LaBRI

université de BORDEAUX

THÈSE

PRÉSENTÉE À

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE
DE MATHÉMATIQUES ET D'INFORMATIQUE

par **Claire Hilaire**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : INFORMATIQUE

Structure de Graphes, Mineurs et Arbres Induits

Structure of Graphs: Minors and Induced Trees

Soutenue le 4 juillet 2023

Membres du jury :

M. Éric Colin de Verdière	Directeur de Recherche, CNRS, Université Gustave Eiffel	Examinateur
M. David Eppstein	Distinguished Professor, University of California	Rapporteur
M. Cyril Gavoille	Professeur, Université de Bordeaux	Directeur
M. Gwenaël Joret	Chargé de Recherche, Université Libre de Bruxelles	Rapporteur
Mme Zi-Xia Song	Professeure, University of Central Florida	Rapportrice
M. Marc Zeitoun	Professeur, Université de Bordeaux	Examinateur

Membre invitée :

Mme Marthe Bonamy	Chargée de Recherche, CNRS, Université de Bordeaux	Co-encadrante

Titre : Structure de graphes, mineurs et arbres induits

Résumé : Cette thèse traite des questions structurelles de la théorie des graphes qui découlent de motivations algorithmiques. En effet, les problèmes de graphes NP-complets sont typiquement plus faciles dès lors qu'on restreint la classe de graphes. Sur les arbres, qui sont des graphes connexes sans cycles, la programmation dynamique donne souvent un algorithme linéaire. Tous les graphes ne sont pas des arbres, mais il est possible de quantifier à quel point un graphe est plus complexe qu'un arbre. Un outil standard pour cela est la largeur arborescente, qui est profondément liée à la théorie des mineurs, comme nous allons voir dans le premier chapitre. Une autre stratégie consiste à obtenir une structure arborescente en supprimant le moins de sommets possible, comme nous le verrons dans le deuxième chapitre.

Le premier chapitre porte sur l'idée de graphes universels pour la relation de mineur. Un graphe est mineur-universel pour une famille de graphes s'il contient comme mineur chaque graphe de la famille. Un résultat célèbre de Robertson, Seymour et Thomas en 1994 montre que la grille $2n \times 2n$ est mineur-universelle pour la famille des graphes planaires à n sommets. Une façon d'affiner ce résultat est de trouver des sous-classes de graphes planaires à n sommets qui admettent une grille mineure-universelle plus petite. On dit que qu'un graphe planaire G admet un dessin sur une grille s'il existe un plongement de G sur une grille, tel que les sommets de G sont des sommets de la grille, et ses arêtes sont des suites de segments dont les jonctions sont placées sur des sommets des la grille. La taille d'un tel dessin est le nombre de sommets de la grille. L'idée est d'établir une relation entre la taille d'un dessin sur une grille d'un graphe H, et la taille de la plus petite grille qui admet H comme mineur. Une autre direction consiste à le généraliser à une classe plus grande: les graphes de genre borné. Plus précisément, nous montrons qu'il existe un graphe de genre g mineur-universel pour les graphes à n sommets de genre g, et dont la taille est polynomiale en n et g.

Le deuxième chapitre se concentre sur la recherche d'une grande forêt (union de plusieurs arbres) induite dans un graphe. Un cas intéressant est celui où la forêt en question est un unique chemin. En 2017, Esperet, Lemoine et Maffray ont conjecturé que tout graphe k-dégénéré avec un chemin à n sommets admet un chemin induit à $\log n^{\Omega_k(1)}$ sommets. Nous prouvons que c'est le cas pour plusieurs classes de graphes dégénérés, y compris la classe des graphes excluant un graphe comme mineur topologique. Dans le cas d'une forêt, le théorème d'Erdős-Pósa stipule que si un graphe n'a pas k cycles disjoints, alors il admet un ensemble de $O(k \log k)$ sommets dont l'élimination produit une forêt, appelé *feedback vertex set*. Nous obtenons une propriété similaire pour les graphes sans sous-graphe $K_{t,t}$ et sans k cycles indépendants (disjoints et sans arête entre eux). La contrepartie de cette généralisation est un *feedback vertex set* dont la taille est logarithmique en la taille du graphe, et cette borne peut être atteinte.

Mots-clés : graphe, structure, mineur, arbre, planaire, surface

Laboratoire d'accueil : 351, cours de la Libération F-33405 Talence cedex

Title: Structure of Graphs: Minors and Induced Trees

Abstract: This thesis deals with structural graph theory questions that stem from algorithmic motivations. Indeed, NP-complete graph problems are typically easier on restricted graph classes. On trees, which are connected graphs without cycles, dynamic programming often yields a linear algorithm. Not every graph is a tree, but we can measure how much more complex the graph is. A standard tool is the treewidth which is deeply connected to minor theory, as discussed in the first chapter. Another strategy is to obtain a much simpler structure by deleting as few vertices as possible, as discussed in the second chapter.

The first chapter concerns the idea of minor-universality. A graph is minor-universal for a family of graphs if it contains as a minor every graph in the family. A celebrated result of Robertson, Seymour, and Thomas in 1994 shows that the $2n \times 2n$-grid is minor-universal for the family of n-vertex planar graphs. One direction to refine this result is to find some subclasses of n-vertex planar graphs that admit a smaller minor-universal grid. A poly-line grid drawing of a planar graph G is a planar embedding of G such that its vertices are mapped on the vertices of the grid and its edges are curves made up of line segments connecting on vertices of the grid. The size of such drawing is the order of the grid. The idea is to establish a relation between the size of a poly-line grid drawing of a graph H and the order of the smallest grid that admits H as a minor. Another direction is to extend the result from planarity to bounded genus. More precisely, we show that there is a graph of genus g minor-universal for the n-vertex graphs of genus g, and whose order is polynomial in n and g.

The second chapter focuses on finding a large induced forest in a graph. An interesting case is when the forest is required to be a single path. In 2017, Esperet, Lemoine, and Maffray conjectured that every k-degenerate graph with an n-path admits an induced $\log n^{\Omega_k(1)}$-path. We prove that this holds for several degenerate classes, including the class of topological-minor-free graphs. In the case of a forest, the Erdős-Pósa theorem states that if a graph does not have k vertex-disjoint cycles, then it admits a set of $O(k \log k)$ vertices whose removal yields a forest, called the feedback vertex set. We obtain a similar property for the graphs with no $K_{t,t}$-subgraph and no k independent cycles (vertex-disjoint and no edge between them). The trade-off for this generalization is a feedback vertex set whose size depends in the logarithm of the order, and this is sharp.

Keywords: graph, structure, minor, tree, planar, surface

Hosting Laboratory: 351, cours de la Libération F-33405 Talence cedex

Contents

Résumé étendu en français

Cette thèse traite des questions structurelles de la théorie des graphes qui découlent de motivations algorithmiques. En effet, les problèmes de graphes NP-complets sont typiquement plus faciles dès lors qu'on restreint la classe de graphes. Sur les arbres, qui sont des graphes connexes sans cycles, la programmation dynamique donne souvent un algorithme linéaire. Tous les graphes ne sont pas des arbres, mais il est possible de quantifier à quel point un graphe est plus complexe qu'un arbre. Un outil standard pour cela est la largeur arborescente, qui est profondément liée à la théorie des mineurs. Nous abordons ici deux grands axes: d'un côté l'étude des mineurs de graphes, et de l'autre, la recherche de structure arborescente dans les graphes.

Définitions

Cette thèse porte sur la théorie structurelle des graphes, des objets théoriques composés de *sommets* (représentés par des points) reliés ou non par des *arêtes*. Lorsqu'une arête lie deux sommets, on dit qu'ils sont *adjacents* ou voisins. Un graphe est *connexe* si pour tout couple de sommets, il y a un chemin (une suite d'arêtes avec à chaque étape un sommet en commun) qui les relie. On mesure ici la *taille* d'un graphe par son nombre de sommets.

On s'intéresse particulièrement aux relations entre les graphes. Un graphe H est un *sous-graphe induit* d'un graphe G si on peut le construire à partir de G en supprimant uniquement des sommets (et les arêtes qui n'ont plus leurs deux extrémités). H est un *sous-graphe* de G si on autorise aussi la suppression d'arêtes de G. Enfin, H est un *mineur* de G si on autorise en plus la contraction d'arêtes (on fusionne les deux extrémités u et v d'une arête en un unique sommet, dont le voisinage est l'ensemble des voisins de u et v). On dit aussi que G est un *majeur* de H. Ces relations sont recapitulées sur la figure 1.

Une *classe* de graphes est un ensemble de graphes avec une caractéristique commune. Cette caractéristique se traduit souvent sous la forme de l'exclusion d'un graphe comme mineur ou sous-graphe.

Une classe de graphes très étudiée est celle des *arbres* (graphes connexes sans

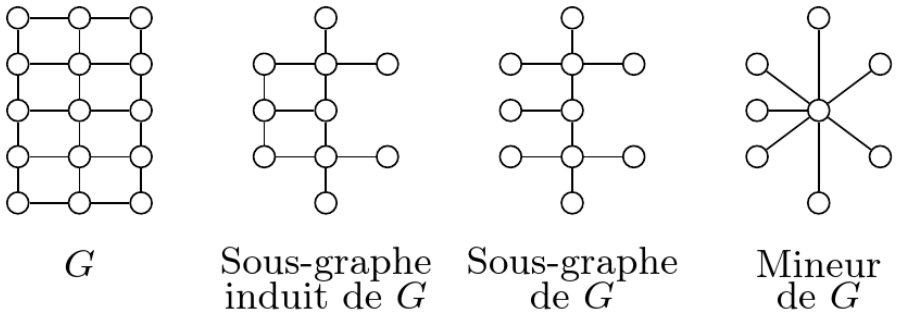


Figure 1: de gauche à droite, un graphe G, un sous-graphe induit, un sous-graphe non induit, et un mineur qui n'est pas un sous-graphe de G.

cycles comme sous-graphe), car beaucoup de problèmes de graphe NP-complets ont une solution linéaire en la taille du graphe lorsque celui-ci est un arbre. De ce fait, trouver une structure arborescente dans un graphe peut permettre d'obtenir des solutions efficaces pour ces problèmes.

Un majeur pour les gouverner tous

Un graphe est *mineur-universel* pour une classe de graphes s'il contient comme mineur chaque graphe de la classe. Trouver un tel graphe permet en quelque sorte de capturer des caractéristiques de toutes la classe en un unique graphe, plus facile à manipuler. Notamment, un résultat célèbre de Robertson, Seymour et Thomas en 1994 montre que la grille $2n \times 2n$ est mineur-universelle pour la classe des graphes planaires (graphes que l'on peut dessiner dans le plan sans croisement) à n sommets [RST94]. Ce résultat a permis par la suite de borner la largeur arborescente[1] d'un graphe qui exclut un planaire quelconque comme mineur, en se focalisant uniquement sur les graphes excluant une grille comme mineur.

Les dessins de graphes

Une façon d'affiner ce résultat est de trouver des sous-classes de graphes planaires à n sommets qui admettent une grille mineure-universelle plus petite. On dit qu'un graphe planaire G admet un dessin sur une grille s'il existe un plongement de G sur une grille, tel que les sommets de G sont des sommets de la grille, et ses arêtes sont des suites de segments dont les jonctions sont placées sur des sommets de la grille. La taille d'un tel dessin est le nombre de sommets de la grille. L'idée est d'établir une relation entre la taille d'un dessin d'un graphe H sur une grille, et

[1] paramètre qui quantifie une certaine complexité du graphe par rapport à un arbre

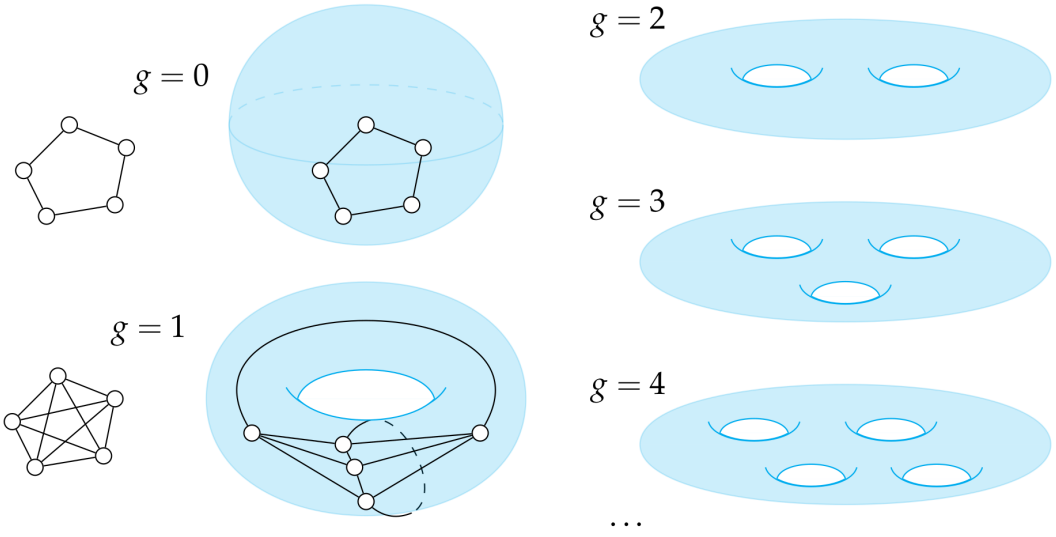


Figure 2: À gauche, un graphe planaire et son plongement sur la sphère (genre 0), et un graphe non planaire plongeable sur le tore (genre 2). À droite, des exemples de surfaces de genre croissant.

la taille de la plus petite grille qui admet H comme mineur, comme cela a été fait par Dieng et Gavoille [DG20].

Les graphes de genre borné

Plutôt que d'affiner le résultat, une autre direction consiste à le généraliser à une classe plus grande qui inclue les graphes planaires, telle que celle de graphes de genre borné. En effet, un graphe est planaire s'il peut être dessiné sans croisement dans le plan ou de façon équivalente sur une sphère. Un graphe est de genre g s'il peut être dessiné sans croisement sur une surface ressemblant à une sphère avec g trous.

Nous prouvons qu'il existe un graphe de genre g mineur-universel pour les graphes à n sommets de genre g, et dont la taille est polynomiale en n et g. Le résultat s'applique également aux surfaces non-orientables.

Les forêts cachées

Le deuxième chapitre se concentre sur la recherche d'une grande forêt (union d'arbres) induite dans un graphe.

Un cas intéressant est celui où la forêt en question est un unique chemin.

Lorsqu'un graphe a un chemin induit de taille n, il a un chemin de taille n, mais la réciproque n'est pas toujours vraie : c'est faux notamment lorsque le graphe a un grand[2] $K_{t,t}$ comme sous-graphe. Une très vaste classe qui évite ce cas de figure est la classe des graphes dits t-dégénérés. En 2017, Esperet, Lemoine et Maffray ont conjecturé que tout graphe k-dégénéré avec un chemin à n sommets admet un chemin induit à $(\log n)^{\Omega_k(1)}$ sommets [ELM17].

Nous prouvons que c'est effectivement le cas pour plusieurs classes de graphes dégénérés, dont une très grande, celle des graphes excluant un graphe comme mineur topologique.

Dans le cas où c'est une forêt induite que l'on cherche, le théorème d'Erdős-Pósa [EP65] stipule que si un graphe n'a pas k cycles disjoints, alors il admet un ensemble de $O(k \log k)$ sommets dont l'élimination produit une forêt, appelé *feedback vertex set*. Nous obtenons une propriété similaire pour les graphes sans sous-graphe $K_{t,t}$ et sans k cycles indépendants (disjoints et sans arête entre eux). La contrepartie de cette généralisation est un *feedback vertex set* dont la taille est logarithmique en la taille du graphe, et cette borne peut être atteinte.

[2] Kt, t: deux ensembles disjoints de t sommets chacun, avec toutes les arêtes possibles entre eux et aucune à l'intérieur de chaque ensemble

Introduction

The results in this thesis all belong to the field of structural graph theory. Let us start with the basics.

Graph theory for problem solving

Imagine you want to organize a gathering with as many of people as possible among your colleagues: Alice, Bob, Caroline, Daniel, Ellen and Frank. However, you know that there are some tensions between Alice and Bob, so you can not invite them both at the party. You also heard that Daniel and Ellen had a fight, and so on. A convenient way to represent the situation is with a schema as in Fig. 3:

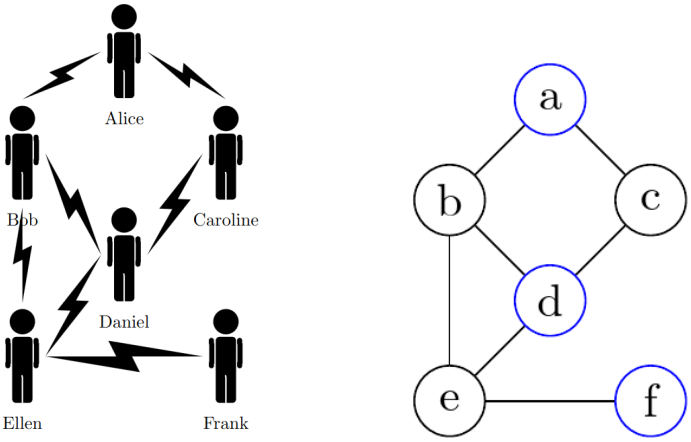


Figure 3: Schema representing conflictual relations between the colleagues (left) and the graph modeling the situation (right), with blue vertices corresponding to a maximum independent set.

To solve our problem, the only thing that matters is the pairwise relations (in conflict or not in conflict) among your colleagues. The mathematical object modeling this kind of relation within a given set is called a *graph*.

A graph is described by a set of objects called *vertices*, here the set of colleagues (using only their initial to simplify the notation), and a set of pair of vertices, called *edges*, that represents a relation between the vertices. In our example that would be the conflicts among your colleagues.

Basic vocabulary and notation

Let G be a graph. We denote by $V(G)$ the set of vertices of G, in our example that would be $\{a, b, c, d, e, f\}$, and by $E(G)$ the set of edges (in the example $\{(a, b), (a, c), (c, d), \dots\}$). The number of vertices in a graph is called its *order* and is denoted $|V(G)|$ or $|G|$ as a shorthand. Two vertices are *adjacent* or *neighbors* if they share an edge, and the set of neighbors of a vertex is called its *neighborhood*. The number of edges incident to a vertex is called its *degree*, thus corresponds to the size of its neighborhood. In our example, the vertex c has degree 3, and its neighborhood is $\{a, d, e\}$.

A *path* in a graph is a sequence of edges which joins a sequence of pairwise distinct vertices. In this example, $c - a - b - d - e - f$ is a path. Observe that this path covers all the vertices of the graph: we say that it is a *Hamiltonian path*. A *cycle* is a path whose extremities are adjacent. In our example, $a - b - d - c$ is a cycle. If the cycle covers all the vertices, it is a *Hamiltonian cycle*, and a graph containing a Hamiltonian cycle is called *Hamiltonian*. Observe that the graph of the example is not Hamiltonian. A *connected component* in a graph is an inclusion-maximal set of vertices such that we can go from a vertex to any other using a path. A graph is said *connected* if it has only one connected component. A set of vertices is *separating* if its removal results in a graph that is not connected.

Furthermore, a set of vertices such that there is no edge between any of them is called an *independent* set. On the example in Figure 3, the graph is connected. However, if we consider only the subset of vertices $\{a, d\}$, the set is not connected and forms in fact an independent set. It is not a maximal independent set as we can obtain a bigger one by adding f. The resulting independent set is not only maximal but also maximum, as there is no bigger independent set.

In the literature, some graphs can be infinite, have an edge with twice the same endpoint, multiple edges between the same pair of vertices, or directed edges (Alice dislikes Bob, but Bob does not mind Alice). However, in this thesis, we do not consider any such extension of the definition of a graph.

With this model of the situation, inviting as many colleagues as possible without tensions is thus equivalent to finding an independent set of maximum size in the graph representing the tensions among them. This problem, Maximum Independent Set (MIS for short), is a very well-known and studied problem[3].

[3] see https://en.wikipedia.org/wiki/Independent_set_(graph_theory).

Therefore, we moved from a very specific problem to a generic one with known strategies. Notice that for this instance, we could easily find a solution by hand (as represented in Fig. 3), but in general, we look for an algorithm that will work in various situations and save time in the future. For example, you could be an event planner and have to solve such problems in different contexts multiple times a week.

Moreover, MIS models various problems, so investing time and energy in finding a good strategy on this abstract model means that you obtain a good strategy for a whole category of problems with no other common point than that they correspond to maximising the size of a set of elements all pairwise "compatible". For example, in computer science, independent sets are useful for parallelizing processes, as we can identify which pairs can run in parallel and which cannot.

MIS is an NP-complete problem in general, so it would be surprising to obtain polynomial-time algorithms for it without further assumptions on the input graph. Fortunately, in most real-life world problems, the input graph is not arbitrary but satisfies some kind of structural properties.

Graph classes and containment relations

Graphs representing real-life situations are generally not completely chaotic, they follow patterns of some type. Let us go back to our party-planning problem. Since we organize it within a company, we can take into account its hierarchy: it could be that people are hesitant to go to a party with their direct boss (let us assume each person has exactly one direct boss). If there is no tension other than this one (say the company is very recent), then the model of the situation is a very simple graph, with no cycle, called a *tree*, as in Fig. 4.

A *graph class* or a *family of graphs* is a possibly infinite set of graphs. In most cases, graph classes are defined with a property shared by all the graphs of tree. The class of trees (or *forests* if the graph is not assumed to be connected: a forest is a graph where every connected component is a tree), is a famous one as a lot of difficult problems become easy to solve when considered within this class. This is a general pattern: knowing some structural properties on the input graph can help design more efficient algorithms to solve problems. Let us go back one final time to our example. Instead of trying to solve MIS on an arbitrary graph (where the problem is NP-complete), we focus now on solving it on a tree, where the problem can be solved in linear time, using a simple induction.

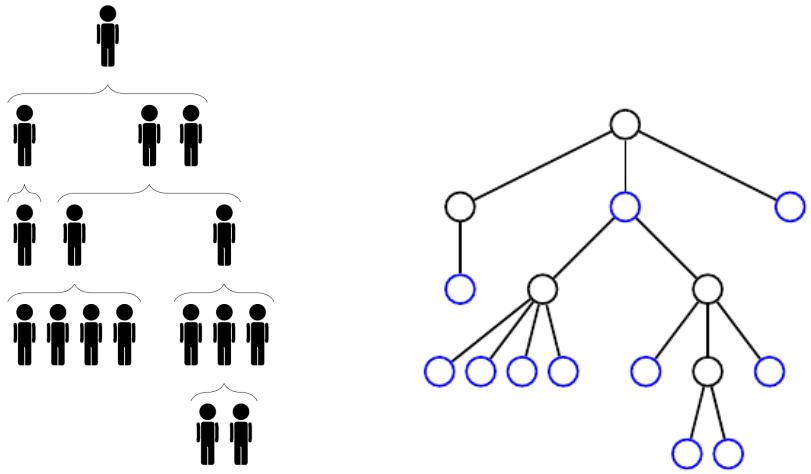


Figure 4: Hierarchy within the company (left) and the graph representing the conflicts (right), with blue vertices for the MIS.

Relations between graphs

Trees are described as connected graphs with no cycle. A common way to define a family of graphs is using a structure they contain or, more commonly, one that they do not contain. What does it mean, then, for a graph to contain a structure or not? In fact, there are several containment relations possible. Here are the four common relations we will use in this thesis, each one a special case of the next (see Fig. 5 for an example of each relation).

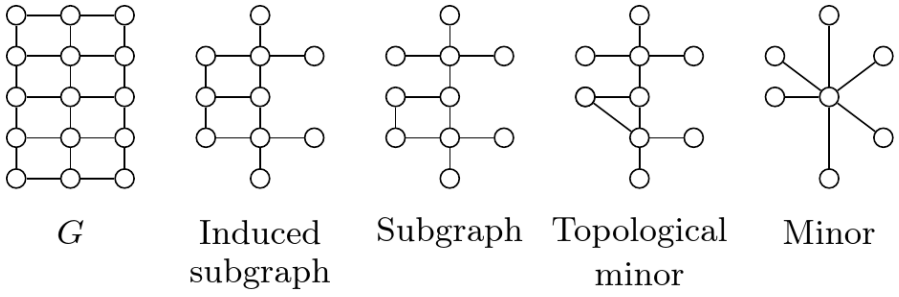


Figure 5: From left to right, a graph G, an induced subgraph of G that is not G, a subgraph of G that is not induced, a topological minor of G that is not a subgraph and a minor of G that is not topological.

- First of all, the *induced subgraph*. We say that H is an induced subgraph of G if $V(H)$ is a subset of $V(G)$ and for each pair of vertices of H, they are adjacent in H if and only if they are adjacent in G. In other words, we can

construct H from G by deleting a (possibly empty) set of vertices (and the edges incident to them).

- Then, if we relax the definition by allowing also deletion of edges, we say that H is a *subgraph* of G.

- An *edge-subdivision* consists in adding a new vertex in the middle of an edge, splitting the edge into two edges. The graph obtained by applying a (possibly empty) sequence of edge-subdivisions to a graph is called a *subdivision* of the initial graph. If G admits a subgraph isomorphic[4] to a subdivision of H, we say that H is a *topological minor* of G.

- Finally, an *edge-contraction* consists in merging the extremities of an edge into a single vertex, whose neighborhood is the union of the neighborhood of both extremities. If we can construct H from G by removing some vertices and edges and contracting some edges, then we say that H is a *minor* of G.

A common way to define a class is with the substructures (in the form of graphs) all the graphs of the class exclude. As we saw earlier, the class of trees can be defined as the graphs that exclude every cycle as (induced) subgraph. Observe that if we take a minor of a tree, we cannot create a cycle, so all the minors of a tree are also trees. We say that the class is *closed under taking minors*.

If we lower the restriction and only forbid cycles of odd length, then we get a larger class of graph: the *bipartite graphs*. A bipartite graph consist in two independent sets with edges between them. This time, if we take a minor of a bipartite graph, we could create an odd cycle, therefore the class is not closed under taking minor. However, any subgraph of a bipartite graph is bipartite, so this class is *closed under taking subgraphs*.

Let n_1, n_2 be two integers. The graph obtained by taking two sets of vertices, one of size n_1 and the other one of size n_2, and adding all edges between the first and the second is a *complete bipartite* graph, denoted K_{n_1,n_2}. The class of complete bipartite graphs is not closed under taking subgraphs, as removing any edge results in a graph that is bipartite but not complete. However, it is closed under taking induced subgraphs, or *hereditary*. A similar common hereditary graph class is that of *complete graphs*, or *cliques*. The complete graph on n vertices, noted K_n, is the graph with all possible edges. We can also define those two classes by forbidding induced subgraphs (see Fig. 6).

Notice that for the last two classes, we described them with a finite number of forbidden induced subgraphs whereas there are infinitely many cycles, one for each size, to exclude for the trees. If we define the trees by excluding minors

[4] equal, up to renaming the vertices.

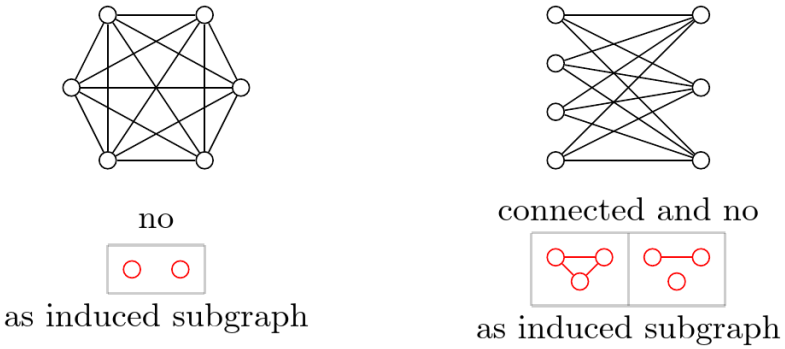


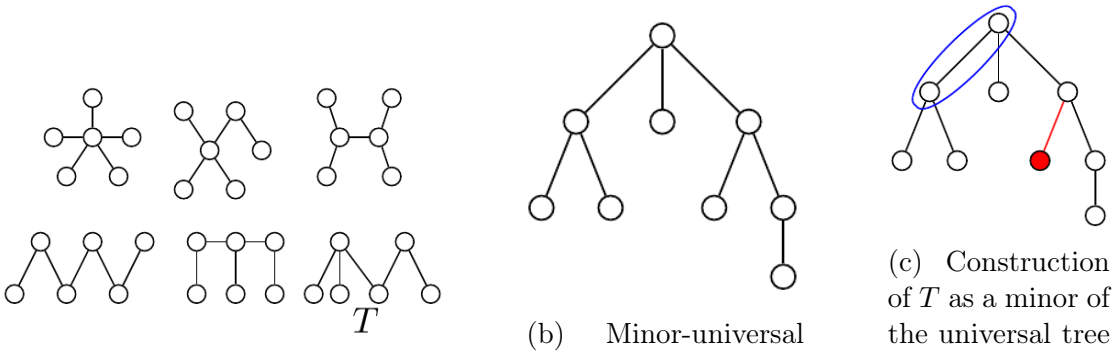
Figure 6: A complete graph K_6 and a bipartite complete graph $K_{4,3}$, with the forbidden induced subgraph.

instead of subgraphs, we reduce this set to a single one: the trees are the graphs that do not contain a triangle as a minor. Indeed, note that any longer cycle contains a triangle as a minor, and that there is no smaller cycle than a triangle. The seminal Robertson-Seymour Theorem [RS04], whose proof spans over twenty papers and two decades, states that every graph class closed under taking minors can be defined by a finite set of forbidden minors. In fact, a lot of classical graph classes are closed under taking minors and can thus be defined this way. Along the way, they got a result related to a topic we will study in more details: minor-universality.

Minor-universality

A graph is *universal* for a family of graphs if it contains all the graphs in the family, for some containment relation (subgraph, induced subgraph, minor, ...). Graph universality plays an important role in graph theory and graph algorithms, as it allows us to manipulate a single graph instead of the whole family to deduce properties. The goal is then to wisely choose the constraints of the universal graph to represent the family. The first chapter of this thesis focuses on the notion of *minor-universality*.

Let us take for example the family of trees on at most 6 vertices, as represented in Fig. 7. We can obtain each one of them as a minor of the tree represented on the left: this tree is minor-universal for the family of trees on at most 6 vertices. Even if it has more than 6 vertices, this minor-universal graph is also a tree. In fact, it is a smallest tree that is minor-universal for this family. Of course, we could have taken a minor-universal graph that is not a tree, like the complete graph K_6; It has fewer vertices, but it does not really capture the properties of the family (here, not having cycles).



(a) Trees on 6 vertices.

(b) Minor-universal tree.

(c) Construction of T as a minor of the universal tree for this family.

Figure 7: A tree (7b) that contains every tree on at most $n = 6$ vertices (7a) as a minor. Only the ones with exactly 6 vertices are represented, as the smaller ones are minors of those ones. For example, Fig. 7c represent how to obtain T as a minor of this universal tree, by removing the red vertex and contracting the edge circled in blue.

In Chapter 1, we explore this question of minor-universal graph for families of graphs closed under taking minors, such that the universal graph has similar properties as the family. To avoid manipulating infinite families (and thus infinite graphs), we parameterize by the number of vertices. Namely, we focus on the graphs that have at most n vertices for some integer n, and observe how the size of the universal graph increases as n increases. The graphs on which we study minor-universality correspond to another common graph class: the planar graphs and their generalizations.

Planar graphs

A graph is *planar* if it can be embedded on a plane surface such that no two edges cross each other. When such embedding is given, we call it a *plane* graph. A famous old puzzle, called the three utilities problem, relies on this notion. Suppose we want to place and connect in a city three houses to three utilities: gas, water and electricity. However, we want to connect them with paths that do not cross each other. This is in fact an impossible puzzle. The situation can be represented by $K_{3,3}$, one set for the houses, one for the utilities, and this graph was proved not planar. In fact, this graph and K_5 (represented Fig. 8) are the smallest non-planar graphs, up to taking minors. In other word, any non-planar graph contains one of those two graphs as minor. The class of planar graphs is closed under taking minors, and can be defined as the $\{K_{3,3}, K_5\}$-minor-free graphs.

Robertson, Seymour, and Thomas [RST94] found in 1984 a planar graph on $4n^2$ vertices that is minor-universal for all the planar graphs on at most n vertices.

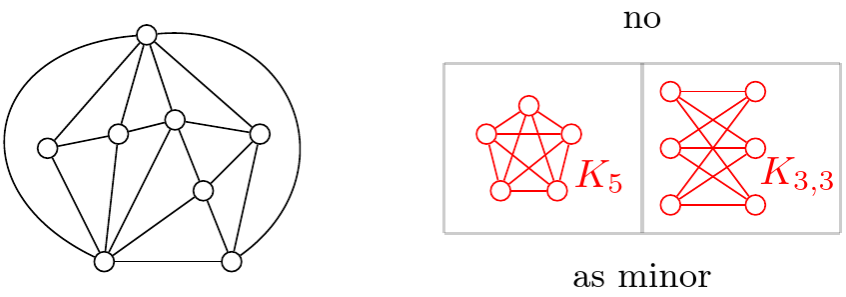


Figure 8: An example of planar graph (left) and the excluded minors K_5, $K_{3,3}$ (right).

There are two directions to refine this result: either by considering subfamilies of planar graphs for which there is a smaller minor-universal graph, or by considering a family that generalizes the notion of planarity. This generalization involves bounding a new parameter of a graph: its genus.

Genus of a graph

Planar graphs are defined as graphs that can be embedded on a plane surface (or equivalently on a sphere). Therefore the natural generalization is to consider graphs that can be embedded on more complex surfaces. Here we will only consider connected compact surfaces without boundary. The parameter to characterize the complexity of a surface is the *genus*.

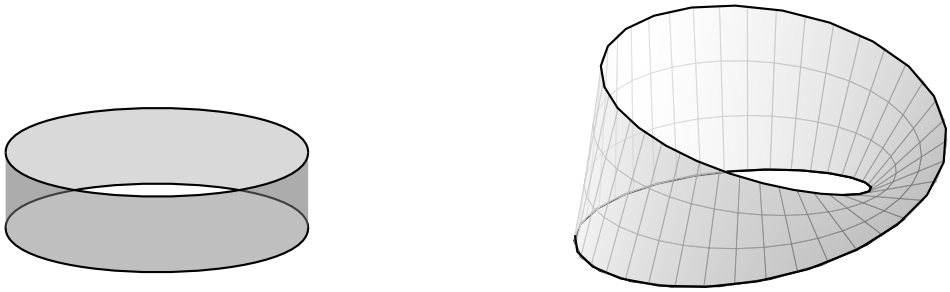
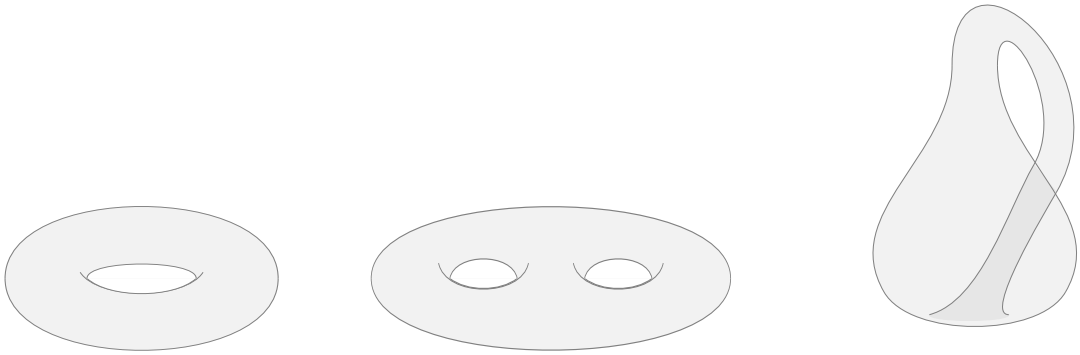


Figure 9: A ring (left) and a Möbius strip (right).

