N° d'ordre : 3083

# THÈSE

PRÉSENTÉE À

# L'UNIVERSITÉ BORDEAUX I

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET D'INFORMATIQUE

Par **Pascal Ochem**

POUR OBTENIR LE GRADE DE

## DOCTEUR

SPÉCIALITÉ : INFORMATIQUE

---

**Graph coloring and combinatorics on words**

---

**Soutenue le :** 24 novembre 2005

**Après avis des rapporteurs :**

| | |
|---|---|
| Jean Berstel ....... | Professeur |
| Jan Kratochvíl .... | Professeur |

**Devant la commission d'examen composée de :**

| | | |
|---|---|---|
| Jean Berstel ....... | Professeur ................. | Rapporteur |
| Robert Cori ....... | Professeur ................. | Président |
| Jan Kratochvíl .... | Professeur ................. | Rapporteur |
| André Raspaud .... | Professeur ................. | Examinateur |
| Gwénaël Richomme | Maître de Conférences HDR | Examinateur |
| Éric Sopena ....... | Professeur ................. | Directeur de thèse |

– 2005 –

# Remerciements

Je tiens tout d'abord à remercier Éric Sopena, mon directeur de thèse, pour avoir fait confiance à Dieter Kratsch, pour ses conseils toujours avisés, et pour m'avoir permis et même encouragé à travailler aussi sur les mots.

I warmly thank Jan Kratochvíl who reviewed my thesis and came from Prague to attend to the defense.

Je remercie Jean Berstel d'avoir été rapporteur et pour ses conseils concernant les exposés oraux.

J'ai apprécié que André Raspaud et Gwénaël Richomme, avec qui j'ai eu l'occasion de travailler, aient participé à mon jury. Je les remercie, ainsi que Robert Cori pour avoir présidé ce jury.

Je n'oublie pas les doctorants et parfois co-auteurs du LaBRI, notamment Daniel, Mickaël, Alexandre, Ismaïl, Abdelaziz, Fabrice et Jérémie.

Enfin, je remercie mes parents et Marie-Thérèse.

# First part

# Graph coloring

# Chapter 1

# Preliminaries

## 1.1 Graphs

A simple graph is a set of vertices and a set of edges. An edge is a subset of the vertex set of size two. An oriented graph is an orientation of an undirected graph, obtained by assigning to every edge one of the two possible orientations. An oriented edge is called an arc. If $G$ is a graph, $V(G)$ denotes its vertex set, $E(G)$ denotes its set of edges (or arcs if $G$ is an oriented graph). A graph is planar if it can be drawn in the plane without edges (or arcs) crossing. A plane graph is particular planar embedding of a planar graph. The set of faces of a plane graph $G$ is denoted $F(G)$. The girth of a graph is the length of its shortest cycle.

Let us now introduce other well-known graph classes. A graph is outerplanar or 1-outerplanar if it has a planar embedding such that every vertex belongs to the outerface. For $k \geqslant 2$, a graph is $k$-outerplanar if it has a planar embedding such that the vertices not contained in the outerface induce a $(k-1)$-outerplanar graph. A complete graph on $k$ vertices is called a clique on $k$ vertices or a $k$-clique. Partial $k$-trees are the subgraphs of $k$-trees, which are defined as follows: the $k$-clique is a $k$-tree and every other $k$-tree $G$ contains a vertex $v$ whose neighborhood induces a $k$-clique and is such that $G \setminus v$ is a $k$-tree. Finally, a graph is $k$-degenerate if all of its non-empty subgraphs contain a vertex of degree at most $k$. We use in this thesis the following notations:

$\mathcal{P}_k$ denotes the class of planar graphs with girth at least $k$.

$out(k)$ denotes the class of $k$-outerplanar graphs.

$\mathcal{S}_k$ denotes the class of graphs with maximum degree at most $k$.

$\mathcal{T}_k$ denotes the class of partial $k$-trees.

$\mathcal{D}_k$ denotes the class of $k$-degenerate graphs.

$bip$ denotes the class of bipartite graphs.

A proper vertex coloring of a simple graph $G$ is an assignment $c$ of colors to the vertices of $G$ such that $c(u) \neq c(v)$ if the vertices $u$ and $v$ are adjacent in $G$. A $k$-coloring is a proper vertex coloring using $k$ colors. The chromatic number $\chi(G)$ is the smallest integer $k$ such that $G$ has a $k$-coloring. A graph is bipartite if it has 2-coloring. In the next sections, we introduce other types of coloring and the corresponding chromatic numbers, if it is defined. For any type of chromatic number $\chi_{\mathbf{x}}$ of a graph, we also define the (possibly infinite) chromatic number $\chi_{\mathbf{x}}(\mathcal{C})$ of a graph class $\mathcal{C}$ as the maximum of $\chi_{\mathbf{x}}(G)$ taken over every graph $G \in \mathcal{C}$.

The four color Theorem and Grötzsch's Theorem state respectively that planar graphs are 4-colorable and triangle-free planar graphs are 3-colorable.

**Theorem 1.1.**

1. *$\chi(\mathcal{P}_3) = 4$ [2, 3].*
2. *$\chi(\mathcal{P}_4) = 3$ [31].*

## 1.2 Oriented colorings

A homomorphism from an oriented graph $G$ to an oriented graph $H$ is a mapping $\varphi$ from $V(G)$ to $V(H)$ which preserves the arcs, that is $(x, y) \in E(G) \Longrightarrow (\varphi(x), \varphi(y)) \in E(H)$. We say that $H$ is a *target graph* of $G$ if there exists a homomorphism from $G$ to $H$. The oriented chromatic number $\chi_o(G)$ of an oriented graph $G$ is defined as the smallest order of a target graph of $G$. The oriented chromatic number $\chi_o(G)$ of an undirected graph $G$ is then defined as the maximum oriented chromatic number of its orientations. We will say that a graph $G$ is *$H$-colorable* if $H$ is a target graph of $G$ and the vertices of $H$ will be called *colors*.

Raspaud and Nešetřil introduced in [55] the *strong oriented chromatic number* of an oriented graph $G$ (denoted $\chi_s(G)$), which definition differs from that of $\chi_o(G)$ by requiring that the target graph is an oriented Cayley graph. Various bounds on the (strong) oriented chromatic number have been found for subclasses of planar graphs:

**Theorem 1.2.**

1. *$\chi_o(\mathcal{P}_3) \leqslant 80$ [64].*
2. *$\chi_s(\mathcal{P}_3) \leqslant 271$ [50].*
3. *$\chi_o(\mathcal{P}_3) \geqslant 17$ [49].*
4. *$\chi_s(\mathcal{P}_5) \leqslant 19$ [10].*
5. *$\chi_s(\mathcal{P}_6) \leqslant 11$ [10].*
6. *$\chi_s(\mathcal{P}_8) \leqslant 7$ [10].*
7. *$\chi_s(\mathcal{P}_{14}) = 5$ [10, 56].*
8. *$\chi_o(out(1)) = \chi_s(\mathcal{T}_2) = 7$ [70].*

## 1.3 Acyclic improper colorings

A vertex coloring $c$ of a graph $G$ is *acyclic* if for every two distinct colors $i$ and $j$, the edges $uv$ such that $c(u) = i$ and $c(v) = j$ induce a forest. A cycle or a path is said to be *alternating* if it is properly colored with two colors. Notice that only even cycles can be alternating and that a coloring is acyclic if and only if there exists no alternating cycle. The acyclic chromatic number $\chi_a(G)$ is the smallest number of colors needed in an acyclic proper coloring of the graph $G$. The next theorem summarizes the known bounds on the acyclic chromatic number of planar graphs with given girth.

**Theorem 1.3.**

1. $\chi_a(\mathcal{P}_3) = 5$ *[8].*
2. $\chi_a(\mathcal{P}_4 \cap bip \cap \mathcal{D}_2) = 5$ *[41].*
3. $\chi_a(\mathcal{P}_5) \leqslant 4$ *[11].*
4. $\chi_a(\mathcal{P}_7) \leqslant 3$ *[11].*

Raspaud and Sopena showed that the oriented chromatic number of a graph is bounded in terms of its acyclic chromatic number:

**Proposition 1.1.** *[64] For every graph $G$ such that $\chi_a(G) = k$, $\chi_o(G) \leqslant k2^{k-1}$.*

Kostochka, Sopena, and Zhu then proved that the acyclic chromatic number of a graph is bounded in terms of its oriented chromatic number :

**Proposition 1.2.** *[42] For every graph $G$ such that $\chi_o(G) = k \geqslant 4$, $\chi_a(G) \leqslant k^2 + k^{3+\lceil \log_2 \log_2 k \rceil}$.*

The result of Borodin that planar graphs are acyclically 5-colorable (i.e. $\chi_a(\mathcal{P}_3) = 5$) thus implies that the oriented chromatic number of a planar graph is at most 80 (i.e. $\chi_o(\mathcal{P}_3) \leqslant 80$), which is yet the best known upper bound. In order to get a better upper bound on $\chi_o(\mathcal{P}_3)$, if possible, it is interesting to study the tightness of Proposition 1.1, in particular for $k = 5$. The previously best known lower bound on the maximum value of $\chi_o(G)$ in terms of $\chi_a(G)$ was given in Vignal's thesis [77] with a family of graphs $G_k$, $k \geqslant 1$ such that $\chi_a(G_k) = k$ and $\chi_o(G_k) = 2^k - 1$.

Improper colorings are defined as follows: A graph $G$ belongs to the class $\mathcal{C}_0 \circ \cdots \circ \mathcal{C}_{k-1}$ if and only if $G$ has a $k$-coloring such that the $i^{th}$ color class induces a graph in $\mathcal{C}_i$, for $0 \leqslant i \leqslant k-1$. Boiron, Sopena, and Vignal introduced the notion of acyclic improper coloring [7]. Let $\mathcal{C}_0, \ldots, \mathcal{C}_{k-1}$ be graph classes. A graph $G$ belongs to the class $\mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-1}$ if and only if $G$ has an acyclic $k$-coloring such that the $i^{th}$ color class induces a graph in $\mathcal{C}_i$, for $0 \leqslant i \leqslant k-1$. For brevity, if $\mathcal{C}_0 = \cdots = \mathcal{C}_{k-1} = \mathcal{C}$ we will denote by $\mathcal{C}^k$ the class $\mathcal{C}_0 \circ \cdots \circ \mathcal{C}_{k-1}$ and by $\mathcal{C}^{(k)}$ the class $\mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-1}$.

The main motivation in the study of acyclic improper colorings is the following generalization of Proposition 1.1.

**Proposition 1.3.** *[7] Let $\mathcal{C}_0, \ldots, \mathcal{C}_{k-1}$ be graph classes such that $\chi_o(\mathcal{C}_i) = n_i$, for $0 \leqslant i < k$. Every graph $G \in \mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-1}$ satisfies $\chi_o(G) \leqslant 2^{k-1} \sum_{i=0}^{i<k} n_i$.*

The bound of Proposition 1.3 is shown to be tight for $k \geqslant 3$ under mild assumptions in Section 2.1.

We know from [69, 71] that $\chi_o(\mathcal{T}_3) = \chi_o(\mathcal{T}_3 \cap \mathcal{P}_3) = 16$. Thus, Boiron et al. [7] pointed out that:

1. $\mathcal{P}_3 \subset \mathcal{T}_3 \odot \mathcal{S}_0 \odot \mathcal{S}_0$ would imply that $\chi_o(\mathcal{P}_3) \leqslant 72$,

2. $\mathcal{P}_3 \subset \mathcal{T}_3 \odot \mathcal{S}_1 \odot \mathcal{S}_0$ would imply that $\chi_o(\mathcal{P}_3) \leqslant 76$.

We will see that the second point is meaningless. Indeed, we have that $\mathcal{P}_3 \subset \mathcal{T}_3 \odot \mathcal{S}_1 \odot \mathcal{S}_0 \iff \mathcal{P}_3 \subset \mathcal{T}_3 \odot \mathcal{S}_0 \odot \mathcal{S}_0$ from a general result on acyclic improper colorings of planar graphs given in Section 2.2.

An acyclic improper coloring is interesting with respect to Proposition 1.3 if every color class has bounded oriented chromatic number. That is why we often use as color classes the graph classes $\mathcal{S}_0$, $\mathcal{S}_1$, and $\mathcal{D}_1$: We have that $\chi_o(\mathcal{S}_0) = 1$, $\chi_o(\mathcal{S}_1) = 2$, and $\chi_o(\mathcal{D}_1) = 3$.

## 1.4 List colorings

For any type $\mathbf{x}$ of vertex coloring with a notion of chromatic number, we can define a list-version of $\mathbf{x}$ coloring. A graph $G$ is $\mathbf{x}$ $L$-colorable if for a given list assignment $L = \{L(v) : v \in V(G)\}$ there exists a coloring $c$ of $G$ such that $c(v) \in L(v)$ for every vertex $v \in V(G)$ and $c$ is an $\mathbf{x}$ coloring of $G$. If $G$ is $\mathbf{x}$ $L$-colorable for every list assignment with $|L(v)| \geqslant k$ for all $v \in V(G)$, then $G$ is said $\mathbf{x}$ $k$-choosable. The parameter $\chi^l_{\mathbf{x}}(G)$ is the smallest integer $k$ such that $G$ is $\mathbf{x}$ $L$-colorable.

For any type of chromatic number $\chi_{\mathbf{x}}$ of a graph, we also define the (possibly infinite) chromatic number $\chi_{\mathbf{x}}(\mathcal{C})$ of a graph class $\mathcal{C}$ as the maximum of $\chi_{\mathbf{x}}(G)$ taken over every graph $G \in \mathcal{C}$.

Thomassen proved that every planar graph is properly 5-choosable [74] and Voigt proved that there exist planar graphs which are not properly 4-choosable [79], thus $\chi^l(\mathcal{P}_3) = 5$. Borodin et al. studied the acyclic choosability of planar graphs and obtained $\chi^l_a(\mathcal{P}_3) \leqslant 7$ [9], they also conjectured that every planar graph is acyclically 5-choosable. The following Proposition provides two examples of coloring results that easily extend to their list-version.

**Proposition 1.4.**

$\chi^l_a(\mathcal{T}_k) = k + 1.$

$\chi^l(\mathcal{D}_k) = k + 1.$

In Section 2.6, we investigate the acyclic choosability of some sparse graphs.

## 1.5 Edge colorings

Many graph parameters in the litterature are defined as the minimum size of a partition of the edges of the graph such that each part belongs to some graph class $C$. The most common is the chromatic index $\chi'(G)$, in this case $C$ is the class of graphs with maximum degree one. Vizing [78] proved that $\chi'(G)$ either equals $\Delta(G)$ or $\Delta(G)+1$. Deciding whether $\chi'(G)=3$ is shown to be NP-complete for general graphs in [36]. The arboricity $a(G)$ is another well studied parameter, for which $C$ is the class of forests. In [54], Nash-Williams proved that:

$$a(G) = \max_{H \subseteq G} \left\lceil \frac{|E_H|}{|V_H|-1} \right\rceil \tag{1.1}$$

with the maximum being over all the subgraphs of $G$. Even with this nice formula, the polynomial algorithm computing the arboricity of a graph is not trivial [35]. Other similar parameters have been studied. A *star* is a tree of diameter at most two. A *caterpillar* is a tree whose non-leaf vertices form a path. For the *star arboricity* $sa(G)$ (resp. *linear arboricity* $la(G)$, *caterpillar arboricity* $ca(G)$), the corresponding class $C$ is the class of star forests (resp. linear forests, caterpillar forests). Since paths are caterpillars and since stars are caterpillars which are trees, we have the following two inequalities for any graph $G$.

$$la(G) \geqslant ca(G) \tag{1.2}$$

$$sa(G) \geqslant ca(G) \geqslant a(G) \tag{1.3}$$

Since trees are easily partitionable into two forests of stars we have that:

$$2 \times a(G) \geqslant sa(G) \tag{1.4}$$

Hakimi et al. [34] showed the following inequality.

$$\chi_a(G) \geqslant sa(G) \tag{1.5}$$

Other interesting graph parameters include the track number $t(G)$ [33, 43] and the sub-chromatic index $\chi'_{sub}(G)$ [23], for which $C$ is respectively the class of interval graphs and the class of forests of stars and triangles. Notice that the class of triangle-free interval graphs is equivalent to the class of caterpillar forests. Thus, if $G$ is triangle-free, then $t(G)=ca(G)$ and $\chi'_{sub}(G)=sa(G)$.

A *$T$-free forest* is a forest without subgraphs isomorphic to $T$. For example, the $P_n$-free forests and the $K_{1,n}$-free forests correspond to, respectively, the forests with diameter at most $n-2$ and to the forests with degree at most $n-1$. We define the $T$-free arboricity $T\text{-}fa(G)$ of a graph $G$ as the minimum number of $T$-free forests needed to cover the edges of $G$. Using this terminology, we can redefine some of the parameters we introduced. The chromatic index, the star arboricity, the linear arboricity, and the caterpillar arboricity correspond to, respectively, the $P_3$-free arboricity, the $P_4$-free arboricity, the $K_{1,3}$-free arboricity, and the $S_3$-free arboricity.

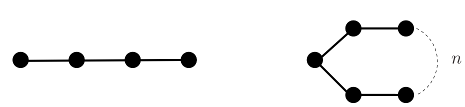


Figure 1.1: The path $P_4$ and the tree $S_n$.

If a tree $T_1$ is a subtree of a tree $T_2$, then $T_1\text{-}fa(G) \geqslant T_2\text{-}fa(G)$. So, the poset of trees produces a poset of arboricities. In this thesis, we focus on the chain $P_4 \subset S_2 \subset S_3 \subset S_n$, for

$n \geqslant 4$. We study the maximum value of these parameters, taken over the graphs of a class $\mathcal{C}$, $T\text{-}fa(\mathcal{C}) = \max\{T\text{-}fa(G),\ G \in \mathcal{C}\}$. We mainly consider the classes of planar graphs with girth at least $g$ (i.e. $\mathcal{P}_g$). Since $\mathcal{P}_{g+1} \subset \mathcal{P}_g$, we have $T\text{-}fa(\mathcal{P}_g) \geqslant T\text{-}fa(\mathcal{P}_{g+1})$ for any tree $T$. For a graph class $\mathcal{C}$, we define $U\text{-}fa(\mathcal{C})$ as the minimum value of $T\text{-}fa(\mathcal{C})$, taken over every finite tree $T$. Thus, for every tree $T$ and every class $\mathcal{C}$, we have that:

$$T\text{-}fa(\mathcal{C}) \geqslant U\text{-}fa(\mathcal{C}) \geqslant a(\mathcal{C}) \tag{1.6}$$

Theorem 3.2 shows in particular that $\mathcal{P}_4$ and $\mathcal{T}_2$ are examples of graph classes $\mathcal{C}$ such that the strict inequality $U\text{-}fa(\mathcal{C}) > a(\mathcal{C})$ holds. By the above relations, for any class $\mathcal{C}$, the parameters listed in Table 3.4 are ordered as follows:

$$\chi_a(\mathcal{C}) \geqslant P_4\text{-}fa(\mathcal{C}) \geqslant S_2\text{-}fa(\mathcal{C}) \geqslant S_3\text{-}fa(\mathcal{C}) \geqslant S_4\text{-}fa(\mathcal{C}) \geqslant U\text{-}fa(\mathcal{C}) \geqslant a(\mathcal{C}) \tag{1.7}$$

## 1.6 Forbidden configurations and maximum average degree

Let us explain in general the method we use to prove coloring results. Such a result has the form "every graph $G \in \mathcal{C}$ has a coloring $c$". We define a partial order $\prec$ on the set of graphs that extends the subgraph partial order. Then we consider a potential counter-example $G \in \mathcal{C}$ which is not colorable and is minimal with this property according to $\prec$. This means that if $H \in \mathcal{C}$, $H \neq G$, and $H \prec G$, then $H$ is colorable. Now, we provide a set $S$ of configurations that $G$ cannot contain due to its minimality property. To show that a configuration $C \in S$ is forbidden, we suppose that $G$ contains $C$ and proceed as follows:

1. We find a suitable graph $H$ such that $H \in \mathcal{C}$, $H \neq G$, and $H \prec G$.
2. We show that any coloring of $H$ induces a coloring of a proper subgraph $G'$ of $G$ that can be extended in a coloring of $G$.

This is a contradiction because $H$ is supposed to be colorable but not $G$. Thus $G$ cannot contain the forbidden configuration $C$. To finish the proof, we show that no counter-example exists because every graph in $\mathcal{C}$ contains at least one configuration in $S$.

There are several ways to do this last step, i.e. the unavoidability of $S$, most often using a *discharging method*. We now present the discharging method used by Borodin et al. [10] to obtain Theorems 1.2.(4) to 1.2.(7). It uses the following graph parameter.

**Definition 1.1.** *Let $G$ be a graph, the maximum average degree of $G$, denoted by $mad(G)$, is:*

$$mad(G) = \max\{2|E(H)|/|V(H)|,\ H \subseteq G\}$$

We assign to every vertex $v$ of the counter-example $G$ an initial charge equal to its degree $d(v)$ and define a *discharging procedure* that preserves the total charge. Then, we show that if the discharging procedure is applied to a graph $K$ avoiding $S$, then the final charge $d^*(v)$ of every vertex $v \in V(K)$ satisfies $d^*(v) \geqslant q$. We thus have

$$mad(K) \geqslant \frac{2|E(K)|}{|V(K)|} = \frac{\sum_{v \in V(K)} d(v)}{|V(K)|} = \frac{\sum_{v \in V(K)} d^*(v)}{|V(K)|} \geqslant \frac{q|V(K)|}{|V(K)|} = q.$$

This implies a statement of the form "every graph $G$ such that $mad(G) < q$ has a coloring $c$". We can get a corollary of such a statement for planar graphs thanks to the following well known observation:

**Observation 1.1.** $G \in \mathcal{P}_g \Longrightarrow mad(G) < \frac{2g}{g-2}$.

*Proof.* If a planar graph $G$ has girth $g$, then $|F(G)| \leqslant \frac{2|E(G)|}{g}$. By Euler's formula we get that $mad(G) = \frac{2|E(H)|}{|V(H)|} \leqslant \frac{2g|E(H)|}{2g+(g-2)|E(H)|} < \frac{2g}{g-2}$, for some graph $H \subseteq G$. $\square$

Let us finish this part with some conventions. We call respectively $k$-vertex, ${}^{\geqslant}k$-vertex, and ${}^{\leqslant}k$-vertex a vertex of degree $k$, $\geqslant k$, and $\leqslant k$. We also define $k$-neighbor, ${}^{\geqslant}k$-neighbor, and ${}^{\leqslant}k$-neighbor in the same way. In a figure representing a forbidden configuration, all the neighbors of "white" vertices are drawn, whereas "black" vertices may have other neighbors in the graph. Two or more black vertices may coincide in a single vertex, provided they do not share a common white neighbor.

# Chapter 2

# Vertex colorings

This chapter is devoted to various positive and negative results in graph coloring. We consider oriented coloring and acyclic coloring, which are closely related, and two variants of the latter, namely acyclic improper colorings and acyclic list coloring.

## 2.1 Acyclic improper coloring versus oriented coloring

The following theorem shows that the upper bound of Proposition 1.3 is best possible in many cases.

**Theorem 2.1.** *Let $k \geqslant 3$. Let $\mathcal{C}_0, \ldots, \mathcal{C}_{k-1}$ be hereditary graph classes closed under disjoint union, and such that $\chi_o(\mathcal{C}_i) = n_i$. Then $\chi_o(\mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-1}) = 2^{k-1} \sum_{i=0}^{i<k} n_i$.*

*Proof.* We construct an oriented graph $G \in \mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-1}$ such that $\chi_o(G) = 2^{k-1} \sum_{i=0}^{i<k} n_i$. Let $u_1, u_2, u_3$ be a directed 2-path with arcs $u_1u_2$ and $u_2u_3$, or $u_3u_2$ and $u_2u_1$. We say that $u_1$ and $u_3$ are the *endpoints* of the directed 2-path. By definition, the endpoints of the directed 2-path get distinct colors in any oriented coloring. Since $\chi_o(\mathcal{C}_i) = n_i$, there exists a witness oriented graph $W^i$ such that $\chi_o(W^i) = n_i$. The graph $G_i$ contains $k-1$ independent vertices $v^i_j$, $0 \leqslant j < k-1$ and $2^{k-1}$ disjoint copies $W^i_l$, $0 \leqslant l < 2^{k-1}$ of $W^i$. We consider the binary representation $l = \sum_{n=0}^{n<k-1} 2^n x_n(l)$ of $l$. For every two vertices $v^i_j$ and $u^i_l \in W^i_l$, we put the arc $v^i_j u^i_l$ (resp. $u^i_l v^i_j$) if $x_j(l) = 1$ (resp. $x_j(l) = 0$). If $l \neq l'$, their binary representations differ at the $n^{th}$ digit, thus $u^i_l \in W^i_l$ and $u^i_{l'} \in W^i_{l'}$ are the endpoints of a directed 2-path $u^i_l, v^i_n, u^i_{l'}$. So the same color cannot be used in distinct copies of $W^i$, which means that at least $2^{k-1} n_i$ colors are needed to color the copies of $W^i$ in any oriented coloring of $G_i$. To show that $G_i \in \mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-1}$, we acyclically color $G_i$ as follows. The $k-1$ vertices $v^i_j$ get pairwise distinct colors in $\{0, \ldots, k-1\} \setminus \{i\}$ and every vertex $u^i_l$ get color $i$ (that is why we need the "closed under disjoint union" assumption). Let $S_i$ denote the set of colors in some oriented coloring of the vertices $u^i_l$ of $G_i$. Now we take one copy of each graph $G_i$ and finish the construction of $G$. For every two vertices $u^i_l \in W^i_l$ and $u^{i'}_{l'} \in W^{i'}_{l'}$, such that $i \neq i'$, we add a new vertex $l$ and create a directed 2-path $u^i_l, l, u^{i'}_{l'}$. So, for $i \neq i'$, we have $S_i \cap S_{i'} = \emptyset$, which means that at least $2^{k-1} \sum_{i=0}^{i<k} n_i$ colors are needed in any oriented coloring of $G$. To obtain an acyclic coloring of $G$, the new vertex $l$ adjacent to $u^i_l$ and $u^{i'}_{l'}$ gets a color in $\{0, \ldots, k-1\} \setminus \{i, i'\}$, which is non-empty if $k \geqslant 3$. $\square$

Notice that Theorem 2.1 cannot be extended to the case $k = 2$ in general. By setting $k = 2$ and $\mathcal{C}_0 = \mathcal{C}_1 = \mathcal{S}_0$, we obtain the class of forests $\mathcal{S}_0^{(2)} = \mathcal{D}_1$. Proposition 1.3 provides the bound $\chi_o(\mathcal{S}_0^{(2)}) \leqslant 4$. This is not a tight bound, since oriented forests have a homomorphism to the oriented triangle, and thus $\chi_o(\mathcal{S}_0^{(2)}) = 3$.

The proof of Theorem 2.1 is constructive, but it does not help for the problem of determining $\chi_o(\mathcal{P}_3)$. Indeed, the graph corresponding to the proper 5-acyclic coloring (i.e. $k = 5$ and $\mathcal{C}_0 = \cdots = \mathcal{C}_4 = \mathcal{S}_0$) is highly non-planar (it contains e.g. $K_{32,48}$ as a minor).

## 2.2 Acyclic improper colorings of planar graphs

In this section, we study which acyclic improper colorings may be able to color every planar graph. The next result implies that in such colorings, a "too small" class is not more useful than an independent set.

**Theorem 2.2.** *Let $2 \leqslant k \leqslant 4$. Suppose that $\chi_o(\mathcal{C}_{k-1}) \leqslant 14$, then $\mathcal{P}_3 \subset \mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-2} \odot \mathcal{C}_{k-1} \Longleftrightarrow \mathcal{P}_3 \subset \mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-2} \odot S_0$.*

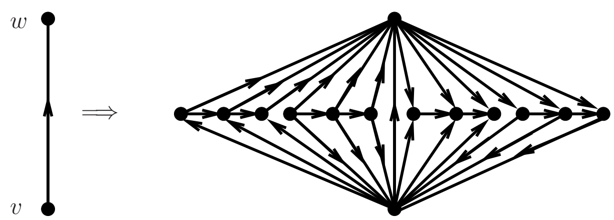


Figure 2.1: The graph we add to an arc.

*Proof.* Let $G$ be an oriented graph. The oriented graph $f(G)$ is obtained from $G$ by adding to every arc 12 vertices as described in Figure 2.1. We also define $f_n(G)$ such that $f_0(G) = G$ and $f_{n+1}(G) = f(f_n(G))$. Notice that if $G$ is planar, then $f(G)$ is also planar. We consider the family $G_k$ from [77] mentioned in Section 1.1, and in particular the oriented planar graph $G_4$ whose oriented chromatic number is 15. It is easy to check that $G_4$ is a subgraph of $f_2(K_2)$. Let us consider now any acyclic improper $k$-coloring $c$ of $f_n(K_2)$ such that $c(v) = c(w) = 0$. To avoid an alternating cycle $vxwy$ for some vertices $x$ and $y$, $f_k(K_2)$ must contain a monochromatic copy of $f_1(K_2)$ colored 0. By induction, $f_{i\times k}(K_2)$ must contain a monochromatic copy of $f_i(K_2)$ for $i \geqslant 1$. The "$\Longleftarrow$" implication of Theorem 2.2 holds by definition. We now prove the "$\Longrightarrow$" implication by contradiction. Suppose there exists an oriented planar witness graph $W$ such that $W \in \mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-2} \odot \mathcal{C}_{k-1}$ and $W \notin \mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-2} \odot S_0$. This means that any $\mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-2} \odot \mathcal{C}_{k-1}$-coloring of $W$ contains a monochromatic edge $vw$ colored $k-1$. So, by previous discussions, the graph $f_8(W)$ contains a monochromatic copy of $G_4$ colored $k-1$, which contradicts the requirement $\chi_o(\mathcal{C}_{k-1}) \leqslant 14$. $\square$

Theorem 2.2 allows us to study which statement of the form "every planar graph belongs to $\mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-1}$" may improve the upper bound $\chi_o(\mathcal{P}_3) \leqslant 80$. If $k = 4$, the "least" candidate class would be $\mathcal{C}_0 \odot S_0 \odot S_0 \odot S_0$ with $\chi_o(\mathcal{C}_0) = 15$, but the corresponding bound is too large:

$2^{4-1}(15+1+1+1) = 144 > 80$. If $k = 3$, there must be exactly one improper color, otherwise the least candidate, $\mathcal{C}_0 \odot \mathcal{C}_1 \odot S_0$ with $\chi_o(\mathcal{C}_0) = \chi_o(\mathcal{C}_1) = 15$, provides a too large bound: $2^{3-1}(15+15+1) = 124 > 80$.
The constant 14 in Theorem 2.2 can be improved to 15 using ideas in [71].
We have not been able to get similar results for planar graphs with larger girth.

## 2.3 Acyclic improper colorings and $k$-outerplanar graphs

A theorem of Boiron et al. [7] states that some planar graphs have no acyclic $\mathcal{T}_3^{(2)}$-coloring, actually they even showed that $out(3) \not\subset \mathcal{D}_3^{(2)}$. This result on acyclic improper colorings of planar graphs still holds with larger color classes.

**Theorem 2.3.** *$out(3) \not\subset \mathcal{F}_1 \odot \mathcal{F}_2$, where for $1 \leqslant i \leqslant 2$, $\mathcal{F}_i = \mathcal{D}_3$, $\mathcal{T}_4$, or $out(2)$.*

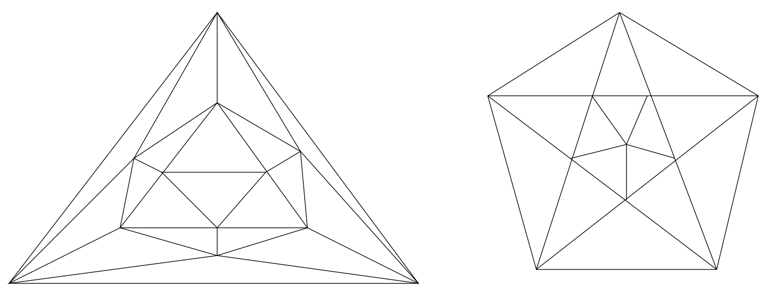

Figure 2.2: The graph $I$ (the icosahedron) and the graph $I^-$.

Let $I$ denote the icosahedron graph depicted in Figure 2.2 (left). The next lemma considers improper acyclic 2-colorings of $I$ without restriction on the color classes. Let $\mathcal{G}$ denote the class of all simple graphs.

**Lemma 2.1.** *Up to symetries, there are only two types of $\mathcal{G}^{(2)}$-coloring of the icosahedron:*

*(i) At most one vertex is colored 1 and all others are colored 2.*

*(ii) Two vertices at distance 3 are colored 1 and all others are colored 2.*

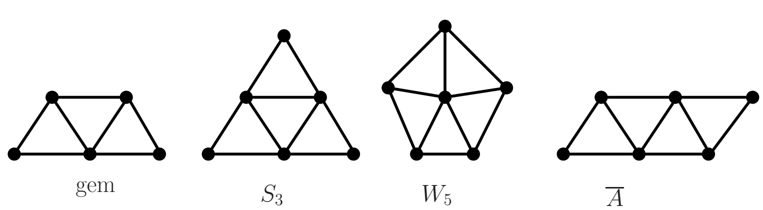


Figure 2.3: Small graphs.

*Proof.* We can assume w.l.o.g. that at most 6 vertices are colored 1. Suppose first that two adjacent vertices are colored 1. They have two common neighbors, so at least one of them must be colored 1 to avoid an alternating $C_4$. Thus we have a 1-monochromatic $K_3$. Three vertices outside of this $K_3$ are adjacent to two vertices of the $K_3$, so at least one of them must be colored 1 to avoid an alternating $C_6$. Thus we have a 1-monochromatic $K_4^-$. Four vertices outside of this $K_4^-$ are adjacent to two vertices of the $K_4^-$, and at least one of them must be colored 1 to avoid an alternating $C_8$. Thus we have a 1-monochromatic gem (see Figure 2.3). Four vertices outside of this gem are adjacent to at least two vertices of the gem, and at least one of them must be colored 1 to avoid an alternating $C_8$. Thus we have a 1-monochromatic subgraph $S$, which is either $S_3$, $W_5$, or $\overline{A}$) (see Figure 2.3). Since $|S| = 6$, $I \setminus S$ must be 2-monochromatic. We easily verify that for each possible $S$ there exists an alternating cycle in $I$. Suppose now that two vertices at distance two are colored 1. They have two common neighbors, and at least one of them must be colored 1 to avoid an alternating $C_4$. Thus we have a 1-monochromatic $P_3$ and we fall in the previous case. □

Let us denote by $I^-$ the graph obtained by deleting one vertex from $I$ (see Figure 2.2 (right)).

**Lemma 2.2.** *$I^-$ is neither 3-degenerate, 2-outerplanar, nor a partial 4-tree.*

*Proof.* Since the minimum degree of $I^-$ is four, it is not 3-degenerate. The graph $I^-$ is 3-connected, so it has a unique embedding on the sphere. Notice that $I^-$ contains four distinct non-equivalent types of faces: one of degree five and three types of triangles. For every face $F$, the graph obtained by removing the vertices incident to $F$ is not outerplanar, thus $I^-$ is not 2-outerplanar. Finally, to prove that $I^-$ is not a partial 4-tree, we show that we cannot obtain the empty graph from $I^-$ by repeatedly deleting a $\leqslant$4-vertex and placing a clique on its neighbors [67]. Any such vertex elimination ordering must start with one of the 4-vertices of the outerface of $I^-$, which all play the same role. Deleting a 4-vertex of $I^-$ and placing a clique on its neighbors gives a graph $J$. Now $J$ has two $\leqslant$4-vertices playing the same role. Deleting one of them and placing a clique on its neighbors gives a graph $K$. Since $K$ has minimum degree five, $I^-$ has no elimination ordering and thus is not a partial 4-tree. □

Lemmas 2.1 and 2.2 proves that the only improper acyclic colorings of $I$ with the color classes of Theorem 2.3 are of type (ii).

