge-disjoint, then removing one edge from each of the $\nu_e(G)$ cycles yields a tree,

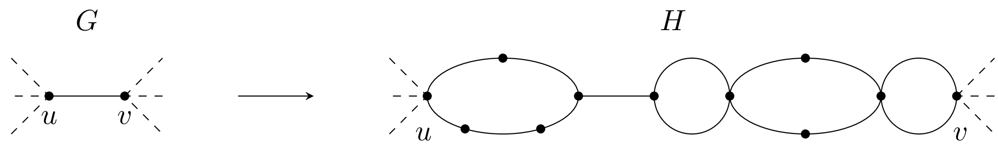


Figure 2.5: Replacing the edge $e \in E(G)$ with a 4-cycle-path

so $\mu(G) = \nu_e(G)$. Conversely, if $\mu(G) = \nu_e(G)$, then there exists a set $\mathcal{C}$ of $\mu(G)$ edge-disjoint cycles. Since removing one edge from each of these cycles yields a tree, no edge in $E(G) \setminus E(\mathcal{C})$ is contained in a cycle, so all edges of G are edge-disjoint.

$\mathcal{G}_e(k)$ For $k \in \mathbb{N}_0$, let $\mathcal{G}_e(k)$ denote the set of 2-connected graphs G with $n(G) \geq 2$ and $\mu(G) - \nu_e(G) = k$. By the above remark on cactus graphs, $\mathcal{G}_e(0)$ contains exactly the 2-connected cactus graphs, i.e. K_2 and the cycles $C_2, C_3, \ldots$. The next lemma shows that it suffices to restrict our attention to 2-connected graphs.

Lemma 4. *Let $k \in \mathbb{N}_0$. If G is a graph with $\mu(G) - \nu_e(G) = k$ whose blocks $B_1, B_2, \ldots, B_l$ satisfy $B_i \in \mathcal{G}_e(k_i)$ for $1 \leq i \leq l$, then $k = \sum_{i=1}^{l} k_i$.*

Proof. Every cycle of G is entirely contained in some block of G. □

Now we need some notation to explain the extension rule for the definition of $\mathcal{P}_e(k)$.

Let $l \in \mathbb{N}_0$. An *l-cycle-path* is a cactus with at most 2 endblocks and exactly l cycles.

An *l-cycle-path-subgraph* of a graph G with *attachment vertices* u and v is an induced l-cycle-path H in G such that u and v are two distinct vertices of H for which $d_G(w) = d_H(w)$ for all $w \in V_H \setminus \{u, v\}$, and the graph obtained from H by adding an edge between u and v is 2-connected, i.e. only the attachment vertices may have neighbours outside of V_H and, if H has more than one block, then the attachment vertices are two non-cutvertices from the two endblocks of H.

Note that a 0-cycle-path-subgraph of G with attachment vertices u and v is an ear of G with endvertices u and v.

A graph H is said to arise from a graph G *by replacing the an edge e incident with $u, v \in V(G)$ with an l-cycle-path*, if H has an l-cycle-path-subgraph Q with attachment vertices u and v such that (cf. Figure 2.5)

$$\begin{aligned} V(G) &= V(H) \setminus \big(V(Q) \setminus \{u, v\}\big), \text{ and} \\ E(G) &= \big(E(H) \setminus E(Q)\big) \cup \{e\}. \end{aligned}$$

A graph H is said to *extend* a graph G, if there is a maximum edge-disjoint cycle
$E(\mathcal{C})$ packing $\mathcal{C}$ of G such that H arises from G by replacing every edge $e \in E(\mathcal{C}) := \bigcup_{C \in \mathcal{C}} E(C)$ with a 0-cycle-path and replacing every edge $e \in E(G) \setminus E(\mathcal{C})$ with an l-cycle-path for some $l \in \mathbb{N}_0$. A graph H is said to be *reduced*, if it extends no graph other than itself.

For $k \in \mathbb{N}_0$, let $\mathcal{P}_e(k)$ denote the set of reduced graphs in $\mathcal{G}_e(k)$. Note that the characterisation of $\mathcal{G}_e(0)$ implies $\mathcal{P}_e(0) = \{K_2, C_2\}$. For $k \geq 1$, a graph in $\mathcal{P}_e(k)$ contains neither vertices of degree at most 2 nor l-cycle-path-subgraphs for $l \geq 2$.

The next lemma summarises some important properties of the above extension notion.

Lemma 5. *If $G_0 \in \mathcal{G}_e(k)$, G_1 extends G_0, and G_2 extends G_1, then*

(i) $G_1 \in \mathcal{G}_e(k)$,

(ii) G_2 *extends* G_0, *and*

(iii) every graph in $\mathcal{G}_e(k)$ *extends a graph in* $\mathcal{P}_e(k)$.

Proof. Let $\mathcal{C}_0$ be a maximum edge-disjoint cycle packing of G_0 such that G_1 arises from G_0 by replacing every edge $e \in E(G_0)$ with an l_e-cycle-path L_e with $l_e = 0$ for $e \in E(\mathcal{C}_0)$. Let $\mathcal{C}_1'$ denote the set of the $\sum_{e\in E(G_0)} l_e$ edge-disjoint cycles contained in the l_e-cycle-paths L_e for $e \in E(G_0)$. Then $\mu(G_1) = \mu(G_0) + |\mathcal{C}_1'|$.

Since the set of the cycles in G_1 that are subdivisions of the cycles in $\mathcal{C}_0$ together with the cycles in $\mathcal{C}_1'$ form an edge-disjoint cycle packing of G_1, we obtain $\nu_e(G_1) \geq \nu_e(G_0) + |\mathcal{C}_1'|$.

Let $\mathcal{C}_1$ be a maximum edge-disjoint cycle packing of G_1 such that G_2 arises from G_1 by replacing every edge $f \in E(G_1)$ with an h_f-cycle-path H_f with $h_f = 0$ for $f \in E(\mathcal{C}_1)$ and such that subject to this condition $|\mathcal{C}_1' \cap \mathcal{C}_1|$ is as large as possible.

If E_1' is an arbitrary set of edges which contains exactly one edge from every cycle in $\mathcal{C}_1'$, then removing the $|\mathcal{C}_1'|$ edges in E_1' from G_1 can destroy at most $|\mathcal{C}_1'|$ cycles in $\mathcal{C}_1$. Since the remaining cycles are subdivisions of edge-disjoint cycles in G_0, we obtain $\nu_e(G_0) \geq \nu_e(G_1) - |\mathcal{C}_1'|$.

In view of the above, this implies that

$$\nu_e(G_1) \quad = \quad \nu_e(G_0) + |\mathcal{C}_1'| \tag{2.1}$$

and hence (i).

Furthermore, this implies that every edge contained in a cycle in $\mathcal{C}_1'$ belongs to $E(\mathcal{C}_1)$, and edges contained in different cycles in $\mathcal{C}_1'$ are contained in different cycles in $\mathcal{C}_1$: Otherwise there would be a choice for E_1' such that removing the edges in E_1' would only delete at most $|\mathcal{C}_1'| - 1$ cycles in $\mathcal{C}_1$, which implies the contradiction $\nu_e(G_0) \geq \nu_e(G_1) - |\mathcal{C}_1'| + 1$.

If follows that $\mathcal{C}_1$ contains all edge-disjoint cycles contained in the l_e-cycle-path L_e for each $e \in E(G_0)$ with $l_e \geq 2$.

Furthermore, if $l_e = 1$ for some $e \in E(G_0)$ and $\mathcal{C}_1$ does not contain the unique cycle C_e within the 1-cycle-path L_e, then there are exactly two cycles C_e' and C_e'' in $\mathcal{C}_1$ which contain $E(C_e)$. Since $\big(E(C_e') \cup E(C_e'')\big) \setminus E(C_e)$ contains the edge set of a cycle C_e''',

$$\tilde{\mathcal{C}}_1 = \big(\mathcal{C}_1 \setminus \{C_e', C_e''\}\big) \cup \{C_e, C_e'''\})$$

is a maximum edge-disjoint cycle packing of G_1 such that $E(\tilde{\mathcal{C}}_1) \subseteq E(\mathcal{C}_1)$ and

$$|\mathcal{C}_1' \cap \tilde{\mathcal{C}}_1| > |\mathcal{C}_1' \cap \mathcal{C}_1|,$$

which contradicts the choice of $\mathcal{C}_1$.

Hence $\mathcal{C}_1' \subseteq \mathcal{C}_1$. By (2.1), the cycles in $\mathcal{C}_1 \setminus \mathcal{C}_1'$ are subdivisions of the cycles in a maximum cycle packing $\mathcal{C}_0'$ of G_0. Clearly, $l_e > 0$ implies $e \notin E(\mathcal{C}_0')$. Since $h_f > 0$ for some $f \in E(G_1) \setminus E(\mathcal{C}_1)$ implies that f is a bridge of an l_e-cycle-path L_e with $e \notin E(\mathcal{C}_0')$, G_2 extends G_0, i.e. (ii) holds.

By definition, for every graph $H \in \mathcal{G}_e(k)$ there is a graph $G \in \mathcal{P}_e(k)$ such that H arises from G by a finite sequence of extensions. By (ii), H extends G, and (iii) follows. □

We proceed to our main result.

Theorem 8. *The set $\mathcal{P}_e(k)$ is finite for every $k \in \mathbb{N}_0$.*

Proof. We prove the result by induction on k. Since $|\mathcal{P}_e(0)| = 2$, we may assume $k \geq 1$. We will argue that every graph in $\mathcal{P}_e(k)$ arises from some graph in $\mathcal{P}_e(k-1)$ by applying a subset of a finite set of operations. Since, by induction, $\mathcal{P}_e(k-1)$ is finite, so is $\mathcal{P}_e(k)$.

Let $H \in \mathcal{P}_e(k)$. Let $H_0, H_1, \ldots, H_t = H$ be an ear decomposition of H, i.e. H_0 is a cycle and, for $1 \leq i \leq t$, the graph H_i arises from H_{i-1} by adding an ear. Clearly, for $1 \leq i \leq t$, $\mu(H_i) = \mu(H_{i-1}) + 1$ and $\nu_e(H_{i-1}) \leq \nu_e(H_i) \leq \nu_e(H_{i-1}) + 1$, which implies

$$\mu(H_{i-1}) - \nu_e(H_{i-1}) \leq \mu(H_i) - \nu_e(H_i) \leq \mu(H_{i-1}) - \nu_e(H_{i-1}) + 1.$$

Therefore, since $H_0 \in \mathcal{G}_e(0)$, $H = H_t \in \mathcal{G}_e(k)$ and $k \geq 1$, there is some $1 \leq i^* \leq t$ such that $H_{i^*-1} \in \mathcal{G}_e(k-1)$ and $H_i \in \mathcal{G}_e(k)$ for $i^* \leq i \leq l$. Setting $l = t - i^* + 1$ and $G_i = H_{i^*+i-1}$ for $0 \leq i \leq l$ yields a sequence of 2-connected graphs

$$G_0, G_1, \ldots, G_l$$

such that

- $G_l = H$,
- G_i arises by adding the ear P_i to G_{i-1} for $1 \leq i \leq l$,
- $\nu_e(G_0) = \nu_e(G_1)$ and
- $\nu_e(G_{i-1}) = \nu_e(G_i) - 1$ for $2 \leq i \leq l$.

We assume that the sequence is chosen to be shortest possible, i.e. l is minimum.

Note that $G_0 \in \mathcal{G}_e(k-1)$ and $G_i \in \mathcal{G}_e(k)$ for $1 \leq i \leq l$.

By Lemma 5 (iii), G_0 extends some graph $G \in \mathcal{P}_e(k-1)$.

Let $\mathcal{C}_l$ be an optimal cycle packing of $H = G_l$.

Since $\nu_e(G_{l-1}) = \nu_e(G_l) - 1$ for $l \geq 2$, and removing the ear P_l from G_l destroys at most one cycle from $\mathcal{C}_l$, the ear P_l is contained in a unique cycle $C_l \in \mathcal{C}_l$, and $\mathcal{C}_{l-1} := \mathcal{C}_l \setminus \{C_l\}$ is an optimal cycle packing of G_{l-1}. Iterating this argument, we obtain that for $i = l, (l-1), (l-2), \ldots, 2$, the ear P_i is contained in a unique cycle $C_i \in \mathcal{C}_i \subseteq \mathcal{C}_l$ and that $\mathcal{C}_{i-1} := \mathcal{C}_l \setminus \{C_i, C_{i+1}, \ldots, C_l\}$ is an optimal cycle packing of G_{i-1}. Note that this argument does not apply to $i = 1$ because $\nu_e(G_0) = \nu_e(G_1)$.

Since each of the ears in $\mathcal{E} := \{P_2, P_3, \ldots, P_l\}$ is contained in a unique different cycle in $\mathcal{C}_l$, no internal vertex of any P_i is contained in any P_j for indices $i \neq j$ with $2 \leq i \leq l$

and $1 \leq j \leq l$. Since H is reduced and hence has no vertex of degree 2, this implies that the ears in $\mathcal{E}$ all have length 1, i.e. they are all edges.

Let $P = v_0 e_1 v_1 e_2 v_2 \ldots e_r v_r$ be a maximal ear of G_1. Since G_1 is 2-connected and $k \geq 1$, the endvertices v_0 and v_r of P have degree at least 3. Let $I = \{v_1, v_2, \ldots, v_{r-1}\}$ be the set of internal vertices of P.

The next claim is obvious.

Claim A *If an ear P_i for $2 \leq i \leq l$ has exactly one endvertex in I, then C_i contains either the edge e_1 or the edge e_r. Therefore, at most two ears in $\mathcal{E}$ have exactly one endvertex in I.*

Claim B *No ear P_i for $2 \leq i \leq l$ has both endvertices in I.*

Proof of Claim B. For contradiction, we assume that the index i with $2 \leq i \leq l$ is minimum such that P_i has the endvertices $v_x, v_y \in I$ for $1 \leq x < y \leq r-1$. Since $\nu_e(G_{i-1}) = \nu_e(G_i) - 1$, the cycle C_i consists of P_i and the subpath P' of P between v_x and v_y. This implies that no internal vertex of P' is an endvertex of an ear $P_j \in \mathcal{E} \setminus \{P_i\}$. Hence P_i is an ear of H, and C_i is a 1-cycle-path-subgraph of H.

Let H' arise from H by removing the ear P_i.

If $\nu_e(H') = \nu_e(H)$, we may choose $\tilde{G}_0 = H'$, $\tilde{P}_1 = P_i$ and $\tilde{G}_1 = H$, contradicting the minimality of the sequence $G_0, G_1, \ldots, G_l$. Hence $\nu_e(H') = \nu_e(H) - 1$. This implies that H' has a maximum edge-disjoint cycle packing that does not use the edges of P' and H is not reduced, which is a contradiction. $\square$

Claim C *G_1 does not contain a 2-cycle-path-subgraph.*

Proof of Claim C. For contradiction, we assume that Q is a 2-cycle-path-subgraph of G_1 with attachment vertices u and v. We may assume that $d_Q(u), d_Q(v) \geq 2$, i.e. the two cycles C' and C'' of Q are the endblocks of Q.

Clearly, for every maximum edge-disjoint cycle packing $\mathcal{C}_1'$ of G_1, we have $E(C') \cup E(C'') \subseteq E(\mathcal{C}_1')$. This implies that $E_{C'} \cup E_{C''} \subseteq E(\mathcal{C}_1)$ and, by Claims A and B, no ear in $\mathcal{E}$ has an endvertex in $V_Q \setminus \{u, v\}$. Hence Q is also a 2-cycle-path-subgraph of H, and H is not reduced, which is a contradiction. $\square$

Since G_1 arises by adding the ear P_1 to G_0, Claim C implies that G_0 does not contain any s-cycle-path-subgraph for $s \geq 6$. Since every s-cycle-path-subgraph for $s \leq 5$ yields at most $2 \cdot 5 + 6 = 16$ maximal ears, the number of maximal ears of G_0 is at most $16m(G)$, so the number of maximal ears of G_1 is at most $16m(G) + 3$.

Since H is reduced and hence has no vertex of degree 2, Claim A implies that no maximal ear of G_1 has more than 2 internal vertices. Therefore, order and size of G_1 are bounded in terms of $m(G)$.

Since all ears in $\mathcal{E}$ are edges between vertices of G_1, the number of ears in $\mathcal{E}$ with different endvertices is bounded in terms of $n(G_1)$, so it is bounded in terms of $m(G)$.

Furthermore, since all ears in $\mathcal{E}$ are contained in different edge-disjoint cycles, the number of ears in $\mathcal{E}$ which have the same endvertices is bounded by $m(G_1)$, so it is bounded in terms of $m(G)$.

Altogether, G_1 arises from G by applying a subset of a set of operations whose cardinality is bounded in terms of $m(G)$, and H arises from G_1 by applying a subset of a set of operations whose cardinality is also bounded in terms of $m(G)$. $\square$

Note that the proof of Theorem 8 yields a — rather inefficient — algorithm which for $k \geq 1$ allows to derive $\mathcal{P}_e(k)$ from $\mathcal{P}_e(k-1)$ and whose running time is bounded in terms of $|\mathcal{P}_e(k-1)|$ and the maximum size of graphs in $\mathcal{P}_e(k-1)$. Therefore the set $\mathcal{P}_e(k)$ can be constructed for every $k \in \mathbb{N}$.

We conclude with another algorithmic consequence of Theorem 8.

Let $k \in \mathbb{N}_0$ be fixed and let G be a fixed graph in $\mathcal{P}(k)$. For a given 2-connected graph H as input, we can decide in linear time whether H extends G. A simple argument for this is to consider all injective maps $V(G) \to V(H)$ and check whether the edges of G can be suitably replaced by cycle-paths in order to obtain H. This can clearly be done in polynomial time.

Therefore, in view of Lemma 4 and Theorem 8, for a given graph H as input, it can be decided in polynomial time whether $\mu(H) - \nu_e(H) = k$. Furthermore, in view of the proof of Lemma 5, we can also efficiently construct an optimal cycle packing of H — even all of them — in this case.

$\mathcal{P}_e(1)$ and $\mathcal{P}_e(2)$

In this subsection we illustrate Theorem 8 and determine $\mathcal{P}_e(1)$ and $\mathcal{P}_e(2)$ explicitly.

The following lemma captures a straightforward yet important observation which was essentially also used in the proof of Theorem 8.

Lemma 6. *Let $k \geq 1$.*

(i) Every graph $H \in \mathcal{P}_e(k)$ arises by adding an edge to a graph G such that either $\nu_e(G) = \nu_e(H)$ and G extends a graph in $\mathcal{P}_e(k-1)$, or $\nu_e(G) = \nu_e(H) - 1$ and G extends a graph in $\mathcal{P}_e(k)$.

(ii) Let $\mathcal{Q} \subseteq \mathcal{P}_e(k)$.

If every graph H in $\mathcal{P}_e(k)$ which arises by adding an edge to a graph G such that either $\nu_e(G) = \nu_e(H)$ and G extends a graph in $\mathcal{P}(k-1)$, or $\nu_e(G) = \nu_e(H) - 1$ and G extends a graph in $\mathcal{Q}$, also belongs to $\mathcal{Q}$, then $\mathcal{Q} = \mathcal{P}_e(k)$.

Proof. (i) Let $H \in \mathcal{P}_e(k)$ and let P be the last ear in some ear decomposition of H. Since H is reduced, P has length 1. Let G arise by removing $E(P)$ from H. Then $\mu(G) = \mu(H) - 1$, and $\nu_e(G) \in \{\nu_e(H), \nu_e(H) - 1\}$.

By the definition of $\mathcal{P}_e(k)$, $\nu_e(G) = \nu_e(H)$ implies that G extends a graph in $\mathcal{P}_e(k-1)$ and $\nu_e(G) = \nu_e(H) - 1$ implies that G extends a graph in $\mathcal{P}_e(k)$.

(ii) Let $H \in \mathcal{P}_e(k)$.

By iteratively deleting edges as in (i) and reducing the constructed graphs, we obtain a finite sequence $G_0, G_1, \ldots, G_l$ such that $G_l = H$, $G_i \in \mathcal{P}_e(k)$ for $1 \leq i \leq l$, $G_0 \in \mathcal{P}_e(k-1)$, and for $1 \leq i \leq l$, G_i contains an edge e_i such that $G_i - e_i$ extends G_{i-1}.

By inductive application of the hypothesis, we obtain that $G_i \in \mathcal{Q}$ for $1 \leq i \leq l$, i.e. $H \in \mathcal{Q}$ which implies $\mathcal{Q} = \mathcal{P}_e(k)$.

□

Note that Lemma 6 (ii) yields a criterion to check whether some subset $\mathcal{Q}$ of $\mathcal{P}_e(k)$ already contains all of $\mathcal{P}_e(k)$. Therefore, the proofs of the following two results reduce to case analysis. The following result is equivalent to a result in [16].

Theorem 9. $\mathcal{P}_e(1) = \{K_2^3\}$, *where* K_2^3 *is the unique graph with two vertices and three parallel edges (cf. Figure 2.6).*

Proof. It is easy to verify that $K_2^3 \in \mathcal{P}_e(1)$.

Let $H \in \mathcal{P}_e(1)$ be a graph that arises by adding an edge to a graph G. If G extends a graph in $\mathcal{P}_e(0)$, then G is a cycle-path. This implies that, since H is reduced, $H \cong K_2^3$.

Furthermore, if $H \in \mathcal{P}_e(1)$ arises by adding an edge to a graph G with $\nu_e(G) = \nu_e(H) - 1$ and G extends K_2^3, then H is not reduced, a contradiction. By Lemma 6 (ii), the proof is complete. □

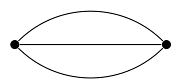


Figure 2.6: $\mathcal{P}_e(1) = \left\{K_2^3\right\}$.