There are two types of surfaces, *orientable* and *non-orientable* surfaces. To understand better this notion, take a strip and attach the extremities together to form a ring (see Fig. 9): we can distinguish the interior and exterior of this newly formed ring. Now, if we twist the band before joining the extremities, we obtain a *Möbius strip* for which there is no distinction between interior and exterior anymore. Notice that the boundary of a Möbius strip is a circle, up to deformation. The surfaces (without boundaries) are there classified as follow:

Theorem (Classification of surfaces)**.** *Every connected surface without boundary is homeomorphic to either:*

- *An orientable surface of genus $g \geqslant 0$, obtained by gluing g handles on the sphere. Its Euler characteristic is $\chi = 2 - 2g$.*
- *A non-orientable surface of genus $g \geqslant 0$, obtained by attaching g structures called "crosscap" (corresponding to a Möbius strip) on the sphere. Its Euler characteristic is $\chi = 2 - g$.*



(a) A torus (oriented genus 1 and Euler genus 2).

(b) A double-torus (oriented genus 2 and Euler genus 4).

(c) A Klein bottle (non-oriented genus 2 and Euler genus 2).

Figure 10: Examples of surfaces of small Euler genus.

When we consider the genus of a graph, it is sometimes more convenient to consider its *Euler genus*, that encompasses both orientable and non-orientable cases. Euler genus is the minimum integer g such that the graph can be embedded on a non-orientable surface of genus g or on an orientable surface of genus $g/2$. A graph G with a given embedding also has a Euler characteristic: $\chi(G) = |V(G)| - |E(G)| + |F(G)|$ (where $F(G)$ are the faces of G). This Euler characteristic corresponds to that of a surface it can be embedded on, so $\chi(G) = 2 - g$. The Euler genus is thus a parameter that measures the complexity of a graph, and how far it is from being planar: when the genus is small, we expect sometimes a behaviour similar to that of a planar graph.

Tree parameters

To study and exploit the properties of a graph class, instead of using a representative of the class, another strategy is to use other parameters, and mostly those

that measure their resemblance to trees. Indeed, as we saw earlier, most problems have an easy algorithmic solution on trees. When we consider graphs that resemble to trees (whatever that means exactly), we would like to adapt the solution on trees into a solution on those graphs. So, how do we properly define this notion of proximity to being a tree?

Feedback vertex set

A first solution is to try to consider a graph as a forest with some additional "bad" vertices. A *feedback vertex set (FVS)* of a graph is a set of vertices whose removal results in an induced forest. The intuition is that whenever the FVS of a graph is small, you can apply an algorithm adapted for trees on most of the graph (the induced forest), then correct the result to take into account the small set of vertices that have been put aside (The FVS). In fact, the size of a smallest FVS is closely related to the number of disjoint cycles in the graph: Erdős-Pósa therorem [EP65] states that if a graph contains fewer than k pairwise vertex-disjoint cycles, then it has a FVS of size $O(k \log k)$.

The size of a smallest FVS is a good parameter to measure the likeness to a tree, but it is rather restrictive. We will see later graphs with a big FVS that behave similarly to a tree, pointing to a need for an other parameter.

Chordal graphs

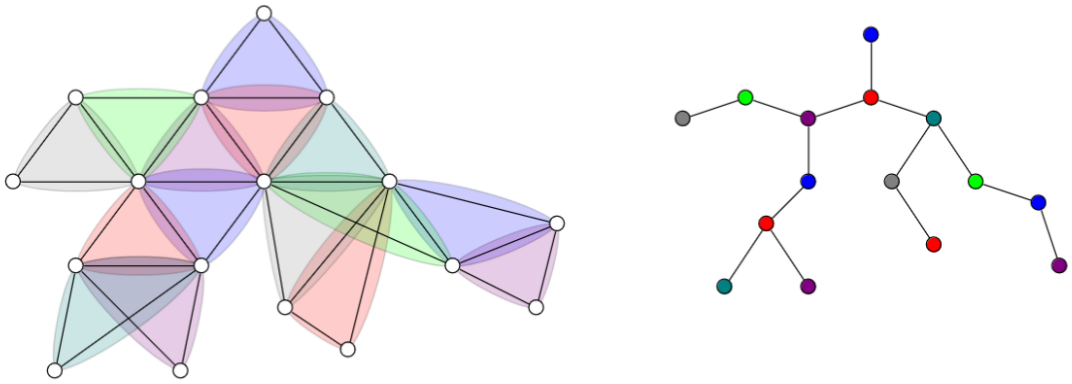


Figure 11: Example of a 2-tree, and a tree representation where every K_3 is represented by a vertex.

A *k-tree*, for some integer $k \geqslant 1$ is graph that can be constructed as follow: start from the complete graph on $k+1$ vertices, and recursively add each other

vertex by connecting it to a clique on k vertices of the current graph. Notice that a tree is a 1-tree. A clique tree can this be seen a collection of K_{k+1} organized in a shape of a tree. Fig. 11 gives side-by-side a $2-tree$ and the tree organization of the K_3s. Intuitively, a k-tree is a tree that is more or less thick, its thickness depending on k. Observe that a k-tree can have a lot of disjoint cycles (thus no small FVS) even if k is small (and thus have a structure close to that of a tree).

More generally, if we allow the cliques to be of different size, the graph obtained is a *chordal* graph. The *clique number* of a graph is the size of a biggest clique. A k-tree is thus a special case of chordal graph, where all the cliques of the construction have size $k+1$. This class is also defined as the class of graphs with no cycle of at least 4 vertices as induced subgraph.

Moreover, a chordal graph such that the cliques are organized into a path are called *interval graphs*.

Treewidth

If we consider an arbitrary graph, it can always be completed into a chordal graph by adding some edges: we can for example add all the possible edges to obtain a complete graph, which is chordal. Now, among all the possible completions, we select one that minimizes the clique number of the resulting chordal graph, say it is k. Then we say that the graph has *treewidth* $k-1$. In other words, a graph has treewidth k if it can be complemented into a chordal graph of clique number $k+1$ and not in a chordal graph of clique number k. Roughly speaking, a graph of treewidth k can be seen as a subgraph of a k-tree, which has an underlying tree structure. In fact, as we will see more precisely in the thesis, whenever a graph has treewidth k, it admits a decomposition, called a *tree decomposition*, into a tree such that each vertex of the tree represents a bag of size $k+1$ (corresponding to cliques in the chordal completion).

Similarly, if a graph can be completed into an interval graph of clique number $k+1$ and not into an interval graph of clique number k, then we say that the graph has *pathwidth* k.

The treewidth, and the tree decomposition associated to it, are the most common tools to generalize algorithms on trees to other graphs. It is NP-complete to find the treewidth of a graph but there is a linear algorithm to check if a graph has treewidth k for a fixed k. Moreover the size of a smallest FVS is an upper bound of the treewidth, so bounding the size of a smallest FVS results in bounding the treewidth of a graph.

A seminal metatheorem using treewidth to bound the complexity of a problem is Courcelle's theorem [Cou90]. It states that every decision problem on graphs that can be expressed in monadic second-order logic (a logic defined on the vertex set of a graph, with the adjacency relation, and for which the quantifiers are

restricted to sets of vertices only) can be solved in linear time. For example, MIS can be expressed in monadic second-order logic, and thus can be solved in linear time on graphs of bounded treewidth.

Results and organization of the thesis

The results of this thesis are distributed in two chapters, each starting with an introduction more specific to the subject of the chapter.

In Chapter 1, we focus on minor-universality. Recall that this is motivated by the fact that there is a planar graph on $O(n^2)$ vertices minor-universal for the planar graphs on at most n vertices. We first discuss the problem of finding a smaller minor-universal graph for planar graphs and subclasses of planar graphs. Then, we extend this theorem to more general graph classes: the graphs of bounded Euler genus.

Then, in Chapter 2, we present our result related to tree structure in graphs and more precisely to induced forest in graphs. The first part of this chapter focuses on a conjecture of Esperet, Lemoine, and Maffray stating that k-degenerate[5] graphs with a path of size n as subgraph also have a path of size $(\log n)^{\Omega(1)}$ as induced subgraph. We prove this to be true for graphs with bounded treewidth and generalize this to the class of graphs excluding any fixed graph as topological minor. In the second part of this chapter, we generalize the Erdős-Pósa Theorem, and showed that sparse graphs with a small number of cycles that are pairwise not connected have an FVS whose size is logarithmic in the number of vertices, which is sharp.

Finally, in Chapter 3 we conclude by recapitulating our results, as well as those that are not presented in this manuscript ([BCH+23, DHK+21]). We then present some questions that arise from our work and that could be interesting to pursue.

[5] a graph is k-degenerate if every subgraph (including the graph itself) admits a vertex of degree at most k.

Chapter 1

One Major to Rule Them All

Contents

1.1 Introduction

A graph is *universal* for a family of graphs if it contains all the graphs in the family, for some containment relation (subgraph, induced subgraph, minor, ...). Graph universality plays an important role in graph theory and graph algorithms. Rather than considering individually each graph of a given family, it is sometimes simpler to manipulate a reduced set of graphs, or even a single graph, whose properties are close enough from those of the graphs of the family. Universal graphs are studied for families of finite and infinite graphs. In this chapter, we focus on families of finite graphs. These families can also be infinite, so, for convenience, we restrict them to the graphs of the family with at most n vertices for some integer n, and observe how the size of the universal graph increases when n increases.

If we consider the relation of subgraph, a trivial example would be that the complete graph on n vertices K_n is the subgraph-universal graph with the fewest

vertices for the family of all n-vertex graphs, since every graph on at most n vertices is a subgraph of K_n. This result, optimal for any subfamily of n-vertex graphs, does not grasp many properties of the family. In that sense, minimizing the number of vertices of a subgraph-universal graph is not particularly insightful. However, subgraph-universality has been well-studied, with more relevant constraints on the desired universal graph.

For instance, the Strong Product Theorem [DJM+20], which states that every planar graph is a subgraph of the strong product[1] of a path and a graph of bounded treewidth, has provided many new insights in planar graph theory.

Thanks to this theorem, it suffices to study such a product to derive properties on all planar graphs. This includes a proof for bounding a useful parameter called the queue number and a variant of chromatic number for planar graphs [DEJ+20]. This construction also led to the construction of a subgraph-universal graph for planar n-vertex graphs minimizing the number of edges. Indeed, Esperet, Joret and Morin [EJM22] constructed such a universal graph with $(1 + o(1))n$ vertices and $n^{1+o(1)}$ edges, whereas the best previous bound on the number of edges, from Babai, Chung, Erdös, Graham and Spencer in 1982, was $O(n^{3/2})$ [BCE+82].

Graph universality is also widely studied for the relation of induced subgraph, especially for classes closed under taking subgraph, called *hereditary*. Indeed, since the universal graph contains all the graphs of the class as induced subgraphs, it also contains their induced subgraphs, so it makes sense to consider a class that includes them. This interest for those universal graphs stems from the equivalence between finding a smallest induced-subgraph-universal graph of a given family and finding an optimal *adjacency labelling scheme* for that family. An adjacency labelling scheme for a family of n-vertex graphs $\mathcal{F}_n$ is the combination of an encoding and a decoding functions, depending only on the family. The former affects some bits of data, called a label, to each vertex, and the latter assesses the adjacency of two vertices thanks to the information contained in their labels only. The quality of the labelling scheme is measured by the maximum number of bits used for a label. The universal graph of $\mathcal{F}_n$ corresponding to a labelling scheme of $\mathcal{F}_n$ is the graph with a vertex for each possible label and whose adjacency respects the labels. The reverse construction is also possible, and the size of the labels in the labelling scheme is the logarithm of the order of the corresponding universal graph. Finding a labelling scheme of $\mathcal{F}_n$ with small labels, or equivalently finding a smallest induced-subgraph-universal graph, gives a compressed description of the graphs of $\mathcal{F}_n$.

Observe that since a labelling scheme of $\mathcal{F}_n$ has to differentiate all the graphs

[1] The *strong product* of two graphs G, H is a way of combining them into a single graph, whose vertex set is the Cartesian product of $V(G) \times V(H)$, and $(u, v), (u', v')$ are adjacent if and only if u, u' are either equal or adjacent in G and v, v' are either equal or adjacent in H.

of the family, the labels have to contain at least $\frac{\log|\mathcal{F}_n|}{n}$ bits, where $|\mathcal{F}_n|$ is the number of distinct graphs, up to isomorphism, in $\mathcal{F}_n$. In other words, any universal graph of $\mathcal{F}_n$ has at least $2^{\frac{\log|\mathcal{F}_n|}{n}}$ vertices. This bound is tight for dense families, i.e. containing at least $2^{\Omega(n^2)}$ graphs, as Bonamy, Esperet, Groenland and Scott proved that there is always such universal graph on $2^{\frac{\log|\mathcal{F}_n|}{n}+o(n)}$ vertices [BEGS21]. However, and contrary to the Implicit Representation Conjecture, this is surprisingly not always the case for sparse families, as Hatami and Hatami exhibited in 2021 [HH22] a family of graphs a most $2^{n\log n}$ graphs such that every universal graph has at least $2^{n^{\Omega(1)}}$ vertices.

Recently, Dujmović et al.[DEG+20] proved that there is a universal graph for planar n-vertex graphs on $n^{1+o(1)}$ vertices. They generalised their construction to other families including the graphs of bounded genus. Interestingly, this result was also obtained using the characterisation of the Strong Product Theorem, and was completely out of reach without it.

1.1.1 Minor-universality

This chapter focuses on *minor-universality*, i.e. finding universal graphs for the minor relation. When H is a minor of G, it is sometimes convenient here to consider the reverse relation and say that G is a *major* of H. Then, a graph is minor-universal for a family if it is a major of every graph of this family.

Observe that K_n is a major of every n-vertex graph, so finding a smallest (number of vertices) minor-universal graph for a family of finite graphs does not give much information on the graphs of the family. However, we can set some constraints for the universal graph to capture more information. For example, Bodini proved that the smallest minor-universal tree for n-vertex trees has between $\Omega(n\log n)$ and $O(n^{1.985})$ vertices [Bod02]. In this result, the minor-universal graph is required to be itself a tree.

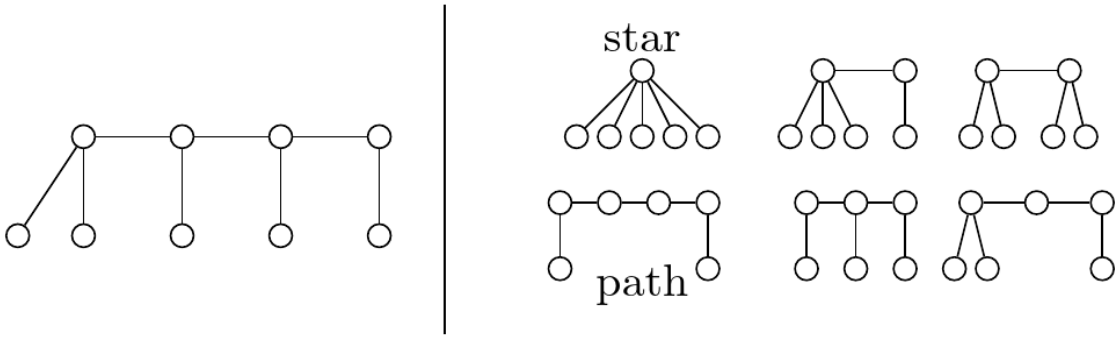


Figure 1.1: $(2n-3)$-vertex caterpillar (left) minor-universal for the n-vertex caterpillars (right), with $n=6$.

A simpler example is considering the family of *caterpillars* on at most n vertices. A caterpillar is a tree whose internal[2] vertices form a path. In fact we can focus on those on exactly n vertices as the smaller ones are minors of them. Observe that a caterpillar minor-universal for the caterpillars has to have at least $n-1$ leaves to contain the star as minor, and at least $n-2$ non-leaf vertices to contain the n-path as minor. Both these lower bounds can be reached and the caterpillar on $2n-3$ vertices presented in Fig. 1.1 is minor-universal for the n-caterpillars, as one can easily check that every n-caterpillar is a minor of this caterpillar.

If we consider the family of planar n-vertex graphs, we saw previously that the Strong Product Theorem gives a subgraph-universal graph, which is also a minor-universal graph for this family. However, this universal graph is far from being planar as its genus is unbounded. If we require the minor-universal graph to be itself planar, there is a fundamental result, related to the celebrate Excluded Grid Theorem of Robertson and Seymour [RS86]. The Excluded Grid Theorem states that the graphs excluding as minor a given planar n-vertex graph have their treewidth bounded by a function of n. This theorem is a part of their seminal work on Graph Minor Theory to characterise minor-closed families by a finite set of forbidden minors. It implies a deep connection between Graph Minor Theory, a branch of Structural Graph Theory, and the treewidth of a graph, which is a parameter that has many algorithmic applications (see Chapter 2 that is more focused on tree structures and treewidth). The bound on the treewidth of a graph G excluding a planar n-vertex graph is enormous, and in a later paper, Robertson, Seymour and Thomas [RST94] give a better bound, using in particular the following theorem:

Theorem 1.1.1 ([RST94, Theorem 1.5])**.** *Every planar n-vertex graph is a minor of the $2n \times 2n$-grid* [3]*.*

In other words, there is a $4n^2$-vertex planar graph minor-universal for the planar n-vertex graphs. Notice that if a graph excludes a fixed planar n-vertex graph as a minor, then in particular it excludes the $2n \times 2n$-grid as a minor. This radically simplifies the search for a bound on the treewidth of those graphs: any bound for graphs excluding the $2n \times 2n$-grid as a minor directly applies to graphs excluding any given n-vertex planar graph.

There are two directions to refine this result: either by considering subfamilies of planar n-vertex graphs for which there is a smaller minor-universal graph, or by considering a family that generalizes the notion of planarity. The next subsection explores the former direction. For the latter direction, a natural family to consider

[2] In a tree, a leaf is a vertex of degree 1, the other vertices are called internal vertices.

[3] The $w \times h$-grid is the plane graph on $w \cdot h$ vertices of the form (i, j) for $1 \leqslant i \leqslant w$ and $1 \leqslant j \leqslant h$, where (i, j) is adjacent to $(i, j+1)$ and $(i+1, j)$.

when generalizing the family of planar graphs is the family of graphs with bounded genus. With Cyril Gavoille [GH23b], we generalize Theorem 1.1.1 to those graphs as follows:

Theorem 1.1.2. *For every n and every surface Σ of Euler genus g, there is a graph $\mathcal{U}_{n,\Sigma}$ embedded on Σ with $O(g^2(n+g)^2)$ vertices that contains as minor every graph embeddable on Σ with n vertices.*

In other words, $\mathcal{U}_{n,\Sigma}$ is a minor-universal graph for n-vertex graphs embeddable on Σ. In fact, as we will see in Theorem 1.3.2, the graph $\mathcal{U}_{n,\Sigma}$ depends only on the Euler genus and orientability of Σ. Moreover, the embedding of the minor is preserved in $\mathcal{U}_{n,\Sigma}$.

Our proof is constructive. The minor-universal $\mathcal{U}_{n,\Sigma}$, as well as the witness of the minor in $\mathcal{U}_{n,\Sigma}$, can be constructed in polynomial time. Our construction is inspired by the construction for planar graphs, and interestingly, along the way, we give an alternative and constructive proof for the non-constructive part of the proof of Theorem 1.1.1.

1.1.2 Grid-drawing and grid-major

Among subfamilies of planar n-vertex graphs, we focus on graphs with a certain property on their planar embedding. We know from Theorem 1.1.1 that those planar graphs are minor of a grid of *area* (number of vertices) $4n^2$, and we wonder what other property enables us to reduce this area.

A *poly-line grid-drawing* of a graph is a special planar embedding where vertices have integral coordinates and edges are poly-line segments such that two consecutive segments of an edge meet in what we call a *bend* that has integral coordinates as well (see Fig. 1.2 for an example of poly-line drawing). These embeddings can be seen as a planar embedding on a grid where vertices and edge bends are mapped to vertices of the grid. Several types of drawings of a graph on a grid are studied (like straight-line if there are no edge bends, orthogonal if each edge is a path of the grid, and so on) but unless specified otherwise, we will refer to a poly-line grid-drawing when we consider a grid-drawing.

Links between graph drawing and graph minor theory are far from new. For instance, pathwidth and height (i.e., the number of rows) of the drawing are two well-known parameters that are linked: the pathwidth of the graph is a lower bound on the height in any grid-drawing [Sud03] (including poly-line one), whereas there are known grid-drawings for outerplanar graphs achieving the pathwidth, up to some constant factor (combining [Bie12] and [BBCR14]). As noticed by [BCE+19], the pathwidth is also a lower-bound for the height of any grid-major of the graph.

Moreover, when H is a minor of a graph G, we can define a *model* of H in G, that is a set of pairwise disjoint connected subgraphs of G, each one called

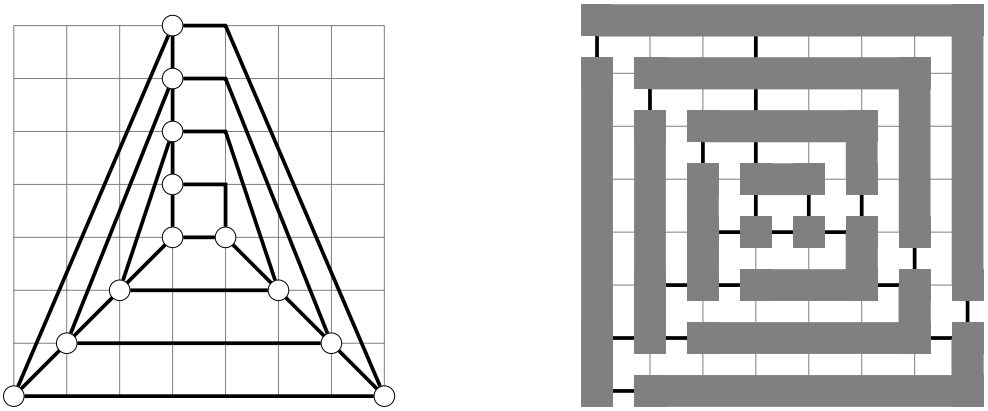


Figure 1.2: A poly-line grid-drawing (left) and a minor-drawing (right) of a graph formed by 4 nested triangles, on an 8×8-grid.

super-node in a one-to-one correspondence with $V(H)$, and a subset of edges of G, each one called *super-edge* in one-to-one correspondence with $E(H)$, such that $(u, v) \in E(H)$ if and only if there exists a super-edge connecting a vertex of the super-node of u and a vertex of the super-node of v. In the case of a graph H minor of a grid, we can view a model of H as a planar embedding on the grid where vertices are mapped on connected subgraphs of the grid and and edges are mapped on edges of the grid. This defines another kind of grid-drawing, that we call a *minor-drawing* on a grid.

It is well-known that any planar graph G with n vertices has a poly-line drawing of area $O(n^2)$. However, the quadratic area is required because of the nested-triangles (see Fig. 1.2) graph among other worst-case examples [dFPP88, Bie11]. Notice that, for this nested-triangles graph, the smallest grid-major also has area $\Omega(n^2)$ [BCE+19, Lemma 5]. The same authors showed an example of a graph, a ladder with a universal vertex inspired from [Bie11], that is minor of a grid of area n and for which any grid-drawing required an area of $\Omega(n^2)$. We observe that, in both contexts, some planar graphs may require quadratic-area grids, for drawing or for major (like the nested triangles graph), whereas no graphs with simultaneously small area grid-drawing and large grid-major are known.

Recently, Dieng and Gavoille [DG20] raised the question of whether any graph having a grid-drawing of area A has a grid-major of area $O(A)$. They show an upper bound of $O(A^{3/2})$ on the grid-major area. More precisely, they show that every graph having a poly-line $w \times h$-grid-drawing is minor of a $O(wh^2)$ area grid, leaving open whether we can hope for a grid of approximately the same area, i.e. $O(wh)$.

Our results. Working on this problem with Cyril Gavoille, we first show that the decisional problem associated to finding the smallest grid-major of a planar graph, even knowing an optimal grid-drawing of the graph, is NP-complete. This contrasts with the variants where only the height (of the drawing or of the grid-major) has to be minimized. From [BCE+19], the complexity status of the "height" versions are open. However, an existential FPT algorithm in the height for the minimum height grid-major problem can be derived from Courcelle's Theorem (see [BCE+19]).

We then consider in particular the family of k-outerplanar graphs, which are plane graphs where the vertices are partitioned into k layers, which can be constructed by successively peeling all the vertices lying on the outerface. We show that k-outerplanar graphs have a grid-major of area $O(kn)$, and that this bound is the best possible. This improves upon the $O(n \log n)$ upper-bound for 1-outerplanar graphs of [DG20].

Finally we consider the reverse question, that is finding a poly-line grid-drawing of a graph when given its model as a minor of a grid. We show that if a graph of max degree Δ is a minor of a $w \times h$-grid, then it has a $\delta w \times \delta h$ poly-line drawing, where $\delta = \max(1, \lfloor \frac{\Delta}{2} \rfloor)$. This implies in particular that for a family of planar graphs of constant degree, if we have a lower bound on the area of their grid-drawing, then we we have, up to a constant, the same lower bound on the area of a minor-universal grid.

1.2 Universal grid for grid-drawings

1.2.1 Overview

The objective of this section is to establish relationships between the area of a smallest grid-major of planar graph and the properties of its other embeddings.

We start in Section 1.2.2 with the necessaries definitions and preliminaries. We remark in Proposition 1.2.2 that, given any integer $\Delta \geqslant 3$, it takes a linear time to compute from G a maximum-degree-Δ major G' of G with the minimum number of vertices.

In contrast, we show in Section 1.2.3 that computing a grid-major (which is a graph of bounded maximum degree) of a graph with smallest possible area is NP-Hard, even if the optimal grid-drawing of the graph is given as input.

In Section 1.2.4, we show that k-outerplanar graphs have a grid-major of area $O(kn)$, which is sharp and improves the $O(n \log n)$ upper-bound of [DG20] for $k = 1$.

Finally, in Section 1.2.5, we consider the question of constructing grid-drawings from minor-drawings with similar area, and this in order to deduce lower bounds

on the smallest grid-major of a graph.

1.2.2 Definitions and preliminaries

Grid-drawings

Recall that a *grid-drawing* can be seen as a (planar) grid embedding where vertices and edge bends are mapped to vertices of the grid (often refer here as *grid-point* to avoid confusion with the vertices of the embedded graph). The drawing is *straight-line* if there are no edge bends, it is *poly-line* otherwise. It is *orthogonal* if each edge is a path of the grid. Only planar graphs of maximum degree at most four can have orthogonal drawings, which are special poly-line drawings.

Consider a poly-line drawing Γ of a given plane graph G. We say that Γ *preserves the embedding* if we have an embedding homeomorphic to the one of G. Most grid-drawing algorithms preserve the embedding, for instance by first triangulating the input embedded graph.

Minor-drawings

A *minor-drawing*[4] of a graph H is a drawing of an *H-model* in some grid-major of H. Therefore, in this kind of drawing, vertices of H are pairwise disjoint connected subgraphs of the grid (the super-nodes), and edges of H are edges of the grid (the super-edges). (See Fig. 1.3 for an example.) To "simplify" our drawings, we will allow super-edges to be internally disjoint grid paths (and not only grid-edges). It is straightforward to check that both definition leads to drawings of same area. Moreover, the second definition (so allowing path-grid edges) generalizes orthogonal drawings. Hence, every graph that has an orthogonal drawing of area A, has a minor-drawing of area at most A.

In order to define the embedding preserving property of a minor-drawing, we need to associate with each super-node a region. More precisely, the *region* of a super-node X is the region of the plane composed of all the vertices and edges in X and of all faces of the grid that are bounded by four edges of X. Note that the region of a super-node may be not homeomorphic to a disc: it may contain "holes" making the region not contractible to a single point without crossing other super-edges and/or super-nodes of the H-model. It is however easy to see that every super-node X having a region with some "holes" can be transformed into a new super-node X' with $V(X') = V(X)$ whose region has no holes and that remains an H-model[5]. A minor-drawing is *simple* if all super-node regions are

[4] The terminology in [BCE+19] is "grid-representation" but this is the same idea.

[5] One can remove each hole of the region of X by deleting from X a shortest path in the dual of the grid limited to edges of X from inside to outside the region of X. Deleting such a path

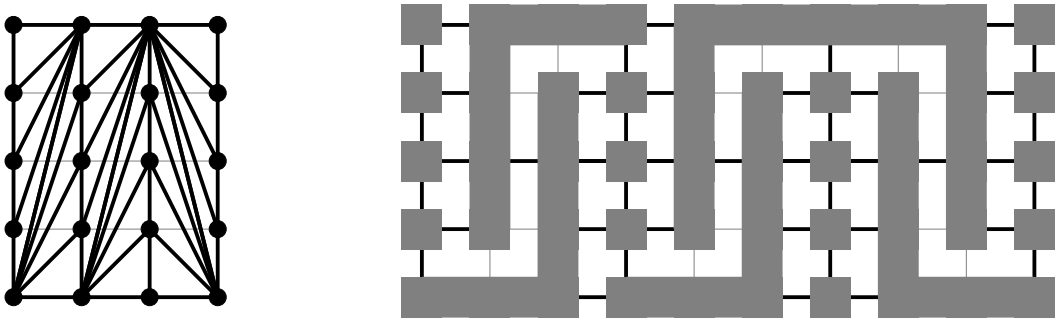


Figure 1.3: A dense drawing on a $w \times h$ grid for a planar graph with $w = 4$ and $h = 5$, and minor-drawing on a $(3w - 2) \times h$ grid of the same graph and with the same embedding.

homeomorphic to a disc.

Consider a plane graph H. We say that the minor-drawing Γ *preserves the embedding* of H if after contracting every super-node to a single point we obtain an embedding that is homeomorphic to the embedding of H. Only simple minor-drawings can preserve the embedding.

Basic properties

We first present two properties.

Proposition 1.2.1. *If H has a grid-major of area A, then every planar graph excluding H as minor has treewidth at most $9\lfloor\sqrt{3A}\rfloor/2 - 4 \approx 7.8\sqrt{A}$.*

Proof. Using [GT12, Theorem 1.4, pp. 419], [DG20] showed that every planar graph excluding an $r \times r$ grid as minor has treewidth at most $9r/2-4$. From [Lei80, Theorem 2], we can derive that every grid of area A is minor of a square grid of area at most $3A$, i.e., a $r \times r$ grid with $r^2 \leqslant 3A$. We conclude by setting $r = \lfloor\sqrt{3A}\rfloor$ that is the largest integer such that $r^2 \leqslant 3A$. □

Proposition 1.2.2. *If G' is a maximum-degree-Δ major with $\Delta \geqslant 3$ of a graph G, then the minimum number n' of vertices for G' is*

$$n' = \sum_{u \in V(G)} \max\{1, f_\Delta(\deg(u))\} = \sum_{d=0}^{\Delta} n_d + \sum_{d>\Delta} n_d \cdot f_\Delta(d)$$

where $f_\Delta(d) = \lceil (d-2)/(\Delta-2) \rceil$ and n_d denotes the number of vertices of degree d in G. Such a major G' can be obtained in linear time from G by expanding every vertex of degree $d > \Delta$ by a path of $f_\Delta(d)$ vertices.

will remove edges of X and will remove a hole from its region while preserving X connected.

Proof. By definition, to each vertex u of G corresponds a super-node X_u in G'. To prove the claim, it is sufficient to show that $|X_u| \geqslant (\deg(u)-2)/(\Delta-2)$ whenever $\deg(u) > \Delta$. Indeed, for $\deg(u) \leqslant \Delta$, then $|X_u| = 1$ is enough and sufficient, and for $\deg(u) > \Delta$, $|X_u| \geqslant (\deg(u)-2)/(\Delta-2)$ implies $|X_u| \geqslant f_\Delta(\deg(u))$. Furthermore, we check that to a path of $f_\Delta(\deg(u))$ vertices we can connect up to $\deg(u)$ vertices outside of the path. Additionally, we check that we can do so without exceeding the degree bound of Δ, thus obtaining $|X_u| \leqslant f_\Delta(\deg(u))$.

So, let us show that $|X_u| \geqslant (\deg(u)-2)/(\Delta-2)$ for $\deg(u) > \Delta$. Given a graph H and $X \subseteq V(H)$, we denote by $N_H(X) = \{(u,v) \in E(H) : u \in X \text{ and } v \notin X\}$ the open neighborhood of X, i.e., the set of neighbors of X not in X. We said that X is connected if the induced subgraph $H[X]$ is connected. Let $\Delta(H)$ denote the maximum degree of H. For every integer $x \geqslant 1$, denote by

$$\eta(x) \;=\; \max_{\substack{\forall H,\, \Delta(H)\leqslant\Delta \\ \forall X\subseteq V(H),\, |X|=x \\ X \text{ connected}}} |N_H(X)|$$

the maximum number of neighbors that can have a connected subset of x vertices in a graph of maximum degree Δ.

Let us show that

$$\eta(x-1) \geqslant \eta(x) - \Delta + 2, \quad \text{for every } x \geqslant 2. \tag{1.1}$$

To show this, consider a connected subset X and a graph H realizing the optimal bound $\eta(|X|)$ with $|X| = x \geqslant 2$. Consider a rooted spanning tree of X, and let v be a leaf in this tree, and let w be its parent. Such vertices exist since X is connected and has $x \geqslant 2$ vertices. Now, consider $Y = X \setminus \{v\}$. Y is connected since v is a leaf. The number of neighbors of Y in $H \setminus Y$ is at least $|N_H(X)| - (\deg_H(v) - 1) + 1$, obtained from all the neighbors of $N_H(X)$ minus at most the $\deg_H(v) - 1$ neighbors of v (the neighbor w of v cannot be removed since it is in Y) and plus v which is not anymore in Y. We have therefore showed that $|N_H(Y)| \geqslant |N_H(X)| - \deg_H(v) + 2$. Note that $|N_H(Y)|$ gives a lower bound on $\eta(|Y|) = \eta(|X|-1) = \eta(x-1)$. Therefore, $\eta(x-1) \geqslant \eta(x) - \Delta + 2$ because $|N_H(X)| = \eta(|X|) = \eta(x)$ and $\deg_H(v) \leqslant \Delta(H) \leqslant \Delta$, and Eq. (1.1) is proved.