**Remark 2.1.** *If $I$ has a coloring of type (ii), then for every 2-monochromatic triangle $t$, there is a vertex colored 1 adjacent to two vertices of $t$.*

Consider now the graph $G$ depicted in Fig 2.4 (left) obtained from $K_4$ by identifying each of the 3 marked faces with the outerface of a copy of an icosahedron.

**Lemma 2.3.** *If $G$ is acyclically 2-colored such that every copy of $I$ has a coloring of type (ii), then the outer-face is monochromatic and there is an alternating path beetwen $a_0$ and $a_1$.*

*Proof.* Suppose the first part of statement is false and assume that $c(a_0) = c(a_1) = 2$ and $c(a_2) = 1$ (see Fig 2.4 (middle)). We must have $c(m) = 2$ to avoid an alternating cycle $a_2, a_0, m, a_1$. By remark 2.1, one $u_i$ must be colored 1, and this creates an alternating $C_4$, a contradiction. Now we check the last part of statement (see Fig 2.4 (right)). By the previous

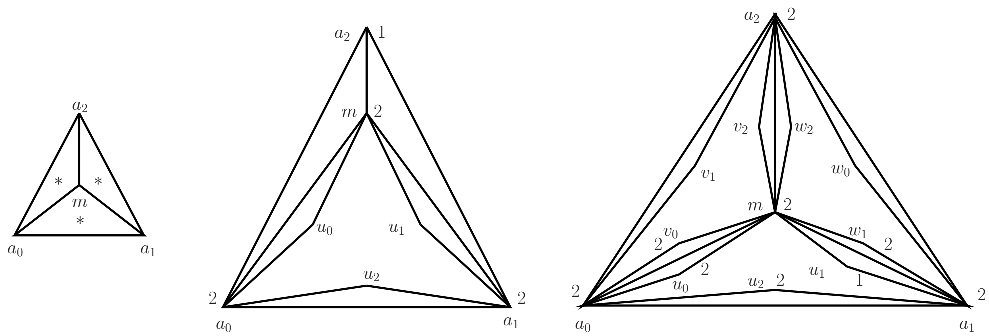


Figure 2.4: The graph G considered in Lemma 2.3.

discussion, the vertices $a_i$ are colored 2, and $c(m) = c(u_2) = 2$ to avoid an alternating path beetwen $a_0$ and $a_1$. By remark 2.1, one of the $u_i$'s must be colored 1 and we suppose w.l.o.g. that $u_1 = 1$. This forces $c(w_1) = 2$ to avoid an alternating $C_4$ and $u_0 = v_0 = 2$ to avoid an alternating path beetwen $a_0$ and $a_1$. By remark 2.1, either $v_1$ or $v_2$ (resp. $w_1$ or $w_0$) must be colored 1. In these 4 cases we have either an alternating cycle or an alternating path beetwen $a_0$ and $a_1$. □

To finish the proof of Theorem 2.3, we take two copies $G'$ and $G''$ of $G$ and we identify $a'_0$ and $a''_0$ (resp. $a'_1$ and $a''_1$) to obtain the 3-outerplanar graph $G^*$. By the previous lemmas, if $G'$ and $G''$ are both colored as in Lemma 2.3, then there exist one alternating path beetwen $a'_0$ and $a'_1$ in $G'$ and another one in $G''$, which create an alternating cycle in $G^*$.

We now consider acyclic improper colorings of partial $k$-trees and show that the equality $\chi_a(\mathcal{T}_k) = k+1$ is best possible in this context.

**Theorem 2.4.** *For every $k \in \mathbb{N}^*$ and for every $G \in \mathcal{T}_k$, $\mathcal{T}_k \not\subset$ ($G$-free)$^{(k)}$.*

*Proof.* The case $k = 1$ is obvious, so assume $k \geqslant 2$ is a fixed integer in the following. Let us call *good* a clique $c$ such that $2 \leqslant |c| \leqslant k$. Now we define the graphs $U_{k,n}$, $n \geqslant 1$, such that:

1. $U_{k,1} = K_2$.
2. For each good clique $c$ of $U_{k,n}$, we add a new vertex adjacent to every vertex of $c$ to obtain $U_{k,n+1}$.

Clearly, every graph in $\mathcal{T}_k$ is a subgraph of $U_{k,n}$ for some $n$. To finish the proof, we will show that in any improper acyclic $k$-coloring, $U_{k,n\times k}$ contains a monochromatic copy of $U_{k,n}$. For $n = 1$, we have that $U_{k,k}$ contains a clique $K_{k+1}$, and thus contains a monochromatic $K_2$. Now assume that $U_{k,n\times k}$ contains a monochromatic copy of $U_{k,n}$. For every good clique $c$ of that copy there are $k$ new vertices adjacent to $c$ in $U_{k,n\times k+k}$, and one of these $k$ new vertices must get the same color as $c$. This implies that $U_{k,(n+1)\times k}$ contains a monochromatic copy of $U_{k,n+1}$. □

Finally, we obtain the following result on acyclic improper colorings of $k$-outerplanar graphs.

**Theorem 2.5.** $out(k+1) \subset \mathcal{S}_0 \odot \mathcal{S}_0 \odot \mathcal{S}_0 \odot out(k)$

*Proof.* We show that there always exists a coloring of a $(k+1)$-outerplanar graph such that any vertex of the outerface gets one of the first three colors and all other vertices are colored with the last color. Let us characterize a counter-example $T$ for the specified coloring with minimal number of vertices. The special type of coloration considered allows us to assume w.l.o.g. that the non outer-vertices of $T$ induce an indepedent set, as a potential alternating cycle contains no monochromatic edge. We can also add a maximal number of edges connecting outer-vertices of $T$ inside the outerface. This way, the neighborhood of any non outer-vertex induces a single vertex, a $K_2$, or a cycle of the outerface. The outerface cannot contain a vertex cut of size two. These two vertices would be adjacent, therefore they would get different colors in a valid coloring of "one part" of $T$. We could extend this coloring to the whole graph, as a potential minimal alternating cycle cannot lie on both parts. The only remaining possibility is that $T$ is a wheel, and a wheel has the specified coloring. $\square$

## 2.4 Paley tournaments and $S_{k,n}$ properties

For a prime power $q \equiv 3 \pmod 4$, the vertices of the Paley tournament $QR_q$ are the elements of the finite field $\mathbb{F}_q$ and $(i,j)$ is an arc in $QR_q$ if and only if $j-i$ is a non-zero quadratic residue of $\mathbb{F}_q$. The condition $q \equiv 3 \pmod 4$ ensures that for every $i \neq 0$, $i$ is a quadratic residue if and only if $-i$ is a quadratic non-residue. Paley tournaments are oriented Cayley graphs with addition in $\mathbb{F}_q$ as operation and non-zero quadratic residues as generators. An interesting property of Paley tournaments is the arc-transitivity. For any arc $(i,j) \in E(QR_q)$, the mapping $x \to \frac{x-i}{j-i}$ is an automorphism of $QR_q$ which maps $(i,j)$ to $(0,1)$. So every arc in $QR_q$ plays the same role. If $K$ is an oriented graph, let $K^R$ be the graph obtain by reversing every arc in $K$. The tournament $QR_q$ is self-reverse in the sense that is $QR_q$ automorphic to $QR_q^R$ by the mapping $x \to -x$. Notice that $q = p^r \equiv 3 \pmod 4$ if and only if $p \equiv 3 \pmod 4$ and $r \equiv 1 \pmod 2$. An *orientation vector* of size $k$ is a sequence $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ in $\{0,1\}^k$. let $G$ be an oriented graph and $X = (x_1, x_2, \ldots, x_k)$ be a sequence of pairwise distinct vertices of $G$. A vertex $y$ of $G$ is said to be an *$\alpha$-successor of* $X$ if for every $i$, $1 \leqslant i \leqslant k$, we have $\alpha_i = 1 \Longrightarrow (x_i, y) \in E(G)$ and $\alpha_i = 0 \Longrightarrow (y, x_i) \in E(G)$. The graph $G$ satisfies property $S_{k,n}$ if for every sequence $X = (s_1, s_2, \ldots, s_k)$ of $k$ pairwise distinct vertices of $G$, and for every orientation vector $\alpha$ of size $k$, there exist at least $n$ vertices in $V(G)$ which are $\alpha$-successors of $X$. The next lemma provides some useful $S_{k,n}$ properties of Paley tournaments.

**Lemma 2.4.**

1. *The Paley tournament $QR_q$ satisfies $S_{1,\lfloor \frac{q}{2} \rfloor}$.*

2. *The Paley tournament $QR_q$ satisfies $S_{2,\lfloor \frac{q}{4} \rfloor}$.*

3. *The Paley tournament $QR_{27}$ satisfies $S_{3,2}$.*

4. *The Paley tournament $QR_{59}$ satisfies $S_{3,5}$.*

*Proof.* We easily check that $QR_q$ satisfies property $S_{1,\lfloor \frac{q}{2} \rfloor}$, which simply means that every vertex has $\lfloor \frac{q}{2} \rfloor$ successors and $\lfloor \frac{q}{2} \rfloor$ predecessors. Proposition 1 in [10] means that $QR_q$ also satisfies property $S_{2,\lfloor \frac{q}{4} \rfloor}$. We now prove the two statements of the form "$QR_q$ satisfies $S_{3,n}$". Note that the order of the vertices in a sequence does not matter. Thus, since any oriented triangle contains a directed 2-path, we only have to consider sequences $(s_1, s_2, s_3)$ in which $(s_1, s_2)$ and $(s_2, s_3)$ are arcs of $QR_q$. Now, by the arc-transitivity of $QR_q$, we only need to check the property on sequences of the form $(0, 1, v)$ such that $v \in \mathbb{F} \setminus \{0, 1\}$ and $v-1$ is a quadratic residue of $\mathbb{F}$. Let us write $\langle v_1, v_2, v_3 \rangle$ if and only if $v_1 \neq v_2$ and there are automorphisms of $QR_q$ mapping $(0, 1, v_1)$ to $(v_2, 0, 1)$ and $(1, v_3, 0)$. We easily see that if $\langle v_1, v_2, v_3 \rangle$ and $(0, 1, v_1)$ is checked, then $(0, 1, v_2)$ and $(0, 1, v_3)$ need no check. In the case of $QR_{27}$, we have $\langle x^5, x^{19}, x^{15} \rangle$ and $\langle x^7, x^{21}, x^{11} \rangle$ (see Figure 2.5 for the elements of $\mathbb{F}_{27}$). For every remaining sequence and for every orientation vector, two $\alpha$-successors are listed in Figure 2.6. In the case of $QR_{59}$, we have $\langle 2, 58, 30 \rangle$, $\langle 6, 47, 50 \rangle$, $\langle 8, 42, 23 \rangle$, $\langle 10, 13, 54 \rangle$, and $\langle 18, 52, 37 \rangle$ (see Figure 2.7 for quadratic residues). For every remaining sequence and for every orientation vector, five $\alpha$-successors are listed in Figure 2.8. $\square$

| element | (mod $3, x^3 - x + 1$) | element | (mod $3, x^3 - x + 1$) |
|---|---|---|---|
| 0 | 0 | | |
| 1 | 1 | $x^{13}$ | -1 |
| $x$ | $x$ | $x^{14}$ | $-x$ |
| $x^2$ | $x^2$ | $x^{15}$ | $-x^2$ |
| $x^3$ | $-1+x$ | $x^{16}$ | $1-x$ |
| $x^4$ | $-x+x^2$ | $x^{17}$ | $x-x^2$ |
| $x^5$ | $-1+x-x^2$ | $x^{18}$ | $1-x+x^2$ |
| $x^6$ | $1+x+x^2$ | $x^{19}$ | $-1-x-x^2$ |
| $x^7$ | $-1-x+x^2$ | $x^{20}$ | $1+x-x^2$ |
| $x^8$ | $-1-x^2$ | $x^{21}$ | $1+x^2$ |
| $x^9$ | $1+x$ | $x^{22}$ | $-1-x$ |
| $x^{10}$ | $x+x^2$ | $x^{23}$ | $-x-x^2$ |
| $x^{11}$ | $-1+x+x^2$ | $x^{24}$ | $1-x-x^2$ |
| $x^{12}$ | $-1+x^2$ | $x^{25}$ | $1-x^2$ |

Figure 2.5: Multiplicative and additive representations of the elements of $\mathbb{F}_{27}$.

| (0,1,$v$) | {0,0,0} | {0,0,1} | {0,1,0} | {0,1,1} | {1,0,0} | {1,0,1} | {1,1,0} | {1,1,1} |
|---|---|---|---|---|---|---|---|---|
| (0,1,$x^2$) | $x, x^3$ | $x^9, x^{17}$ | $x^5, x^{11}$ | $x^7, x^{13}$ | $x^{10}, x^{12}$ | $x^4, x^8$ | $x^6, x^{14}$ | $x^{16}, x^{18}$ |
| (0,1,$x^5$) | $x, x^{23}$ | $x^3, x^9$ | $x^7, x^{11}$ | $x^{13}, x^{15}$ | $x^{10}, x^{12}$ | $x^4, x^8$ | $x^{16}, x^{18}$ | $x^2, x^6$ |
| (0,1,$x^6$) | $x^3, x^9$ | $x, x^{17}$ | $x^5, x^7$ | $x^{11}, x^{13}$ | $x^4, x^{10}$ | $x^8, x^{12}$ | $x^{14}, x^{16}$ | $x^2, x^{22}$ |
| (0,1,$x^7$) | $x^3, x^9$ | $x, x^{17}$ | $x^{13}, x^{21}$ | $x^5, x^{11}$ | $x^{12}, x^{20}$ | $x^4, x^8$ | $x^2, x^{14}$ | $x^6, x^{16}$ |
| (0,1,$x^{13}$) | $x, x^3$ | $x^{17}, x^{23}$ | $x^5, x^{15}$ | $x^7, x^{11}$ | $x^8, x^{20}$ | $x^4, x^{10}$ | $x^2, x^6$ | $x^{14}, x^{16}$ |
| (0,1,$x^{14}$) | $x^{17}, x^{23}$ | $x, x^3$ | $x^5, x^{11}$ | $x^7, x^{19}$ | $x^8, x^{12}$ | $x^4, x^{10}$ | $x^{18}, x^{22}$ | $x^2, x^6$ |
| (0,1,$x^{16}$) | $x^{17}, x^{27}$ | $x, x^3$ | $x^7, x^{13}$ | $x^5, x^{11}$ | $x^{10}, x^{20}$ | $x^4, x^8$ | $x^2, x^{14}$ | $x^6, x^{18}$ |
| (0,1,$x^{18}$) | $x, x^9$ | $x^3, x^{23}$ | $x^{15}, x^{19}$ | $x^5, x^7$ | $x^4, x^{12}$ | $x^8, x^{10}$ | $x^2, x^{16}$ | $x^6, x^{14}$ |
| (0,1,$x^{22}$) | $x^{23}, x^{27}$ | $x, x^3$ | $x^5, x^{13}$ | $x^7, x^{11}$ | $x^4, x^8$ | $x^{10}, x^{12}$ | $x^6, x^{16}$ | $x^2, x^{14}$ |

Figure 2.6: The table of $\alpha$-successors in $QR_{27}$.

| |
|---|
| 0 1 3 4 5 7 9 12 15 16 17 19 20 21 22 25 26 27 28 29 35 36 41 45 46 48 49 51 53 57 |

Figure 2.7: The quadratic residues modulo 59.

| $(0,1,v)$ | {0,0,0}<br>{1,0,0} | {0,0,1}<br>{1,0,1} | {0,1,0}<br>{1,1,0} | {0,1,1}<br>{1,1,1} |
|---|---|---|---|---|
| (0,1,2) | 32.33.34.39.40<br>12.15.25.35.41 | 11.14.24.31.38<br>3.7.9.19.48 | 8.10.13.42.52<br>4.16.20.26.36 | 6.18.23.30.37<br>5.17.21.22.27 |
| (0,1,4) | 14.34.38.43.44<br>3.12.15.35.41 | 11.24.31.32.33<br>7.9.19.25.45 | 6.10.18.37.42<br>17.22.27.28.36 | 2.8.13.23.30<br>5.16.20.21.26 |
| (0,1,5) | 11.38.39.43.44<br>7.15.19.35.45 | 14.24.31.32.33<br>3.9.12.25.41 | 2.13.18.23.37<br>4.16.28.29.36 | 6.8.10.30.50<br>17.20.21.22.26 |
| (0,1,6) | 14.24.38.39.40<br>3.12.19.45.48 | 11.31.32.33.34<br>7.9.15.25.35 | 2.8.30.37.50<br>5.16.17.20.29 | 10.13.18.23.42<br>4.21.22.26.27 |
| (0,1,8) | 14.31.32.38.39<br>3.7.19.41.45 | 11.24.33.34.43<br>9.12.15.25.35 | 10.18.42.47.50<br>4.5.16.21.22 | 2.6.13.23.30<br>17.20.27.28.29 |
| (0,1,10) | 24.33.34.40.43<br>3.7.9.12.41 | 11.14.31.32.38<br>15.19.25.35.45 | 6.18.23.42.47<br>5.16.20.21.28 | 2.8.13.30.37<br>4.17.22.26.27 |
| (0,1,16) | 11.24.34.39.40<br>7.9.12.15.48 | 14.31.32.33.38<br>3.19.25.35.41 | 13.18.30.47.50<br>4.22.26.27.29 | 2.6.8.10.23<br>5.17.20.21.28 |
| (0,1,17) | 14.31.40.55.56<br>12.19.25.35.41 | 11.24.32.33.34<br>3.7.9.15.45 | 2.8.10.13.23<br>5.16.27.28.49 | 6.18.37.42.52<br>4.20.21.22.26 |
| (0,1,18) | 11.14.24.31.32<br>3.9.15.41.48 | 33.34.38.39.40<br>7.12.19.25.35 | 2.6.13.42.50<br>17.20.26.28.29 | 8.10.23.30.37<br>4.5.16.21.22 |
| (0,1,20) | 11.31.33.34.38<br>3.15.19.51.53 | 14.24.32.39.40<br>7.9.12.25.35 | 8.13.30.50.52<br>4.5.16.17.22 | 2.6.10.18.23<br>21.27.29.36.46 |
| (0,1,21) | 14.31.32.34.39<br>9.12.35.45.51 | 11.24.33.38.40<br>3.7.15.19.25 | 2.6.18.23.52<br>4.5.16.17.20 | 8.10.13.30.37<br>22.26.28.36.46 |
| (0,1,22) | 24.32.33.40.55<br>3.7.15.19.35 | 11.14.31.34.38<br>9.12.25.41.48 | 2.6.10.13.18<br>5.17.21.28.36 | 8.23.37.42.47<br>4.16.20.26.27 |
| (0,1,26) | 11.14.32.34.39<br>7.9.19.25.57 | 24.31.33.38.43<br>3.12.15.35.41 | 6.10.23.37.50<br>4.5.17.21.22 | 2.8.13.18.30<br>16.20.27.29.46 |
| (0,1,27) | 11.24.33.38.40<br>7.12.15.35.41 | 14.31.32.34.39<br>3.9.19.25.48 | 2.6.8.10.18<br>5.20.22.26.29 | 13.30.42.47.52<br>4.16.17.21.28 |
| (0,1,28) | 11.24.34.38.39<br>3.7.9.12.19 | 14.31.32.33.40<br>15.35.45.48.53 | 2.6.8.13.23<br>16.21.27.36.46 | 10.18.37.47.50<br>4.5.17.20.22 |
| (0,1,29) | 14.24.31.39.40<br>3.7.9.12.25 | 11.32.33.34.38<br>15.19.41.45.48 | 2.8.10.13.37<br>4.17.20.22.26 | 6.18.23.30.50<br>5.16.21.27.36 |
| (0,1,36) | 11.14.24.31.32<br>7.9.15.19.35 | 34.39.40.43.55<br>3.12.25.41.45 | 8.10.42.47.50<br>16.17.20.21.27 | 2.6.13.18.23<br>4.5.22.26.28 |
| (0,1,46) | 11.24.31.34.39<br>19.25.41.45.48 | 14.32.33.38.40<br>3.7.9.12.15 | 10.18.30.37.42<br>5.17.20.21.26 | 2.6.8.13.23<br>4.16.22.28.36 |
| (0,1,49) | 14.24.32.33.34<br>3.45.48.51.57 | 11.31.38.39.43<br>7.9.12.15.19 | 8.13.23.30.37<br>4.20.21.22.27 | 2.6.10.18.47<br>5.16.17.26.36 |
| (0,1,50) | 14.24.31.33.34<br>9.15.25.35.41 | 11.32.39.40.44<br>3.7.12.19.48 | 2.23.30.47.52<br>4.5.21.22.28 | 6.8.10.13.18<br>16.17.20.26.27 |

Figure 2.8: The table of $\alpha$-successors in $QR_{59}$.

## 2.5 Oriented colorings of triangle-free planar graphs

In this section, we focus on triangle-free planar graphs and prove the following theorem:

**Theorem 2.6.**

1. $\chi_o(\mathcal{P}_4 \cap bip \cap out(2) \cap \mathcal{D}_2) \geqslant 11$.
2. $\chi_s(\mathcal{P}_4) \leqslant 59$.
3. $\chi_s(\mathcal{P}_4 \cap out(2)) \leqslant 27$.

### 2.5.1 The lower bound

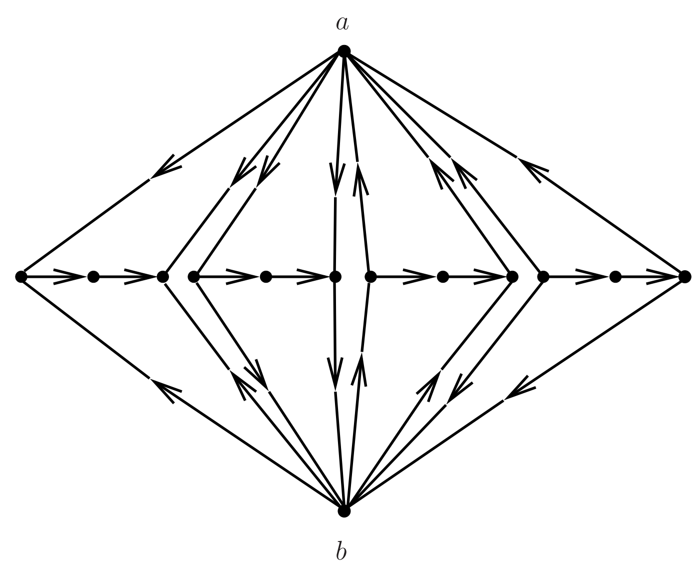


Figure 2.9: How to force a good pair in a target graph.

Let $N^+(x)$ and $N^-(x)$ be respectively the out-neighborhood and in-neighborhood of the vertex $x$. We say that a pair $(x, y)$ of distinct vertices forms a *good pair* if the sets $N^+(x) \cap N^+(y)$, $N^+(x) \cap N^-(y)$, $N^-(x) \cap N^+(y)$ and $N^-(x) \cap N^-(y)$ are all of size at least 2. A triplet $(x, y, z)$ is a *good triplet* if $(x, y)$, $(y, z)$ and $(z, x)$ are all good pairs. Consider the graph $G$ in Figure 2.9. We remark that every two distinct $^{\geqslant}$3-vertices are joined by an arc or a directed 2-path. Therefore every two $^{\geqslant}$3-vertices must be assigned distinct colors in any oriented coloring of $G$. This provides a simple proof that $\chi_o(G) = 10$. It also implies that the colors of $a$ and $b$ form a good pair in any target graph of $G$. We now construct the graph $G^*$ by taking 3 copies $G_1$, $G_2$, $G_3$ of $G$ and identifying $a_1$ and $b_2$, $a_2$ and $b_3$, $a_3$ and $b_1$. Similarly, the colors of $a_1$, $a_2$, $a_3$ form a good triplet in any target graph of $G^*$. A computer check shows that no tournament of order 10 contains a good triplet. We can see that $G^*$ has a homomorphism to $QR_{11}$. First, we give pairwise distinct colors to $a_1$, $a_2$, $a_3$. Notice that $QR_{11}$ satisfies $S_{2,2}$ by Lemma 2.4.(2). Thus, in each copy of $G$, we can give pairwise distinct colors to the eight 3-vertices. Then we can easily color the 2-vertices. So $G^*$ is a 2-degenerate 2-outerplanar bipartite graph such that $\chi_o(G^*) = 11$, which proves Theorem 2.6.(1).

### 2.5.2 Triangle-free planar graphs

We use the method of reducible configurations to show that every triangle-free planar graph is $QR_{59}$-colorable. We define the partial order $\prec$ for the set of all graphs. Let $n_3(G)$ be the number of $\geqslant$3-vertices in $G$. For any two graphs $G_1$ and $G_2$, we have $G_1 \prec G_2$ if and only if at least one of the following conditions hold:

- $G_1$ is a proper subgraph of $G_2$.
- $n_3(G_1) < n_3(G_2)$.

Note that this partial order is well-defined, since if $G_1$ is a proper subgraph of $G_2$, then $n_3(G_1) \leqslant n_3(G_2)$. So $\prec$ is a partial linear extension of the subgraph poset.

**Lemma 2.5.** *Let $QR_q$ be a Paley tournament satisfying $S_{3,n}$ for some $n \geqslant 1$. Let $G$ be a graph having no homomorphism to $QR_q$ which is minimal with this property according to $\prec$.*

1. *The graph $G$ is 2-connected and its cut sets of size two consist of non-adjacent vertices.*
2. *For every 2-cut $\{u, v\}$ of $G$, the graph $G \setminus \{u, v\}$ has exactly two connected components, and one of them is a single vertex.*
3. *The graph $G$ contains no vertex adjacent to at most $(n-1)$ 2-vertices and at most three other vertices.*
4. *The graph $G$ contains no 3-vertex.*

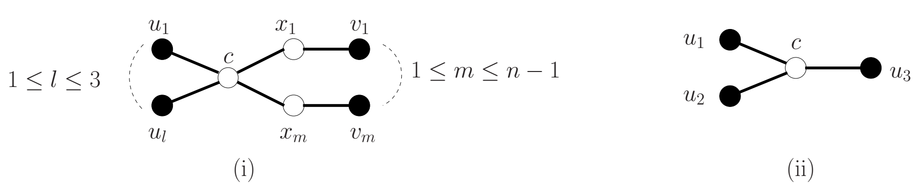


Figure 2.10: Forbidden configurations for Lemma 2.5.

*Proof.*

1. If $G$ is not 2-connected, then we can obtain a $QR_q$-coloring of $G$ from the coloring of its 2-connected components since $QR_q$ is a circular tournament. Moreover $G$ cannot contain a cut set consisting of two adjacent vertices since $QR_q$ is an arc-transitive tournament.
2. Suppose that $G$ contains a 2-cut $\{u, v\}$ contradicting Lemma 2.5. Let $A_1, \dots A_n$ denote the connected components of $G \setminus \{u, v\}$ of size at least two. For $1 \leqslant i \leqslant n$, we construct the graph $B_i$ from the graph induced by $V(A_i) \cup \{u, v\}$ by adding a directed 2-path between $u$ and $v$. Notice that if $G$ is triangle-free and planar (resp. 2-outerplanar), then $A_i$ is triangle-free and planar (resp. 2-outerplanar). Moreover $G' \prec G$ since $n_3(A_i) < n_3(G)$. Since $QR_q$ is arc transitive and self-reverse, $A_i$ has a $QR_q$-coloring $f$ such that $f(u) = 0$ and $f(v) = 1$. These colorings of the $A_i$'s induce a coloring of the $B_i$'s that can be extended to $G$.

3. Consider configuration (i) in Figure 2.10. Let $f$ be any $QR_q$-coloring of $G \setminus \{x_1, \ldots, x_m\}$. By property $S_{3,n}$, we can choose $f$ such that $f(c) \notin \{f(v_1), \ldots, f(v_m)\}$ and extend this coloring to $G$.

4. Consider configuration (ii) in Figure 2.10. Notice that $u_1$, $u_2$, and $u_3$ are ${}^{\geqslant}3$-vertices since configuration (i) with $l = 2$, $m = 1$ is forbidden. Since $QR_q$ is self-reverse, we assume w.l.o.g. that $d^-(c) \leqslant d^+(c)$ by considering either $G$ or $G^R$. We have $d^-(c) \neq 0$ since otherwise we can extend any $QR_q$-coloring of $G \setminus \{c\}$ to $G$. Suppose now $d^-(c) = 1$, which is the only remaining case. Let us set $N^-(c) = \{u_1\}$, $N^+(c) = \{u_2, u_3\}$. We now consider the graph $G'$ obtained from $G \setminus \{c\}$ by adding directed 2-paths joining respectively $u_1$ and $u_2$, and $u_1$ and $u_3$. Notice that if $G$ is triangle-free and planar (resp. 2-outerplanar), then $G'$ is triangle-free and planar (resp. 2-outerplanar). Moreover $G' \prec G$ since $n_3(G') = n_3(G) - 1$. Any $QR_q$-coloring $f$ of $G'$ induces a coloring of $G \setminus \{c\}$ such that $f(u_1) \neq f(u_2)$ and $f(u_1) \neq f(u_3)$, which can be extended to $G$.

□

Euler's formula $|V(G)| + |F(G)| = |E(G)| + 2$ and

$$\sum_{v \in V(G)} d(v) = \sum_{f \in F(G)} d(f) = 2|E(G)|$$

show that

$$\sum_{v \in V(G)} (d(v) - 4) + \sum_{f \in F(G)} (d(f) - 4) = -8.$$

We set an *initial charge* $ch$ to every vertex and every face:

$$\forall x \in V(G) \cup F(G),\ ch(x) = d(x) - 4$$

Then we use a discharging procedure consisting of the following two rules, and we get a *final charge* $ch^*$.

**Rule 1.** Every ${}^{\geqslant}4$-vertex $v$ gives $\frac{1}{2}$ to each face $f$ incident to both $v$ and a 2-neighbor of $v$.
**Rule 2.** Every face $f$ gives 1 to each 2-vertex incident to $f$.

Since the above procedure preserves the total charge, we have:

$$\sum_{x \in V(G) \cup F(G)} ch(x) = \sum_{x \in V(G) \cup F(G)} ch^*(x) = -8.$$

We now prove the following to get a contradiction:

$$\forall x \in V(G) \cup F(G),\ ch^*(x) \geqslant 0.$$

**case** $x \in V(G)$

$d(x) = 2$: By **Rule 2**, $x$ receives exactly 1 from each of the two faces incident to $x$ and thus $ch^*(x) = -2 + 2 \times 1 = 0$.

$d(x) = 3$: $G$ contains no 3-vertex by Lemma 2.5.(4).

$d(x) = k,\ 4 \leqslant k \leqslant 7$: By Lemma 2.5.(3) with $l = 3,\ m = k-3$, $x$ has at most $(k-4)$ 2-neighbors, so $x$ gives $\frac{1}{2}$ to at most $2\times(k-4)$ faces and thus $ch^*(x) \geqslant k-4-2\times(k-4)\times\frac{1}{2} = 0$.

$d(x) = k \geqslant 8$: $x$ gives $\frac{1}{2}$ to at most $k$ faces and thus $ch^*(x) \geqslant k-4-k\times\frac{1}{2} \geqslant 0$.

**case** $x \in F(G)$

$d(x) = 3$: $G$ is triangle-free, so it contains no face of degree 3.

$d(x) = 4$: By Lemma 2.5.(3) with $l = m = 1$, $x$ is not incident two adjacent 2-vertices. The face $x$ cannot be incident to two non-adjacent 2-vertices, since otherwise the $^{>}2$-vertices incident to $x$ would create a 2-cut that would contradict Lemma 2.5.(2). Thus $x$ is incident to at most one 2-vertex. If $x$ is incident to a 2-vertex then $ch^*(x) = 0+2\times\frac{1}{2}-1 = 0$, otherwise $ch^*(x) = ch(x) = 0$.

$d(x) = k \geqslant 5$: Let $n$ be the number of 2-vertices incident to $x$. Since $x$ is not incident two adjacent 2-vertices, we have $n \leqslant \lfloor\frac{k}{2}\rfloor$ and $x$ receives $\frac{1}{2}$ from at least $n$ vertices, thus $ch^*(x) \geqslant k-4+n\times\frac{1}{2}-n\times 1 = k-4-\frac{n}{2} \geqslant k-4-\lfloor\frac{k}{4}\rfloor = \lceil\frac{3k}{4}\rceil - 4 \geqslant 0$.

This proves Theorem 2.6.(3).

### 2.5.3 Triangle-free 2-outerplanar graphs

**Lemma 2.6.** *Let $G$ be a connected triangle-free outerplanar graph on at least three vertices. Then $G$ contains at least one of the following:*

1. *A $^{\geqslant}2$-vertex adjacent to at least $(d(u)-1)$ 1-vertices.*
2. *A $^{\geqslant}3$-vertex contained in a cycle adjacent to $(d(u)-2)$ 1-vertices.*
3. *Two adjacent 2-vertices which are contained in a cycle.*

*Proof.* Suppose $G$ does not contain any of the two first items. Since $G$ is outerplanar, its 2-connected components form a tree-like structure. A 2-connected triangle-free outerplanar contains at least two times two adjacent 2-vertices. Thus a 2-connected component of $G$ which is a leaf in the tree-like structure contains at least one time two adjacent 2-vertices, which is the third item. □

We consider a potential counter-example $G$ to Theorem 2.6.(3) which is minimal according to the partial order $\prec$. Let us fix a 2-outerplanar embedding of $G$. Lemma 2.5.(1) implies that the outerface induces a chordless cycle. Let $C$ denote the outerface of $G$. The graph $H = G \setminus C$ is thus an outerplanar triangle-free graph. Moreover, $H$ must be connected, otherwise $G$ would contain a 2-cut in $C$ that contradicts Lemma 2.5.(2). In what follows we consider $C$ and $H$ together with their embeddings, as implied by the embedding of $G$.

A vertex $v \in H$ is *very special* if and only if $d_H(v) = 1$. A vertex $v \in G$ is *special* if and only if either $v$ is very special or $v \in C$.

**Lemma 2.7.** *Every very special vertex $v$ satisfies $d_G(v) = 2$.*

*Proof.* By Lemma 2.5.(4), $d_G(v) \neq 3$. Now, if $d_G(v) \geqslant 4$, then $d_C(v) \geqslant 3$ thus there exists a vertex $x \in N_C(v)$ such that $d_G(x) = d_C(x) + d_H(x) = 2 + 1 = 3$, which contradicts Lemma 2.5.(4). □

**Lemma 2.8.** *A vertex $v \in H$ has no three consecutive special neighbors.*

*Proof.* We suppose that $H$ contains such a vertex $v$ and we note $v_1, v_2, v_3$ its three consecutive special neighbors. We consider the vertices $w_1, w_2, w_3$ defined as follows: For $i = 1, 2, 3$, $w_i = v_i$ if $v_i \in C$ and $w_i$ is the neighbor of $v_i$ which belongs to $H$ if is $v_i$ very special. Since $G$ does not contain triangles nor configuration (iii), we have $w_1 \neq w_2 \neq w_3 \neq w_1$. This implies that $d_G(w_2) = 3$, which contradicts Lemma 2.5.(4). □

Let $u$ be a vertex in $H$ and $N^1_H(u) = \{x \in N_H(u) \mid d_H(x) = 1\}$ (i.e. $N^1_H(u)$ is the set of very special neighbors of $u$). We consider now four cases:

1. $H$ is an isolated vertex. In this case we have clearly a 3-vertex $v$ in $G$, which is a contradiction.