We say that the graphs which arise from one of the two graphs G_1 or G_2 in Figure 2.7 by contracting a subset of the edges indicated by dashed lines are *generated from* G_1 or G_2, respectively.

Theorem 10. $\mathcal{P}_e(2)$ *consists of* K_4 *and all graphs which are generated from* G_1 *or* G_2.

Proof. It is easy to verify that K_4 and all graphs which are generated from G_1 or G_2 belong to $\mathcal{P}_e(2)$.

Let $H \in \mathcal{P}_e(2)$. We consider different cases.

Case 1 *H arises by adding an edge uv to a graph G with $\nu_e(G) = \nu_e(H) = 1$ such that G extends K_2^3.*

In this case, G is a subdivision of K_2^3. Since $\nu_e(H) = 1$, the vertices u and v are not contained in a common maximal ear of G. As H is reduced, this implies that $H = K_4$.

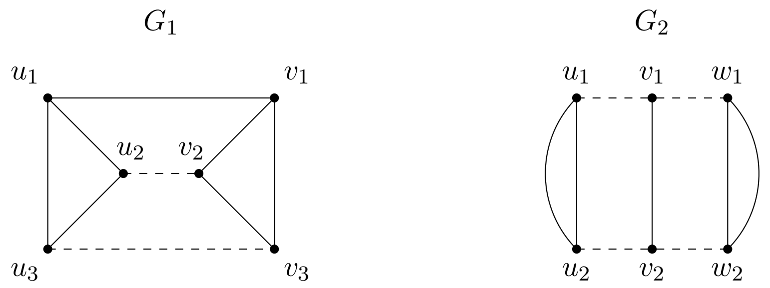


Figure 2.7: The graphs $G_1, G_2 \in \mathcal{P}_e(2)$.

Case 2 *H arises by adding an edge uv to a graph G with $\nu(G) = \nu(H) \geq 2$ such that G extends K_2^3.*

In this case G has a unique optimal cycle packing $\mathcal{C}$.

If $d_G(u) = d_G(v) = 2$ and u and v are inner vertices of some maximal ear contained in a cycle in $\mathcal{C}$, then $H = G_2$.

If $d_G(u) = d_G(v) = 2$ and u and v are inner vertices in different maximal ears contained in a cycle in $\mathcal{C}$, then H extends K_4. Since $H \neq K_4$, H is not reduced, which is a contradiction.

If $d_G(u) = d_G(v) = 2$ and u and v are contained in different cycles in $\mathcal{C}$, then H is generated from G_1.

If $d_G(u) \geq 3$, $d_G(v) = 2$, and v is contained in a cycle in $\mathcal{C}$, then H extends K_4. Since $H \neq K_4$, H is not reduced, which is a contradiction.

In all remaining subcases, H is generated from G_2.

Case 3 *H arises by adding an edge uv to a graph G with $\nu_e(G) = \nu_e(H) - 1$ such that G extends K_4.*

Let v_1, v_2, v_3, v_4 denote the vertices of K_4. We may assume that G arises by replacing the edges v_iv_j with $l_{i,j}$-cycle-paths $Q_{i,j}$.

Since H is reduced and $\nu_e(G) = \nu_e(H) - 1$, the vertices u and v are not both contained in one of the cycle-paths $Q_{i,j}$ and we obtain that H is generated from G_1.

Case 4 *H arises by adding an edge uv to a graph G with $\nu_e(G) = \nu_e(H) - 1$ such that G extends a graph generated from G_1.*

It is easy to verify that $\nu_e(G) = \nu_e(H) - 1$ implies that H is generated from G_1.

Case 5 *H arises by adding an edge uv to a graph G with $\nu_e(G) = \nu_e(H) - 1$ such that G extends a graph generated from G_2.*

It is easy to verify that $\nu_e(G) = \nu_e(H) - 1$ implies that H is generated from K_4 or G_2.

By Lemma 6 (ii), the proof is complete. $\square$

2.2.2 Graphs G with $\mu(G) - \nu_v(G) = k$

Since every vertex-disjoint cycle packing is also an edge-disjoint cycle packing, the inequality $\nu_v(G) \leq \nu_e(G) \leq \mu(G)$ holds for all graphs G.

In analogy to the results of the previous subsection, we are now going to prove the existence of a finite set $\mathcal{P}_v(k)$ of graphs for all $k \in \mathbb{N}_0$ such that every 2-connected graph G with $\mu(G) - \nu_v(G) = k$ arises by applying a simple extension rule to a graph in $\mathcal{P}_v(k)$. As an algorithmic consequence we describe an algorithm that calculates $\min\{\mu(G) - \nu_v(G), k+1\}$ in linear time for fixed k. Unlike in the edge-disjoint case, the problem to find many vertex-disjoint cycles in a graph can not be reduced to its blocks.

We start by giving a constructive characterization of the graphs in $\mathcal{G}_v(k)$

$$\mathcal{G}_v(k) \;=\; \{G \mid n(G) \geq 2,\ \mu(G) - \nu_v(G) = k, \text{ and } G \text{ is 2-connected}\}.$$

In order to define the extension rule, we need similar definitions as in the previous subsection. For $l \in \mathbb{N}_0$, a graph P is an *l-cycle-chain between u and v*, if

- P is cactus with at most two endblocks,
- the set $\mathcal{C}(P)$ of cycles of P consists of l vertex-disjoint cycles,
- $u \neq v$ and $d_P(u) = d_P(v) = 1$.

If G is a graph and $e \in E(G)$ is an edge with endvertices u and v, then the graph H is said *to arise from G by replacing the edge e with an l-cycle-chain P* (cf. Figure 2.8), if H arises from the disjoint union of G and an l-cycle-chain P between u' and v' by removing the edge e and identifying u with u' and v with v'. In this case H is said to *contain the l-cycle-chain P*. Note that subdividing an edge is the same as replacing it with a 0-cycle-chain.

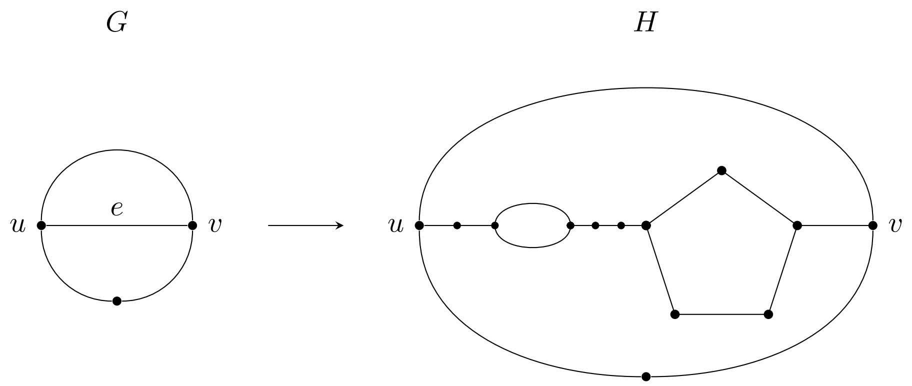


Figure 2.8: Replacing the edge $e = uv \in E(G)$ with a 2-cycle-chain

We say that a graph H *extends* a graph G, if H arises from G by replacing every edge $e \in E(G)$ with an l_e-cycle-chain P_e such that $\mu(H) - \nu_v(H) = \mu(G) - \nu_v(G)$. A graph H is called *reduced*, if H does not extend a graph G different from H.

Let $\mathcal{P}_v(k)$ be the set of the reduced elements of $\mathcal{G}_v(k)$. $\mathcal{P}_v$

The next lemma summarises some important properties of the above extension notion.

Lemma 7. *Let H arise from G by replacing every edge $e \in E(G)$ with an l_e-cycle-chain P_e. Let*

$$l = \sum_{e \in E(G)} l_e \quad \text{and} \quad \mathcal{C} = \bigcup_{e \in E(G)} \mathcal{C}(P_e).$$

(i) If H extends G, then $\mu(H) - \mu(G) = \nu_v(H) - \nu_v(G) = l$, and every maximum vertex-disjoint cycle packing of H contains all l cycles in $\mathcal{C}$.

(ii) H extends G if and only if G has a maximum vertex-disjoint cycle packing $\mathcal{C}(G)$ such that $l_e = 0$ for all $e \in E(\mathcal{C}(G))$.

Proof. Let $\mathcal{C}(H)$ be a maximum vertex-disjoint cycle packing of H. Let E be a set of l edges that intersects every cycle in $\mathcal{C}$. Removing the edges in E destroys at most l cycles in $\mathcal{C}(H)$, so

$$\nu_v(H) - \nu_v(G) \leq l. \tag{2.2}$$

Clearly, $\mu(H) - \mu(G) = l$.

(i) Since H extends G, we have $\mu(H) - \nu(H) = \mu(G) - \nu(G)$, so $\nu(H) - \nu(G) = l$. Furthermore, since (2.2) holds with equality for every choice of E, we obtain $E(\mathcal{C}) \subseteq E(\mathcal{C}(H))$. By the definition of a cycle-chain, this implies $\mathcal{C} \subseteq \mathcal{C}(H)$.

(ii) If H extends G, then, by (i), the cycles in $\mathcal{C}(H) \setminus \mathcal{C}$ are subdivisions of the cycles in an optimal cycle packing $\mathcal{C}(G)$ of G. Clearly, $l_e = 0$ for all $e \in E(\mathcal{C}(G))$.

Conversely, if $\mathcal{C}(G)$ is a maximum vertex-disjoint cycle packing of G such that $l_e = 0$ for all $e \in E(\mathcal{C}(G))$, then the cycles in H which are subdivisions of the cycles in $\mathcal{C}(G)$ together with the cycles in $\mathcal{C}$ form a cycle packing of H, so $\nu_v(H) - \nu_v(G) \geq l$. Together with (2.2), it follows that $\nu_v(H) - \nu_v(G) = l$ and H extends G.

□

As for the extension operation for the definition of $\mathcal{P}_e$, iterated extensions as used for the definition of $\mathcal{P}_v$ are not more powerful than single extensions. The proof is simpler than that of the corresponding Lemma 5.

Lemma 8. *(i) If G_2 extends G_1 and G_1 extends G_0, then G_2 extends G_0.*

(ii) For $k \in \mathbb{N}_0$, every graph in $\mathcal{G}_v(k)$ extends a graph in $\mathcal{P}_v(k)$.

Proof. (i) For $i \in \{1, 2\}$, let G_i extend G_{i-1} by replacing every edge $e \in E(G_{i-1})$ with an $l_e^{(i)}$-cycle-chain $P_e^{(i)}$. If $e \in E(G_0)$, $f \in E\left(P_e^{(1)}\right)$, and $l_f^{(2)} \geq 1$, then, by Lemma 7 (i), f is a bridge of $P_e^{(1)}$. Therefore, G_2 extends G_0 by replacing every edge $e \in E(G_0)$ with an l_e-cycle-chain, where

$$l_e = l_e^{(1)} + \sum_{f \in E\left(P_e^{(1)}\right)} l_f^{(2)}.$$

(ii) Let $H \in \mathcal{G}_v(k)$. By definition, there is a finite sequence $G_0, G_1, \ldots, G_s \in \mathcal{G}_v(k)$ such that G_i extends G_{i-1} for $1 \leq i \leq s$, $G_0 \in \mathcal{P}_v(k)$ and $H = G_s$. Repeated application of (i) implies that H extends G_0.

□

In view of the observation on graphs G with $\mu(G) = \nu_v(G)$ made in the introduction it is easy to determine $\mathcal{G}_v(0)$ and $\mathcal{P}_v(0)$.

Lemma 9. *(i) No reduced graph H contains a vertex u with $d_H(u) = |N_H(u)| = 2$ or a 2-cycle-chain.*

(ii) $\mathcal{G}_v(0) = \{K_2\} \cup \{C_n \mid n \geq 2\}$ and $\mathcal{P}_v(0) = \{K_2, C_2\}$.

Proof. (i) If $u \in V(H)$ is a vertex with $d_H(u) = |N_H(u)| = 2$, then contracting an edge incident with u results in a graph G such that H extends G, so H is not reduced. If H contains a 2-cycle-chain P, then every maximum vertex-disjoint cycle packing of H contains both cycles of P. Therefore, if G arises from H by contracting one cycle C in P together with one further edge incident with C (cf. Figure 2.9), then H extends G, so H is not reduced.

(ii) Let $G \in \mathcal{G}_v(0)$. Since $\nu_v(G) \leq \nu_e(G) \leq \mu(G)$, we have $\mathcal{G}_v(0) \subseteq \mathcal{G}_e(0)$. Indeed, $\mathcal{G}_v(0) = \mathcal{G}_e(0) = \{K_1, K_2\} \cup \{C_n \mid n \geq 2\}$. By (i), K_0, K_1, K_2 and C_2 are the only reduced graphs in $\mathcal{G}_v(0)$.

□

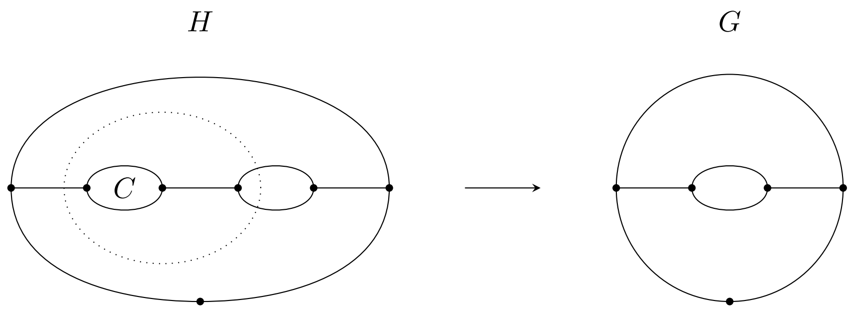


Figure 2.9: Contraction in the proof of Lemma 9 (i)

In analogy to Theorem 8, we can now prove the main result of this section.

Theorem 11. *$\mathcal{P}_v(k)$ is finite for every $k \in \mathbb{N}_0$.*

Proof. We prove the result by induction on k. For $k = 0$, the result follows from Lemma 9 (ii).

For positive k, we are going to show that the number of edges in any graph $H \in \mathcal{P}_v(k)$ is bounded in terms of the number of edges in some graph in $\mathcal{P}_v(k-1)$.

Since H is 2-connected and has order at least 2, it has an ear decomposition, i.e. it arises from a chordless cycle by iteratively adding ears. Since removing an ear from H

reduces $\mu(H)$ by exactly 1 and $\nu_v(H)$ by at most 1, iteratively removing the ears of an ear decomposition of H yields a sequence of 2-connected graphs $G_0, G_1, \ldots, G_l = H$, such that

- for each $i \in \{1, \ldots, l\}$, G_i arises by adding the ear Ear_i to G_{i-1},
- $\nu_v(G_{i-1}) = \begin{cases} \nu(G_i) & , \text{if } i = 1 \\ \nu_v(G_i) - 1 & , \text{if } i > 1. \end{cases}$

The second condition implies that $G_0 \in \mathcal{G}_v(k-1)$ and $G_i \in \mathcal{G}(k)$ for $i \in \{1, \ldots, l\}$. By Lemma 8 (ii), G_0 extends some graph $G \in \mathcal{P}_v(k-1)$.

Let $\mathcal{C}_l$ be an optimal cycle packing of G_l. If $l \geq 2$, then the ear E_l is contained in a unique cycle Cyc_l of $\mathcal{C}_l$, and $\mathcal{C}_l \setminus \{Cyc_l\}$ is a maximum vertex-disjoint cycle packing of G_{l-1}. By repeated applications of this argument to indices from l down to 2, we obtain vertex-disjoint cycles $Cyc_2, \ldots, Cyc_l \in \mathcal{C}_l$ such that Ear_i is contained in Cyc_i for $2 \leq i \leq l$. Since H is reduced, Lemma 9 (i) implies that $\mathcal{E} := \{Ear_2, \ldots, Ear_l\}$ is a set of subgraphs which are isomorphic to K_2.

Claim: *The graph G_1 does not contain a 2-cycle-chain.*

Proof of the Claim: For contradiction, we assume that G_1 contains a 2-cycle-chain P. It suffices to show that G_2 contains a 2-cycle-chain. Repeating this argument we obtain that $G_l = H$ contains a 2-cycle-chain, so by Lemma 9, it is not reduced, which is a contradiction.

Clearly, any maximum vertex-disjoint cycle packing $\mathcal{C}_1$ of G_1 contains both cycles C' and C'' of P. Let P' denote the path in P between C' and C''. Recall that Ear_2 is contained in the cycle Cyc_2 which is vertex-disjoint to all cycles in $\mathcal{C}_1$. Therefore, if Ear_2 has no endvertex in P', then G_2 contains a 2-cycle-chain contained in P, and, if Ear_2 has an endvertex in P', then Ear_2 has both its endvertices in P' and G_2 even contains a 3-cycle-chain. □

Since G_1 arises from G_0 by adding the ear Ear_1, the claim implies that the graph G_0 does not contain a 6-cycle-chain. Since every l-cycle-chain for $l \leq 5$ contains at most $2 \cdot 5 + 6 = 16$ maximal ears, the number of maximal ears of G_0 is at most $16m(G)$. Hence the number of maximal ears of G_1 is at most $16m(G) + 3$.

Since H is reduced, all internal vertices of a maximal ear P of G_1 must be endvertices of edges in $\mathcal{E}$. At most two internal vertices can be contained in some $Ear_i \in \mathcal{E}$ such that Cyc_i contains an endvertex of P. Each further internal vertex must be incident with the edge of an ear $Ear_i \in \mathcal{E}$ such that Cyc_i consists of Ear_i and a subpath of P. Hence, since H is reduced, Lemma 9 (i) implies that each maximal ear of G_1 contains at most four internal vertices. Therefore, each maximal ear contributes at most five edges to G_1, i.e. $m(G_1) \leq 5(16m(G) + 3)$. Finally, since the edges in $\mathcal{E}$ are vertex-disjoint and $n(G_1) \leq m(G_1)$, we obtain $|\mathcal{E}| \leq \frac{5}{2}(16m(G) + 3)$, which implies $m(H) \leq 8(16m(G) + 3)$. □

We proceed to an algorithmical consequence of Theorem 11:

Input: A graph G
Output: $\min\{\mu(G) - \nu_v(G), k + 1\}$

begin
    while G *contains a bridge* $e \in E(G)$ **do**
        Delete e;
    end
    while G *contains a vertex* u *with* $d_G(u) = |N_G(u)| = 2$ **do**
        Contract one of the edges incident with u;
    end
    while G *contains a* 2*-cycle-chain* P **do**
        Contract one cycle C in P together with one further edge incident with C;
    end
    while G *contains a component* C *isomorphic to* K_1 *or* C_2 **do**
        Delete C;
    end
    if $V(G) = \emptyset$ **then return** 0;
    Select an endblock B of G;
    if $\mu(B) - \nu_v(B) \geq k + 1$ **then return** $k + 1$;
    if B *contains a cutvertex* **then**
        Let $u \in V(B)$ be the cutvertex;
    else
        Let $u \in V(B)$ be any vertex;
    end
    Let u be contained in s blocks of G;
    $\Delta k \longleftarrow \mu(B) - \nu_v(B)$;
    if u *is contained in every optimal cycle packing of* B **then**
        $\Delta k \longleftarrow \Delta k + d_{G-E(B)}(u) - (s - 1)$;
        $G' \longleftarrow G - V(B)$;
        if $\Delta k \geq k + 1$ **then return** $k + 1$;
    else
        $G' \longleftarrow G - (V(B) \setminus \{u\})$;
    end
    Let k' be the output of DIFFERENCE$(k - \Delta k)$ applied to G';
    return $\min\{\Delta k + k', k + 1\}$;
end

Algorithm 2: DIFFERENCE(k)

Theorem 12. *For every $k \in \mathbb{N}_0$, Algorithm 2 works correctly and has linear running time.*

Proof of correctness: By induction on the recursive depth, we may assume that the output of the recursive call performed in line 31 is correct.

Up to line 13, G is modified such that the difference $\mu(G) - \nu_v(G)$ does not change (cf. the argument in the proof of Lemma 9 (i)). Note that after these preprocessing steps, G contains neither a bridge, nor a vertex u with $d_G(u) = |N_G(u)| = 2$, nor a 2-cycle-chain, nor a component which is an isolated vertex or a chordless cycle.

Clearly, it is correct to return 0 in line 14.

Since $\mu(G) - \nu_v(G) \geq \mu(B) - \nu_v(B)$, it is correct to return $k + 1$ in line 16.

If u is contained in every optimal cycle packing of B, then there is an optimal cycle packing of G which is the union of an optimal cycle packing of $G - V(B)$ and an optimal cycle packing of B. Since

$$\mu(G) \;=\; \mu(G - V(B)) + \mu(B) + d_{G-E(B)}(u) - (s - 1),$$

we obtain

$$\mu(G) - \nu_v(G) \;=\; \mu(G - V(B)) - \nu_v(G - V(B)) \;+\; \mu(B) - \nu(B) \;+\; d_{G-E(B)}(u) - (s - 1),$$

and the value returned in line 27 or line 32 is correct.