Clearly, $\eta(1) = \Delta$. From Eq. (1.1), it follows that

$$\begin{aligned} \eta(x) &\leqslant \eta(x-1) + \Delta - 2 \\ &\leqslant \eta(x-i) + i \cdot (\Delta - 2) \\ &\leqslant \eta(1) + (x-1) \cdot (\Delta - 2) \\ &\leqslant (\Delta - 2) \cdot x + 2\ . \end{aligned} \tag{1.2}$$

Let us come back with vertex u and its super-node X_u in G'. Obviously, we need X_u connected and $|N_{G'}(X_u)| \geqslant \deg(u)$. Since $\Delta(G') \leqslant \Delta$, we must

have $|N_{G'}(X_u)| \leqslant \eta(|X_u|)$. Combining with Eq. (1.2), it follows that $\deg(u) \leqslant (\Delta - 2) \cdot |X_u| + 2$, or equivalently $|X_u| \geqslant (\deg(u) - 2)/(\Delta - 2)$, as desired. This completes the proof of Proposition 1.2.2. □

1.2.3 NP-Completeness for grid-major problems

We consider two variants of the Grid-Major Problem.

The Grid-Major Problem *for a graph* is, given a planar graph H and an integer A, to decide whether there is a grid of area at most A containing H as minor.

The Grid-Major Problem *for a drawing* is, given the optimal poly-line drawing Γ of a planar graph H and an integer A, to decide whether there is a grid of area at most A containing H as minor.

Theorem 1.2.3. *The* Grid-Major Problem *for a graph and for a drawing are both NP-complete.*

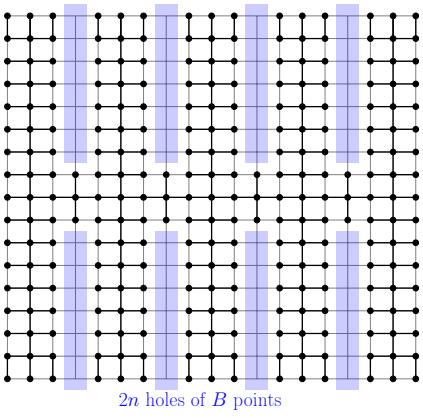
In fact, the problems remain NP-complete even if the input graph is a forest of maximum degree four, and, for the drawing variants, even if the input drawing has optimal area. The proof of this theorem is an adaptation of the NP-completeness proof of the Forest Layout Problem, by Dolev, Leighton, and Trickey in [DLT94], as reviewed in the next paragraph.

Sketch of the proof for Forest Layout Problem

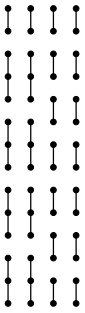
Lemma 1.2.4 ([DLT94, Theorem 7])**.** *Given a forest and an integer A, deciding whether the forest has an orthogonal drawing with area at most A, i.e., the* Forest Layout Problem*, is NP-complete.*

As many subsequent NP-completeness results in graph drawings, their proof is a reduction from the 3-Partition Problem. For this latter problem, we are given two integers B, m and a set of integers $X = \{x_1, \ldots, x_{3m}\}$ summing to mB and with $B/4 < x_i < B/2$. We ask whether X can be partitioned into m disjoint sets each summing to B. This problem is strongly NP-complete, i.e., even if integers B and m are encoded in unary.

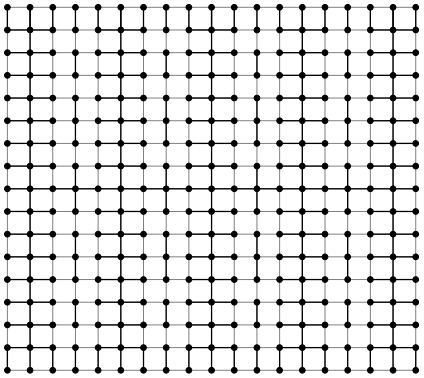
The reduction is the following: given an instance of the 3-Partition Problem, construct a forest F such that for each $x_i \in X$, there is a path p_i on x_i vertices. Suppose that $m = 2n$ (if m is odd, then $n = \lceil m/2 \rceil$ and we add a path on m vertices to the forest), and let $A = (4n + 3) \times (2B + 3)$. Add to this forest a frame tree T_f, which is a tree of $(2B + 3) \cdot (4n + 3) - 2nB$ vertices, such as represented in Fig. 1.4a.



(a) Orthogonal drawing of the frame tree T_f,



(b) Rest of the forest,



(c) Orthogonal drawing of F (solution of the forest layout problem on this instance),

Figure 1.4: The forest F for $n = 4$, $B = 7$, and a $(4n + 3) \times (2B + 3)$-grid.

The Fig. 1.4 represents the instance of the forest layout problem associated to the 3-Partition Problem with $B = 7, m = 6 = 2n$ and $x_1 = \cdots = x_6 = 3$, $x_7 = \cdots = x_{18} = 2$.

It is proved in [DLT94] that the 3-Partition Problem instance has a solution if and only if F has an orthogonal drawing on a grid of area A. They proved that the only orthogonal drawing of T_f on the grid is the one represented in the Fig. 1.4a, leaving $m = 2n$ holes of B points, which implies that in order to draw F, it is necessary to distribute the paths in those holes, like in Fig. 1.4c.

NP-completeness of the Grid-Major Problem

The same reduction from 3-Partition Problem can be done for the Grid-Major Problem, because in the situation presented, orthogonal drawing and minor-drawing are the same. In fact, for the graph version, it can be directly reduced from Forest Layout Problem.

Claim 1.2.5. *Let H be a simple planar graph with $\Delta(H) \leqslant 4$, then a minor-drawing of H on a grid of area $|V(H)|$, if it exists, is equivalent to an orthogonal drawing of area $|V(H)|$.*

Proof. Let H be a planar graph on n vertices with degree $\leqslant 4$.

Suppose that there is a minor-drawing Γ of H with area $N = |V(H)|$. Then there are N distinct super-nodes and N grid-points thus each super-nodes contains exactly 1 grid-point, and (by definition of a minor-drawing) every *edge* of H is a path in the grid. Therefore Γ is an orthogonal drawing of H.

As we saw in Section 1.2.2, an orthogonal drawing can be seen as a special case of minor-drawing, so the reverse is trivial. □

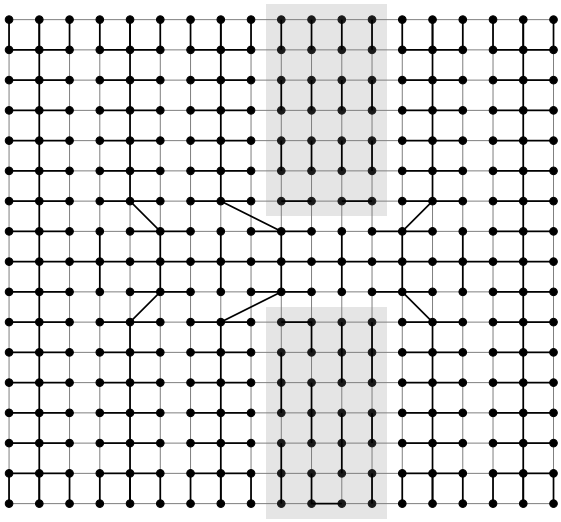


Figure 1.5: Optimal poly-line drawing of F (for $n = 4$ and $B = 7$) on a $(4n+3) \times (2B+3)$-grid.

Therefore, Grid-Major Problem for a forest F with $\Delta \leqslant 4$ and with $A = |V(F)|$ is equivalent by Claim 1.2.5 to Forest Layout Problem with the same inputs, which is NP-complete. In order to prove that Grid-Major Problem for drawing is NP-complete, we can do the a similar reduction from 3-Partition Problem.

Given an instance of the 3-Partition Problem, construct the exact same forest F and area A (with the same notation) as in Section 1.2.3. Take the following straight-line (thus poly-line) drawing Γ of F, represented on Fig. 1.5: the frame tree T_f is drawn almost the same way as in the Fig. 1.4a, but the vertical branches are pushed on the sides, leaving in the middle two empty $n \times B$ subgrids (drawn in gray in the figure), and all the paths of F are drawn in those subgrids. This drawing is optimal since $|V(F)| = A$.

Solving Grid-Major Problem for a drawing with input Γ and A, is deciding if there is a minor-drawing of F of area A. Since F has A vertices, by Claim 1.2.5, finding a minor-drawing of F on a grid of area A is equivalent to finding an orthogonal drawing of F, which is itself equivalent to finding a solution to the 3-Partition Problem. Thus the Grid-Major Problem for a drawing is also NP-complete.

1.2.4 Outerplanar graphs

In this section we show that n-vertex k-outerplanar graphs have $O(kn)$ grid-major area, which improves the previous $O(n \log n)$ upper bound for $k = 1$ [DG20]. This is complemented by an $\Omega(kn)$ lower bound.

Recall that a planar graph is k-outerplanar if it has a plane embedding where the vertices partition into k layers, obtained by successively peeling all the vertices lying on the outerface. If the embedding is given, we call it a *k-outerplane* graph. Outerplanar graphs are 1-outerplanar graphs.

By expanding along each layer each vertex of degree $d > 4$ by a path of $f_4(d) = \lceil (d-2)/2 \rceil$ vertices on which we attach the d neighbors in a balanced way as in Proposition 1.2.2, we can show that (see also [Bod98, Lemma 82] for a non-quantitative variant[6] and [DK99, Lemma 3] for infinite planar graphs):

Proposition 1.2.6. *Every k-outerplane graph H with $n \geqslant 4$ vertices is a minor of a k-outerplane graph G of maximum degree four, with $O(n)$ vertices, and preserving the embedding of H. Such a major G can be constructed in $O(n)$ time.*

From previous propositions, we can derive an upper bound for grid-major area of k-outerplanar graphs, as follows.

Consider a k-outerplane graph H with n vertices. Apply first Proposition 1.2.6 to obtain a new k-outerplane graph H' with $O(n)$ vertices and maximum degree four. Then, apply the orthogonal drawing of [DLT94] to obtain an orthogonal drawing of H' of area $O(k|V(H')|) = O(kn)$ while preserving the embedding of H'. As seen in Section 1.2.2, it is easy to check that any orthogonal drawing actually a minor-drawing of same area. Thus, H' has a minor-drawing of same area. It follows that H is a minor of a grid of area $O(kn)$. We remark that all the steps take polynomial time, and so a minor-drawing of such area can be constructed in polynomial time and preserve the initial embedding of H.

To summarize:

Corollary 1.2.7. *Every n-vertex k-outerplanar graph has a grid-major of area $O(kn)$. Moreover, a minor-drawing of the graph into such a grid and preserving the embedding can be constructed in polynomial time.*

The bound of Corollary 1.2.7 is tight thanks to the next Proposition 1.2.8.

Proposition 1.2.8. *For every $k \geqslant 1$, there exists a k-outerplanar graph with at most n vertices such that every grid-major has area $\Omega(kn)$.*

Proof. Consider a minor-drawing of a graph: the sugbraph induced by all the grid-points of the grid that are either in a super-node or in a super-edge has an

[6] The variant presents therein a sub-cubic major without bounds on its number of vertices.

outerface. The grid-points on this outerface form the *border* of the drawing, and all the grid-points that are inside this border (in other words the grid-points that are not strictly in the outerface) form the *internal area* of the drawing.

For every $k \geqslant 1$ we denote by H_k the 3-connected graph on $3k$ vertices that consists in k nested triangles, as represented in Fig. 1.6.

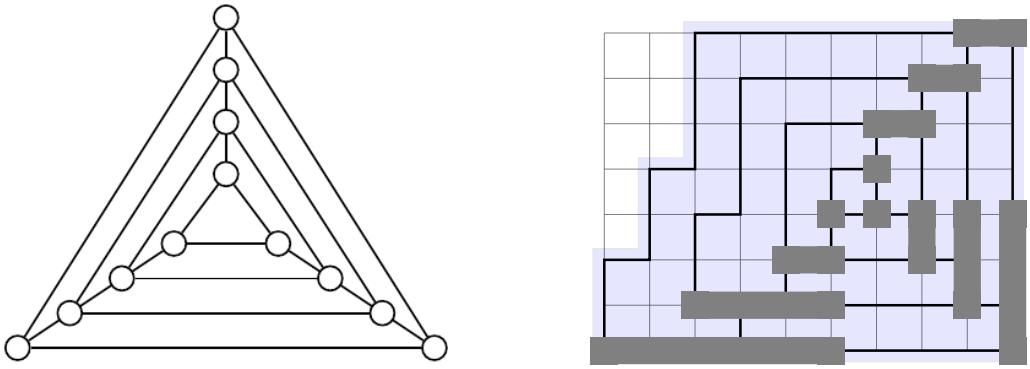


Figure 1.6: H_4 (left) and a minor-drawing of H_4 (right): the blue area is the internal area of the drawing.

Claim 1.2.9. *Every grid minor-drawing of H_k preserving its embedding has internal area more than k^2.*

Proof. We prove this by induction on k. Suppose first that $k = 1$. Then the smallest minor-drawing of the triangle H_1 is on the 2×2-grid and has internal area 4, which is more than 1.

Suppose now that the claim is true for some k. Let Γ be a minor-drawing of H_{k+1} that minimizes the internal area. Let F be the triangle forming the outerface of H_{k+1}: observe that the subgraph of the grid induced by grid-points contained in the super-nodes and super-edges of F contains a cycle C. Since Γ preserves the embedding, the k other nested triangles are drawn inside this cycle, thus Γ contains a subdrawing Γ' of H_k preserving its embedding, which has internal area more than k^2.

Let B be the border of Γ'. We remark that a closed walk on this border is a sequence of more than $2k$ grid-edges. Indeed, let $w, h \geqslant 2$ be such that the Γ' spans over h rows and w columns. The closed walk requires using $h - 1$ vertical edges to go from the highest grid-point to the lowest, and the same amount for the return. Similarly, it requires $2(w - 1)$ horizontal edges to go from the leftmost grid-point to the rightmost and return. Moreover, since the area of Γ is more that k^2, then either h or w, say h, is more than k. Then the whole closed walk requires more than $2(k - 1) + 2(w - 1) = 2k$ grid-edges.

Since C is a cycle of the grid that goes around B, it requires more edges than this closed walk, thus at least $2k+1$ of them. The internal area of Γ contains the grid-points of the subdrawing of H_k, plus those on C, thus is more than $k^2 + (2k+1) = (k+1)^2$.

⌟

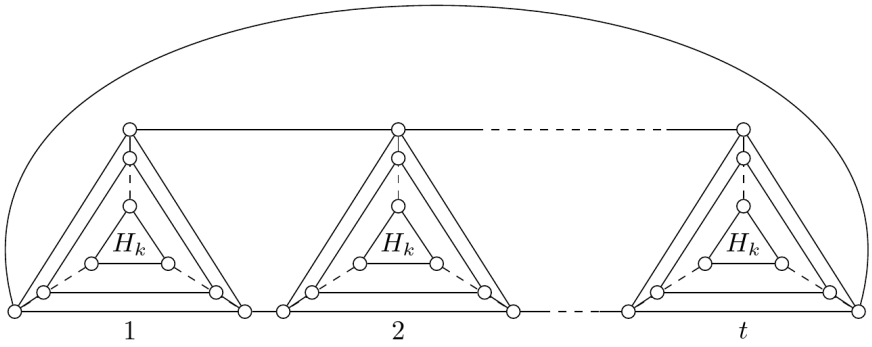


Figure 1.7: A representation of G.

Let $k \geqslant 1$ and $n \geqslant 6k$. Let G be the graph composed of $t = \left\lfloor \frac{n}{3k} \right\rfloor \geqslant 2$ copies of H_k attached forming a chain as in Fig. 1.7. The graph G is k-outerplanar and has at most n vertices. Moreover, G is 3-connected so choosing the outerface fixes the embedding. Let Γ be a minor-drawing of G on a grid of smallest area A, not necessary preserving the embedding. As we saw in Section 1.2.2, we can suppose that the minor-drawing is simple. Observe that no matter the choice of the outerface, Γ preserves the embedding of at least $t-1$ copies of H_k that are subgraphs of G. Thus the minor-drawing of G contains $t-1$ distinct subdrawings, each one is a minor-drawing of H_k preserving its embedding, thus by Claim 1.2.9 of internal area at least k^2. Therefore A is at least $(t-1) \cdot k^2 = \Omega(nk)$.

□

1.2.5 From grid-major to grid-drawing

In this section, we consider the reverse question: considering a graph H that is a minor of the $w \times h$-grid, we want to deduce a grid-drawing of H.

Theorem 1.2.10. *Let H be a graph of maximum degree Δ with a $w \times h$ minor-drawing and let $\delta = \max(1, \lfloor \Delta/2 \rfloor)$. Then, H has a $\delta w \times \delta h$ poly-line drawing, and a $3w \times 3h$ orthogonal drawing if $\Delta \leqslant 4$. Such drawings can be constructed in polynomial time.*

Before proving the theorem, observe that it implies that any lower bound on the area of orthogonal drawing, or poly-line drawing for a graph of bounded max

degree, gives a lower bound on the area of grid-major. Note that [BCE+19] showed an example of a graph, a ladder with a universal vertex inspired from [Bie11], that is minor a grid of area n and for which any grid-drawing required an area of $\Omega(n^2)$. This theorem implies that any such graph for which the area of the grid-major is small compared to the area of the grid-drawing has a vertex of high degree.

Proof. Let H be a planar graph of max degree Δ with a $w \times h$ minor-drawing Γ. Each vertex u of H is represented in Γ by the super-node X_u and each edge (u, v) is represented by a grid-edge whose extremities lie respectively in X_u and X_v.

Suppose first that H has maximum degree $\Delta \leqslant 1$, then H is the disjoint union of isolated edges and isolated vertices. Thus we can trivially deduce a $w \times h$ orthogonal drawing Γ' from Γ: for each edge (u, v) of H, we keep as representatives of u and v only the two vertices respectively in X_u and X_v that are incident to the grid-edge representing (u, v), and for each isolated vertex u, we keep an arbitrary vertex of X_u. For the rest of the proof we can suppose $\Delta \geqslant 2$ and denote $\delta = \left\lfloor \frac{\Delta}{2} \right\rfloor$.

For each vertex u of H of degree $d \leqslant \Delta$, each edge incident to u is represented by a grid-edge $e_i = (u_i, v_i)$ for some $i \in \{0, \dots, d-1\}$, where u_i is the extremity in X_u and v_i the extremity in the super-node corresponding to the neighbor of u. Note that for $i \neq j$, v_i and v_j are distinct but it is possible that $u_i = u_j$.

We define $X'_u = X_u \cup \{e_0, \dots, e_{d-1}\}$, the *extended super-node* of u. Note the extended super-nodes overlap on the edges between super-nodes: for each edge (u, v) of H, there is an edge e in Γ between X_u and X_v and $e = X'_u \cap X'_v$.

Let us first prove the following claim:

Claim 1.2.11. *Let u be a vertex of H of degree d and extended super-node X'_u. There is a rooted tree T_u of root r_u in X'_u such that:*

- *The leaves are exactly $v_0, \dots, v_{d-1}$,*
- *No subtree induced by the descendant of a child of r_u has more than $\left\lfloor \frac{\Delta}{2} \right\rfloor$ leaves.*

Proof. If $d = 0$, then T_u is reduced to r_u, an arbitrary vertex in X_u. If $d = 1$, it suffices to take $V(T_u) = \{v_0, r_u = u_0\}$ and $E(T_u) = \{e_0\}$. If $d = 2$, let P be a path from u_0 to u_1 inside X'_u, take $r_u = u_0$ and $T_u = P \cup \{e_0, e_1\}$. We can suppose for the rest of the proof that $d \geqslant 3$.

Take an unrooted subtree T of X'_u that connects $v_0, \dots, v_{d-1}$: T exists since X'_u is connected. The vertices $v_0, \dots, v_{d-1}$ have degree 1 in X'_u so they are necessary leaves in T, and any other leaf can be deleted from T thus the leaves of T are exactly $\{v_0, \dots, v_{d-1}\}$.

Let us now choose a root for T. Let r be an arbitrary internal node of T and choose r as the root of T. We define $T(n)$ to be the subtree of T induced by the descendants of a node n and nleaves(n) the number of leaves in the subtree $T(n)$.

While r has a child c such that $\text{nleaves}(c) > \lfloor \frac{\Delta}{2} \rfloor$, choose c as the new root of T. Let us prove that this computes a rooted tree with the desired properties. If $T(c)$ has more than $\lfloor \frac{\Delta}{2} \rfloor$ leaves, then $T \setminus T(c)$ has $\ell \leqslant \lfloor \frac{\Delta}{2} \rfloor$ leaves. When we choose c as the new root, c is not a leaf thus c has t children $c_1, \dots, c_t = r$ with $2 \leqslant t \leqslant 4$, and each subtree contains at least one leaf. Observe that $\text{nleaves}(c) = \sum_{i=1}^{t-1} \text{nleaves}(c_i)$, thus for each i, $\text{nleaves}(c_i) \leqslant \text{nleaves}(c)$ and $\text{nleaves}(r) = \ell \leqslant \lfloor \frac{\Delta}{2} \rfloor < \text{nleaves}(c)$. Therefore, each step can only decrease the number of leaves in the biggest subtree, and since we progress toward the leaves, it has to eventually stop. Thus after the final step, we obtain a root such that none of its children induce a subtree with more than $\lfloor \frac{\Delta}{2} \rfloor$ leaves. Then let r_u be this final root and T_u the tree T rooted in r_u. The rooted tree T_u has all the required properties. □

Thanks to Claim 1.2.11, we can create a new representation Γ_1 of H on a $w \times h$-grid where:

- each vertex u of H is associated to the root r_u of the tree T_u,
- each edge (u, u') of H is associated to a path from r_u to $r_{u'}$ in $T_u \cup T_{u'}$.

We define a weight function ω on the edges of T_u such that for an edge (x, y) where x is the parent of y, the number of leaves of $T_u(y)$ is $\omega((x, y))$, where $T_u(y)$ is the subtree of T_u rooted in y. Note that Claim 1.2.11 implies that $\omega((x, y)) \leqslant \delta$ for each edge (x, y).

Notice that in the representation Γ_1, each vertex of H is represented by one vertex of the grid, but the paths representing the edges of H are not edge-disjoint. More precisely, each edge e of $T_u \cup T_{u'}$ is used by $\omega(e)$ such paths. In order to obtain a poly-line drawing, we want to represent the edges of H by vertex-disjoint paths between representatives of the vertices of H, where the edges of the paths are not necessary edges of the grid but are straight lines between vertices of the grid. In order to do that, we want first to make a new drawing Γ_2 on a bigger grid where each T_u of Γ_1 corresponds to a star S_u of center r_u and each branch is a path that corresponds to a path from r_u to a leaf of T_u.

Let us now draw the new representation Γ_2 on a $\delta w \times \delta h$-grid. Number the columns $1, 1 + \frac{1}{\delta}, 1 + \frac{2}{\delta}, \dots, 1 + \frac{\delta-1}{\delta}, 2, 2 + \frac{1}{\delta}, \dots, w + \frac{\delta-1}{\delta}$, and number the rows the same way $1, 1 + \frac{1}{\delta}, 1 + \frac{2}{\delta}, \dots, h + \frac{\delta-1}{\delta}$. In order to create correspondence between the two drawings on grids of different sizes, we will need the following tools to characterize Γ_1.

Notation inside a tree: Consider a vertex u of H: we can use the previous notation inside T_u: the leaves of T_u are $(v_i)_{i \in \{0,\dots,d-1\}}$ where d is the degree of u, $(e_i = (u_i, v_i))_{i \in \{0,\dots,d-1\}}$ are the *final edges* of the tree and are the only edges of T_u

that are in common with other trees of Γ_1. Upon renumbering $(v_i)_{i\in\{0,\dots,d-1\}}$, we can suppose that they are numbered in their order of appearance in an Euler tour of T_u that respects the embedding of T_u and encounters the children anticlockwise.

Port number: Consider a final edge $e_i = (u_i, v_i)$ for some $i \in \{0, \dots, d-1\}$. Let x be the child of r_u on the path from r_u to v_i. The edge (r_u, x) leads to a subtree with $\omega_x = \omega((r_u, x)) \leqslant \delta$ leaves, and there is a constant c_x such that the leaves of $T(x)$ are $v_{c_x}, v_{c_x+1}, \dots, v_{c_x+\omega_x-1}$. Since v_i is a leaf of $T(x)$ there is an integer $p \in \{0, \dots, \omega_x - 1\}$ such that $i = c_x + p$. We define the *port number* of e_i as $port(i) = p = i - c_x$. Observe that the port number of e_i depends only on i and on which subtree at the root it belongs to and that $0 \leqslant port(i) \leqslant \delta - 1$. Note also that in each strict subtree, the port numbers are increasing with respect to the order of the Euler tour.

Direction of an edge: Let x be a vertex of T_u and e an edge of T_u incident to x: since x is a vertex of a grid, it has at most 4 incident edges coming from each direction. We say that the edge e has a *direction with respect to* x that can be `up`, `down`, `left` or `right`. For example, if e has direction `left` with respect to x, it means that e is on the left of x in the representation. Observe that for each $e = (x, x')$, the direction of e with respect to x is the opposite direction of that of e with respect to x'.

Switch: Let x be a vertex of T_u of coordinate (i, j) in Γ_1, such that x is not r_u nor a leaf of T_u. In Γ_2, we define the *switch* of x, noted switch(x), the subgrid induced by the vertices of coordinates $(i + \frac{k}{\delta}, j + \frac{k'}{\delta})_{0\leqslant k,k'\leqslant\delta-1}$, with what we call *gates* on the vertices of the border and that we detail right after. There are four kinds of switches, one for each direction. The *direction of the switch* of x is the direction with respect to x of the edge coming from the parent of x. The border of the switch has four sides (one in each direction), and for each side $c \in \{\texttt{up}, \texttt{down}, \texttt{left}, \texttt{right}\}$, for each integer $p \in \{0, \dots, \delta - 1\}$ we define a *gate* of the switch(x), denoted gate(x, c, p) as follow:

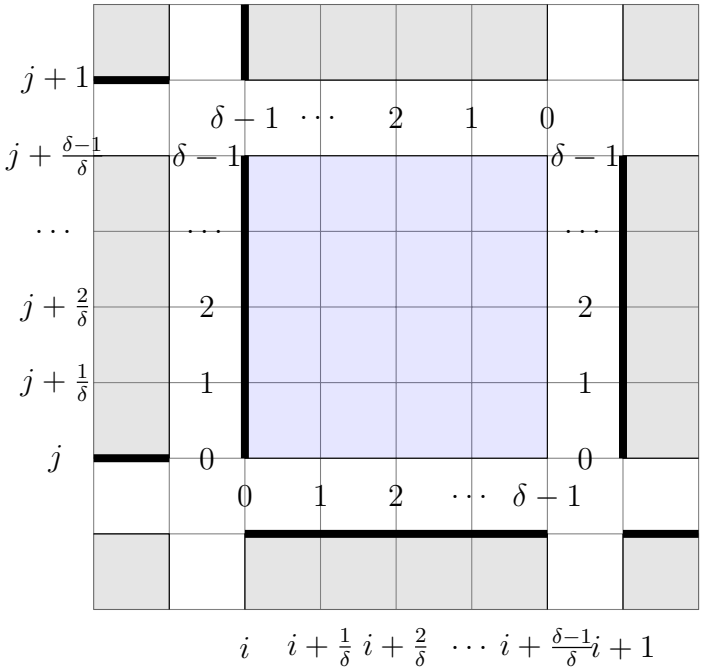


Figure 1.8: Representation of a `left` switch for a vertex of coordinates (i, j) in Γ_1.

Suppose that the direction of the switch of x is `left`, as represented on Fig. 1.8.

$$\begin{aligned}
\text{gate}(x, \texttt{left}, p) &= (i, j + \frac{p}{\delta}) \\
\text{gate}(x, \texttt{down}, p) &= (i + \frac{p}{\delta}, j) \\
\text{gate}(x, \texttt{right}, p) &= (i + \frac{\delta - 1}{\delta}, j + \frac{p}{\delta}) \\
\text{gate}(x, \texttt{up}, p) &= (i + \frac{\delta - 1 - p}{\delta}, j + \frac{\delta - 1}{\delta})
\end{aligned}$$

We say that p is the *gate number*. Observe that on the side of the direction of the switch, the gates (that will be called *in-gates*) are ordered with increasing number clockwise around the switch, and on each other side, the gates (that can be called *out-gates*) are ordered anti-clockwise. All the switches respect this property no matter their direction: any kind of switch can be obtained by rotating the one described here so that the particular side matches the direction of the switch.

Claim 1.2.12. *Let $e = (x, x')$ be an edge of T_u where x is the parent of x', incident to x and x' with direction respectively c and c'. Then for all $p \in \{0, \dots, \delta - 1\}$, $\text{gate}(x, c, p)$ and $\text{gate}(x', c', p)$ are adjacent.*

Proof. Let $e = (x, x')$ be an edge of T_u where x is the parent of x'. We have seen earlier that e is incident to x and x' with opposite directions c and c'. Since x and x' are adjacent in Γ, their switches are next to each other in Γ_2, such that the side c of switch(x) is facing side c' of switch(x'). Since e is the edge coming from the parent of x', switch(x') has direction c' and switch(x) has direction different than c. This implies that the gates on the side c of switch(x) are ordered anti-clockwise around the center of the switch and the gates on the side c' of switch(x') are ordered clockwise. Therefore, for all $p \in \{0, \dots, \delta - 1\}$, gate$(x, c, p)$ and gate(x', c', p) are adjacent. □

Let R_u be the vertex in Γ_2 with the same coordinates as r_u in Γ_1. At this point, we have a treelike disposition of the switches in Γ_2, with a R_u as the root. The aim is to deduce from this representation the star S_u whose branches are constructed using the switches.

For each $k \in \{0, \dots, d-1\}$ let Q_k be the path in T_u from r_u to v_k, and let $t_k = R_u$ if $u_k = r_u$, $t_k = \text{gate}(u_k, c, p)$ otherwise, with p the port number of e_k and c the direction of e_k with respect to u_k.

Claim 1.2.13. *There is a collection of paths $(P_k)_{k\in\{0,\dots,d-1\}}$, each from R_u to t_k, that intersect each other only in R_u.*

Proof. For each $k \in \{0, \dots, d-1\}$, construct P_k as follow. Let p be the port number of k, let ℓ be the length of Q_k, and $Q_k = x_0 - x_1 - \cdots - x_\ell$ with $r_u = x_0$ and $v_k = x_\ell$. By construction, $\ell > 0$.

If $\ell = 1$, then $u_k = r_u$ and $Q_k = r_u - v_k$, thus $t_k = R_u$ and P_{v_k} is reduced to R_u.

If $\ell \geqslant 2$, for each $s \in \{1, \dots, \ell - 1\}$ we put $c_{s,in}$ and $c_{s,out}$ the directions of respectively (x_{s-1}, x_s) and (x_s, x_{s+1}) with respect to x_s. We construct by induction the path P_k^s from R_u to gate$(x_s, c_{s,out}, p)$.

The path P_k^1 is the path $R_u -$ gate$(x_1, c_{1,in}, p) -$ gate$(x_1, c_{1,out}, p)$. Let $s \in \{1, \dots, \ell - 1\}$ and suppose that P_{s-1} is constructed from R_u to gate$(x_{s-1}, c_{s-1,out}, p)$. Then P_s is constructed as the union of P_{s-1} and the path gate$(x_{s-1}, c_{s-1,out}, p) -$ gate$(x_s, c_{s,in}, p) -$ gate$(x_s, c_{s,out}, p)$. At the end of this construction $s = \ell - 1$ thus gate$(x_{\ell-1}, c_{\ell-1,out}, p) = t_k$ and $P_k = P_k^{\ell-1}$.

Let us show now that these paths intersect each other only in R_u. Suppose that there exist k, k' such that P_k and $P_{k'}$ intersects.

Case 1: Suppose first that they intersect in a common vertex $v \neq R_u$ of the paths. v is a vertex of $P_k \setminus R_u$ thus there is a vertex x in $Q_k \setminus r_u$ and a direction c such that $v = \text{gate}(x, c, p)$ where p is the port number of e_k. Similarly, v is a vertex of $P_{k'} \setminus R_u$ thus $x \in Q_{k'} \setminus r_u$ and p is also the port number of $e_{k'}$. By definition, two edges have the same port number if and only if $k = k'$, or the leaves v_k and

$v_{k'}$ are not on the same subtree rooted in a child of r_u, which is impossible since they have x as common ancestor.

Case 2: Suppose now that they intersect in a point that is not a vertex of the paths. Thus there are two edges $f_k \in E(P_k)$ $f_{k'} \in E(P_{k'})$ that cross each other. By Claim 1.2.12, an edge between two switches is an edge of the grid, thus $f_k, f_{k'}$ cannot be both edges between switches. This implies that one of them, w.l.o.g. f_k, is either incident to R_u or inside a switch.

Suppose first that f_k is incident to R_u. Let (i, j) be the coordinates of R_u and $N(r_u)$ the list of children of r_u and for each child $x \in N(r_u)$, c_x is the direction of the switch of x. Let G_u be the subgraph of the grid induced by $(i + \frac{i'}{\delta}, j + \frac{j'}{\delta})_{0 \leqslant i', j' \leqslant \delta - 1} \cup (\text{gate}(x, c_x, p))_{x \in N(r_u), p \in \{0, \dots, \delta - 1\}}$: it is the part of the grid around R_u where vertices are in no switch, except the ones on the border that are the in-gates of the switches of the children of r_u (Fig. 1.9 gives a representation of G_u with $\delta = 3$). Therefore, f_k is an edge from R_u to an in-gate on the border of G_u, and crosses $f_{k'}$ somewhere in G_u thus where there is no switch. However, by construction and Claim 1.2.12, there is no edge outside a switch between two gates that are not adjacent in the grid. Thus $f_{k'}$ has to be also an edge incident to R_u. Thus f_k and $f_{k'}$ intersect in R_u, which contradicts the hypothesis.

Suppose now that there is vertex x such that f_k and $f_{k'}$ intersect inside switch(x). Let $p = port(k)$ and $p' = port(k')$, and c_{in} the direction of the switch of x. Since f_k and $f_{k'}$ are inside switch(x), by construction of P_k and $P_{k'}$, there are c and c' such $f_k = (\text{gate}(x, c_{in}, p), \text{gate}(x, c, p))$ and $f_{k'} = (\text{gate}(x, c_{in}, p'), \text{gate}(x, c', p'))$. Suppose w.l.o.g. that $k < k'$. Then $p < p'$ since v_k and $v_{k'}$ have x as common ancestor and thus are on the same subtree at the root. This implies that $\text{gate}(x, c_{in}, p)$ is before $\text{gate}(x, c_{in}, p')$ clockwise $\text{gate}(x, c, p)$ is before $\text{gate}(x, c', p')$ anticlockwise. Indeed, if $c = c'$, the out-gates on a same side are ordered anti-clockwise, and if $c \neq c'$, since the leaves of T_u are ordered in their order of appearance in an Euler tour, c is before c' anticlockwise thus $\text{gate}(x, c, p)$ is also before $\text{gate}(x, c', p')$ anticlockwise in this situation. Therefore, $f_k = (\text{gate}(x, c_{in}, p), \text{gate}(x, c, p))$ and $f_{k'} = (\text{gate}(x, c_{in}, p'), \text{gate}(x, c', p'))$ do not intersect. □

At this point, for each vertex u of degree d of H, there is a star S_u in Γ_2 whose center is R_u and whose branches are $(P_k)_{1 \leqslant h \leqslant d}$. It remains to connect the stars together to obtain the poly-line drawing of the H.