2. $H$ contains a vertex $u$ such that $|N^1_H(u)| \geqslant d_H(u) - 1 \geqslant 1$. Let $v$ be a vertex in $N^1_H(u)$. Since $v$ is a very special vertex then according to Lemma 2.7, $d_G(v) = 2$. Now, if $d_G(u) \leqslant 4$, then we have a 2-vertex $v$ adjacent to a $^{\leqslant}4$-vertex, which contradicts Lemma 2.5.(3). If $d_G(u) \geqslant 5$, then $u$ has at least three consecutive special neighbors, which contradicts Lemma 2.8.

3. $H$ contains a vertex $u$ such that $|N^1_H(u)| = d_H(u) - 2 \geqslant 1$ and $u$ is contained in a cycle in $H$. Let $v$ be a vertex in $N^1_H(u)$. Since $v$ is a very special vertex in $H$ then $d_G(v) = 2$ according to Lemma 2.7. Now, if $d_G(u) \leqslant 4$, then we have a 2-vertex $v$ adjacent to a $^{\leqslant}4$-vertex, which contradicts Lemma 2.5.(3). If $d_G(u) \geqslant 5$, then $u$ has at least three consecutive special neighbors, which contradicts Lemma 2.8.

4. $H$ contains two adjacent 2-vertices $u$ and $v$ which are contained in a cycle in $H$. We assume w.l.o.g. that $d_G(u) \leqslant d_G(v)$ and consider the following subcases:

   (a) If $d_G(v) \geqslant 5$, then $v$ has at least three consecutive special neighbors and this contradicts Lemma 2.8. So, we have $d_H(u) \leqslant d_H(v) \leqslant 4$.

   (b) If $d_G(u) = 2$, we have a 2-vertex adjacent to a $^{\leqslant}4$-vertex, which is forbidden by Lemma 2.5.(3).

   (c) The case $d_G(u) = 3$ is forbidden by Lemma 2.5.(4).

   (d) If $d_G(u) = d_G(v) = 4$, there exists a vertex $x \in N_C(u)$ such that $d_G(x) = 3$.

This proves Theorem 2.6.(3).

Concerning 2-outerplanar graphs in general, we know from Theorem 2.5 that $out(2) \subset \mathcal{S}_0 \odot \mathcal{S}_0 \odot \mathcal{S}_0 \odot out(1)$. By Proposition 1.3, the oriented chromatic number of a 2-outerplanar graph is thus at most $2^{4-1} \times (1 + 1 + 1 + 7) = 80$. This is not an improvement over the bound $\chi_o(out(2)) \leqslant \chi_o(\mathcal{P}_3) \leqslant 80$, but it provides another target graph of size 80 for 2-outerplanar graphs.

## 2.6 Acyclic choosability

In this section, we study the list version of (proper) acyclic coloring on graphs with bounded maximum average degree.

**Theorem 2.7.**

1. *Every graph $G$ with $mad(G) < \frac{8}{3}$ is acyclically 3-choosable.*
2. *Every graph $G$ with $mad(G) < \frac{19}{6}$ is acyclically 4-choosable.*
3. *Every graph $G$ with $mad(G) < \frac{24}{7}$ is acyclically 5-choosable.*

For planar graphs, thanks to Observation 1.1, we get:

**Corollary 2.1.**

1. $\chi_a^l(\mathcal{P}_8) = 3$.
2. $\chi_a^l(\mathcal{P}_6) \leqslant 4$.
3. $\chi_a^l(\mathcal{P}_5) \leqslant 5$.

A $d(k)$-vertex is a $d$-vertex adjacent to at least $k$ 2-vertices. Every item of Theorem 2.6 is proved using the method described in Section 1.6. The partial order on graphs considered here is the subgraph partial order. Let $H$ be graph that is not acyclically $n$-choosable and is minimal with this property. To prove that a configuration $C$ is forbidden, we suppose that $H$ contains $C$ and show that an acyclic coloring $c$ chosen from $L$ of some proper subgraph of $H$ can be extended in a acyclic coloring chosen from $L$ of the whole graph $H$, for every $L$.

### 2.6.1 Proof of Theorem 2.7.(1)

We prove now that every graph $G$ with $mad(G) < 8/3$ is acyclically 3-choosable.

**Lemma 2.9.** *Let $n \geqslant 3$ and let $H$ be a minimal graph such that $\chi_a^l(H) > n$. Then $H$ does not contain*

1. *a $d$-vertex adjacent to the vertices a $d$-clique* $(0 \leqslant d \leqslant n-1)$,
2. *a $d(d)$-vertex* $(2 \leqslant d \leqslant n^2 - 1)$,
3. *a $d(d-1)$-vertex* $(2 \leqslant d \leqslant (n-1)^2)$,
4. *a $d(2)$-vertex* $(2 \leqslant d \leqslant n)$,
5. *a $d(1)$-vertex* $(2 \leqslant d \leqslant n-1)$.

*Proof.*

1. Suppose that $H$ contains such a vertex $w$. Any coloring $c$ of $H \setminus \{w\}$ can be extended to $H$ by giving $w$ a color distinct from those of its neighbors.

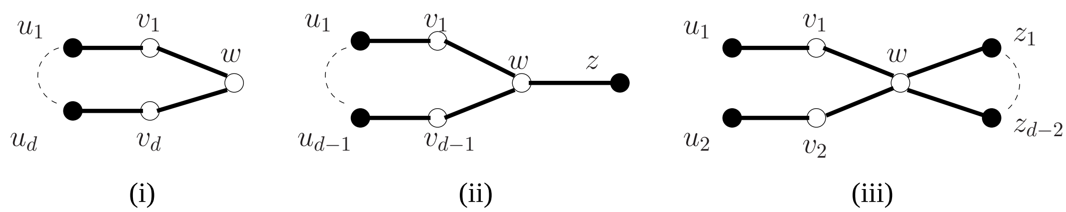


Figure 2.11: $(i)$: A $d(d)$-vertex. $(ii)$: A $d(d-1)$-vertex. $(iii)$: A $d(2)$-vertex.

2. Suppose that $H$ contains a $d(d)$-vertex $w$ adjacent to $d$ 2-vertices $v_1, \dots, v_d$. Each vertex $v_i$ is adjacent to $w$ and to another vertex $u_i$, $1 \leqslant i \leqslant d$ (see Figure 2.11(i)). Let $c$ be a coloring of $H \setminus \{w, v_1, \dots, v_d\}$. Since $d \leqslant n^2 - 1$ and $|L(w)| = n$, the pigeonhole principle ensures that some $j \in L(w)$ is used at most $n-1$ times to color the $u_i$'s. We set $c(w) = j$. If $c(u_i) \neq j$, we can choose $c(v_i)$ in $L(v_i) \setminus \{c(u_i), j\}$ since $|L(v_i)| = n \geqslant 3$. The number of $v_i$ such that $c(u_i) = j$ is at most $n-1$, so we can give these $v_i$ distinct colors different from $j$.

3. Suppose that $H$ contains a $d(d-1)$-vertex $w$ adjacent to $(d-1)$ 2-vertices $v_1, \dots, v_{d-1}$ and to another vertex $z$. Each vertex $v_i$ is adjacent to $w$ and to another vertex $u_i$, $1 \leqslant i \leqslant d-1$ (see Figure 2.11(ii)). Let $c$ be a coloring of $H \setminus \{w, v_1, \dots, v_{d-1}\}$. Note that we have $|L(w) \setminus \{c(z)\}| \geqslant n-1$ and $d - 1 \leqslant (n-1)^2 - 1$. We set $c(w) = j$ where $j \in L(w) \setminus \{c(z)\}$ is used at most $n-2$ times to color the $u_i$'s. If $c(u_i) \neq j$, we can choose $c(v_i)$ in $L(v_i) \setminus \{c(u_i), j\}$ since $L(v_i) = n \geqslant 3$. The number of $v_i$ such that $c(u_i) = j$ is at most $n-2$, so we can give these $v_i$ distinct colors different from $j$ and $c(z)$.

4. Suppose that $H$ contains a $d(2)$-vertex $w$ adjacent to $z_1, \dots, z_{d-2}$, and to two 2-vertices $v_1, v_2$ that are adjacent respectively to $u_1, u_2$ (see Figure 2.11(iii)). We assume $n \geqslant 4$ since the case $n = 3$ is implied by Lemma 2.9.(3). Let $c$ be a coloring of $H \setminus \{w, v_1, v_2\}$.

   4.1 If the $c(z_i)$ are pairwise distinct, we choose $c(w) \in L(w) \setminus \{c(z_1), \dots, c(z_{d-2}), c(u_1)\}$ and $c(v_1) \in L(v_1) \setminus \{c(w), c(u_1)\}$. If $c(w) = c(u_2)$, we choose $c(v_2) \in L(v_2) \setminus \{c(z_1), \dots, c(z_{d-2}), c(w)\}$; otherwise we choose $c(v_2) \in L(v_2) \setminus \{c(w), c(u_2)\}$.

   4.2 If the $c(z_i)$ are not pairwise distinct, we consider a coloring $c$ of $H \setminus \{v_1, v_2\}$ and assume w.l.o.g. that $c(z_1) = c(z_2)$. If $c(w) = c(u_1)$, we choose $c(v_1) \in L(v_1) \setminus \{c(z_2), \dots, c(z_{d-2}), c(w)\}$, otherwise we choose $c(v_1) \in L(v_1) \setminus \{c(u_1), c(w)\}$. If $c(w) = c(u_2)$, we choose $c(v_2) \in L(v_2) \setminus \{c(z_2), \dots, c(z_{d-2}), c(v_1), c(w)\}$, otherwise we choose $c(v_2) \in L(v_2) \setminus \{c(u_2), c(w)\}$.

5. The proof is similar (and simpler) to that of Lemma 2.9.(4).

□

It follows that the minimum degree of $H$ is at least 2 and that no 2-vertex is in a triangle.

We use the following discharging rule: Each vertex gives $\frac{1}{3}$ to each of its 2-neighbors. Let us check that for every $v \in V(H)$, $d^*(v) \geqslant \frac{8}{3}$:

- If $d(v) = 2$, then $d^*(v) = 2 + 2\frac{1}{3} = \frac{8}{3}$, since $v$ has no 2-neighbor by Lemma 2.9.(3) and $v$ receives $\frac{1}{3}$ from each neighbor.
- If $d(v) = 3$, then $d^*(v) \geqslant 3 - \frac{1}{3} = \frac{8}{3}$, since $v$ has at most one 2-neighbor by Lemma 2.9.(3), so it gives at most $\frac{1}{3}$.
- If $d(v) = k \geqslant 4$, then $d^*(v) \geqslant k - k\frac{1}{3} = \frac{2k}{3} \geqslant \frac{8}{3}$ because $v$ gives at most $k$ times $\frac{1}{3}$.

### 2.6.2 Proof of Theorem 2.7.(2)

We prove now that every graph $G$ with $mad(G) < 19/6$ is acyclically 4-choosable.

**Lemma 2.10.** *Let $n \geqslant 4$ and let $H$ be a minimal graph such that $\chi_a^l(H) > n$. Then $H$ does not contain*

1. *a 5(3)-vertex adjacent to a 3-vertex,*
2. *a 3-vertex adjacent to two 3-vertices.*

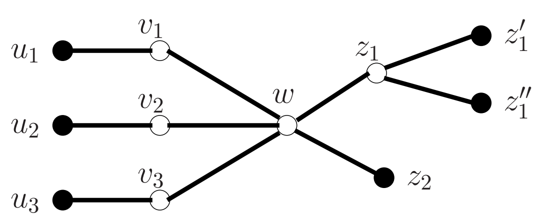


Figure 2.12: A 5(3)-vertex adjacent to a 3-vertex.

*Proof.* 1. Suppose that $H$ contains a 5(3)-vertex $w$ adjacent to three 2-vertices $v_1, v_2, v_3$ (each adjacent to another vertex $u_i$), a 3-vertex $z_1$ (adjacent to $z_1'$ and $z_1''$) and another vertex $z_2$ (see Figure 2.12). Let $c$ be a coloring of $H \setminus \{v_1\}$. If $c(u_1) \neq c(w)$, we give a proper color to $v_1$. Now, we assume that $c(u_1) = c(w) = 1$:

1.1 If $c(z_1) \neq c(z_2)$, we erase the colors of $v_2, v_3$ and we modify the color of $w$: In $L(w) \setminus \{c(z_1), c(z_2)\}$, there is a color which appears on at most one of $u_1, u_2, u_3$; we choose this color for $w$. Then, we give a color different from $c(z_1), c(z_2), c(w)$ to the vertex $v_j$ (if it exists) whose neighbors have the same color ($c(w)$) and we give a proper color to the other $v_i$.

1.2 If $c(z_1) = c(z_2)$ and w.l.o.g., $c(z_1) = 2$. Observe that $L(v_1)$ contains 1 and 2; otherwise, we can color $v_1$ with a color different from 1,2 and $c(v_2), c(v_3)$. We assume w.l.o.g. that $L(v_1) = \{1, 2, 3, 4\}$. If we cannot color $v_1$ this implies that $c(u_1) = c(u_2) = c(u_3) = 1$, $c(v_2) = 3$, $c(v_3) = 4$ and $c(z_1) = 2$.

1.2.1 If $c(z_1') \neq c(z_1'')$, we modify the colors of $z_1$, $w$ and give a proper color to $v_1, v_2, v_3$: $c(z_1) \in L(z_1) \setminus \{c(z_1'), c(z_1''), 2\}$, $c(w) \in L(w) \setminus \{c(z_1), c(z_2), 1\}$.

1.2.2 If $c(z_1') = c(z_1'')$, we modify the color of $w$ with a color different from 1, 2, $c(z_1')$ and give proper colors to $v_1, v_2, v_3$.

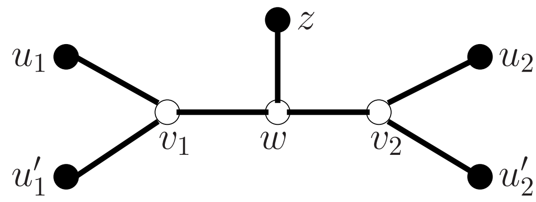


Figure 2.13: A 3-vertex having two 3-neighbors.

2. First suppose that $H$ contains a 3-vertex adjacent to two adjacent 3-vertices (see Figure 2.13, left). Let $c$ be a coloring of $H \setminus \{v_1, v_2, v_3\}$. We can choose $c(v_1)$ in $L(v_1) \setminus \{c(u_1), c(u_2), c(u_3)\}$, $c(v_2)$ in $L(v_2) \setminus \{c(v_1), c(u_2), c(u_3)\}$, and then $c(v_3)$ in $L(v_3) \setminus \{c(v_1), c(v_2), c(u_3)\}$. Now suppose that $H$ contains a 3-vertex $w$ adjacent to two 3-vertices $v_1, v_2$ (each adjacent to $u_1, u'_1$ and $u_2, u'_2$) and to another vertex $z$ (see Figure 2.13, right). Let $c$ be a coloring of $H \setminus \{w\}$. We have to consider the following cases:

   2.1 $c(v_1)$, $c(v_2)$ and $c(z)$ are pairwise distinct. We color $w$ with a proper color.

   2.2 $c(v_1) = c(v_2) \neq c(z)$. W.l.o.g., suppose that $c(v_1) = c(v_2) = 2$ and $c(z) = 1$. Observe that $L(w)$ contains 1 and 2; otherwise, we color $w$ with a color different from 1 or 2 and different from $c(u_1), c(u'_1)$. Assume that $L(w) = \{1, 2, 3, 4\}$. If we cannot color $w$, this implies that $\{c(u_1), c(u'_1)\} = \{c(u_2), c(u'_2)\} = \{3, 4\}$. As well, observe that $L(v_1) = L(v_2) = \{1, 2, 3, 4\}$; otherwise, we modify the color of $v_1$ (or $v_2$) with a color different from 1,2,3,4 to get case 2.1. Hence, we recolor $v_1$ and $v_2$ with 1 and color $w$ with 2.

   2.3 $c(v_1) = c(z) \neq c(v_2)$. W.l.o.g., suppose that $c(v_1) = c(z) = 1$ and $c(v_2) = 2$. With the same argument as above, we can assume that $L(w) = \{1, 2, 3, 4\}$ and $L(v_1) = \{1, 2, 3, 4\}$ (for the same reasons). We recolor $v_1$ with 2 to get case 2.2.

   2.4 $c(v_1) = c(v_2) = c(z)$. Observe that $c(u_1) = c(u'_1)$; otherwise, we modify the color of $v_1$ to get a previous case. We have $c(u_2) = c(u'_2)$ for the same reason and we can choose $c(w) \in L(w) \setminus \{c(u_1), c(u_2), c(z)\}$.

□

We use the following discharging rule: Each $\geqslant$4-vertex gives $\frac{7}{12}$ to each of its 2-neighbors and $\frac{1}{12}$ to each of its 3-neighbors. Let us check that for every $v \in V(H)$, $d^*(v) \geqslant \frac{19}{6}$:

- If $d(v) = 2$, then $v$ has two $\geqslant$4-neighbors by Lemma 2.9.(5), so $d^*(v) = 2 + 2\frac{7}{12} = \frac{19}{6}$.

- If $d(v) = 3$, then $v$ has at least two $\geqslant$4-neighbors by Lemma 2.9.(5) and Lemma 2.10.(2), so $d^*(v) \geqslant 3 + 2\frac{1}{12} = \frac{19}{6}$.

- If $d(v) = 4$, then $v$ has at most one 2-neighbor by Lemma 2.9.(4), so $d^*(v) \geqslant 4 - \frac{7}{12} - 3\frac{1}{12} = \frac{19}{6}$.

- If $d(v) = 5$, then $v$ has at most three 2-neighbors by Lemma 2.9.(3). If $v$ is a 5(3)-vertex, then it has no 3-neighbor by Lemma 2.10.(1), so $d^*(v) = 5 - 3\frac{7}{12} = \frac{13}{4} > \frac{19}{6}$. Otherwise, $d^*(v) \geqslant 5 - 2\frac{7}{12} - 3\frac{1}{12} = \frac{43}{12} > \frac{19}{6}$.

- If $d(v) = k, 6 \leqslant k \leqslant 7$, then $v$ has at most $(k-2)$ 2-neighbors by Lemma 2.9.(3), so $d^*(v) \geqslant k - (k-2)\frac{7}{12} - 2\frac{1}{12} = \frac{5k}{12} + 1 \geqslant \frac{7}{2} > \frac{19}{6}$.
- If $d(v) = k \geqslant 8$, then $d^*(v) \geqslant k - k\frac{7}{12} = \frac{5k}{12} \geqslant \frac{10}{3} > \frac{19}{6}$.

### 2.6.3 Proof of Theorem 2.7.(3)

We prove now that every graph $G$ with $mad(G) < 19/6$ is acyclically 5-choosable. A vertex is said *weak* if it is either a 3-vertex or a 6(4)-vertex.

**Lemma 2.11.** *Let $n \geqslant 5$ and let $H$ be a minimal graph such that $\chi_a^l(H) > n$. Then $H$ does not contain*

1. *a $d(d-2)$-vertex adjacent to a weak vertex, $3 \leqslant d \leqslant 10$,*
2. *a 6(3)-vertex adjacent to three weak vertices,*
3. *a 6(4)-vertex adjacent to a $^{\leqslant}4$-vertex,*
4. *a 4-vertex adjacent to three 3-vertices.*

*Proof.* 1. Suppose that $H$ contains a $d(d-2)$-vertex $w$ adjacent to $(d-2)$ 2-vertices $v_i$, $1 \leqslant i \leqslant d-2$ (each adjacent to another vertex $u_i$), a 3-vertex $z$ (adjacent to two other vertices $z_1, z_2$) and a vertex $y$, $3 \leqslant d \leqslant 10$ (see Figure 2.14).

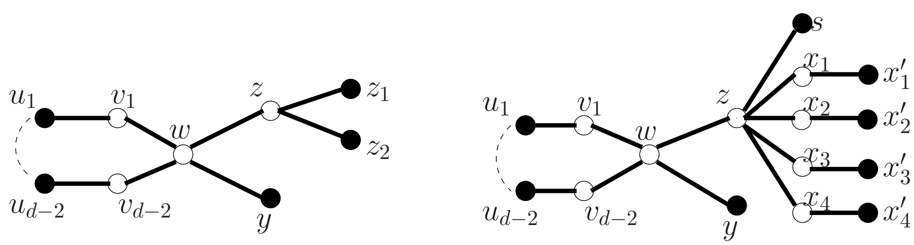


Figure 2.14: A $d(d-2)$-vertex adjacent to a weak vertex.

Let $c$ be a coloring of $H \setminus \{v_i, 1 \leqslant i \leqslant d-2\}$.

- $c(z) \neq c(y)$. We recolor $w$ with a color, different from $c(z), c(y)$, which appears on at most two of the $u_i$'s, $1 \leqslant i \leqslant d-2$. If $c(u_i) \neq c(w)$, we color $v_i$ with a proper color. At most two of the $u_i$'s (say $u_1, u_2$) satisfy $c(u_i) = c(w)$. We can choose $c(v_1) \in L(v_1) \setminus \{c(w), c(y), c(z)\}$ and $c(v_2) \in L(v_2) \setminus \{c(w), c(y), c(z), c(v_1)\}$.
- $c(z) = c(y)$. Observe that $c(z_1) = c(z_2)$; otherwise, we replace the color of $z$ with a color different from $c(z_1), c(z_2), c(y), c(w)$ and we are in the previous case. Now, we recolor $w$ with a color, different from $c(z), c(z_1)$, which appears on at most wo of the $u_i$'s, $1 \leqslant i \leqslant d-2$. As above, it is easy then to color $v_i$, $1 \leqslant i \leqslant d-2$.

Now, we consider the case where the $d(d-2)$-vertex $w$ is adjacent to a 6(4)-vertex $z$ adjacent to four 2-vertices $x_j$, $1 \leqslant j \leqslant 4$ and another vertex $s$ (see Figure 2.14). Observe that $x_i \neq u_j$ for all $i, j$ since there is no 2(1)-vertex by Lemma 2.9.(3). Let $c$ be a coloring of $H \setminus \{x_1\}$.

- $c(w) \neq c(s)$. We erase the colors of the vertices $z, x_2, x_3, x_4$. We recolor $z$ with a color, different from $c(s), c(w)$, which appears on at most one of $x'_i, 1 \leqslant i \leqslant 4$. Then, we give a proper color to $x_i$ for each index $i$ such that $c(x'_i) \neq c(z)$ and give a color different from $c(z), c(w), c(s)$ to the vertex $x_j$ such that $c(z) = c(x'_j)$.
- $c(w) = c(s)$. If $c(x'_1) \neq c(z)$, we color $x$ properly, which suffices. If $c(z) \neq c(x'_i)$ for some $i$, we color $x_1$ avoiding $c(w), c(z)$, and all $c(x_j)$ for $j \neq i$, $j > 1$, which suffices.

Thus we may assume that $c(x'_1) = c(x'_2) = c(x'_3) = c(x'_4) = c(z) = 1$ and $c(s) = c(w) = 2$. Now, we erase the colors of the vertices $x_i$ $(1 \leqslant i \leqslant 4)$, $v_j$ $(1 \leqslant j \leqslant d-2)$, $w$ and $z$. We recolor $w$ with a color different from $c(y)$ and 2, which appears on at most 2 of the $u_j$'s. So, $c(s) \neq c(w)$ and we recolor $z$ with a color different from 1,2, $c(w), c(y)$, then we color each $x_i$ with a proper color. Finally, we recolor the $v_i$'s as in the case $c(z) \neq c(y)$.

2. Suppose that $H$ contains a 6(3)-vertex $w$ adjacent to three 2-vertices $v_1, v_2, v_3$ (each adjacent to another vertex $u_i$) and three weak vertices $z_1, z_2, z_3$. Let $c$ be a coloring of $H \setminus \{v_1, v_2, v_3\}$.

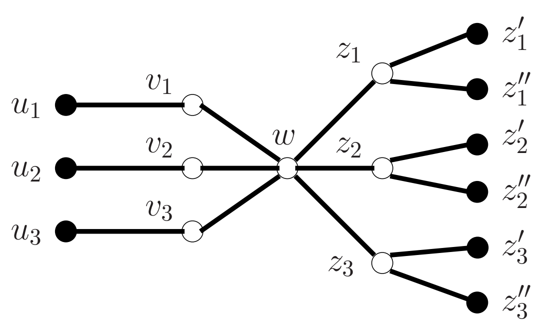


Figure 2.15: A 6(3)-vertex $w$ adjacent to three 3-vertices.

First, observe that if $c(z_1), c(z_2), c(z_3)$ are all different, we can color $v_1, v_2, v_3$: We recolor $w$ with a color different from $c(z_1), c(z_2), c(z_3)$, which appears on at most one of $u_1, u_2, u_3$. Then, we give proper color to $v_i$ for each index $i$ for which $c(u_i) \neq c(w)$ and a color different from $c(w), c(z_1), c(z_2), c(z_3)$ otherwise.

Second, observe that if $c(z_1) = c(z_2) = c(z_3)$, we can color $v_1, v_2, v_3$: If $c(u_i) \neq c(w)$, we give a proper color to $v_i$. In the worst case, we have $c(u_1) = c(u_2) = c(u_3) = c(w)$ and we color $v_1$ with $c(v_1) \in L(v_1) \setminus \{c(w), c(z_1)\}$, $v_2$ with $c(v_2) \in L(v_2) \setminus \{c(w), c(z_1), c(v_1)\}$ and $v_3$ with $c(v_3) \in L(v_3) \setminus \{c(w), c(z_1), c(v_1), c(v_2)\}$.

Consider now the case where two of $z_1, z_2, z_3$ have the same color. W.l.o.g., we assume that $c(w) = 1, c(z_1) = c(z_2) = 2, c(z_3) = 3$.

Third, observe that if $c(u_1) \neq 1$, we can color $v_1, v_2, v_3$: We color $v_1$ and $v_2$ such that $c(v_1) \in L(v_1) \setminus \{1, c(u_1)\}$ and $c(v_2) \in L(v_2) \setminus \{1, 2, 3, c(u_2)\}$. Then if $c(u_3) = 1$, we choose $c(v_3)$ in $L(v_3) \setminus \{1, 2, 3, c(v_2)\}$ and otherwise we choose $c(v_3)$ in $L(v_3) \setminus \{1, c(u_3)\}$.

So, suppose now that $c(u_1) = c(u_2) = c(u_3) = c(w) = 1, c(z_1) = c(z_2) = 2, c(z_3) = 3$. The idea is to consider the neighborhood of the two vertices of $z_1, z_2, z_3$ which have the same color ($z_1, z_2$ in our case) and modify if necessary the color of one of these two vertices to get a previous case.

By permuting indices, we have only two cases to study:

2.1 $z_1$ is a 6(4)-vertex. The 6(4)-vertex $z_1$ is adjacent to $w$, to four 2-vertices $x_i$ (each adjacent to another vertex $x'_i$) and another vertex $s$. Observe that since there is no 2(1)-vertex by Lemma 2.9.(3), $x_i \neq u_j$ for all $i, j$. We erase the colors of $w, z_1, x_1, x_2, x_3, x_4$. We recolor $z_1$ with a color, different from 2, 3, $c(s)$, which appears on at most two of $x'_i$, $1 \leqslant i \leqslant 4$. We recolor now $w$ with a color different from 1, 2, 3, $c(z_1)$ and give proper colors to $v_1, v_2, v_3$. Finally, we color the $x_i$, $1 \leqslant i \leqslant 4$: For the two or fewer vertices whose neighbors have the same color, we give distinct colors different from $c(s), c(w), c(z_1)$ and give proper colors to the other vertices $x_i$.

2.2 $z_1$ and $z_2$ are 3-vertices. The vertex $z_1$ is adjacent to $w$ and two other vertices $z'_1, z''_1$ and the vertex $z_2$ is adjacent to $w$ and two other vertices $z'_2, z''_2$ (see Figure 2.15). It may be that $z_i, z'_j, z''_k$ are not distinct, but it will not matter. If $c(z'_1) \neq c(z''_1)$ we can recolor $z_1$ and $w$ such that $c(z_1) \in L(z_1) \setminus \{2, 3, c(z'_1), c(z''_1)\}$ and $c(w) \in L(w) \setminus \{1, 2, 3, c(z_1)\}$, and then give proper colors to the $v_i$'s, $1 \leqslant i \leqslant 3$. Thus $c(z'_1) = c(z''_1)$ and, for the same reason, $c(z'_2) = c(z''_2)$. Now we can recolor $w$ with a color different from 1, 2, 3, $c(z'_1)$ and we give proper colors to the $v_i$'s, $1 \leqslant i \leqslant 3$.

3. Suppose that $H$ contains a 6-vertex $w$ adjacent to four 2-vertices $v_1, v_2, v_3, v_4$ (each adjacent to another vertex $u_i$), a ${}^{\leqslant}4$-vertex $z$ and another vertex $y$ (see Figure 2.16). Notice that if $d(z) < 4$ then the configuration is forbidden by Lemma 2.10.(1) and Lemma 2.9.(3). So suppose $z$ is a 4-vertex adjacent to $z_1, z_2, z_3$ (see Figure 2.16).

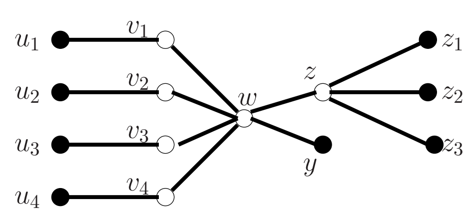


Figure 2.16: A 6(4)-vertex adjacent to a 4-vertex

Let $c$ be a coloring of $H \setminus \{v_1, v_2, v_3, v_4\}$. If $c(y) \neq c(z)$, we recolor $w$ with a color from $L(w) \setminus \{c(z), c(y)\}$ that appears on at most one $u_i$, then properly color each $v_i$ avoiding $c(u_i), c(w), c(z)$, and $c(y)$. Suppose that $c(y) = c(z)$. If $c(u_i) \neq c(w)$, we properly color $v_i$ and then may ignore it, so the worst case is $c(u_1) = c(u_2) = c(u_3) = c(u_4) = c(w)$. Assume that $c(u_1) = 1$ and $c(z) = 2$. Consider the following three cases:

3.1 If $c(z_1) \neq c(z_2) \neq c(z_3) \neq c(z_1)$, we modify the color of $z$, then we recolor $w$ with a color different from 1, $c(z), c(y)$, then we color $v_i$ $(i = 1, \ldots, 4)$ with proper colors.

3.2 If $c(z_1) = c(z_2) \neq c(z_3)$, we recolor $w$ such that $c(w) \in L(w) \setminus \{1, 2, c(z_1), c(z_3)\}$ and give proper colors to $v_i$.

3.3 If $c(z_1) = c(z_2) = c(z_3)$. We modify the color of $w$. We color $w$ with $c(w) \in L(w) \setminus \{1, 2, c(z_1)\}$ and give proper colors to $v_i$.

4. Suppose that $H$ contains a 4-vertex $w$ adjacent to three 3-vertices $x_1, x_2, x_3$ (each adjacent to $x_i', x_i''$) and to another vertex $z$ (see Figure 2.17). Although $x_i, x_j', x_k''$ may not all be distinct, it will not matter.

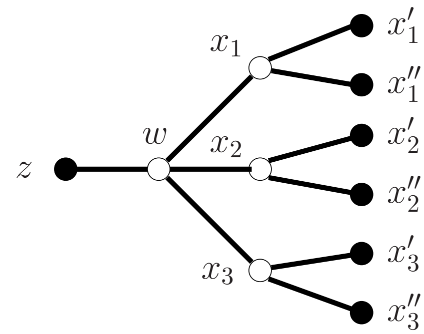


Figure 2.17: A 4-vertex adjacent to three 3-vertices

Let $c$ be a coloring of $H \setminus \{w\}$. We consider the following cases:

4.1 If $c(x_1), c(x_2), c(x_3), c(z)$ are all different, then we color $w$ with a proper color.

4.2 Suppose that two neighbors of $w$ have the same color, and no color is shared by three neighbors of $w$.

4.2.1 Suppose that $c(x_1) = c(x_2) \neq c(x_3)$. W.l.o.g. we assume that $c(x_1) = 1$.

4.2.1.1 If $c(x_3) \neq c(z)$ and $c(x_3) \neq 1, c(z) \neq 1$, we assume that $c(x_3) = 2$ and $c(z) = 3$. Necessarily, $L(w)$ contains $1, 2, 3$; otherwise, we can color $w$ with a color different from 1, 2, 3, $c(x_1')$ and $c(x_1'')$. W.l.o.g., we suppose that $L(w) = \{1, 2, 3, 4, 5\}$. If we cannot color $w$, this implies that $\{c(x_1'), c(x_1'')\} = \{c(x_2'), c(x_2'')\} = \{4, 5\}$. Observe now that $L(x_1) = L(x_2) = \{1, 2, 3, 4, 5\}$; otherwise, we can recolor $x_1$ with a color different from 1, 2, 3, 4, 5 to get case 4.1. So, we recolor $x_1$ and $x_2$ with 3 and color $w$ with 1.

4.2.1.2 If $c(x_3) = c(z)$ and $c(x_3) \neq 1$, we assume that $c(x_3) = 2$. Observe first that $c(x_3') = c(x_3'')$; otherwise, we can recolor $x_3$ with a color different from $1, 2, c(x_3'), c(x_3'')$ to get case 4.2.1.1. So, suppose that $c(x_3') = c(x_3'') = 3$ ($c(x_3') = c(x_3'') = 1$ is an easier case). Necessarily, $L(w)$ contains 1, 2, 3; otherwise, we can color $w$ with a color different from 1, 2, 3, $c(x_1')$ and $c(x_1'')$. W.l.o.g., $L(w) = \{1, 2, 3, 4, 5\}$, and $\{c(x_1'), c(x_1'')\} = \{c(x_2'), c(x_2'')\} = \{4, 5\}$. So, we recolor $x_1$ and $x_2$ with a color different from 1, 2, 4, 5 and we color $w$ with 1.

4.2.2 Suppose that $c(x_1) = c(z)$. W.l.o.g. we assume that $c(x_1) = 1$. Observe that $c(x_2) \neq c(x_3)$; otherwise, we get case 4.2.1.2. We assume that $c(x_2) = 2$ and $c(x_3) = 3$. Observe that $c(x_1') = c(x_1'')$; otherwise we can recolor $x_1$ with a color different from $1, c(x_1'), c(x_1'')$ to get case 4.1 or 4.2.1.1. Hence, we color $w$ with a color different from 1, 2, 3, $c(x_1')$.

4.3 Suppose that exactly three neighbors of $w$ have the same color.

4.3.1 We assume that $c(x_1) = c(x_2) = c(x_3) = 1$ and $c(z) = 2$. Observe that $c(x'_1) = c(x''_1)$; otherwise, we can recolor $x_1$ with a color different from 1, 2, $c(x'_1), c(x''_1)$ to get case 4.2.1.1. by the same way, $c(x'_i) = c(x''_i)$, $i = 1, 2, 3$. Then $L(w) = \{1, 2, c(x'_1), c(x'_2), c(x'_3)\}$ with $c(x'_1) \neq c(x'_2) \neq c(x'_3) \neq c(x'_1)$; otherwise, we color $w$ with a color different from $1, 2, c(x'_1), c(x'_2), c(x'_3)$. So, we color $w$ with $c(x'_1)$.

4.3.2 We assume that $c(z) = c(x_1) = c(x_2) = 1$ and $c(x_3) = 2$. As above, observe that $c(x'_1) = c(x''_1)$ and $c(x'_2) = c(x''_2)$; otherwise we can recolor $x_1$ or $x_2$ to obtain a previous case. Hence, we color $w$ with a color different from 1, 2, $c(x'_1), c(x'_2)$.