If u is not contained in every optimal cycle packing of B, then there exists a maximum vertex-disjoint cycle packing of G that is the union of a maximum cycle packing of $G' := G - (V(B) \setminus \{u\})$ and a maximum cycle packing of $B' := G[V(B) \setminus \{u\}]$. Since $\mu(G) = \mu(G') + \mu(B)$ and $\nu_v(B) = \nu_v(B')$, we obtain

$$\mu(G) - \nu_v(G) \;=\; \mu(G') - \nu_v(B') \;+\; \mu(B) - \nu_v(B),$$

and the value returned in line 32 is correct. □

Proof of linear running time: If B is a component of G or u is not contained in every optimal cycle packing of B, then, by Lemma 9 (ii) and the preprocessing, $\mu(B) - \nu_v(B) > 0$. If B is contained in $s \geq 2$ blocks of G, then, by the preprocessing, G has no bridge and hence $d_{G-E(B)}(u) - (s - 1) > 0$. This implies that $\Delta k > 0$ in line 31. Therefore, the recursive depth is at most k, and it suffices to show that all steps until line 30 can be done in linear time.

Since the block-cutvertex tree of G can be determined in linear time [71], the deletion of bridges (line 3), the deletion of trivial components (line 12), the selection of B (line 15) and the selection of u (line 22) can be done in linear time. Furthermore, it is easy to see that the contractions in the preprocessing (lines 6 and 9) can be done in linear time.

By Lemma 8 (ii), if $\mu(B) - \nu_v(B) \leq k$, then there exists a graph $B' \in \mathcal{P} := \bigcup_{i=0}^{k} \mathcal{P}_v(i)$ such that B extends B'. Since B contains at most one vertex v with $d_G(v) = |N_G(v)| = 2$ — the cutvertex u — and since G contains no 2-cycle-chain after the preprocessing, B contains no 4-cycle-chain. Therefore, in order to obtain B, each edge of B' is replaced by a subgraph with at most 11 edges. Since, by Theorem 11, $\mathcal{P}$ is finite, $\mu(B) - \nu_v(B) \leq k$ can only hold, if B belongs to a finite set of graphs depending on k and lines 16, 23, and 24 can be done in constant time. □

It is easy to modify Difference(k) such that it also returns a maximum vertex-disjoint cycle packing of G in linear time provided that $\mu(G) - \nu(G) \leq k$. Such a packing would consist of the cycles contracted in line 9, the cycles of length 2 deleted in line 12, an optimal cycle packing of B which, if possible, avoids u and an optimal cycle packing of G' obtained recursively.

3 Cycle Spectrum of Hamiltonian Graphs

The *cycle spectrum* of a graph G is the set of cycle lengths of G. In this chapter, we will study lower bounds on the size $s(G)$ of the cycle spectrum of Hamiltonian graphs. $s(G)$

In the study of Hamiltonian graphs, interest in cycle spectra developed due to Bondy's "Metaconjecture" (based on [9]) that sufficient conditions for the existence of Hamiltonian cycles usually also imply that a graph is pancyclic, with possibly a small family of exceptional graphs. In particular, the result of [9] showed that the sufficient condition for the existence of a Hamiltonian cycle due to Ore [55] — every two nonadjacent vertices x and y have degrees $d_G(x)$ and $d_G(y)$ summing to at least the order n — implies further that G is pancyclic or is the complete bipartite graph $K_{\frac{n}{2},\frac{n}{2}}$. Schmeichel and Hakimi [64] advanced our understanding of Bondy's Metaconjecture by studying conditions for a Hamiltonian graph to be pancyclic. They showed that if a spanning cycle in a graph G of order n has two consecutive vertices x and y with $d_G(x) + d_G(y) \geq n$, then G is pancyclic or bipartite or lacks only a cycle of length $n-1$ from the spectrum, and if $d_G(x) + d_G(y) \geq n+1$, then G is pancyclic. Bauer and Schmeichel [5] used this to give unified proofs that the sufficient conditions of Bondy [10], Chvátal [14], and Fan [22] for the existence of Hamiltonian cycles in fact imply pancyclicity, except for a small family of exceptions. Further results about what is forced into the cycle spectrum by conditions on vertex degrees of selected vertices on a spanning cycle appear in [24] and [65].

At the 1999 conference "Paul Erdős and His Mathematics", Jacobson and Lehel initiated the study of a related question: *Under weaker conditions than those that guarantee the existence of a Hamiltonian cycle, how small can the cycle spectrum of a Hamiltonian graph be?*

By Bondy's result [9], $\lceil n(G)/2 \rceil$-regular graphs G are both Hamiltonian and pancyclic except for $K_{\frac{n(G)}{2},\frac{n(G)}{2}}$. Graphs that are 2-regular and Hamiltonian have only the cycle length $n(G)$. For regular graphs with degree greater than 2, the question becomes interesting. Jacobson and Lehel asked for the minimum size of the cycle spectrum of a k-regular Hamiltonian graph of given order, in particular for $k = 3$. They observed that for some positive constant c, $c \log n(G)$ is a lower bound. For an upper bound, they constructed the following regular examples G with $s(G) = \frac{(k-2)n(G)}{2k} + k$ for sufficiently large orders divisible by $2k$.

Example 1 (Jacobson, Lehel)**.** *For an integer $k > 1$ and an integer n which is a multiple of $2k$, arrange $\frac{n}{2k}$ disjoint copies of $K_{k,k}$ in cyclic order, with vertex sets V_1 up to $V_{\frac{n}{2k}}$. Remove one edge from each copy, and replace it by an edge to the next copy to restore regularity (see Figure 3.1 for the case $k = 3$). If the order of the construction is sufficiently large, then a cycle of length greater than $2k$ must pass through each V_i, and in each V_i it uses an even number of vertices ranging from 4 to $2k$. Hence the cycle lengths*

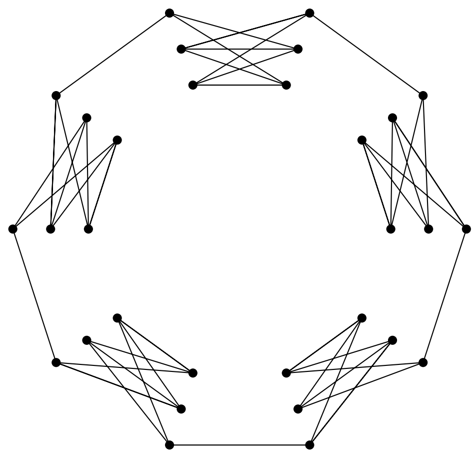


Figure 3.1: Example of regular graphs with $s(G) = \frac{(k-2)n(G)}{2k} + k$ by Jacobson and Lehel

are the even numbers from 4 *up to* $2k$*, and each even integer from* $\frac{2n}{k}$ *through* n*.*

Main result The new results of this chapter are motivated by the question of Jacobson and Lehel. As the main result of this chapter, we prove that every Hamiltonian graph satisfies $s(G) \geq \sqrt{\frac{4}{7}(m(G) - n(G))}$. The following example shows that there exist graphs with $s(G) = \Theta\left(\sqrt{m(G) - n(G)}\right)$ and that the factor $\sqrt{\frac{4}{7}}$ cannot be increased above 1.

Example 2. *For* $t \leq \frac{n}{2}$*, the graph* G *of order* n *and size* $m = n + t^2 - 2t$ *which arises by subdividing one edge of* $K_{t,t}$ *exactly* $n - 2t$ *times is bipartite and Hamiltonian. Its cycle spectrum equals* $\{4, 6, \ldots, 2t\} \cup \{n - 2t + 4, n - 2t + 6, \ldots, n\}$ *which implies* $s(G) \leq 2(t-1) = 2\sqrt{m - n + 1}$ *with equality for* $n > 4(t-1)$ *and* $s(G) = \sqrt{m - n + 1}$ *for* $n = 2t$*.*

An important tool in our proof of the main result is a lemma due to Faudree et al. (Lemma 3 in [23]). Since the original proof of this lemma given in [23] implies a slightly weaker statement, we first present a corrected proof of this lemma and some extensions in Section 3.1. In Section 3.2 we derive the consequences concerning the cycle spectrum. The results of this chapter are based on [52].

We conclude with an open question: The graphs from Example 1 satisfy $s(G) = \frac{m(G)-n(G)}{k} + k$. Although there exist some graphs with even smaller cycle spectrum (e.g. the cubic bipartite graph of order 14 and girth 6), it seems possible that for any fixed k, sufficiently large k-regular graphs satisfy $s(G) \geq \frac{m(G)-n(G)}{k} + k$. However, no lower bounds on the cycle spectrum that exploit regularity are known, so we only know that the size of the cycle spectrum is at least linear in $\sqrt{n(G)}$.

Question 1. *Does there exist a constant* $c > 0$ *such that* $s(G) \geq cn(G)$ *for all cubic Hamiltonian graphs* G*?*

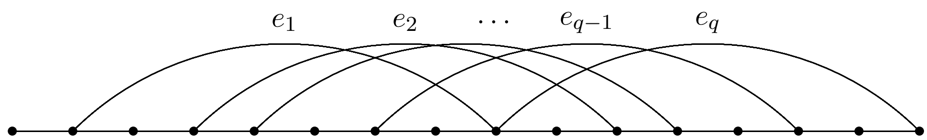


Figure 3.2: Chords in Lemma 10

3.1 Chords of a Hamiltonian Path

Throughout this section, let G be a graph with a Hamiltonian path $P = v_1 v_2 \dots v_n$ such that

$$E(G) \setminus E(P) = \{e_j \mid j \in \{1, \dots, q\}\},$$

where $e_j = v_{i_j} v_{i_j+l}$ for some $l \in \{2, \dots, n-1\}$ and the sequence i_j is monotonously increasing. We call the edges in $E(G) \setminus E(P)$ *chords* (of length l). For any $1 \leq a \leq b \leq n$, we define l, q, e_j, i_j

$$P[a, b] := P[\{v_i \colon a \leq i \leq b\}].$$

We say that two chords e_j and e_k *overlap*, if $P[i_j, i_j + l]$ and $P[i_k, i_k + l]$ contain a common edge. By a *route* in G, we denote a path with endpoints v_1 and v_n. Finally, let r be the number of lengths of routes in G. r

Lemma 3 in [23] claims that in this setting, $r \geq q/3 + 1$. However, the argument in [23] produces only $q/6 + 1$ path lengths in the following example.

Example 3. *Let G be the cubic graph of order $12k$ that is obtained from a Hamiltonian path $P = v_0 v_1 \dots v_{12k}$ by adding three chords $\{v_{6i}, v_{6i+3}\}$, $\{v_{6i+1}, v_{6i+4}\}$, and $\{v_{6i+2}, v_{6i+5}\}$ of length three for each $i \in \{0, \dots, 2k-1\}$. Any route passes through the $2k+1$ vertices $v_0, v_6, v_{12}, \dots, v_{12k}$ in increasing order. Each subpath from v_{6i} to $v_{6(i+1)}$ has either length 4 or length 6, so the lengths of routes are the $2k + 1 = q/3 + 1$ even numbers from $4 \cdot 2k$ up to $6 \cdot 2k$.*

The argument in [23] discards either the chords indexed by odd or by even values of i. By the same argument as above, the remaining graph admits only routes of $k+1 = q/6+1$ different lengths.

Theorem 13 below will provide a lower bound on r that is always at least as large as $q/3 + 1$. The graph in Example 3 demonstrates sharpness.

Lemma 10. *If any two of the chords $e_1, \dots, e_{q-1}$ overlap, then $r \geq q - 1$, with equality possible only if l is odd.*

Proof. Suppose first that e_q overlaps e_1. For $2 \leq j \leq q$, let P_j be the unique route that uses the chords e_1 and e_j and no other chords. It traverses, in this order, $P[1, i_1]$, e_1, $P[i_j, i_1 + l]$, e_j, $P[i_j, n]$, so its length is $n + 1 - 2(i_j - i_1)$, and $P_2, \dots, P_q$ have distinct lengths. Furthermore, the route Q that contains e_1 and no other chord has length $n - l$.

Since $P_2, \dots, P_q$ have distinct lengths, $r \geq q-1$. If l is even, then the length of Q has opposite parity from the lengths of $P_2, \dots, P_q$, and hence $r \geq q$.

Now suppose that e_q does not overlap e_1. Let P' be the route using e_1 and e_q and no other chords. It has length $n+1-2l$, so it is shorter than any of the paths $P_2, \dots, P_{q-1}$ or Q. □

Lemma 11. *Let $Q_1, \dots, Q_t$ be pairwise edge-disjoint subpaths of P, and let H_j be the subgraph of G induced by $V(Q_j)$. If there are paths of r_j different lengths in H_j that join the endpoints of Q_j, then G has routes of at least $1 + \sum_{j=1}^{t}(r_j - 1)$ different lengths.*

Proof. For each j, the route P can be shortened $r_j - 1$ times while replacing only edges in H_j. Combining these modifications for different values of j, we can produce paths of $\sum_{j=1}^{t}(r_j-1)$ different lengths, each of which is shorter than P. □

Theorem 13. *If G is a graph consisting of a Hamiltonian path P with vertices $v_1, \dots, v_n$ and q chords of length l, then the number r of lengths of routes in G is at least*

$$\max\left\{\frac{q}{2} - \frac{n-1}{2l} + 1,\ \frac{q}{3} + 1\right\}.$$

If l is even, then $r \geq q/2 + 1$.

Proof. We start by greedily selecting "independent chords" in order to decompose G: Choose chords $c_1, \dots c_k$ such that for $j \in \{1, \dots, k\}$, c_j is the first chord in the sequence $e_1, \dots, e_q$ that overlaps no chord c_i with $i < j$ and such that all chords coincide with or overlap one of the chords $c_1, \dots, c_k$.

Let $Q_0 := P[1, \max c_1]$, $Q_j := P[\min c_j, \max c_{j+1}]$ for $j \in [1, k-1]$ and $Q_q := P[\min c_k, n]$. For $j \in [0,k]$, let H_j be the subgraph of G induced by $V(Q_j)$, and let r_j be the number of lengths of paths in H_j that connect the endvertices of Q_j.

Since the chord c_j is contained both in H_j and H_{j-1} for $1 \leq j \leq k$ and every other chord of G belongs to exactly one of these subgraphs, $\sum_{j=0}^{k} q_j = q + k$, where q_j is the number of chords in H_j. Each H_j has the form discussed in Lemma 10. Hence $r_j \geq q_j - 1$ for $1 \leq j \leq k$, and we have $r_0 = q_0 + 1 = 2$ since H_0 contains only one chord.

The odd-indexed subgraphs among $H_0, \dots, H_k$ are pairwise disjoint, as are the even-indexed subgraphs. By applying Lemma 11 separately to the graphs arising from P by adding all chords in $\bigcup E(H_{2j})$ resp. $\bigcup E(H_{2j+1})$ and summing the resulting two inequalities, we obtain

$$2r \ \geq\ 2 + \sum_{j=0}^{k}(r_j - 1) \ \geq\ 2 + q_0 + \sum_{j=1}^{k}(q_j - 2) \ =\ 2 + q - k,$$

so $r \geq (q-k)/2 + 1$. Since no two chords among $c_1, \dots, c_k$ overlap, $n - 1 \geq kl$, so

$$r \geq \frac{q}{2} - \frac{n-1}{2l} + 1. \tag{3.1}$$

Furthermore, considering routes that use no chords other than $c_1, \ldots, c_k$, we obtain $r \geq 1 + k$. Hence

$$r \geq \max\left\{1 + k, \frac{q-k}{2} + 1\right\}.$$

Optimizing k yields $r \geq q/3 + 1$.

If l is even, then Lemma 10 yields $r_i \geq q_i$ for $1 \leq i \leq k$, hence

$$2r \;\geq\; 2 + \sum_{j=0}^{k}(r_j - 1) \;\geq\; 2 + q_0 + \sum_{i=1}^{k}(q_i - 1) \;=\; 2 + q.$$

Therefore, $r \geq q/2 + 1$ in this case. □

Corollary 1. *In the setting of Theorem 13, if $l \leq n/2$, then*

$$r \geq \frac{q}{3}\left(1 + \frac{l}{n}\right).$$

Proof. By using (3.1) to improve upon the second bound in Theorem 13, we obtain $r \geq \max\{f_1(q), f_2(q)\}$ with $f_1(x) = \frac{x}{2} - \frac{n}{2l} + 1$ and $f_2(x) = \frac{x}{3} + 1$.

Note that f_1 and f_2 are linear functions of x that intersect at a point $(x_0, y_0) = \left(\frac{3n}{l}, \frac{n}{l} + 1\right)$. Since $f_1(0) < 0 < f_2(0)$, the line $y = \frac{x}{3}\left(1 + \frac{l}{n}\right)$ that passes through $(0, 0)$ and (x_0, y_0) provides a uniform lower bound on $\max\{f_1(x), f_2(x)\}$. □

3.1.1 Chords of length greater than three

It is natural to ask in which sense the above bounds are tight. Although this question has no impact on the discussions in Section 3.2, we are going to give improved bounds in this subsection.

Tight examples for the inequality $r \geq q/3 + 1$ include Example 3 and the path of order n with $n - 3$ chords of length 3. Similarly, paths of order n with $n - 2$ chords of length 2 show that $r \geq q/2 + 1$ is best possible for even values of l. In fact, it is easy to check that for any value of l and any even value of q, the graph that arises by adding chords of length l with start vertices v_{2lj} and v_{2lj+1} for $j \in \{1, \ldots, \frac{q}{2}\}$ to the path $v_1 v_2 \ldots v_{(q+2)l}$ satisfies $r = q/2 + 1$. However, the following theorem shows that a very weak condition that excludes similar special configurations allows to improve the bounds for all values of $l > 3$.

Theorem 14. *Let $P = v_1 v_2 \ldots v_n$ be a Hamiltonian path of a graph G, and define a total order on $V(G)$ by $v_i < v_j :\Leftrightarrow i < j$. Let l be such that all other edges are chords of length l, i.e. of the form $v_x v_{x+l}$. Let r be the number of lengths of paths with endvertices v_1 und v_n in G and let q the number of chords in G. If G contains two overlapping chords $v_x v_{x+l}$ and $v_y v_{y+l}$ with $y - x \geq 2$, then*

$$r(G) \geq \begin{cases} \frac{l-1}{l} q(G) - 14l & , \textit{if } l \textit{ is even,} \\ \frac{l-1}{2l} q(G) - 6l & , \textit{if } l \textit{ is odd.} \end{cases}$$

Proof. Let

$$\lambda = \begin{cases} l-1 & \text{, if } l \text{ is even,} \\ \frac{l-1}{2} & \text{, if } l \text{ is odd.} \end{cases}$$

Let Q_0 denote the set of chords in G. Let $c_1, c_2, \dots, c_k$ be a sequence of pairwise non-overlapping chords that is obtained by the following construction: If the sequence has been constructed up to the chord c_{i-1}, we select $c_i \in Q_{i-1}$ such that

$$L_i := \{c_i\} \cup \{c' \in Q_{i-1} \colon \min c' < \min c_i < \max c'\}$$

has maximum cardinality. If $|L_i| < 3$, then we stop the construction and set $k := i-1$. Otherwise, we define Q_i to be the set of all chords from Q_{i-1} that do not overlap any chord in L_i. We stop the construction and set $k := i$, if

$$S := \sum_{j=1}^{i} \left\lfloor \frac{|L_j|-1}{2} \right\rfloor \geq \lambda - 1.$$

First, we consider the case $S < \lambda - 1$. In this case,

$$\begin{aligned} |Q_k| &= q(G) - \sum_{i=1}^{k} |Q_{i-1} \setminus Q_i| \\ &\geq q(G) - \sum_{i=1}^{k} 3|L_i| \\ &\geq q(G) - \sum_{i=1}^{k} 12 \left\lfloor \frac{|L_i|-1}{2} \right\rfloor \\ &> q(G) - 12\lambda. \end{aligned}$$

If $k = 0$, then G contains no triple of pairwise overlapping chords. Since

$$2 \leq 2(y-x) - 2 \leq 2l - 4,$$

the lengths of the route P that uses no chord and of the route P' that uses the two chords $v_x v_{x+l}$ and $v_y v_{y+l}$ do not differ by a multiple of $(l-1)$. Now G contains a set I of $\frac{q(G)-4}{2}$ pairwise non-overlapping chords that neither coincide with nor overlap the chords $v_x v_{x+l}$ and $v_y v_{y+l}$, and for each $0 \leq i \leq |I|$ both paths P and P' can be shortened by $i \cdot (l-1)$ using i chords in I. Hence $r(G) \geq 2(|I|+1) \geq q(G) - 2$, which is larger than the desired bound.

If $k > 0$, then G contains a route P' that uses only chords from L_1 such that the lengths of P and P' do not differ by a multiple of $(l-1)$. Since $Q_k \subseteq Q_1$, $|Q_k| \geq q(G) - 12\lambda$, and Q_k contains no triple of pairwise overlapping chords, Q_1 contains a set I of $\left\lceil \frac{q(G)-12\lambda}{2} \right\rceil$ pairwise non-overlapping chords, and we obtain $r(G) \geq q(G) - 12\lambda + 2$ as above.