Claim 1.2.14. *For each edge (u, u') in H, there is a path $P_{u,u'}$ from R_u to $R_{u'}$ in Γ_2 and all those paths do not intersect each other except in their extremities.*

Proof. Let (u, u') be an edge of H. There are k, k' such that:

- T_u has a leaf v_k such that $e_k = (u_k, v_k)$ corresponds to (u, u') in the minor-drawing, with $u_k \in X_u, v_k \in X_{u'}$;
- $T_{u'}$ has a leaf $v_{k'}$ such that $e'_{k'} = (u'_{k'}, v'_{k'})$ corresponds to (u, u') in the minor-drawing, with $u_{k'} \in X_{u'}, v'_{k'} \in X_u$

Observe that e_k and $e'_{k'}$ both represent (u, u') thus $u_k = v'_{k'}$ and $v_k = u'_{k'}$. Let p be the port number of e_k (resp. p' the port number of $e'_{k'}$), and P_{v_k} (resp. $P_{v'_{k'}}$) be the path as described in Claim 1.2.13. Let us show that we can add an edge e (not necessarily a grid-edge) to the drawing Γ_2 so that $P_{v_k} \cup e \cup P_{v'_{k'}}$ is a path from R_u to $R_{u'}$.

Assume first that $P_{v_k} \neq \{R_u\}$ and $P_{v'_{k'}} \neq \{R_{u'}\}$. The last vertices of those paths are respectively gate(u_k, c, p) and gate$(u'_{k'}, c', p')$, with c and c' the direction of e_k with respect to respectively u_k and $u'_{k'}$. Since u_k and $u'_{k'}$ are the extremities of e_k, the side c of switch(u_k) and the side c' of switch$(u'_{k'})$ are facing, even if the gates are not facing gates with same number (as they are all out-gates). Here we add the edge e that connects gate(u_k, c, p) and gate$(u'_{k'}, c', p')$ without necessarily $p = p'$. There is no other edge between those two switches, thus e intersects no other edge, and from Claim 1.2.13, neither P_{v_k} nor $P_{v'_{k'}}$ intersect another path. Thus $P_{u,u'} = P_{v_k} \cup e \cup P_{v'_{k'}}$ is a path from R_u to $R_{u'}$ that does not intersect other path.

Suppose now w.l.o.g. that $P_{v_k} = \{R_u\}$ and $P_{v'_{k'}} \neq \{R_{u'}\}$, then gate$(u'_{k'}, c', p') =$ gate(v_k, c', p') is a gate on the border of G_u (as described in the proof of Claim 1.2.13), thus the edge $e = (R_u, \text{gate}(u'_{k'}, c', p'))$ intersects no other path and $P_{u,u'} = e \cup P_{v'_{k'}}$ is a path from R_u to $R_{u'}$ that does not intersect other path.

Finally, if $P_{v_k} = \{R_u\}$ and $P_{v'_{k'}} = \{R_{u'}\}$, then r_u and $r_{u'}$ are adjacent thus $P_{u,u'} = (R_u, R_{u'})$ is a path from R_u to $R_{u'}$ that does not intersect other paths. □

For each vertex u of H, there is a corresponding vertex R_u in Γ_2 and for each edge (u, u') of H, there is a path from R_u to $R_{u'}$ in Γ_2, and those paths are vertex-disjoint. Thus Γ_2 is a poly-line drawing of H on a grid $\lfloor \Delta/2 \rfloor\, w \times \lfloor \Delta/2 \rfloor\, h$.

Adaptation of the construction to an orthogonal drawing. Suppose now that the maximum degree of H is $\Delta \leqslant 4$, let us show that we can obtain an orthogonal drawing of H on a $3w \times 3h$ grid. Let us first construct Γ_1 the same way as previously and Γ_2 almost the same way, but considering $\delta = 3$ instead of $\delta = \left\lfloor \frac{\Delta}{2} \right\rfloor$, with the same notation. Observe that the port numbers are in $\{0, 1\}$ even if the gates are constructed with gate numbers in $\{0, 1, 2\}$: the gates numbered 2 are not used.

In order to construct an orthogonal drawing of H from Γ_2, we need to replace every oblique edge in Γ_2 by a path of the grid, which is why we need a bigger grid.

Let e be an oblique edge, there are 3 possibilities:

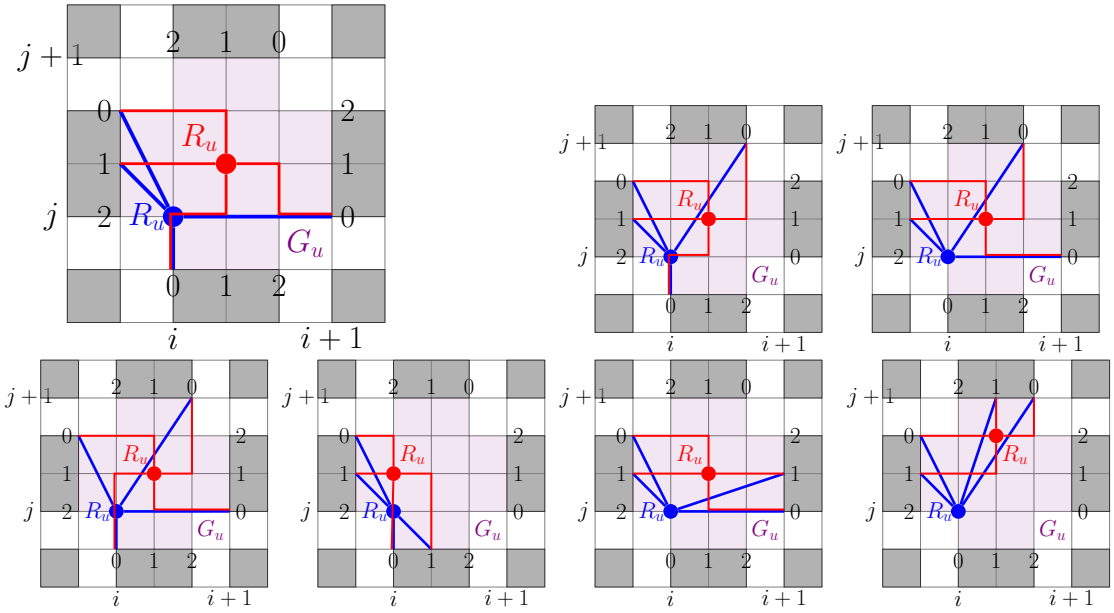


Figure 1.9: The seven principal configurations of the edges incident to R_u (in blue) and their corresponding transformations (in red).

1. e is an edge incident to R_u for some u in H;
2. e is inside a switch;
3. e is an edge that connects two stars (as defined in Claim 1.2.14);

Case 1: Let R_u be the vertex corresponding to u in Γ_2. Let us modify the edges inside G_u (the part of the grid around R_u) so that there is no oblique edge incident to R_u.

Since $\Delta \leqslant 4$, there are at most 4 edges incident to R_u and at most two of them incident to a same switch, thus there are only 7 configurations to consider. Any other configuration is a rotation of one of those (with a modification of the initial emplacement of R_u inside G_u that does not affect the result), or a simplification by considering fewer edges. Those configurations are depicted (in blue) in Fig. 1.9 as well as how to modify them (in red) to get rid of the oblique edges. The gray areas are parts of the potential switches around R_u, at coordinates (i, j), and the violet area represents G_u. Observe that the modifications are restricted to G_u and thus do not affect the rest of the drawing.

Case 2: Let x be a vertex of Γ_1, let us modify the edges inside switch(x) to get rid of oblique edges. We can consider w.l.o.g. that the direction of the switch is `left` (any other direction is a rotation of this situation), and that there are

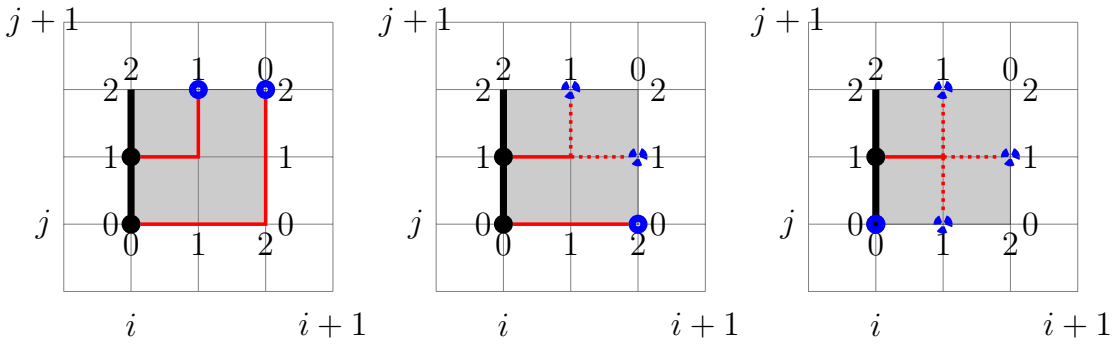


Figure 1.10: Different possibilities for two paths to go through a switch `up`, and how it is transformed into an orthogonal path.

two paths passing through the switch (otherwise we consider only the edge that actually exists). Recall that a path enters and gets out of a switch through gates of same number. The different configurations are represented in Fig. 1.10, one figure for each choice of out-gate for the path using gates numbered 0. The black points are the in-gates used, the blue points the out-gates, and how to connect them without oblique edges in red (when there are multiple choices for the out-gate numbered 1, they are represented with dotted lines).

Case 3: Let e be an edge added between two stars: that is the only case where an edge between two switches can be oblique, because that is the only case where an edge can be between out-gates with potentially different port numbers. There are two vertices x, x' such that $e = (\text{gate}(x, c, p), \text{gate}(x', c', p'))$ with c, c' being the two opposite directions of (x, x') in Γ_1. If $p = p' = 1$, then e is not oblique, thus w.l.o.g. $p = 0$ and p' is either 0 or 1. Recall that e is an edge added to connect stars thus x and x' were each a leaf of a tree of Γ_1 thus $\omega((x, x')) = 1$ and there is no other edge between the switch of x and the switch of x'. Therefore $\text{gate}(x, c, 1)$ is unused and e can be replaced by the path of the grid $\text{gate}(x, c, 0) - \text{gate}(x, c, 1) - \text{gate}(x', c', 1)$ if $j = 1$. If $j = 0$, $\text{gate}(x', c', 1)$ is also unused and e can be replaced by the path $\text{gate}(x, c, 0) - \text{gate}(x, c, 1) - \text{gate}(x', c', 1) - \text{gate}(x', c', 0)$.

By repeating these transformations for all the oblique edges of the drawing, we obtain a new drawing where each vertex u of H is represented by a vertex R_u of the grid and each edge of H is represented by a path of the grid, thus an orthogonal drawing of H. □

1.3 Universal graph for bounded genus graphs

1.3.1 Overview

In this section, we show the following:

Theorem 1.1.2. *For every n and every surface Σ of Euler genus g, there is a graph $\mathcal{U}_{n,\Sigma}$ embedded on Σ with $O(g^2(n+g)^2)$ vertices that contains as minor every graph embeddable on Σ with n vertices.*

This is a generalisation of the theorem of Robertson, Seymour, and Thomas [RST94], stating that there is a $4n^2$-vertex planar graph minor-universal for planar graphs, to the graphs on surfaces.

We start with Section 1.3.2 by giving the necessary definitions in order to manipulate those graphs and surfaces. Then, in Section 1.3.3, we sketch the proof and how it splits into two lemmas. The proof of the first one, simplifying the graph class, is distributed into two parts: Section 1.3.4 gives a part of the proof considering only plane graph, which result into a constructive alternative proof of that of Robertson, Seymour and Thomas, then in Section 1.3.5 we reuse the arguments and adapt them in the case of graph on surfaces. Finally, in Section 1.3.6, we prove that all the graphs of this class are minors of the minor-universal graph we describe.

1.3.2 Preliminaries

We consider compact connected surfaces without boundary. A *canonical system of loops* for a surface Σ of Euler genus g is a one-vertex graph embedded on Σ with g loop-edges such that by cutting Σ along these loops results in a surface homeomorphic to a polygon P with $2g$ sides. The way the pairs of sides of P are reattached to form Σ can be described by a *signature*, a word associating clockwise one symbol with each side of P. The polygon P and its signature form a *canonical polygonal schema*. Note that the loops are either two-sided (the surface Σ is orientable with oriented genus $g/2$) or one-sided (Σ is non-orientable with non-oriented genus g).

For $g > 0$, the signature is $a_1a_2\bar{a_1}\bar{a_2}\dots a_{g-1}a_g\bar{a}_{g-1}\bar{a}_g$ if Σ is orientable, and $a_1a_1\dots a_ga_g$ if Σ is non-orientable. For the sphere, $g = 0$, the signature is $a_0\bar{a_0}$ by convention. Sides are reattached if they have the same (or opposite) symbol in the signature, given the convention that positive symbols correspond to a clockwise orientation of the sides, and negative symbols to an anticlockwise orientation.

Consider a graph G embedded[7] on some surface Σ with a canonical system of

[7] We only consider simple graphs and cellular embeddings, i.e., where each face is homeomorphic to a disc.

loops L. Denote by $G \uplus L$ the graph embedded on Σ obtained from the union of G and L, and by adding a new vertex at each intersection between G and L.

A *polygonal embedding* for G is the planar embedding Π into a polygon obtained from $G \uplus L$ by cutting Σ along the edges of L. (We refer to Fig. 1.11 for an illustration.) The edges and vertices on L appear duplicated on the boundary of outerface of Π. More precisely, the boundary of the outerface of Π is a cycle that can be cut into a clockwise sequence of paths sharing their extremities. These paths, called *sides* of Π, correspond to the edges of the polygon P, and their extremities, called *corners* of Π, corresponds to the vertices of P. Note that corners have degree two in Π. The ordered sequence of corners is called the *border* of Π. The *signature* of Π, denoted by $\sigma(\Pi)$, describes how the sides of Π are reattached to form $G \uplus L$ on Σ. More precisely, the ith symbol of the signature is associated with the ith side of Π, i.e., the path between the ith and $(i+1)$th corners. These sides are merged according to the orientation given by the symbols, and there are called *twin sides* of Π. Obviously, twin sides must contain the same number of vertices. Furthermore, the vertices and edges that are identified in this process are called respectively *twin vertices* and *twin edges*.

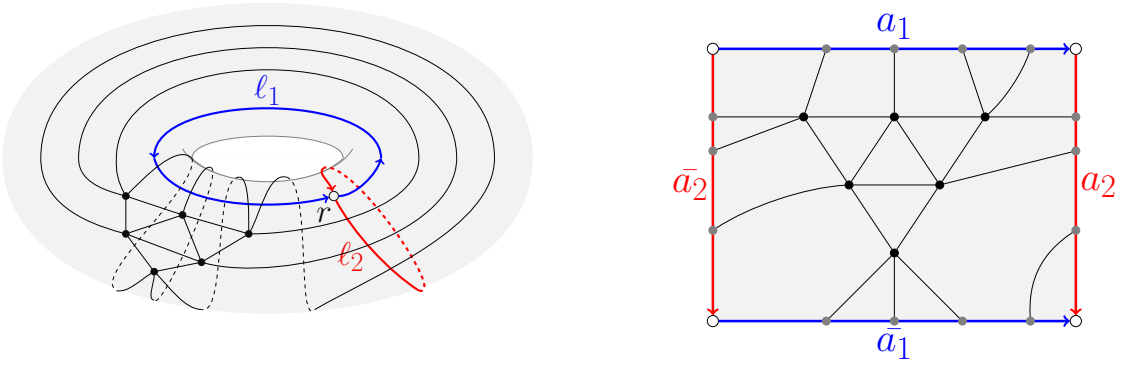


Figure 1.11: The graph K_6 embedded on the torus (of Euler genus 2) with the system of loops $L = (\{r\}, \{\ell_1, \ell_2\})$, and a polygonal embedding Π of K_6 of signature $\sigma = a_1 a_2 \bar{a_1} \bar{a_2}$. By construction, K_6 is a minor of the graph sewing$(\Pi) = K_6 \uplus L$.

Given a polygonal embedding Π, the *sewing* of Π, denoted by sewing(Π), is the graph obtained by reattaching the sides of Π according to its border and its signature. It is clear that, if Π comes from $G \uplus L$ embedded on Σ by cutting along the edges of L, then sewing(Π) is isomorphic to $G \uplus L$. In particular, sewing(Π) embeds on Σ and contains G as minor.

1.3.3 Main result

To prove Theorem 1.1.2, we consider a graph G embedded on the surface Σ of Euler genus g and having a canonical systems of loops L with vertex r and signature σ. Thus G has a polygonal embedding Π_G such that G is a minor of sewing(Π_G), a graph that embeds on Σ. Note that $|V(\Pi_G)| \leqslant 2|V(G \uplus L)| + |\sigma|$ since every vertex of $G \uplus L$ appears on at most two sides of Π_G, and each corner is a copy of r. In fact, Π_G has at most n internal vertices due to G, and a certain number of vertices lying on the sides of Π_G due to r and to the intersections between G and L. We can bound this number thanks to the following result[8]:

Lemma 1.3.1 ([LPVV01, FHdM22]). *Given an embedding of G on a surface Σ of Euler genus g, there is a polynomial time algorithm that computes a canonical system of g loops such that each loop intersects any edge of G in at most 4 points if Σ is orientable [LPVV01], and at most 30 points if Σ is non-orientable [FHdM22].*

It follows that $|V(\Pi_G)| = O(|V(G)| + g|E(G)|) = O(g(n+g))$ since each of the g loops of L crosses $O(1)$ times each of the at most $3n + 3g - 6$ edges[9] of G. Our strategy then is to transform step by step the initial polygonal embedding Π_G into some unique polygonal embedding $\mathbb{U}$ that preserves signature and minor containment of its sewing. The number of vertices of this final embedding is $|V(\mathbb{U})| = O(|V(\Pi_G)|^2)$. Our minor-universal graph $\mathcal{U}_{n,\Sigma}$ as in Theorem 1.1.2 is nothing else than sewing$(\mathbb{U})$ that has therefore less than $|V(\mathbb{U})| = O(g^2(n+g)^2)$ vertices as claimed.

To formalize signature and minor containment preserving, we introduce the following relation. A polygonal embedding Π is a *p-minor* of a polygonal embedding Π' if they have the same signature and if sewing(Π) is a minor of sewing(Π'). Moreover, we say that Π has *size* (m, n) if each side has at most m vertices, the corners excluded, and at most n internal vertices (those that are not lying on the boundary of the outerface of Π). In particular, every polygonal embedding Π of size (m, n) has at most $n + |\sigma(\Pi)|(m+1)$ vertices[10], and sewing(Π) has at most $n + |\sigma(\Pi)|(m+1)/2$ vertices. This is even at most $n + |\sigma(\Pi)|m/2 + 1$ vertices if $\sigma(\Pi)$ is *minimal*, i.e., comes from a system of loops (see Section 1.3.6). A polygonal embedding of size of $(m, 0)$ is a Hamiltonian outerplane graph.

The final polygonal embedding is denoted hereafter by $\mathbb{U}_{\sigma,m}$, where σ is any signature and m an integral parameter. We will see later in Section 1.3.6 that $\mathbb{U}_{\sigma,m}$ has signature σ and size $(m, O((|\sigma|m)^2))$. Roughly speaking, it is a square

[8] We refer to [CdV21][Theorem 8.1] for a reformulation of the original statement of [LPVV01].

[9] From Euler's Formula in simple connected graphs of Euler genus g, that is $n - m + f = \chi(\Sigma) = 2 - g$, and from the fact that $3f \geqslant 2m$.

[10] $|\sigma(\Pi)|$ denotes the number of symbols in $\sigma(\Pi)$.

half-grid cut by its diagonal. The vertices on the boundary of the outerface are those of the diagonal, see Fig. 1.12 for an example.

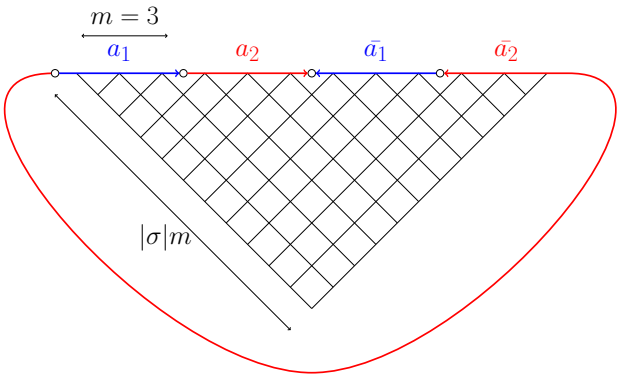


Figure 1.12: The polygonal embedding $\mathbb{U}_{\sigma,m}$, with $m = 3$ and signature $\sigma = a_1 a_2 \bar{a_1} \bar{a_2}$.

The technical theorem is the following. Note that it applies to any polygonal embedding, not only those of canonical signature.

Theorem 1.3.2. *Every polygonal embedding Π of size (m, n) is a p-minor of $\mathbb{U}_{\sigma(\Pi),m+2n}$.*

Before proving Theorem 1.3.2, let us show that it implies Theorem 1.1.2. Consider a graph G on n vertices embedded on a surface Σ of Euler genus g. By Lemma 1.3.1, G has a polygonal embedding Π_G whose signature σ depends only on Σ (as it corresponds to that of the canonical polygonal schema of Σ), and such that Π_G has size (m, n) with $m = O(n + g)$. Moreover, G is a minor[11] of sewing(Π_G). By Theorem 1.3.2, Π_G is a p-minor of $\mathbb{U}_{\sigma,m+2n}$, which by definition of p-minor implies that sewing(Π_G) is a minor of sewing$(\mathbb{U}_{\sigma,m+2n})$. It follows that G is a minor of sewing$(\mathbb{U}_{\sigma,m+2n}) = \mathcal{U}_{n,\Sigma}$, that has the required properties: it is embedded on Σ and has $O(g^2(n + g)^2)$ vertices.

To prove Theorem 1.3.2, we proceed in two steps, which are summarized and formalized by the next two lemmas whose proofs are sketched hereafter and fully proved later:

Lemma 1.3.3. *Every polygonal embedding Π of size (m, n) is a p-minor of a polygonal embedding of size $(m + 2n, 0)$.*

Intuition is that all internal vertices of Π can be somehow pushed to the sides of the outerface, without increasing much the number of vertices on each side. It

[11] In fact Π_G is a minor of sewing$(\Pi_G) \setminus \{r\}$ as the unique vertex r of system of loops is not part of G and can be removed.

follows we can consider that all the vertices belong to the outerface of the polygonal embedding making it outerplanar. This is done by first cutting along the edges of some spanning forest of the inner graph (rooted in a vertex of the outerface), creating an empty space and transforming the graph into an outerplanar graph. Along this process, each vertex that is not a leaf appears more than once (the number of new vertices is bounded by the number of edges of the forest, thus by n). To keep track of this vertex duplication, we add some edges that could be contracted later to get back to the original polygonal embedding. Other transformations are then needed to embed those new edges and obtain an outerplanar polygonal embedding of the right size and containing Π as a p-minor.

Lemma 1.3.4. *Every polygonal embedding Π of size $(m, 0)$ is a p-minor of $\mathbb{U}_{\sigma(\Pi),m}$.*

Note that Theorem 1.3.2 is a straightforward combination of Lemma 1.3.3 and Lemma 1.3.4.

The proof of Lemma 1.3.4 is inspired from the embedding into a $n \times n$-grid for hamiltonian planar graphs with n vertices [RST94, Theorem (1.3)]. From this embedding, one can easily show that outerplanar graphs with n vertices are minors of a half-$n \times n$-grid. In our context, the diagonal of this half-grid is the place for the sides of our polygonal embedding, as depicted in Fig. 1.12. Interestingly, along the way, we give in Proposition 1.3.8 an alternative proof of a result of [RST94, Theorem (1.4)] that states that every planar graph with n vertices is minor of a Hamiltonian planar graph with $2n$ vertices.

Organization of the proofs

Similarly to the notion of major, we say that polygonal embedding Π_G is a *p-major* of Π_G, if Π_H is a p-minor of Π_G.

Because pthe lane graphs we will consider are actually polygonal embeddings, the boundary of the outerface is a cycle. Whenever we talk about outerface we refer to this cycle. We say that two plane graphs H, G such that H is minor of G have the *same outerface* if the minor H has been obtained from G without removing vertices or edges of the outerface of G, nor contracting edges with both endpoints on the outerface of G. It follows:

Property 1.3.5. *If a polygonal embedding Π_H is a minor of polygonal embedding Π_G with same outerface, border and signature, then Π_H is a p-minor of Π_G.*

Recall that a polygonal embedding is nothing else than a plane graph with a given cycle outerface, border and signature. Property 1.3.5 allows us to consider that we work on plane graphs, as long as we keep the same outerface and border when constructing a major.

In Section 1.3.4, we consider plane graphs with a fixed cycle outerface, i.e., with an outerface whose boundary is a cycle. We give tools to construct a plane major that preserves this outerface. These tools will be used later in the construction of the p-major of Π in Section 1.3.5.

1.3.4 Considering plane graphs

Let G be a planar graph G with a given planar embedding and having a cycle outerface O. We denote by $n = |V(G)| - |V(O)|$ the number of internal vertices of G.

Because the boundary of the outerface of G is a connected graph (a cycle), we can assume w.l.o.g. that G is connected as well. If not, we can for instance triangulate all the faces except the outerface. By doing this we get a plane major of G without altering its number of vertices nor its cycle outerface.

Let $\mathcal{F}$ be a forest of G, composed of k trees denoted by $T_1, \ldots, T_k$, and constructed as follows:

- Start with a spanning tree T of G;
- Remove one edge of any path in T that connects two vertices of O, if any;
- Remove all remaining isolated vertices that are in O; and
- Root each T_i at the only one vertex of $T_i \cap O$, denoted by r_i.

Observe that $\mathcal{F}$ spans all internal vertices of G, thus $|V(F)| = n + k$ and $|E(F)| = n$. Because we consider a planar major of G, it is quite tempting to add edges to force $\mathcal{F}$ to be a single tree. However, we cannot do that and suppose that the inner vertices induced a connected subgraph of G, as there could be an edge between two vertices of the outerface separating the internal vertices into two parts.

We now construct a major of G, denoted by G_1, that consists in *blowing up* each tree T_i into a graph $T_i' = C_i \cup E_i$, composed of a cycle C_i plus a set E_i of extra edges. More precisely, blowing up T_i consists in traversing the tree according to a plane Euler tour from r_i (or in other words, a walk along the boundary of the outerface of T_i considered here as a single plane graph). The cycle C_i is constructed iteratively by adding a new vertex at each vertex of T_i visited along this tour, two consecutive vertices on this tour being connected by an edge. See Fig. 1.13 for an illustration. Since $\mathcal{F}$ has no isolated vertices, T_i has at least one edge. If T_i has exactly one edge, then C_i consists of one single edge. The set E_i is composed of all the edges connecting any two vertices u, v of C_i if they corresponds to the same visited vertex of T_i and that, among them, appear consecutive during the visit. In other words, we add an edge between u, v in C_i if they corresponds to same visited vertex w of T_i and that none of the vertices between u and v in C_i corresponds to w.

This completes the description of blowing up T_i into T'_i.

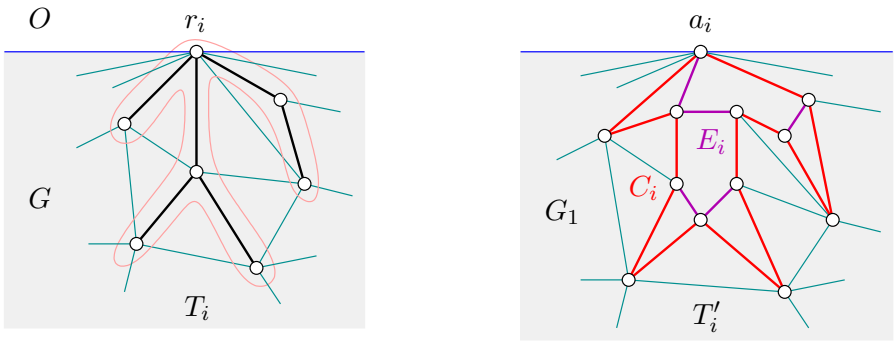


Figure 1.13: Blowing up the tree T_i (with black edges) into $T'_i = C_i \cup E_i$ (with red and violet edges). The root r_i of T'_i becomes an anchor a_i of C_i in T'_i.

The graph G_1 is then obtained from G where each T_i is replaced by T'_i which is outerplanar. This is possible because the walk along the boundary of the outerface of T_i and T'_i are isomorphic, so T'_i can be plugged into T_i by keeping the plane embedding of G.

Claim 1.3.6. *The graph G_1 has the following properties:*

- *G_1 is a plane major of G with same outerface O;*
- *$C_1, \dots, C_k$ form a partition of internal vertices of G_1.*
- *Each C_i contains exactly one vertex in O, called its* anchor*;*
- *G_1 has $2n - k$ internal vertices;*

Proof. It is easy to check that T'_i is a major of T_i, as it suffices to contract all edges of E_i to obtain a graph isomorphic to T_i. By doing this for every T'_i in G_1 we get exactly G (up to some isomorphism) since these are the only differences.

We have seen that the trees T_i's of $\mathcal{F}$ form a partition of the internal vertices of G. It follows that T'_i forms a partition of the internal vertices of G_1, and C_i as well since $V(T'_i) = V(C_i)$.

Each T'_i intersects O in exactly one vertex, its anchor, the vertex corresponding to r_i in T_i. In particular, no edge of T'_i belongs to O. So contracting edges of E_i to get T_i cannot affect the outerface, and thus G_1 and G have the same outerface O.

To obtained C_i from T_i, each edge of T_i is traversed twice. It follows that $|V(C_i)| = 2|E(T_i)|$. We have seen that $|E(F)| = \sum_{i=1}^{k} |E(T_i)| = n$. Therefore, the number of internal vertices of G_1 is $\sum_{i=1}^{k}(|V(C_i)| - 1) = 2n - k$ as we need to remove the anchors. □

We now construct from G_1 a major G_2, that consists in applying a *splitting anchor* operation at every anchor of G_1. More precisely, consider a cycle C_i with anchor a_i in G_1. Let u, w_1, w_2, v be the neighbors of a_i, taken in this cyclically ordered around a_i, such that u, v belong to O and w_1, w_2 to T'_i. By construction u, a_i, v belong to the boundary of a common face, as well as w_1, a_i, w_2. The edge $a_i - w_1$ belongs to C_i, and $a_i - w_2$ belongs to either C_i or E_i, depending whether r_i has one or several children, and $w_1 = w_2$ if C_i consists in one edge. The splitting operation consists in replacing a_i by an edge $a'_i - a''_i$, called *edge-anchor* of a_i, and reconnecting to a'_i all the a_i's neighbors going from u to w_1, and reconnecting to a''_i all the a_i's neighbors going from w_2 to v. We denote by P_i the path going from a'_i to a''_i in C_i, and by F_i the new set of edges E_i in G_2. Note that C_i and P_i differ by a_i, a'_i, a''_i, and that E_i and F_i differ by a''_i in G_2 that plays the role of a_i in G_1. See Fig. 1.14 for an illustration.

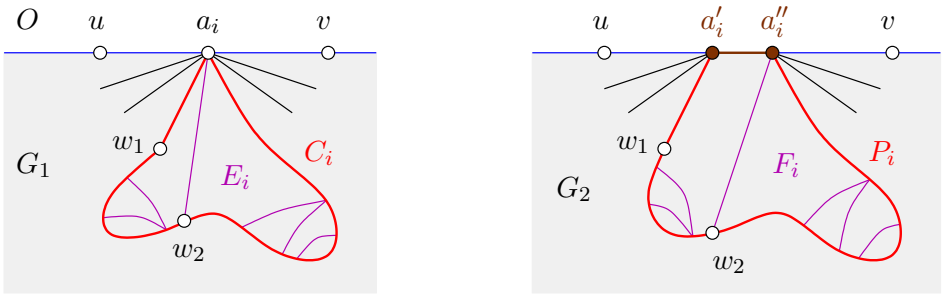


Figure 1.14: Splitting anchor a_i of C_i into the edge-anchor $a'_i - a''_i$.

Claim 1.3.7. *The graph G_2 is a Hamiltonian plane major of G_1 with $|V(O)|+2n$ vertices.*

Proof. Clearly, G_2 is a plane major of G_1, as it suffices to contract every edge-anchor of G_2 to obtained G_1.

To construct a Hamiltonian cycle in G_2, we start with the cycle outerface of G_2. Then, each edge-anchor $a'_i - a''_i$ is replaced by P_i. Inherited from Claim 1.3.6 and from the fact that C_i's partition all internal vertices of G_1, all the internal vertices of G_2 are spanned by the P_i's, which are pairwise disjoint. Thus it forms an Hamiltonian cycle for G_2.

The graph G_1 has $|V(O)|$ vertices on its outerface O and $2n-k$ internal vertices (by Claim 1.3.6). Splitting each of the k anchors adds k vertices in G_2. Therefore, $|V(G_2)| = |V(G_1)| + k = (|V(O)| + (2n - k)) + k = |V(O)| + 2n$. □

Note that the splitting anchor operation alters the outerface. So, for the proof of Lemma 1.3.3 that needs to consider polygonal embeddings whose twin sides

have to be of same length (among other things), we will need to make extra transformation on the outerface as explained later in Section 1.3.5.

Before, we observe that the two previous transformations (by Claim 1.3.6 and Claim 1.3.7) provide an alternative proof of [RST94, Theorem (1.4)] that states that every n-vertex planar graph is a minor of a planar Hamiltonian graph with at most $2n$ vertices.

Recall that a *circuit* in a graph extends the notion of cycle to subgraph composed of one single vertex or edge. A circuit is *separating* if its deletion increases the number of connected components of the graph, and it is *non-separating* otherwise. It is not difficult to see that every graph has a non-separating circuit with at least one vertex.

Proposition 1.3.8. *Every planar graph with n vertices and with a non-separating circuit of k vertices is minor of a Hamiltonian planar graph with at most $2n-k$ vertices. In particular, every triangulation with $n \geqslant 4$ vertices is minor of a Hamiltonian planar graph with at most $2n-4$ vertices.*

Proof. Let G be a plane graph with n vertices and with a non-separating circuit of k vertices, denoted by O. W.l.o.g. we can assume that O is a cycle, i.e., $k \geqslant 3$, since otherwise we can triangulate G and consider its triangle outerface as non-separating circuit. And, obviously the triangulation is a major of G and we will get a minor of $2n-3$ vertices that is less than $2n-k$ if $k < 3$.

First, we embed G in the plane such that all the vertices of $G \setminus O$ are inside O. Note that O is not necessarily the boundary of the outerface of this embedding, since there can be some chords outside of O. Denote by C the set of these chords. Let G' be the plane graph obtained from $G \setminus C$ such that O is the boundary of its outerface which is a cycle.

Since G' is a plane graph with cycle outerface O, we can construct the major G_1 as in Claim 1.3.6, and the Hamiltonian plane major G_2 from G_1 (by Claim 1.3.7). Due to the edge-anchor, the outerface of G_2 is a subdivision of O. Therefore, we can add all the chords of C in G_2 that we can embedded outside the outerface of G_2, while preserving its planarity. Thus, $G_2 \cup C$ is the desired major of G.

The graph G' has $n' = n - k$ internal vertices by construction. So, by Claim 1.3.7, G_2 (and $G_2 \cup C$ as well) has $k + 2n' = 2n - k$ vertices as claimed.