4.4 All the neighbors of $w$ have the same color. Suppose that $c(x_1) = c(x_2) = c(x_3) = c(z) = 1$. As above, for $i = 1, 2, 3$, $c(x'_i) = c(x''_i)$ (otherwise we can get a previous case). We color $w$ with a color different from 1, $c(x'_1), c(x'_2), c(x'_3)$.

□

We use the following discharging rule: Each $\geqslant$4-vertex gives $\frac{5}{7}$ to each of its 2-neighbors, $\frac{3}{14}$ to each of its 3-neighbors and $\frac{1}{7}$ to each of its 6(4)-neighbors. Let us check that for every $v \in V(H)$, $d^*(v) \geqslant \frac{24}{7}$:

- If $d(v) = 2$, then $v$ has two $\geqslant$5-neighbors by Lemma 2.9.(5), so $d^*(v) = 2 + 2\frac{5}{7} = \frac{24}{7}$.

- If $d(v) = 3$, then $v$ has at least two $\geqslant$4-neighbor by Lemma 2.9.(5) and Lemma 2.10.(1), so $d^*(v) \geqslant 3 + 2\frac{3}{14} = \frac{24}{7}$.

- If $d(v) = 4$, then $v$ has no 2-neighbor by Lemma 2.9.(5), no 6(4)-neighbor by Lemma 2.11.(3), and at most two 3-neighbors by Lemma 2.11.(4), so $d^*(v) \geqslant 4 - 2\frac{3}{14} = \frac{25}{7} > \frac{24}{7}$.

- If $d(v) = 5$, then $v$ has at most one 2-neighbor by Lemma 2.9.(4), so $d^*(v) \geqslant 5 - \frac{5}{7} - 4\frac{3}{14} = \frac{24}{7}$.

- If $d(v) = 6$, by Lemma 2.9.(3), $v$ has at most four 2-neighbors. If $v$ is a 6(4)-vertex, then it has no weak neighbor by Lemma 2.11.(1), so $d^*(v) = 6 - 4\frac{5}{7} + 2\frac{1}{7} = \frac{24}{7}$. If $v$ has three 2-neighbors, then it has at most two weak neighbors by Lemma 2.11.(2), so $d^*(v) \geqslant 6 - 3\frac{5}{7} - 2\frac{3}{14} = \frac{24}{7}$. Otherwise, $v$ has at most two 2-neighbors, so $d^*(v) \geqslant 6 - 2\frac{5}{7} - 4\frac{3}{14} = \frac{26}{7} > \frac{24}{7}$.

- If $d(v) = k, 7 \leqslant k \leqslant 10$, then $v$ has at most $(k-2)$ 2-neighbors by Lemma 2.9.(3). If $v$ is a $k(k-2)$-vertex, then it has no weak neighbor by Lemma 2.11.(1) and $d^*(v) = k - (k-2)\frac{5}{7} = \frac{2k+10}{7} \geqslant \frac{24}{7}$. Otherwise, $d^*(v) \geqslant k - (k-3)\frac{5}{7} - 3\frac{3}{14} = \frac{4k+21}{14} \geqslant \frac{7}{2} > \frac{24}{7}$.

- If $d(v) = 11$, then $v$ has at most nine 2-neighbors by Lemma 2.9.(3), so $d^*(v) \geqslant 11 - 9\frac{5}{7} - 2\frac{3}{14} = \frac{29}{7} > \frac{24}{7}$.

- If $d(v) = k \geqslant 12$, then $d^*(v) \geqslant k - k\frac{5}{7} = \frac{2k}{7} \geqslant \frac{24}{7}$.

### 2.6.4 Optimality of Theorem 2.7

In order to study the tightness of Theorem 2.7, we introduce two measuring functions.

**Definition 2.1.** *Let* $f : \mathbb{N} \to \mathbb{R}$ *be the function defined by* $f(n) = \inf\{mad(H) \mid \chi_a(H) > n\}$.

**Definition 2.2.** *Let* $f^l : \mathbb{N} \to \mathbb{R}$ *be the function defined by* $f^l(n) = \inf\{mad(H) \mid \chi_a^l(H) > n\}$.

It is easy to check that $f^l(2) = f(2) = 2$. By Theorem 2.7, we have lower bounds on $f^l(n)$ for $3 \leqslant n \leqslant 5$, namely $f^l(3) \geqslant \frac{8}{3}$, $f^l(4) \geqslant \frac{19}{6}$, and $f^l(5) \geqslant \frac{24}{7}$. We now give graphs that provide upper bounds on these quantities.

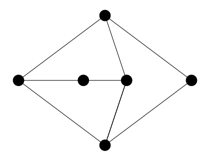

Figure 2.18: A graph $G$ with $mad(G) = \frac{8}{3}$ such that $\chi_a(G) = \chi_a^l(G) = 4$.

The graph $G$ with $mad(G) = \frac{8}{3}$ depicted in Figure 2.18 is acyclically 4-choosable by Theorem 2.7.(2). To see that $G$ is not acyclically 3-colorable, consider its four 3-vertices: Any two of them are either adjacent or have three common neighbors. Thus, different colors must be assigned to four vertices in any acyclic 3-coloring of $G$. This contradiction shows that:

$$f^l(3) = f(3) = \frac{8}{3}$$

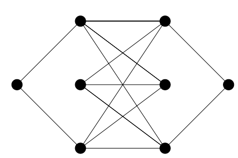

Figure 2.19: A graph $G$ with $mad(G) = \frac{13}{4}$ such that $\chi_a(G) = \chi_a^l(G) = 5$.

The graph $G$ with $mad(G) = \frac{13}{4}$ depicted in Figure 2.19 is acyclically 5-choosable: First, we assign five distinct colors to the four 4-vertices and to one of the 3-vertex, then we assign proper colors to the other vertices. To see that $G$ is not acyclically 4-colorable, consider its four 4-vertices: Any two of them are either adjacent or have four common neighbors. Thus, different colors are assigned to the 4-vertices in any acyclic 4-coloring of $G$. Now, observe that properly coloring the 3-vertices produces a bicolored $C_4$ in every case. This contradiction shows that:

$$\frac{19}{6} \leqslant f^l(4) \leqslant f(4) \leqslant \frac{13}{4}$$

The graph $G$ with $mad(G) = \frac{11}{3}$ depicted in Figure 2.20 is acyclically 6-choosable: First, we assign distinct colors to the six 7-vertices, then we assign proper colors to the 2-vertices. To see that $G$ is not acyclically 5-colorable, consider its six 7-vertices: Any two of them are

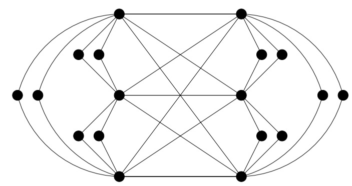

Figure 2.20: A graph $G$ with $mad(G) = \frac{11}{3}$ such that $\chi_a(G) = \chi_a^l(G) = 6$.

either adjacent or have five common neighbors. Thus, different colors must be assigned to six vertices in any acyclic 5-coloring of $G$. This contradiction shows that:

$$\frac{24}{7} \leqslant f^l(5) \leqslant f(5) \leqslant \frac{11}{3}$$

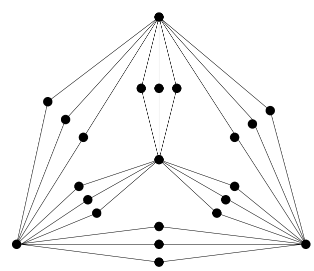

Figure 2.21: The graph $G_n$ is such that $mad(G_n) = 4 - \frac{8}{n^2+2}$ and $\chi_a(G_n) = \chi_a^l(G_n) = n + 1$.

We now use the construction proposed in [41] to obtain an asymptotic upper bound on $f(n)$. Let $G_n$ be the graph defined as follows: $G_n$ is a $(n + 1)$-clique in which each edge is replaced by $n$ paths with length 2 (see the graph $G_3$ depicted in Figure 2.21). It is easy to see that $mad(G_n) = 4 - \frac{8}{n^2+2}$. The graph $G_n$ is acyclically $(n+1)$-choosable: First, we assign distinct colors to the $^>2$-vertices, then we assign proper colors to the 2-vertices. To see that $G_n$ is not acyclically $n$-colorable, consider its $^>2$-vertices: Any two of them have $n$ common neighbors. Thus, different colors must be assigned to $n + 1$ vertices in any acyclic $n$-coloring of $G_n$. This contradiction shows that:

$$f(n) \leqslant 4 - \frac{8}{n^2 + 2}.$$

**Problem 2.1.**

- *What are the values of* $f^l(n)$ *and* $f(n)$ *for* $n > 3$ *?*
- *Does the equality* $f^l(n) = f(n)$ *hold also for every* $n > 3$ *?*

We remark that we cannot reach the results of [11] applied to the acyclic choosability without using some contraints of planarity: Indeed, to imply Theorem 1.3.(4), we should have proven that every graph $G$ with $mad(G) < \frac{10}{3}$ is acyclically 4-choosable, which is false since there exists a graph $G$ with $mad(G) = \frac{13}{4} < \frac{10}{3}$ which is not acyclically 4-colorable (see Figure

2.19). Similarly, it is impossible to prove that every graph $G$ with $mad(G) < \frac{14}{5}$ is acyclically 3-choosable to imply Theorem 1.3.(3), since there exists a graph $G$ with $mad(G) = \frac{8}{3} < \frac{14}{5}$ which is not acyclically 3-colorable (see Figure 2.18).

## 2.7 Relationship between $\chi_a$, $\chi^l$ and $\chi_a^l$

We first consider the relationship between $\chi_a$ and $\chi^l$. The graph $G_n$ above satisfies $\chi_a(G_n) = \chi_a^l(G_n) = n+1$ (we assign $n+1$ distinct colors to the $n+1$ vertices of the initial $(n+1)$-clique; then we color the remaining vertices with proper colors) and $\chi^l(G_n) = 3$, thus we cannot bound $\chi_a(G)$ by a function of $\chi^l(G)$ for a general graph $G$. On the other hand, we can show that $\chi^l(G) \leqslant 2\chi_a(G) - 2$ by using the following lemma:

**Lemma 2.12.** *[82] Every maximal acyclically $k$-colorable graph with $n$ vertices has exactly $(k-1)(n-\frac{k}{2})$ edges.*

Suppose $k \geqslant 2$: Lemma 2.12 implies that if a graph $G$ is acyclically $k$-colorable, then $G$ has arboricity $k-1$, so $G$ is $(2k-3)$-degenerate, and thus $G$ is $(2k-2)$-choosable. The previous result is best possible for $k = 2$ since $\chi_a(K_2) = \chi^l(K_2) = 2$.

We now consider the relationship between $\chi_a$ and $\chi_a^l$.

**Lemma 2.13.** *Let $G$ be a graph which is not $l$-choosable. Let $G'$ be the graph obtained by replacing every edge $uv$ of $G$ by $l$ 2-vertices, each adjacent to $u$ and $v$. Then $G'$ is bipartite and 2-degenerate, but not acyclically $l$-choosable. Moreover, if $G$ is bipartite then $G'$ belongs to $\mathcal{S}_0 \odot \mathcal{D}_1$, otherwise we have $\chi_a(G') = \chi(G)$.*

*Proof.* The graph $G'$ is clearly bipartite and 2-degenerate. A vertex of $G'$ that is also in $G$ is called *old*, and for each edge $uv$ of $G$, the non-old vertices of $G'$ adjacent to $u$ and $v$ are called $(u,v)$-vertices. If $G$ is bipartite, we have a bipartition $V(G) = V_0 \cup V_1$. In $G'$, we color an old vertex $v$ with $c(v) = i$ if and only if $v$ belongs to $V_i$ in $G$, and we color 1 the non old vertices. This gives a $\mathcal{S}_0 \odot \mathcal{D}_1$-coloring of $G'$. If $\chi(G) = p \geqslant 3$, we can get an acyclic $p$-coloring of $G'$ as follows: The colors of the old vertices of $G'$ are induced by a proper $p$-coloring of $G$. To color the $(u,v)$-vertices, we use a color in $S$ distinct from $c(u)$ and $c(v)$: Such a color exists since $p \geqslant 3$. We check easily that this coloring is acyclic. Finally we have to show that $\chi_a^l(G') > l$. Let $L$ be a list assignment of the old vertices with lists of size $l$. For each edge $uv$ of $G$, pick one endpoint $u$, and assign the list $L(u)$ to every $(u,v)$-vertex. Suppose $c(u) = c(v)$. To avoid a bicolored $C_4$, no two $(u,v)$-vertices can get the same color. There are $l$ such vertices but only $l-1$ colors in the set $L(u) \setminus c(u)$. This contradiction shows that $c(u) \neq c(v)$. Given a non-colorable list assignment of $V(G)$ with lists of size $l$, we can thus produce a list assignment of $V(G')$ with lists of size $l$ that is not acyclically colorable. □

It is well known that, for any fixed $k$, there exist bipartite graphs which are not $k$-choosable. There also exist 3-colorable non-4-choosable planar graphs, see [53, 81]. We can use Lemma 2.13 with these graphs to obtain the following corollary.

**Corollary 2.2.**

1. *For any fixed $k$, there exist bipartite 2-degenerate graphs in $\mathcal{S}_0 \odot \mathcal{D}_1$ which are not acyclically $k$-choosable.*

2. *There exist bipartite 2-degenerate planar graphs which are acyclically 3-colorable but not acyclically 4-choosable.*

# Chapter 3

# Edge colorings

Recall that a $T$-$fa$ coloring is an improper coloring of the edges such that every color class induces a forest which does not contain the tree $T$. In this chapter, we consider certain $T$-$fa$ colorings of some subclasses of planar graphs.

## 3.1 Introduction

A *foot* is an edge of a tree $T$ incident to a leaf of $T$.

If we remove zero or more leaves from a tree $T$, we obtain a *skeleton* $T'$ of $T$. If we remove every leaf of a tree $T$, we obtain the *basic skeleton* $T'$ of $T$.

**Lemma 3.1.** *For any tree $P$ with basic skeleton $P'$ and any tree $T$ with skeleton $T'$, if $T'$ is $P'$-free then $T$ is $P$-free.*

*Proof.* We show the contraposition. If $T$ contains $P$, then every leaf of $T$ is either a leaf of $P$ or does not belong to $P$. So, by removing some leaves of $T$, we do not remove non-leaves of $P$. Thus any skeleton $T'$ contains the basic skeleton $P'$. □

**Definition 3.1.** *For any tree $P$ with basic skeleton $P'$, a $k$-$P$-$fa$ coloring of $G$ is a $k$-coloring and a partial orientation of its edges such that:*

- *The graph induced by a color class is a $P$-free forest.*
- *If the edge $uv$ is colored $i$ and is oriented towards $v$, then $v$ is a leaf in the $i^{th}$ forest.*
- *The graph induced by the unoriented edges of a color class is a $P'$-free forest.*

By Lemma 3.1, a graph $G$ has $P$-$fa(G) \leqslant k$ if and only if it has a $k$-$P$-$fa$ coloring. Notice that the basic skeletons of the trees $P_4$ and $S_n$, with $n > 1$, are respectively $K_2$ and $K_{1,n}$. This implies that the forest induced by the unoriented edges in a $k$-$P_4$-$fa$ coloring (resp. $k$-$S_n$-$fa$ coloring) has maximum degree 0 (resp. $n-1$). If the graph $G$ is $k$-$P$-$fa$ colored, for each of its $k$ forests, we distinguish two types of vertices: The *leaves*, which have an incident arc in this forest oriented toward them, and the *inner vertices*. A $k$-$P$-$fa$ coloring is said to be *suitable* if every vertex is an inner vertex in at least $k-1$ forests.

**Theorem 3.1.** *Positive results:*

1. *Every graph $G$ of girth at least 5 with $mad(G) < \frac{10}{3}$ has a suitable 3-$S_4$-$fa$ coloring.*
2. *Every graph $G$ of girth at least 6 with $mad(G) < 3$ has a suitable 3-$S_3$-$fa$ coloring.*
3. *Every graph $G$ of girth at least 10 with $mad(G) < \frac{5}{2}$ has a suitable 2-$S_3$-$fa$ coloring.*
4. *Every graph $G$ of girth at least 14 with $mad(G) < \frac{7}{3}$ has a suitable 2-$S_2$-$fa$ coloring.*

By Observation 1.1, we get:

**Corollary 3.1.**

1. *If $G \in \mathcal{P}_5$, then $S_4$-$fa(G) \leqslant 3$.*
2. *If $G \in \mathcal{P}_6$, then $S_3$-$fa(G) \leqslant 3$.*
3. *If $G \in \mathcal{P}_{10}$, then $S_3$-$fa(G) \leqslant 2$.*
4. *If $G \in \mathcal{P}_{14}$, then $S_2$-$fa(G) \leqslant 2$.*

**Theorem 3.2.** *Negative results:*

1. *For every tree $T$, there exist 2-degenerate planar bipartite graphs having no edge partition into a forest and a $T$-free forest.*
2. *For every tree $T$, there exist partial 2-trees having no edge partition into a forest and a $T$-free forest.*

## 3.2 Proofs of the positive results

Every item of Theorem 3.1 is proved using the method described in Section 1.6. The partial order on graphs considered here is the subgraph partial order. Let us first present a general lemma which is used in the next subsections.

**Lemma 3.2.** *Let $n \geqslant 2$, $k \geqslant 2$ and let $H$ be a minimal graph having no suitable $k$-$S_n$-$fa$ coloring. Then $H$ does not contain*

1. *a $^{\leqslant}1$-vertex (for $n \geqslant 2$, $k \geqslant 2$),*
2. *a 2-vertex adjacent to a $(2n-1)$-vertex (for $n \geqslant 2$, $k = 3$).*

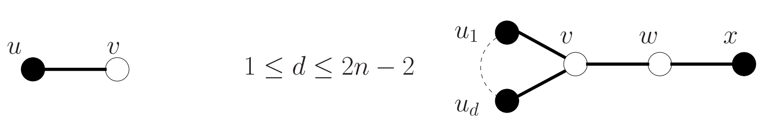


Figure 3.1: Forbidden configurations for Lemma 3.2.

*Proof.*

1. Suppose $H$ contains the configuration depicted in Figure 3.1 (left). Consider a suitable $k$-$S_n$-$fa$ coloring of $H \setminus \{v\}$. Since $k-1 \geqslant 1$, $u$ is an inner vertex in at least one $S_n$-free forest, say $F_i$. We can extend this coloring to a suitable $k$-$S_n$-$fa$ of $H$ by orienting the edge $uv$ towards $v$ and coloring $uv$ with $i$.

2. Suppose $H$ contains the configuration depicted in Figure 3.1 (right). Consider a suitable coloring of $H \setminus \{w\}$ into three forests $F_1$, $F_2$, $F_3$. If $v$ is an inner vertex in all three forests, then there is a forest $F_i$ such that $v$ is incident to at most $\lfloor \frac{2n-2}{3} \rfloor \leqslant n-2$ edges in $F_i$. If $v$ is an inner vertex in two forests, say $F_2$, $F_3$, then exactly one edge, say $u_1v$, is a foot edge for $F_1$. So there is an edge forest $F_i$ such that $v$ is incident to at most $\lfloor \frac{2n-3}{2} \rfloor \leqslant n-2$ unoriented edges in $F_i$. The vertex $x$ is an inner vertex in at least two forests, say $F_j$, $F_k$, and thus either $j$ or $k$ is distinct from $i$ (say $j$). Now we can color $vw$ with $i$ and $wx$ with $j$. We check that the coloring is suitable: $w$ may not be an inner vertex in $F_j$ but is an inner vertex in both $F_i$ and the third forest.

□

### 3.2.1 Proof of Theorem 3.1.(1)

**Lemma 3.3.** *Let $H$ be a minimal graph of girth at least 5 having no suitable 3-$S_4$-$fa$ coloring. Then $H$ does not contain*

1. *a 3-vertex adjacent to two 3-vertices,*
2. *a $d$-vertex adjacent to $(d-1)$ 2-vertices and one $^{\leqslant}$3-vertex (for $d \leqslant 9$).*

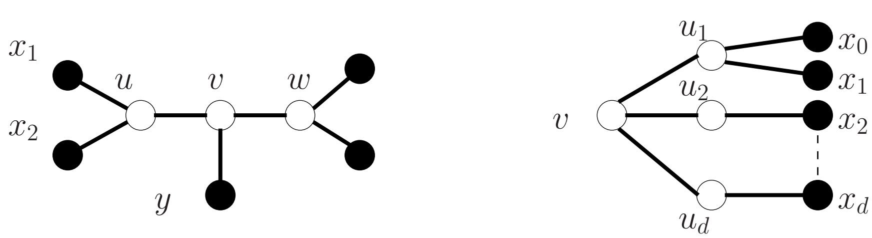


Figure 3.2: Forbidden configurations for Lemma 3.3.

*Proof.*

1. Suppose $H$ contains the configuration depicted in Figure 3.2 (left). Consider a suitable 3-$S_4$-$fa$ coloring of $H \setminus \{v\}$ into three forests $F_1$, $F_2$, $F_3$. Let $w$ be an inner vertex in the forests $F_1$ and $F_2$. If $u$ is an inner vertex in $F_3$, let $uv$ be unoriented and colored 3, and let $yv$ be oriented toward $v$ and colored 1 or 2 (according to the status of $v$ in $F_1$ and $F_2$), say 1. Now, we just let $vw$ be unoriented and colored 2, and obtain a suitable 3-$S_4$-$fa$ coloring of $H$. If $u$ is a leaf in $F_3$, one of its incident edges, say $x_1u$, is oriented

toward $u$ and colored 3. If the edge $x_2u$ is colored 1 (resp. 2), let the edges $uv$ and $vw$ be unoriented and colored 2 (resp. 1). We orient $yv$ toward $v$ and color it 1 or 3 (resp. 2 or 3), thus we obtain a suitable 3-$S_4$-$fa$ coloring of $H$.

2. Suppose $H$ contains the configuration depicted in Figure 3.2 (right). Consider a suitable 3-$S_4$-$fa$ coloring of $H \setminus \{v, u_2, \ldots, u_d\}$ into three forests $F_1$, $F_2$, $F_3$. W.l.o.g. consider that $u_1$ is an inner vertex in the forests $F_1$. Let all the edges $vu_i$, for $1 \leqslant i \leqslant d$, be unoriented and color them 1 if $1 \leqslant i \leqslant 3$, 2 if $4 \leqslant i \leqslant 6$, and 3 if $7 \leqslant i \leqslant 9$. We orient edges $x_iu_i$ toward $u_i$ and color them with a color distinct from the color of $vu_i$. Thus, we obtain $H$ has a suitable 3-$S_4$-$fa$ coloring of $H$.

We use the following discharging rule: each ${}^{\geqslant}$4-vertex gives $\frac{2}{3}$ to each of its 2-neighbors and $\frac{1}{6}$ to each of its 3-neighbors. Let us check that for every $v \in V(H)$, $d^*(v) \geqslant \frac{10}{3}$:

- $d(v) = 2$: $v$ has two ${}^{\geqslant}$8-neighbors by Lemma 3.2.(2), so $d^*(v) = 2 + 2\frac{2}{3} = \frac{10}{3}$.
- $d(v) = 3$: $v$ has no 2-neighbor and at least two ${}^{\geqslant}$4-neighbor by Lemma 3.3.(1), so $d^*(v) \geqslant 3 + 2\frac{1}{6} = \frac{10}{3}$.
- $d(v) = k, 4 \leqslant k \leqslant 7$: $v$ has no 2-neighbor so $d^*(v) \geqslant k - k\frac{1}{6} = k\frac{5}{6} \geqslant \frac{10}{3}$.
- $d(v) = k, 8 \leqslant k \leqslant 9$, then $v$ has at most $(k-1)$ 2-neighbors by Lemma 3.3.(2), so $d^*(v) \geqslant k - (k-1)\frac{2}{3} = \frac{k+2}{3} \geqslant \frac{10}{3}$.
- $d(v) = k \geqslant 10$: $d^*(v) \geqslant k - k\frac{2}{3} = \frac{k}{3} \geqslant \frac{10}{3}$.

□

### 3.2.2 Proof of Theorem 3.1.(2)

We use the following discharging rule: each vertex gives $\frac{1}{2}$ to each of its 2-neighbors. Let us check that for every $v \in V(H)$, $d^*(v) \geqslant 3$:

- $d(v) = 2$: $v$ has two ${}^{\geqslant}$6-neighbors by Lemma 3.2.(2), so $d^*(v) = 2 + 2\frac{1}{2} = 3$.
- $d(v) = k, 3 \leqslant k \leqslant 5$: $v$ has no 2-neighbor by Lemma 3.2.(2), so $d^*(v) = k \geqslant 3$.
- $d(v) = k \geqslant 6$: $d^*(v) \geqslant k - k\frac{1}{2} = \frac{k}{2} \geqslant 3$.

### 3.2.3 Proof of Theorem 3.1.(3)

**Lemma 3.4.** *Let $H$ be a minimal graph of girth at least 10 having no suitable 2-$S_3$-$fa$ coloring. Then $H$ does not contain any of the configurations depicted in Figure 3.3.*

*Proof.*

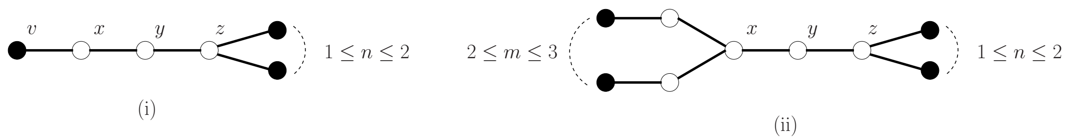


Figure 3.3: Forbidden configurations for Lemma 3.3.

(i) Suppose $H$ contains the configuration (i) depicted in Figure 3.3. Consider a suitable 2-$S_3$-$fa$ coloring of $H \backslash \{y\}$. In every case, $z$ is an inner vertex in some forest $F_i$ such that $z$ is incident to at most one non-oriented edge colored $i$. We can extend this coloring to $H$ such that $xy$ and $yz$ are non-oriented, $vx$ and $xy$ get different colors, and $yz$ gets color $i$.

(ii) Suppose $H$ contains the configuration (ii) depicted in Figure 3.3. Consider a suitable 2-$S_3$-$fa$ coloring of the graph $H'$ obtained from $H$ by deleting the edge $yz$. we can always modify this coloring into a suitable 2-$S_3$-$fa$ coloring of $H'$ such that $xy$ is non-oriented and there exist no monochromatic path connecting $y$ to any $u_i$. In every case, $z$ is an inner vertex in some forest $F_i$ such that $z$ is incident to at most one non-oriented edge colored $i$. We can extend this coloring to $H$ such that $yz$ is non-oriented and gets color $i$.

□

A 3-vertex is *weak* if it has three 2-neighbors. A 2-vertex is *weak* if is adjacent to a 2-vertex or a weak 3-vertex. We use the following discharging rule: each $^{\geqslant}$4-vertex gives $\frac{1}{2}$ to its weak 2-neighbors and $\frac{1}{4}$ to its non-weak 2-neighbors, each non-weak 3-vertex gives $\frac{1}{4}$ to its 2-neighbors. Let us check that for every $v \in V(H)$, $d^*(v) \geqslant 3$:

- $d(v) = 2$: if $v$ is weak, then $v$ has a $^{\geqslant}$4-neighbor (see Figure 3.3.(i) and Figure 3.3.(ii) with $m = 2$), so $d^*(v) = 2 + \frac{1}{2} = \frac{5}{2}$. Otherwise $v$ receives $\frac{1}{4}$ from each neighbor, so $d^*(v) \geqslant 2 + 2\frac{1}{4} = \frac{5}{2}$.

- $d(v) = 3$: if $v$ is weak, then $d^*(v) = d(v) = 3 > \frac{5}{2}$. Otherwise $v$ has at most two 2-neighbors, so $d^*(v) \geqslant 3 - 2\frac{1}{4} = \frac{5}{2}$.

- $d(v) = 4$: if $v$ has four 2-neighbors, then its 2-neighbors are not weak (see Figure 3.3.(ii)with $m = 3$), so $d^*(v) \geqslant k - k\frac{1}{4} = \frac{3k}{4} \geqslant 3 > \frac{5}{2}$. Otherwise $v$ has at most three 2-neighbors, so $d^*(v) \geqslant 4 - 3\frac{1}{2} = \frac{5}{2}$.

- $d(v) = k \geqslant 5$: $d^*(v) \geqslant k - k\frac{1}{2} = \frac{k}{2} \geqslant \frac{5}{2}$.

### 3.2.4 Proof of Theorem 3.1.(4)

**Lemma 3.5.** *Let $H$ be a minimal graph of girth at least 14 having no suitable 2-$S_2$-$fa$ coloring. Then $H$ does not contain the configuration depicted in Figure 3.4.*

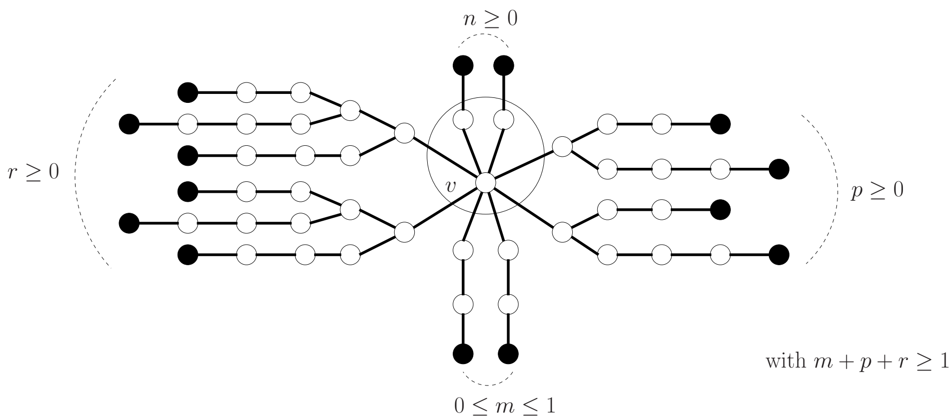


Figure 3.4: Forbidden configuration for Lemma 3.5.

*Proof.* Suppose that a $H$ contains this configuration. The graph $H'$ is obtained from $H$ by removing the white vertices of the configuration, except those inside the circle. The condition $m+p+r \geqslant 1$ ensures that $H'$ is a proper subgraph of $H$. Consider a suitable 2-$S_2$-$fa$ coloring of $H'$. If necessary, we can always modify this coloring into a suitable 2-$S_2$-$fa$ coloring of $H'$ such that $v$ is an inner vertex in both forests. Then this modified coloring can be extended in a suitable 2-$S_2$-$fa$ coloring of the whole graph $H$. □

A $q$-chain is a path of $q$ consecutive 2-vertices. The configuration in Figure 3.4 with $n = m = 1,\ p = r = 0$ ensures that $H$ contains no 4-chain. A 3-vertex is *very weak* if is adjacent to a 3-chain and a $^{\geqslant}$2-chain. A 3-vertex is *weak* if is adjacent to a 3-chain and a very weak vertex. A $^{\geqslant}$4-vertex or 3-vertex which is neither very weak nor weak is said to be *non-weak*. We use the following discharging rule: each $^{\geqslant}$3-vertex $v$ gives:

- $\frac{1}{6}$ to each 2-vertex that belong to a $q$-chain adjacent to $v$.
- $\frac{1}{6}$ to each weak vertex adjacent to $v$.
- $\frac{1}{3}$ to each very weak vertex adjacent to $v$.

Let us check that for every $v \in V(H)$, $d^*(v) \geqslant \frac{7}{3}$:

- $d(v) = 2$: $v$ gets $\frac{1}{6}$ from the two $^{\geqslant}$3-vertices adjacent to the $q$-chain containing $v$, so $d^*(v) = 2 + 2\frac{1}{6} = \frac{7}{3}$.
- $d(v) = 3$ and $v$ is very weak: the third neighbor of $v$ is a $^{\geqslant}$3-vertex (take $n = 2,\ m = 1,\ p = r = 0$) which is neither very weak (take $n = 2,\ p = 1,\ m = r = 0$) nor weak (take $n = 2,\ r = 1,\ m = p = 0$). So $v$ has a non-weak neighbor and $d^*(v) \geqslant 3 + \frac{1}{3} - 2 \times 3\frac{1}{6} = \frac{7}{3}$.
- $d(v) = 3$ and $v$ is weak: the third neighbor of $v$ is a $^{\geqslant}$3-vertex (take $n = 2,\ p = 1,\ m = r = 0$) which is neither very weak (take $n = 2,\ r = 1,\ m = p = 0$) nor weak (take $m = 0,\ n = p = r = 1$). So $v$ has a non-weak neighbor and $d^*(v) \geqslant 3 + \frac{1}{6} - 3\frac{1}{6} - \frac{1}{3} = \frac{7}{3}$.

- $d(v) = 3$ and $v$ is non-weak: If all the neighbors of $v$ are 2-vertices, then they belong to 1-chains (take $n = 2,\ m = 1,\ p = r = 0$), so $d^*(v) = 3 - 3\frac{1}{6} = \frac{5}{2} > \frac{7}{3}$. Otherwise, $v$ has a non-weak neighbor (take any combination with $n < 3,\ m = 0,\ n + p + r = 3$). Since $v$ is non-weak, its two other neighbors are in the worst cases either two 2-chains, a 2-chain and a very weak vertex, two very weak vertices, a 3-chain and a 1-chain, or a 3-chain and a weak vertex. In every case $v$ gives at most $\frac{2}{3}$, so $d^*(v) \geqslant 3 - \frac{2}{3} = \frac{7}{3}$.

- $d(v) = k \geqslant 4$: If $v$ has $k$ 2-neighbors, then they belong to 1-chains (take $n = k - 1,\ m = 1,\ p = r = 0$), so $d^*(v) = k - k\frac{1}{6} = \frac{5k}{6} \geqslant \frac{10}{3} > \frac{7}{3}$. Otherwise, $v$ has a non-weak neighbor (take any combination with $n < k,\ m = 0,\ n + p + r = k$), so $d^*(v) \geqslant k - 3(k-1)\frac{1}{6} = \frac{k+1}{2} \geqslant \frac{5}{2} > \frac{7}{3}$.

## 3.3 Proofs of the negative results

### 3.3.1 Proof of Theorem 3.2.(1)

We define the family of rooted trees $T_n,\ n \geqslant 0$. Take $T_0 = K_2$. For $n \geqslant 1$, $T_n$ is obtained from $n$ copies of $T_{n-1}$ and another vertex, the root of $T_n$, which is adjacent to the root of every copy of $T_{n-1}$. Let $G_0$ be the 4-cycle $(u_0 x v_0 y)$. For $n \geqslant 1$, $G_n$ is obtained from $2n$ copies of $G_{n-1}$, $G^i_{n-1}$ for $1 \leqslant i \leqslant 2n$, and two vertices, $u_n$ and $v_n$, which are adjacent to the vertices $u^i_{n-1}$ and $v^i_{n-1}$ of every copy of $G_{n-1}$.