It remains to consider the case $S \geq \lambda - 1$. In this case,

$$|Q_k| \geq q(G) - 12\lambda - |Q_{k-1} \setminus Q_k| \geq q(G) - 12\lambda - 3l,$$

and the set

$$\bigcup_{i=1}^{k} L_i \setminus \{e_i\}$$

contains $\lambda - 1$ disjoint sets $C_1, C_2, \ldots, C_{\lambda-1}$, such that each C_j contains two consecutive chords e_{x_j} and e_{x_j+1} for some index x_j. Let $I \subseteq Q_k$ be a maximum set of pairwise non-overlapping chords. Clearly,

$$|I| \geq \frac{q(G) - 12\lambda - 3l}{l} = \frac{q(G)}{l} - 12\frac{\lambda}{l} - 3.$$

Now we can construct a route of length

$$n - 1 + 2a - (l-1)(k+b)$$

for each pair

$$(a,b) \in \{0,1,\ldots,\lambda-1\} \times \{0,1,\ldots,|I|\}$$

as follows: Let P_b be a route of length $n-1-(l-1)(k+b)$ in G that uses the chords $c_1, c_2, \ldots, c_k$ and b chords from I. If $C_i = \{v_r v_{r+l}, v_t v_{t+l}\}$ for some $1 \leq i \leq \lambda - 1$, then replacing the subpath of P_b with endvertices $\min e_{x_i}$ and $\min e_{x_i+1}$ with the chords e_{x_i} and e_{x_i+1} and the subpath of P with endvertices $\max e_{x_i}$ and $\max e_{x_i+1}$ increases the length of P_b by exactly 2 (cf. Figure 3.3). Executing such a replacement a times for

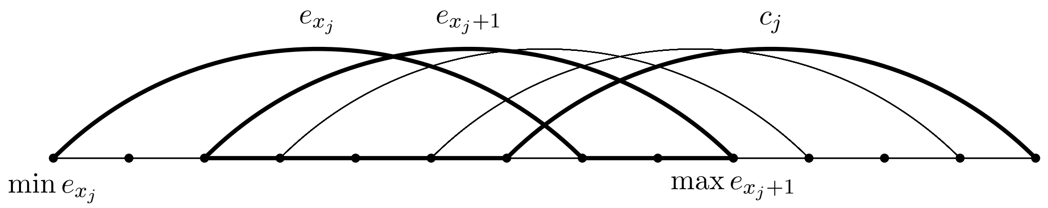


Figure 3.3: Increasing the length by 2

$0 \leq a \leq \lambda - 1$ results in a path of the desired length. Note that such replacements can be combined without conflict.

Since $0, 2, 4, \ldots, 2\lambda - 2$ are members of different residue classes modulo $l-1$, the path lengths corresponding to the pairs (a,b) are pairwise different. Hence

$$r(G) \geq \lambda(|I|+1) \geq \frac{\lambda q(G)}{l} - 12\frac{\lambda^2}{l} - 2\lambda,$$

which again implies the desired statement. $\square$

Note that, if G contains the maximum possible number $q(G) = n - l$ of chords of length l, then

$$r(G) = \begin{cases} \frac{l-1}{l} q(G) + O(l) & \text{, if } l \text{ is even,} \\ \frac{l-1}{2l} q(G) + O(l) & \text{, if } l \text{ is odd,} \end{cases}$$

i.e. up to the $O(l)$ term Theorem 14 is best-possible. Furthermore, the graphs which are excluded by the hypothesis of Theorem 14 satisfy $r(G) \geq \frac{q(G)}{2} + 1$.

3.2 Cycle Lengths in Hamiltonian Graphs

In the following discussion, G is a graph with a distinguished Hamiltonian cycle C. The edges in $E(G) \setminus E(C)$ are called *chords*. The *length* of a chord $\{u, v\}$ is $\mathrm{dist}_C(u, v)$. The length of any chord is at least 2 and at most $\lfloor n/2 \rfloor$. Let the *normalised length* of a chord $\{u, v\}$ be $\frac{\mathrm{dist}_C(u,v)}{n/2}$. Two chords $\{u, v\}$ and $\{x, y\}$ *cross*, if each of the two paths in C with endvertices u and v contains exactly one of the vertices x and y as an inner vertex.

The bound $r \geq \frac{q}{3} + 1$ of Lemma 3 in [23] and of Theorem 13 easily implies a lower bound on $s(G)$ of the right order of magnitude.

Corollary 2. *Every Hamiltonian graph G satisfies $s(G) \geq \sqrt{\frac{m(G)-n(G)}{3}}$.*

Proof. Let C be a Hamiltonian cycle of G and let $S = E(G) \setminus E(C)$ be the set of all $q := m(G) - n(G)$ chords. If S contains a subset S' of at least $\sqrt{\frac{q}{12}}$ chords of pairwise different lengths, then G contains cycles of $2q \geq \sqrt{\frac{q}{3}}$ different lengths that use at most one chord, chosen from S'. Otherwise, G contains a set L of more than $\frac{q}{\sqrt{\frac{q}{12}}} = \sqrt{12q}$ chords of some length l. Let e be an arbitrary edge of C and let $P := C - e$ be a distinguished Hamiltonian path of $G - e$. Then the elements of L are chords of P of length l or $n - l$, so a subset $L' \subseteq L$ of at least $\sqrt{3q}$ of them are P-chords of the same length. Therefore, $G - e$ contains paths of $\frac{\sqrt{3q}}{3} + 1 = \sqrt{\frac{q}{3}} + 1$ different lengths. Adding e to these paths yields $\sqrt{\frac{q}{3}} + 1$ cycles in G of different lengths. $\square$

In the remainder of the section, we are going to improve the constant by more careful arguments: Lemma 12 applies an independent argument to improve upon the naive argument above, if the average length of the chords is greater than $\frac{n}{12}$, while Lemma 13 refines the argument of Corollary 2. A combination of these bounds yields our main result.

Lemma 12. *Let G be graph with a Hamiltonian cycle C. If the average normalised length of the chords of C is β, then $s(G) > \sqrt{\beta(m(G) - n(G))}$.*

Proof. We seek a large set of chords in one of two special configurations.

If G contains a set I of q pairwise noncrossing chords, then we obtain $s(G) \geq q + 1$ by considering cycles in G that contain at most one chord, chosen from I.

If G contains a set X of q pairwise crossing chords, then we obtain $s(G) \geq q - 1$ by considering cycles that use two chords, both chosen from X, one of which is a fixed chord $x \in X$.

For any choice of a path S in C, let G_S be the graph whose vertices are the chords joining $V(S)$ and $V(C) \setminus V(S)$ and in which two vertices of G_S are adjacent, if they are crossing chords. This graph is called a *permutation graph*, since — after embedding C into the Euclidean plane such that its edges are arcs of a circle — the vertices of G_S correspond to straight lines that join two points of the circle and cross the straight line connecting the two endvertices of S. By [58], such a graph is *perfect*, so it satisfies

$\chi(G_S) = \omega(G_S)$ and, since $\chi(G)\alpha(G) \geq n$ for any graph G, $\omega(G_S) \cdot \alpha(G_S) \geq n(G_S)$. This implies that $\alpha(G_S) > \sqrt{n(G_S)} - 1$ or $\omega(G_S) > \sqrt{n(G_S)} + 1$, both of which imply $s(G) > \sqrt{n(G_S)}$ by the above arguments.

Choose one of the n sets of $\lfloor n/2 \rfloor$ consecutive vertices of C at random with equal probability. The probability that a chord of length l has exactly one endpoint in S is $\frac{2l}{n}$, i.e. the normalised length of the chord. Therefore, the expected number of chords with one endpoint in S is the sum of the normalised lengths, i.e. $\beta(m(G)-n(G))$. This implies that, for some choice of S, $n(G_S) \geq \beta(m(G)-n(G))$ and thus $s(G) > \sqrt{\beta(m(G)-n(G))}$. $\square$

Lemma 13. *Let G be a graph with a Hamiltonian cycle C. If the average normalised length of the chords of C is β, then*

$$s(G) \geq \sqrt{\frac{2}{3}\left(1 - \frac{\beta}{4}\right)(m(G) - n(G))}.$$

Proof. For any given chord of length l, the lengths of the two cycles using this and no other chord are $l+1$ and $n-l+1$. Both cycles are shorter than C, and their lengths coincide only if $l = n/2$. If t is the number of different lengths of chords in G, then the bound $s(G) \geq 2t$ is achieved by cycles using at most one chord.

If t is small, then many chords have equal length. In order to benefit from the improvement in Corollary 1, we assign to each chord of C with length l the weight $w(l) := (1+l/n)/3$. Choose an edge $e \in E(C)$ chosen uniformly at random and consider $P := C - e$ to be a distinguished Hamiltonian path of G. Any chord of C is also a chord of P; let the *P-length* of a chord $\{x,y\}$ of C be $\mathrm{dist}_P(x,y)$. The P-length of a chord e of length l is equal to l with probability $1 - l/n$. Let W be the expected value of the total weight of all chords whose length and P-length coincide. The expected number of chords of length l that contribute to W is $a_l(1-l/n)$, where a_l is the number of chords of length l. Now

$$\begin{aligned}
W &= \sum_{l \geq 2} \frac{1}{3}\left(1 + \frac{l}{n}\right) a_l \left(1 - \frac{l}{n}\right) = \frac{1}{3}\sum_{l\geq 2} a_l \left(1 - \frac{l^2}{n^2}\right) \\
&= \frac{1}{3}\left(\sum_{l\geq 2} a_l - \frac{1}{4}\sum_{l \geq 2} a_l \left(\frac{l}{n/2}\right)^2\right) \geq \frac{1}{3}\left((m(G)-n(G)) - \frac{1}{4}\sum_{l\geq 2} a_l \frac{l}{n/2}\right) \\
&= \frac{1}{3}\left((m(G)-n(G)) - \frac{1}{4}\beta(m(G)-n(G))\right) = \frac{1}{3}(m(G)-n(G))\left(1 - \frac{\beta}{4}\right).
\end{aligned}$$

For some choice of $e \in E(C)$, the total weight of the chords whose length and P-length coincide is at least W. If t is the number of lengths of chords, some particular length contributes at least W/t to this total weight. Let l be this length. We have $a_l \geq \frac{W}{tw(l)}$, so by Corollary 1 the chords of this length contribute at least W/t cycle lengths.

We now have

$$s(G) \geq \max\left\{2t,\ \frac{(m(G)-n(G))}{3t}\left(1 - \frac{\beta}{4}\right)\right\} \geq \sqrt{\frac{2}{3}\left(1 - \frac{\beta}{4}\right)(m-n)},$$

where the final inequality chooses t to minimise the maximum. $\square$

Theorem 15. *If G is a Hamiltonian graph, then $s(G) \geq \sqrt{\frac{4}{7}(m(G) - n(G))}$.*

Proof. By Lemmas 12 and 13, $s(G) \geq \sqrt{(m(G) - n(G)) \max\left\{\beta,\ \frac{2}{3}(1 - \frac{\beta}{4})\right\}}$. Choosing $\beta = 4/7$ minimises the larger lower bound. $\square$

4 Forbidden Cycles and the Independence Ratio

The independence number is a fundamental and well-studied graph parameter [49]. Calculation of $\alpha(G)$ is computationally difficult: Berman and Fujito [6] have shown that even its restriction to subcubic graphs is APX-hard; incidentally, Lemma 2 and Lemma 4 imply that this also holds if restricted further to graphs of girth at least 6. The aim of this chapter is to derive lower bounds on the *independence ratio* $\frac{\alpha(G)}{n(G)}$ in terms of the average degree $d(G)$ for connected graphs with given (odd) girth.

Graphs of given order While the research of this chapter only differentiates graphs by their density, not by their order or size alone, we mention two results in which the order of the graph is fixed. For graphs with a nontrivial odd girth condition and fixed order, increasing the number of edges forces the graph to become more structured. Therefore, unlike in our setting, in which the order is unrestricted, high average degree forces a high independence number in graphs of fixed order. This is reflected by Shearer's short proof of the inequality $\alpha(G) \geq m(G)$ for graphs of odd girth 7 in [68].

The minimum order of a triangle-free graph that forces the existence of an independent set of order t is the *Ramsey-number* $R(3,t)$. While Ajtai, Kómlos and Szemerédi proved in [2] that $R(3,t) = O\left(\frac{t^2}{\log(t)}\right)$, Kim showed in [44] that $R(3,t) = \Omega\left(\frac{t^2}{\log(t)}\right)$. Together, this roughly determines the minimum possible independence number of triangle-free graphs of given order.

Since the independence ratio $\frac{\alpha(G)}{n(G)}$ of triangle-free graphs of given odd girth and unbounded order can become arbitrarily small, we are going to discuss lower bounds of this ratio in terms of the density in the remainder of this chapter.

Bounds for arbitrary graphs Caro [13] and Wei [78] proved

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{d_G(v)+1} \tag{4.1}$$

for every graph G. In [3], Alon and Spencer show that this bound is achieved by the naive algorithm that obtains an independent set I by randomly selecting vertices of the graph as members of I and removing their closed neighbourhoods. Indeed, let G_i be the graph that arises after i neighbourhood removals for some $i \in \mathbb{N}_0$. If $v \in V(G_i)$ for some $v \in V(G)$, then the conditional probability that v is selected in the next step, provided that v is removed in the next step, is reciprocal to $|N_{G_i}(v)| \leq |N_G(v)| = d_G(v) + 1$.

Note that this argument also shows that the bound is only tight, if $N_{G_i}(v) = N_G(v)$ whenever $v \in V(G_i)$, so the only tight examples for inequality (4.1) are disjoint unions of complete graphs.

We can derive a tight bound on the independence ratio in terms of the average degree, as illustrated in Figure 4.1.

Corollary 3. *For every $r \in \mathbb{N}$ and every graph G,*

$$\frac{\alpha(G)}{n(G)} \geq \frac{2r - d(G)}{r(r+1)}.$$

Each bound is tight for disjoint unions of K_r and K_{r+1}.

Proof. Since K_r and K_{r+1} satisfy the bound assigned to r with equality, so do their disjoint unions because of its linearity. The bound assigned to a given r is stronger than all other bounds for $d(G) \in \,]r-1, r[$, so it is sufficient to prove each bound for

$$d(G) \in \; [r-1, r]. \tag{4.2}$$

For graphs of fixed order and size, the bound (4.1) is minimal if and only if no two vertex degrees differ by more than one: Indeed, if a graph G contains two vertices v and w with $d_G(w) > d_G(v) + 1$, then replacing an edge $\{w, x\}$ for $x \in N_G(w) \setminus N_G[v]$ with the edge $\{v, x\}$ lowers the value of the bounding function due to the strict convexity of the map $x \mapsto \frac{1}{x+1}$. Therefore, under the assumption (4.2), the bound (4.1) is minimal, if all vertices have degree $r-1$ or r.

Using the representation $d(G) = pr + (1-p)(r-1)$ for $p = d(G) - r + 1$, (4.1) implies

$$\frac{\alpha(G)}{n(G)} \geq p \cdot \frac{1}{r+1} + (1-p) \cdot \frac{1}{r} = \frac{d(G) - r + 1}{r+1} + \frac{r - d(G)}{r} = \frac{2r - d(G)}{r(r+1)}.$$

□

Although this lower bound is best-possible for all rational average degrees, it is not tight for any graph that contains a component of order at least $d+2$, so it appears natural to ask for the best lower bound for large connected graphs. More precisely, we

$\alpha(\mathcal{P}, d)$ define, for a class $\mathcal{P}$ of graphs, the *asymptotic independence ratio* by

$$\alpha(\mathcal{P}, d) = \lim_{n \to \infty} \left(\inf \left\{ \frac{\alpha(G)}{n(G)} \;\middle|\; G \in \mathcal{P}, d(G) \leq d, n(G) \geq n \right\} \right).$$

By the presence of tight examples of arbitrary order and rational average degree, Corol-

$\mathcal{G}$ lary 3 implies that $\alpha(\mathcal{G}, d)$ is the linear interpolation of the map $d \mapsto \frac{1}{d+1}$, where $\mathcal{G}$ is the class of all graphs.

Improvement for connected graphs As an improvement of an earlier result due to Harant and Schiermeyer [33], Harant and Rautenbach [31] proved that for each connected graph G, there exists a positive integer $k \in \mathbb{N}$ and a function $f\colon V(G) \to \mathbb{N}_0$ such that $f(u) \leq d_G(u)$ for $u \in V(G)$,

$$\alpha(G) \geq k \geq \sum_{u \in V(G)} \frac{1}{d_G(u) + 1 - f(u)}, \quad \text{and} \quad \sum_{u \in V(G)} f(u) = 2(k-1). \tag{4.3}$$

Corollary 4. *For every $r \in \mathbb{N} \setminus \{1\}$ and every connected graph G, $\frac{\alpha(G)}{n(G)} \geq b_r\big(\tilde{d}(G)\big)$ with*

$$b_r(d) = \frac{2r - d}{(r-1)(r+2)}$$

and

$$\tilde{d}(G) = \frac{2(m(G)+1)}{n(G)}.$$

This bound is satisfied with equality for connected graphs that arise from disjoint unions of K_{r-1} and K_r components by adding bridges.

Proof. The tightness result can be easily verified. For given r, the lower bound b_r is at least as strong as all other bounds b_i for $d \in \left[r - 1 + \frac{2}{r},\ r + \frac{2}{r+1}\right]$, while b_{r+1} is a stronger bound, if $d > r + \frac{2}{r+1}$ and b_{r-1} is stronger, if $r > 2$ and $d < r - 1 + \frac{2}{r}$. Hence we may assume that $\tilde{d}(G)$ is contained in that interval.

In equation (4.3), we may assume $k < n(G)\frac{2r - \tilde{d}(G)}{(r-1)(r+2)}$ because otherwise the desired bound follows immediately, so we obtain $\alpha(G) \geq \sum_{i=1}^{n} \frac{1}{x_i}$ with $x_i \in \mathbb{N}$ and

$$\sum_{i=1}^{n} x_i = n(G) + 2m(G) - 2(k-1) = n(G)\left(1 + \tilde{d}(G) - 2\frac{2r - \tilde{d}(G)}{(r-1)(r+2)}\right).$$

For $\tilde{d}(G) = r - 1 + \frac{2}{r}$, the sum of the denominators is $n(G)r$, so the bound in (4.3) is minimal if all denominators are r, which implies the desired bound in this case. Similarly, if $\tilde{d}(G) = r + \frac{2}{r+1}$, the sum of the denominators is $n(G)(r+1)$, which again implies the bound. By the convexity of the function $x \mapsto \frac{1}{x}$, the bound in (4.3) is minimised, if no two values of x_i differ by more than one, so for any $\tilde{d}(G) \in \left[r - 1 + \frac{2}{r},\ r + \frac{2}{r+1}\right]$, all summands of this bound are either $\frac{1}{r}$ or $\frac{1}{r+1}$. Since the fraction of those summands that are of the form $\frac{1}{r}$ is linear in $\tilde{d}(G)$, the resulting bound is linear. Since it coincides with the linear function b_r at both margins of the interval $\left[r - 1 + \frac{2}{r},\ r + \frac{2}{r+1}\right]$, the proof is complete. ☐

Let $\mathcal{G}_{\text{conn}}$ be the class of connected graphs. The corollary implies that $\alpha(\mathcal{G}_{\text{conn}}, d)$ is the linear interpolation of the values $\alpha(\mathcal{G}_{\text{conn}}, d) = \frac{1}{r}$ for $d = r - 1 + \frac{2}{r}$ with $r \in \mathbb{N} \setminus \{1\}$. $\mathcal{G}_{\text{conn}}$

Bounds for triangle-free graphs For triangle-free graphs G, Shearer [67] has proved

$$\alpha(G) \geq \sum_{u \in V(G)} f_{\mathrm{Sh}}(d_G(u)) \tag{4.4}$$

where $f_{\mathrm{Sh}}(0) = 1$ and $f_{\mathrm{Sh}}(d) = \frac{1+(d^2-d)f_{\mathrm{Sh}}(d-1)}{d^2+1}$ for $d \in \mathbb{N}$.

The function f_{Sh} has the best-possible order of magnitude $f_{\mathrm{Sh}}(d) = \Omega\left(\frac{\log d}{d}\right)$. It also serves as a lower bound for the expected cardinality of an independent vertex set obtained by iteratively choosing a random vertex and deleting its closed neighbourhood. Denley [17] and Shearer [68] also gave bounds with the best-possible order of magnitude for graphs with specified odd girth. By a similar argument as in Corollary 3, using the convexity of f_{Sh} (cf. Lemma 1 in [67]), we obtain that $\alpha(\mathcal{G}_{\Delta\text{-free}}, d)$ is at least the linear
$\mathcal{G}_{\Delta\text{-free}}$ interpolation of the values $\alpha(\mathcal{G}_{\Delta\text{-free}}, d) = f_{\mathrm{Sh}}(d)$ for integral d, where $\mathcal{G}_{\Delta\text{-free}}$ denotes the class of triangle-free graphs. Suitable disjoint unions of complete graphs of orders 1 and 2 and cycles of length 5 imply that $\alpha(\mathcal{G}_{\Delta\text{-free}}, d)$ is exactly this lower bound for $d \leq 2$. Results by Kreher and Radziskowski [43] imply $\alpha(\mathcal{G}_{\Delta\text{-free}}, d) = \frac{10-d}{20}$ for $d \in \left[2, \frac{10}{3}\right]$ and $\alpha(\mathcal{G}_{\Delta\text{-free}}, d) = \frac{12-d}{26}$ for $d \in \left[\frac{10}{3}, 4\right]$. In [38], Jones and Locke have given an efficient algorithm that achieves the independence ratio of $\frac{12-d}{26}$ based on the earlier work of Jones in [37]. Tight examples have been independently discovered by Kreher and Radziskowski in [43] and Jones in [37]. Some of them appear in the following as *difficult blocks*.