If G is a triangulation with at least four vertices, then G has a non-separating circuit of length $k+1$. Indeed, the outerface is a non-separating triangle that can be increased by one using any incident internal face. So the above property construction applies with $k = 4$. □

Actually, the proof of [RST94, Theorem (1.4)] gives an upper bound of $2n-4$ vertices (assuming $n \geqslant 4$), and relies on Whitney's Theorem[12] [Whi31]. In contrast

[12] Every planar triangulation without a separating triangle has a Hamiltonian cycle.

Proposition 1.3.8 gives an explicit and direct construction with the same upper bound. We believe that it gives an interesting construction of the major where the Hamiltonian cycle cuts the graph into two sides. Inside the cycle are the copies of edges that where initially in G, and outside of the cycle are the edges introduced by the operations of blowing up trees and splitting anchors: contracting those edges results in G.

1.3.5 Proof of Lemma 1.3.3

The goal of this subsection is to show that:

Lemma 1.3.3. *Every polygonal embedding Π of size (m, n) is a p-minor of a polygonal embedding of size $(m + 2n, 0)$.*

Let Π be a polygonal embedding of size (m, n). Up to taking a p-major of Π, we assume that each corner of Π is on the boundary of an inner triangle: if not, we can connect by an edge the two neighbors of each corner. This does not alter the size of Π. This is to prevent any further edge insertion that could increase the degree of corners which must be exactly two in any polygonal embedding. The boundary of the outerface of Π is a cycle, so the construction of G_2 as in Claim 1.3.6 applies. Let Π_1 be the polygonal embedding obtained from Claim 1.3.6, with Π in the role of G, Π_1 in the role of G_1. The embedding Π_1 admits Π as a minor while preserving the cycle outerface (border and signature does not matter for Claim 1.3.6). Thus Π is a p-minor of Π_1. Keeping the same notation as in Claim 1.3.6, Π_1 has exactly $2n - k$ internal vertices that are partitioned by the cycles $C_1, \ldots, C_k$, each C_i intersecting the outerface in one vertex, its anchor a_i.

Unfortunately, as said previously, the transformation of G_1 into G_2 as in Claim 1.3.7 does not preserve a polygonal embedding: twin edges in G_1 are not twin anymore in G_2 because of the new created edge-anchors. So it breaks the p-major sequence we are constructing. To preserve a polygonal embedding, we need to modify accordingly twin sides whenever we apply a splitting anchor so that each vertex of one side has a twin vertex on its twin side.

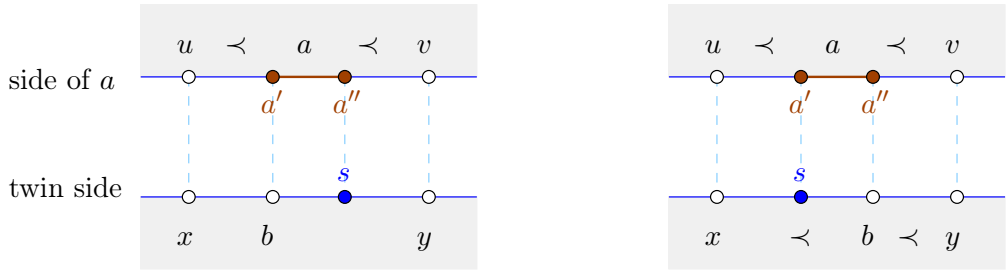
For this purpose we define a *twin splitting* operation that applies to each anchor of Π_1. Roughly speaking, we apply a splitting anchor operation while subdividing one of the two twin edges that are incident to the twin vertex of the anchor. The edge to subdivide depends on a total ordering $\prec$ defined on the vertices of the cycle outerface. It is based on their rank, starting with the very first corner of the border of Π_1. So $u \prec v$ if vertex u is visited before v when traversing the cycle outerface from the first corner of the border.

Each step t of this twin process consists in applying one twin splitting operation which results into a new polygonal embedding Π_1^t obtained from Π_1^{t-1}, starting with $\Pi_1^0 = \Pi_1$. More precisely, consider $u - a - v$ be three consecutive vertices on

a side of Π_1^{t-1} where a is an anchor and $u \prec a \prec v$. Let x, b, y be the twin vertices respectively of u, a, v. Because an anchor cannot be a corner, vertices u, v, x, y are well defined. The twin splitting operation consists in applying a splitting anchor operation on a followed by (see Fig. 1.15):

- a subdivision of the edge $b - y$, if $a \prec b$ or $y \prec b$; or
- a subdivision of the edge $x - b$, if $b \prec a$ and $b \prec y$.

In the former case, the edge $b-y$ is replaced by the path $b-s-y$ and $b-s$ becomes the twin edge of $a' - a''$, the edge-anchor of a. In the latter case, the edge $x - b$ is replaced by the path $x - s - b$ and $s - b$ becomes the twin of $a' - a''$. All the other twin relations are unchanged. See Fig. 1.15 for an illustration. It is important to observe that b may be an anchor that will be eventually split at some step $t' > t$.



case 1: $a \prec b$ or $y \prec b$ case 2: $b \prec a$ and $b \prec y$

Figure 1.15: The twin splitting operation for an anchor a: it combines a splitting anchor operation on a (in brown) and an edge subdivision (in blue) incident to its twin b. Twin relations are with dotted lines. In this representation, the focus on the twin sides creates the illusion that the inside part of the embedding (in gray) is a disconnected region, which is not true. However, in this representation choice, the inside part, that is locally planar, may appear twisted if fully represented in the non-oriented surface case.

After this twin process, denote by Π_2 the final embedding obtained from Π_1, i.e., the embedding obtained from Π_1 by applying successively all twin splitting operations.

Claim 1.3.9. *The embedding Π_2 has the following properties:*

- *Π_2 is a polygonal embedding that is p-major of Π_1;*
- *Π_2 has size $(m + k, 2n - k)$;*
- *$P_1, \ldots, P_k$ form a partition of the internal vertices of Π_2;*
- *During the twin process, no edge-anchor is subdivided and no two edge-anchors are twins.*

Proof. Each twin splitting operation inserts edges and split anchors on the outer-face of Π_1. So, by contracting these new edges, we get back to Π_1 which shows that Π_2 is a major of Π_1. Moreover, thanks to the twin splitting, eventually every vertex (and edge) of a side in Π_2 of have a well defined twin. So, Π_2 is a polygonal embedding that is p-major of Π_1.

At each twin splitting operation, the number of vertices of a side increases by at most one while the number of internal vertices does not change. Since Π_1 contains k anchors, after applying the twin process, each side of Π_2 contains at most $m+k$ vertices (corners excluded) and $2n-k$ internal vertices. So, the size of Π_2 is $(m+k, 2n-k)$.

The internal vertices in Π_1 and in Π_2 are the same. And after each splitting anchor operation the paths P_i's span the same internal vertices than C_i's. So by Claim 1.3.6, P_i's also form a partition of the internal vertices of Π_2.

It remains to show that during the twin splitting process no two edge-anchors can be twin or subdivided. So, consider any step t' that performs a twin splitting operation at some anchor b of $\Pi_1^{t'-1}$. We need to check that indeed the operation does not create twin edge-anchors or subdivide an edge-anchor. If the twin of b does not belong to an edge-anchor, then it is fine since the twin splitting of b will not result in a subdivision or a twin edge of an edge-anchor. So we only need to check the situation where the twin of b belongs to an edge-anchor. This occurs only if b and its twin a were both anchors in Π_1^{t-1} for some $t < t'$. W.l.o.g. assume that t is the step where a is split into $a' - a''$, and let s be the neighbor of b in Π_1^t resulting of the twin splitting of a in Π_1^{t-1}.

We remark that two edge-anchors cannot be incident, and that on a side with path $x-b-y$ that has been subdivided into $x-b-s-y$, then neither $x-b$ nor $b-s$ can be subdivided anymore (because the twin of b cannot be an anchor anymore). The path $x-b-s$ can only be replaced by an edge-anchor $x-b'-b''-s$ if b is an anchor.

We use the notation as above and as in Fig. 1.15. Let $u-a-v$ be the vertices on the twin side of b in Π_1^{t-1}, where a is an anchor and twin of b. W.l.o.g. assume $u \prec a \prec v$, and let x, y be the twins of u, v respectively. In Π_1^t, we have $u \prec a' \prec a'' \prec v$, and there is some vertex s adjacent to b coming from an edge subdivision incident to b (of either $b-y$ or $x-b$). Let us apply step t' on b which produces the edge-anchor $b'-b''$ in $\Pi_1^{t'}$.

To simplify notation we will reuse names u, v, x, y to denote the neighbors of a', a'', b, s in Π_1^t as well as in $\Pi_1^{t'-1}$.

Assume that $b \prec a$ and $b \prec y$, i.e., the case 2 holds for a. In that case, we have the path $x-s-b-y$ in $\Pi_1^{t'-1}$, and also $u \prec a' \prec a''$. By exchanging the roles of $s-b-y$ with $a'-a''-v$ the case 1 holds for b. This is because the condition $a \prec b$ rewrites in $b \prec a$ which is true (a'' plays the role of a). Therefore, the twin

splitting of b subdivides the edge $a'' - v$ which cannot be an edge-anchor by the previous remark. The edge-anchor $b' - b''$ is twin with a non edge-anchor as well. Thus, step t' does not twin or subdivide any edge-anchor in this case.

Assume that $a \prec b$ or $y \prec b$, i.e., the case 1 holds for a. In that case, we have the path $x - b - s - y$ in $\Pi_1^{t'-1}$, and also that $u \prec a' \prec a''$.

If $y \prec b$, then we have $y \prec s \prec b \prec x$. By exchanging the roles of $s - b - x$ with $u - a' - a''$ the case 1 holds for b. This is because the condition $u \prec a'$ rewrites in $y \prec b$ which is true.

If $b \prec y$, then we have $x \prec b \prec s$. By exchanging $x - b - s$ with $u - a' - a''$ the case 1 holds again for b. This is because the condition $a' \prec a''$ rewrites in $b \prec y$ which is true (s plays the role of y).

Thus, in both cases, the twin splitting of b subdivides the edge $u - a'$ which cannot be an edge-anchor by the previous remark. The edge-anchor $b' - b''$ is twin with a non edge-anchor as well. And therefore, in all the cases, step t' does not twin or subdivide any edge-anchor. □

The last property of Claim 1.3.9 (no edge-anchor is subdivided) ensures that the k edge-anchors in Π_1 are still existing in Π_2 after applying the twin process on Π_1.

In order to transform Π_2 into an outerplanar embedding, we will apply at each of its edge-anchor a *swapping* operation. So, each step t of this process consists in applying one such operation which results into a new polygonal embedding Π_2^t obtained from Π_2^{t-1}, starting with $\Pi_2^0 = \Pi_2$.

The swapping operation is defined as follows. (We reuse the same notation as in Fig. 1.14 describing the splitting anchor operation which has been used to make Π_2.) Let $a_i' - a_i''$ be an edge-anchor of Π_2^{t-1}, P_i be the path connecting a_i' to a_i'', and F_i be the edges not in P_i and embedded inside the cycle $P_i \cup \{a_i' - a_i''\}$. Finally, let b', b'' be the twin vertices of respectively a_i', a_i''.

To obtain Π_2^t, we first delete the edge $a_i' - a_i''$ and all edges of F_i. By this way, the side containing $a_i' - a_i''$ is extended by P_i. Then, on the outerface of Π_2^{t-1}, we connect b' to b'' by a copy P_i' of P_i and including a copy F_i' of all the edges of F_i such a way that the direction of P_i is preserved. It means that if u is traversed before v when going from a_i' to a_i'' on P_i, then u' is traversed before v' when going from b' in b'' in P_i' where u', v' are the copies in P_i' of u, v respectively. Furthermore, the twin of any vertex u in P_i is its copy of u' in P_i'. By this way, the side containing $b' - b''$ is extended by P_i'. See Fig. 1.16 for an illustration.

The last property of Claim 1.3.9 (no edge-anchors are twins) ensures that $b' - b''$ was not an edge-anchor in Π_2^{t-1}. So the swapping operation on the edge-anchor $a_i' - a_i''$ leaves unchanged all the other edge-anchors.

We denote by Π_3 the final embedding obtained from Π_2 by applying successively a swapping operation on each of its edge-anchors. Let us denote by $A \cong B$

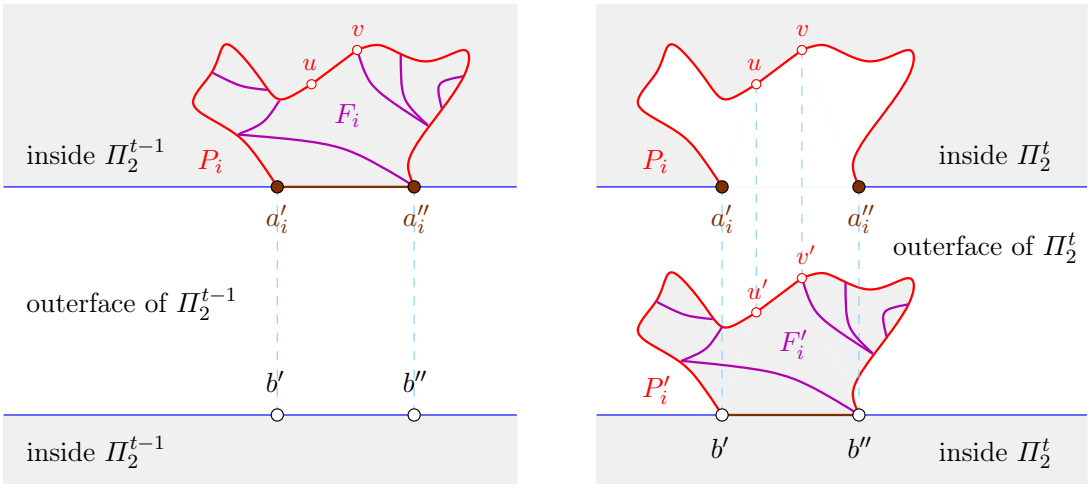


Figure 1.16: A swapping operation for the edge-anchor $a'_i - a''_i$ in Π_2^{t-1} (on the left) leading to a new embedding Π_2^t (on the right). Vertex u is traversed before v when going from a'_i to a''_i on P_i, and so for their twin vertices u', v' when going from b' to b'' on P'_i. The vertices of P_i that were internal in Π_2^{t-1} belongs to the outerface of Π_2^t.

if A, B are isomorphic graphs or homemorphic embeddings.

Claim 1.3.10. *The embedding Π_3 is a polygonal embedding p-major of Π_2 with* $\text{sewing}(\Pi_3) \cong \text{sewing}(\Pi_2)$ *and size* $(m + 2n, 0)$.

Proof. Up to a permutation of the indices, we can assume that the swapping operation for $a'_i - a''_i$ is performed at step i from Π_2^{i-1}. It is easy to see that each operation does not create edge crossings, that the outerface remains a cycle, and that each vertex (edge) of each side has a well-defined twin vertex (edge) in its twin side. Therefore, Π_2^i is a polygonal embedding, and also Π_3, that is Π_2^k, by transitivity.

To prove it has the expected size, let $n_i = |V(P_i) \setminus \{a'_i, a''_i\}|$ be the number of internal vertices of P_i in Π_2^{i-1}. Note that this number is the same as the number of vertices of P_i in $\Pi_2^0 = \Pi_2$ because P_i is not altered by any of the previous steps $j < i$. Therefore, from Claim 1.3.9, $\sum_{j=1}^{k} n_j = 2n - k$.

The swapping operation for $a'_i - a''_i$ increases by n_i the number of vertices of its side and its twin side, and, at the same time, decreases by n_i the number of internal vertices of Π_2^{i-1}. By Claim 1.3.9, Π_2 has size $(m + k, 2n - k)$. So after step i, the size of Π_2^i is $(m + k + \sum_{j=1}^{i} n_j, n - 2k - \sum_{j=1}^{i} n_j)$. It follows that Π_3, which is Π_2^k, has size $(m + 2n, 0)$ since $\sum_{j=1}^{k} n_j = 2n - k$.

It is clear that signature and border, i.e., the sequence of corners, are not altered by any swapping operation. So, it remains to show that $\text{sewing}(\Pi_3) \cong \text{sewing}(\Pi_2)$

(and by this way it will also show that Π_3 is a p-major of Π_2). By transitivity, it suffices to show that sewing$(\Pi_2^i) \cong$ sewing(Π_2^{i-1}).

Let R_i be the plane graph composed of $P_i \cup \{a'_i - a''_i\} \cup F_i$ in Π_2^{i-1} such that the boundary of its outerface is the cycle $P_i \cup \{a'_i - a''_i\}$. Similarly, let R'_i be the plane graph composed of $P'_i \cup \{b' - b''\} \cup F'_i$ in Π_2^i with cycle outerface $P'_i \cup \{b' - b''\}$. Note that $V(R_i) = V(P_i)$ and $V(R'_i) = V(P'_i)$. For convenience, denote by $G^j =$ sewing(Π_2^j). We want to show that $G^{i-1} \cong G^i$.

Observe that in G^{i-1} and in G^i, whenever the twin sides are merged, $a'_i = b'$, $a''_i = b''$ and thus the edge $a'_i - a''_i = b' - b''$ exists in both graphs (whereas $a'_i - a''_i$ exists only in Π_2^{i-1}). Also, the way the vertices of P_i are twins with the vertices P'_i (by preserving the order when going from a'_i to a''_i in P_i) ensures that $P'_i = P_i$ and $F'_i = F_i$ in G^i. It follows that $R_i \cong R'_i$, and since the swapping operation alters only R_i, $V(G^{i-1}) = V(G^i)$. Thus, if there is a difference between G^{i-1} and G^i, it must be an edge.

From previous equalities, if an edge $x - y$ lies in R_i then the edge exists in R'_i and thus in G^i. Conversely, if $x - y$ lies in R'_i, the edge exists in R_i and thus in G^{i-1}. If $x - y$ lies outside R_i (x, y may be both in $V(R_i)$ with an embedding not inside R_i), then $x - y$ lies also in Π_2^i since the swapping operation alters only R_i (and its inside). So $x - y \in E(G^{i-1})$ implies $x - y \in E(G^i)$ (even if $x - y$ belongs a some side of Π_2^{i-1}). Conversely, if $x - y$ lies outside R'_i in Π_2^i, then $x - y$ lies also in Π_2^{i-1} (and outside of R_i). It follows that the edge sets of G^{i-1} and G^i are the same, proving that $G^{i-1} \cong G^i$. □

By transitivity of the p-minor relation, every polygonal embedding Π of size (m, n) is a p-minor of Π_2 (by Claim 1.3.9) that is a p-minor of Π_3 of size $(m+2n, 0)$ (by Claim 1.3.10), which completes the proof of Lemma 1.3.3.

1.3.6 Proof of Lemma 1.3.4

The goal of this subsection is to show that:

Lemma 1.3.4. *Every polygonal embedding Π of size $(m, 0)$ is a p-minor of $\mathbb{U}_{\sigma(\Pi),m}$.*

The proof is an adaptation of the proof of [RST94, Theorem (1.3)] to the case of outerplanar polygonal embeddings. The original proof uses grid major whereas we need to use $\mathbb{U}_{\sigma,m}$.

We first define the polygonal embedding $\mathbb{U}_{\sigma,m}$, which is defined for all signatures σ and $m \in \mathbb{N}$. To construct $\mathbb{U}_{\sigma,m}$, we start from a square half-grid of dimension $|\sigma|m$, that is with $|\sigma|m$ rows and $|\sigma|m$ columns. The vertices are column-row pairs of integers (i, j) where $0 \leqslant j \leqslant i < |\sigma|m$, and (i, j) and (i', j') are adjacent if and only if $|i - i'| + |j - j'| = 1$. Then, the vertices of the diagonal are connected by

a cycle, (i, i) being connected to $(i-1, i-1)$ modulo $|\sigma|m$. This cycle forms the cycle outerface of $\mathbb{U}_{\sigma,m}$. Finally, every m edges along this cycle is subdivised by one vertex, a corner. More precisely, for each $i \in \{0, \ldots, |\sigma| - 1\}$, we add a vertex c_i between (im, im) and $(im - 1, im - 1)$ modulo $|\sigma|m$. The border of $\mathbb{U}_{\sigma,m}$ is the sequence $(c_0, \ldots, c_{|\sigma|-1})$, and the ith side of $\mathbb{U}_{\sigma,m}$, associated with the ith symbol of σ, is the path between c_i and $c_{i+1 \bmod |\sigma|}$ on the cycle outerface. See Fig. 1.17 and Fig. 1.12 for illustrations.

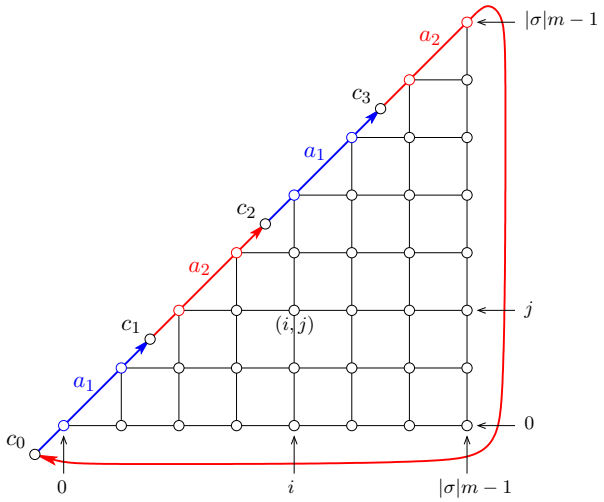


Figure 1.17: The polygonal embedding $\mathbb{U}_{\sigma,m}$ for $m = 2$ and $\sigma = a_1a_2a_1a_2$, a non-canonical signature for a non-orientable surface of genus 2. Its size is $(2, 28)$. The cycle outerface is $c_0 - (0, 0) - (1, 1) - c_1 - (2, 2) - (3, 3) - c_2 - (4, 4) - (5, 5) - c_3 - (6, 6) - (7, 7) - c_0$.

Clearly, the number of internal vertices of $\mathbb{U}_{\sigma,m}$ is $n = 1 + 2 + \cdots + |\sigma|m - 1$. By merging the twin sides of $\mathbb{U}_{\sigma,m}$ to obtain sewing$(\mathbb{U}_{\sigma,m})$, we destroy at least half of the vertices of the cycle outerface of $\mathbb{U}_{\sigma,m}$. It follows that sewing$(\mathbb{U}_{\sigma,m})$ has no more than $n + |\sigma|(m + 1)/2$ vertices. However, depending on the signature, this number can be lower as the set of corners may collaps even more when merging the sides. E.g., if the signature is minimal, i.e., comes from a system of loops, sewing$(\mathbb{U}_{\sigma,m})$ has only $n + |\sigma|m/2 + 1$ vertices since such a signature tell us that all the corners are twins, resulting into the single vertex of the system of loops.

To summarize, we have:

Claim 1.3.11. *The embedding $\mathbb{U}_{\sigma,m}$ is a polygonal embedding of signature σ with $\frac{1}{2}|\sigma|m \cdot (|\sigma|m - 1))$ internal vertices, each of the $|\sigma|$ sides having m non-corner vertices. In particular,* sewing$(\mathbb{U}_{\sigma,m})$ *has no more than $\frac{1}{2}(|\sigma|^2m^2 + |\sigma|)$ vertices, and at most $\frac{1}{2}|\sigma|^2m^2 + 1$ vertices if σ is minimal.*

Let Π be a polygonal embedding of size $(m, 0)$. Up to taking a p-major of Π, we can suppose that each side of Π is composed of exactly m non-corner vertices. If needed, one can subdivide an edge and its twin of any too small side, resulting in a p-major of Π. So, w.l.o.g., the cycle outerface of Π is composed of $|\sigma(\Pi)|$ sides of m non-corner vertices each.

Since Π is outerplane, its cycle outerface contains all its vertices. The embeddings Π and $\mathbb{U}_{\sigma(\Pi),m}$ have same signature and also m non-corner vertices per side. So both embeddings have isomorphic border and a cycle outerface of $|\sigma(\Pi)|(m+1)$ vertices.

We denote by $(u_0, \dots, u_{|\sigma(\Pi)|-1})$ the border of Π, and by $v_0, \dots, v_{|\sigma(\Pi)|m-1}$ its non-corners vertices ordered clockwise around its cycle outerface so that this cycle is

$$u_0 - v_0 - v_1 - \cdots - v_{m-1} - u_1 - v_m - \cdots - v_{im-1} - u_i - v_{im} - \cdots$$

Up to taking a p-major of Π, we will assume that all inner faces of Π are triangulated such that corners have degree two. For each non-corner v_i, we define the two indices $a(i), b(i)$ such that $\{v_k : k \in [a(i), b(i)]\}$ is the minimal subset of vertices containing v_i and all its non-corner neighbors. Note that $a(i) = i$ or $b(i) = i$ is possible. For technical reasons, we increase $b(0)$ to $b(0) = |\sigma(\Pi)|m$.

In order to show that Π is a minor of $\mathbb{U}_{\sigma(\Pi),m}$, we construct a *witness* of Π in $\mathbb{U}_{\sigma(\Pi),m}$ with the property that it preserves its outerface and its border. A witness for a minor Π in $\mathbb{U}_{\sigma(\Pi),m}$ is a collection $\{W(u)\}_{u \in V(\Pi)}$ of nonempty pairwise disjoint subsets of vertices of $\mathbb{U}_{\sigma(\Pi),m}$ such that each set $W(u)$ induces a connected subgraph of $\mathbb{U}_{\sigma(\Pi),m}$, and for each edge $u-v$ of Π, $W(u) \cup W(v)$ induces a connected component in $\mathbb{U}_{\sigma(\Pi),m}$.

The witness of Π in $\mathbb{U}_{\sigma(\Pi),m}$ is defined by:

- $W(u_i) = \{c_i\}$ for each corner u_i; and
- $W(v_i) = \{(i, j)\} \cup C(v_i) \cup R(v_i)$ for each non-corner v_i,
  where $C(v_i) = \{(i, k) : k \in (a(i), i]\}$ and $R(v_i) = \{(k, i) : k \in [i, b(i))\}$.

Clearly, for corner vertices, $W(u_i) = \{c_i\}$ fulfills all desired properties: nonemptyness, disjointness and connectedness. For non-corner vertices, $W(v_i)$ induces a connected components of $\mathbb{U}_{\sigma(\Pi),m}$ since it is the union of a column subpath $C(v_i)$ and of a row subpath $R(v_i)$ of $\mathbb{U}_{\sigma(\Pi),m}$ intersecting in (i, i) (in the case where $C(v_i)$ and $R(v_i)$ are not empty).

Let us check that $W(v_i)$ and $W(v_j)$ are pairwise disjoint (witness for corner and non-corner are clearly disjoint). Assume $j < i$. A non-empty intersection is only possible between $C(v_i)$ and $R(v_j)$. And, this can only occur at (i, j). Note that the case $j = 0$ provides disjoint witnesses, since $(i, 0) \notin C(v_i)$. If $(i, j) \in C(v_i)$, then $j \in (a(i), i]$, and thus v_i has a neighbor $v_{a(i)}$ with $a(i) < j$. If $(i, j) \in R(v_j)$, then $i \in [j, b(j))$, and thus v_j has a neighbor $v_{b(j)}$ with $i < b(j)$. It follows that

$a(i) < j < i < a(j)$ and that $v_{a(i)} - v_i$ and $v_j - v_{b(j)}$ is a pair of crossing edges, which is impossible in the outerplanar embedding Π.

It remains to check that, for each $u - v \in E(\Pi)$, $W(u) \cup W(v)$ induces a connected graphs in $\mathbb{U}_{\sigma(\Pi),m}$. We can restrict our attention to non-corner neighbors, since for a corner u_i, $c_i \in W(u_i)$ have exactly two neighbors that are $(im-1, im-1) \in W(v_{im-1})$ and $(im, im) \in W(v_{im})$ (indices modulo $|\sigma(\Pi)|m$).

So consider an edge $v_i - v_j$ of Π with $j < i$. Let $W(v_i, v_j)$ be the set of vertices of the path connecting $(i,i) \in W(v_i)$ to $(j,j) \in W(v_j)$ inside the grid part of $\mathbb{U}_{\sigma(\Pi),m}$, and defined by

$$W(v_i, v_j) \;=\; \{(i,i), (i,i+1), \ldots,\; (i,j-1)\;, (i,j), (i-1,j), \ldots, (j,j)\}\;.$$

To show that $W(v_i) \cup W(v_j)$ induces a connected component in $\mathbb{U}_{\sigma(\Pi),m}$, it suffices to show that $W(v_i, v_j) \subseteq C(v_i) \cup R(v_j)$ since $W(v_i, v_j)$ induces a path in $\mathbb{U}_{\sigma(\Pi),m}$, and $C(v_i) \subset W(v_i)$ and $R(v_j) \subset W(v_j)$. If $(i', j') \notin W(v_i, v_j)$ for some $i' \in [j, i)$ or $j' \in [i, j)$, then the set $\{v_k : k \in [a(i), b(i)]\}$ cannot cover the non-corner neighbors of v_i or the set $\{v_k : k \in [a(j), b(j)]\}$ cannot cover the non-corner neighbors of v_j. It remains to check that $(i,j) \in C(v_i) \cup R(v_j)$, even if $(i,j) \notin C(v_i)$ or $(i,j) \notin R(v_j)$ is possible. If $j = 0$, then we are done because $b(0) = |\sigma(\Pi)|m$ and thus $(i,0) \in R(0)$. We are also done if $|i-j| = 1$ since $(i,i) \in W(v_i)$ and $(j,j) \in W(v_j)$ are neighbors in $\mathbb{U}_{\sigma(\Pi),m}$. Since the inner faces of Π are triangulated, the edge $v_i - v_j$ (that lies inside Π since v_i and v_j are not consecutive in the cycle outerface) share a triangle $v_i - v_j - v_k$ such that $k \notin [i, j]$, the case $j = 0$ and $i = |\sigma(\Pi)|m - 1$ being excluded. If $k > i$, then $a(j) > i$ and $(i,j) \in C(v_j)$. If $k < j$, then $b(i) > j$ and $(i,j) \in R(v_i)$.

We have therefore proved that $W(v_i, v_j) \subseteq W(v_i) \cup W(v_j)$ and that $W(v_i) \cup W(v_j)$ induces a connected component in $\mathbb{U}_{\sigma(\Pi),m}$. It follows that Π is a minor of $\mathbb{U}_{\sigma(\Pi),m}$ with same outerface, border and signature. By Property 1.3.5, Π is a p-minor of $\mathbb{U}_{\sigma(\Pi),m}$ that completes the proof of Lemma 1.3.4.

Chapter 2

Finding Induced Trees in Graphs

Contents

2.1 Introduction

A well-studied algorithmic problem on graph is the Maximum Independent Set (MIS). A set of vertices is an *independent set* if all the vertices in the set are pairwise non-adjacent. The MIS problem consists in finding an independent set of maximum size in an input graph, as in Fig. 2.1.

This problem is one of the fundamental graph problems that are NP-complete [Har82]. Moreover, MIS is one of the first problem that has been shown to be NP-hard even to approximate [FGL+96].

However, this problem can be polynomially solved when restricted to certain classes of graphs. A simple class where MIS, as well as most hard problems, is polynomial is the class of trees.

Figure 2.1: Example of a maximal (left) and a maximum (right) independent set on the same graph.

Here are the ideas to design a simple linear algorithm that computes a maximum independent set in a tree, using dynamic programming. Let T be a tree and choose a vertex r to be its root: for each vertex u, we denote by T_u the subtree of T rooted in u. Let S_u be a maximum independent set in T_u containing u, and $\bar{S_u}$ a maximum independent set in T_u not containing u. Let S'_u be the largest set between S_u and $\bar{S_u}$. Note that, with those definitions, S'_r is a maximum independent of T.

Observe that if u is a leaf, $S_u = \{u\}$ and $\bar{S_u} = \varnothing$. If u is not a leaf, let $v_1, \ldots, v_k$ for some k be the children of u. Then $S_u = \{u\} \cup \bar{S_{v_1}} \cup \cdots \cup \bar{S_{v_k}}$, and $\bar{S_u} = S'_{v_1} \cup \cdots \cup S'_{v_k}$. With those relations it is easy to deduce a linear algorithm that starts from the leaves and computed recursively $\bar{S_u}, S'_u$ for each u. This algorithm ends with the computation of S'_r, which is a maximum independent set of the whole tree T.

From this simple algorithm, we can derive a linear algorithm for graphs with treewidth bounded by a constant as their structure is similar to that of a tree.

A natural question is to identify the other graph classes for which MIS is easy. This question has been largely studied on hereditary [1] graph classes that exclude a graph H (or a set of graphs $\mathcal{H}$) as induced subgraph.

We take this in two main directions in this chapter: excluding certain cycles as induced subgraph (see 2.1.1 and later 2.3), and excluding a small path as induced subgraph (see 2.1.2 and 2.2).

2.1.1 Excluding cycle structures

Observe that the class of trees, and more generally the class of forests, is the class of graph excluding all cycles as subgraph. Since MIS is easy in those classes, a natural direction is to study graphs that exclude some cycles as induced subgraphs.

Recently, a polynomial-time algorithm for Maximum Independent Set has been found for graphs that exclude all cycles of length at least 5 as subgraphs [ACP+18]. In fact, excluding only the odd cycles is sufficient to make MIS solvable in polynomial time. This result has been extended to the graphs that exclude a constant number of vertex-disjoint odd cycles [FJWY21].

[1] closed under taking induced subgraphs

The Erdős-Pósa theorem [EP65] implies a similar result for the graphs that exclude a constant number of disjoint-cycles. More precisely, this theorem states that if a graph G does not have k vertex-disjoint cycles, then G admits a *feedback vertex set* (i.e., a subset of vertices whose removal yields a forest) of size $O(k \log k)$, hence treewidth $O(k \log k)$. As we saw previously, MIS is polynomial for graphs of bounded treewidth.

The intuition behind the Erdős-Pósa theorem is that if there are not many disjoint cycles, all the cycles are packed together and share vertices, so it does not require many vertices to hit every cycle.

In what extent does this hold if we loosen the condition on the cycles? More precisely, we say that two cycles are *independent* if they are vertex-disjoint and there are no edges between them. A graph is said to be O_k*-free* if it does not contain k independent cycles. In other words, the O_k-free graphs are the graphs who do not have a disjoint union of k cycles as an induced subgraph.

Observe that the complete bipartite graphs are O_2-free but have treewidth linear in their order. We thus focus on O_k-free graphs that are *sparse* (in this context, this means no $K_{t,t}$ subgraph for some t).

Our result: Section 2.3 describes the results obtained with Marthe Bonamy, Edouard Bonnet, Hugues Déprés, Louis Esperet, Colin Geniet, Stéphan Thomassé and Alexandra Wesolek [BBD+23].

Our main theorem is that sparse O_k-free graphs have treewidth at most logarithmic in the number of vertices. This is sharp as we exhibit family of O_2-free graphs without $K_{3,3}$-subgraph and whose treewidth is logarithmic. In fact for our main theorem, we prove something even stronger, we show that the size of the feedback vertex set is logarithmic.