**Proposition 3.1.** *In any edge partition of $G_n$ into two forests $F_1$ and $F_2$, either $u_n$ or $v_n$ is the root of a tree $T_n \subseteq F_2$.*

*Proof.* It is clear for $G_0$, since $F_1$ cannot cover all the edges of the cycle. So, we prove the proposition by induction on $n$. Consider an edge partition of $G_n$ into two forests $F_1$ and $F_2$. Since $u^i_{n-1}$ and $v^i_{n-1}$ play similar role in $G^i_{n-1}$, we assume w.l.o.g. that $u^i_{n-1}$ is the root of a tree $T_{n-1} \subseteq F_2 \cap G^i_{n-1}$ for $1 \leqslant i \leqslant 2n$. There is at most one vertex $u^i_{n-1}$ such that the edges $u_n u^i_{n-1}$ and $v_n u^i_{n-1}$ are in $F_1$, otherwise there would be a 4-cycle in $F_1$. Now if we consider the remaining $2n - 1$ vertices $u^i_{n-1}$, then the pigeonhole principle ensures that either $u_n$ or $v_n$, say $u_n$, has $n$ neighbors $u^i_{n-1}$ in $F_2$. This means that $u_n$ is the root of a tree $T_n \subseteq F_2$ and proves the proposition. □

It is easy to see that every tree is a subtree of $T_n$ for some $n$. So for every tree $T$, there exist some $G_n$ having no edge partition into a forest and a $T$-free forest. Since every $G_n$ is clearly 2-degenerate, planar, and bipartite, this proves Theorem 3.2.(1).

### 3.3.2 Proof of Theorem 3.2.(2)

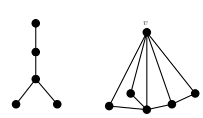

Figure 3.5: The tree $T$ and the graph $H_T$.

We define another family of rooted trees $T_n$, $n \geqslant 0$. Take $T_0 = K_1$. For $n \geqslant 1$, $T_n$ is obtained from two copies of $T_{n-1}$, such that the root of one copy is the root of $T_n$ and is adjacent to the root of the root of the other copy.

Let $T$ be a tree. We construct the graph $H_T$ by adding to $T$ a new vertex $v$ adjacent to every vertex of $T$ (see Figure 3.5). We now define a family $G_n$ of graphs, depending on $T$, with a specified vertex $s$. Let $G_0 = K_1$. For $n \geqslant 1$, $G_n$ is obtained from a copy $H$ of $H_T$, by identifying every vertex of $H$ with the specified vertex of copy of $G_{n-1}$. The specified vertex of $G_n$ is the vertex $v$ of $H$.

**Proposition 3.2.** *In any edge partition of $G_n$ into a forest $F_1$ and a $T$-free forest $F_2$, the specified vertex $s$ is the root of a 2-monochromatic copy of $T_n$.*

*Proof.* It is clear that $G_0$ contains a 2-monochromatic copy $T_0$. We prove the proposition by induction on $n$. Let $v$ denote the specified vertex of $G_n$. There exist two vertices $x, y \in H \setminus \{s\}$ such that $xy$ is colored 1, since otherwise $H \setminus \{s\}$ would be a 2-monochromatic copy of $T$. The edges $sx$ and $sy$ cannot be both colored 1, since it would induce a 1-monochromatic triangle $(sxy)$. So we assume w.l.o.g. that $sx$ is colored 2. By induction, both $s$ and $x$ are the root of a 2-monochromatic copy of $T_{n-1}$ in the associated copy of $G_{n-1}$. So $s$ is the root of a 2-monochromatic copy of $T_n$. □

It is easy to see that every tree is a subtree of $T_n$ for some $n$. So for every tree $T$, we can construct a family of partial 2-trees $G_i$ such that $G_i$ have no edge partition into a forest and a $T$-free forest if $i \geqslant n$ for some $n$. This proves Theorem 3.2.(2).

## 3.4 Summary of results

The table below summarizes what is now known about the $T$-free arboricity of planar graphs with given girth. Some bounds in this table are given by known results from the litterature. The bounds on the acyclic chromatic number are reported from Theorem 1.3 and the values of the arboricity are given by Nash-Williams' formula (1.1). Gonçalves proved that planar graphs have caterpillar arboricity (i.e. $S_3$-$fa$) at most four [28], and that there exist planar graphs with bistar arboricity (i.e. $S_2$-$fa$) five [29]. Corollary 3.1 gives new upper bounds on the $T$-free arboricity of some graph classes for some trees $T$. New lower bounds are derived from Theorem 3.2.(1) and Theorem 4.4.

| girth | $\chi_a$ | $P_4$ | $S_2$ | $S_3$ | $S_4$ | $U$ | arb |
|---|---|---|---|---|---|---|---|
| 3 | 5 | 5 | 5 | 4 | 4 | 3-4 | 3 |
| 4 | 5 | 4 | 4 | 4 | 4 | 3-4 | 2 |
| 5 | 3-4 | 3-4 | 3-4 | 3-4 | 3 | 2-3 | 2 |
| 6 | 3-4 | 3-4 | 3-4 | 3 | 3 | 2-3 | 2 |
| 7-8 | 3 | 3 | 3 | 2-3 | 2-3 | 2-3 | 2 |
| 9 | 3 | 3 | 2-3 | 2-3 | 2-3 | 2-3 | 2 |
| 10-13 | 3 | 3 | 2-3 | 2 | 2 | 2 | 2 |
| $\geqslant$14 | 3 | 3 | 2 | 2 | 2 | 2 | 2 |

Figure 3.6: Table of known results for planar graphs.

# Chapter 4

# NP-complete colorings

In this chapter, we expose NP-completeness results concerning the colorings considered in the two previous chapters. We chose to group them in a same chapter since their proofs often use similar polynomial reductions.

## 4.1 Introduction

If $\mathcal{C}_1$ and $\mathcal{C}_2$ are graph classes, we note $(\mathcal{C}_1 : \mathcal{C}_2)$ the problem of deciding whether a given graph $G \in \mathcal{C}_1$ belongs to $\mathcal{C}_2$. If $P_1$ and $P_2$ are decision problems, we note $P_1 \propto P_2$ if there is a polynomial reduction from $P_1$ to $P_2$. Notice that if $(\mathcal{C}_1 : \mathcal{C}_2)$ is NP-complete, then it implies in particular that $\mathcal{C}_1$ is not a subclass of $\mathcal{C}_2$. Actually, whenever we obtained a graph-theoretic negative result of the form "$\mathcal{C}_1 \not\subseteq \mathcal{C}_2$", we have tried to strengthen it into a complexity result of the form "$(\mathcal{C}_1 : \mathcal{C}_2)$ is NP-complete".

Kratochvil proved that PLANAR $(3,^{\leqslant} 4)$-SAT is NP-complete [44]. In this restricted version of SAT, the graph of incidences variable-clause of the input formula must be planar, every clause is a disjonction of exactly three literals, and every variable occurs in at most four clauses. A subcoloring is an improper coloring of the vertex set such that each color class induces a union of disjoint cliques. The problem 2-SUBCOLORABILITY is NP-complete on triangle-free planar graphs with maximum degree 4 [22, 27]. Notice that on triangle-free graphs, a 2-subcoloring corresponds to a vertex partition into two graphs with maximum degree 1. Finally, the problem 3-COLORABILITY is shown to be NP-complete on planar graphs with maximum degree 4 in [26]. Thus, using our notations, we have that $(\mathcal{P}_4 \cap \mathcal{S}_4 : \mathcal{S}_1^2)$ and $(\mathcal{P}_3 \cap \mathcal{S}_4 : \mathcal{S}_0^3)$ are NP-complete.

## 4.2 Acyclic and/or improper colorings

**Theorem 4.1.** PLANAR $(3,^{\leqslant} 4)$-SAT $\propto (\mathcal{P}_6 \cap \mathcal{S}_3 : \mathcal{S}_0 \circ \mathcal{S}_1)$

*Proof.* In a $\mathcal{S}_0 \circ \mathcal{S}_1$ or $\mathcal{S}_0 \odot \mathcal{S}_1$ coloring $c$, a vertex $v$ gets color $c(v) = i$ if $v$ is in the color class $\mathcal{S}_i$, $0 \leqslant i \leqslant 1$. We observe that the graph depicted in Figure 4.1(i) has no $\mathcal{S}_0 \circ \mathcal{S}_1$ coloring such that $c(u) = c(v) = 1$. This implies that the vertex $u$ in the graph depicted in Figure 4.1(ii) must be colored 1. Given an instance $I$ of PLANAR $(3,^{\leqslant} 4)$-SAT, we build a graph $G$ as

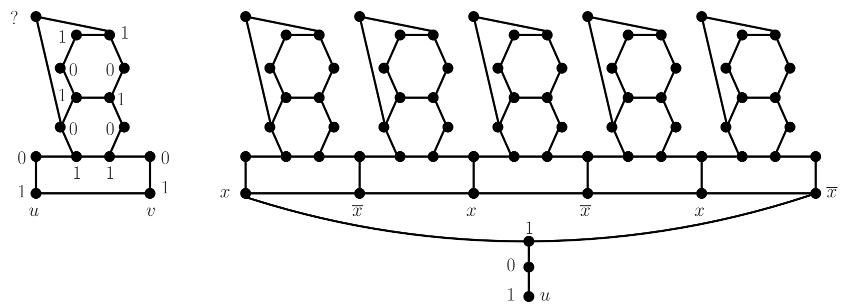


Figure 4.1: The forcing gadget for the reduction of Theorem 4.1.

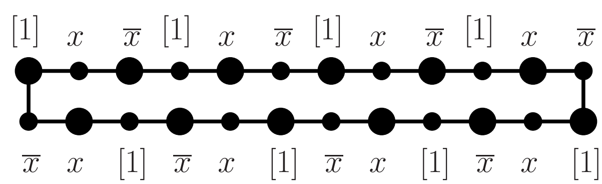


Figure 4.2: The variable gadget for the reduction of Theorems 4.1 and 4.2.

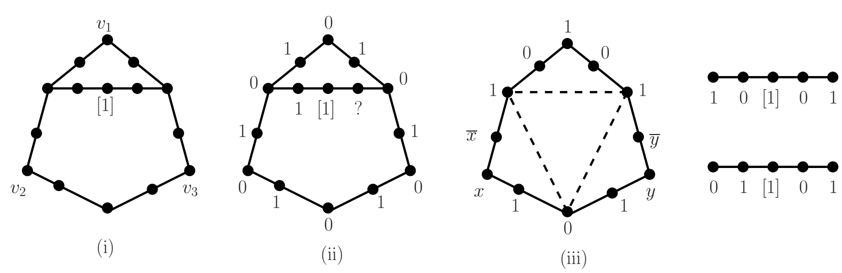


Figure 4.3: The clause gadget for the reduction of Theorems 4.1

follows. We replace every variable of $I$ by a copy the variable gadget depicted in Figure 4.2. We replace every clause of $I$ by a copy the clause gadget depicted in Figure 4.3(i). The way we link variables to clauses is best explained with an example: for a clause gadget $C = (x, \overline{y}, z)$ and variables gadgets $X, Y, Z$, we add an edge between a big vertex $\overline{x}$ of $X$ and the vertex $v_1$ of $C$, between a big vertex $y$ of $Y$ and the vertex $v_2$ of $C$, and between a big vertex $\overline{z}$ of $Z$ and the vertex $v_3$ of $C$. The boolean value true (resp. false) is associated with the color 0 (resp. 1). We see in figure 4.3(ii) that an unsatisfied clause is not $\mathcal{S}_0 \circ \mathcal{S}_1$ colorable, whereas any satisfied clause (i.e. such that at least one $v_1, v_2, v_3$ is colored 1) is colorable, see figure 4.3(ii). This means that $I$ is satisfiable if and only if $G$ belongs to $\mathcal{S}_0 \circ \mathcal{S}_1$. We easily check that $G$ is indeed planar, with girth 6, and maximum degree 3. □

Notice that on triangle-free graphs, the $\mathcal{S}_0 \circ \mathcal{S}_1$ coloring correspond to the $(1, 2)$-subcoloring defined in [47]. Theorem 4.1 improves a result in [47] stating that $(\mathcal{P}_4 \cap \mathcal{S}_3 : \mathcal{S}_0 \circ \mathcal{S}_1)$ is NP-complete.

**Theorem 4.2.** PLANAR $(3,^{\leqslant} 4)$-SAT $\propto (\mathcal{P}_{10} \cap \mathcal{S}_3 \cap bip : \mathcal{S}_0 \odot \mathcal{S}_1)$

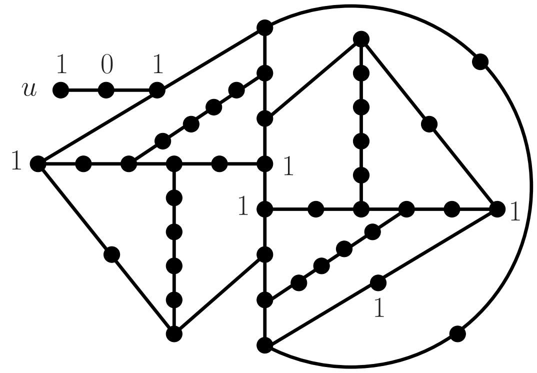


Figure 4.4: The forcing gadget for the reduction of Theorem 4.2.

*Proof.* The proof is similar to the previous one, with the following two changes. We use another forcing gadget depicted in Figure 4.4. In any $\mathcal{S}_0 \odot \mathcal{S}_1$ coloring of the forcing gadget, the vertex $u$ must be colored 1. The clause gadget is obtained from the one in Figure 4.3(i) by deleting the vertex forced to be colored 1 and its two 2-neighbors. If a clause is unsatisfied, then its clause gadget is not colorable (an alternating cycle $C_{12}$ is forbidden). If a clause is satisfied, then its clause gadget is colorable (the coloring in Figure 4.3(iii) is acyclic). □

**Theorem 4.3.** $(\mathcal{P}_4 \cap \mathcal{S}_4 : \mathcal{S}_1^2) \propto (\mathcal{P}_8 \cap \mathcal{S}_4 \cap bip : \mathcal{S}_1^{(2)})$

*Proof.* Consider the graph depicted in Figure 4.5(i). Any $\mathcal{S}_1^2$ coloring such that the vertex $u$ is colored 1 and has no neighbor colored 1 contains an alternating cycle $C_8$. So in every $\mathcal{S}_1^{(2)}$ coloring, both $u$ and one neighbor of $u$ must get the same color. Now we use three copies of this graph in the forcing gadget depicted in Figure 4.5(ii). Since the cycle of this forcing gadget cannot be alternating, two types of special edges are forced: monochromatic edges and edges whose endpoints have distinct colors and have a neighbor of the same color. Given a

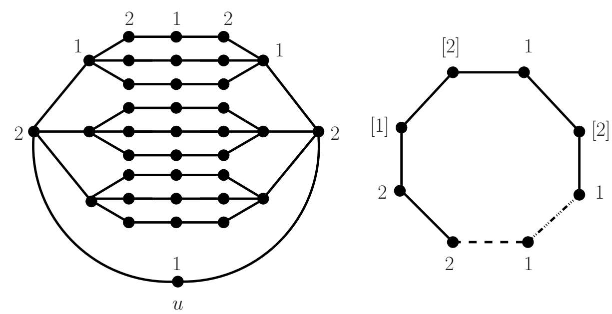


Figure 4.5: The forcing gadget for the reduction of Theorem 4.3.

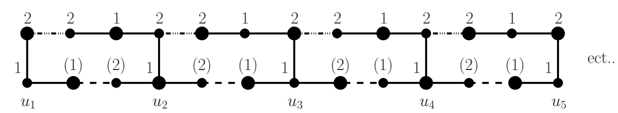


Figure 4.6: The vertex gadget for the reduction of Theorem 4.3.

planar graph $G$, we construct the graph $G'$ as follows. We replace every vertex of $G$ by a copy of the vertex gadget depicted in Figure 4.6 and for every edge $vw$ we link a big vertex $u_i$ in the gadget of $v$ to a small vertex $u_i$ in the gadget of $w$. In any $\mathcal{S}_1^{(2)}$ coloring of the vertex gadget, all $u_i$'s get the same color, say 1, and there exists no alternating path between distinct $u_i$'s. This common color in the gadget of a vertex $v$ corresponds to the color of $v$ in a $\mathcal{S}_1^2$ coloring of $G$. Notice that if one of the $u_i$'s, say $u_2$, has a neighbor colored 1 not in the gadget, then every other $u_i$ has a neighbor colored 1 in the gadget. Thus we can obtain a $\mathcal{S}_1^{(2)}$ coloring of $G'$ from a $\mathcal{S}_1^2$ coloring of $G$ and vice-versa. □

**Theorem 4.4.** $(\mathcal{P}_3 \cap \mathcal{S}_4 : \mathcal{S}_0^3) \propto (\mathcal{P}_4 \cap \mathcal{S}_4 \cap bip \cap \mathcal{D}_2 : \mathcal{S}_0^{(3)})$

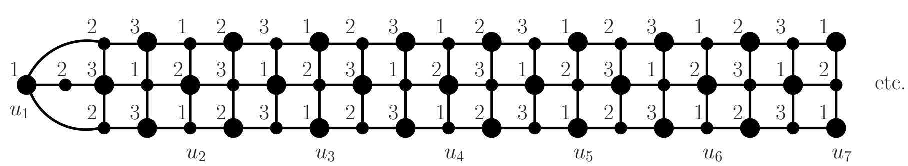


Figure 4.7: The vertex gadget for the reduction of Theorem 4.4.

*Proof.* Given a planar graph $G$, we construct the graph $G'$ as follows. We replace every vertex of $G$ by a copy of the vertex gadget depicted in Figure 4.7 and for every edge $vw$ we link a big vertex $u_i$ in the gadget of $v$ to a small vertex $u_i$ in the gadget of $w$. The given 3-acyclic coloring of the vertex gadget is the unique one up to permutation of colors. Notice that all $u_i$'s get the same color and there exists no alternating path between distinct $u_i$'s. This common color in the gadget of a vertex $v$ corresponds to the color of $v$ in a 3-coloring of $G$. Thus $G'$ is acyclically 3-colorable if and only if $G$ is 3-colorable. □

**Theorem 4.5.** $(\mathcal{P}_3 \cap \mathcal{S}_4 : \mathcal{S}_0^3) \propto (\mathcal{P}_4 \cap \mathcal{S}_8 \cap bip \cap \mathcal{D}_2 : \mathcal{S}_0^{(4)})$

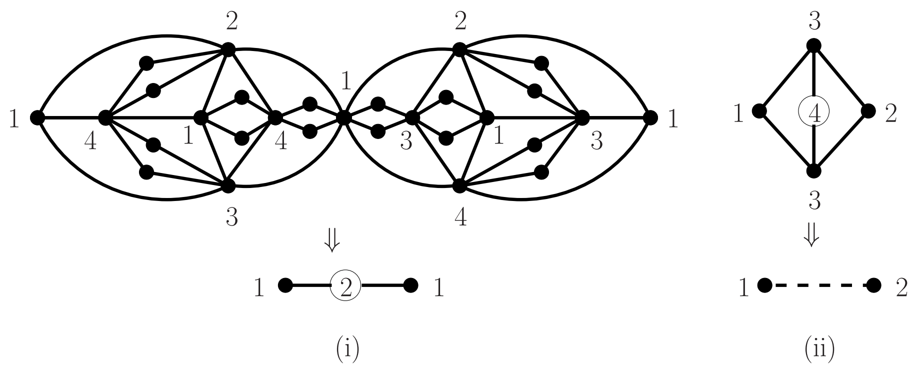


Figure 4.8: The forcing gadgets for the reduction of Theorem 4.5.

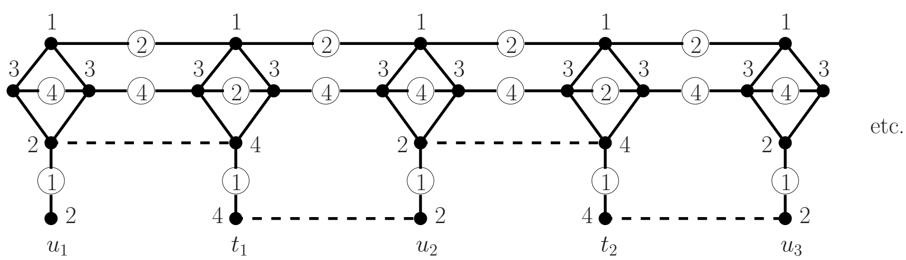


Figure 4.9: The vertex gadget for the reduction of Theorem 4.5.

*Proof.* Consider the graph depicted in Figure 4.8(i). Any $\mathcal{S}_0^{(4)}$ coloring is such that $x$ and $y$ get the same color. Moreover, this gadget has a coloring such that there exists only one alternating path between $x$ and $y$ (the path colored 1 and 2 in Figure 4.8(i)). In the graph depicted in Figure 4.8(ii), the vertices $x$ and $y$ must have distinct colors and there is no alternating path between $x$ and $y$. Given a planar graph $G$, we construct the graph $G'$ as follows. We replace every vertex of $G$ by a copy of the vertex gadget depicted in Figure 4.9 and for every edge $vw$ we connect a vertex $u_i$ (resp. $t_i$) in the gadget of $v$ to a vertex $u_i$ (resp. $t_i$) in the gadget of $w$ using the forcing gadget (ii) (resp. (i)). In any $\mathcal{S}_0^{(4)}$ coloring of the vertex gadget, all $u_i$'s get the same color and all $t_i$'s get a same color distinct from the color of the $u_i$'s. The color of the $u_i$'s in the gadget of a vertex $v$ corresponds to the color of $v$ in a 3-coloring of $G$. The color of the $t_i$'s is common to every vertex gadget in $G'$, assuming that $G$ is connected. Thus $G'$ is acyclically 4-colorable if and only if $G$ is 3-colorable. □

## 4.3 Acyclic list colorings

**Theorem 4.6.** *It is NP-complete to decide whether a planar graph $G$, given with an acyclic 4-coloring and 4-list assignment $L$, is $L$-colorable.*

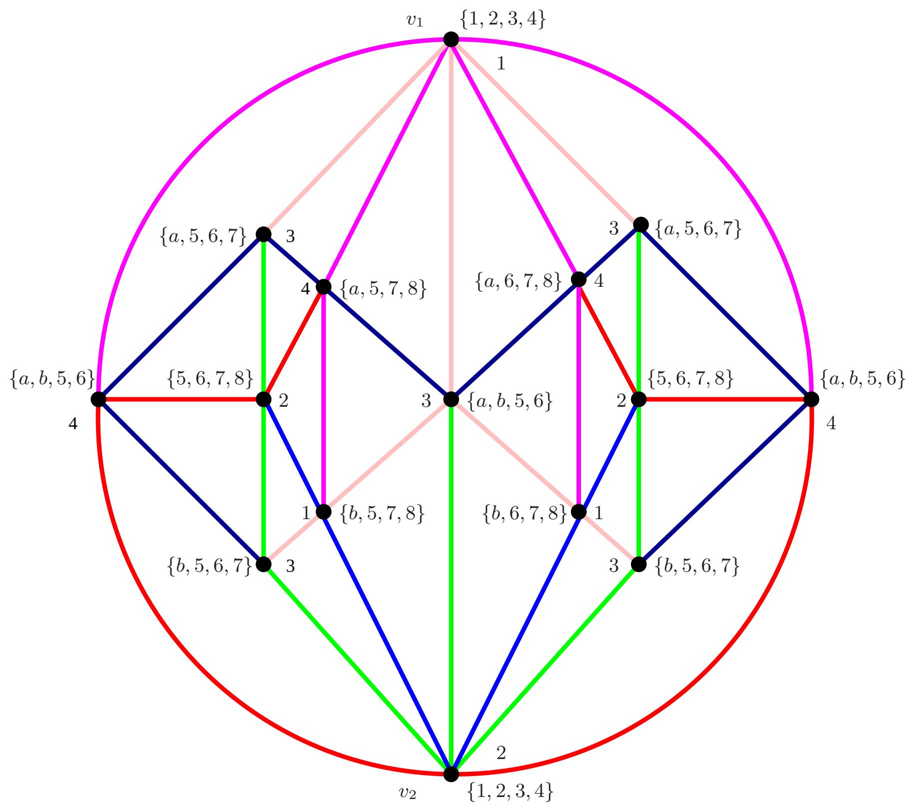


Figure 4.10: The forcing gadget $F_{a,b}$: 4-list assignment and acyclic 4-coloring.

*Proof.* We use again a reduction from $(\mathcal{P}_3 \cap \mathcal{S}_4 : \mathcal{S}_0^3)$. The proof for the restriction to graphs with maximum degree 4 [26] shows that we can assume that the input graph $G$ is acyclically 4-colorable and that an acyclic 4-coloring can be constructed in polynomial time.

Let $p_1$ and $p_2$ be adjacent vertices with lists $L_{p_1} = L_{p_2} = \{1, 2, 3, 4\}$. We take three copies of the graph $F_{a,b}$ depicted in Figure 4.10: $F_{4,1}$, $F_{4,2}$, $F_{4,3}$. We identify all the vertices $v_1$ (resp. $v_2$) to the vertex $p_1$ (resp. $p_2$) to obtain the forcing gadget. We check that $F_{a,b}$ is not colorable if $c(v_1) = a$ and $c(v_2) = b$. So, the forcing gadget is colorable if and only if $c(v_1) \neq 4$. Given an input graph $G$ for $(\mathcal{P}_3 \cap \mathcal{S}_4 : \mathcal{S}_0^3)$, we assign to every vertex $v$ in $G$ the list $\{1, 2, 3, 4\}$ and identify $v$ to the vertex $p_1$ of copy of the forcing gadget. We thus obtain a graph $G'$ with an assignment $L$ of lists of size four. Notice that any 3-coloring of $G$ extends to an $L$-coloring of $G'$ and conversely any $L$-coloring of $G'$ induces a 3-coloring of $G$. □

## 4.4 $T$-free arboricities

**Theorem 4.7.** *The following problems are NP-complete:*

1. *Deciding whether a graph $G \in \mathcal{P}_4 \cap \mathcal{D}_2 \cap bip$ satisfies $P_4$-$fa(G) \leqslant 3$.*
2. *For every $n \geqslant 2$, deciding whether a graph $G \in \mathcal{P}_4 \cap \mathcal{D}_2 \cap bip$ satisfies $S_n$-$fa(G) \leqslant 3$.*
3. *For every $n \geqslant 3$, deciding whether a graph $G \in \mathcal{P}_6 \cap bip$ satisfies $S_n$-$fa(G) \leqslant 2$.*
4. *Deciding whether a graph $G \in \mathcal{P}_8 \cap bip$ satisfies $S_2$-$fa(G) \leqslant 2$.*
5. *For every $n$, deciding whether a graph $G \in \mathcal{P}_n \cap \mathcal{S}_3 \cap bip$ satisfies $P_4$-$fa(G) \leqslant 2$.*
6. *Deciding whether a graph $G \in \mathcal{P}_3$ satisfies $P_4$-$fa(G) \leqslant 4$.*

Let us denote by $L(G)$ the line graph of $G$ and by $\mathcal{L}$ the class of line graphs of "planar bipartite graphs with maximum degree three and girth at least six". Notice that graphs in $\mathcal{L}$ are planar with maximum degree four and line graphs of bipartite graphs, thus perfect [38]. Alternatively, we can say that they are planar $(K_{1,3}, K_4, K_4^-, C_4,$odd-hole)-free (this implies the maximum degree four). As we already mentioned, if $G$ is triangle-free, then $P_4$-$fa(G) = \chi'_{sub}(G) = \chi_{sub}(L(G))$. Thus, EDGE SUBCOLORABILITY is NP-complete on planar graphs. This answers an open question of Fiala and Le [23].

**Corollary 4.1.** *Deciding whether a graph $G \in \mathcal{L}$ satisfies $\chi_{sub}(G) \leqslant 2$ is NP-complete.*

Shermer proved that recognizing graphs with caterpillar arboricity two is NP-complete [73], but their reduction produces non-planar graphs that contain copies of $K_4$. Theorem 4.7.(3) shows in particular that this result still holds if the input graph is restricted to the class $G \in \mathcal{P}_6 \cap bip$.

Theorem 4.7.(2) implies the existence of bipartite planar graphs with caterpillar arboricity four and, as we already mentioned, the track number of a triangle-free graph equals its caterpillar arboricity (i.e. $S_3$-$fa$). This answers an open question of Gyárfás and West [33].

**Corollary 4.2.** *There exist bipartite planar graphs with track number four.*

Theorems 4.7.(1), 4.7.(2), and 4.7.(6) are each obtained by a polynomial reduction from $(\mathcal{P}_3 \cap \mathcal{S}_4 : \mathcal{S}_0^3)$. Theorems 4.7.(3), 4.7.(4), and 4.7.(5) are each obtained by a polynomial reduction from $(\mathcal{P}_4 \cap \mathcal{S}_4 : \mathcal{S}_1^2)$. Let us now describe the reductions for Theorems 4.7.(1) and 4.7.(2) (resp. Theorems 4.7.(3) and 4.7.(4)). Given a planar graph (resp. a triangle-free planar graph) $G$, we construct a graph $G'$ that belongs to the class specified in the theorem as follows: we add a "vertex gadget" to every vertex $v$ of $G$ and replace every edge $uv$ of $G$ by an "edge gadget". The vertex gadget forces the vertex $v$ to be an inner vertex in at most one forest $F_i$ for any $k$-$T$-$fa$ coloring (with $k$ and $T$ as mentioned in the theorem). The edge gadget is such that $G'$ is $k$-$T$-$fa$ colorable if and only if $G$ is 3-colorable (resp. 2-subcolorable). More precisely if $G$ has a vertex coloring $c$, then $G'$ has $k$-$T$-$fa$ coloring such that every original vertex $v$ is an inner vertex in $F_{c(v)}$, and conversely, if $G'$ has a $k$-$T$-$fa$ coloring such that every original vertex $v$ is an inner vertex of $F_i$, then taking $c(v) = i$ gives a vertex coloring of $G$.

### 4.4.1 Proof of Theorem 4.7.(1)

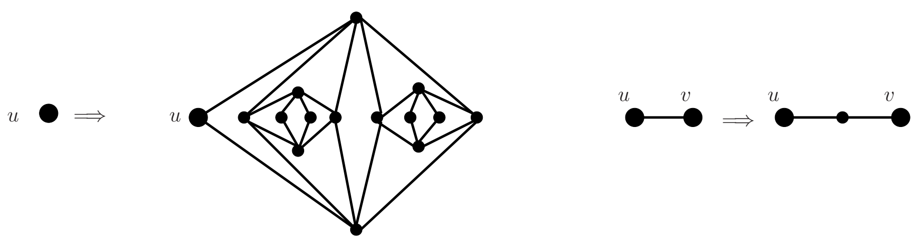


Figure 4.11: The vertex gadget and the edge gadget for the reduction of Theorem 4.7.(1).

Given a planar graph $G$, we construct $G'$ by adding to every vertex of $G$ the vertex gadget depicted in Figure 4.11 and by subdividing every edge of $G$.

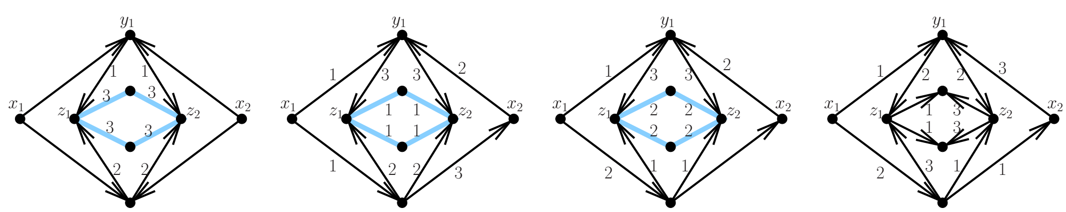

Figure 4.12: 3-$P_4$-$fa$ coloring with conditions on $x_1$ and $x_2$.

Given a planar graph $G$, we construct $G'$ by adding to every vertex of $G$ the vertex gadget depicted in Figure 4.11 and by subdividing every edge of $G$. First, we comment on how to 3-$P_4$-$fa$ color the graph depicted in Figure 4.12. Notice that if a vertex has two incoming edges colored 1 and 2, all its remaining incident edges have to be colored 3. In the first drawing, we impose that all the edges $x_iy_j$ are oriented toward $y_j$. This implies that all the edges $y_iz_j$ are oriented toward $z_j$, and that we just used two colors for these edges. This finally implies that all the remaining edges incident to the $z_i$'s have the same color, which is not allowed. In the second drawing, we impose that only one edge $x_iy_j$ is oriented toward $x_2$ and that the edges incident to $x_1$ have the same color, 1. The edges $x_2y_1$ and $y_2x_2$ have to be respectively colored 2 and 3. This implies that the edges $y_1z_i$ are oriented toward $z_i$ and colored 3. This implies that the edges $y_2z_i$ are oriented toward $z_i$ and colored 2. This finally implies that all the remaining edges incident to the $z_i$'s have the same color, which is not allowed. In the third drawing, we impose that just one edge $x_iy_j$ is oriented toward $x_2$, that the edges incident to $x_1$ have distinct colors, 1 and 2, and that the edges $x_1y_2$ and $x_2y_1$ have the same color, 1. This implies that the edges $y_1z_i$ are oriented toward $z_i$ and colored 3. This implies that the edges $y_2z_i$ are oriented toward $z_i$ and colored 1. This finally implies that all the remaining edges incident to the $z_i$'s have the same color, which is not allowed. In the last drawing, we see a 3-$P_4$-$fa$ coloring of the graph.

This implies that there is not much flexibility for coloring the vertex gadget in Figure 4.11. Actually, in any 3-$P_4$-$fa$ coloring of the vertex gadget, the two edges incident to $u$ have to be oriented toward $u$. So $u$ is an inner vertex in at most one forest, say $F_1$. In any 3-$P_4$-$fa$ coloring of $G'$, an edge incident to $u$ that belongs to an edge gadget must be colored 1 and must be oriented toward the subdivision vertex. Thus the edge gadget forces the vertices $u$ and $v$ to be inner vertices in distinct forests.

Assume that $G$ has a 3-coloring $c$, then for any vertex $u \in V_G$ we color its vertex gadget in $G'$ so that $u$ is an inner vertex in $F_{c(u)}$ and we extend this 3-$P_4$-$fa$ coloring to $G'$. Conversely, suppose $G'$ has a 3-$P_4$-$fa$ coloring, we color the vertices of $G$ accordingly to the forest for which they are an inner vertex in the 3-$P_4$-$fa$ coloring of $G'$, and then we obtain a 3-coloring of $G$.

### 4.4.2 Proof of Theorem 4.7.(2)

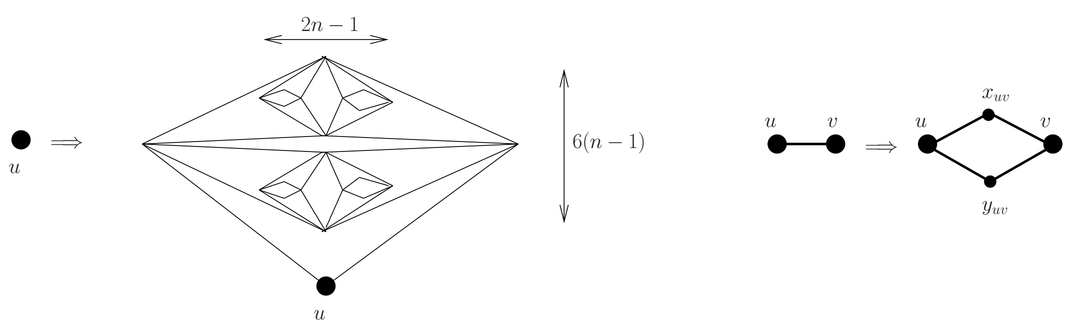


Figure 4.13: The vertex gadget and the edge gadget for the reduction of Theorem 4.7.(2).