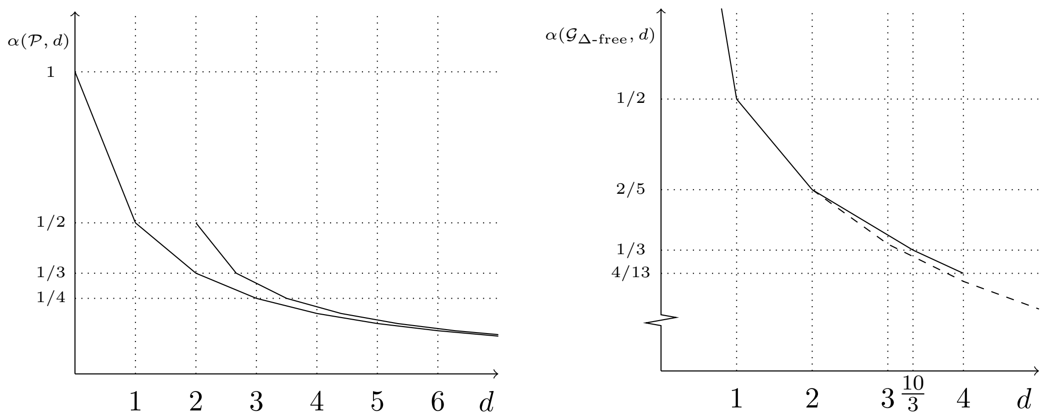


Figure 4.1: The left graph shows the $\alpha(\mathcal{P}, d)$ for $\mathcal{P} \in \{\mathcal{G}, \mathcal{G}_{\mathrm{conn}}\}$. The right graph shows the exact value of $\alpha(\mathcal{G}_{\Delta\text{-free}}, d)$ for $d \leq 4$ and the lower bound based on (4.4).

Improvements for connected triangle-free graphs For triangle-free graphs G with maximum degree at most 3, Heckman and Thomas [35] proved the best-possible inequality

$$\alpha(G) \geq \frac{1}{7}\left(4n(G) - m(G) - \lambda(G)\right), \tag{4.5}$$

where $\lambda(G)$ counts the number of so-called *difficult components* of G, which will be defined later. Their result implies

$$\alpha(G) \geq \frac{5}{14} n(G)$$

for triangle-free graphs G of maximum degree at most 3, which was originally conjectured by Albertson, Bollobás, and Tucker [1] and first proved by Staton [70] (cf. also [4, 22, 37, 30, 34]). In [80], Zhu proved a theorem that generalises the bound of Heckman and Thomas and implies that every connected triangle-free graph with maximum degree at most 3 contains an induced bipartite subgraph of order at least $\frac{2}{7}(4n(G) - m(G) - 1)$.

In the next section, we are going to prove that — after a suitable modification — inequality (4.5) still holds even if we drop the maximum degree condition. As a consequence, we determine the exact value of $\alpha(\mathcal{G}_{\Delta\text{-free}} \cap \mathcal{G}_{\text{conn}}, d)$ for $d \leq \frac{10}{3}$.

Most results of this chapter are based on [48].

4.1 Triangle-free Graphs

4.1.1 Average degrees below 10/3

In order to state the result of Heckman and Thomas [35], we need to define $\lambda(G)$ for triangle-free graphs G of maximum degree at most 3.

Heckman and Thomas call a graph a *difficult block*, if it is one of the two graphs G_2 and G_3 in Figure 4.2. Furthermore, they call a graph G *difficult*, if every block of G is either difficult or is an edge between two difficult blocks. For a graph G, $\lambda(G)$ counts the number of components of G that are difficult graphs.

Theorem 16 (Heckman and Thomas [35])**.** *If G is a triangle-free graph of maximum degree at most* 3*, then*

$$\alpha(G) \geq \frac{1}{7}\left(4n(G) - m(G) - \lambda(G)\right).$$

We will show that, in a suitably modified form, Theorem 16 remains true without the bound on the maximum degree. Our approach will closely follow the method from [35]. A main ingredient of our proof are further difficult blocks, which are special cases of *general pentagons* as defined in [72].

Since we need to extend the definition of Heckman and Thomas to graphs that are not necessarily subcubic and are going to use various difficult blocks for subsequent results, we define

Definition 7. *Let $\mathcal{S}$ be a set of* 2*-connected graphs. We refer to the elements of $\mathcal{S}$ as* $\mathcal{S}$-difficult blocks*. A connected graph which after the removal of all its bridges is the union of vertex-disjoint $\mathcal{S}$-difficult blocks is called an* $\mathcal{S}$-difficult component*. A graph in which each component is $\mathcal{S}$-difficult is called an* $\mathcal{S}$-difficult graph*. For a graph G, let $\lambda_{\mathcal{S}}(G)$ denote the number of $\mathcal{S}$-difficult components of G.* $\lambda_{\mathcal{S}}(G)$

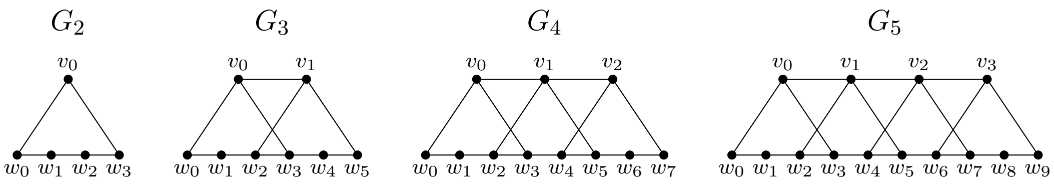


Figure 4.2: The smallest four $\mathcal{G}_1$-difficult blocks.

For the extension of Theorem 16, we define a class $\mathcal{G}_1 = \{G_k : k \in \mathbb{N} \setminus \{1\}\}$ of graphs, which has been independently defined in [29], [37], and [43] before. The first four members of this sequence are shown in Figure 4.2. In general, these difficult blocks are defined by $\mathcal{G}_1$, G_k

$$V(G_k) = \{w_i : i \in \{0, 1, \ldots, 2k-1\}\} \cup \{v_i : i \in \{0, 1, \ldots, k-2\}\} \quad \text{and}$$

$$E(G_k) = \{w_i w_{i+1} : i \in \{0, \ldots, 2k-2\}\} \cup \{v_i v_{i+1} : i \in \{0, \ldots, k-3\}\}.$$

Note that for $k \geq 3$, G_{k-1} is an induced subgraph of G_k with

$$V(G_k) = V(G_{k-1}) \cup \{v_{k-2}, w_{2k-2}, w_{2k-1}\} \quad \text{and}$$

$$E(G_k) = E(G_{k-1}) \cup \{w_{2k-3}w_{2k-2},\ w_{2k-2}w_{2k-1},\ w_{2k-1}v_{k-2},\ v_{k-2}v_{k-3}, v_{k-3}w_{2k-2}\}.$$

In analogy to Theorem 16, our result is as follows.

$f(G)$ **Theorem 17.** *If G is a triangle-free graph, then*

$$\alpha(G) \geq f(G) := \frac{1}{7}\left(4n(G) - m(G) - \lambda_{\mathcal{G}_1}(G)\right).$$

The bound in Theorem 17 is best-possible for all $\mathcal{G}_1$-difficult graphs (cf. Lemma 15 (i) below). Furthermore, it is clearly also best-possible for all graphs for which the bound in Theorem 16 is best-possible. These graphs have been characterised by Heckman [34].

Before we prove Theorem 17, we establish some useful properties of the $\mathcal{G}_1$-difficult blocks. Properties (i) and (ii) have already been shown in [43], and we include the proofs for the sake of completeness.

Lemma 14. *Let $k \geq 2$.*

(i) For any $k \geq 2$ and any two vertices $v, w \in G_k$ of degree two, there is an automorphism that maps v to w.

(ii) G_k has order $3k-1$, size $5k-5$, and independence number k.

(iii) For every two vertices u and v of G_k, the graph G_k has a maximum independent set containing neither u nor v.

(iv) If $abcd$ is an induced path of G_k such that the vertices b and c have degree 2 and $u \notin \{a, d\}$ is a vertex of G_k, then the graph G_k has a maximum independent set containing a and d but not u.

Proof. (i) For $k = 2$, this is obvious. Otherwise, let f_1 be the "mirror" automorphism defined by $v_i \mapsto v_{k-2-i}$ and $w_i \mapsto w_{2k-1-i}$, and let f_2 be the automorphism that exchanges w_0 with w_1 and v_0 with w_2 and is identical on all other vertices. Then the images of any of the four vertices of degree two under the four automorphisms id, $f_1, f_2, f_1 \circ f_2$ are pairwise different.

(ii) Order and size of G_k are obvious from the definition, and the vertices w_{2i} form an independent set of order k. Suppose that k is the minimum integer for which G_k contains a larger independent set I. By (i), $G' := G_k - N_{G_k}[w] \cong G_{k-1}$ for any vertex w of degree 2, so by the minimality of k, $|I \cap N_{G_k}[w]| = 2$. For $w = w_0$, this implies $\{v_0, w_1\} \subseteq I$. For $w = w_1$, it implies $\{w_0, w_2\} \subseteq I$, a contradiction to the independence of I.

(iii) We prove this by induction on k. For $k = 2$, the statement is easily verified. Therefore, let $k \geq 3$. Let w be a vertex of degree 2 other than u or v and G' be defined as in (ii). By the induction hypothesis, G' contains an independent set I' of order $k-1$ that contains neither u nor v. Then we may choose $I := I' \cup \{w\}$.

(iv) Since the result is easy to check for $k = 2$, we assume that $k \geq 3$. By (i), we may assume $b = w_{2k-1}$, which implies $a = v_{k-2}$ and $d = w_{2k-3}$. Since deleting $N_{G_k}[a] \cup N_{G_k}[d]$ from G_k results in G_{k-2}, the statement follows from (iii).

□

Now we can prove Theorem 17 in analogy to the proof of Heckman and Thomas [35] by showing that each counterexample could be reduced to a smaller instance by excising a vertex set whose choice depends on the order and degree sum of a smallest neighbourhood.

We begin with some generic observations that can be applied to prove similar bounds with respect to other classes of difficult blocks. The property in part (iii) of the Lemma can be reformulated as $r(\mathcal{G}_1) \geq 2$, where the *robustness* $r(G)$ of a graph G is defined by $r(G)$

$$r(G) := \max\{k \in \mathbb{N}_0 \colon \forall M \subseteq V(G) \colon |M| \leq k \Rightarrow \alpha(G) = \alpha(G - M)\},$$

and the *robustness* of a graph class $\mathcal{G}$ is the minimum over the robustness of its members

$$r(\mathcal{G}) := \min\{r(G) \colon G \in \mathcal{G}\}.$$

Note that for each vertex v in a triangle-free graph G, removal of $N_G[v]$ destroys $s_G(v) := \sum_{w \in N_G(v)} d_G(w)$ edges. $s_G(v)$

Lemma 15. *Let $a, b \in \mathbb{N}$ with $a \leq b + 1$ and let $\mathcal{G}$ be a set of 2-connected graphs with $r(\mathcal{G}) \geq 1$ for which the inequality* $f(G)$

$$\alpha(G) \geq f(G) := (an(G) - m(G) - \lambda_{\mathcal{G}}(G))/b \tag{4.6}$$

holds with equality. If G is a vertex-minimal triangle-free counterexample to inequality (4.6)*, then*

(i) G is 2-edge-connected and not $\mathcal{G}$-difficult.

(ii) If H is a connected $\mathcal{G}$-difficult induced subgraph of G, then the number $\varphi(H)$ of edges with exactly one endvertex in $V(H)$ is at least $(r(\mathcal{G})-1)x+2$, where x is the number of difficult blocks in H. $\varphi(H)$

(iii) If $G' = G - X$ for some vertex set $X \subseteq V(G)$, $\alpha(G) - \alpha(G') \geq \alpha_X$ and $m(G) - m(G') \leq m_X$, then

$$\varphi(X) \geq (r(\mathcal{G})+1) \cdot (b \cdot \alpha_X + m_X - a|X| + 1).$$

(iv) For each $v \in V(G)$, $s_G(v) \leq d_G(v) + \frac{r(\mathcal{G})+1}{r(\mathcal{G})}\big((a-1)d_G(v) + a - b - 1\big)$.

Proof. In the following proof, we write difficult instead of $\mathcal{G}$-difficult and λ instead of $\lambda_{\mathcal{G}}$. Note that, as a minimum counterexample, G is connected.

(i) Suppose that G contains a bridge $e = \{u, v\}$. Since G is vertex-minimal, neither of the components that arises by removing e violates (4.6), so both u and v are contained in every maximum independent set of their respective components, in particular $\alpha(G-u) = \alpha(G) - 1$. Removing the vertex u increases the number of difficult components by at most $d_G(u)-1$, reduces the size by $d_G(u)$ and reduces the order by one, so it reduces the value of f by at most $\frac{a-d_G(u)+(d_G(u)-1)}{b} = \frac{a-1}{b} \leq 1$, contradicting the minimality of G. Therefore, G is 2-edge-connected. By the definition of a difficult graph, this implies that G can only be difficult, if it is 2-connected, but all 2-connected graphs satisfy (4.6).

Since a minimum counterexample is 2-connected and the 2-connected difficult graphs are elements of $\mathcal{G}$, which satisfy (4.6) with equality, it cannot be a difficult graph.

(ii) For $x = 1$, let S be the set of vertices of H that are adjacent to a vertex outside H. Suppose for contradiction that $\varphi(H) \leq r(\mathcal{G})$. Then $\alpha(G[V(H) \setminus S]) = \alpha(H)$, so removing H from G reduces the independence number by $\alpha(H)$. On the other hand, this removal introduces at most $\varphi(H) - 1$ difficult components, reduces the order by $n(H)$ and the size by $m(H) + \varphi(H)$, so the bound drops by at most

$$\frac{a \cdot n(H) - (m(H) + \varphi(H)) + (\varphi(H) - 1)}{b} = f(H) = \alpha(H),$$

which contradicts the minimality of G.

For general x, the above argument implies that each of the x blocks of H has at least $r(\mathcal{G})+1$ outgoing edges; $2(x-1)$ of these edges are bridges of H (edges are counted twice iff they are bridges), so $\varphi(H) \geq (r(\mathcal{G})+1)x - 2(x-1) = (r(\mathcal{G})-1)x + 2$.

(iii) If the inequality does not hold, then $\left\lfloor \frac{\varphi(X)}{r(\mathcal{G})+1} \right\rfloor \leq b \cdot \alpha_X + m_X - a \cdot |X|$. Since, by (ii), each difficult component of G' contributes at least $r(\mathcal{G})+1$ edges to the cut induced by X in G, we have $\lambda(G') \le b \cdot \alpha_X + m_x - a \cdot |X|$. Hence,

$$\begin{aligned}
\alpha(G') &\le \alpha(G) - \alpha_X \;<\; f(G) - \alpha_X \;=\; \frac{an(G) - m(G)}{b} - \alpha_X \\
&\le \frac{a \cdot n(G') - m(G') + a \cdot |X| - b \cdot \alpha_X - m_X}{b} \;\le\; \frac{an(G') - m(G') - \lambda(G')}{b} \\
&= f(G'),
\end{aligned}$$

which contradicts the minimality of G.

(iv) Removing $N_G[v]$ from G reduces the independence number by at least one, destroys $d_G(v) + 1$ vertices and $s_G(v)$ edges, and introduces at most $\left\lfloor \frac{s_G(v) - d_G(v)}{r(\mathcal{G})+1} \right\rfloor$ difficult components, since $\varphi(H) \ge r(\mathcal{G})+1$ for every difficult component H of the remaining graph by (ii). Therefore, the bound drops by at most

$$\begin{aligned}
&\frac{a\,(d_G(v)+1) - s_G(v) + \left\lfloor \frac{s_G(v)-d_G(v)}{r(\mathcal{G})+1} \right\rfloor}{b} \\
=\; &\frac{a + (a-1)d_G(v) - \left\lceil \frac{r(\mathcal{G})}{r(\mathcal{G})+1} \big(s_G(v) - d_G(v)\big) \right\rceil}{b},
\end{aligned}$$

which contradicts the minimality of G, if

$$\begin{aligned}
&(a-1)d_G(v) + (a-b) \le \left\lceil \frac{r(\mathcal{G})}{r(\mathcal{G})+1} \big(s_G(v) - d_G(v)\big) \right\rceil \\
\Leftrightarrow\; &\frac{r(\mathcal{G})}{r(\mathcal{G})+1} (s_G(v) - d_G(v)) > (a-1)d_G(v) + (a-b) - 1 \\
\Leftrightarrow\; &s_G(v) > d_G(v) + \frac{r(\mathcal{G})+1}{r(\mathcal{G})} \big((a-1)d_G(v) + a - b - 1\big).
\end{aligned}$$

□

We can now apply this result to prove the main theorem, Theorem 17. Throughout the proof, we write λ for $\lambda_{\mathcal{G}_1}$ and difficult for $\mathcal{G}_1$-difficult.

Proof of Theorem 17: In order to obtain a contradiction, we assume that G is a counterexample of minimum order. Since $r(\mathcal{G}_1) = 2$ by Lemma 14 (iii), Lemma 15 (ii) implies that for any difficult subgraph H of G, the size of the cut induced by H satisfies $\varphi(H) \ge 3$ with equality only possible if H is a difficult block. Lemma 15 (i) implies that G is 2-edge-connected and not difficult. In particular, $\delta(G) \ge 2$.

Case 1: $\delta(G) = 2$. Let $v \in V(G)$ be a vertex of degree 2. Lemma 15 (iii) implies $s_G(v) \le 5$.

Subcase 1.A: $s_G(v) = 4$. In this case, both neighbours of v have degree 2. Let $N_G(v) = \{u, w\}$, $N_G(u) = \{v, u'\}$ and $N_G(w) = \{v, w'\}$.

If $u' = w'$, then the removal of $X := N_G[u] \cup N_G[w]$ reduces the independence number by $\alpha_X = 2$ and the size by $m_X = 2 + d_G(w')$. Lemma 15 (iii) yields $\varphi(X) \geq 3(7\alpha_X + m_X - 4|X| + 1) = 3(d_G(w') + 1)$, a contradiction to $\varphi(X) = d_G(w') - 2$. Therefore, $u' \neq w'$.

If u' and w' are not adjacent, then the graph $G' := G - \{u, w\} + \{vu', vw'\}$ is triangle-free. Since $\alpha(G') = \alpha(G) - 1$ and $f(G') \leq f(G) - \frac{4 \cdot 2 - 2 + 1}{7}$, G' is a smaller counterexample, a contradiction. Therefore, u' and v' are adjacent.

Now the removal of $X := N_G[u] \cup N_G[w]$ reduces the independence number by $\alpha_X = 2$ and the size by $m_X = d_G(u') + d_G(w') + 1$, so Lemma 15 (iii) yields $\varphi(X) \geq 3(7\alpha_X + m_X - 4|X| + 1) = 3(d_G(u') + d_G(w') - 4)$. As $\varphi(X) = d_G(u') + d_G(w') - 4$, this implies $d_G(u') = d_G(w') = 2$, i.e. $G \cong G_2$, a contradiction to Lemma 15 (i).

Subcase 1.B: $s_G(v) = 5$. Let $N_G(v) = \{a, c\}$ with $d_G(a) = 3$ and $d_G(c) = 2$. If $G' := G - N_G[v]$ contains no difficult component, then the $f(G') = f(G) - \frac{4 \cdot 3 - 5}{7} = f(G) - 1$, a contradiction to the minimality of G. Therefore, G' contains a difficult component, and $\alpha(G') = \alpha(G) - 1$. Since $\varphi(N_G[v]) = 3$, Lemma 15 (ii) implies that G' is a difficult block. If $G' \cong G_2$, then $G \cong G_3$, a contradiction to Lemma 15 (i). Therefore, we may assume that $G' = G_k$ for some $k \geq 3$, and we are going to show that this implies $G \cong G_{k+1}$.

By the contradiction in Subcase 1.A, the neighbour of c different from v has degree 3 in G, so it has degree 2 in G'. By Lemma 14 (i), we may assume $N_G(c) = \{v, w_{2k-1}\}$. Since $s_G(w_{2k-2}) \geq 6$, w_{2k-2} cannot have degree two in G by the contradiction in Subcase 1.A, so $w_{2k-2} \in N_G(a)$. Since G is triangle-free, $w_{2k-3} \notin N_G(a)$.

Suppose that a is not adjacent to v_{k-2} in G. Then Lemma 14 (iv) — applied to G', the induced path $w_{2k-3}w_{2k-2}w_{2k-1}v_{k-2}$ of G', and the unique vertex $u \in N_G(a) \setminus \{v, w_{2k-2}\}$ — implies the existence of a maximum independent set I' of G' that contains w_{2k-3} and v_{k-2} but not u. Now $I' \cup \{a, c\}$ is an independent set of G, which implies the contradiction $\alpha(G) \geq 2 + \alpha(G')$. Hence a is adjacent to f in G, i.e. $N_G(a) = \{v, w_{2k-2}, v_{k-2}\}$.

The identification of a with v_{2k-1}, of v with w_{2k+1} and of c with w_{2k} shows that G is isomorphic to G_{k+1}, again contradicting Lemma 15 (i).

Case 2: $\delta(G) \geq 3$. In this case, every difficult induced subgraph H of G satisfies $\varphi(H) \geq 4$, since H has at least four vertices of degree 2 while G has none.

Suppose first that $\delta(G) = d \geq 4$ and let u be a vertex of minimum degree. Clearly, $s_G(u) \geq d^2$. Removing the closed neighbourhood $N_G[u]$ of u reduces the independence number by at least 1. On the other hand, it destroys $d + 1$ vertices and $s_G(u)$ edges, while it introduces at most $\frac{s_G(u) - d}{4}$ difficult components, since the cut induced by each of these components contains at least four edges. Therefore

$$\begin{aligned} f(G) - f(G') &\leq \frac{4(d+1) - s_G(u) + \frac{s_G(u) - d}{4}}{7} = \frac{3d + 4 - \frac{3}{4}(s_G(u) - d)}{7} \\ &\leq \frac{3d + 4 - \frac{3}{4}(d^2 - d)}{7} \leq 1, \end{aligned}$$

which contradicts the minimality of G.