Since the treewidth is not constant, Courcelle's theorem is not sufficient to deduce that the Maximum Independent Set is polynomial on those graphs. However, Pilipczuk gave a refinement of Courcelle's theorem, and showed that any problem expressible in Existential Counting Modal Logic (ECML) admits a single-exponential fixed-parameter algorithm in treewidth [Pil11]. This implies that a problem expressible in ECML can be solved in polynomial time on any class with logarithmic treewidth. In particular, Maximum Independent Set is expressible in ECML thus is polynomial for the sparse O_k-free graphs, as well as several fundamental problems such as Minimum Vertex Cover, Minimum Dominating Set, Minimum Coloring.

For the O_k-free graphs (without the sparse assumption), we come short to prove that MIS is polynomial but not by much. We obtain a quasi-polynomial time algorithm for MIS in general O_k-free graphs, which implies that this problem is not NP-complete (assuming there is no complexity-theoretic collapse).

A consequence to our main theorem is that testing if a graph is O_k-free with no $K_{t,t}$-subgraph is polynomial. A recent work of Nguyen, Scott and Seymour [NSS22] extends this results to O_k-freeness.

2.1.2 Finding a long induced path

Another direction in the study of MIS for H-induced-free graph is the study of graphs excluding a small path. Indeed, MIS is polynomial for P_k-induced-free graphs when $k < 7$ and quasi-polynomial for every fixed k [GL20]. This leads us to study the length of a longest induced path in a graph.

Every induced path is a path, but does the existence of a long path in a graph imply that of a long induced path? More precisely, for a graph family $\mathcal{G}$, we want to know if there is an increasing function f such that every graph of $\mathcal{G}$ with an n-path has an induced path on $f(n)$ vertices.

In general, this is false as shown by cliques and bicliques. However, Galvin, Rival, and Sands showed in 1982 that there is such a function as soon as the graphs exclude a biclique as a subgraph [GRS82, Theorem 4]. Unfortunately, this function increase very slowly as their proof relies on the infinite Ramsey's theorem for 4-tuples, hence can be used mostly as an existential result, rather than one providing accurate and tight bounds.

For a positive integer k, a graph G is said to be *k-degenerate* if every subgraph of G (including G itself) has a vertex of degree at most k. Observe that G excludes $K_{k,k}$ as subgraph. A better lower-bound was given by Nešetřil and Ossona de Mendez [NDM12, Lemma 6.4] in the case of k-degenerate graphs: there, a path of order n implies the existence of an induced path of order at least $\frac{\log\log n}{\log(k+1)}$.

Arocha and Valencia exhibited an infinite family of outerplanar graphs (2-degenerate) where the order of a longest induced path is at most logarithmic in the order of the graph [AV00]. For several subclasses of $\mathcal{G}$ of k-degenerate graph (bounded genus, some bounded treewidth subclasses), Esperet, Lemoine, and Maffray found a $(\log n)^{\Omega(1)}$ lower bound for the order of a longest induced graph in a graph of $\mathcal{G}$ with an n-path [ELM17]. Those results are recapitulated later in Table 2.1. They conjectured that a $(\log n)^{\Omega(1)}$ lower bound should hold for the whole family of k-degenerate graphs. This conjecture is widely open, even for $k = 2$.

ADDENDUM: During the writing of this thesis, Defrain and Raymond [DR23] disproved the conjecture for $k = 2$, by finding a family of 2-degenerate graphs such that for each n if a graph has an n-path, then all its induced paths have order $O((\log\log n)^2)$.

Our result: The Section 2.2 relates the advances made with Jean-Florent Raymond [HR23] on this Conjecture.

Our first result is that every graph of pathwidth less than k that has a path of order n also has an induced path of order at least $\frac{1}{3}n^{1/k}$.

This result is then used to find $(\log n)^{\Omega(1)}$ lower bounds for the graphs of treewidth less than k and then for the class of topological-minor-free graphs. The latter class generalizes a lot of known classes, such as minor-closed graphs, graphs of bounded genus, graphs of bounded treewidth, but also immersion-closed graph class (like graphs of bounded cutwidth, carving width, or tree-cut width), and graphs of bounded degree.

2.2 Longest induced path

2.2.1 Overview

This section gives results on the following conjecture:

Conjecture 2.2.1 ([ELM17, Conjecture 1.1]). *For every integer k there is a constant d such that every k-degenerate graph that has a path of order n also has an induced path of order at least $(\log n)^d$.*

This conjecture is widely open, so we focus on subclasses of the class of k-degenerate graphs.

Let $\mathcal{G}$ be a hereditary class of k-degenerate graphs. There is a function $f_{\mathcal{G}}$ such that every graph G with an n-path admits an induced path of order at least $f_{\mathcal{G}}(n)$, and $f_{\mathcal{G}}(n) \geqslant \frac{\log\log n}{\log(k+1)}$[NDM12]. The problem here is finding the most accurate lower bound and upper bound for $f_{\mathcal{G}}$. For the lower bond, this can be formalized as finding the biggest function f such that this statement holds:

($\star$) *For every $n \in \mathbb{N}$, if a graph in $\mathcal{G}$ has a path of order n, then it has an induced path of order at least $f(n)$.*

The upper bond is obtained by exhibiting the smallest function f and an infinite family of graphs in $\mathcal{G}$, such that for each $n \in \mathbb{N}$, there is a graph of this family with an n-path and with no induced path of order $f(n)$.

Class of graphs $\mathcal{G}$	Lower bound	Upper bound
Outerplanar	2-connected: $\Omega(\log n)$ [ELM17]	$O(\log n)$ [AV00]
bounded genus	$\Omega((\log n)^{\frac{1}{2}})$ [ELM17]	$\frac{3\log n}{\log\log n}$ [ELM17]
graphs with $pw \leqslant k$	intervals graphs: $\Omega((\log n)^{\frac{1}{k^2}})$ [ELM17] $\to \frac{1}{3}n^{\frac{1}{k+1}}$ (Theorem 2.2.2)	$\to n^{\frac{2}{k}} + 1$ (Theorem 2.2.11)
graphs with $tw \leqslant k$	k-tree: $\frac{1}{k\log k}\log n$ [ELM17] $k = 2$: $\left(\frac{1}{2} - o(1)\right)\log n$ [ELM17] $\to \frac{1}{4}(\log n)^{\frac{1}{k}}$ (Theorem 2.2.3)	$(k+2)(\log n)^{\frac{2}{k}}$ [ELM17]
Topological-minor-free	$\to (\log n)^{\frac{1}{O(1)}}$ (Theorem 2.2.4)	$(\log n)^{\frac{1}{\Omega(1)}}$ [ELM17]
k-degenerate	$\frac{\log\log n}{\log(k+1)}$[NDM12]	$(k+2)(\log n)^{\frac{2}{k}}$ [ELM17]

Table 2.1: Known upper and lower bounds on the order of a longest induced path in a graph of $\mathcal{G}$ that has a path of length n.

The Table 2.1 sum up the known upper and lower bound of $f_{\mathcal{G}}$ for several classes $\mathcal{G}$ that interest us. The blue arrowed bounds are the new ones presented in this section. Notice that since an upper bound for a class is obtained by exhibiting a family of graphs from this class, then the upper bound holds for classes generalizing

this class. In particular the upper bound for the graphs of bounded treewidth holds for the class of topological-minor-free graphs and of degenerates graphs, for some k.

In their paper, Esperet, Lemoine, and Maffray gave a $\Omega((\log n)^{\frac{1}{(k)^2}})$ lower bound for interval graphs of pathwidth a most k [ELM17, Theorem 4.1] Our first result is a generalization of this result to graphs of bounded pathwidth, with an exponential improvement of the bound.

Theorem 2.2.2. *For every $k \in \mathbb{N}$, if G is a graph of pathwidth less than k that has a path of order n, then G has an induced path of order at least $\frac{1}{3}n^{1/k}$.*

Its proof consists in a simple win/win strategy. We identify an induced path whose removal decreases the pathwidth of G. Then either this path is at least as long as the bound promised by the statement and we are done, or its removal decreases the pathwidth without decreasing the number of vertices much, and we conclude by applying the induction hypothesis. Theorem 2.2.2 is complemented by an upper-bound (that even holds for interval graphs) of $n^{2/k}+1$ (Theorem 2.2.11), showing that the exponential dependency in $1/k$ in our lower-bound above is unavoidable.

We then show that in a graph of small treewidth that has a large path there is always (as a contraction) a graph of small pathwidth that has a long path. This statement is used to obtain the following polylogarithmic bound for graphs of bounded treewidth.

Theorem 2.2.3. *For every $k \in \mathbb{N}$, if G is a graph of treewidth less than k that has a path of order n, then G has an induced path of order at least $\frac{1}{4}(\log n)^{1/k}$.*

Esperet et al. [ELM17] constructed chordal graphs of clique number k (thus treewidth $k-1$) that have a path of order n and where no induced path has order more than $(k+1)(\log n)^{\frac{2}{k-1}}$ (see upper bound for graphs of bounded treewidth in Table 2.1). Therefore neither the logarithmic dependency in n nor the exponential dependency in $1/k$ could be improved in our lower bound above.

The ideas developed in the proofs of Theorem 2.2.2 and Theorem 2.2.3 allowed us to generalize their statement to this much more general setting. A graph class is *non-trivial* if it is not the class of all graphs.

Theorem 2.2.4. *For every non-trivial graph class $\mathcal{G}$ that is closed under taking topological minors there is a constant $d \in (0,1)$ such that if a graph $G \in \mathcal{G}$ has a path of order n, then G has an induced path of order at least $(\log n)^d$.*

The proof deals separately with the different parts that compose graphs excluding a topological minor (as given by the structure theorem of Grohe and Marx [GM15]) and then shows how they can be combined together. We actually prove

a stronger statement than Theorem 2.2.4, namely that the same outcome holds for all graphs that admit a tree decomposition where every torso is either almost embeddable in a surface of bounded genus or has almost bounded degree (see Theorem 2.2.28 for the formal statement).

In subsection 2.2.2 we give the necessary definitions, and preliminary results. We prove Theorem 2.2.2 in Section 2.2.3 and Theorem 2.2.3 in Section 2.2.4. Finally, Section 2.2.5 is dedicated to the proof Theorem 2.2.4.

Before giving the details of the proof, let us present some perspectives for those results.

Open problems. A first direction for future work is to investigate how widely the results proved in this paper could be generalized. What are the most general graph classes where the conjecture holds?

Note that the bound we obtained in Theorem 2.2.2 is polynomial. An interesting task could be to characterize hereditary classes where such a property holds. Also, as mentioned in the introduction, Esperet et al. fond an $\Omega(\log n)$ bound in k-trees, while their upper-bound for graphs of treewidth at most k (see Table 2.1) shows that such a bound where k does not appear in the exponent of $\log n$ does not hold for graphs of treewidth at most k. This suggests that our results could be improved in the restricted setting of edge-maximal graphs from the considered classes.

Finally, a natural research direction about this problem is to obtain tight bounds for our theorems, especially in the cases of bounded pathwidth or treewidth. To the best of our knowledge, this question is also open for planar graphs, and more generally graphs of bounded Euler genus (see Table 2.1).

2.2.2 Preliminaries

Remark 2.2.5. If G is a graph that has a path P of order n and $X \subseteq V(G)$ is not empty, then $G - X$ has

1. at most $|X| + 1$ connected components that contain a vertex of P; and
2. a connected component that has a path of order at least $\frac{n-|X|}{|X|+1} \geqslant \frac{n}{2|X|} - 1$.

Long paths versus Hamiltonian paths. A *Hamiltonian path* in a graph G is a path that visits all the vertices of G.

The statements of the results in subsection 2.2.1 follow the general form of ($\star$). In contrast, in the rest of the paper we work with statements of the following form (for $\mathcal{G}$ a class of graphs):

($\star\star$) *For every $n \in \mathbb{N}$, if a graph of $\mathcal{G}$ has a Hamiltonian path and order n then it has an induced path of order at least $f(n)$.*

Remark 2.2.6. For a hereditary graph class $\mathcal{G}$, the statements ($\star$) and ($\star\star$) are equivalent. Given a graph $G \in \mathcal{G}$ with a path P of order n (as required by ($\star$)), it suffices to consider the induced subgraph $G[V(P)] \in \mathcal{G}$ to be able to apply statement ($\star\star$). This shows that ($\star\star$) implies ($\star$) and the other direction is trivial.

However the form ($\star\star$) is more convenient for the proofs because the induced paths that we construct will never use vertices other than those of the path P whose existence is assumed, so by using form ($\star\star$) we do not need to explicitly say that we restrict our attention to $G[V(P)]$.

Representations. To easily deal with graphs of bounded pathwidth or treewidth (defined hereafter) we find it convenient to define tree representations, which are objects that are closely related to tree decompositions, as we explain below. Formally, a *tree representation* of a graph G is a pair $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ such that

1. T is a tree;
2. for every $v \in V(G)$, T_v is a subtree of T, called the *model* of v;
3. for every edge (u, v) of G the subtrees T_u and T_v intersect.

When T is a path, we call $\mathcal{T}$ a *path representation*. If item (3) is strengthened as follows

3'. (u, v) is an edge of G if and only if T_u and T_v intersect

then G is the intersubsection graph of the vertex sets of the subtrees $\{T_v\}_{v \in V(G)}$ and it is a *chordal graph*[2]; if furthermore T is a path then G is an *interval graph*. In this latter case we call $\mathcal{T}$ an *interval representation* of G to stress that it is an interval graph. To avoid confusion between the vertices of G and T, we use the synonym *nodes* to refer to vertices of T.

Tree representations and tree decompositions are closely linked, as we explain now. For a tree representation $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ of a graph G, we define for every $t \in V(T)$ the *bag at t* as the following subset of $V(G)$

$$\beta_{\mathcal{T}}(t) = \{v \in V(G) \mid t \in T_v\}.$$

[2]Chordal graphs are usually defined as graphs with no induced cycle of order 4 or more. They can be seen as intersubsection graphs of subtrees of a tree, as proved by Gavril [Gav74, Theorem 3].

We drop the subscript when there is no ambiguity. The *width* of a tree representation $\mathcal{T}$ is $\max_{t\in V(T)} |\beta_{\mathcal{T}}(t)| - 1$. The *treewidth* of G is the minimum width of any of its tree representations and the *pathwidth* of G is the minimum width of any of its path representations. It can easily be seen that these definitions coincide with the usual definitions for treewidth and pathwidth as, with the notation above, $(T, \{\beta(t)\}_{t\in V(T)})$ is a tree decomposition of G. We respectively denote by $\mathbf{tw}(G)$ and $\mathbf{pw}(G)$ the treewidth and pathwidth of G.

The following is a consequence of items (2) and (3) of the definition of a tree representation.

Remark 2.2.7. Let G be a graph with a tree representation $(T, \{T_v\}_{v\in V(G)})$ and let H be a connected subgraph of G. Then $\bigcup_{v\in V(H)} T_v$ is a (connected) subtree of T.

2.2.3 Induced paths in graphs of bounded pathwidth

In this subsection we show that the maximum function f_k such that property ($\star$) holds for graphs of pathwidth less than k is such that $\frac{1}{3}n^{1/k} \leqslant f_k(n) \leqslant n^{2/k} + 1$ (Theorem 2.2.2 and Theorem 2.2.11).

A *caterpillar* is a tree in which all the vertices are at distance at most one of some path.

Lemma 2.2.8 (See [PT99, subsection 6])**.** *Every connected graph of pathwidth at most one is a caterpillar.*

We will also use the following consequence of the Helly property of intervals (see for instance [GGL95, subsection 2.5]).

Lemma 2.2.9. *Suppose $\mathcal{P} = (P, \{P_v\}_{v\in V(G)})$ is a path representation of a graph G and K is a clique of G. Then there is a node $t \in V(P)$ such that $V(K) \subseteq \beta_{\mathcal{P}}(t)$.*

We first prove Theorem 2.2.2, that we restate here for convenience.

Theorem 2.2.2. *For every $k \in \mathbb{N}$, if G is a graph of pathwidth less than k that has a path of order n, then G has an induced path of order at least $\frac{1}{3}n^{1/k}$.*

Proof. We actually prove the following by induction on n and k, which is equivalent to the desired statement according to Remark 2.2.6.

> *If G is a graph of order at least n and pathwidth less than k that has a Hamiltonian path, then G has an induced path of order at least $\frac{1}{3}n^{\frac{1}{k}}$.*

Note that the statement is vacuously true when $k = 1$ and $n > 1$ as there is no connected graph with $|G| \geqslant 2$ and $\mathbf{pw}(G) = 0$. When $\mathbf{pw}(G) = n-1$ (thus $k = n$), the graph is a clique and the order of its longest induced path is $2 \geqslant \frac{1}{3}n^{\frac{1}{k}}$. In the

cases where $n \geqslant 2$ and $k = 2$, G is a caterpillar, by Lemma 2.2.8. Every caterpillar with a Hamiltonian path is a path, so the statement holds as G is already an induced path and $n \geqslant \frac{1}{3}\sqrt{n}$.

So we now assume that $k \geqslant 3$, $n > k$ and that the statement holds for all smaller values of k and n. We now show that it also true for k and n. We may also assume that $\frac{1}{3}n^{\frac{1}{k}} > 2$ as otherwise any edge is an induced path of the required order.

Let G be a graph that has a Hamiltonian path and such that $|G| \geqslant n$ and $\mathbf{pw}(G) < k$. We fix a path representation $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ of G of width less than k and with no empty bag. Let u be a vertex of G such that R_u contains one endpoint of R and let v be a vertex whose model contains the other endpoint; possibly $u = v$. Let Q be an induced path in G between u and v. Observe that Q may consist of a single vertex (when $u = v$) or of two vertices (when u and v are adjacent). Otherwise, Q can be constructed by taking a shortest path between u and v.

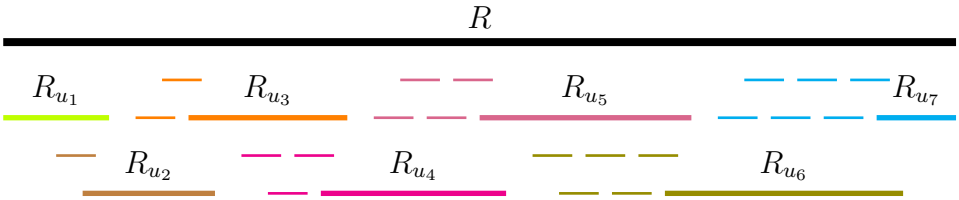
In the case where $|Q| \geqslant \frac{1}{3}n^{\frac{1}{k}}$ we are done and Q is the desired path. So in the rest of the proof we may assume that $|Q| < \frac{1}{3}n^{\frac{1}{k}}$. Let P be a Hamiltonian path of G. The path Q intersects P in $|Q|$ vertices thus removing Q from G cuts P into at most $|Q| + 1$ subpaths, and the longest of them, that we call P', has an order n' that is at least $\frac{n-|Q|}{|Q|+1}$. Let us consider G' the graph induced by P' in G.

We now show that $\mathbf{pw}(G') \leqslant k-1$. For this we consider the path representation $\mathcal{R}' = (R, \{R_w\}_{w \in V(G')})$ of G'. Let $r \in R$ be a node such that $|\beta_{\mathcal{R}'}(r)|$ is maximum. From the definition we have $\beta_{\mathcal{R}'}(r) \subseteq \beta_{\mathcal{R}}(r)$. By Remark 2.2.7, the definition of Q and the fact that it is connected, we know that the union of the sets $\{R_w\}_{w \in V(Q)}$ is equal to $V(R)$. Therefore $\beta_{\mathcal{R}}(r)$ contains a vertex of Q. This implies $|\beta_{\mathcal{R}}(r)| > |\beta_{\mathcal{R}'}(r)|$ and the claimed bound on the pathwidth of G' follows.

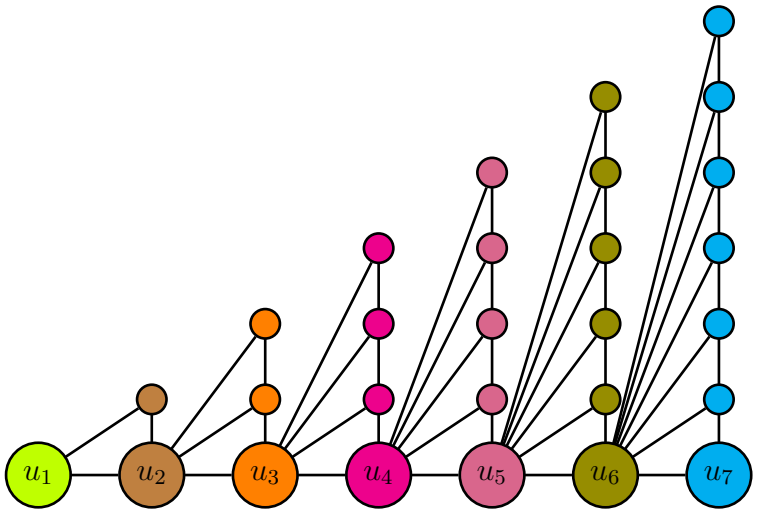
The graph G' has a Hamiltonian path (by definition), order $n' < n$ and pathwidth at most $k - 1$. By induction, G' admits an induced path Q' of order at least $\frac{1}{3}n'^{\frac{1}{k-1}}$. Recall that $n' \geqslant \frac{n-|Q|}{|Q|+1}$. Since $|Q| < \frac{1}{3}n^{\frac{1}{k}}$, we have $n' > \frac{n}{\frac{1}{3}n^{\frac{1}{k}}+1} - 1$. As we assume $\frac{1}{6}n^{\frac{1}{k}} \geqslant 1$ we deduce $n' > 2n^{\frac{k-1}{k}} - 1 \geqslant n^{\frac{k-1}{k}}$. Therefore we have $|Q'| \geqslant \frac{1}{3}(n^{\frac{k-1}{k}})^{\frac{1}{k-1}} = \frac{1}{3}n^{\frac{1}{k}}$. Since G' is an induced subgraph of G, Q' is also an induced path of G so we are done. □

As every interval graph of clique number k admits a path representation of width less than k (given by its interval representation), we have the following improvement of the bound given by Esperet et al. in Table 2.1

Corollary 2.2.10. *For every $k, n \in \mathbb{N}$, if G is an interval graph of order at least n and clique number at most k that has a Hamiltonian path, then G has an induced path of order at least $\frac{1}{3}n^{\frac{1}{k}}$.*



(a) The interval representation of the graph $G_{n,3}$ with $q = 7$.



(b) The graph $G_{n,3}$ with $q = 7$.

Figure 2.2: The construction of Theorem 2.2.11.

The following statement complements Theorem 2.2.2 and Corollary 2.2.10 by giving an upper-bound on the order of induced paths one can guarantee in (interval) graphs of bounded pathwidth.

Theorem 2.2.11. *For every $k, n \in \mathbb{N}$ with $2 \leqslant k \leqslant n$, there exists an interval graph $G_{n,k}$ with a Hamiltonian path of order at least n and clique number at most k, such that every induced path of $G_{n,k}$ has order at most $n^{\frac{2}{k}} + 1$.*

Proof. The proof is by induction on n and k. When $k = n$, $G_{n,n}$ is the clique on n vertices, where every induced path has order at most 2, which is less than $n^{\frac{2}{n}} + 1$. When $k = 2$, $G_{n,2}$ is the (induced) path on n vertices, which again satisfies the desired statement for every $n \geqslant 2$.

For any $n \geqslant k$, let $q = \left\lfloor n^{\frac{2}{3}} + 1 \right\rfloor$. When $k = 3$, $G_{n,3}$ is constructed from a collection $\{P_i\}_{i \in \{1,\dots,q\}}$ of q paths, where for every $i \in \{1, \dots, q\}$ the path P_i has order i and an endpoint called u_i, by connecting u_i to all the vertices of P_{i+1}

for every $i \in \{1, \dots, q-1\}$. See Figure 2.2b for a depiction of a small case and Figure 2.2a for an interval representation $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ of it, where each R_v is drawn below R respecting the x-axis.

We can see that at most 3 intervals intersect (R_{u_i} and the intervals representing two consecutive vertices of P_{i+1} for each i), thus $G_{n,3}$ is an interval graph with clique number 3.

The number of vertices of $G_{n,3}$ is $\sum_{i=1}^{q} |P_i| = \frac{q(q+1)}{2} \geqslant \frac{n^{\frac{2}{3}}(n^{\frac{2}{3}}+1)}{2}$ since $q \geqslant n^{\frac{2}{3}}$, which is greater than n. $G_{n,3}$ admits a Hamiltonian path starting in u_q that, for each i from q to 2, follows P_i from u_i to its other endpoint, goes to u_{i-1}, and repeats the same process.

Let us now bound the maximum order of an induced path in $G_{n,3}$. Let Q be an induced path of $G_{n,3}$ and let i denote the minimum integer such that Q has a vertex from P_i; clearly Q has at most i vertices from this path. Let $j \in \{i+1, \dots, q\}$ and observe that every vertex of P_j has u_{j-1} as unique neighbor in $\{P_{j'}\}_{j'<j}$. As Q is induced we deduce that it contains at most one vertex of P_j. This holds for each of the $q-i$ paths of $\{P_j\}_{j \in \{i+1,\dots,q\}}$, so we get the bound $|Q| \leqslant q$, as desired. This concludes the proof for the case $k = 3$.

So we now take $n > k \geqslant 4$, and assume that $G_{n',k'}$ is defined and satisfies the statement for every $k' < k$ and $n' < n$ such that $2 \leqslant k' \leqslant n'$. To construct $G_{n,k}$, we proceed as follows.

Let $q = \left\lfloor n^{\frac{2}{k}} + 1 \right\rfloor$. If $q \geqslant n$ then the graph $G_{n,k} = P_q$ clearly satisfies the desired statement. Otherwise, we set $n' = \left\lceil \frac{n-q}{q-1} \right\rceil$; observe that $n' \geqslant 2$. We construct $G_{n,k}$ from the disjoint union of a path $Q = v_1 - \cdots - v_q$ and $q-1$ copies $H_1, \dots, H_{q-1}$ of $G_{n',k-2}$ by connecting v_i and v_{i+1} to all vertices of H_i, for every $i \in \{1, \dots, q-1\}$.

To show that $G_{n,k}$ is an interval graph, we now provide an interval representation of it (see Figure 2.3 for an illustration). By our induction hypothesis, H_i admits an interval representation $\mathcal{M}^i = (M^i, \{M^i_v\}_{v \in V(H_i)})$ for every $i \in \{1, \dots, q-1\}$. Let M denote the concatenation of the paths $M^1, \dots, M^{q-1}$. That is, M is obtained from the disjoint union of these paths by adding an edge between an endpoint of M^1 and one endpoint of M^2, from the other endpoint of M^2 to one endpoint of M^3, and so on. Clearly $\left(M, \bigcup_{i=1}^{q-1} \{M^i_v\}_{v \in V(H_i)}\right)$ is an interval representation of the disjoint union of the H_i's. For every $i \in \{2, \dots, q-1\}$, let $M_{v_i} = M[V(M^{i-1}) \cup V(M^i)]$ and let $M_{v_1} = M^1$ and $M_{v_q} = M^{q-1}$. Observe that $(M, \{M_{v_i}\}_{i \in \{1,\dots,q\}})$ is an interval representation of Q. Furthermore, $\mathcal{M} = (M, \{M_{u_i}\}_{i \in \{1,\dots,q\}} \cup \bigcup_{i=1}^{q-1} \{M^i_v\}_{v \in V(H_i)})$ is an interval representation of $G_{n,k}$. This proves that $G_{n,k}$ is an interval graph.

The graph $G_{n,k}$ is composed of the path Q on q vertices and $q-1$ copies of

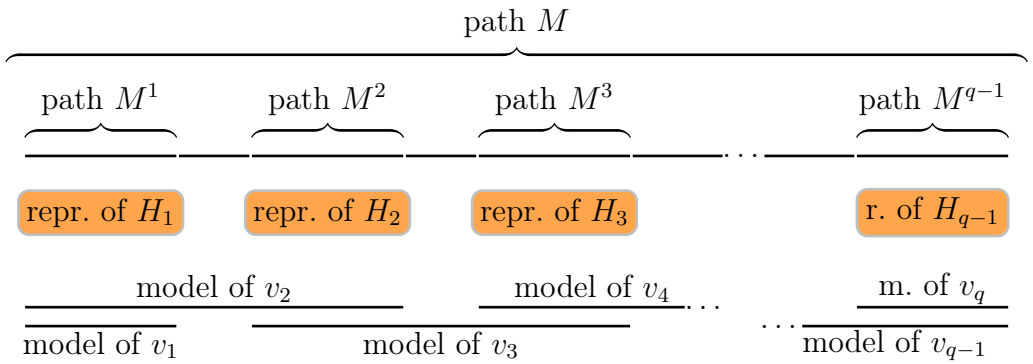


Figure 2.3: The construction of an interval representation for $G_{n,k}$ when $k > 3$.

$G_{n',k-2}$ which have (by induction) at least n' vertices each, so we have

$$\begin{aligned} |G_{n,k}| &\geqslant q + (q-1)n' \\ &= q + (q-1)\left\lceil \frac{n-q}{q-1} \right\rceil \\ &\geqslant n. \end{aligned}$$

We now show that the clique number of $G_{n,k}$ is at most k. Let K be a maximum clique of $G_{n,k}$. By Lemma 2.2.9, there exists a node x of M such that K consists of all vertices of G whose model (in $\mathcal{M}$) contains x. Let j be such that $x \in M^j$ (such a vertex exists by definition of M), we then have

$$V(K) = \quad (\, V(K) \cap V(Q)\,) \quad \cup \quad (\, V(K) \cap V(H_j)\,).$$

The first intersubsection has size at most 2 since Q is an induced path. Recall that H_j is a copy of $G_{n',k'}$ thus its clique number is at most $k-2$. Therefore the second intersubsection has size at most $k-2$. We deduce $|K| \leqslant (k-2)+2 = k$.

Let us show that $G_{n,k}$ has a Hamiltonian path. For every i, let R_i denote a Hamiltonian path of H_i, which exists by induction hypothesis. By construction, every vertex of R_i (in particular its endpoints) is adjacent to both v_{i-1} and v_i in $G_{n,k}$. Therefore $v_1 - R_1 - v_1 - R_2 \dots v_{q-1} - R_{q-1} - v_q$ is a (Hamiltonian) path in G.

We proved that $G_{n,k}$ is an interval graph on at least n vertices, with clique number at most k, and with a Hamiltonian path. In order to conclude the proof, it remains to prove that G does not have an induced path longer than $n^{2/k}+1$. Let P be an induced path of G of maximum length.

We first consider the case where $|V(P) \cap V(H_i)| \geqslant 2$ for some i. As for every $w \in V(H_i)$, $N(w) \setminus V(H_i) = \{v_{i-1}, v_i\}$, we deduce that $V(P) \subseteq V(H_i)$, otherwise P would not be induced. By induction hypothesis we get $|P| \leqslant n'^{\frac{2}{k-2}} + 1$. Observe

that

$$\begin{aligned} n' &\leqslant \frac{n-q}{q-1} + 1 && \text{(from the definition)} \\ &\leqslant \frac{n}{q-1} \\ &\leqslant n^{\frac{k-2}{k}}. \end{aligned}$$

So $|P| \leqslant n^{\frac{2}{k}} + 1$, as required. We now consider the remaining case where $|V(P) \cap V(H_i)| \leqslant 1$ for all i. Notice that an internal vertex w of P cannot belong to H_i for some i. Indeed the only neighbors of w outside H_i are v_{i-1} and v_i, which are adjacent. We then get the three following cases:

- either $V(P)$ does not intersect any H_i, in which case $P = Q$ so $|P| \leqslant n^{\frac{2}{k}} + 1$;
- or there is an i such that P starts at some vertex $w \in V(H_i)$ and does not intersect H_j for all $j \neq i$, in which case $P = w - v_i - v_{i+1} \dots v_q$ or $P = w - v_{i-1} - v_{i-2} \dots v_1$ and $|P| \leqslant |Q| \leqslant n^{\frac{2}{k}} + 1$;
- or there are two integers i, j with $i < j$ such that P starts from some vertex $w_i \in V(H_i)$ and ends at some vertex w_j of $V(H_j)$, in which case $P = w_i - v_i - v_{i+1} \dots v_{j-1} - w_j$ and again $|P| \leqslant |Q| \leqslant n^{\frac{2}{k}} + 1$.

□

2.2.4 Induced paths in graphs of bounded treewidth

In this subsection we show that the maximum function f_k such that property ($\star$) holds for graphs of treewidth less than k is such that $f_k(n) \geqslant \frac{1}{4} n^{1/k}$ (Theorem 2.2.3).

Let G be a graph and let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a tree representation of G. The *weight of a path* P of T is the number of vertices v of G such that T_v intersects P. The *weight of a node* x of T is the maximum weight of a path from x to a leaf minus $|\beta_{\mathcal{T}}(x)|$, and is noted $w_{\mathcal{T}}(x)$ (or $w(x)$ when there is no ambiguity on the tree representation).

It is well-known that the order of a tree is upper-bounded by a function of its height and maximum degree. The following lemma extends this statement to graphs of bounded treewidth that have a Hamiltonian path.

Lemma 2.2.12. *Let $k, w \in \mathbb{N}$ and let G be a graph that has a Hamiltonian path. If there is a tree representation of G of width less than k that has a node of weight at most w, then $|G| \leqslant (k+1)^{w+1} - 1$.*

Proof. Let $k \in \mathbb{N}$. The proof is by induction on w.

If $w = 0$ then for every graph G and tree representation $\mathcal{T}$ as in the statement of the lemma and x node of weight zero we have $V(G) \subseteq \beta_{\mathcal{T}}(x)$ and there are at most k vertices in $\beta_{\mathcal{T}}(x)$ so the claimed bound holds. So we now suppose that $w \geqslant 1$ and that the statement is true for every weight $w' < w$.

We consider a graph G with a Hamiltonian path P and a tree representation $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ of width less than k that has a node x such that $w_{\mathcal{T}}(x) \leqslant w$. Let us consider the graph $G \setminus \beta_{\mathcal{T}}(x)$. Removing the at most k vertices of $\beta_{\mathcal{T}}(x)$ cuts P into $t \leqslant k+1$ subpaths $P_1, \dots P_t$. For each i, let G_i be the graph induced by P_i and let T_i be the union of the T_v's for $v \in V(G_i)$. Then T_i is a subgraph of T and by Remark 2.2.7 it is connected. Observe that $\mathcal{T}_i = (T_i, \{T_v\}_{v \in V(G_i)})$ is a tree representation of G_i.

Let x_i be the node of T_i that is the closest to x in T. Let w_i be the weight of x_i in T_i and Q_i a path in T_i of maximum weight from x_i to a leaf l. Let us consider the path Q in T from x to l; note that Q_i is a subpath of Q. So x_i belongs to Q, and by construction of T_i, $\beta_{\mathcal{T}_i}(x_i)$ is not empty and belongs to $G \setminus \beta_{\mathcal{T}}(x)$ thus there is at least one vertex in $\beta_{\mathcal{T}}(x_i) \setminus \beta_{\mathcal{T}}(x)$. This implies that $w_{\mathcal{T}}(x) \geqslant w_i + 1$, and thus $w_i \leqslant w_{\mathcal{T}}(x) - 1 \leqslant w - 1$.