Given a planar graph $G$, we construct $G'$ by adding to every vertex of $G$ the vertex gadget depicted in Figure 4.13 and by replacing every edge $uv$ of $G$ by a cycle $(ux_{uv}vy_{uv})$. Using similar arguments as in the previous proof, we show that in any 3-$S_n$-$fa$ coloring of the vertex gadget, the vertex $u$ is an inner vertex in exactly one forest, say $F_1$. This implies that an edge incident to $u$ that belongs to an edge gadget must get color 1. Consider an edge $uv \in E_G$. Since in a 3-$S_n$-$fa$ coloring of $G'$, the edges $ux_{uv}$ and $uy_{uv}$ (resp. $vx_{uv}$ and $vy_{uv}$) have the same color, the vertices $u$ and $v$ have to be inner vertices in distinct forests in order to avoid a monochromatic 4-cycle $(ux_{uv}vy_{uv})$. Assume that $G$ has a 3-coloring $c$, then for any vertex $u \in V_G$ we color its vertex gadget in $G'$ so that $u$ is an inner vertex in $F_{c(u)}$ and then extend this 3-$P_4$-$fa$ coloring to $G'$. Conversely, suppose $G'$ has a 3-$S_n$-$fa$ coloring, we color the vertices of $G$ according to the forest in which they are inner vertices in the 3-$S_n$-$fa$ coloring of $G'$. Thus we obtain a 3-coloring of $G$.

### 4.4.3 Proof of Theorem 4.7.(3)

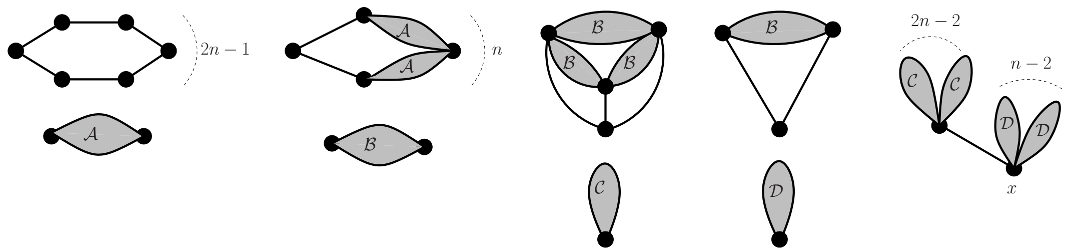


Figure 4.14: The vertex gadget for the reduction for the reduction of Theorem 4.7.(3).

Given a triangle-free planar graph $G$, we construct $G'$ by adding to every vertex of $G$ the vertex gadget depicted in Figure 4.14 and by subdividing every edge of $G$. A 2-$S_n$-$fa$ coloring of the vertex gadget forces an original vertex of $G$ to be an inner vertex in at most one forest, say $F_1$, and to be incident to at least $n-2$ unoriented edges of $F_1$. We consider now 2-$S_n$-$fa$ colorings of the edge gadget of an edge $uv$ of $G$. If $u$ and $v$ are inner vertices in distinct forests, then we can 2-$S_n$-$fa$ color the edges of the edge gadget and orient them toward the subdivision vertex. If $u$ and $v$ are inner vertices in the forest $F_1$, then both edges of the edge gadget have to be unoriented edges colored 1. Thus $u$ and $v$ are now incident to $n-1$ unoriented edges of $F_1$. This shows that $G$ has a 2-subcoloring if and only if $G'$ has a 2-$S_n$-$fa$ coloring.

### 4.4.4 Proof of Theorem 4.7.(4)

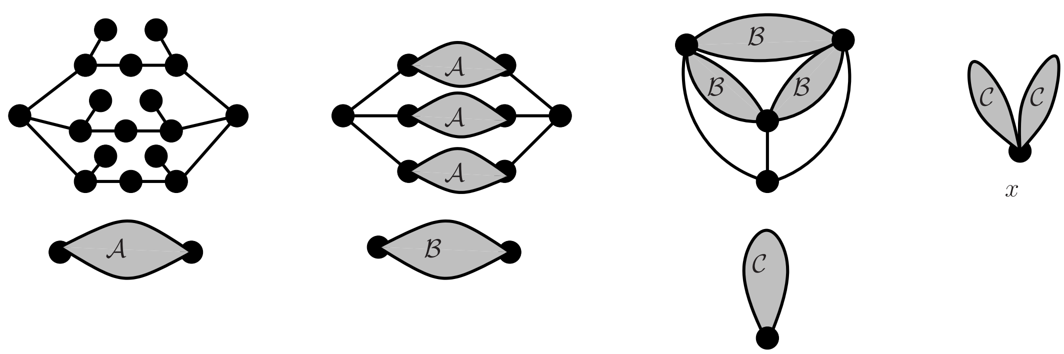


Figure 4.15: The vertex gadget for the reduction for the reduction of Theorem 4.7.(4).

Given a triangle-free planar graph $G$, we construct $G'$ by adding to every vertex of $G$ a copy of the vertex gadget depicted in Figure 4.15 and by replacing every edge $uv$ of $G$ by a copy of the graph depicted in Figure 4.16, left. A 2-$S_2$-$fa$ coloring of the vertex gadget forces

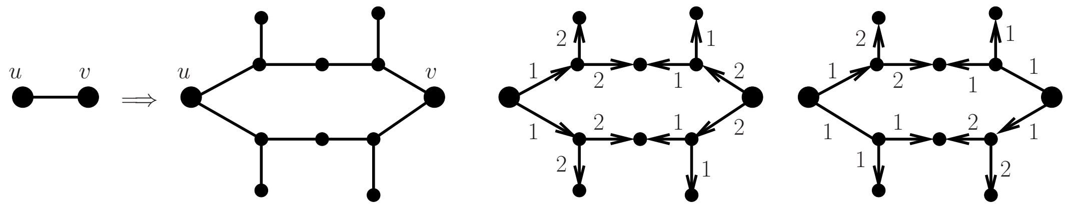


Figure 4.16: The edge gadget for the reduction for the reduction of Theorem 4.7.(4).

that the original vertex of $G$ is an inner vertex in at most one forest, say $F_1$, but may not be incident to an unoriented edge of $F_1$. We consider now 2-$S_2$-$fa$ colorings of the edge gadget of an edge $uv$ of $G$. If $u$ and $v$ are inner vertices in distinct forests, then we can 2-$S_2$-$fa$ color the edge gadget such that neither $u$ nor $v$ is incident to an unoriented edge of the edge gadget (see Figure 4.16, middle). If $u$ and $v$ are inner vertices in the forest $F_1$, then in any 2-$S_2$-$fa$ coloring of the edge gadget, both $u$ and $v$ are incident to an unoriented edge colored 1 (see Figure 4.16, right). This shows that $G$ has a 2-subcoloring if and only if $G'$ has a 2-$S_2$-$fa$ coloring.

### 4.4.5 Proof of Theorem 4.7.(5)

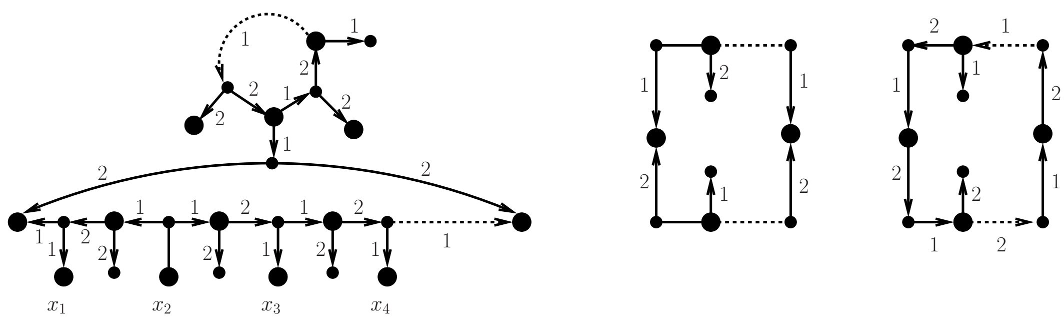


Figure 4.17: The vertex gadget and the edge gadget for the reduction of Theorem 4.7.(5).

Given a triangle-free planar graph $G$, we construct $G'$ by replacing every vertex of $G$ by the vertex gadget depicted in Figure 4.17, left. For every edge $uv$ of $G$, we identify two vertices $x_i$ of the gadget of $u$ with two vertices $x_i$ of the gadget of $v$ as in Figure 4.17, middle. In any 2-$P_4$-$fa$ coloring of the vertex gadget, all edges incident to a vertex $x_i$ have the same color and are oriented toward $x_i$, except at most one *special* edge (the one incident to $x_2$ in Figure 4.17). This common color in the vertex gadget of a vertex $u$ corresponds to the color of $u$ in a 2-subcoloring of $G$. We consider now 2-$P_4$-$fa$ colorings of the edge gadget of an edge $uv$ of $G$. If the gadgets of $u$ and $v$ have the distinct common colors, then we can 2-$P_4$-$fa$ color the edge gadget without using any special edge (see Figure 4.17, middle). If the gadgets of $u$ and $v$ have the same common color 1, then in any 2-$P_4$-$fa$ coloring of the edge gadget uses the special edge of both vertex gadget (see Figure 4.17, right). This shows that $G$ has a 2-subcoloring if and only if $G'$ has a 2-$P_4$-$fa$ coloring. We easily check that the resulting graph $G'$ is planar, bipartite, with maximum degree 3, and may have arbitrary girth (see the dotted lines in Figure 4.17).

### 4.4.6 Proof of Theorem 4.7.(6)

The proof is similar to the reduction $(\mathcal{P}_3 \cap \mathcal{S}_4 : \mathcal{S}_0^3) \propto (\mathcal{P}_4 \cap \mathcal{S}_8 \cap bip \cap \mathcal{D}_2 : \mathcal{S}_0^{(4)})$ (Theorem 4.5).

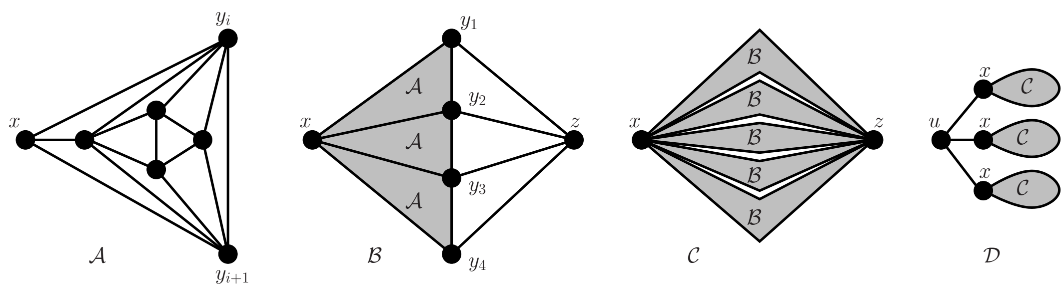


Figure 4.18: The forcing gadgets for the reduction of Theorem 4.7.(6).

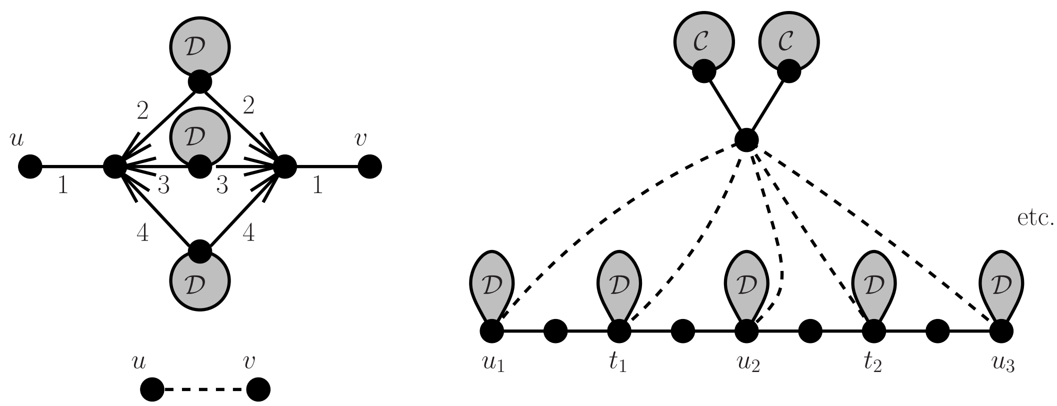


Figure 4.19: The vertex gadget for the reduction of Theorem 4.7.(6).

Consider the forcing gadgets depicted in Figure 4.18. For any 4-$P_4$-$fa$ coloring of $\mathcal{C}$, the vertex $x$ is incident to edges of all four forests and $x$ is a center in at least one forest. Thus in any 4-$P_4$-$fa$ coloring of $\mathcal{D}$, the vertex $u$ is a center in at most one forest. The gadget represented by a dotted edge in Figure 4.19 forces that the vertices $u$ and $v$ are centers in a common forest. The vertex gadget is such that the vertex $x$ is a center in at most two forests, the $u_i$'s and the $t_i$'s are centers in at most one forest. The 2-path between $u_i$ and $t_i$ (resp. $t_i$ and $u_{i+1}$) implies that $u_i$ and $t_i$ (resp. $t_i$ and $u_{i+1}$) must be centers in distinct forests. Now, because of the dotted edges, all the $u_i$'s are centers in a same forest and all the $t_i$'s are centers in a same forest distinct from that of the $u_i$'s.

Given a planar graph $G$, we construct the graph $G'$ as follows. We replace every vertex of $G$ by a copy of the vertex gadget and for every edge $vw$ of $G$ we connect a vertex $u_i$ (resp.

$t_i$) in the gadget of $v$ to a vertex $u_i$ (resp. $t_i$) in the gadget of $w$ using a copy of the gadget represented by a dotted edge (resp. a 2-path). In any 4-$P_4$-$fa$ coloring of $G'$, the color of the $t_i$'s is common to every vertex gadget in $G'$, assuming that $G$ is connected. Moreover if $vw \in E(G)$, then the $u_i$'s in the gadget of $v$ and the $u_i$'s in the gadget of $w$ cannot be centers in the same forest. Thus $G'$ has a 4-$P_4$-$fa$ if and only if $G$ is 3-colorable.

# Chapter 5

# Conclusion

We first studied acyclic improper colorings, oriented colorings, and their relationship. In particular, we tried to see when a result about acyclic improper colorings could lower the known upper bound on the oriented chromatic number of a graph class. This situation occurs for planar graphs, where the implication $\chi_a(\mathcal{P}_3) \leqslant 5 \Longrightarrow \chi_o(\mathcal{P}_3) \leqslant 80$ provides the best known upper bound on $\chi_o(\mathcal{P}_3)$. A first negative result is that the inequality $\chi_o(\mathcal{C}_0 \odot \cdots \odot \mathcal{C}_{k-1}) \leqslant 2^{k-1} \sum_{i=0}^{i<k} \chi_o(\mathcal{C}_i)$ is tight for $k \geqslant 3$. A second negative result states that if every planar graph has an acyclic improper coloring with at most four colors, then the oriented chromatic number of a color class distinct from $\mathcal{S}_0$ is at least 15.

We were able to improve the upper bound on the oriented chromatic number of two significant graph classes, but the proof do not use any acyclic improper coloring. Precisely, we obtained that the strong oriented chromatic number of triangle-free planar graphs (resp. triangle-free 2-outerplanar graphs) is at most 59 (resp. 27) via an homomorphism to the Paley tournament $QR_{59}$ (resp. $QR_{27}$). The proofs required to formalize the notion of $S_{k,n}$ property of an oriented graph and to check these properties by computer in the case of Paley tournaments. Notice that Marshall used a generalization of $S_{k,n}$ properties in his impressive proof of $\chi_s(\mathcal{P}_3) \leqslant 271$.

There remain interesting cases where acyclic improper colorings could be useful: A proof that $\mathcal{P}_5 \subseteq \mathcal{S}_0^{(3)}$ or $\mathcal{P}_5 \subseteq \mathcal{D}_1^{(2)}$ would lower the known upper bound on $\chi_o(\mathcal{P}_5)$ from 19 to 12. Similarly, a proof that $\mathcal{P}_6 \subseteq \mathcal{S}_0 \odot \mathcal{D}_1$ would lower the known upper bound on $\chi_o(\mathcal{P}_6)$ from 11 to 8.

A proof that $\mathcal{P}_5 \subseteq \mathcal{S}_0^{(3)}$ would also have important implications in edge coloring: $\chi_a(\mathcal{P}_5) = 3$ implies $P_4\text{-}fa(\mathcal{P}_5) = 3$. Our $T$-free arboricities not only relate to acyclic colorings: we were able to obtain (negative) results concerning other parameters, namely the track number and the subchromatic number, via NP-completeness proofs for some $T\text{-}fa$ colorings.

We also studied the list version of acyclic colorings and obtained somewhat surprising results. There exist graphs $G$ with acyclic chromatic number and list chromatic number at most 3, and arbitrarily large acyclic list chromatic number. For planar graphs, on the other hand, $\chi_a(\mathcal{P}_3) = \chi^l(\mathcal{P}_3) = 5$ and $\chi_a^l(\mathcal{P}_3) = 5$ is conjectured by Borodin et al. [9].

# Second part

# Combinatorics on words

# Chapter 6

# Preliminaries

## 6.1 Repetitions, patterns, and letter frequencies

In this part, we consider words over a finite alphabet.

Let $\Sigma_i$ denote the $i$-letter alphabet $\{0, 1, \ldots, i-1\}$ and let $\varepsilon$ denote the empty word.

A *morphism* is an application $h : \Sigma_s^* \longrightarrow \Sigma_e^*$ such that $h(xy) = h(x)h(y)$. So a morphism $h : \Sigma_s^* \longrightarrow \Sigma_e^*$ satisfies $h(\varepsilon) = \varepsilon$ and is completely defined by the couples $(a, h(a))$ for $a \in \Sigma_s$. For $q \geqslant 2$, a morphism $h : \Sigma_s^* \longrightarrow \Sigma_e^*$ is said $q$-uniform if $|h(a)| = q$ for every $a \in \Sigma_s$. We say that a word $w \in \Sigma_s^*$ avoids a set $F \subset \Sigma_s^*$ of words if $w$ does not contain a word in $F$ as a factor, i.e. $w$ cannot be written as $w = lsr$ with $l, r \in \Sigma_s^*,\ s \in F$

A *square* is a repetition of the form $xx$, where $x$ is a nonempty word; an example in English is `hotshots`. It is easy to see that every word of length at least 4 over $\Sigma_2$ must contain a square, so squares cannot be avoided in infinite binary words. However, Thue showed [6, 75, 76] that there exist infinite words over $\Sigma_3$ that avoid squares.

Squares can be seen as repetitions of exponent 2. For $\alpha > 1$ a rational number, we say that $y$ is an $\alpha$-repetition if we can write $y = x^n x'$ with $x'$ a prefix of $x$ and $|y| = \alpha|x|$. For example, the French word `entente` is a $\frac{7}{3}$-repetition and the English word `tormentor` is a $\frac{3}{2}$-repetition. An $\alpha^+$-repetition is a $\alpha'$-repetition for some $\alpha' > \alpha$. Brandenburg [12] and (implicitly) Dejean [19] considered the problem of determining the *repetition threshold*; that is, the smallest exponent $R_k$ such that there exist an infinite word over $\Sigma_k$ that avoids $R_k^+$-repetitions. Dejean proved that $R_3 = \frac{7}{4}$. She also conjectured that $R_4 = \frac{7}{5}$ and $R_k = \frac{k}{k-1}$ for $k \geqslant 5$. In its full generality, this conjecture is still open, although Pansiot [61] proved that $R_4 = \frac{7}{5}$ and Moulin-Ollagnier [52] proved that Dejean's conjecture holds for $5 \leqslant k \leqslant 11$. For more information, see [14].

Instead of avoiding *all* squares, one interesting variation is to avoid *all sufficiently large* squares. Let us define the length of a square $x^2$ as $|x|$. Entringer, Jackson, and Schatz [21] showed that there exist infinite binary words avoiding all squares of length at least three. For some other papers about avoiding sufficiently large squares, see [20, 24, 62, 63, 68].

In Chapter 7, we study a generalization of the repetition threshold of Dejean which handles avoidance of all sufficiently large repetitions. Pansiot suggested such a generalization at the end of his paper [61].

Let $\alpha > 1$ be a rational number, and let $\ell \geqslant 1$ be an integer. A word $w$ is a $(\alpha, \ell)$-*repetition* if we can write it as $w = x^n x'$ where $x'$ is a prefix of $x$, $|x| = \ell$, and $|w| = \alpha|x|$. Notice that an $\alpha$-repetition is an $(\alpha, \ell)$-repetition for some $\ell$. We say a word is $(\alpha, \ell)$-*free* if it contains no factor that is a $(\alpha', \ell')$-repetition for $\alpha' \geqslant \alpha$ and $\ell' \geqslant \ell$. We say a word is $(\alpha^+, \ell)$-*free* if it is $(\alpha', \ell)$-free for all $\alpha' > \alpha$.

For integers $k \geqslant 2$ and $\ell \geqslant 1$, we define the *generalized repetition threshold* $R(k, \ell)$ as the real number $\alpha$ such that, over $\Sigma_k$, there exists

(a) either an $(\alpha^+, \ell)$-free infinite word, but all $(\alpha, \ell)$-free words are finite;

(b) or an $(\alpha, \ell)$-free infinite word, but for all $\epsilon > 0$, all $(\alpha - \epsilon, \ell)$-free words are finite.

Notice that $R(k, 1) = R_k$ corresponds to the repetition threshold.

**Theorem 6.1.** *The generalized repetition threshold $R(k, \ell)$ exists and is finite for all integers $k \geqslant 2$ and $\ell \geqslant 1$. Furthermore, $1 + \ell/k^\ell \leqslant R(k, \ell) \leqslant 2$.*

*Proof.* Define $S$ to be the set of all real numbers $\alpha \geqslant 1$ such that there exists an $(\alpha, \ell)$-free infinite word over $\Sigma_k$. Since Thue proved that there exists an infinite word over a two-letter alphabet (and hence over larger alphabets) avoiding $(2^+)$-repetitions, we have that $\beta = \inf S$ exists and $\beta \leqslant 2$. If $\beta \in S$, we are in case (b) above, and if $\beta \notin S$, we are in case (a). Thus $R(k, \ell) = \beta$.

For the lower bound, note that any word of length at least $k^\ell + \ell$ contains at least $k^\ell + 1$ factors of length $\ell$. Since there are only $k^\ell$ distinct factors of length $\ell$, such a word contains at least two occurrences of some word of length $\ell$, and hence is not $(1 + \frac{\ell}{k^\ell}, \ell)$-free. $\square$

It may be worth noting that we do not know any instance where case (b) of the definition of generalized repetition threshold above actually occurs, but we have not been able to rule it out.

Squares can be seen as occurrences of the pattern $AA$. A pattern $P$ is a finite word over the alphabet $\{A, B, \dots\}$. An occurrence of $P$ is the image of $P$ by a non-erasing morphism $\phi : \{A, B, \dots\}^* \longrightarrow \Sigma_e^*$. Non-erasing means that for all $a \in \{A, B, \dots\}$, $\phi(a) \neq \varepsilon$. A word $w \in \Sigma_e^*$ avoids a pattern $P$ if avoids the set of occurrences of $P$. A pattern $P$ is $k$-avoidable if there exists an infinite word $w \in \Sigma_k^*$ that avoids $P$. The avoidability index $\mu(P)$ of a pattern $P$ is the smallest integer $k$ such that $P$ is $k$-avoidable.

The avoidability index of some ternary patterns (i.e. patterns over $\{A, B, C\}$) was left open in Cassaigne's thesis. In particular, $ABCBABC$ was the only avoidable ternary pattern not known to be 3-avoidable. We obtain that $\mu(ABCBABC) = 2$, which gives the following result.

**Theorem 6.2.** *Every ternary pattern is either unavoidable or 3-avoidable.*

In Chapter 8, we complete the determination of the avoidability index of all ternary patterns, i.e. patterns over $\{A, B, C\}$.

Given a factorial language $L$ defined by an alphabet size and a set of forbidden repetitions, we denote by $f_{\min}$ (resp. $f_{\max}$) the minimal (resp. maximal) letter frequency in an infinite word that belong to $L$. Letter frequencies have been studied for binary repetition-free words [40] and ternary square-free words [65, 72]. We consider here the frequency of the letter 0. Let $|w|_0$ denote the number of occurrences of 0 in the finite word $w$. So the letter frequency in $w$ is $\frac{|w|_0}{|w|}$. In Chapter 9, we improve the estimation of the minimal letter frequencies given in [40, 72]. We also provide new conjectures and results about extremal letter frequencies for other languages.

## 6.2 A method to produce morphisms

This section is devoted to the method [58] we used to find all the positive results in this part.

A $q$-uniform morphism $h : \Sigma_s^* \to \Sigma_e^*$ is *synchronizing* if for any $a, b, c \in \Sigma_s$ and $v, w \in \Sigma_e^*$, if $h(ab) = vh(c)w$, then either $v = \varepsilon$ and $a = c$ or $w = \varepsilon$ and $b = c$.

**Lemma 6.1.** *Let $\alpha, \beta \in \mathbb{Q}$, $1 < \alpha < \beta < 2$ and $n \in \mathbb{N}^*$. Let $h : \Sigma_s^* \to \Sigma_e^*$ be a synchronizing $q$-uniform morphism (with $q \geqslant 1$). If $h(w)$ is $(\beta^+, n)$-free for every $\alpha^+$-free word $w$ such that $|w| < \max\left(\frac{2\beta}{\beta-\alpha}, \frac{2(q-1)(2\beta-1)}{q(\beta-1)}\right)$, then $h(t)$ is $(\beta^+, n)$-free for every (finite or infinite) $\alpha^+$-free word $t$.*

*Proof.* Suppose $w$ is an $\alpha^+$-free word such that $h(w)$ is not $(\beta^+, n)$-free and $w$ is of minimum length with this property. Thus $h(w)$ contains a $\beta^+$-repetition, that is, a factor $uvu$ such that $\frac{|uvu|}{|uv|} > \beta$. Denote $x = |u|$ and $y = |v|$. Since $\frac{|uvu|}{|uv|} = \frac{2x+y}{x+y} > \beta$, we have $y < \frac{2-\beta}{\beta-1}x$. If $x \geqslant 2q-1$, then each occurrence of $u$ contains at least one full $h$-image of a letter. As $h$ is synchronizing, the two occurrences of $u$ in $uvu$ contain the same $h$-images and in the same positions. Let $U$ be the factor of $w$ that contains all letters whose $h$-images are contained in $u$, and let $V$ be the factor of $w$ that contains all letters whose $h$-images intersect $v$. Denoting $X = |U|$ and $Y = |V|$, we have $Yq < y + 2q$ and $Xq > x - 2q$, or equivalently $x < (X+2)q$. Since $UVU$ is a factor of $w$ and $w$ is $\alpha^+$-free, then $\frac{2X+Y}{X+Y} \leqslant \alpha$, which gives $X \leqslant \frac{\alpha-1}{2-\alpha}Y$. Now we have

$$Yq < y + 2q < \frac{2-\beta}{\beta-1}x + 2q < \frac{2-\beta}{\beta-1}(X+2)q + 2q \leqslant \frac{2-\beta}{\beta-1}\left(\frac{\alpha-1}{2-\alpha}Y + 2\right)q + 2q,$$

implying that $Y < \frac{2(2-\alpha)}{\beta-\alpha}$. By the minimality of $w$ we get

$$|w| \leqslant 2 + Y + 2X \leqslant 2 + Y\left(1 + 2\frac{\alpha-1}{2-\alpha}\right) < 2 + \frac{2(2-\alpha)}{\beta-\alpha}\frac{\alpha}{2-\alpha} = \frac{2\beta}{\beta-\alpha}.$$

If $x \leqslant 2q-2$, then $y < \frac{2-\beta}{\beta-1}(2q-2)$ and thus $2x + y < \frac{2\beta}{\beta-1}(q-1)$. The minimality of $w$ implies that $(|w|-2)q \leqslant |uvu| - 2 = 2x + y - 2$. By the above we get that $|w| < \frac{2(q-1)(2\beta-1)}{q(\beta-1)}$, which completes the proof. □

For convenience, let us denote the maximum in Lemma 6.1 by $\max_{\alpha,\beta,q}$. Using Lemma 6.1 with $R_s < \alpha$) allows us to check that the $h$-image of any (infinite) $\alpha^+$-free $s$-ary word is $(\beta^+, n)$-free with a finite amount of computation.

We want to find a $q$-uniform morphism $h : \Sigma_s^* \to \Sigma_e^*$ such that for every $\alpha^+$-free word $t \in \Sigma_s^*$, $h(t)$ satisfies a property $\mathcal{P}$, where $\mathcal{P}$ consists in some $(\beta^+, n)$-freeness properties and the avoidance of a finite set $S$ of words. We fix $\alpha, \beta \in \mathbb{R}$, $1 < \alpha < \beta < 2$ and $e, n, q \in \mathbb{N}$. We use depth-first search to find an $e$-ary word $w$ of size $s \times q$ which defines the morphism $h$ by posing $w = h(0)h(1)\dots h(s-1)$. Obviously, we can restrict the search to words satisfying $h(s-1) \prec \cdots \prec h(1) \prec h(0)$, where $\prec$ is the lexicographic order of $\Sigma_e^q$. We prune the search tree by checking property $\mathcal{P}$ on the prefixes of a potential $w$. If no morphism is found, we increase the value of $q$ and try again.

## 6.3 Exponential growth

Let $t(n)$ be the number of words of length $n$ satisfying a property $\mathcal{P}$. We say that there exist polynomially (resp. exponentially) many words satisfying $\mathcal{P}$ if there exists a constant $c > 1$ such that $t(n) \leqslant n^c + c$ (resp. $t(n) \geqslant c^n$) for every $n$.

The repetition threshold for binary words is 2, and this result is tight in the following senses:

1. There are only polynomially many $2^+$-free binary words.
2. There exist arbitrarily large squares in any large enough $2^+$-free binary word.

In this section we show that no similar situation occurs for ternary and 4-ary words. We use the following easy lemma, which is already implicitely used in [39].

**Lemma 6.2.** *There are at least $2^{\lceil \frac{n}{k} \rceil}$ $R_k^+$-free words of length $n$ over $\Sigma_{k+1}$.*

*Proof.* Consider an $R_k^+$-free word $w$ in $\Sigma_k^n$. At least one letter in $\Sigma_k$, say 0, occurs at least $\lceil \frac{n}{k} \rceil$ times in $w$. The letter $k$ belongs to $\Sigma_{k+1}$ but does not belong to $\Sigma_k$. Notice that replacing zero or more occurrences of 0 by an occurrence of $k$ in $w$ produces an $R_k^+$-free word of length $n$ over $\Sigma_{k+1}$, and that we can obtain at least $2^{\lceil \frac{n}{k} \rceil}$ such words. □

**Theorem 6.3.**

1. *There exist exponentially many $\frac{7}{4}^+$-free ternary words with no large $\frac{7}{4}$-repetition.*
2. *There exist exponentially many $\frac{7}{5}^+$-free 4-ary words with no large $\frac{7}{5}$-repetition.*

*Proof.* By Lemma 6.2, there exist exponentially many $\frac{7}{5}^+$-free words over $\Sigma_5$ and exponentially many $\frac{5}{4}^+$-free words over $\Sigma_6$.

The following 59-uniform morphism $h$ is such that for any $\frac{7}{5}^+$-free word $t \in \Sigma_5^*$, $h(t) \in \Sigma_3^*$ is $\frac{7}{4}^+$-free and $\left(\frac{3}{2}^+, 10\right)$-free.

$$
\begin{aligned}
0 &\mapsto 01020121021201210120102101202102012102120210120102101210212\\
1 &\mapsto 01020120210201210212012101201021202102012021201210201021012\\
2 &\mapsto 01020120210201210120102101202120121021202102012021201021012\\
3 &\mapsto 01020120210201021201210201202101201021012102120210201021012\\
4 &\mapsto 01020120210121020102120121012021020120212012102010212021012
\end{aligned}
$$

The following 132-uniform morphism $h$ is such that for any $\frac{5}{4}^+$-free word $t \in \Sigma_6^*$, $h(t) \in \Sigma_4^*$ is $\frac{7}{5}^+$-free and $\left(\frac{61}{44}^+, 11\right)$-free.

$0 \mapsto$ 0120310213201230210320130231203213012310213203123013210312013021
032012310213012032102312013210312301320310213012310320130231203210 23

$1 \mapsto$ 0120310213201230210320130231032130123102132031230132102312013021
0320123021320310230120321023120132103123021320123103213023120132102 3

$2 \mapsto$ 0120310213201230210320130231032130120310230132103120130231032012
3021320310230120321023120132103123013203102130123103213023120132102 3

$3 \mapsto$ 0120310213201230210312013210231203213023103201230213203102301321
0312013021032012310213203123013210231203213023103201302103120132102 3

$4 \mapsto$ 0120310213201230210312013023103213012310213203123013210312013021
0320123021320310230120321302312013210312302132012310321302312032102 3

$5 \mapsto$ 0120310213201230210312013023103213012031023013210312013021032012
3021320310230120321023120130210312301320310213012310321302312032102 3

□

We have not been able to extend Theorem 6.3 to $\Sigma_5$. However, we believe that the following strong form of Dejean's conjecture holds.

**Conjecture 6.1.** *For every $k \geqslant 5$, there exist exponentially many $\frac{k}{k-1}^+$-free words over $\Sigma_k$.*

The C sources of the programs and the morphisms discussed in this part are available at: `http://dept-info.labri.fr/~ochem/morphisms/`.

# Chapter 7

# Generalized repetition thresholds

In this chapter, we provide new values and conjectured values of some generalized repetition thresholds $R(k, \ell)$.

## 7.1 Lower bounds

Figure 7.1 gives the established and conjectured values of $R(k, \ell)$. Entries in **bold** have been proved; the others (with question marks) are merely conjectured. However, in either case, if the entry for $(k, \ell)$ is $\alpha$, then we have proved, using the usual tree-traversal technique discussed below, that there is no infinite $(\alpha, \ell)$-free word over $\Sigma_k$.

The proved results are as follows:

- $R(2, 1) = R(2, 2) = 2$ follows from Thue's proof of the existence of infinite overlap-free binary words, together with the observation of Entringer, Jackson and Schatz [21] that squares of size at least two are unavoidable over $\Sigma_2$;
- The values of $R(k, 1) = R_k$, $3 \leqslant k \leqslant 11$ correspond to the proven cases of Dejean's conjecture.
- $R(2, 3) = \frac{8}{5}$, $R(2, 4) = \frac{3}{2}$, $R(2, 5) = \frac{7}{5}$, $R(2, 6) = \frac{4}{3}$, $R(3, 2) = \frac{3}{2}$ and $R(3, 3) = \frac{4}{3}$ are new and are proved in Section 7.2.

We now explain how the conjectured results were obtained. We used the usual tree-traversal technique, as follows: suppose we want to determine if there are only finitely many words over the alphabet $\Sigma_k$ that avoid a certain set of words $S$. We construct a certain tree $T$ and traverse it using breadth-first or depth-first search. The tree $T$ is defined as follows: the root is labelled $\varepsilon$ (the empty word). If a node $w$ has a factor contained in $S$, then it is a leaf. Otherwise, it has children labelled $wa$ for all $a \in \Sigma_k$. It is easy to see that $T$ is finite if and only if there are finitely many words avoiding $S$.