Suppose now that G is not cubic. Then there exists a vertex u of degree 3 with $s_G(u) \geq 10$. Removing the closed neighbourhood $N_G[u]$ reduces the independence number by at least 1. On the other hand, it destroys 4 vertices and $s_G(u)$ edges while introducing at most $\left\lfloor \frac{s_G(u)-3}{4} \right\rfloor$ difficult components. Therefore,

$$f(G) - f(G') \leq \frac{16 - s_G(u) + \left\lfloor \frac{s_G(u)-3}{4} \right\rfloor}{7} \leq \frac{13 - \left\lceil \frac{3}{4}(10-3) \right\rceil}{7} = 1,$$

which contradicts the minimality of G.

Since, by Theorem 16, the statement holds for cubic graphs, the proof of Theorem 17 is complete.[1] $\square$

Heckman and Thomas [35] have described a linear time algorithm that determines an independent set of an order as guaranteed by Theorem 16 in a given triangle-free graph of maximum degree at most 3. The proof of Theorem 17 easily yields a polynomial time algorithm that determines an independent set of an order as guaranteed by Theorem 17 in a given triangle-free graph, since the excision arguments correspond to reduction steps in an obvious recursive procedure and it is possible to check in polynomial time whether a given graph is $\mathcal{G}_1$-difficult.

Finally, Theorem 17 allows to determine the following bound.

Corollary 5.

$$\alpha(\mathcal{G}_{\Delta\text{-free}} \cap \mathcal{G}_{conn}, d) \geq \frac{8-d}{14}$$

with equality for $\frac{12}{5} \leq d \leq \frac{10}{3}$.

Proof. Theorem 17 implies that $\alpha(\mathcal{G}_{\Delta\text{-free}} \cap \mathcal{G}_{\text{conn}}, d)$ has at least the given values. That $\alpha(\mathcal{G}_{\Delta\text{-free}} \cap \mathcal{G}_{\text{conn}}, d)$ is not larger for $\frac{12}{5} \leq d \leq \frac{10}{3}$ follows by considering connected $\mathcal{G}_1$-difficult graphs. $\square$

Figure 4.3 illustrates the result from Corollary 5 (values for $d \leq \frac{12}{5}$ are due to Corollary 3 below).

4.1.2 Average degrees beyond 10/3

As mentioned in the introduction, results by [43] and [38] imply

$$\forall d \in \left[\tfrac{10}{3}, 4\right]: \quad \alpha(\mathcal{G}_{\Delta\text{-free}}, d) = \frac{12-d}{26}.$$

This bound is achieved by disjoint unions of the quartic Ramsey graph $Ci_{13}[1,5]$ and a family $\mathcal{X} := \{X_k : k \geq 4\}$ of graphs of average degree $\frac{10}{3}$, where X_k is the graph that arises from two cycles $v_0v_1 \ldots v_{k-1}$ and $w_0w_1 \ldots w_{2k-1}$ by adding the edges $\{v_i, w_{2i}\}$ and $\{v_i, w_{2i+3 \mod 2k-1}\}$ for each $i \in \{0, \ldots, k-1\}$ (cf. Figure 4.4). Note that the difficult $\mathcal{X}$, X_k

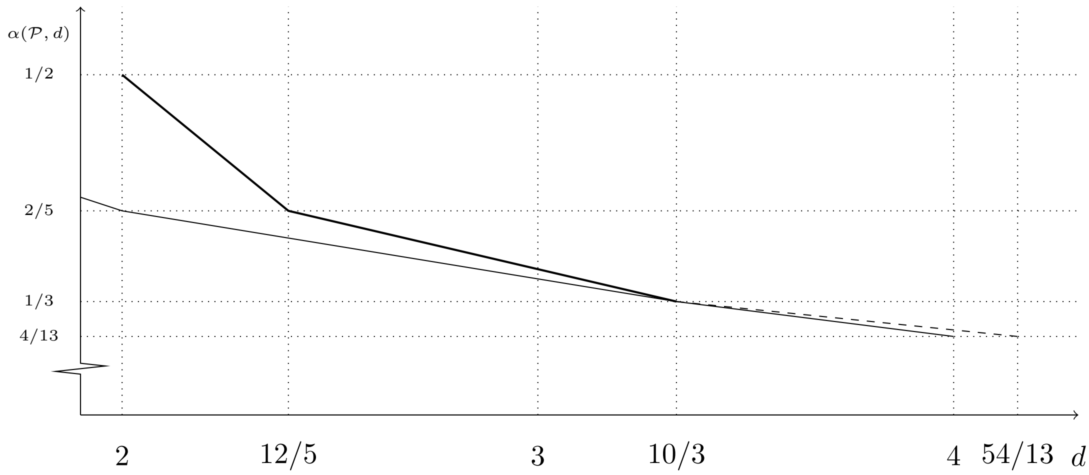


Figure 4.3: The upper line shows the value of $\alpha(\mathcal{G}_{\Delta\text{-free}} \cap \mathcal{G}_{\text{conn}}, d)$ for $d \in \left[2, \frac{10}{3}\right]$. For comparison, the lower line shows the value of $\alpha(\mathcal{G}_{\Delta\text{-free}}, d)$ for $d \leq 4$, and the dashed line shows the expected value of $\alpha(\mathcal{G}_{\Delta\text{-free}} \cap \mathcal{G}_{\text{conn}}, d)$ for $d \in [\frac{10}{3}, \frac{54}{13}]$ according to Conjecture 1.

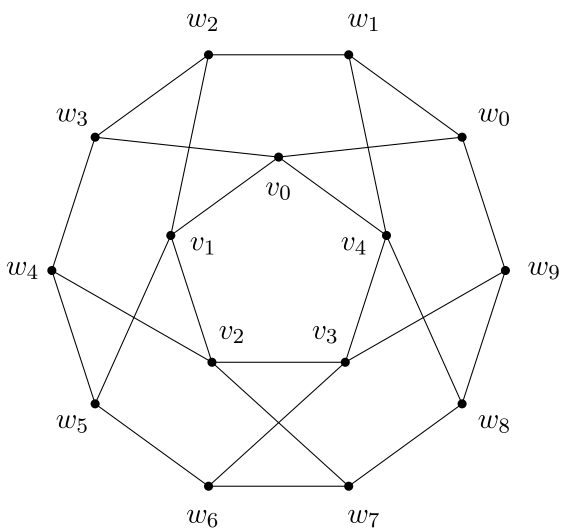


Figure 4.4: X_5

block G_k in Figure 4.2 is the subgraph of X_k that arises by removing the vertex v_{k-1} and the edge $\{w_0, w_{2k-1}\}$.

In view of the asymptotic independence ratio of connected graphs without girth restriction and Corollary 17, it seems natural to assume that the asymptotic independence ratio of connected triangle-free graphs is governed by connected graphs that arise from the disjoint union of $Ci_{13}[1,5]$ and elements of $\mathcal{X}$ by adding bridges: Let $\mathcal{G}_2 = \{Ci_{13}[1,5]\}$. $\mathcal{G}_2$

Conjecture 1. *If G is a triangle-free graph, then*

$$\alpha(G) \geq \frac{1}{16}\left(7n(G) - m(G) - \lambda_{\mathcal{G}_2}(G)\right).$$

This statement would determine the asymptotic independence ratio for connected graphs for average degrees $d \in \left[\frac{10}{3}, 4\frac{2}{13}\right]$ (cf. Figure 4.3), since it is tight for both aforementioned graph classes. The following Lemma supports the conjecture.

Lemma 16. *$r(\mathcal{G}_2) = 4$, and any minimal counterexample H to Conjecture 1 satisfies $\delta(H) \geq 3$.*

Proof. Since $G := Ci_{13}[1,5]$ satisfies $r(G) \leq \alpha(G) = 4$, it suffices to show that for each choice of a set X of four vertices in G, there exists an independent set of size 4 that avoids X. Denote the vertex set of G by $\mathbb{Z}/13\mathbb{Z}$ with $\{x,y\} \in E(G) \;\Leftrightarrow\; x - y \in \{\pm 1, \pm 5\}$. We may assume w.l.o.g. that $0 \in X$.

Consider the three disjoint independent sets $I_j := \{j, j+3, j+6, j+9\}$ for $j \in \{1,2,3\}$. If there is a j such that I_j and X are disjoint, then we can select I_j as the desired independent set, so we may assume that each I_j contains exactly one element of X. Similarly, the three independent sets $J_j := \{j, j+2, j+4, j+6\}$ for $j \in \{2,5,10\}$ are disjoint, so we may assume that each J_j contains exactly one element of X.

Let x be the element of $I_2 \cap X$. Since the automorphism of $G[\mathbb{Z}/13\mathbb{Z}^*]$ given by $i \mapsto 13 - i$ maps 2 to 11 and 5 to 8, we may assume $x \in \{2,5\}$.

Suppose first $x = 2$. Then X is disjoint to $Y_2 := (I_2 \cup J_2) \setminus \{2\} = \{4,5,6,8,11\}$. Therefore, one of the two remaining elements of X is contained in $I_1 \setminus Y_2 = \{1,7,10\}$ and the other one in $I_3 \setminus Y_2 = \{3,9,12\}$. Similarly, one of them is contained in $J_5 \setminus Y_2 = \{7,9\}$ and the other one in $J_{10} = \{10,12,1,3\}$. This implies $X \setminus \{0,2\} \in \big\{\{7,3\},\{7,12\},\{9,10\},\{9,1\}\big\}$. In each of these four cases, X is disjoint to one of the independent sets $\{4,6,8,10\}$ and $\{1,3,5,7\}$.

Suppose now $x = 5$. Then X is disjoint to $Y_5 := (I_2 \cup J_5) \setminus \{5\} = \{2,7,8,9,11\}$. Therefore, one of the remaining elements of X is contained in $I_1 \setminus Y_5 = \{1,4,10\}$ and the other one in $I_3 \setminus Y_5 = \{3,6,12\}$. Also, one of them is contained in $J_2 \setminus Y_5 = \{4,6\}$ and the other one in $J_{10} = \{10,12,1,3\}$. This implies $X \setminus \{0,5\} \in \big\{\{4,3\},\{4,12\},\{6,1\},\{6,10\}\big\}$. In each of these four cases, X is disjoint to one of the independent sets $\{6,8,10,12\}$, $\{1,3,7,10\}$ and $\{2,4,8,11\}$. This finishes the proof of $r(G) \geq 4$.

In order to prove the statement on the minimum degree, note that by Lemma 15 (iii), $s_H(v) \leq \frac{17d_H(v)-25}{2}$. H contains no vertex of v degree less than 2, since in this case the

[1]The only argument missing for a proof that does not rely on Theorem 16 is given by Claim 6 in [35].

inequality yields $s_H(v) < 0$. For a vertex of degree 2, we obtain $s_H(v) \leq 4$, so H is a cycle, but cycles other than the triangle satisfy the conjecture. □

Instead of proving Conjecture 1 in general, it may be easier to restrict it to graphs of maximum degree at most four. Note that in this context, excisions never produce difficult components.

Lemma 17. *Let G be a minimum counterexample to Conjecture 1 with $\Delta(G) \leq 4$.*

(i) Every vertex of degree 3 has at most two neighbours of degree 4.

(ii) Every C_4 subgraph in G contains two adjacent vertices of degree 4.

(iii) If G contains an $K_{2,3}$ subgraph H with partite sets $\{v_1, v_2\}$ and $\{w_1, w_2, w_3\}$, then each v_i has a fourth neighbour v_i' such that v_1' and v_2' are adjacent vertices of degree three, and exactly one element of $\{w_1, w_2, w_3\}$ has degree three.

Proof. (i) If $d_G(v) = 3$ and $s_G(v) = 12$, then removing $N_G[v]$ lowers the bounding function by $\frac{4 \cdot 7 - 12}{16} = 1$ and the independence number by at least 1, contradicting the minimality of G.

(ii) Let C be such a cycle, and let $v_1 v_2 v_3 v_4$ be its vertices. Suppose $d_G(v_1) = d_G(v_3) = 3$. If $N(v_1) \neq N(v_3)$, then removing $X := N_G[v_1] \cup N_G[v_3]$ destroys six vertices and at least 11 edges, so the bound drops by at most $\frac{7 \cdot 6 - 11}{16} \leq 2$, contradicting the minimality of G. Otherwise, removing the same set destroys five vertices and at least nine edges, so the bound drops by at most $\frac{7 \cdot 5 - 9}{16} \leq 2$, again a contradiction to the minimality of G.

(iii) Since G is triangle-free, H is an induced subgraph.

It suffices to show that the removal of the vertex set $X := N_G[v_1] \cup N_G[v_2]$ results in a smaller counterexample. By an application of (ii) to the three C_4 subgraphs of H, we may assume that $d_G(v) = 4$ for all vertices of H with the possible exception of v_1 and w_1.

If $d_G(v_1) = 3$, then removing X destroys $|X| = 6$ vertices and $s_G(v_2) \geq 14$ edges, i.e. it lowers the bound by at most $\frac{7 \cdot 6 - 14}{16} \leq 2$, contradicting the minimality of G. Suppose now $d_G(v_1) = 4$. If v_1 and v_2 have a common neighbour outside H, then removing X destroys six vertices and $s_G(v_2) \geq 15$ edges, a contradiction as above. If v_1 and v_2 have two different neighbours v_1' resp. v_2', then removing X destroys $|X| = 7$ vertices and $s_G(v_1) + d_G(v_2') - \varepsilon$ vertices, where $\varepsilon = \begin{cases} 1 & \text{, if } v_1' \in N_G(v_2') \\ 0 & \text{, if } v_1' \notin N_G(v_2') \end{cases}$. Now $s_G(v_1) + d_G(v_2') \geq 17$ with equality only if $d_G(v_1') = d_G(v_2') = d_G(w_1) = 3$, so unless equality holds and v_1' and v_2' are adjacent, removing X destroys at least 17 edges and therefore lowers the bound by at most $\frac{7 \cdot 7 - 17}{16} = 2$.

□

The best lower bound on the independence ratio of possibly disconnected triangle-free graphs seems to be unknown for average degrees beyond four. It is natural to ask whether the situation is similar to the case without odd girth restriction, i.e. whether $\alpha(\mathcal{G}_{\Delta\text{-free}}, \cdot)$ continues to be piecewise-linear for large average degrees, and whether there exists a graph X of average degree greater than four such that all disjoint unions of X and $Ci_{13}[1,5]$ minimise the independence ratio among all graphs with the same average degree. A natural candidate for such a graph X is the unique sparsest instance among the triangle-free graphs with $\alpha = 6$ and maximum order, which contains 22 vertices and 60 edges. Denser graphs with even lower independence ratio include a graph or order 27 and size 85 with $\alpha = 7$ and the graph $Ci_{35}[1,7,11,16]$ with $\alpha = 8$. The following bound would be best possible for $Ci_{13}[1,5]$ and R_6:

Question 2. *Do all triangle-free graphs G satisfy $\frac{\alpha(G)}{n(G)} \geq \frac{84-5d(G)}{208}$?*

Unfortunately, a proof by simple reduction arguments as applied in this chapter appears to be very difficult for graphs with average degree greater than four.

4.2 Graphs with odd girth 7

It seems possible that a result similar to Theorem 17 holds for graphs of odd girth at least 7.

Question 3. *Does there exist a set $\mathcal{H}$ of 2-connected graphs such that*

$$\alpha(G) \geq \frac{1}{9}\left(5n(G) - m(G) - \lambda_{\mathcal{H}}(G)\right)$$

for each graph G of odd girth at least 7?

The answer of this question is unknown even when restricted to the structurally simpler class of subcubic graphs of girth 7. If it is positive, then $\mathcal{H}$ is a richer family of difficult blocks than the family $\mathcal{G}_1$ used for Theorem 17: $\mathcal{G}_1$ can be constructed from C_5 by repeated application of an extension operation that turns a given difficult block into a unique supergraph with three additional vertices, five additional edges and an independence number that rises by one. Similarly, $\mathcal{H}$ contains graphs that can be constructed from C_5 by repeated application of an extension operation that turns a given difficult block into a supergraph with five additional vertices, seven additional edges and an independence number that rises by two.

There are two circumstances that make the construction of $\mathcal{H}$ more difficult than that of $\mathcal{G}_1$: First, the extension does not always yield a *unique* supergraph. C_7 itself for example allows the two extensions shown in Figure 4.5. While H_{12} and its extensions give rise to a family of uniquely extendable blocks similar to the family $\mathcal{G}_1$ (see Figure 4.7), there are several ways to extend F_{12} and its extensions. One choice of such extensions is shown in Figure 4.6, but the general structure of these blocks has yet to be explored.

A second problem is that not all difficult blocks are extensions of C_7 as above. There are two cubic graphs of girth 7 and order 26 that must be contained in $\mathcal{H}$ because their

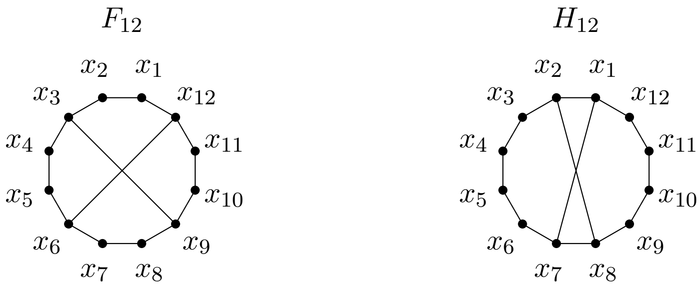


Figure 4.5: The two $\mathcal{H}$-difficult blocks F_{12} and H_{12}.

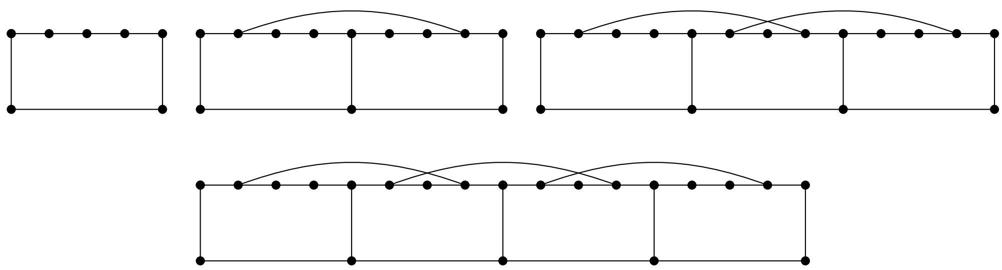


Figure 4.6: First members of an infinite sequence of $\mathcal{H}$-difficult blocks starting with C_7 and F_{12}.

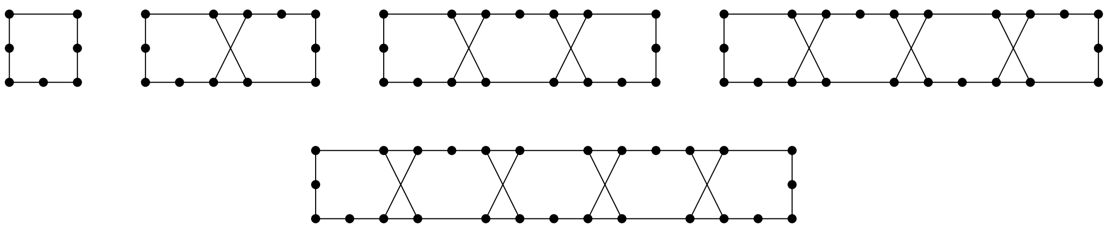


Figure 4.7: First members of a second infinite sequence of $\mathcal{H}$-difficult blocks starting with C_7 and H_{12}.

independence number is 10, but all extensions have an order of residue 2 modulo 5. One of them is the graph which arises from the disjoint union of two cycles $v_0v_1 \dots v_{12}$ and $w_0w_1 \dots w_{12}$ of order 13 by adding the edges $\{v_i, w_{5i}\}$ for $i \in \{0, \dots, 12\}$. While all elements of $\mathcal{G}_\infty$ satisfy $\alpha(G) = \frac{n(G)-\frac{m(G)}{5}}{2}$ and all extensions of C_7 satisfy $\alpha(G) = \frac{n(G)-\frac{m(G)}{7}}{2}$, these two graphs are more dense than all extensions of C_7 and have independence number smaller than $\frac{n(G)-\frac{m(G)}{7}}{2}$, which may raise doubts that Question 3 has a positive answer.

Using a computer, we have verified Question 3 for small graphs with the result that the answer is positive for all subcubic triangle-free graphs of order at most 23. If the bound holds, then every connected subcubic graph G of odd girth at least 7 would have an independent set of order at least $(5n(G)-m(G)-1)/9$. In particular, every connected subcubic graph G of odd girth at least 7 would have an independent set of order at least $(7n(G)-2)/18$ which would be best-possible in view of the two cubic graphs of order 26 mentioned above.

We now give two theorems which give positive answers to restrictions of Question 3.

Let $\mathcal{H}_0 = \{C_7, F_{12}, H_{12}\}$. From now on, we will write "difficult" for "$\mathcal{H}_0$-difficult" and $\lambda(G)$ for $\lambda_{\mathcal{H}_0}(G)$. Note that the elements of $\mathcal{H}_0$ (and in fact all extensions of C_7) satisfies the hypotheses of Lemma 15: Removal of at most two vertices from an element of $\mathcal{H}_0$ does not reduce its independence number, i.e. $r(\mathcal{H}_0) = 2$, and all extensions of C_7 satisfy $\alpha(G) = \frac{5n(G)-m(G)-1}{9}$.

The following two theorems show that the statement holds when restricted to graphs in which high degree vertices are sparse as expressed by the conditions $(\star)$ and $(\star\star)$.