By construction, G_i admits a Hamiltonian path and we just proved that in the tree representation $(T_i, \{T_v\}_{v \in V(G_i)})$ of width less than k there is a node x_i of weight at most $w-1$. By induction, G_i has at most $(k+1)^w - 1$ vertices. As $V(G) = \beta_{\mathcal{T}}(x) \cup \bigcup_{i=1}^{t} V(G_i)$ with $t \leqslant k+1$ and $|\beta_{\mathcal{T}}(x)| \leqslant k$ we get $|G| \leqslant (k+1)((k+1)^w - 1) + k = (k+1)^{w+1} - 1$. □

Corollary 2.2.13. *In every tree representation of width less than k of a graph of order n there is a node of weight at least* $\log_{k+1}(n+1) - 1$.

In a graph G, the *contraction* of an edge (u, v) is the operation that creates a new vertex w adjacent to the neighbors of u and v and then deletes u and v. We say that a graph H is a *contraction* of a graph G if H can be obtained from G after a (possibly empty) sequence of edge contractions. In the sequel we use Corollary 2.2.13 to extract a graph with a long path and bounded pathwidth from a graph with a large path and bounded treewidth. The obtained graph will be a contraction of the original one, which is interesting for us because of the following property.

Remark 2.2.14. Let H be an induced subgraph or a contraction of a graph G. If H has an induced path of order n, then so does G.

Lemma 2.2.15. *Let $k, w \in \mathbb{N}$ and let G be a graph that has a Hamiltonian path and order n. If G admits a tree representation of width less than k that has a path of weight w, then there is a contraction of G that has a Hamiltonian path and is of order w and pathwidth less than k.*

Proof. Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be the tree representation as in the statement of the lemma and R the path of T of weight w. Let P be a Hamiltonian path of G. We prove the statement by induction on the number p of vertices v of G such that $V(T_v) \cap V(R) = \varnothing$. In the case $p = 0$, for every vertex $v \in V(G)$ the subtree T_v intersects R. By the properties of tree representations, we have:

- for every $v \in V(G)$, $V(T_v) \cap V(R)$ induces a (connected) subpath of R, that we call R_v; and
- for every $u, v \in V(G)$, T_u and T_v share a vertex if and only if they share a vertex of R.

Therefore $(R, \{R_v\}_{v \in V(G)})$ is a path representation of G. Clearly it has w vertices and width less than k hence we are done.

So we may assume in the sequel that $p > 0$ and that the statement holds for all values smaller than p. As $p > 0$ and P is a Hamiltonian path, there is an edge $(u, v) \in E(P)$ such that T_u intersects R while T_v does not. Let us consider the graph G' obtained by contracting (u, v) into a new vertex y and setting $T_y = T_u \cup T_v$ (which is connected as T_u and T_v intersect). We call $\mathcal{T}' = (T, \{T_z\}_{z \in V(G')})$ the corresponding tree representation. Observe that for every $t \in V(T)$, we have $\beta_{\mathcal{T}'}(t) = \beta_{\mathcal{T}}(t)$ if $t \notin T_u \cup T_v$ and $\beta_{\mathcal{T}'}(t) = \beta_{\mathcal{T}}(t) \setminus \{u, v\} \cup \{y\}$ otherwise. Therefore the width of $\mathcal{T}'$ is less than k. Also, observe that the weight of R in $\mathcal{T}'$ is still w. Applying the induction hypothesis on G' yields the desired result. □

We are now ready to prove Theorem 2.2.3, that we restate here for convenience.

Theorem 2.2.3. *For every $k \in \mathbb{N}$, if G is a graph of treewidth less than k that has a path of order n, then G has an induced path of order at least $\frac{1}{4}(\log n)^{1/k}$.*

Proof. The statement that we actually prove is the following, which implies the desired statement according to Remark 2.2.6.

> *For every $k, n \in \mathbb{N}$, if G is a graph of order at least n and treewidth less than k that has a Hamiltonian path, then G has an induced path of order at least $\frac{1}{4}(\log n)^{1/k}$.*

Let k, n, and G be as in the statement above. The cases $n \leqslant 2$ or $k \leqslant 2$ are trivial and the case $n = k$ is handled by Theorem 2.2.2 (as then G has pathwidth at most k), so we suppose $n > k \geqslant 3$.

By combining Corollary 2.2.13 and Lemma 2.2.15 we obtain a contraction G' of G of order at least $\log_{k+1}(n + 1)$ with pathwidth less than k and that has a Hamiltonian path.

By Theorem 2.2.2, G' admits an induced path Q of order

$$|Q| \geqslant \frac{1}{3}(\log_{k+1}(n+1))^{\frac{1}{k}} \geqslant \frac{1}{3 \cdot c}(\log(n+1))^{1/k}$$

where $c < 1.3$ is the maximum of the function $k \mapsto \log(k+1)^{1/k}$. By Remark 2.2.14 we deduce that G has an induced path of the same order and we are done. □

2.2.5 Induced paths in topological minor-closed classes

In this subsection we use the decomposition theorem of Grohe and Marx for graphs excluding a topological minor (Theorem 2.2.29 in this paper) in order to prove Theorem 2.2.4. According to Grohe and Marx' result, such graphs admit tree decompositions where the bags[3] are required to come from some prescribed graph classes.

We first show that ($\star$) holds with a polylogarithmic bound for graphs from these classes (subsections 2.2.5 and 2.2.5) and then that it does too in tree representations where, intuitively, the interaction between bags is low (subsection 2.2.5). These results are combined in subsection 2.2.5 to finally prove Theorem 2.2.4.

Almost bounded degree graphs

Let $\Delta, k \in \mathbb{N}$. We say that a graph G has (k, Δ)-*almost bounded degree* if G has a set of at most k vertices whose removal yields a graph of maximum degree at most Δ. In this subsection we show that such graphs satisfy property ($\star$) with a logarithmic bound (that depends on k and Δ), that is the following lemma.

Lemma 2.2.16. *For every $\Delta, k \in \mathbb{N}$ there is a constant $c \in \mathbb{R}^+$ such that if a graph with (k, Δ)-almost bounded degree has a path of order n, then it has an induced path of order at least $c \log n$.*

This result is a direct consequence of Corollary 2.2.18 (for deleting a constant number of vertices) and Corollary 2.2.20 (for graphs of bounded degree) that we prove below. They actually follow from more general statements, that we leave here as they may be useful in order to prove Conjecture 2.2.1.

Lemma 2.2.17. *Let $\mathcal{G}$ be a hereditary class of graphs such that for some $c, d \in \mathbb{R}^+$, if a graph $G \in \mathcal{G}$ has a Hamiltonian path and order n, then G has an induced path of order at least $c(\log n)^d$.*

Let $\varepsilon \in (0,1)$ and let $\mathcal{G}_\varepsilon$ denote the class of graphs such that for every $G \in \mathcal{G}_\varepsilon$ there is a subset $X \subseteq V(G)$ of at most n^ε vertices such that $G - X \in \mathcal{G}$. Then

[3] Actually the torsos, to be defined in one of the following subsections.

there is a constant $c' \in \mathbb{R}^+$ depending on c, d, and ε such that if a graph $G \in \mathcal{G}_\varepsilon$ has a Hamiltonian path and order n, then G has an induced path of order at least $c'(\log n)^d$.

Proof. Let $G \in \mathcal{G}_\varepsilon$ and let P be a Hamiltonian path of G. We may assume $n \geqslant 3$ otherwise P is already an induced path of the desired length. Let $X \subseteq V(G)$ be such that $G - X \in \mathcal{G}$ and $1 \leqslant |X| \leqslant n^\varepsilon$, which exists by definition of $\mathcal{G}_\varepsilon$. By Remark 2.2.5, $G - X$ has a component H with a path of order at least $n' = \frac{n-|X|}{|X|+1} \geqslant \frac{n}{2|X|} - 1$. As $\mathcal{G}$ is hereditary, $H \in \mathcal{G}$ so it has an induced path of order at least

$$\begin{aligned} c(\log n')^d &\geqslant c\left(\log\left(\frac{n^{1-\varepsilon}}{2} - 1\right)\right)^d \\ &\geqslant c'(\log n)^d \end{aligned}$$

for a suitable choice of the constant $c' > 0$ (depending on c, d, and ε). This path is an induced subgraph of G (by Remark 2.2.14) so we are done. □

Corollary 2.2.18. *Let $\mathcal{G}$ be a hereditary class of graphs such that for some $c, d \in \mathbb{R}^+$, if a graph $G \in \mathcal{G}$ has a Hamiltonian path and order n, then G has an induced path of order at least $c(\log n)^d$.*

Let $k \in \mathbb{N}$ and let $\mathcal{G}_k$ denote the class of graphs such that for every $G \in \mathcal{G}_k$ there is a subset $X \subseteq V(G)$ of order at most k vertices such that $G - X \in \mathcal{G}$. Then there is a constant $c' \in \mathbb{R}^+$ depending on c, d, and k such that if a graph $G \in \mathcal{G}_k$ has a Hamiltonian path and order n, then G has an induced path of order at least $c'(\log n)^d$.

Lemma 2.2.19. *Let $c \in \mathbb{R}^+$, $d \in [0, 1)$. Let $\mathcal{G}$ be the class containing every graph G with maximum degree at most $2^{c(\log |G|)^d}$. If a graph $G \in \mathcal{G}$ has a Hamiltonian path and order n, then G has an induced path of order at least $\frac{1}{c}(\log n)^{1-d}$.*

Proof. Let P be a Hamiltonian path of G. Let Q be an induced path of G of maximum order and let u be one of its endpoints. Let $q = |Q|$. For every $i \in \{0, \ldots, q-1\}$, let D_i denote the set of vertices at distance exactly i from u. As

$D_i \subseteq N(D_{i-1})$ for every $i \in \{1, \dots, q-1\}$ we get $|D_i| \leqslant |D_{i-1}| \cdot 2^{c(\log n)^d}$. Therefore

$$\begin{aligned} n &= \sum_{i=0}^{q-1} |D_i| && \text{(by maximality of } Q\text{)} \\ &\leqslant \sum_{i=0}^{q-1} \left(2^{c(\log n)^d}\right)^i \\ &\leqslant \left(2^{c(\log n)^d}\right)^q \\ \text{so} \quad q &\geqslant \frac{1}{c}(\log n)^{1-d}. \end{aligned}$$

□

The following corollary about graph classes of degree bounded by a constant Δ can be obtained from the previous lemma by taking $c = \log(\Delta)$ and $d = 0$.

Corollary 2.2.20. *Let $\Delta \in \mathbb{N}$. Every graph with maximum degree at most Δ that has a Hamiltonian path and order n has an induced path of order at least $\frac{\log n}{\log \Delta}$.*

Escaping the vortices

Similar to the concept of tree representation, we can define a *cycle representation* of a graph G as a pair $\mathcal{C} = (C, \{C_v\}_{v \in V(G)})$ such that:

1. C is a cycle;
2. for every $v \in V(G)$, C_v is a connected subgraph of C; and
3. for every edge (u, v) of G the subgraphs C_u and C_v intersect.

The notions of *bag* and *width* of a cycle representation are defined similarly as for tree representations.

Let G_0 be a graph embedded in a surface Σ. Let C be a facial cycle of G_0. A *C-vortex* is a cycle representation $(C, \{C_v\}_{v \in V(H)})$ of a graph H such that $V(H) \cap V(G_0) = V(C)$ and $v \in C_v$ for every $v \in V(C)$. Note that C is both a subgraph of H and the graph where the representation of H is defined.

For $g, p, a, k \in \mathbb{N}$, a graph G is *(g, p, a, k)-almost-embeddable* if for some set $A \subseteq V(G)$ with $|A| \leqslant a$ there are graphs $G_0, \dots, G_s$ with $s \leqslant p$ such that

1. $G - A = G_0 \cup G_1 \cup \cdots \cup G_s$;
2. $G_1, \dots, G_s$ are vertex-disjoint;
3. G_0 can be embedded in a surface of Euler genus at most g;

4. there are s pairwise vertex-disjoint facial cycles $F_1, \dots, F_s$ of G_0 in this embedding, and

5. for every $i \in \{1, \dots, s\}$, G_i has an F_i-vortex of width less than k.

This notion was introduced for the purpose of the proof of the Graph Minor Structure Theorem of Robertson and Seymour [RS03] and is also used in the decomposition theorem of Grohe and Marx on which we rely.

For $(g, 0, 0, 0)$-almost embeddable graphs, which by definition are graphs of Euler genus at most g, property $(\star)$ is known to hold with a polylogarithmic bound as proved by Esperet et al. (Table 2.1). We state here their theorem for convenience:

Theorem 2.2.21 ([ELM17, Theorem 3.8])**.** *For every $g \in \mathbb{N}$ there is a constant $c = \frac{1}{2\sqrt{6}} - o(1)$ such that for every graph G embeddable in a surface with Euler genus at most g, if G has a Hamiltonian path and order n, then G has an induced path of order at least $c\sqrt{\log n}$.*

We use this result as a base case to show the following more general statement.

Lemma 2.2.22. *For every $g, p, a, k \in \mathbb{N}$ with $k \geqslant 2$ there is a constant c such that the following holds. If G is a (g, p, a, k)-almost-embeddable graph that has a Hamiltonian path and order n, then G has an induced path of order at least $c\,(\log n)^{\frac{1}{k}}$.*

Proof. We may assume that $n \geqslant 3$ otherwise the statement is trivial. Let us first assume that $a = 0$. In the case where $p = 0$ then G has Euler genus at most g and the result follows from Theorem 2.2.21. So we assume that $p \geqslant 1$. Let $G_0, \dots, G_s$ and $F_1, \dots, F_s$ be defined as above.

Suppose first that for some $i \in \{1, \dots, s\}$, $|F_i| \geqslant (k+1)(\log n + 1)$. Let $\mathcal{C} = (F_i, \{C_v\}_{v \in V(G_i)})$ be an F_i-vortex of G_i of width less than k (given by the definition of (g, p, a, k)-almost-embeddable graphs). Let $u \in V(F_i)$. The bag $\beta_{\mathcal{C}}(u)$ has size at most k so some connected component of $F_i \setminus \beta_{\mathcal{C}}(u)$ contains a subpath F of F_i of order at least $\log n$. Let $H = G[V(F)]$. This graph has F as a Hamiltonian path. Observe that by definition of F, no model (in $\mathcal{C}$) of a vertex of F contains u. In other words, the models of the vertices of F are all subpaths of $F_i \setminus \{u\}$. Hence $(F_i \setminus \{u\}, \{C_v\}_{v \in V(H)})$ is a path representation of width less than k of H. By Theorem 2.2.2 we get an induced path of order at least $\frac{1}{3}(\log n)^{1/k}$ in H hence in G. We are not completely done yet but let us now consider the second case before concluding.

In the second case we assume that $|F_i| < (k+1)(\log n + 1)$ for every $i \in \{1, \dots, s\}$. This implies that $|G_i| < k(k+1)(\log n + 1)$. Let $X = \bigcup_{i=1}^{s} V(G_i)$, then $1 \leqslant |X| < pk(k+1)(\log n + 1)$.

By Remark 2.2.5, $G-X$ has a connected component G' that has a Hamiltonian path and order n' where $n' \geqslant \frac{n}{2|X|} - 1$. So $n' \geqslant n^{\varepsilon}$ for some constant $\varepsilon \in (0,1)$ that depends on p and k only.

Because it is a subgraph of G_0, the graph G' has Euler genus at most g. Applying Theorem 2.2.21 there is a constant c_g depending on g only such that G' (hence G) has an induced path of order q at least

$$\begin{aligned} q &\geqslant c_g\sqrt{\log n'} \\ &\geqslant c_g\sqrt{\varepsilon}\sqrt{\log n} \\ &\geqslant c_g\sqrt{\varepsilon}(\log n)^{1/k} && \text{as } k \geqslant 2. \end{aligned}$$

Let $c = \min(1/3, c_g\sqrt{\varepsilon})$. In both cases we obtained an induced path of order at least $c(\log n)^{1/k}$, as claimed.

The case where $a \geqslant 1$ follows from Corollary 2.2.18 applied to the case where $a = 0$. This concludes the proof. □

Representations of bounded adhesion

In this subsection we build upon the ideas developed in subsections 2.2.3 and 2.2.4 to show that, roughly, if $\mathcal{G}$ is a class of graphs where ($\star$) holds with a polylogarithmic bound then the same can be said of graphs obtained by gluing together in a tree-like fashion graphs from $\mathcal{G}$ (Lemma 2.2.27). This is a crucial step towards the proof of Theorem 2.2.4.

Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a tree representation of a graph G. The *adhesion set* of an edge (t,t') of T is defined as the following subset of $V(G)$:

$$\mathrm{adh}_{\mathcal{T}}(t,t') = \{v \in V(G) \mid \ \{t,t'\} \subseteq V(T_v)\}.$$

Equivalently, it can be defined as the intersubsection of the bags at t and t', that is $\mathrm{adh}_{\mathcal{T}}(t,t') = \beta_{\mathcal{T}}(t) \cap \beta_{\mathcal{T}}(t')$. We drop the subscript when it is clear from the context. The *adhesion* of $\mathcal{T}$ is the maximum size of the adhesion set of an edge of T. For every $t \in V(T)$, the *torso* of $\mathcal{T}$ at t is the graph obtained from $G[\beta(t)]$ by adding all edges (u,v) such that $u,v \in \mathrm{adh}_{\mathcal{T}}(t,t')$ for some neighbor t' of t. If for every $t \in V(T)$, the torso of $\mathcal{T}$ at t belongs to some graph class $\mathcal{G}$, we say that $\mathcal{T}$ *has torsos from* $\mathcal{G}$.

Torsos and adhesion sets provide two different ways to restrict tree representations. Observe that graphs of treewidth (respectively pathwidth) less than k are simply graphs that admit a tree representation (respectively path representation) with torsos from the class of graphs of order at most k. For every graph class $\mathcal{G}$ and integer a, we respectively denote by $\mathrm{PR}(\mathcal{G},a)$ and $\mathrm{TR}(\mathcal{G},a)$ the class of graphs that admit a path representation or a tree representation with torsos from $\mathcal{G}$ and

adhesion less than a. We denote by $\mathrm{TR}(\mathcal{G})$ the class of graphs that admit a tree representation with torsos from $\mathcal{G}$ and no restriction on the adhesion. Notice that if the graphs in $\mathcal{G}$ have cliques of bounded order, the adhesion of tree representations of graphs from $\mathrm{TR}(\mathcal{G})$ is implicitly bounded (i.e. $\mathrm{TR}(\mathcal{G}) \subseteq \mathrm{TR}(\mathcal{G}, a)$ for some $a \in \mathcal{N}$).

The next remark easily follows from the definition of tree representations.

Remark 2.2.23. Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a tree representation of a graph G, let $(t, t') \in E(T)$ and let F, F' be the two connected components of $T \setminus \{(t, t')\}$. Let u and u' be such that $V(T_u) \cap V(F) \neq \varnothing$ and $V(T_{u'}) \cap V(F') \neq \varnothing$. Then either one of u, u' belongs to $\mathrm{adh}_{\mathcal{T}}(t, t')$, or $G \setminus \mathrm{adh}_{\mathcal{T}}(t, t')$ has no path from u to u'.

For the purpose of the proof of Lemma 2.2.27 we need to relate the order of a graph with the length of a path where it is represented. It is not true in general that the existence of a path representation with a long path implies that the represented graph is large; for instance all the bags in this representation could be identical and small. We show below (Lemma 2.2.24) that such a statement holds if we require the considered path representation to satisfy an extra property, being varied, that we define now.

Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a tree representation of a graph G. For an edge (t, t') in T, the tree representation $(T', \{T'_v\}_{v \in V(G)})$ obtained from $\mathcal{T}$ by *contracting* (t, t') is defined as follows:

- T' is obtained from T by contracting (t, t');
- for every $v \in V(G)$, $T'_v = T_v$ if $\{t, t'\} \cap T_v = \varnothing$, and $T'_v = T_v \setminus \{t, t'\} \cup \{t''\}$ otherwise.

Intuitively, we merge the nodes t and t' both in the tree of the representation and in the models of the vertices.

Let us say that $\mathcal{T}$ is *varied* if no bag is a subset of a neighboring bag, i.e. for every $(t, t') \in E(T)$, $\beta(t) \not\subseteq \beta(t')$. In particular, unless G has no vertex, no bag is empty.

Given a tree representation $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$, it is possible to produce a varied tree representation by iteratively contracting in $\mathcal{T}$ the edges (t, t') such that $\beta(t) \subseteq \beta(t')$. Observe that this process changes neither the width or the adhesion of $\mathcal{T}$ nor the fact that it has bags from some specific class of graphs.

Lemma 2.2.24. *Let G be a graph on at least one vertex. If $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ is a varied path representation of G, then $|G| \geqslant |R|$.*

Proof. We prove the following statement by induction on ℓ.

> *For every $\ell \in \mathbb{N}^+$, if a graph G on at least one vertex admits a varied path representation on a path of order ℓ, then $|G| \geqslant \ell$.*

The case $\ell = 1$ is trivial as we require that G is not empty. So let us assume that $\ell > 1$ and that the statement holds for smaller values. Let G be a graph that admits a varied path representation $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ with $|R| = \ell$, and let $r_1 \dots r_\ell$ be the vertices of R in the order of the path. Let $G^- = G[\beta(r_1) \cup \dots \cup \beta(r_{\ell-1})]$. As $\mathcal{R}$ is varied, $\beta(r_1) \neq \varnothing$ so G^- has at least one vertex. Observe that it admits a varied path representation with $\ell - 1$ nodes (for instance $(R \setminus \{r_\ell\}, \{R_v\}_{v \in V(G^-)})$), so by induction it has at least $\ell - 1$ vertices. As $\mathcal{R}$ is varied, there is a vertex $v \in \beta(r_\ell)$ that does not belong to $\beta(r_{\ell-1})$. This vertex does not belong to G^- either (as R_v is connected) so $|G| \geqslant \ell$, as claimed. □

The above lemma allows us to prove the following variant of Theorem 2.2.2 on varied path representations of bounded adhesion.

Lemma 2.2.25. *Let G be a graph that has a Hamiltonian path. If G admits a varied path representation with adhesion less than a and at least ℓ nodes, then G has an induced path of order at least $\frac{1}{3}\ell^{\frac{1}{2a}}$.*

Proof. Let $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ be a path representation as in the statement. For every $r \in V(R)$, let Z_r be a subset of $V(G)$ of minimum size such that for every neighbor r' of r, $\operatorname{adh}(r, r') \subseteq Z_r$ and there is a vertex $v \in Z_r$ that does not belong to $\beta(r')$. Let us show that this set is well-defined and small. Either the bag at r has a vertex v that does not appear in the bags of any of its neighbors, in which case $\{v\} \cup \bigcup_{r' \in N(r)} \operatorname{adh}(r, r')$ satisfies the above properties, or it does not and then $\beta(r)$ is suitable, as $\mathcal{R}$ is varied. Observe that in this case $\beta(r) = \bigcup_{r' \in N(r)} \operatorname{adh}(r, r')$. Recall that in R the vertex r has up to two neighbors. So in both cases we have $|Z_r| \leqslant 2(a-1) + 1 = 2a - 1$.

Let P be a Hamiltonian path of G. An edge of P is called *superfluous* if it has at least one endpoint outside $\bigcup_{r \in V(R)} Z_r$.

Remark 2.2.26. Suppose that (x, y) is a superfluous edge and let G' be the graph obtained by contracting (x, y) into a new vertex z. Then $\mathcal{R}' = (R, \{R_v\}_{v \in V(G')})$ is a path representation of G' (with $R_z = R_x \cup R_y$). We show that additionally $\mathcal{R}'$ is a varied representation. Towards a contradiction, let us assume that there are $r, r' \in V(R)$ such that $\beta_{\mathcal{R}'}(r) \subseteq \beta_{\mathcal{R}'}(r')$. Then for one of x and y, say x, we have $\beta_{\mathcal{R}}(r) \setminus \beta_{\mathcal{R}}(r') = \{x\}$, which implies $y \in \operatorname{adh}_{\mathcal{R}}(r, r')$ (as (x, y) is an edge). We just proved that $x, y \in Z_r$, which is a contradiction with the choice of (x, y). So $\mathcal{R}'$ is varied.

Let H be the graph obtained from G after iteratively contracting all the superfluous edges and let $\mathcal{R}_H = (R, \{R_v\}_{v \in V(H)})$ be the corresponding varied path representation constructed as in the above remark. By Lemma 2.2.24, $|H| \geqslant |R| \geqslant \ell$. All the edges that were contracted did belong to P so H has a Hamiltonian path. No edge of H is superfluous (by definition) so for every $r \in V(R)$, $\beta_{\mathcal{R}_H}(r) = Z_r$.

Therefore $\mathcal{R}_H$ is a path representation of H of width at most $2a-1$. Applying Theorem 2.2.2 to H we get an induced path of order at least $\frac{1}{3}\ell^{\frac{1}{2a}}$. As H is a contraction of G, such a path also exists in G (Remark 2.2.14) hence we are done. □

A graph class is said to be *closed under taking subgraphs* if every subgraph of a graph of the class also belongs to the class. We are now ready to prove the main result of this subsection.

Lemma 2.2.27. *Let $a \in \mathbb{N}^+$, $c \in (0, 1/3]$ and $d \in (0, 1]$. Let $\mathcal{G}$ be a class of graphs that is closed under taking subgraphs and such that for every $n \in \mathbb{N}$, every $G \in \mathcal{G}$ that has a Hamiltonian path and order n has an induced path of order at least $c(\log n)^d$.*

If a graph G of $\mathrm{TR}(\mathcal{G}, a)$ has a Hamiltonian path and order n, then it has an induced path of order at least $c(\log n)^{\frac{1}{4a+\frac{1}{d}}}$.

Proof. Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a varied tree representation of G witnessing that $G \in \mathrm{TR}(\mathcal{G}, a)$. Let us fix $\varepsilon = \frac{4ad}{4ad+1} < 1$. Observe that we may assume

$$c(\log n)^{\frac{1}{4a+\frac{1}{d}}} > 2 \tag{2.1}$$

as otherwise the statement holds trivially. Let P be the Hamiltonian path of G.

Let us first suppose that T has a node t such that the bag $\beta(t)$ has order at least $n^{\frac{1}{(\log n)^{\varepsilon}}}$. Let H denote the graph obtained from G by iteratively contracting every edge of P that has at least one endpoint outside $\beta(t)$. By construction, H is a supergraph of $G[\beta(t)]$ of the same order and every edge $(u, v) \in E(H) \setminus E(G[\beta(t)])$ is such that $u, v \in \mathrm{adh}_{\mathcal{T}}(t, t')$ for some neighbor t' of t. Therefore, H is a subgraph of the torso of $\mathcal{T}$ at t. As $\mathcal{G}$ is subgraph-closed, we deduce $H \in \mathcal{G}$. Besides, these contractions were applied to edges of P so they yield a Hamiltonian path in H. By the properties of $\mathcal{G}$ we deduce that H has an induced path of order at least

$$\begin{aligned} c\left(\log n^{\frac{1}{(\log n)^{\varepsilon}}}\right)^d &= c(\log n)^{(1-\varepsilon)d} \\ &= c(\log n)^{\frac{1}{4a+\frac{1}{d}}}. \end{aligned}$$

Such a path also exists in G (Remark 2.2.14) so we are done.

So we may assume now that at every node t of T we have

$$|\beta(t)| < n^{\frac{1}{(\log n)^{\varepsilon}}}, \tag{2.2}$$

which implies

$$|T| \geqslant \frac{n}{n^{\frac{1}{(\log n)^{\varepsilon}}}} = n^{1-\frac{1}{(\log n)^{\varepsilon}}}. \tag{2.3}$$

Let $t \in V(T)$. As $\mathcal{T}$ is varied, for every neighbor t' of t there is a vertex $v_{t'} \in \beta(t')$ of G such that $v_{t'} \notin \beta(t)$. Let $G_{t'}$ denote the connected component of $G \setminus \beta(t)$ that contains $v_{t'}$. Then for every neighbor $t'' \neq t'$ of t, the components $G_{t'}$ and $G_{t''}$ are distinct, by the properties of tree representations. So $G \setminus \beta(t)$ has at least $\deg_T(t)$ connected components. By (2.2) and Remark 2.2.5, $G \setminus \beta(t)$ has at most $|\beta(t)| + 1$ connected components. We deduce that the maximum degree Δ of T is bounded as follows

$$\begin{aligned} \Delta &\leqslant n^{\frac{1}{(\log n)^{\varepsilon}}} + 1 \\ &\leqslant 2n^{\frac{1}{(\log n)^{\varepsilon}}}. \end{aligned} \tag{2.4}$$

Let R denote a path of maximum order in T. From the classic inequality $|T| \leqslant \Delta^{|R|-1}$ we get

$$\begin{aligned} |R| - 1 &\geqslant \frac{\log |T|}{\log \Delta} \\ &\geqslant \frac{\log n - (\log n)^{1-\varepsilon}}{(\log n)^{1-\varepsilon} + 1} && \text{from (2.3) and (2.4)} \\ &\geqslant \frac{1}{2}\left((\log n)^{\varepsilon} - 1\right) && \text{from (2.1)} \\ \text{so} \quad |R| &\geqslant \frac{1}{2}(\log n)^{\varepsilon} \\ &\geqslant (\log n)^{\varepsilon/2} && \text{from (2.1).} \end{aligned}$$

As in the first part of the proof we iteratively contract the edges of P that do not have both endpoints in $\bigcup_{t \in V(R)} \beta(t)$ in order to produce a graph H that has a Hamiltonian path and a varied path representation $(R, \{R_v\}_{v \in V(H)})$ with adhesion less than a. We can now apply Lemma 2.2.25 to get an induced path of order at least

$$\begin{aligned} \frac{1}{3}|R|^{\frac{1}{2a}} &\geqslant \frac{1}{3}(\log n)^{\frac{\varepsilon}{4a}} \\ &\geqslant c(\log n)^{\frac{1}{4a+\frac{1}{d}}} && \text{as } c \leqslant \frac{1}{3}. \end{aligned}$$

Again such a path also exists in G and we are done. □

Piecing things together

We are now ready to prove the following theorem from which will follow Theorem 2.2.4.

Theorem 2.2.28. *Let $k \in \mathbb{N}$ and let $\mathcal{G}_k$ denote the class of graphs that either are (k,k,k,k)-almost embeddable or have (k,k)-almost bounded degree. There are constants $c \in \mathbb{R}^+$ and $d \in (0,1)$ such that if a graph $G \in \mathrm{TR}(\mathcal{G}_k)$ has a Hamiltonian path and order n, then G has an induced path of order at least $c(\log n)^d$.*

Proof. Observe that a (k,k)-almost bounded degree graph does not contain a clique of order $2k+2$. Also there is a $k' \in \mathbb{N}$ such that no (k,k,k,k)-almost embeddable graph contains a clique of order k' (see for instance [DMW17, Lemma 21] for a linear upper-bound in terms of k). So for $a = \max(2k+2, k')$ we have $\mathrm{TR}(\mathcal{G}_k) \subseteq \mathrm{TR}(\mathcal{G}_k, a)$. By Lemma 2.2.22 and Lemma 2.2.16 the class $\mathcal{G}_k$ satisfies ($\star$) with the function $n \mapsto c'(\log n)^{d'}$ for some constants $c' \in \mathbb{R}^+$ and $d' \in (0,1)$ depending on k. Also, notice that $\mathcal{G}_k$ is closed under taking subgraphs. Together with Lemma 2.2.27 this implies that $\mathrm{TR}(\mathcal{G}_k, a)$ (hence $\mathrm{TR}(\mathcal{G}_k)$) satisfies ($\star$) with the function $n \mapsto c(\log n)^d$ for some constants $c \in \mathbb{R}^+$ and $d \in (0,1)$ depending on k. □

Theorem 2.2.4 is a consequence of Theorem 2.2.28 and the structure theorem of Grohe and Marx for graphs excluding a topological minor, stated hereafter in the setting of tree representations.

Theorem 2.2.29 ([GM15])**.** *For every graph H there is an integer $k \in \mathbb{N}$ such that every graph not containing H as a topological minor has a tree representation $(T, \{T_v\}_{v \in V(G)})$ such that for every $t \in V(T)$ the torso at t is either (k,k,k,k)-almost embeddable or has (k,k)-almost bounded degree.*

2.3 Largest induced forest

2.3.1 Overview

The main technical contribution in this section is the following.

Theorem 2.3.1. *Any t-biclique-free O_k-free graph on n vertices has a feedback vertex set of size $O_{t,k}(\log n)$.*

Since a graph with a feedback vertex set of size k has treewidth at most $k+1$, this implies a corresponding result on treewidth.

Corollary 2.3.2. *Every t-biclique-free O_k-free graph on n vertices has treewidth $O_{t,k}(\log n)$.*

We complement this result by exhibiting an infinite family of 3-biclique-free O_2-free graphs with treewidth logarithmic in the order of the graph.

Corollary 2.3.2 implies that a number of fundamental problems, in particular MIS, can be solved in polynomial time in sparse O_k-free graphs. In particular, this proves that MIS can be solved in polynomial time in the sparse regime.

When dropping the sparse assumption, we obtain a quasi-polynomial time algorithm:

Theorem 2.3.3. *For every positive integer k,* Maximum Independent Set *can be solved in quasi-polynomial time $n^{O(k^2 \log n)}$ in n-vertex O_k-free graphs.*

Organization of the section. In Section 2.3.2, we prove that Theorem 2.3.1 and Corollary 2.3.2 are tight already for $k = 2$ and $t = 3$. Section 2.3.3 solves MIS in O_k-free graphs in quasi-polynomial time, among other algorithmic applications of Corollary 2.3.2.

The proof of our main structural result, Theorem 2.3.1, spans from Section 2.3.4 to Section 2.3.8. After some preliminary results (Section 2.3.4), we show in Section 2.3.5 that it suffices to prove Theorem 2.3.1 when the graph G has a simple structure: a cycle C, its neighborhood N (an independent set), and the remaining vertices R (inducing a forest). Instead of directly exhibiting a logarithmic-size feedback vertex set, we rather prove that every such graph contains a vertex of degree linear in the so-called "cycle rank" (or first Betti number) of the graph. For sparse O_k-free graphs, the cycle rank is at most linear in the number of vertices and decreases by a constant fraction when deleting a vertex of linear degree. We then derive the desired theorem by induction, using as a base case that if the cycle rank is small, we only need to remove a small number of vertices to obtain a tree.

To obtain the existence of a linear-degree vertex in this simplified setting, we argue in Section 2.3.6 that we may focus on the case where the forest $G[R]$ contains only paths or only large "well-behaving" subdivided stars. In Section 2.3.7, we discuss how the O_k-freeness restricts the adjacencies between theses stars/paths and N. Finally, in Section 2.3.8, we argue that the restrictions yield a simple enough picture, and derive our main result.

2.3.2 Sparse O_2-free graphs with unbounded treewidth

In this subsection, we show the following.

Theorem 2.3.4. *For every natural k, there is an O_2-free 3-biclique-free (2^k+k-1)-vertex graph with treewidth at least $k-1$.*

In particular, for infinitely many values of n, there is an O_2-free 3-biclique-free n-vertex graph with treewidth at least $\log_2 n - 2$.

Construction of G_k. To build G_k, we first define a word w_k of length 2^k-1 on the alphabet $[k]$. We set $w_1 = 1$, and for every integer $i > 1$, $w_i = i\ w_{i-1}[1]\ i\ w_{i-1}[2]\ i\ \ldots\ i\ w_{i-1}[2^{i-1}-2]\ i\ w_{i-1}[2^{i-1}-1]\ i$. It is worth noting that equivalently $w_i = \mathrm{incr}(w_{i-1})\ 1\ \mathrm{incr}(w_{i-1})$, where incr adds 1 to every letter of the word. Let Π_k be the (2^k-1)-path where the ℓ-th vertex of the path (say, from left to right) is denoted by $\Pi_k[\ell]$.