We can take advantage of various symmetries in $S$ to speed traversal. For example, if $S$ is closed under renaming of the letters (as is the case in the examples we study), we can label the root with an arbitrary single letter (instead of $\varepsilon$) and deduce the number of leaves in the full tree by multiplying by $k$.

| $R(k,\ell)$ | $\ell$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 2 | $\mathbf{2}$ | $\mathbf{2}$ | $\mathbf{\frac{8}{5}}$ | $\mathbf{\frac{3}{2}}$ | $\mathbf{\frac{7}{5}}$ | $\mathbf{\frac{4}{3}}$ | $\frac{31}{24}?$ | $\frac{24}{19}?$ |
| 3 | $\mathbf{\frac{7}{4}}$ | $\mathbf{\frac{3}{2}}$ | $\mathbf{\frac{4}{3}}$ | $\frac{5}{4}?$ | $\frac{6}{5}?$ | $\frac{7}{6}?$ | $\frac{8}{7}?$ | $\frac{9}{8}?$ |
| 4 | $\mathbf{\frac{7}{5}}$ | $\frac{5}{4}?$ | $\frac{6}{5}?$ | $\frac{7}{6}?$ | $\frac{8}{7}?$ | $\frac{9}{8}?$ | $\frac{10}{9}?$ | $\frac{11}{10}?$ |
| 5 | $\mathbf{\frac{5}{4}}$ | $\frac{6}{5}?$ | $\frac{8}{7}?$ | $\frac{9}{8}?$ | $\frac{10}{9}?$ | | | |
| 6 | $\mathbf{\frac{6}{5}}$ | $\frac{36}{31}?$ | | | | | | |
| 7 | $\mathbf{\frac{7}{6}}$ | $\frac{9}{8}?$ | | | | | | |
| 8 | $\mathbf{\frac{8}{7}}$ | | | | | | | |
| 9 | $\mathbf{\frac{9}{8}}$ | | | | | | | |
| 10 | $\mathbf{\frac{10}{9}}$ | | | | | | | |
| 11 | $\mathbf{\frac{11}{10}}$ | | | | | | | |
| 12 | $\frac{12}{11}?$ | | | | | | | |
| 13 | $\frac{13}{12}?$ | | | | | | | |

Figure 7.1: Known and conjectured values of $R(k, \ell)$.

Furthermore, if we use depth-first search, we can in some cases dramatically shorten the search using the following observation: if at any point some suffix of the current string strictly precedes the prefix of the same length of the same string in lexicographic order, then this suffix must have already been examined. Hence we can immediately abandon consideration of this node.

If the tree is finite, then certain parameters about the tree give useful information about the set of finite words avoiding $S$:

- If $h$ is the height of the tree, then any word of length at least $h$ over $\Sigma_k$ contains a factor in $S$.
- If $M$ is the length of a longest word avoiding $S$, then $M = h - 1$.
- If $I$ is the number of internal nodes, then there are exactly $I$ finite words avoiding $S$. Furthermore, if $L$ is the number of leaves, then (as usual), $L = 1 + (k-1)I$.
- If $I'$ is the number of internal nodes at depth $h-1$, then there are $I'$ words of maximum length avoiding $S$.

Figure 2 gives the value of some of these parameters. Here $\alpha$ is the established or conjectured value of $R(k, \ell)$ from Figure 1. “NR” indicates that the value was not recorded by our program.

| $k$ | $\ell$ | $\alpha$ | $L$ | $I$ | $h$ | $M{=}h{-}1$ | $I'$ |
|---|---|---|---|---|---|---|---|
| 2 | 1 | **2** | 8 | 7 | 4 | 3 | 2 |
| 2 | 2 | **2** | 478 | 477 | 19 | 18 | 2 |
| 2 | 3 | **8/5** | 5196 | 5195 | 34 | 33 | 12 |
| 2 | 4 | **3/2** | 13680 | 13679 | 54 | 53 | 4 |
| 2 | 5 | **7/5** | 40642 | 40641 | 60 | 59 | 4 |
| 2 | 6 | **4/3** | 21476 | 21475 | 40 | 39 | 4 |
| 2 | 8 | 24/19 | 3480734274 | 3480734273 | 452 | 451 | NR |
| 3 | 1 | **7/4** | 6393 | 3196 | 39 | 38 | 18 |
| 3 | 2 | **3/2** | 11655 | 5827 | 31 | 30 | 6 |
| 3 | 3 | **4/3** | 4037361 | 2018680 | 228 | 227 | 6 |
| 3 | 4 | 5/4 | 188247 | 94123 | 63 | 62 | 24 |
| 3 | 5 | 6/5 | 493653 | 246826 | 63 | 62 | 12 |
| 3 | 6 | 7/6 | 782931 | 391465 | 60 | 59 | 24 |
| 3 | 7 | 8/7 | 2881125 | 1440562 | 68 | 67 | 24 |
| 3 | 8 | 9/8 | 6987903 | 3493951 | 62 | 61 | 24 |
| 4 | 1 | **7/5** | 709036 | 236345 | 122 | 121 | 48 |
| 4 | 2 | 5/4 | 10324 | 3441 | 17 | 16 | 24 |
| 4 | 3 | 6/5 | 153724 | 51241 | 24 | 23 | 96 |
| 4 | 4 | 7/6 | 2501620 | 833873 | 35 | 34 | 24 |
| 4 | 5 | 8/7 | 30669148 | 10223049 | 40 | 39 | 864 |
| 4 | 6 | 9/8 | 340760884 | 113586961 | 50 | 49 | NR |
| 5 | 1 | **5/4** | 1785 | 446 | 7 | 6 | 120 |
| 5 | 2 | 6/5 | 453965 | 113491 | 23 | 22 | 240 |
| 5 | 3 | 8/7 | 7497345 | 1874336 | 34 | 33 | 720 |
| 5 | 4 | 9/8 | 1521535445 | 380383861 | 52 | 51 | NR |
| 6 | 1 | **6/5** | 13386 | 2677 | 8 | 7 | 720 |
| 6 | 2 | 36/31 | 17372138466 | 3474427693 | 751 | 750 | NR |
| 7 | 1 | **7/6** | 112441 | 18740 | 9 | 8 | 5040 |
| 7 | 2 | 9/8 | 345508219 | 57584703 | 32 | 31 | NR |
| 8 | 1 | **8/7** | 1049448 | 149921 | 10 | 9 | 40320 |

Figure 7.2: Tree statistics for various values of $k$ and $l$

We have seen how to prove computationally that only finitely many $(\alpha, \ell)$-free words exist. But what is the evidence that suggests we have determined the smallest possible $\alpha$? For this, we explore the tree corresponding to avoiding $(\alpha^+, \ell)$-repetitions using depth-first search. If we are able to construct a "very long" word avoiding $(\alpha^+, \ell)$-repetitions, then we suspect we have found the optimal value of $\alpha$. For each unproven $\alpha$ given in Figure 1, we were able to construct a word of length at least 20000 avoiding the corresponding repetitions. This constitutes weak evidence of the correctness of our conjectures, but it is evidently not conclusive.

Based on the data in Figure 1, we propose the following conjectures.

**Conjecture 7.1.**

1. $R(3,\ell) = 1 + \frac{1}{\ell}$ *for* $\ell \geqslant 2$.
2. $R(4,\ell) = 1 + \frac{1}{\ell+2}$ *for* $\ell \geqslant 2$.

These conjectures are weakly supported by the numerical evidence above.

## 7.2 New results

In this section, we prove six results of the form $R(k,\ell) = \alpha$. From the numerical results reported in Figure 2, we know in each case that there exist no infinite $(\alpha, \ell)$-free words over

$\Sigma_k$. It therefore suffices to exhibit an infinite $(\alpha^+, \ell)$-free word over $\Sigma_k$.

**Theorem 7.1.** $R(2,3) = \frac{8}{5}$.

*Proof.* Consider the 992-uniform morphism $h : \Sigma_4^* \longrightarrow \Sigma_2^*$ defined by

$0 \mapsto$ 0000101011110000111010100011110000101011110100001111000101011100001111010000101111
0000111010100011110000101111010100001111000101011100001111010100001011110000111010100011
1100001010111101000011110001010111000011110100001010111100001110101000111100001011110101
0000111100010101110000111101000010111100001110101000111100001011110100001111000101011100
0011110101000010111100001110101000111100001010111101000011110001010111000011110100001011
1100001110101000111100001011110101000011110001010111000011110100001010111100001110101000
1111000010111101000011110001010111000011110100001011110000111010100011110000101011110100
0011110001010111000011110101000010111100001110101000111100001011110100001111000101011100
0011110100001010111100001110101000111100001011110101000011110001010111000011110101000010
1111000011101010001111000010101111010000111100010101110000111101000010101111000011101010
0011110000101111010000111100010101110000111101010000101111000011101010001111000010111101
0100001111000101011100001111101,

$1 \mapsto$ 0000101011110000111010100011110000101011110100001111000101011100001111010000101111
0000111010100011110000101111010100001111000101011100001111010100001011110000111010100011
1100001010111101000011110001010111000011110100001010111100001110101000111100001011110101
0000111100010101110000111101000010111100001110101000111100001011110100001111000101011100
0011110100001010111100001110101000111100001010111101000011110001010111000011110101000010
1111000011101010001111000010111101010000111100010101110000111101000010101111000011101010
0011110000101111010000111100010101110000111101000010111100001110101000111100001010111101
0000111100010101110000111101010000101111000011101010001111000010111101000011110001010111
0000111101000010101111000011101010001111000010111101010000111100010101110000111101000010
1111000011101010001111000010101111010000111100010101110000111101000010101111000011101010
0011110000101111010000111100010101110000111101010000101111000011101010001111000010111101
0100001111000101011100001111101,

$2 \mapsto$ 0000101011110000111010100011110000101011110100001111000101011100001111010000101111
0000111010100011110000101111010000111100010101110000111101010000101111000011101010001111
0000101011110100001111000101011100001111010000101011110000111010100011110000101111010100
0011110001010111000011110101000010111100001110101000111100001011110100001111000101011100
0011110100001010111100001110101000111100001010111101000011110001010111000011110101000010
1111000011101010001111000010111101010000111100010101110000111101000010111100001110101000
1111000010101111010000111100010101110000111101000010101111000011101010001111000010111101
0000111100010101110000111101010000101111000011101010001111000010101111010000111100010101
1100001111010000101111000011101010001111000010111101010000111100010101110000111101000010
1011110000111010100011110000101111010000111100010101110000111101000010111100001110101000
1111000010101111010000111100010101110000111101010000101111000011101010001111000010111101
0100001111000101011100001111101,

$3 \mapsto$ 0000101011110000111010100011110000101011110100001111000101011100001111010000101111
0000111010100011110000101111010000111100010101110000111101010000101111000011101010001111
0000101011110100001111000101011100001111010000101011110000111010100011110000101111010100
0011110001010111000011110100001011110000111010100011110000101111010000111100010101110000
1111010000101011110000111010100011110000101011110100001111000101011100001111010100001011
1100001110101000111100001011110101000011110001010111000011110100001011110000111010100011
1100001010111101000011110001010111000011110100001010111100001110101000111100001011110101
0000111100010101110000111101010000101111000011101010001111000010101111010000111100010101
1100001111010000101111000011101010001111000010111101010000111100010101110000111101000010
1011110000111010100011110000101111010000111100010101110000111101000010111100001110101000
1111000010101111010000111100010101110000111101010000101111000011101010001111000010111101
0100001111000101011100001111101.

By a result of Pansiot [61], there exist $\frac{7}{5}^+$-free infinite words over $\Sigma_4$. Consider one such word $\mathbf{x}$. A computer check shows that $h$ is synchronizing and that for every $\frac{7}{5}^+$-free word $t \in \Sigma_4^*$ such that $|t| < \max_{\frac{7}{5},\frac{8}{5},992} = 16$, $h(t)$ is $\left(\frac{8}{5}^+, 3\right)$-free. By Lemma 6.1, this proves that $h(\mathbf{x})$ is an infinite binary $\left(\frac{8}{5}^+, 3\right)$-free word. □

**Theorem 7.2.** $R(2,4) = \frac{3}{2}$.

*Proof.* Consider the 19-uniform morphism $h : \Sigma_4^* \longrightarrow \Sigma_2^*$ defined by

$$0 \mapsto 0000110100100111110, \qquad 1 \mapsto 0000011011001010111,$$
$$2 \mapsto 0000011010100111111, \qquad 3 \mapsto 0000010110111110010.$$

We again consider an infinite $\frac{7}{5}^+$-free word $\mathbf{x}$ over $\Sigma_4$. A computer check shows that $h$ is synchronizing and that for every $\frac{7}{5}^+$-free word $t \in \Sigma_4^*$ such that $|t| < \max_{\frac{7}{5},\frac{3}{2},19} = 30$, $h(t)$ is $\left(\frac{3}{2}^+, 4\right)$-free. By Lemma 6.1, this proves that $h(\mathbf{x})$ is an infinite binary $\left(\frac{3}{2}^+, 4\right)$-free word. □

**Theorem 7.3.** $R(2,5) = \frac{7}{5}$.

*Proof.* Consider the 45-uniform morphism $h : \Sigma_5^* \longrightarrow \Sigma_2^*$ defined by

$$0 \mapsto 000000101011111100100000001101010100111111011,$$
$$1 \mapsto 000000101010111100010011011101000001111111011,$$
$$2 \mapsto 000000010101011111100100101101100010001110111,$$
$$3 \mapsto 000000010101011001100111111010001000101110111,$$
$$4 \mapsto 0000000100111101010100000001100111111101010011.$$

By a result of Moulin-Ollagnier [52], there exist $\frac{5}{4}^+$-free infinite words over $\Sigma_5$. Consider one such word $\mathbf{x}$. A computer check shows that $h$ is synchronizing and that for every $\frac{5}{4}^+$-free word $t \in \Sigma_5^*$ such that $|t| < \max_{\frac{5}{4},\frac{7}{5},45} = \frac{56}{3} < 19$, $h(t)$ is $\left(\frac{7}{5}^+, 5\right)$-free. By Lemma 6.1, this proves that $h(\mathbf{x})$ is an infinite binary $\left(\frac{7}{5}^+, 5\right)$-free word. □

**Theorem 7.4.** $R(2,6) = \frac{4}{3}$.

*Proof.* Consider the 71-uniform morphism $h : \Sigma_5^* \longrightarrow \Sigma_2^*$ defined by

$$0 \mapsto 0000000101010111111111000010001011101111010000000001101101010010011110111,$$
$$1 \mapsto 00000000101011111111000110010100101111011000010001111111101010101001100111,$$
$$2 \mapsto 0000000010101101111111111000110011010101000000011111110010001011010110 0111,$$
$$3 \mapsto 000000001010101111111000110010100101110111100001000110110101001001110111,$$
$$4 \mapsto 00000000100101011011111111100110010100000000011111101011010000100111101 11.$$

We again consider an infinite $\frac{5}{4}^+$-free word $\mathbf{x}$ over $\Sigma_5$. A computer check shows that $h$ is synchronizing and that for every $\frac{5}{4}^+$-free word $t \in \Sigma_5^*$ such that $|t| < \max_{\frac{5}{4},\frac{4}{3},71} = 32$, $h(t)$ is $\left(\frac{4}{3}^+, 6\right)$-free. By Lemma 6.1, this proves that $h(\mathbf{x})$ is an infinite binary $\left(\frac{4}{3}^+, 6\right)$-free word. □

**Theorem 7.5.** $R(3,2) = \frac{3}{2}$.

*Proof.* Consider the 3-uniform morphism $h : \Sigma_4^* \longrightarrow \Sigma_3^*$ defined by

$$0 \mapsto 021, \qquad 1 \mapsto 100, \qquad 2 \mapsto 122, \qquad 3 \mapsto 201.$$

We again consider an infinite $\frac{7}{5}^+$-free word $\mathbf{x}$ over $\Sigma_4$. A computer check shows that $h$ is synchronizing and that for every $\frac{7}{5}^+$-free word $t \in \Sigma_4^*$ such that $|t| < \max_{\frac{7}{5},\frac{3}{2},3} = 30$, $h(t)$ is $\left(\frac{3}{2}^+, 2\right)$-free. By Lemma 6.1, this proves that $h(\mathbf{x})$ is an infinite ternary $\left(\frac{3}{2}^+, 2\right)$-free word. □

**Theorem 7.6.** $R(3, 3) = \frac{4}{3}$.

*Proof.* Consider the 14-uniform morphism $h : \Sigma_5^* \longrightarrow \Sigma_3^*$ defined by

$$\begin{array}{lll} 0 \mapsto 00011112122220, & 1 \mapsto 00101112202021, & 2 \mapsto 01012111102120, \\ 3 \mapsto 10002212102020, & 4 \mapsto 10100222112020. & \end{array}$$

We again consider an infinite $\frac{5}{4}^+$-free word $\mathbf{x}$ over $\Sigma_5$. A computer check shows that $h$ is synchronizing and that for every $\frac{5}{4}^+$-free word $t \in \Sigma_5^*$ such that $|t| < \max_{\frac{5}{4},\frac{4}{3},14} = 32$, $h(t)$ is $\left(\frac{4}{3}^+, 3\right)$-free. By Lemma 6.1, this proves that $h(\mathbf{x})$ is an infinite ternary $\left(\frac{4}{3}^+, 3\right)$-free word. □

# Chapter 8

# Avoiding patterns and large squares

In this chapter, we show how the method in Section 6.2 can be used to construct infinite words avoiding some pattern. We also provide simpler proofs of known results about some infinite binary words avoiding large squares.

## 8.1 Pattern avoidance

We consider here the ternary patterns whose 2-avoidability was left open in Cassaigne's thesis [13]. We add to that list the binary pattern $AABBA$ (resp. $ABAAB$) which was already known to be 2-avoidable, and is here 2-avoided together with its reverse $ABBAA$ (resp. $ABBAB$). In particular, the 2-avoidability of $ABCACB$ was one of Currie's open problems [18], which was mentioned mainly because $ABCACB$ and its reverse are not simultaneously 2-avoidable.

**Lemma 8.1.** *Any factor $uvu$ of a $(\beta^+, n)$-free word $w$, with $\beta < 2$, is such that*

$$|u| \leqslant \max\left(n - 1 - |v|, \left\lfloor \frac{\beta - 1}{2 - \beta}|v| \right\rfloor\right).$$

*Proof.* If $|uv| < n$ then $|u| \leqslant n - 1 - |v|$ and we are done, so suppose $|uv| \geqslant n$. Since $w$ is $(\beta^+, n)$-free, we have $\frac{|uvu|}{|uv|} \leqslant \beta \Longrightarrow |u| \leqslant \frac{\beta-1}{2-\beta}|v|$. □

**Theorem 8.1.** *For every ternary pattern $P$ listed in Table 8.1, there exist exponentially many binary words avoiding $P$.*

*Proof.* Suppose that we are given a synchronizing morphism $h : \Sigma_s^* \to \Sigma_e^*$ and a pattern $P$ over the alphabet $\{A, B, \ldots\}$. We can try to prove that $h(t)$ avoids $P$ for every $\alpha^+$-free word $t \in \Sigma_s^*$ in three steps:

1. Check useful $(\beta^+, n)$-freeness properties of $h(t)$ thanks to Lemma 6.1.
2. Consider a potential occurrence $\phi(P)$ of $P$ and obtain upper bounds on the quantities $a = |\phi(A)|,\ b = |\phi(B)|,\ \ldots$ using Lemma 8.1 and the results of step (1).

| Pattern | $s$ | $\alpha$ | $q$ | Comments |
|---|---|---|---|---|
| AABAACBAAB | 3 | 7/4 | 8 | self-reverse |
| AABACCB | 3 | 7/4 | 24 | avoided with its reverse |
| AABBA | 3 | 7/4 | 21 | avoided with its reverse |
| AABBCABBA | 3 | 7/4 | 102 | unavoidable with its reverse |
| AABBCAC | 3 | 7/4 | 86 | avoided with its reverse |
| AABBCBC | 3 | 7/4 | 34 | avoided with its reverse |
| AABBCC | 3 | 7/4 | 52 | self-reverse |
| AABCBC | 3 | 7/4 | 46 | avoided with its reverse |
| AABCCAB | 3 | 7/4 | 34 | avoided with its reverse |
| AABCCBA | 3 | 7/4 | 56 | avoided with its reverse |
| ABAAB | 3 | 7/4 | 10 | avoided with its reverse |
| ABAACBC | 4 | 7/5 | 17 | avoided with its reverse |
| ABAACCB | 3 | 7/4 | 74 | avoided with its reverse |
| ABACACB | 3 | 7/4 | 12 | avoided with its reverse |
| ABACBC | 4 | 7/5 | 29 | self-reverse |
| ABACCAB | 3 | 7/4 | 19 | avoided with its reverse |
| ABACCBA | 3 | 7/4 | 14 | avoided with its reverse |
| ABBACCA | 3 | 7/4 | 12 | self-reverse |
| ABBACCB | 3 | 7/4 | 42 | avoided with its reverse |
| ABBCACB | 4 | 7/5 | 16 | avoided with its reverse |
| ABBCBAC | 4 | 7/5 | 14 | avoided with its reverse |
| ABBCBCA | 3 | 7/4 | 22 | avoided with its reverse |
| ABBCCCAB | 3 | 7/4 | 20 | avoided with its reverse |
| ABCAACB | 3 | 7/4 | 24 | avoided with its reverse |
| ABCACAB | 3 | 7/4 | 10 | avoided with its reverse |
| ABCACB | 6 | 5/4 | 810 | unavoidable with its reverse |
| ABCBABC | 4 | 7/5 | 19 | self-reverse |
| ABCBBAC | 4 | 7/5 | 18 | avoided with its reverse |

Figure 8.1: Table of 2-avoidable ternary patterns.

3. Use the bounds of step (2) to exhaustively check by computer that no occurrence of $P$ appears in $h(t)$.

Let $P^R$ denote the reverse of the pattern $P$. We have $\mu(P^R) = \mu(P)$ since a word $w$ avoids $P$ if and only the reverse of $w$ avoids $P^R$. Notice that the bounds obtained in step (2) using Lemma 8.1 that hold for potential occurrences of a pattern $P$ also hold for those of $P^R$. Thus we try, when possible, to avoid simultaneously $P$ and $P^R$. The method discussed in Section 6.2 is thus used so that the set $S$ contains the small occurrences of $P$ (and maybe $P^R$). Each line of Table 8.1 contains one of these pattern $P$ and informations about the morphism $h$ we used to show that $\mu(P) = 2$: the $q$-uniform morphism $h : \Sigma_s^* \to \Sigma_2^*$ is such that for every $\alpha^+$-free word $t \in \Sigma_s^*$, $h(t)$ avoids $P$. We also precise whether such a word $h(t)$ also avoids $P^R$. Binary words avoiding $ABCACB$ are constructed from $\frac{5}{4}^+$-free words over $\Sigma_5$, and there are exponentially many such words by Lemma 6.2. For every other pattern $P$ listed in Table 8.1, binary words avoiding $P$ are constructed from either $\frac{7}{4}^+$-free words over $\Sigma_3$, or $\frac{7}{5}^+$-free words over $\Sigma_4$. In both cases, there are exponentially many such words by Theorem 6.3. Now, for each pattern, we give the bounds obtained in step (2) of the proof and the morphism found with the method in Section 6.2.

The following 8-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABAACBAAB$.
The word $h(t)$ is $\left(\frac{24}{13}^+, 3\right)$-free. It contains no square of length at least three, and since the square $AA$ occurs in the pattern, we have that $a \leqslant 2$. By Lemma 8.1, the factor $BAAB$ implies $b \leqslant 11a$. For each occurrence of $BAAB$ appearing in $h(t)$, we have checked that the corresponding occurrence of $AABAA$ does not appear. For instance, the occurrence $\phi(BAAB) = 0110$ (where $\phi(A) = 1,\ \phi(B) = 0$) appears in $h(t)$, but the factor $\phi(AABAA) = 11011$ does not.

$$\begin{aligned}0 &\mapsto 01101011\\ 1 &\mapsto 00111010\\ 2 &\mapsto 00101110\end{aligned}$$

The following 24-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABACCB$ and its reverse.
The word $h(t)$ is $\left(\frac{31}{16}^+, 3\right)$-free, so $a \leqslant 2$ and $c \leqslant 2$. The factor $BACCB$ implies $b \leqslant 15(a+2c)$.

$$\begin{aligned}0 &\mapsto 000101001101011001010111\\ 1 &\mapsto 000101000111010111001011\\ 2 &\mapsto 000101000110100111001011\end{aligned}$$

The following 21-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABBA$ and its reverse.
The word $h(t)$ is $\left(\frac{33}{17}^+, 4\right)$-free, so $a \leqslant 3$ and $b \leqslant 3$.

$$\begin{aligned}0 &\mapsto 001001011100011101101\\ 1 &\mapsto 001000111000101101101\\ 2 &\mapsto 000111000100101101101\end{aligned}$$

The following 102-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABBCABBA$.
The word $h(t)$ is $\left(\frac{31}{16}^+, 27\right)$-free, so $a \leqslant 26$ and $b \leqslant 26$. For each occurrence of $AABB$ we have checked that the corresponding occurrence of $ABBA$ does not appear. Notice that the $k$-avoidability of $AABBCABBA$ implies the $k$-avoidability of $AABBA$. A simple backtracting algorithm shows that $AABBCABBA$ and $ABBAA$ (i.e. the reverse of $AABBA$) are not simultaneously 2-avoidable, so that the two previous results are tight, in a way.

$$\begin{aligned}0 &\mapsto 000100010110111011100010110001000101101110110001011011100010110\,1\\ &\quad\ 110111000101101110110001011011100010111\\ 1 &\mapsto 000100010110111011000101101110111000101100010001011011100010110\,1\\ &\quad\ 110111000101101110110001011011100010111\\ 2 &\mapsto 000100010110111000101101110111000101100010001011011101100010110\,1\\ &\quad\ 110111000101101110110001011011100010111\end{aligned}$$

The following 86-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABBCAC$ and its reverse.
The word $h(t)$ is $\left(\frac{43}{22}^+, 3\right)$-free, so $a \leqslant 2$ and $b \leqslant 2$.
The factor $CAC$ implies $c \leqslant \max(2-a, 21a) = 21a$.

$$\begin{aligned}
0 &\mapsto 00010101100101001101010110010101001101011001010011010\\
&\qquad 100010111010100110101100101011101\\
1 &\mapsto 00010101100101001101010001011101010011010110010101001\\
&\qquad 101010110010100110101100101011101\\
2 &\mapsto 00010101100101001101010001011101010011010110010100110\\
&\qquad 101011001010100110101100101011101
\end{aligned}$$

The following 34-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABBCBC$ and its reverse.
The word $h(t)$ is $\left(\frac{33}{17}^+, 3\right)$-free, so $a \leqslant 2$ and $b = c = 1$.

$$\begin{aligned}
0 &\mapsto 1110101110001010001110001010100011\\
1 &\mapsto 1110101110001010100011100010100011\\
2 &\mapsto 1110101110001110101000111000101000
\end{aligned}$$

The following 52-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABBCC$.
The word $h(t)$ is $\left(\frac{59}{30}^+, 3\right)$-free, so $a \leqslant 2$, $b \leqslant 2$, and $c \leqslant 2$.

$$\begin{aligned}
0 &\mapsto 1101001110001101000101100011010011100101100011100101\\
1 &\mapsto 1101001110001101000101100011100101100011010011100101\\
2 &\mapsto 1101001110001101000101100011100101110100111001011000
\end{aligned}$$

The following 46-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABCBC$ and its reverse.
The word $h(t)$ is $\left(\frac{367}{184}^+, 3\right)$-free, so $a \leqslant 2$ and $b = c = 1$. Notice that the only occurrences of $AABB$ in $h(t)$ are 0011 and 1100.

$$\begin{aligned}
0 &\mapsto 0011010011100011001011000110100110001110010110\\
1 &\mapsto 0011010011100011001011000110100111001011000111\\
2 &\mapsto 0011010011100011001011000111001101001110010110
\end{aligned}$$

The following 34-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABCCAB$ and its reverse.
The word $h(t)$ is $\left(\frac{257}{136}^+, 4\right)$-free, so $a \leqslant 3$ and $c \leqslant 3$. The factor $ABCCAB$ implies $a + b \leqslant \left\lfloor \frac{242}{15} c \right\rfloor$, thus $b \leqslant \left\lfloor \frac{242}{15} c \right\rfloor - a$.

$$\begin{aligned}
0 &\mapsto 0000101111110100001111110100010111 1\\
1 &\mapsto 0000101110100001111010001011111100\\
2 &\mapsto 0000101110100000011110100010111111
\end{aligned}$$

The following 56-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $AABCCBA$ and its reverse.
The word $h(t)$ is $\left(\frac{36}{19}^+, 3\right)$-free, so $a \leqslant 2$ and $c \leqslant 2$. The factor $BCCB$ implies $b \leqslant 17c$.

$$\begin{aligned} 0 &\mapsto 00010110001110010111010011100011010001011000111001011101 \\ 1 &\mapsto 00010110001101001110010110001101000101110100111001011101 \\ 2 &\mapsto 00010110001101000101110100111001011000110100111001011101 \end{aligned}$$

The following 10-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABAAB$ and its reverse.
The word $h(t)$ is $\left(\frac{39}{20}^+, 3\right)$-free, so $a \leqslant 2$. The factor $BAAB$ implies $b \leqslant 38a$.

$$\begin{aligned} 0 &\mapsto 0001110101 \\ 1 &\mapsto 0000111101 \\ 2 &\mapsto 0000101111 \end{aligned}$$

The following 17-uniform morphism $h$ is such that for any $\frac{7}{5}^+$-free word $t \in \Sigma_4^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABAACBC$ and its reverse.
The word $h(t)$ is $\left(\frac{13}{7}^+, 7\right)$-free, so $a \leqslant 6$. The word $h(t)$ is $\left(\frac{23}{16}^+, 20\right)$-free. Suppose $b+c \geqslant 20$, then the factor $CBC$ implies $c \leqslant \left\lfloor \frac{7}{9}b \right\rfloor$ and the factor $BAACB$ implies $b \leqslant \left\lfloor \frac{7}{9}(2a+c) \right\rfloor$. From these relations we can deduce $b \leqslant 22$ and $c \leqslant 17$. Thus we have $b \leqslant 22$ and $c \leqslant 18$.

$$\begin{aligned} 0 &\mapsto 01110000110000111 \\ 1 &\mapsto 01011100110011101 \\ 2 &\mapsto 01000110011000101 \\ 3 &\mapsto 00011110011110001 \end{aligned}$$

The following 74-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABAACCB$ and its reverse.
The word $h(t)$ is $\left(\frac{193}{104}^+, 3\right)$-free, so $a \leqslant 2$ and $c \leqslant 2$. The factor $BAACCB$ implies $b \leqslant \left\lfloor \frac{89}{15}(a+c) \right\rfloor$.

$$\begin{aligned} 0 &\mapsto 00000101100111010011010110010111001101000001111101001100\\ &\qquad 010100011001011111 \\ 1 &\mapsto 00000101100101010011000101100101001101000001111101001100\\ &\qquad 010100011001011111 \\ 2 &\mapsto 00000101100101001101000110010101001101000001111101001100\\ &\qquad 010100011001011111 \end{aligned}$$

The following 12-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABACACB$ and its reverse.
The word $h(t)$ is $\left(\frac{15}{8}^+, 3\right)$-free, so $a = c = 1$. The factor $BACACB$ implies $b \leqslant 14(a+c) = 28$.

$$0 \mapsto 001010011111 \qquad 1 \mapsto 000110100111 \qquad 2 \mapsto 000001101011$$

The following 29-uniform morphism $h$ is such that for any $\frac{7}{5}^+$-free word $t \in \Sigma_4^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABACBC$.
The word $h(t)$ is $\left(\frac{41}{29}^+, 291\right)$-free. Suppose $a + b \geqslant 291$ and $b + c \geqslant 291$. The factors $ABA$, $CBC$, and $BACB$ respectively imply that $17a \leqslant 12b$ $(i)$, $17c \leqslant 12b$ $(ii)$, and $17b \leqslant 12(a+c)$ $(iii)$. The combination $17 \times (i) + 17 \times (ii) + 24 \times (iii)$ gives $a+c \leqslant 0$, a contradiction. So we can suppose without loss of generality that $b+c \leqslant 290$ $(iv)$. If $291 \leqslant a+b$ $(v)$ then $(i)$ and $(iii)$ still hold and the combination $2324 \times (i) + 1649 \times (iii) + 19788 \times (iv) + 19720 \times (v)$ gives $213b \leqslant 0$. This contradiction shows that $a+b \leqslant 290$ and $b+c \leqslant 290$.

$$
\begin{aligned}
0 &\mapsto 00011010110000111100101001110\\
1 &\mapsto 00011010110000011100101001111\\
2 &\mapsto 00001101011000111101110011110\\
3 &\mapsto 00001101011000011110011101111
\end{aligned}
$$

The following 19-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABACCAB$ and its reverse.
The word $h(t)$ is $\left(\frac{35}{19}^+, 5\right)$-free, so $c \leqslant 4$. The factor $ACCA$ implies $a \leqslant \max\left(4 - 2c, \left\lfloor \frac{32}{3}c \right\rfloor\right) = \left\lfloor \frac{32}{3}c \right\rfloor$. The factor $ABACCAB$ implies $a + b \leqslant \left\lfloor \frac{16}{3}(a+2c) \right\rfloor$, thus $b \leqslant \left\lfloor \frac{13a+32c}{3} \right\rfloor$.

$$
\begin{aligned}
0 &\mapsto 0101110010100000111\\
1 &\mapsto 0101100000010011100\\
2 &\mapsto 0100010111010100011
\end{aligned}
$$

The following 14-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABACCBA$ and its reverse.
If $w$ is an occurrence of $ABACCBA$ such that $a > 1$, then the suffix of $w$ of size $|w| - a + 1$ is a smaller occurrence of $ABACCBA$ such that $a = 1$. So we assume without loss of generality that $a = 1$. The word $h(t)$ is $\left(\frac{23}{12}^+, 3\right)$-free, so $c \leqslant 2$. The factor $BACCBA$ implies $a+b \leqslant 22c$, thus $b \leqslant 22c - 1$.

$$
\begin{aligned}
0 &\mapsto 10101100001110\\
1 &\mapsto 01010111100011\\
2 &\mapsto 01010000111110
\end{aligned}
$$

The following 12-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABBACCA$.
The word $h(t)$ is $\left(\frac{31}{16}^+, 4\right)$-free, so $b \leqslant 3$ and $c \leqslant 3$. The factor $ABBA$ implies $a \leqslant \max(3 - 2b, 30b) = 30b$.

$$
\begin{aligned}
0 &\mapsto 000111001011\\
1 &\mapsto 000101111010\\
2 &\mapsto 000100111011
\end{aligned}
$$

The following 42-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABBACCB$ and its reverse.