Theorem 18. *If G is a graph of odd girth at least 7 such that*

$(\star)$ every vertex of degree more than 2 in G has at most one neighbour of degree more than 2,

then $\alpha(G) \geq \frac{5n(G)-m(G)-\lambda(G)}{9}$.

Theorem 19. *If G is a* subcubic *graph of odd girth at least 7 such that*

$(\star\star)$ for every pair of vertices u and v of degree 3 and at distance 3 in G, u or v has at most one neighbour of degree 3,

then $\alpha(G) \geq \frac{5n(G)-m(G)-\lambda(G)}{9}$.

Note that if a graph satisfies $(\star)$, then it does not contain H_{12} as a subgraph. The condition $(\star)$ is equivalent to requiring that the vertices of degree more than 2 induce a subgraph of maximum degree at most 1. By Lemma 1 and Lemma 2, the problem to find a maximum independent set remains APX-hard when restricted to the classes of graphs considered in Theorem 18 and Theorem 19.

The proof of Theorem 19 is significantly more complicated than that of Theorem 18, but relies on similar reduction techniques. Therefore, we give the proof of Theorem 18 here and refer the interested reader to [57] for a proof of Theorem 19.

Proof of Theorem 18. Note that any induced subgraph H of G satisfies ($\star$) with G replaced by H. In order to obtain a contradiction, we assume that G is a counterexample to $\alpha(G) \leq b(G) := \frac{5n(G)-m(G)-\lambda(G)}{9}$ of minimum order. By Lemma 15 (i), G is 2-edge-connected and not difficult. In particular, the minimum degree of G is at least 2. Since it is easy to verify $r(\mathcal{H}_0) = 2$, Lemma 15 (ii) implies that removal of a vertex set X which induces a cut of size $\varphi(X)$ creates at most $\frac{\varphi(X)}{3}$ difficult components. $b(G)$

Let v be a vertex of maximum degree in G, let $N_G(v) = \{v_1, v_2, \ldots, v_k\}$ for $k := d_G(v)$. By ($\star$), we may assume $d_G(v_i) = 2$ for all $i < k$. Since the bound holds for all cycles other than C_3 and C_5, we have $k \geq 3$. We claim that removing the vertex set $X := N_G[v_1] \cup N_G[v_2] \cup \cdots \cup N_G[v_{k-1}]$ decreases the independence number by at least $k-1$.

Note first that $G[X]$ is a tree: It cannot contain a cycle of length at least six, since the distance of any vertex to v is at most two, cycles of length three or five are forbidden by the odd girth condition, so suppose that $G[X]$ contains a C_4. It contains the vertex v, w.l.o.g. the vertices v_1 and v_2, and a common neighbour w of v_1 and v_2. Removing $Y := N_G[v_1] \cup N_G[v_2] = \{v, v_1, w, v_2\}$ decreases the independence number by two, the order by four and the size by $4 + \varphi(Y)$, while creating at most $\frac{\varphi(Y)}{3}$ difficult components. Therefore the bound decreases by at most

$$\frac{5 \cdot 4 - (4 + \varphi(X)) - \frac{\varphi(Y)}{3}}{9} < \frac{5 \cdot 4 - 4}{9} < 2,$$

contradicting the minimality of G and thus the assumption that $G[X]$ contains a cycle.

Let $G' := G - X$. The removal of X destroys $|X| = 2k-1$ vertices and $m_X = \varphi(X) + 2(k-1)$ edges, where $\varphi(X) \geq k$ because all k elements of $X \setminus \{v_1, \ldots, v_{k-1}\}$ have at least one neighbour outside X. Altogether, the removal of X decreases the bound by at most

$$\begin{aligned} b(G) - b(G') &\leq \frac{5|X| - m_X + \lfloor \frac{1}{3}\varphi(X) \rfloor}{9} = \frac{5|X| - (m_X - \varphi(X)) - \lceil \frac{2}{3}\varphi(X) \rceil}{9} \\ &\leq \frac{5(2k-1) - 2(k-1) - \left\lceil \frac{2\varphi(X)}{3} \right\rceil}{9} = \frac{8k - \left\lceil \frac{2\varphi(X)}{3} \right\rceil - 3}{9}. \end{aligned}$$

For $k \geq 4$, we have $\frac{8k - \left\lceil \frac{2\varphi(X)}{3} \right\rceil - 3}{9} \leq \frac{8k - \frac{2k}{3} - 3}{9} \leq k-1$, which contradicts the minimality of G because $\alpha(G) - \alpha(G') \geq k-1$.

Therefore, we may assume $k = 3$. If $\varphi(X) > 3$, then

$$b(G) - b(G') \leq \frac{8 \cdot 3 - \left\lceil \frac{2\varphi(X)}{3} \right\rceil - 3}{9} = \frac{24 - 3 - 3}{9} = 2,$$

contradicting the minimality of G. Therefore, $\varphi(X) = 3$, i.e. the neighbours of v_1 and v_2 other than v have only degree two, and G' is a difficult block by Lemma 15 (ii).

First, suppose $G' \simeq C_7$. In view of the freedom to choose v among the vertices of degree 3, we may assume that every vertex of degree 3 has exactly one neighbour of degree 3 and that there are no two vertices of degree 3 at distance 2 in G. Since G has odd girth at least 7, this implies that y' and z' are adjacent to adjacent vertices of G' and that G is isomorphic to F_{12}, which contradicts the fact that G is not difficult.

Finally, suppose $G' \simeq F_{12}$. By ($\star$), all the vertices of degree 2 in G' also have degree 2 in G. This implies that G has maximum degree at least 4, which is a contradiction. $\square$

4.3 Bipartite ratio

For a graph G, let $\alpha\alpha(G)$ denote the maximum order of an induced bipartite subgraph graph of G, i.e. the maximum size of the union of two disjoint independent sets. Note that this parameter is closely related to $n(G) - \alpha\alpha(G)$, the minimum size of an *odd cycle transversal*, i.e. of a vertex set that meets every odd cycle in G. $\alpha\alpha(G)$

In [46], Lewis and Yannakakis show that it is NP-hard to determine the minimum number of nodes that have to be deleted in order to obtain a graph that belongs to some nontrivial hereditary class, which implies that it is NP-hard to find a minimum odd cycle transversal or to determine $\alpha\alpha(G)$. On the other hand, Reed, Smith and Vetta show in [61] that the determination of this parameter is fixed-parameter-tractable: They introduce a method called *iterative compression* and show that for any fixed value of k, the inequality $\alpha\alpha(G) \geq k$ can be verified in $O(mn)$. The running time of this verification is improved to $O(m\,\alpha(m,n))$, where $\alpha(m,n)$ denotes the inverse Ackermann function, by Reed and Kawarabayashi in [40]. A linear time algorithm for the determination of $\alpha\alpha(G)$ on planar graphs is given by Fiorini et al. in [25].

In this section, we discuss best possible lower bounds for the *bipartite ratio* $\frac{\alpha(G)}{n(G)}$ for graphs of given odd girth and given average degree. In analogy to the definition of the asymptotic independence ratio, we define, for a class $\mathcal{P}$ of graphs, the *asymptotic bipartite ratio* by $\alpha\alpha(\mathcal{P}, d)$

$$\alpha\alpha(\mathcal{P}, d) = \lim_{n\to\infty} \left(\inf \left\{ \frac{\alpha\alpha(G)}{n(G)} \;\middle|\; G \in \mathcal{P}, d(G) \leq d, n(G) \geq n \right\} \right).$$

Since $\alpha(G) \leq \alpha\alpha(G) \leq 2\alpha(G)$ for any graph G, the results cited in the introduction imply that the asymptotic bipartite ratio of arbitrary and connected graphs has the order of magnitude $\Theta\left(\frac{1}{d}\right)$, while the asymptotic bipartite ratio of the class of graphs with $g_{\text{odd}} = k$ is $\Theta\left(\frac{\log d}{d}\right)$ for any fixed odd integer $k > 3$.

In [80], Zhu gives a comprehensive answer to the question of best lower bounds for the bipartite ratio of subcubic triangle-free graphs. It implies that every connected subcubic triangle-free graph satisfies

$$\alpha\alpha(G) = \frac{2}{7}\left(4n(G) - m(G) - 1\right) \tag{4.7}$$

and all instances with $\alpha\alpha(G) < \frac{2}{7}\left(4n(G) - m(G)\right)$ can be constructed from the disjoint union of elements of a set of ten graphs by adding bridges. These ten graphs allow to determine the asymptotic bipartite ratio of arbitrary and connected *subcubic* triangle-free graphs.

4.3.1 Odd girth 3

Proposition 1. *For every $t \in \mathbb{N} \setminus \{1,2\}$ and every graph G, $\frac{\alpha\alpha(G)}{n(G)} \geq b_t(d(G))$ with $b_t(d) = \frac{4t-2d}{t(t+1)}$.*

Proof. For a contradiction, suppose that t is minimal such that the bound $\alpha\alpha(G) \geq n(G)b_t(d(G))$ does not hold, and G is a vertex-minimal counterexample. Then $\Delta(G) < t$, since a graph G' that arises by removing a vertex of degree at least t satisfies $\alpha\alpha(G') \leq \alpha\alpha(G)$ and $n(G')b_t(G') \geq n(G)b_t(G)$, a contradiction to the minimality of G. Similarly $\delta(G) \geq 2$, since removing a vertex of degree at most 1 decreases $\alpha\alpha$ by exactly one and b_t by at most 1.

For $t = 3$ this implies that G is 2-regular, i.e. a disjoint union of cycles. Since every cycle C contains an induced bipartite subgraph of order at least $n(C) - 1$, and its length is at least 3, this implies $\alpha\alpha(C) \geq \frac{2}{3}n(C) = n(C) \cdot b_3(2)$, i.e. cycles and thus their disjoint unions are no counterexamples.

For $t > 3$ note that $\alpha\alpha(G) < n(G)b_t(d(G))$ and $\alpha\alpha(G) \geq n(G)b_{t-1}(d(G))$ implies $b_t(d(G)) > b_{t-1}(d(G))$, which is equivalent to $d(G) > t - 1$. Therefore, G contains a vertex of degree at least t, which contradicts the initial observation $\Delta(G) < t$. □

Since $\alpha(G) \geq \frac{\alpha\alpha(G)}{2}$ for all graphs G, Proposition 1 generalises Corollary 3 for $d \geq 2$. It allows to determine the asymptotic bipartite ratio of arbitrary graphs (see Figure 4.8):

Corollary 6.

$$\alpha\alpha(\mathcal{G}, d) = \begin{cases} 1 - \frac{d}{6} & d \leq 3 \\ \frac{4\lfloor d+1 \rfloor - 2d}{\lfloor d+1 \rfloor \lfloor d+2 \rfloor} & d \geq 2 \end{cases}$$

Proof. In both cases, the given value coincides with $n(G)\min_{t \in \mathbb{N}\setminus\{1,2\}} b_t(d)$ on the intervals $[0,3]$ resp. $[2,\infty)$, so by Proposition 1 it suffices to show that there are tight examples for all rational average degrees. Indeed, for each $t \in \mathbb{N} \setminus \{1,2\}$, the bound $\alpha\alpha(G) = n(G)b_t(d(G))$ holds for disjoint unions of K_t and K_{t+1}, and $\alpha\alpha(G) = n(G)b_3(d(G))$ is satisfied by disjoint unions of K_1, K_3 and K_4. □

By a similar argument, we obtain a slightly stronger statement which generalises the bound (4.1) of Caro and Wei for graphs without isolated vertices:

Proposition 2. *Every graph G satisfies $\alpha\alpha(G) \geq \sum_{v \in V(G)} \frac{2}{\max\{1, d_G(v)\}+1}$.*

Proof. Suppose G is a counterexample of minimum order. If G contains a vertex of degree at most one, then removing this vertex decreases $\alpha\alpha$ by exactly one and the bound by at most one, so it yields a smaller counterexample. Therefore, $\delta(G) \geq 2$. Let v be a vertex

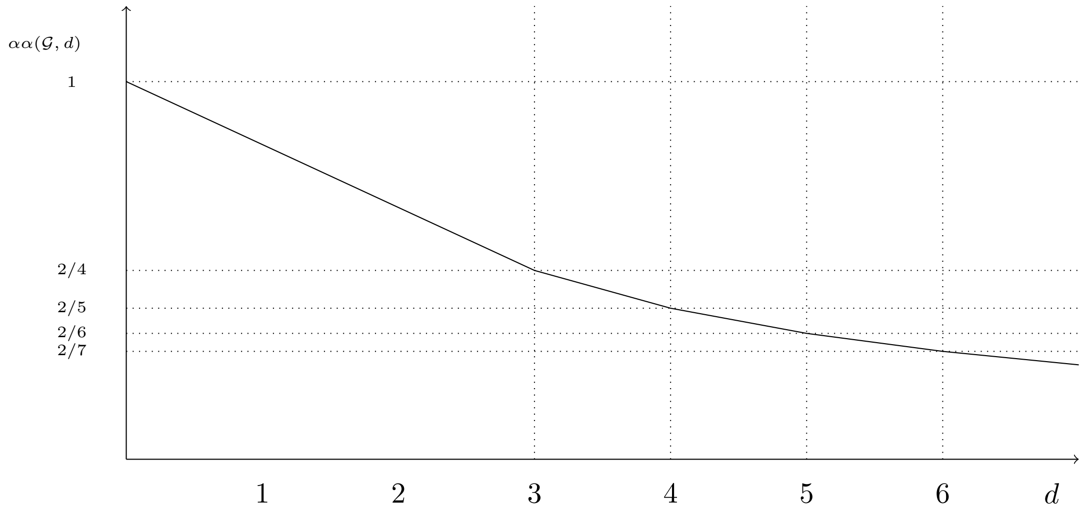


Figure 4.8: Asymptotic bipartite ratio of arbitrary graphs

of maximum degree and let $G' = G - v$. Then $\delta(G') \geq 1$, and thus

$$\begin{aligned}\alpha\alpha(G') &= \alpha\alpha(G) - \frac{2}{d_G(v)+1} + \sum_{w \in N_G(v)} \left(\frac{2}{d_G(w)} - \frac{2}{d_G(w)+1} \right) \\ &= \alpha\alpha(G) - \frac{2}{d_G(v)+1} + \sum_{w \in N_G(v)} \frac{2}{d_G(w)(d_G(w)+1)} \\ &\geq \alpha\alpha(G) - \frac{2}{d_G(v)+1} + |N_G(v)| \frac{2}{d_G(v)(d_G(v)+1)} = \alpha\alpha(G),\end{aligned}$$

a contradiction to the minimality of G. $\square$

Connected graphs The bound $\alpha\alpha(G) \geq 2n(G) - m(G) - 1$ holds for all connected graphs because it is satisfied with equality by trees and remains true whenever an edge is added. For odd values of g, tight instances of odd girth g include the connected graphs that arise from disjoint unions of C_g by adding bridges.

Let $\mathcal{G}_c(g)$ be the class of connected graphs of odd girth at least g. The above observation determines the asymptotic bipartite ratio of this class for low average degrees: $\mathcal{G}_c(g)$

Proposition 3. $\forall d \in [2, 2 + \frac{2}{g}]\colon\ \alpha\alpha(\mathcal{G}_c(g), d) = 2 - \frac{d}{2}$.

This agrees with the following natural metaconjecture.

Conjecture 2. $\alpha\alpha(\mathcal{G}_c(g), d) = 2\alpha(\mathcal{G}_c(g), d)$ *for all* $d \geq 2$.

For larger average degrees, it is still open if the bound implied by Conjecture 2 and

Corollary 4 holds. In Theorem 4 of [48][2], we show that for graphs G with $d(G) \geq 2 + \frac{2}{g_{\text{odd}}(G)}$,

$$\alpha\alpha(G) \geq \frac{(3g_{\text{odd}}(G) - 1)\, n(G) - g_{\text{odd}}(G)(m(G) + 1)}{2g_{\text{odd}}(G)},$$

which implies the following lower bound.

Corollary 7. $\forall d \geq 2 + \frac{2}{g}$: $\alpha\alpha(\mathcal{G}_c(g), d) \geq \frac{(6-d)\,g-2}{4g}$.

4.3.2 Triangle-free graphs

No connectivity restriction

The exact value of $\alpha\alpha(\mathcal{G}_{\Delta\text{-free}}, d)$ is still unknown for any positive value of d. It seems natural to conjecture that $\alpha\alpha(\mathcal{G}_{\Delta\text{-free}}, d) = 2\alpha(\mathcal{G}_{\Delta\text{-free}}, d)$ for sufficiently high values of d (since disjoint unions of C_5 and K_1 of degree $d \in (0, 2)$ satisfy $\alpha\alpha(G) < 2\alpha(\mathcal{G}_{\Delta\text{-free}}, d)$, this does not hold for $d < 2$). This leads to the conjecture

Conjecture 3. *For all triangle-free graphs G, $\alpha\alpha(G) \geq b_1(G) := n(G) - \frac{m(G)}{5}$.*

If this bound holds, then tight examples includes the set $\mathcal{G}_1$ of all *difficult blocks* used in Theorem 17 and the set $\mathcal{X}$ defined in Section 4.1.2, the Petersen graph X_4 (cf. Figure 4.10) and the graphs D_1, D_2, D_3 in Figure 4.9.

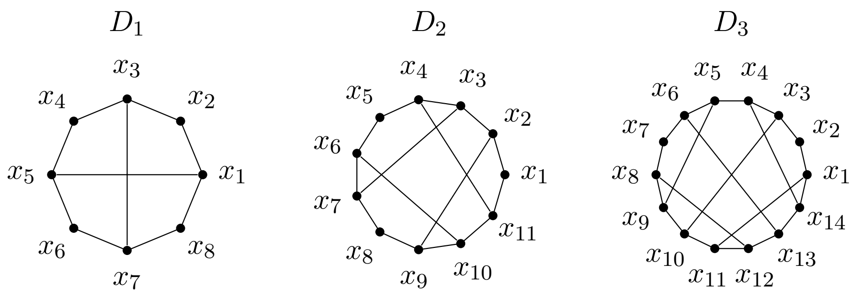


Figure 4.9: Some tight examples for Conjecture 3

Conjecture 4. *If G is a connected graph with $n(G) > 1$ and $\alpha\alpha(G) = n(G) - \frac{m(G)}{5}$, then*

$$G \in \mathcal{G}_1 \cup \mathcal{X} \cup \{X_4, D_1, D_2, D_3\}.$$

The following Lemma restricts the class of possible counterexamples:

[2]The precise statement is slightly stronger: The bound given in the paper is

$$\alpha\alpha(G) \geq \left\lceil \frac{(g_{\text{odd}}(G) - 1)\, n(G)}{g_{\text{odd}}(G)} \right\rceil - \frac{1}{2}\left(m(G) - \left(\left\lfloor \frac{(g_{\text{odd}}(G) + 1)\, n(G)}{g_{\text{odd}}(G)} \right\rfloor - 1 \right) \right).$$

Lemma 18. *Any vertex-minimal counterexample G to Conjecture 3 satisfies $\delta(G) \geq 3$ and $\Delta(G) = 4$.*

Proof. Since Theorem 2.1 in [80] implies that Conjectures 3 and 4 hold for subcubic graphs, it suffices to show that every vertex $v \in V(G)$ has degree three or four. Let $G' = G - v$. If $d_G(v) \geq 5$, then $\alpha\alpha(G') \leq \alpha\alpha(G) < b_1(G) \leq b_1(G')$. If $d_G(v) \leq 1$, then $\alpha\alpha(G') = \alpha\alpha(G) - 1 < b_1(G) - 1 \leq b_1(G')$. In both cases, we obtain a contradiction to the minimality of G, so we may assume that G is a connected graph with $\delta(G) \geq 2$ and $\Delta(G) = 4$. Let v be a vertex of degree two such that $N_G(v) = \{u, w\}$ with $d_G(w) > 2$.

Claim 1 $d_G(u) = d_G(w) = 3$

Indeed, if $d_G(u) = 2$, then let $G' = G - N_G[v]$. Now $n(G) - n(G') = 3$, $m(G) - m(G') = d_G(w) + 2$ and $\alpha\alpha(G') \leq \alpha\alpha(G) - 2$ because for every set $S \subseteq V(G')$ such that $G'[S]$ is bipartite, $G[S \cup \{v, u\}]$ is bipartite. Hence

$$\alpha\alpha(G') \leq \alpha\alpha(G) - 2 < b_1(G) - 2 = b_1(G') + \left(3 - \frac{d_G(w)}{5}\right) - 2 \leq b_1(G').$$

If u or w has degree four, w.l.o.g. $d_G(w) = 4$, then let $G' = G - \{v, w\}$. Now $n(G) - n(G') = 2$, $m(G) - m(G') = 5$ and $\alpha\alpha(G') \leq \alpha\alpha(G) - 1$ because for every $S \subseteq V(G')$ such that $G'[S]$ is bipartite, $G[S \cup \{v\}]$ is bipartite. Hence

$$\alpha\alpha(G') \leq \alpha\alpha(G) - 1 < b_1(G) - 1 = b_1(G') + (2 - 1) - 1 = b_1(G'),$$

again contradicting the minimality of G.