The word $h(t)$ is $\left(\frac{53}{28}^+,3\right)$-free, so $b\leqslant 2$ and $c\leqslant 2$. The factor $ABBA$ implies $a\leqslant\left\lfloor\frac{50}{3}b\right\rfloor$.

$$\begin{aligned}0&\mapsto 00001011110001011101000111101000011101011 1\\1&\mapsto 0000101111000101011001010011010110010101 11\\2&\mapsto 00001011110001010001111010000111010001011 1\end{aligned}$$

The following 16-uniform morphism $h$ is such that for any $\frac{7}{5}^+$-free word $t\in\Sigma_4^*$, $h(t)\in\Sigma_2^*$ avoids the pattern $ABBCACB$ and its reverse.
The word $h(t)$ is $\left(\frac{9}{5}^+,4\right)$-free, so $b\leqslant 3$. The word $h(t)$ is $\left(\frac{233}{160}^+,49\right)$-free. Suppose $a+c\geqslant 49$. The factors $ABBCA$ and $CAC$ respectively imply that $a\leqslant\left\lfloor\frac{73}{87}(2b+c)\right\rfloor$ and $c\leqslant\left\lfloor\frac{73}{87}a\right\rfloor$. From these relations we can deduce $a\leqslant 15$ and $c\leqslant 12$. This contradiction shows that $a+c\leqslant 48$.

$$\begin{aligned}0&\mapsto 0010000001101111\\1&\mapsto 0000111010001111\\2&\mapsto 0000100111111011\\3&\mapsto 0000001001101011\end{aligned}$$

The following 14-uniform morphism $h$ is such that for any $\frac{7}{5}^+$-free word $t\in\Sigma_4^*$, $h(t)\in\Sigma_2^*$ avoids the pattern $ABBCBAC$ and its reverse.
The word $h(t)$ is $\left(\frac{18}{11}^+,7\right)$-free, so $b\leqslant 6$. The word $h(t)$ is $\left(\frac{29}{20}^+,43\right)$-free. Suppose $a+b+c\geqslant 43$. The factors $ABBCBA$ and $CBAC$ respectively imply that $a\leqslant\left\lfloor\frac{9}{11}(3b+c)\right\rfloor$ and $c\leqslant\left\lfloor\frac{9}{11}(a+b)\right\rfloor$. From these relations we can deduce $a\leqslant 54$ and $c\leqslant 49$.

$$\begin{aligned}0&\mapsto 00101010101011\\1&\mapsto 00010001110111\\2&\mapsto 00000101011111\\3&\mapsto 00000010111111\end{aligned}$$

The following 22-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t\in\Sigma_3^*$, $h(t)\in\Sigma_2^*$ avoids the pattern $ABBCBCA$ and its reverse.
The word $h(t)$ is $\left(\frac{173}{88}^+,3\right)$-free, so $b=c=1$. The factor $ABBCBCA$ implies $a\leqslant\left\lfloor\frac{85}{3}(3b+2c)\right\rfloor=141$.

$$\begin{aligned}0&\mapsto 0001101011001010100111\\1&\mapsto 0001101010110010100111\\2&\mapsto 0001101010011100101011\end{aligned}$$

The following 20-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t\in\Sigma_3^*$, $h(t)\in\Sigma_2^*$ avoids the pattern $ABBCCAB$ and its reverse.
The word $h(t)$ is $\left(\frac{15}{8}^+,4\right)$-free, so $b\leqslant 3$ and $c\leqslant 3$. The factor $ABBCCAB$ implies $a+b\leqslant 7(b+2c)$, thus $a\leqslant 6b+14c$.

$$\begin{aligned}0&\mapsto 00010100100101011111\\1&\mapsto 00010010001110110111\\2&\mapsto 00000101011011010111\end{aligned}$$

The following 24-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABCAACB$ and its reverse.
The word $h(t)$ is $\left(\frac{187}{96}^+, 4\right)$-free, so $a \leqslant 3$. The word $h(t)$ is $\left(\frac{355}{192}^+, 97\right)$-free. The factor $CAAC$ implies $c \leqslant \max\left(96 - 2a, \left\lfloor \frac{326a}{29} \right\rfloor\right) = 96 - 2a$. The factor $BCAACB$ implies $b \leqslant \max\left(96 - 2a - 2c, \left\lfloor \frac{326(a+c)}{29} \right\rfloor\right)$.

$$
\begin{aligned}
0 &\mapsto 000001011111001000110111 \\
1 &\mapsto 000001011111000100111011 \\
2 &\mapsto 000001010111110010011011
\end{aligned}
$$

The following 10-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABCACAB$ and its reverse.
The word $h(t)$ is $\left(\frac{79}{40}^+, 3\right)$-free, so $a = c = 1$. The word $h(t)$ is $\left(\frac{149}{80}^+, 41\right)$-free. The factor $ABCACAB$ implies $a + b \leqslant \max\left(40 - a - 2c, \left\lfloor \frac{69}{11}(a + 2c) \right\rfloor\right)$, thus $b \leqslant 40 - 2a - 2c = 36$.

$$
0 \mapsto 0001110101 \qquad 1 \mapsto 0001011101 \qquad 2 \mapsto 0001010111
$$

The 810-uniform morphism $h = m_{4,2} \circ m_{6,4}$ is such that for any $\frac{5}{4}^+$-free word $t \in \Sigma_6^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABCACB$.
The word $h(t)$ is $\left(\frac{1073}{810}^+, 3241\right)$-free. Suppose $a + c \geqslant 3241$. The factors $ABCA$, $BCACB$, and $CAC$ respectively imply that $547a \leqslant 263(b + c)$ $(i)$, $547b \leqslant 263(a + 2c)$ $(ii)$, and $547c \leqslant 263a$ $(iii)$. The combination $2 \times (i) + (ii) + 2 \times (iii)$ gives $305a + 568b + 42c \leqslant 0$. This contradiction shows that $a + c \leqslant 3240$ $(iv)$. The word $h(t)$ is $\left(\frac{29}{14}^+, 4\right)$-free, so the factor $CAC$ implies $c \leqslant 14a$ $(v)$. Suppose now $3238 \leqslant b$ $(vi)$, so that $a + b + 2c \geqslant 3241$ and $(ii)$ still holds. The combination $15 \times (ii) + 7627 \times (iv) + 263 \times (v) + 7632 \times (vi)$ gives $573b + 936 \leqslant 0$. This contradiction shows that $b \leqslant 3237$.

The 135-uniform morphism $m_{4,2}$ is given by:

0 $\mapsto$ 000001001111011000101111000000101101000011111011000010011110000010
1101000011111010010111100001001111101100000100111010000111110100101111
1 $\mapsto$ 000001001111011000101111000000101101000011111011000010011110000010
1101000011110110000010011111011000101111000000101101000011111010010111 1
2 $\mapsto$ 000001001111011000011111010010111100001001111101100000100111010000
0111110100101111000000101101000011111011000010011101000011111010010 1111
3 $\mapsto$ 000001001111011000011111010010111100000010110100001111011000001001
111101100010111100000010110100001111101100001001110100001111101001 01111

The 6-uniform morphism $m_{6,4}$ is given by:

$$
\begin{array}{lll}
0 \mapsto 032131 & 1 \mapsto 031232 & 2 \mapsto 023121 \\
3 \mapsto 021323 & 4 \mapsto 013212 & 5 \mapsto 012313
\end{array}
$$

An occurrence of $ABCBABC$ is a $\left(\frac{3}{2}^+, 4\right)$-repetition since $|\phi(ABCB)| \geqslant 4$ and $\frac{|\phi(ABCBABC)|}{|\phi(ABCB)|} = \frac{2a+3b+2c}{a+2b+c} = \frac{3}{2} + \frac{a+c}{2(a+2b+c)} > \frac{3}{2}$. So $ABCBABC$ is avoided by all the infinite $\left(\frac{3}{2}^+, 4\right)$-free binary words, in particular those constructed in the proof of Theorem 7.2.

The following 18-uniform morphism $h$ is such that for any $\frac{7}{5}^+$-free word $t \in \Sigma_4^*$, $h(t) \in \Sigma_2^*$ avoids the pattern $ABCBBAC$ and its reverse.
The word $h(t)$ is $\left(\frac{8}{5}^+, 7\right)$-free, so $b \leqslant 6$. The word $h(t)$ is $\left(\frac{527}{378}^+, 181\right)$-free. Suppose $a+2b+c \geqslant 181$. The factors $CBBAC$ and $ABCBBA$ respectively imply that $c \leqslant \left\lfloor \frac{149}{229}(a+2b) \right\rfloor$ and $a \leqslant \left\lfloor \frac{149}{229}(3b+c) \right\rfloor$. From these relations we can deduce $a \leqslant 28$ and $c \leqslant 26$. This contradiction shows that $a + 2b + c \leqslant 180$.

$$\begin{aligned}
0 &\mapsto 010000011011011011\\
1 &\mapsto 001001001001111101\\
2 &\mapsto 001000000111111011\\
3 &\mapsto 000000101010111111
\end{aligned}$$

□

## 8.2 Binary words avoiding large squares

Fraenkel and Simpson [24] constructed an infinite binary word containing only three squares. Another construction using uniform morphisms is given in [63]. Shallit [68] also gives uniform morphisms for binary words avoiding:

- squares of length at least 3 and $3^+$-repetitions (10-uniform),
- squares of length at least 4 and $\frac{5}{2}^+$-repetitions (1560-uniform),
- squares of length at least 7 and $\frac{7}{3}^+$-repetitions (252-uniform).

In this section we give small $\Sigma_3^* \to \Sigma_2^*$ uniform morphisms producing words having these properties.

The following 50-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ contains only the squares in $\{0^2, 1^2, (01)^2\}$ and is $\left(\frac{37}{19}^+, 3\right)$-free.

$$\begin{aligned}
0 &\mapsto 00011001011000111001011001110001011100101100010111\\
1 &\mapsto 00011001011000101110010110011100010110001110010111\\
2 &\mapsto 00011001011000101110010110001110010111000101100111
\end{aligned}$$

The following 8-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ is $3^+$-free, $\left(\frac{5}{2}^+, 2\right)$-free, and $\left(\frac{59}{32}^+, 3\right)$-free.

$$0 \mapsto 01101011 \qquad 1 \mapsto 00111010 \qquad 2 \mapsto 00101110$$

The following 103-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$

is $\frac{5}{2}^+$-free, $\left(\frac{7}{3}^+, 3\right)$-free, and $\left(\frac{823}{412}^+, 4\right)$-free.

$$
\begin{aligned}
0 &\mapsto 0010011010010110010011011001010011010110010011011001011010011011\\
&\qquad 001001101011001010011011001011010011011\\
1 &\mapsto 0010011010010110010011011001010011010110010011011001011010011011\\
&\qquad 001001101001011001001101011001010011011\\
2 &\mapsto 0010011010010110010011011001010011010110010011010010110010011011\\
&\qquad 001011010011011001001101011001010011011
\end{aligned}
$$

The following 30-uniform morphism $h$ is such that for any $\frac{7}{4}^+$-free word $t \in \Sigma_3^*$, $h(t) \in \Sigma_2^*$ is $\frac{7}{3}^+$-free and $\left(\frac{79}{40}^+, 7\right)$-free.

$$
\begin{aligned}
0 &\mapsto 001011001011010011011001001101\\
1 &\mapsto 001011001011010011001011001101\\
2 &\mapsto 001011001001101100101101001101
\end{aligned}
$$

# Chapter 9

# Letter frequencies

In this chapter, we consider the extremal frequency of a letter in infinite words over a finite alphabet avoiding some repetitions. The main motivation for this kind of problem is best explain with an example. We know that squares are avoidable over $\Sigma_3$ but are not over $\Sigma_2$. A natural question is thus whether this third letter is essential or just barely needed to avoid squares. We can formalize this question and ask for the minimal letter frequency in an infinite ternary square-free word.

## 9.1 Statement of main results

For ternary square-free words, Tarannikov [72] showed that $f_{\min} \in \left[\frac{1780}{6481}, \frac{64}{233}\right] = [0.27464897\ldots, 0.27467811\ldots]$. According to [65], he also proved that $f_{\max} \leqslant \frac{469}{1201} = 0.39050791\cdots$. Our next result provides better estimations of these constants:

**Theorem 9.1.** *For ternary square-free words, we have*

1. $f_{\min} \in \left[\frac{1000}{3641}, \frac{883}{3215}\right] = [0.27464982\ldots, 0.27465007\ldots]$.
2. $f_{\max} = \frac{255}{653} = 0.39050535\cdots$.

**Theorem 9.2.** *For $(\frac{5}{3}, 3)$-free binary words, we have $f_{\min} = \frac{1}{2}$.*

Theorem 9.2 implies that infinite $(\beta, 3)$-free binary words have equal letter frequency for $\beta \in \left[\frac{8}{5}^+, \frac{5}{3}\right]$. A similar result in [40] says that infinite $(\beta, 1)$-free binary words have equal letter frequency for $\beta \in \left[2^+, \frac{7}{3}\right]$, i.e. $\rho(2^+) = \rho\left(\frac{7}{3}\right) = \frac{1}{2}$. It is noticeable that these two cases of equal letter frequency have different kind of growth function. Karhumäki and Shallit have shown there exist polynomially many $\frac{7}{3}$-free binary words [39], whereas there exist exponentially many $(\frac{8}{5}^+, 3)$-free binary words. To see this, notice that the 992-uniform morphism $h : \Sigma_4^* \to \Sigma_2^*$ given in the proof of Theorem 7.1 produces a $\left(\frac{8}{5}^+, 3\right)$-free binary word $h(w)$ for every $\frac{7}{5}^+$-free word $w \in \Sigma_4^*$, and that there exist exponentially many $\frac{7}{5}^+$-free words over $\Sigma_4$ by Theorem 6.3.(2).

Kolpakov et al. [40] proved that the function $\rho$ is discontinuous at $\frac{7}{3}$, more precisely they obtained that $\rho\left(\frac{7}{3}\right) = \frac{1}{2}$ and $\rho\left(\frac{7}{3}^+\right) \leqslant \frac{10}{21} = 0.47619047\cdots$.

The next result provides new points of discontinuity for $\rho$ in the range $\left[\frac{7}{3}^+, 3\right]$, namely $\frac{17}{7}$, $\frac{5}{2}$, $\frac{23}{9}$, $\frac{41}{16}$, $\frac{18}{7}$, and $\frac{8}{3}$.

**Theorem 9.3.**

1. $\rho\left(\frac{7}{3}^+\right) \leqslant \frac{47}{101} = 0.46534653\cdots$.
2. $\rho\left(\frac{17}{7}\right) \geqslant \frac{467}{1004} = 0.46513944\cdots$.
3. $\rho\left(\frac{17}{7}^+\right) \leqslant \frac{81}{175} = 0.46285714\cdots$.
4. $\rho\left(\frac{5}{2}\right) \geqslant \frac{54286}{117293} = 0.46282386\cdots$.
5. $\rho\left(\frac{5}{2}^+\right) \leqslant \frac{23}{52} = 0.44230769\cdots$.
6. $\rho\left(\frac{23}{9}\right) > \frac{205}{464} = 0.44181034\cdots$.
7. $\rho\left(\frac{23}{9}^+\right) \leqslant \frac{91}{206} = 0.44174757\cdots$.
8. $\rho\left(\frac{41}{16}\right) > \frac{322}{729} = 0.44170096\cdots$.
9. $\rho\left(\frac{41}{16}^+\right) \leqslant \frac{143}{324} = 0.44135802\cdots$.
10. $\rho\left(\frac{18}{7}\right) \geqslant \frac{79}{179} = 0.44134078\cdots$.
11. $\rho\left(\frac{18}{7}^+\right) \leqslant \frac{41}{93} = 0.44086021\cdots$.
12. $\rho\left(\frac{8}{3}\right) > \frac{339}{769} = 0.44083224\cdots$.
13. $\rho\left(\frac{8}{3}^+\right) \leqslant \frac{24}{59} = 0.40677966\cdots$.
14. $\rho(3) > \frac{115}{283} = 0.40636042\cdots$.

## 9.2 Method for negative results

Let $L$ be a factorial language. A word $w$ is said to be $k$-biprolongable in $L$ if there exists a word $lwr \in L$ such that $|l| = |r| = k$. A *suffix cover* of $L$ is a set $S$ of finite words in $L$ such that every finite word that is $k$-biprolongable in $L$ and of length at least $\max_{u\in S}|u|$ has a suffix that belongs to $S$, for some finite number $k$. Taking $k = 20$ is sufficient for every negative result in this paper. For a word $u \in S$, let

$$A_u(q) = \left\{ w \in L \mid uw \in L \text{ and for every prefix } w' \text{ of } w,\ \frac{|w'|_0}{|w'|} < q \right\}.$$

**Lemma 9.1.** *Let $L$ be a factorial language and $S$ one of its suffix covers. Let $q \in \mathbb{Q}$. If $A_u(q)$ is finite for every word $u \in S$, then $f_{\min} \geqslant q$.*

*Proof.* Assume $A_u(q)$ is finite for every word $u \in S$. Then any infinite word $w \in L$ has a decomposition into finite factors $w_0 w_1 w_2 \dots$ such that $|w_0| = k + \max_{u\in S}|u|$ and $\frac{|w_i|_0}{|w_i|} \geqslant q$ for every $i \geqslant 1$. □

Lemma 9.1 enables us to obtain bounds of the form $f_{\min} \geqslant q$ by choosing an explicit suffix cover and checking by computer that every set $A_u(q)$ is finite. It is easy to see that Lemma 9.1 and the definition of $A_u(q)$ can be modified to provide bounds of the form $f_{\min} > q$, $f_{\max} \leqslant q$, or $f_{\max} < q$. This method is a natural generalization of the one in [72], where the suffix cover consists of the empty word, and of the one in [45], where the suffix cover consists of all binary words of length three. Since we study here the frequency of the letter 0 in repetition-free words, every letter other than 0 play the same role. Let us say that two words $u$ and $u'$ in $\Sigma_s$ are equivalent if and only if $u$ can be obtained from $u'$ by a permutation of the letters in $\Sigma_s \setminus \{0\}$. Notice that for two equivalent words $u$ and $u'$, $A_{u'}(q)$ is finite if and only if $A_u(q)$ is finite. We define the reduced suffix cover of a suffix cover $S$ as the quotient of $S$ by this equivalence relation. To prove the negative part of Theorem 9.1.1 we used the reduced suffix cover $\{1, 01210, 0210, 2010\}$, the computation took about 20 days on a XEON 2.2Gh. For Theorem 9.1.2 we used the reduced suffix cover $\{0, 01, 021, 0121\}$. For Theorem 9.2 we used the suffix cover $\{01, 111, 000, 1110, 1010, 0001111000010, 0111101000010, 1110101000010, 0111100010\}$. We omit the computer proof that this is indeed a suffix cover for $(\frac{5}{3}, 3)$-free binary words. The negative statements of Theorem 9.3 (even items) were obtained using the suffix cover $\{1, 10, 100\}$.

## 9.3 Method for positive results

Let $L$ be a factorial language over $\Sigma_s^*$. To construct an infinite word $w \in L$ with a given letter frequency $q \in \mathbb{Q}$, we use again the method described in Section 6.2. We write $q = \frac{a}{b}$ with $a$ coprime to $b$. For increasing values of $k$, we look for a $(k \times b)$-uniform morphism $h : \Sigma_e^* \to \Sigma_s^*$ producing (infinite) words in $L$ such that $|h(i)|_0 = k \times a$ for every $i \in \Sigma_e$.

Consider the 8-uniform morphism $m : \Sigma_3^* \longrightarrow \Sigma_4^*$ defined by

$$m(0) = 01232103,$$
$$m(1) = 01230323,$$
$$m(2) = 01210321.$$

To get the bound $f_{\min} \leqslant \frac{883}{3215}$ in Theorem 9.1, we found a square-free morphism $h^+ : \Sigma_3^* \longrightarrow \Sigma_3^*$ such that $h^+ = m^+ \circ m$ where $m^+ : \Sigma_4^* \longrightarrow \Sigma_3^*$ is a 3215-uniform morphism. To get the bound $f_{\max} \geqslant \frac{255}{653}$ in Theorem 9.1, we found a square-free morphism $h^- : \Sigma_3^* \longrightarrow \Sigma_3^*$ such that $h^- = m^- \circ m$ where $m^- : \Sigma_4^* \longrightarrow \Sigma_3^*$ is a 9142-uniform morphism ($9142 = 14 \times 653$). We need a result of Crochemore [17] saying that a uniform morphism is square-free if the image of every square-free word of length 3 is square-free. The software **mreps** [46] written by Kucherov et al. can test if a word is square-free in linear time. We used it to prove that $h^-$ and $h^+$ are square-free by checking that $h^-(w)$ and $h^+(w)$ are square-free, where $w = 010201210120212$ is square-free and contains every ternary square-free words of length 3 as factors. Checking the image of $w$ is faster than checking the images of the 12 ternary square-free words of length 3 because **mreps** runs in linear time. Since the morphisms $h^-$ (resp. $h^+$) are square-free, we obtain an exponential lower bound for ternary square-free words with letter frequency $\frac{883}{3215}$ (resp. $\frac{255}{653}$).

For each positive statement in Theorem 9.3 (odd items), we found a uniform morphism $h : \Sigma_3^* \longrightarrow \Sigma_2^*$ such that for every $\left(\frac{7}{4}^+\right)$-free ternary word $w$, $h(w)$ has the corresponding properties of repetition-freeness and letter frequency.

## 9.4 Dejean's conjecture and letter frequencies

Based on numerical evidences, we propose another strong form of Dejean's conjecture involving letter frequencies:

**Conjecture 9.1.**

1. *For every $k \geqslant 5$, there exists an infinite $\left(\frac{k}{k-1}^+\right)$-free word over $\Sigma_k$ with letter frequency $\frac{1}{k+1}$.*

2. *For every $k \geqslant 6$, there exists an infinite $\left(\frac{k}{k-1}^+\right)$-free word over $\Sigma_k$ with letter frequency $\frac{1}{k-1}$.*

It is easy to see that the values $\frac{1}{k+1}$ and $\frac{1}{k-1}$ in Conjecture 9.1 would be best possible. For $\left(\frac{5}{4}^+\right)$-free words over $\Sigma_5$, we obtain $f_{\max} < \frac{103}{440} = 0.23409090\cdots < \frac{1}{4}$ using the reduced suffix cover $\{0, 01, 012, 0123, 012341, 401234, 4301234\}$. That is why Conjecture 9.1.(2) is stated with $k \geqslant 6$.

# Chapter 10

# Conclusion

In this part, we mainly presented positive results in combinatorics on words. More precisely, we construct infinite words on small alphabets with properties such as avoiding different kind of repetitions and/or patterns. We also consider the problem of minimizing and maximizing the frequency of one letter in an infinite repetition-free word.

The general idea is to obtain an infinite word with the desired property as the image of an infinite repetition-free word by a uniform morphism. We wrote a computer program that finds these morphisms.

The most natural application of this method is to find upper bounds on generalized repetition thresholds. We obtained the exact value of six generalized repetition threshold $R(k,\ell)$ with small alphabet size $k$ and prefix size $\ell$. Proving lower bounds on $R(k,\ell)$ is rather easy using backtracking, and we can be quite confident that our lower bounds are tight. Our upper bounds always follow from results about the (classical) repetition threshold $R_k = R(k,1)$. This sounds like a limitation to the method, since Dejean's conjecture is settled for $k \leqslant 11$ only. Actually the main problem is the computing power needed to find a morphism, which grows very fast with $k$ and/or $l$.

Then we used this method in pattern avoidance, and completed the determination of the avoidability index of all ternary patterns. A major open problem in pattern avoidance is whether the avoidability index of a pattern is computable. Unfortunately our method does not solve this problem, since it must deal with repetitions. For example, it cannot prove that $\mu(ABWACXBAYBCZCA) = 4$. It would be interesting to know if this method can decide the 2-avoidability of a given pattern, assuming Dejean's conjecture is true.

Concerning the extremal letter frequencies in repetition-free words, we obtained some precise precise bounds and two exact values: The maximal letter frequency in infinite ternary square-free words is $\frac{255}{653}$ and infinite $(\frac{8}{5}^+, 3)$-free binary words have equal letter frequencies. We also found six new values $x$ such that the minimal letter frequencies in $x$-free and $x^+$-free binary words are distinct. We did not have enough time to program a more automated way of detecting such values.

To conclude, every result in this part is due to some progress in algorithmic devoted to combinatorics on words. Once we had an efficient enough algorithm and an implementation, we ran it on a maximum of instances. This gave us ideas about when and why the method fails or becomes impracticable.

# Bibliography


[1] I. Algor and N. Alon. The star arboricity of graphs, *Discrete Math.* **75** (1989) 11–12.

[2] K. Appel and W. Haken. Every planar map is four colorable. I. Discharging, *Illinois J. Math.* **21(3)** (1977) 429–490.

[3] K. Appel, W. Haken, and J. Koch. Every planar map is four colorable. II. Reducibility, *Illinois J. Math.* **21(3)** (1977) 491–567.

[4] K.A. Baker, G.F. McNulty, and W. Taylor. Growth Problems for Avoidable Words, *Theoret. Comput. Sci.* **69** (1989) 319–345.

[5] D.R. Bean, A. Ehrenfeucht, and G.F. McNulty. Avoidable Patterns in Strings of Symbols, *Pacific J. of Math.* **85** (1979) 261–294.

[6] J. Berstel. Axel Thue's Papers on Repetitions in Words: a Translation. Number 20 in Publications du Laboratoire de Combinatoire et d'Informatique Mathématique. Université du Québec à Montréal, February 1995.

[7] P. Boiron, E. Sopena, and L. Vignal. Acyclic improper colourings of graphs, *J. Graph Theory* **32** (1999) 97–107.

[8] O.V. Borodin. On acyclic colorings of planar graphs, *Discrete Math.* **25** (1979) 211–236.

[9] O.V. Borodin, D.G. Fon-Der Flaass, A.V. Kostochka, A. Raspaud, and E. Sopena. Acyclic list 7-coloring of planar graphs, *J. Graph Theory* **40(2)** (2002) 83–90.

[10] O.V. Borodin, A.V. Kostochka, J. Nešetřil, A. Raspaud, and E. Sopena. On the maximum average degree and the oriented chromatic number of a graph, *Discrete Math.* **206** (1999) 77–89.

[11] O.V. Borodin, A.V. Kostochka, and D.R. Woodall. Acyclic colorings of planar graphs with large girth, *J. London Math. Soc.* **60** (1999) 344–352.

[12] F.-J. Brandenburg. Uniformly growing $k$-th power-free homomorphisms, *Theoret. Comput. Sci.* **23** (1983) 69–82.

[13] J. Cassaigne. Motifs évitables et régularité dans les mots, Thèse de Doctorat, Université Paris VI, Juillet 1994.

[14] C. Choffrut and J. Karhumäki. Combinatorics of words, In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, Vol. 1, pp. 329–438. Springer-Verlag, 1997.

[15] R.J. Clark. Avoidable formulas in combinatorics on words, Ph.D. Thesis, University of California, Los Angeles (2001).

[16] B. Courcelle. The monadic second order logic of graphs VI: on several representations of graphs by relational structures, *Discrete Appl. Math.* **54(2-3)**, 117–149.

[17] M. Crochemore. Sharp characterizations of squarefree morphisms, *Theoret. Comput. Sci.* **18** (1982) 221–226.

[18] J.D. Currie. Open problems in pattern avoidance, *Amer. Math. Monthly* **100** (1993) 790–793.

[19] F. Dejean. Sur un théorème de Thue, *J. Combin. Theory. Ser. A* **13** (1972) 90–99.

[20] F.M. Dekking. On repetitions of blocks in binary sequences, *J. Combin. Theory. Ser. A* **20** (1976) 292–299.

[21] R. C. Entringer, D. E. Jackson, and J. A. Schatz. On nonrepetitive sequences, *J. Combin. Theory. Ser. A* **16** (1974) 159–164.

[22] J. Fiala, K. Jansen, V.B. Le, and E. Seidel. Graph Subcolorings: Complexity and Algorithm, *SIAM J. Discrete Math.* **16(4)** (2003) 635–650.

[23] J. Fiala and V.B. Le. The subchromatic index of graphs, *KAM Series* **655** (2004). `http://kam.mff.cuni.cz/~kamserie/serie/clanky/2004/s655.ps`

[24] A.S. Fraenkel and R.J. Simpson. How many squares must a binary sequence contain? *Electronic J. Combinatorics* **2** (1995) #R2.

[25] E. Fried. On homogeneous tournaments, *Combinatorial theory and its applications* **II** (1970) 467–476.

[26] M.R. Garey, D.S. Johnson, and L.J. Stockmeyer. Some simplified NP-Complete graph problems, *Theoret. Comput. Sci.* **1(3)** (1976) 237–267.

[27] J. Gimbel and C. Hartman. Subcolorings and the subchromatic number of a graph, *Discrete Math.* **272** (2003) 139–154.

[28] D. Gonçalves. Caterpillar arboricity of planar graphs, *Discrete Math.* - To appear.

[29] D. Gonçalves. Étude de différents problèmes de partition de graphes, Ph.D. Thesis, LaBRI, Université Bordeaux I (2005).

[30] D. Gonçalves and P. Ochem. On some arboricities in planar graphs, In *Proceedings of the 7th International Colloquium on Graph Theory (ICGT'05)*, Hyères, France, July 12-16 2005.

[31] H. Grötzsch. Ein Dreifarbensatz fur dreikreisfreie Netze auf der Kugel, *Math. Nat. Reihe* **8** (1959) 390–408.

[32] B. Grünbaum. Acyclic colorings of planar graphs, *Israel J. Math.* **14** (1973) 390–408.

[33] A. Gyárfás and D.B. West. Multitrack Interval Graphs, *Congressus Numerantium* **109** (1995) 109–116.

[34] S.L. Hakimi, J. Mitchem, and E. Schmeichel. Star arboricity of graphs, *Discrete Math.* **149** (1996) 93–98.

[35] A.M. Hobbs. Computing edge-toughness and fractional arboricity, *Contemp. Math.* **89** (1989) 89–106.

[36] I. Holyer. The NP-completeness of edge coloring, *SIAM J. Comput.* **10** (1981) 718–720.

[37] L. Ilie, P. Ochem, and J.O. Shallit. A generalization of repetition threshold, *Theoret. Comput. Sci.* **345** (2005) 359–369.

[38] Information System on Graph Class Inclusions,
`http://wwwteo.informatik.uni-rostock.de/isgci/index.html`

[39] J. Karhumäki and J.O. Shallit. Polynomial versus exponential growth in repetition-free binary words, *J. Combin. Theory Ser. A* **105** (2004) 335–347.

[40] R. Kolpakov, G. Kucherov, and Y. Tarannikov. On repetition-free binary words of minimal density, *Theoret. Comput. Sci.* **218** (1999) 161–175.

[41] A.V. Kostochka and L.S. Mel'nikov. Note to the paper of Grünbaum on acyclic colorings, *Discrete Math.* **14** (1976) 403–406.

[42] A.V. Kostochka, E. Sopena, and X. Zhu. Acyclic and oriented chromatic numbers of graphs, *J. Graph Theory* **24(4)** (1997) 331–340.

[43] A.V. Kostochka and D.B. West. Every outerplanar graph is the union of two interval graphs, *Congressus Numerantium* **139** (1999) 5–8.

[44] J. Kratochvíl. A Special Planar Satisfiability Problem and a Consequence of its NP-Completeness, *Discrete Appl. Math.* **52** (1994) 233-252.

[45] G. Kucherov, P. Ochem, and M. Rao. How many square occurrences must a binary sequence contain ? *Electron. J. Comb.* **10(1)** (2003) #R12.

[46] `http://mreps.loria.fr/`

[47] H. Le and V.B. Le. The NP-completeness of $(1, r)$-subcolorability of cubic graphs, *Inform. Process. Lett.* **81(3)** (2002) 157–162.

[48] M. Lothaire. *Algebraic Combinatorics on Words*. Cambridge Univ. Press, 2002.

[49] T.H. Marshall. On $\mathcal{P}$-universal Graphs, *KAM Series* **510** (2001). `http://kam.mff.cuni.cz/~kamserie/serie/clanky/2001/s510.ps`

[50] T.H. Marshall. Antisymmetric flows on planar graphs, *KAM Series* **652** (2003). `http://kam.mff.cuni.cz/~kamserie/serie/clanky/2003/s652.ps`

[51] M. Montassier, P. Ochem, and A. Raspaud. On the acyclic choosability of graphs, *J. Graph Theory* **51(4)** (2006) 281–300.

[52] J. Moulin-Ollagnier. Proof of Dejean's conjecture for alphabets with $5, 6, 7, 8, 9, 10$ and 11 letters, *Theoret. Comput. Sci.* **95** (1992) 187–205.

[53] M. Mirzakhani. A small non-4-choosable planar graph, *Bull. Inst. Combin. Appl.* **17** (1996) 15–18.

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

[55] J. Nešetřil and A. Raspaud. Antisymmetric Flows and Strong Colorings of oriented planar graphs, *Ann. Inst. Fourier* **49** (1999) no.3, 1037–1056.

[56] J. Nešetřil, A. Raspaud, and E. Sopena. Colorings and girth of oriented planar graphs, *Discrete Math* **165-166** (1997) 519–530.

[57] P. Ochem. Oriented colorings of triangle-free planar graphs, *Inform. Process. Lett.* **92(2)** (2004) 71–76.

[58] P. Ochem. A generator of morphisms for infinite words, In *Proceedings of the Workshop on Word Avoidability, Complexity, and Morphisms*, Turku, Finland, July 17 2004. LaRIA Technical Report 2004-07, pp. 9–14.
*RAIRO - Theoretical Informatics and Applications* **40** (2006) 427–441.

[59] P. Ochem. Negative results on acyclic improper colorings, In *Proceedings of the 2005 European Conference on Combinatorics, Graph Theory and Applications* (EuroComb '05) Berlin, Germany, September 5-9 2005, *DMTCS Proceedings* pp. 357–362.

[60] P. Ochem. Letter frequency in infinite repetition-free words, *Theoret. Comput. Sci.* 380 (2007) 388–392.

[61] J.-J. Pansiot. A propos d'une conjecture de F. Dejean sur les répétitions dans les mots, *Disc. Appl. Math.* **7** (1984) 297–311.

[62] H. Prodinger and F.J. Urbanek. Infinite 0–1-sequences without long adjacent identical blocks, *Discrete Math.* **28** (1979) 277–289.

[63] N. Rampersad, J.O. Shallit, and M.-w. Wang. Avoiding large squares in infinite binary words, *Theoret. Comput. Sci.* **339** (2005) 19–34.

[64] A. Raspaud and E. Sopena. Good and semi-strong colorings of oriented planar graphs, *Inform. Process. Lett.* **51(4)** (1994) 171–174.

[65] C. Richard and U. Grimm. On the entropy and letter frequencies of ternary square-free words, *Electron. J. Comb.* **11** (2004) #R14

[66] P. Roth. Every Binary Pattern of Length Six is Avoidable on the Two-Letter Alphabet, *Acta Inf.* **29(1)** (1992) 95–107.

[67] D.P. Sanders. Linear Algorithms for Graphs of Tree-width at Most Four, Program of Algorithms, Combinatorics, and Optimization, PhD thesis, Georgia Institute of Technology, Atlanta, Georgia, June 1993.

[68] J.O. Shallit. Simultaneous avoidance of large squares and fractional powers in infinite binary words, *Int. J. Found. Comput. Sci.* **15** (2004) 317–327. Preprint available at `http://www.arxiv.org/abs/math.CO/0304476`, 2003.

[69] E. Sopena. The chromatic number of oriented graphs, *J. Graph Theory* **25** (1997) 191–205.

[70] E. Sopena. Oriented graph coloring, *Discrete Math.* **229(1-3)** (2001) 359–369.

[71] E. Sopena. There exist oriented planar graphs with oriented chromatic number at least sixteen, *Inform. Process. Lett.* **81(6)** (2002) 309–312.

[72] Y. Tarannikov. The minimal density of a letter in an infinite ternary square-free word is 0.2746..., *J. Integer Sequences* 5(2):Article 02.2.2 (2002).

[73] T.C. Shermer. On rectangle visibility graphs. III. External visibility and complexity, In *Proceedings of the 8th Canadian Conference on Computational Geometry (CCCG'96)*, Turku, Finland, July 17 2004. (1996) 234–239

[74] C. Thomassen. Every planar graph is 5-choosable, *J. Combin. Theory. Ser. B* **62** (1994) 180–181.

[75] A. Thue. Über unendliche Zeichenreihen, *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **7** (1906) 1–22. Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell, editor, Universitetsforlaget, Oslo (1977) 139–158.

[76] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **1** (1912) 1–67. Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell, editor, Universitetsforlaget, Oslo (1977) 413–478.

[77] L. Vignal. Homomorphismes et colorations de graphes orientés, Ph.D. Thesis, LaBRI, Université Bordeaux I (1997).

[78] V.G. Vizing. On an estimate of the chromatic class of a p-graph, *Diskret. Analiz* **3** (1964) 23–30. (in Russian)

[79] M. Voigt. List colourings of planar graphs, *Discrete Math.* **120** (1993) 215–219.

[80] M. Voigt. A non-3-choosable planar graph without cycles of length 4 and 5, Manuscript.

[81] M. Voigt and B. Wirth. On 3-colorable non-4-choosable planar graphs, *J. Graph Theory* **24** (1997) 233–235.

[82] R. Xu. Note on acyclic colorings of graphs, *Combinatoria* **72** (2004) 235–239