Claim 2 *No vertex x with $\mathrm{dist}_G(v, x) = 2$ has degree two.*

Suppose $x \in N_G(u)$ and $d_G(x) = 2$. Let $G' = G - \{v, u, x\}$. Then $n(G) - n(G') = 3$, $m(G) - m(G') = 5$ and $\alpha\alpha(G) - \alpha\alpha(G') \geq 2$ because for every $S \subseteq V(G')$ such that $G'[S]$ is bipartite, $G[S \cup \{v, x\}]$ is bipartite. Hence we obtain the contradiction

$$\alpha\alpha(G') \leq \alpha\alpha(G) - 2 < b_1(G) - 2 = b_1(G') + \left(3 - \frac{5}{5}\right) - 2 = b_1(G').$$

Claim 3 *v is not contained in a 4-cycle.*

Suppose u and w have more than one common neighbour. We may assume $N_G(u) = \{v, x, u'\}$ and $N_G(w) = \{v, x, w'\}$ for vertices x, u' and w' with $x \notin \{u', w'\}$. Let $G' = \{v, u, w, x\}$. Then $n(G) - n(G') = 4$ and $m(G) - m(G') = d_G(x) + 4 > 5$, so $b_1(G) - b_1(G') < 4 - \frac{5}{5} = 3$. Since G is no minimal counterexample, this implies $\alpha\alpha(G) - \alpha\alpha(G') \leq 2$. This implies that for every maximum subset $S \subseteq V(G')$ such that $G'[S]$ is bipartite and every set $X \subset V(G) \setminus V(G')$ of cardinality three, $G[S \cup X]$ is not bipartite. For $X = \{u, v, w\}$ this implies that $G'[S]$ contains neighbours of u and w in different colour classes, i.e. u' and w' are contained in opposite partite sets of $G'[S]$. For $X = \{v, u, x\}$ this implies that x has a neighbour in the partite set of $G'[S]$ containing u', and for $X = \{v, w, x\}$ it shows that x has a neighbour in the partite set of $G'[S]$ containing w'. We conclude that $d_G(x) = 4$.

Let $G'' = G - \{v, u, w, x, u'\}$. Then $n(G) - n(G'') = 5$, $m(G) - m(G') = 7 + d_G(u') \geq 10$, and for every set $S \subseteq V(G')$ such that $G'[S]$ is bipartite, $G[S \cup \{u, v, w\}]$ is bipartite, so $\alpha\alpha(G) \geq \alpha\alpha(G'') + 3$. This leads to the contradiction

$$\alpha\alpha(G') \leq \alpha\alpha(G) - 3 < b_1(G) - 3 \leq b_1(G') + \left(5 - \frac{10}{5}\right) - 3 = b_1(G'')$$

and concludes the proof of Claim 3.

By Claim 3, we may assume $N_G(u) = \{v, u_1, u_2\}$. Let $G' = G - \{v, u, w, u_1, u_2\}$. Then $n(G) - n(G') = 5$, $m(G) - m(G') = 4 + d_G(u_1) + d_G(u_2) \geq 10$, and $\alpha\alpha(G) - \alpha\alpha(G') \geq 3$ because for every set $S \subseteq V(G')$ such that $G'[S]$ is bipartite, $G[S \cup N_G[v]]$ is bipartite. This implies the contradiction

$$\alpha\alpha(G') \leq \alpha(G) - 3 < b_1(G) - 3 \leq b_1(G') + \left(5 - \frac{10}{5}\right) - 3 = b_1(G'),$$

which concludes the proof. □

Note that under the assumption that Conjecture 3 holds, the argument of the proof implies that every graph G of minimum degree 2 with $\alpha\alpha(G) = n(G) - \frac{m(G)}{5}$ contains a proper induced subgraph of order at least $n(G) - 5$ that satisfies the same inequality. Therefore, a computer search shows that the list of graphs in Conjecture 4 does not miss any graph with minimum degree 2, if Conjecture 3 holds.

It seems natural to ask whether $\alpha\alpha(G) \geq n(G) - \frac{m(G)}{g_{\text{odd}}(G)}$ holds for graphs of arbitrary odd girth, which would agree with the bound $\alpha\alpha(G) \geq n(G) - \frac{m(G)}{3}$ in Proposition 1, Conjecture 3 and the fact that bipartite graphs satisfy $\alpha\alpha(G) = n(G)$. However, the answer is negative, since the two cubic graphs of order 26 and girth 7 mentioned as tight examples to 3, satisfy $\alpha\alpha(G) = 20 < 20 + \frac{3}{7} = n(G) - \frac{m(G)}{7}$.

Connected graphs

The asymptotic bipartite ratio of connected triangle-free graphs up to an average degree of $\frac{12}{5}$ is given by Corollary 3. Conjecture 2 suggests that this value can be obtained for average degrees in the interval $\left[\frac{12}{5}, \frac{10}{3}\right]$ by doubling the bound of Theorem 17. Here, we conjecture the statement of such a possible generalisation, but its validity remains open.

Conjecture 5. *Let G be a triangle-free graph and let G' be the graph that arises from G by removing all bridges. Then $\alpha\alpha(G) \geq b_2(G) := \frac{8n(G) - 2m(G') - \lambda_1(G') - 2\lambda_2(G')}{7}$, where λ_i counts the number of components that are contained in some class $\mathcal{L}_i$ to be defined below.*

The removal of bridges and the distinction between G and G' are a slight strengthening of the statement, which is possible because bridges have no influence on the parameter $\alpha\alpha$. We call the elements of the sets $\mathcal{L}_i$ *difficult blocks*, specifically those of $\mathcal{L}_1$ *weakly difficult* and those of $\mathcal{L}_2$ *strongly difficult*.

The class $\mathcal{L}_2$ should contain the class $\mathcal{G}_1$ of the *difficult blocks* used in Theorem 17,
$\mathcal{L}_2$ since $\forall k \geq 2\colon \alpha\alpha(G_k) = 2\alpha(G_k)$. We define

$$\mathcal{L}_2 := \mathcal{G}_1 \cup \{X_1, X_2, X_3\},$$

where the additional elements are drawn in Figure 4.9.

Most elements of $\mathcal{L}_1$ can be obtained from an element of $\mathcal{L}_2$ by a simple operation: For a graph G, let $f(G)$ be the set of isomorphism classes of 2-connected graphs $G'(u,v,y)$ $f(G)$ that can be constructed from G as follows: Let $\{u,v\} \in E(G)$, and let $y \in N_G(v) \setminus \{u\}$ be a vertex that is contained in every maximum induced bipartite subgraph of $G - \{u,v\}$. Then $G'(u,v,y)$ is the graph obtained from G by removing $\{u,v\}$ and adding a new vertex x with $N_G(x) = \{u,y\}$. The following Lemma quantifies how this operation "reduces the difficulty" of a strongly difficult block:

Lemma 19. *Let G be a graph that is edge-minimal with respect to $\alpha\alpha$, i.e. satisfies $\alpha\alpha(G-e) > \alpha\alpha(G)$ for each $e \in E(G)$. Then all graphs $G' \in f(G)$ are bridgeless and satisfy $\alpha\alpha(G') = \alpha\alpha(G)+1$, $n(G') = n(G)+1$ and $m(G') = m(G)+1$.*

Proof. Since G is edge-minimal, it is bridgeless, and so is G'. The statements on order and size are obvious from the construction. It remains to show that the graph G' that is obtained from G by removing the edge $\{u,v\}$ and adding a vertex x with $N_G(x) = \{u,y\}$ satisfies $\alpha\alpha(G') = \alpha\alpha(G)+1$.

By the edge-minimality of G, $\alpha\alpha(G - \{u,v\}) = \alpha\alpha(G)+1$, and both u and v are contained in the same partite set of every maximum induced bipartite subgraph of G. By assumption, y is also contained in every such subgraph, but since it is adjacent to v, it belongs to the opposite partite set. Therefore, x has neighbours in both partite sets of every maximum induced bipartite subgraph of G, and $\alpha\alpha(G') = \alpha\alpha(G'-x) = \alpha\alpha(G - \{u,v\})$. $\square$

It turns out that $f(G_2) = f(X_1) = \emptyset$, and $\forall G \in \mathcal{L}_2 \setminus \{G_2, X_1\}\colon\ |f(G)| = 1$. Since all elements of $\mathcal{L}_2$ are edge-minimal with respect to $\alpha\alpha$, the class $\mathcal{L}_1$ should contain $f(\mathcal{L}_2)$. We define $\mathcal{L}_1$

$$\mathcal{L}_1 := f(\mathcal{L}_2) \cup \{K_1, X_4, X_5\},$$

where the graphs X_4 and X_5 are given in Figure 4.10.

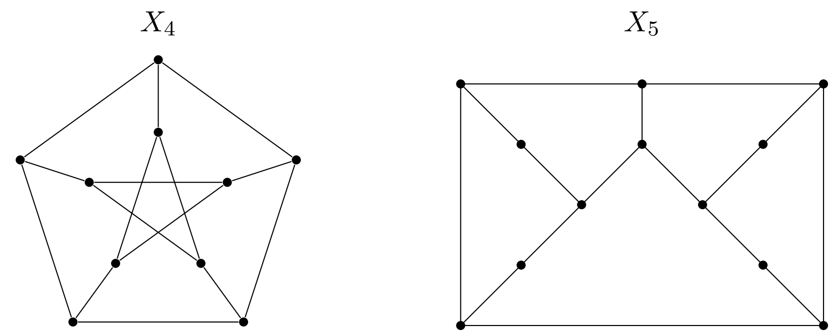


Figure 4.10: Two weakly difficult graphs

Bibliography

[1] M. Albertson, B. Bollobás, and S. Tucker, The independence ratio and the maximum degree of a graph, *Congressus Numerantium* **17** (1976), 43–50.

[2] M. Ajtai, J. Komlós and E. Szemerédi, A note on Ramsey numbers, *Journal of Combinatorial Theory, Series A*, vol. 29 (1980), pp. 354–360.

[3] N. Alon, J. Spencer, The Probabilistic Method, John Wiley 1992.

[4] B. Bajnok and G. Brinkmann, On the Independence Number of Triangle-Free Graphs with Maximum Degree Three, *J. Combin. Math. Combin. Comput.* **26** (1998), 237–254.

[5] D. Bauer and E. F. Schmeichel, Hamiltonian degree conditions which imply a graph is pancyclic. *J. Combin. Theory Ser. B* 48 (1990), 111–116.

[6] P. Berman and T. Fujito, On approximation properties of the Independent set problem for degree 3 graphs, *Lecture Notes in Computer Science*, vol. 955 (1995), pp. 449–460.

[7] P. Berman, M. Karpinski, On some tighter inapproximability results (1998). Lecture Notes In Computer Science, vol. 1644 (1999), pp. 200–209.

[8] B. Bollobás, Extremal graph theory, L. M. S. Monographs. 11. London - New York - San Francisco: Academic Press. XX, 488 p. (1978).

[9] J.A. Bondy, Pancyclic Graphs. *J. Combin. Theory Ser. B* 11 (1971), 80–84.

[10] J.A. Bondy, Longest paths and cycles in graphs of high degree, Res. Rep. No. CORR 80-16, Univ. Waterloo, Waterloo, ON, 1980.

[11] B. Bollobás and A. Thomason, Weakly pancyclic graphs. *J. Combin. Theory Ser. B* 77 (1999), 121–137.

[12] A. Caprara and R. Rizzi, Packing triangles in bounded degree graphs, *Inf. Process. Lett.* **84** (2002), 175–180.

[13] Y. Caro, New Results on the Independence Number, Technical Report, Tel-Aviv University, 1979.

[14] V. Chvátal, On Hamilton's ideals. *J. Combin. Theory Ser. B* 12 (1972), 163–168.

[15] J. Chuzhoy and S. Khanna, New Hardness Results for Undirected Edge-Disjoint Paths. Manuscript, 2005.

[16] J. Degenhardt and P. Recht, On a relation between the cycle packing number and the cyclomatic number of a graph, *manuscript* (2008).

[17] T. Denley, The Independence Number of Graphs with Large Odd Girth, *Electron. J. Comb.* **1** (1994), R9.

[18] R. Diestel, Graphentheorie, *Springer-Verlag*

[19] P. Erdős, Some of my favourite problems in various branches of combinatorics. *Matematiche (Catania)* 47 (1992), 231-240.

[20] P. Erdős, R. Faudree, C. Rousseau, and R. Schelp, The number of cycle lengths in graphs of given minimum degree and girth. *Discrete Math.* 200 (1999), 55-60.

[21] P. Erdős and L. Pósa, On the maximal number of disjoint circuits of a graph. *Publ. Math. Debrecen* **9** (1962), 3-12.

[22] G. Fan, New sufficient conditions for cycles in graphs. *J. Combin. Theory Ser. B* 37 (1984), 221-227.

[23] R.J. Faudree, E. Flandrin, M.S. Jacobson, J. Lehel, and R.H. Schelp, Even cycles in Graphs with Many Odd Cycles. *Graphs and Combinatorics* 16 (2000), 399-410.

[24] M. Ferrara, A. Harris, and M. Jacobson, Cycle lengths in Hamiltonian graphs with a pair of vertices having large degree sum. *Graphs and Combinatorics* 26 (2010), 215-223.

[25] S. Fiorini, N. Hardy, B. Reed, A. Vetta, Planar Graph Bipartization in Linear Time, *Discrete Appl. Math.* **156** (2008), 1175–1180.

[26] Z. Friggstad, M. Salavatipour, Approximability of Packing Disjoint Cycles, In Proceedings of ISAAC 2007, *Lecture Notes in Computer Science* **4835** (2007), pp. 304–315.

[27] M. R. Garey, D. S. Johnson, Computers and intractability: a guide to the theory of incompleteness, Freeman, San Francisco (1979).

[28] F. Göring, J. Harant, D. Rautenbach, and I. Schiermeyer, Locally Dense Independent Sets in Regular Graphs of Large Girth, In: Research Trends in Combinatorial Optimization, Springer Berlin Heidelberg, 2009, 163-183.

[29] J. E. Graver, J. Yackel, Some graph theoretic results associated with Ramsey's theorem, *Journal of Combinatorial Theory* 4 (1968), 125–175.

[30] J. Harant, M.A. Henning, D. Rautenbach, and I. Schiermeyer, Independence Number in Graphs of Maximum Degree Three, *Discrete Math.* **308** (2008), 5829-5833.

[31] J. Harant and D. Rautenbach, Independence in Connected Graphs, manuscript 2009.

[32] J. Harant, D. Rautenbach, F. Regen, and P. Recht, Packing edge-disjoint cycles in graphs and the cyclomatic Number, *Discrete Math.* 310 (2010), 1456-1462.

[33] J. Harant and I. Schiermeyer, On the independence number of a graph in terms of order and size, *Discrete Math.* **232** (2001), 131-138.

[34] C.C. Heckman, On the Tightness of the 5/14 Independence Ratio, *Discrete Math.* **308** (2008), 3169-3179.

[35] C.C. Heckman and R. Thomas, A New Proof of the Independence Ratio of Triangle-Free Cubic Graphs, *Discrete Math.* **233** (2001), 233-237.

[36] I. Holyer, The NP-completeness of some edge-partition problems, SIAM Journal of Computing, vol. 10, No. 4 (November 1981), pp. 713–717.

[37] K.F. Jones, Size and independendence in triangle-free graphs with maximum degree three, *J. Graph Theory* **14** (1990), 525-535.

[38] K.F. Jones, S.C. Locke, Finding Independent Sets in Triangle-free Graphs, *SIAM J. Discrete Math.* **9(4)** (1996), 674-681.

[39] R. M. Karp, Reducibility among combinatorial problems, *Complexity of Computer Computations*, Plenum Press, New York (1972), pp. 85–103.

[40] K. Kawarabayashi, B. Reed, An (almost) Linear-Time Algorithm for Odd Cycle Transversal, SIAM Symp. on Disc. Alg. (2010), pp. 365–378.

[41] B. Korte, and J. Vygen (2005), Combinatorial Optimization, 3$^{\text{rd}}$ edition, Springer-Verlag

[42] M. Krivelevich, Z. Nutov, M.R. Salavatipour, J. Verstraëte, and R. Yuster, Approximation algorithms and hardness results for cycle packing problems, *ACM Trans. Algorithms* **3** (2007), Article No. 48.

[43] D. Kreher, S. Radziszowski, Minimum triangle-free graphs, *Ars Combin.* **31** (1991), 65-92.

[44] J. H. Kim, The Ramsey number $R(3,t)$ has order of magnitude $\frac{t^2}{\log t}$, Random Structures & Algorithms, vol. 7, pp. 173–207.

[45] J. Lauer and N. Wormald, Large independent sets in regular graphs of large girth, *J. Comb. Theory, Ser. B* **97** (2007), 999–1009.

[46] J. M. Lewis and M. Yannakakis, The node-deletion problem for hereditary properties is NP-complete, *J. Comput. Syst. Sci.* **20(2)** (1980), 219–230.

[47] S.C. Locke, Bipartite Density and the Independence Ratio, *J. Graph Theory* **10** (1986), 47-53.

[48] C. Löwenstein, A.S. Pedersen, D. Rautenbach, F. Regen, Independence, Odd Girth, and Average Degree, to appear in *J. Graph Theory*

[49] V. Lozin and D. de Werra, Foreword: Special issue on stability in graphs and related topics, *Discrete Appl. Math.* **132** (2003), 1-2.

[50] G. J. Minty, On maximal independent sets of vertices in claw-free graphs. Journal of Combinatorial Theory Series B, 28 (1980), pp. 284–304.

[51] B. McKay, *Ramsey Graphs*, `http://cs.anu.edu.au/∼bdm/data/ramsey.html`

[52] K. Milans, D. Rautenbach, F. Regen, D. West, Cycle Spectra of Hamiltonian Graphs, manuscript 2009.

[53] J.W. Moon, On edge-disjoint cycles in a graph, *Can. Math. Bull.* **7** (1964), 519-523.

[54] L. Moser, J.R. Pounder, and J. Ridell, On the cardinality of h-bases for n, *J. London Math. Soc.* **44** (1969), 397-407.

[55] O. Ore, Note on Hamilton circuits. *Amer. Math. Monthly* 67 (1960), 55.

[56] G. Oriolo, U. Pietropaoli, G. Stauffer, A New Algorithm for the Maximum Weighted Stable Set Problem in Claw-Free Graphs. Lecture Notes in Computer Science, Volume 5035/2008.

[57] A. S. Pedersen, D. Rautenbach, F. Regen, Lower bounds on the independence number of certain graphs of odd girth at least seven, submitted to *Discrete Appl. Math.*.

[58] A. Pnueli, A. Lempel, and S. Even, Transitive orientation of graphs and identification of permutation graphs. *Canad. J. Math.* 23 (1971), 160–175.

[59] D. Rautenbach and F. Regen, On packing shortest cycles in graphs, *Inf. Process. Lett.* **109** (2009), 816–821.

[60] D. Rautenbach and F. Regen, Graphs with many vertex-disjoint cycles, manuscript 2009.

[61] B. Reed, K. Smith, A. Vetta, Finding odd cycle transversals, *Oper. Res. Lett.* **32** (2004), 299–301.

[62] M.R. Salavatipour and J. Verstraëte, Disjoint cycles: integrality gap, hardness, and approximation, In Proceedings of IPCO 2005, *Lecture Notes in Computer Science* **3509** (2005), 51–65.

[63] N. Sbihi, Algorithme de recherche d'un stable de cardinalité maximum dans un graphe sans étoile, *Discrete Mathematics* 29 (1), 53–76.

[64] E.F. Schmeichel and S.L. Hakimi, A cycle structure theorem for Hamiltonian graphs. *J. Combin. Theory Ser. B* 45 (1988), 99–107.

[65] E.F. Schmeichel and J. Mitchem, Bipartite graphs with cycles of all even lengths. *J. Graph Theory* 6 (1982), 429–439.

[66] A. Schrijver, Combinatorial Optimization Polyhedra and Efficiency, Springer-Verlag Berlin Heidelberg 2004.

[67] J.B. Shearer, A note on the independence number of triangle-free graphs. II, *J. Comb. Theory, Ser. B* **53** (1991), 300–307.

[68] J.B. Shearer, The independence number of dense graphs with large odd girth, *Electron. J. Comb.* **2** (1995), N2.

[69] M. Simonovits, A new proof and generalizations of a theorem of Erdős and Posa on graphs without $k+1$ independent circuits, *Acta Math. Acad. Sci. Hung.* **18** (1967), 191–206.

[70] W. Staton, Some Ramsey-type numbers and the independence ratio, *Trans. Amer. Math. Soc.* **256** (1979), 353–370.

[71] R. Tarjan, Depth-first search and linear graph algorithms, *SIAM J. Computing* **1** (1972), 146–160.

[72] C. Thomassen, On the Chromatic Number of Triangle-Free Graphs of Large Minimum Degree, *Combinatorica* **22** (2002), 591–596.

[73] H.-J. Voss, Über die Taillenweite in Graphen, die genau k knotenunabhängige Kreise enthalten, und über die Anzahl der Knotenpunkte, die in solchen Graphen alle Kreise repräsentieren, Dissertationsschrift TH Ilmenau 1966.

[74] H. Walther and H.-J. Voss, Über Kreise in Graphen, Berlin: VEB Deutscher Verlag der Wissenschaften. 271 S. m. 99 Abb. (1974).

[75] H. Wang, Large vertex-disjoint cycles in a bipartite graph, *Graphs Comb.* **16** (2000), 359–366.

[76] H. Wang, On independent cycles in a bipartite graph, *Graphs Comb.* **17** (2001), 177–183.

[77] H. Wang, Maximal total length of k disjoint cycles in bipartite graphs, *Combinatorica* **25** (2005), 367-377.

[78] V.K. Wei, A Lower Bound on the Stability Number of a Simple Graph, Technical memorandum, TM 81 - 11217 - 9, Bell laboratories, 1981.

[79] H. Whitney, Non-separable and planar graphs, *Trans. Amer. Math. Soc.* **34** (1932), 339–362.

[80] X. Zhu, Bipartite subgraphs of triangle-free subcubic graphs, *J. Comb. Theory, Ser. B* **99** (2009), 62-83.