The graph G_k is obtained by adding to Π_k an independent set of k vertices $v_1, v_2, \ldots, v_k$, and linking by an edge every pair $v_i, \Pi_k[\ell]$ such that $i \in [k]$ and $w_k[\ell] = i$.

Observe that we can also define the graph G_k directly, rather than iteratively: it is the union of a path $u_1, \ldots, u_{2^k-1}$ and an independent set $\{v_0, \ldots, v_{k-1}\}$, with an edge between v_i and u_j if and only if i is the 2-order of j (the maximum k such that 2^k divides j).

See Fig. 2.4 for an illustration.

G_k is O_2-free and 3-biclique-free. The absence of $K_{3,3}$ (even $K_{2,3}$) as a subgraph is easy to check. At least one vertex of the $K_{3,3}$ has to be some v_i, for $i \in [k]$. It forces that its three neighbors x, y, z are in Π_k. In turn, this implies that a common neighbor of x, y, z (other than v_i) is some $v_{i'} \neq v_i$; a contradiction since distinct vertices of the independent set have disjoint neighborhoods.

We now show that G_k is O_2-free. Assume towards a contradiction that $G_k[C_1 \cup C_2]$ is isomorphic to the disjoint union of two cycles $G_k[C_1]$ and $G_k[C_2]$. As C_1 and C_2 each induce a cycle, they each have to intersect $\{v_1, \ldots, v_k\}$. Assume without loss of generality that C_1 contains v_i, and C_2 is disjoint from

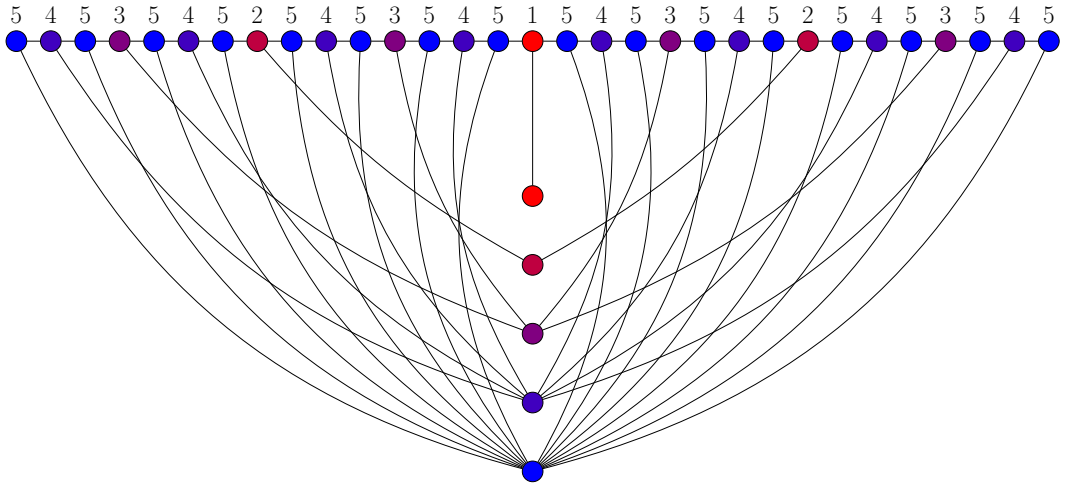


Figure 2.4: The graph G_k for $k = 5$: an O_2-free graph without $K_{3,3}$ subgraph, $k + 2^k - 1$ vertices, and treewidth at least $k - 1$.

$\{v_i, v_{i+1}, \dots, v_k\}$. Consider a subpath S of C_2 with both endpoints in $\{v_1, \dots, v_k\}$, thus in $\{v_1, \dots, v_{i-1}\}$, and all the other vertices of S form a set $S' \subseteq V(\Pi_k)$. It can be that the endpoints are in fact the same vertex $v_{i'}$, and in that case S is the entire C_2.

Let $v_{i'}, v_{i''}$ be the two (possibly equal) endpoints. Observe that S' is a subpath of Π_k whose two endpoints have label $i', i'' < i$. In particular there is a vertex labeled i somewhere along S'. This makes an edge between $v_i \in C_1$ and C_2, which is a contradiction.

G_k has treewidth at least $k - 1$. We build a K_k-minor as follows. We set $V_1 = \{v_1\} \cup R_1$ where $R_1 \subseteq V(\Pi_k)$ consists of the unique neighbor x_1 of v_1 together with all the vertices of Π_k to the right of x_1. Then, for every $i \in [2, k]$, we iteratively define V_i as $\{v_i\} \cup R_i$ where R_i is made of x_i the unique neighbor of v_i in $V(\Pi_k) \setminus \bigcup_{1 \leqslant j \leqslant i-1} V_j$ and all the vertices of $V(\Pi_k) \setminus \bigcup_{1 \leqslant j \leqslant i-1} V_j$ to the right of x_i.

By construction, it is clear that the V_i's are disjoint and that $G_k[V_i]$ is a path for every $i \in [k]$, hence is connected. It can also be observed that there is an edge between R_i and v_j with $j > i$. Thus the V_i's are the branch sets of a K_k-minor. Therefore $\mathbf{tw}(G_k) \geqslant \mathbf{tw}(K_k) = k - 1$. (It is easy to see that the treewidth of G_k is at most k, since $\{v_2, \dots v_k\}$ is a feedback vertex set.)

2.3.3 Algorithmic applications

This subsection presents algorithms on O_k-free graphs based on our main result, specifically using the treewidth bound.

Corollary 2.3.2. *Every t-biclique-free O_k-free graph on n vertices has treewidth $O_{t,k}(\log n)$.*

Single-exponential parameterized $O(1)$-approximation algorithms exist for treewidth. Already in 1995, Robertson and Seymour [RS95] present a $2^{O(\mathbf{tw})}n^2$-time algorithm yielding a tree-decomposition of width $4(\mathbf{tw}+1)$ for any input n-vertex graph of treewidth $\mathbf{tw}$. Run on n-vertex graphs of logarithmic treewidth, this algorithm outputs tree-decompositions of width $O(\log n)$ in polynomial time. We thus obtain the following.

Corollary 2.3.5. Maximum Independent Set, Hamiltonian Cycle, Minimum Vertex Cover, Minimum Dominating Set, Minimum Feedback Vertex Set, *and* Minimum Coloring *can be solved in polynomial time $n^{g(t,k)}$ in t-biclique-free O_k-free graphs, for some function g.*

Proof. Let $h(t,k)$ be the implicit function in Corollary 2.3.2 such that every t-biclique-free O_k-free n-vertex graph has treewidth at most $h(t,k)\log n$.

Algorithms running in time $2^{O(\mathbf{tw})}n^{O(1)} = 2^{h(t,k)\log n}n^{O(1)} = n^{h(t,k)+O(1)} = n^{g(t,k)}$ exist for all these problems but for Minimum Coloring. They are based on dynamic programming over a tree-decomposition, which by Corollary 2.3.2 has logarithmic width and by [RS95] can be computed in polynomial time. For Maximum Independent Set, Minimum Vertex Cover, Minimum Dominating Set, and q-Coloring (for a fixed integer q) see for instance the textbook [CFK+15, Chapter 7.3]. For Hamiltonian Cycle and Minimum Feedback Vertex Set, deterministic parameterized single-exponential algorithms require the so-called rank-based approach; see [CFK+15, Chapter 11.2].

By Corollary 2.3.14, t-biclique-free O_k-free graphs have bounded chromatic number. Thus a polynomial time algorithm for Minimum Coloring is implied by the one for q-Coloring. □

In a scaled-down refinement of Courcelle's theorem [Cou90], Pilipczuk showed that any problem expressible in Existential Counting Modal Logic (ECML) admits a single-exponential fixed-parameter algorithm in treewidth [Pil11]. In particular:

Theorem 2.3.6 ([Pil11])**.** *ECML model checking can be solved in polynomial time on any class with logarithmic treewidth.*

In a nutshell, this logic allows existential quantifications over vertex and edge sets followed by a counting modal formula that should be satisfied from every vertex v. Counting modal formulas enrich quantifier-free Boolean formulas with $\Diamond^S\varphi$, whose semantics is that the current vertex v has a number of neighbors satisfying φ in the ultimately periodic set S of non-negative integers. Another consequence of Corollary 2.3.2 (and Theorem 2.3.6) is that testing if a graph is O_k-free can be done in polynomial time among sparse graphs, further indicating that the general case could be tractable.

Corollary 2.3.7. *For any fixed k and t, deciding whether a t-biclique-free graph is O_k-free can be done in polynomial time.*

Proof. One can observe that O_k-freeness is definable in ECML. Indeed, one can write

$$\varphi = \exists X_1 \exists X_2 \dots \exists X_k \left(\bigwedge_{1 \leqslant i \leqslant k} X_i \to \Diamond^{\{2\}} X_i \right) \wedge \left(\bigwedge_{1 \leqslant i < j \leqslant k} \neg(X_i \wedge X_j) \wedge (X_i \to \Diamond^{\{0\}} X_j) \right).$$

Formula φ asserts that there are k sets of vertices $X_1, X_2, \dots, X_k$ such that every vertex has exactly two neighbors in X_i if it is itself in X_i, the sets are pairwise disjoint, and every vertex has no neighbor in X_j if it is in some distinct X_i (with $i < j$). Thus G is O_k-free if and only if φ does not hold in G. □

We now show the main algorithmic consequence of our structural result. This holds for any (possibly dense) O_k-free graph, and uses the sparse case (Corollary 2.3.5) at the basis of an induction on the size of a largest collection of independent 4-vertex cycles. It should be noted that this result (as well as the previous result on MIS above) also works for the weighted version of the problem, with minor modifications.

Theorem 2.3.3. *For every positive integer k,* Maximum Independent Set *can be solved in quasi-polynomial time $n^{O(k^2 \log n)}$ in n-vertex O_k-free graphs.*

Proof. Let G be our n-vertex O_k-free input. Let q be the maximum integer such that G admits q independent 4-vertex cycles (the cycles themselves need not be induced). Clearly $q < k$. We show the theorem by induction on q, namely that MIS can be Turing-reduced in time $n^{c(q+1)^2 \log n}$ for some constant c (specified later) to smaller 2-biclique-free instances (hence such that $q = 0$). We first examine what happens with the latter instances. Let $f(k) = h(2, k)$ with $h(t, k)$ the hidden dependence of Corollary 2.3.2. If $q = 0$, G is 2-biclique-free, so we can solve MIS in polynomial time $n^{f(k)+O(1)}$ by Corollary 2.3.5.

We now assume that $q \geqslant 1$, $n \geqslant 4$, and that the case $q - 1$ of the induction has been established (or $q - 1 = 0$). Let C be a 4-vertex cycle part of a $4q$-vertex subset consisting of q independent 4-vertex cycles. Let $\mathcal{S}$ be the set of all

$4q$-vertex subsets consisting of q independent 4-vertex cycles in the current graph (at this point, G), and $s = |\mathcal{S}|$. Thus $1 \leqslant s \leqslant n^{4q}$. By assumption, the closed neighborhood of C, $N[C]$, intersects every subset in $\mathcal{S}$. In particular, there is one of the four vertices of C, say, v, such that $N[v]$ intersects at least $s/4$ subsets of $\mathcal{S}$.

We branch on two options: either we put v in (an initially empty set) I, and remove its closed neighborhood from G, or we remove v from G (without adding it to I). With the former choice, the size of $\mathcal{S}$ drops by at least $s/4$, whereas with the latter, it drops by at least 1.

Even if fully expanded while $s > 0$, this binary branching tree has at most

$$\sum_{0 \leqslant i \leqslant 4q \log_{4/3} n} \binom{n}{i} = n^{O(q \log n)} \text{ leaves,}$$

since including a vertex in I can be done at most $4q \log_{4/3} n$ times within the same branch; thus, leaves can be uniquely described as binary words of length n with at most $4q \log_{4/3} n$ occurrences of, say, 1.

We retrospectively set $c \geqslant 1$ such that the number of leaves is at most $n^{cq \log n}$, running the algorithm thus far (when $q \geqslant 1$) takes at most time $n^{c+cq \log n}$. At each leaf of the branching, $s = 0$ holds, which means that the current graph does not admit q independent 4-vertex cycles. By the induction hypothesis, we can Turing-reduce each such instance in time $n^{cq^2 \log n}$. Thus the overall running time is

$$n^{c+cq \log n} + n^{cq \log n} \cdot n^{cq^2 \log n} \leqslant n^{c+cq \log n} \cdot (n^{cq^2 \log n} + 1) \leqslant n^{c(q+1)^2 \log n - cq \log n - c \log n + c + \frac{1}{\log n}}.$$

Note that $n^{cq^2 \log n} \geqslant 1$ thus we could upperbound $n^{cq^2 \log n} + 1$ by $2n^{cq^2 \log n} = n^{cq^2 \log n + \frac{1}{\log n}}$. Since $c, q \geqslant 1$ and $\log n \geqslant 1$, it holds that $-cq \log n - c \log n + c + \frac{1}{\log n} \leqslant -2c + c + 1 \leqslant 0$. Hence we get the claimed running time of $n^{c(q+1)^2 \log n}$ for the reduction to $q = 0$, and the overall running time of $n^{c(q+1)^2 \log n + f(k) + O(1)} = n^{O(k^2 \log n + f(k))}$. □

One may wonder if some other problems beside MIS become (much) easier on O_k-free graphs than in general. As $2K_2$-free graphs are O_2-free, one cannot expect a quasi-polynomial time algorithm for Minimum Dominating Set [Ber84, CP84], Hamiltonian Cycle [Gol04], Maximum Clique [Pol74], and Minimum Coloring [KKTW01] since these problems remain NP-complete on $2K_2$-free graphs. Nevertheless we give a quasi-polynomial time algorithm for 3-Coloring.

Theorem 2.3.8. *There exists a function f such that for every positive integer k,* 3-Coloring *can be solved in quasi-polynomial time $n^{O(k^2 \log n + f(k))}$ in n-vertex O_k-free graphs.*

Proof. We solve the more general List 3-Coloring problem, where, in addition, every vertex v is given a *list* $L(v) \subseteq \{1, 2, 3\}$ from which one has to choose its color. Note that when $L(v) = \varnothing$ for some vertex v, one can report that the instance is negative, and when $|L(v)| = 1$, v has to be colored with the unique color in its list, and this color has to be deleted from the lists of its neighbors (once this is done, v might as well be removed from the graph). These reduction rules are performed as long as they apply, so we always assume that the current instance has only lists of size 2 and 3.

We follow the previous proof, and simply adapt the branching rule, and the value of s. Now s is defined as the sum taken over all vertex sets X consisting of q independent 4-vertex cycles (the cycles themselves need not be induced), of the sum of the list sizes of the vertices of X. Hence $8 \leqslant s \leqslant 12 \cdot n^{4q}$. There is a vertex $v \in C$ and a color $c \in L(v)$ such that c appears in at least $\frac{1}{2} \cdot \frac{1}{12} \cdot \frac{s}{4} = \frac{s}{96}$ of the lists of its neighbors. This is because all the lists have size at least 2, and are subsets of $\{1, 2, 3\}$, thus pairwise intersect. (Note that this simple yet crucial fact already breaks down for List 4-Coloring.)

We branch on two options: either we color v with c, hence we remove color c from the lists of its neighbors or we commit to not color v by c, and simply remove c from the list of v. With the former choice, the size of $\mathcal{S}$ drops by at least $s/96$, whereas with the latter, it drops by at least 1. The rest of the proof is similar with a possibly larger constant c. □

2.3.4 Preliminary results

An important property of t-biclique-free graphs is that they are not dense (in the sense that they have a subquadratic number of edges).

Theorem 2.3.9 (Kővári, Sós, and Turán [KST54])**.** *For every integer $t \geqslant 2$ there is a constant c_t such that any n-vertex t-biclique-free graph contains at most $c_t\, n^{2-1/t}$ edges.*

The following lemma shows that for O_k-free graphs, being t-biclique-free is equivalent to a much stronger 'large girth' condition, up to the removal of a bounded number of vertices.

Lemma 2.3.10. *There is a function f such that for any integer $\ell \geqslant 3$ and any t-biclique-free and O_k-free graph G, the maximum number of vertex-disjoint cycles of length at most ℓ in G is at most $f(\ell, t, k)$.*

Proof. Let $N := (2c_t k\ell^2)^t$, where c_t is the constant of Theorem 2.3.9.

Assume for the sake of contradiction that G contains N vertex-disjoint cycles of length at most ℓ, which we denote by $C_1, \ldots, C_N$. Let H be the graph with

vertex set $v_1, \dots, v_N$, with an edge between v_i and v_j in H if and only if there is an edge between C_i and C_j in G. Since G is O_k-free, H has no independent set of size k. By Turán's theorem [Tur41], H contains at least $\frac{N^2}{2k-2} - \frac{N}{2} \geqslant \frac{N^2}{2k} - \frac{N}{2}$ edges.

Consider the subgraph G' of G induced by the vertex set $\bigcup_{i=1}^{N} C_i$. The graph G' has $n \leqslant \ell N$ vertices, and $m \geqslant 3N + \frac{N^2}{2k} - \frac{N}{2} > \frac{N^2}{2k}$ edges. Note that by the definition of N, we have

$$m > \frac{N^2}{2k} = \frac{1}{2k} \cdot N^{2-1/t} \cdot N^{1/t} \geqslant \frac{1}{2k\ell^{2-1/t}} \cdot n^{2-1/t} \cdot 2c_t k \ell^2 \geqslant c_t\, n^{2-1/t},$$

which contradicts Theorem 2.3.9, since G' (as an induced subgraph of G) is t-biclique-free.

□

The *girth* of a graph G is the minimum length of a cycle in G (if G is acyclic, its girth is set to be infinite). We obtain the following immediate corollary of Lemma 2.3.10.

Corollary 2.3.11. *There is a function g such that for any integer $\ell \geqslant 3$, any t-biclique-free and O_k-free graph G contains a set X of at most $g(\ell, t, k)$ vertices such that $G - X$ has girth at least ℓ.*

Proof. Let f be the function of Lemma 2.3.10, and let $g(\ell, t, k) := (\ell - 1) \cdot f(\ell - 1, t, k)$. Consider a maximum collection of disjoint cycles of length at most $\ell - 1$ in G. Let X be the union of the vertex sets of all these cycles. By Lemma 2.3.10, $|X| \leqslant (\ell - 1) f(\ell - 1, t, k) = g(\ell, t, k)$, and by definition of X, the graph $G - X$ does not contain any cycle of length at most $\ell - 1$, as desired. □

We now state (without proof) a slight variant of Lemma 2.3.10, which will be particularly useful at the end of the proof of our main result. A *banana* in a graph G is a pair of vertices joined by at least 2 disjoint paths whose internal vertices all have degree 2 in G.

Lemma 2.3.12. *There is a function f' such that any t-biclique-free and O_k-free graph G contains a set X of at most $f'(t, k) = O_t(k^t)$ vertices such that all bananas of G intersect X.*

The proof is the same as that of Lemma 2.3.10. We can take $f'(t, k) = 2f(2, t, k) = O(k^t)$, where f is the function of Lemma 2.3.10 and the implicit multiplicative constant in the $O(\cdot)$ depends on t. In all the applications of this lemma, t will be a small constant (2 or 3).

The *average degree* of a graph $G = (V, E)$, denoted by $\mathrm{ad}(G)$, is defined as $2|E|/|V|$. Let us now prove that t-biclique-free and O_k-free graphs have bounded average degree. This can also be deduced from the main result of [KO04], but we

include a short proof for the sake of completeness.[4] Moreover, the decomposition used in the proof will be used again in the proof of our main result.

Lemma 2.3.13. *Every O_k-free graph G of girth at least 11 has average degree at most $2k$.*

Proof. We proceed by induction on k. When $k = 1$, G is a forest, with average degree less than 2. Otherwise, let C be a cycle of minimal length in G. Let N be the neighborhood of C, let S the second neighborhood of C, and let $R = V(G)\backslash(C \cup N)$. Thus $V(G)$ is partitioned into C, N, R, and we have $S \subseteq R$. Observe that there are no edges between C and R in G, so it follows that $G[R]$ is O_{k-1}-free, and thus $\mathrm{ad}(G[R]) \leqslant 2k-2$ by induction. Observe also that since G has girth at least 11 and C is a minimum cycle, the two sets N and S are both independent sets. Moreover each vertex of N has a unique neighbor in C, and each vertex in S has a unique neighbor in N. Indeed, in any other case we obtain a path of length at most 5 between two vertices of C, contradicting the minimality of C. It follows that C is the only cycle in $G[C \cup N \cup S]$, hence this graph has average degree at most 2. As a consequence, G has a partition of its edges into two subgraphs of average degree at most $2k - 2$ and at most 2, respectively, and thus $\mathrm{ad}(G) \leqslant 2k - 2 + 2 = 2k$, as desired. □

It can easily be deduced from this result that every t-biclique-free O_k-free graph has average degree at most $h(t, k)$, for some function h (and thus chromatic number at most $h(t, k) + 1$).

Corollary 2.3.14. *There is a function h such that every t-biclique-free O_k-free graph has average degree at most $h(t, k)$, and chromatic number at most $h(t, k)+1$.*

Proof. Let G be a t-biclique-free O_k-free graph. By Corollary 2.3.11, G has a set X of at most $g(11, t, k)$ vertices such that $G - X$ has girth at least 11. Note that $\mathrm{ad}(G) \leqslant \mathrm{ad}(G - X) + |X| \leqslant \mathrm{ad}(G - X) + g(11, t, k) \leqslant 2k + g(11, t, k)$, where the last inequality follows from Lemma 2.3.13.

Let $h(t, k) = 2k + g(11, t, k)$. As the class of t-biclique-free O_k-free graphs is closed under taking induced subgraphs, it follows that any graph in this class is $h(t, k)$-degenerate, and thus $(h(t, k) + 1)$-colorable. □

A *feedback vertex set* (FVS) X in a graph G is a set of vertices of G such that $G - X$ is acyclic. The minimum size of a feedback vertex set in G is denoted by $\mathrm{fvs}(G)$. The classical Erdős-Pósa theorem [EP65] states that graphs with few vertex-disjoint cycles have small feedback vertex sets.

[4] Using the more recent result of [Dvo18], it can be proved that the class of t-biclique-free and O_k-free graphs actually has *bounded expansion*, which is significantly stronger than having bounded average degree. This can also be deduced from our main result, as it implies that sparse O_k-free graphs have logarithmic separators, and thus polynomial expansion.

Theorem 2.3.15 (Erdős and Pósa [EP65])**.** *There is a constant $c > 0$ such that if a multigraph G contains less than k vertex-disjoint cycles, then* $\mathrm{fvs}(G) \leqslant ck \log k$.

We use this result to deduce the following useful lemma.

Lemma 2.3.16. *There is a constant $c > 0$ such that the following holds. Let G consist of a cycle C, together with ℓ paths $P_1, \ldots, P_\ell$ on at least 2 edges*

- *whose endpoints are in C, and*
- *whose internal vertices are disjoint from C, and*
- *such that the internal vertices of each pair of different paths P_i, P_j are pairwise distinct and non-adjacent.*

Suppose moreover that G is O_k-free (with $k \geqslant 2$) and has maximum degree at most $d + 2$. Then

$$\ell \leqslant c\, d\, k \log k.$$

Proof. Observe that each path P_i intersects or is adjacent to at most $2(d-1) + 4d < 6d$ other paths P_j: indeed, if P_i has endpoints x, y in C, then there are at most $2(d-1)$ paths P_j which intersect P_i by sharing x or y as endpoint, and at most $4d$ paths P_j which are adjacent to P_i because some endpoint of P_j is adjacent to either x or y. It follows that there exist $s \geqslant \frac{\ell}{6d}$ of these paths, say $P_1, \ldots, P_s$ without loss of generality, that are pairwise non-intersecting and non adjacent.

Consider the subgraph G' of G induced by the union of C and the vertex sets of the paths $P_1, \ldots, P_s$. Since the paths P_i, $1 \leqslant i \leqslant s$, are non-intersecting and non-adjacent, and since G' does not contain k independent cycles, the graph G' does not contain k vertex-disjoint cycles. Let G'' be the multigraph obtained from G' by suppressing all vertices of degree 2 (i.e., replacing all maximal paths whose internal vertices have degree 2 by single edges). Observe that since G' does not contain k vertex-disjoint cycles, the graph G'' does not contains k vertex-disjoint cycles either. Observe also that G'' is cubic and contains $2s$ vertices. It was proved by Jaeger [Jae74] that any cubic multigraph H on n vertices satisfies $\mathrm{fvs}(H) \geqslant \frac{n+2}{4}$. As a consequence, it follows from Theorem 2.3.15 that $\frac{2s+2}{4} \leqslant \mathrm{fvs}(G'') \leqslant c'k \log k$ (for some constant c'), and thus $\ell \leqslant 12dc'k \log k = cdk \log k$ (for $c = 12c'$), as desired.

□

A *strict* subdivision of a graph is a subdivision where each edge is subdivided at least once.

Lemma 2.3.17. *There is a constant $c > 0$ such that for any integer $k \geqslant 2$, any strict subdivision of a graph of average degree at least $c\, k \log k$ contains an induced O_k.*

Proof. Note that if a graph G contains k vertex-disjoint cycles, then any strict subdivision of G contains an induced O_k. Hence, it suffices to prove that any graph with less than k vertex-disjoint cycles has average degree at most some function of k. By Theorem 2.3.15, there is a constant c' such that any graph G with less than k vertex-disjoint cycles contains a set X of at most $c'k\log k$ vertices such that $G - X$ is acyclic. In this case $G - X$ has average degree at most 2, and thus G has average degree at most $c'k\log k + 2 \leqslant ck\log k$ (for some constant c), as desired. □

2.3.5 Logarithmic treewidth of sparse O_k-free graphs

Recall our main result.

Theorem 2.3.1. *Any t-biclique-free O_k-free graph on n vertices has a feedback vertex set of size $O_{t,k}(\log n)$.*

The proof of Theorem 2.3.1 relies on *the cycle rank*, which is defined as $r(G) = |E(G)| - |V(G)| + |C(G)|$ where $C(G)$ denotes the set of connected components of G. The cycle rank is exactly the number of edges of G which must be deleted to make G a forest, hence it is a trivial upper bound on the size of a minimum feedback vertex set. Remark the following simple properties.

Lemma 2.3.18. *The cycle rank is invariant under the following operations:*

1. *Deleting a vertex of degree 0 or 1.*

2. *Deleting a connected component which is a tree.*

We call *reduction* the operation of iteratively deleting vertices of degree 0 or 1, which preserves cycle rank by the above lemma. A graph is *reduced* if it has minimum degree at least 2, and the *core* of a graph G is the reduced graph obtained by applying reductions to G as long as possible. The inclusion-wise minimal FVS of G and of its core are exactly the same.

In a graph G, a vertex x is called *ε-rich* if $d(x) \geqslant \varepsilon \cdot r(G)$. Our strategy to prove Theorem 2.3.1 is to iteratively reduce the graph, find an ε-rich vertex, add it to the FVS and delete it from the graph. The following lemma shows that the cycle rank decreases by a constant factor each iteration, implying that the process terminates in logarithmically many steps.

Lemma 2.3.19. *In a reduced O_k-free graph, deleting a vertex of degree d decreases the cycle rank by at least $\frac{d-k+1}{2}$.*

Proof. In any graph G, deleting a vertex x of degree d decreases the cycle rank by $d - c$, where c is the number of connected components of $G - x$ which contain a neighbor of x. If G is O_k-free, then all but at most $k - 1$ components of $G - x$ are trees. Furthermore, if T is a connected component of $G - x$ which is a tree, then T must be connected to x by at least two edges, as otherwise T must contain a vertex of degree 1 in G, which should have been deleted during reduction. Thus we have

$$2c - (k - 1) \leqslant d. \tag{2.5}$$

Therefore the cycle rank decreases by at least $d - \frac{d+k-1}{2} = \frac{d-k+1}{2}$ as desired. □

The existence of rich vertices is given by the following result.

Theorem 2.3.20. *For any k, there is some $\varepsilon_k > 0$ such that any O_k-free graph with girth at least 11 has an ε_k-rich vertex.*

Let us first prove Theorem 2.3.1 using Theorem 2.3.20.

Proof of Theorem 2.3.1. Fix k and t. Given a graph G which is O_k-free and t-biclique-free, we apply Lemma 2.3.10 to obtain a set X of size at most $f(11, t, k)$ such that $G' \stackrel{\text{def}}{=} G - X$ has girth at least 11. Thus, it suffices to prove the result for G', and finally add X to the resulting FVS of G'. Since $\log r(G') \leqslant \log \binom{|V(G')|}{2} \leqslant 2 \log |V(G')|$, we have reduced the problem to the following.

Claim 2.3.21. *For any k, there is a constant c_k such that if G is an O_k-free graph with girth at least 11, then* $\mathrm{fvs}(G) \leqslant c_k \cdot \log r(G)$.

Let us now assume that G is as in the claim, and consider its core H, for which $r(H) = r(G)$ and $\mathrm{fvs}(H) = \mathrm{fvs}(G)$. Consider an ε_k-rich vertex x in H with ε_k as in Theorem 2.3.20. If $r(G) \geqslant 2k \cdot \varepsilon_k^{-1}$, then $d(x) \geqslant 2k$, hence by Lemma 2.3.19, deleting x decreases the cycle rank of G by at least

$$\frac{d(x) - k + 1}{2} \geqslant \frac{d(x)}{4} \geqslant \frac{\varepsilon_k}{4} r(G). \tag{2.6}$$

Thus, as long as the cycle rank is more than $2k \cdot \varepsilon_k^{-1}$, we can find a vertex whose deletion decreases the cycle rank by a constant multiplicative factor. After logarithmically many steps, we have $\mathrm{fvs}(G) \leqslant r(G) \leqslant 2k \cdot \varepsilon_k^{-1}$. In the end, the feedback vertex set consists of at most $f(11, t, k)$ vertices in X, logarithmically many rich vertices deleted in the induction, and at most $2k \cdot \varepsilon_k^{-1}$ vertices for the final graph. □

We now focus on proving Theorem 2.3.20. Let G be an O_k-free graph with girth at least 11. Consider C a shortest cycle of G, N the neighborhood of C, and $R \stackrel{\text{def}}{=} G - (C \cup N)$ the rest of the graph (see Figure 2.5). Remark that there is

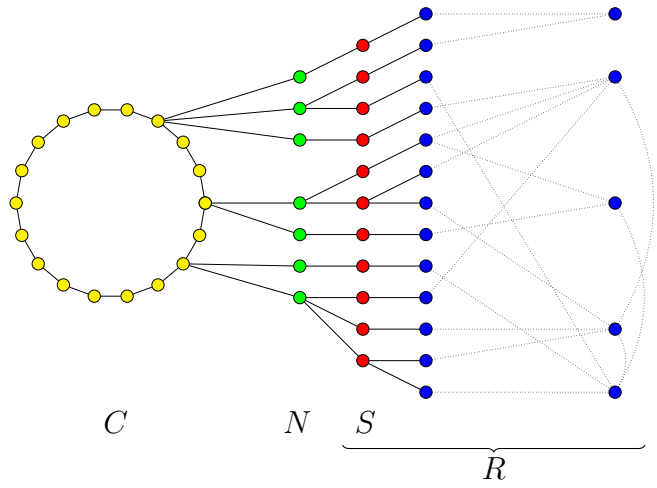


Figure 2.5: Subgraph of an O_4-free graph G. $V(G)$ is partitioned into three sets C, N, R, where C is a shortest cycle, N is an independent set and first neighborhood of C, and R is O_3-free. S is the second neighborhood of N. Gray lines correspond to induced paths where all internal vertices have degree 2.

no edge between C and R, hence $G[R]$ is an O_{k-1}-free graph. As a special case, if $k = 2$, then $G[R]$ is a forest. We will show that in general, it remains possible to reduce the problem to the case where $G[R]$ is a forest, which is our main technical theorem.

Theorem 2.3.22. *For any k, there is some $\delta_k > 0$ such that if G is a connected O_k-free graph with girth at least 11, and furthermore $G[R]$ is a forest where R is as in the decomposition described above, then G has a δ_k-rich vertex.*

Theorem 2.3.22 will be proved in Section 2.3.8. In the remainder of this subsection, we assume Theorem 2.3.22 and explain how Theorem 2.3.20 can be deduced from it.

Proof of Theorem 2.3.20. The proof is by induction on k. Let $\delta_k > 0$ be as in Theorem 2.3.22, and let $\varepsilon_{k-1} > 0$ be as in Theorem 2.3.20, obtained by induction hypothesis. We fix

$$\varepsilon_k \stackrel{\text{def}}{=} \min\left\{\frac{\varepsilon_{k-1}}{20}, \frac{\delta_k}{20}, \frac{\delta_k}{5(k+1)}, \frac{1}{30(k-2)}\right\}. \tag{2.7}$$

Let G be any O_k-free graph with girth at least 11. Reductions preserve all the hypotheses of the claim, and the value of $r(G)$, hence we can assume G to be reduced. Consider the decomposition C, N, R as previously described. We construct a subset $F \subset R$ inducing a rooted forest in G such that the only edges

from F to $R \setminus F$ are incident to roots of F, and each root of F is incident to at most one such edge.

Claim 2.3.23. *If $F \subset R$ has the former property and $F' \subset R \setminus F$ induces a forest in G, then $F \cup F'$ induces a forest in G.*

Proof. Each connected component of $G[F]$ has a single root, which is the only vertex which can be connected to F'. ⌟

We construct F inductively, starting with an empty forest, and applying the following rule as long as it applies: if $x \in R \setminus F$ is adjacent to at most one vertex in $R \setminus F$, we add x to F, and make it the new root of its connected component in F. The condition on F obviously holds for $F = \varnothing$. When adding x, by Claim 2.3.23, $F \cup \{x\}$ is still a forest. Furthermore, if $y \in F \cup \{x\}$ is adjacent to $R \setminus (F \cup \{x\})$, then either $y = x$ or y was a root before the addition of x, and is not adjacent to x, and therefore x and y are in distinct connected components of $F \cup \{x\}$. In either case, y is a root of $F \cup \{x\}$ as required.

We now denote by F the forest obtained when the previous rule no longer applies, and let $R' = R \setminus F$. Remark that we may have as edge cases that $F = R$, meaning that $G[R]$ is a forest (and we fall in the case of Theorem 2.3.22), or inversely $F = \varnothing$, which means that $G[R]$ has minimum degree at least 2.

Claim 2.3.24. *All vertices in $G[R']$ have degree at least 2.*

Proof. A vertex of degree less than 2 in $G[R']$ should have been added to F. ⌟

Claim 2.3.25. *The graph $G[C \cup N \cup F]$ is connected.*

Proof. It suffices to show that each connected component T of $G[F]$ is connected to N. Each such component T is a tree. If T consists of a single vertex v, then v is the root of T and has at most one neighbor in R' by definition. Since G is reduced, v has degree at least 2 in G, hence it must be connected to N.

If T contains at least two vertices, then it contains at least 2 leaves, and in particular at least one leaf v which is not the root of T. The vertex v has a single neighbor in R (its parent in T), and thus similarly as above it must have a neighbor in N. ⌟

Define B as the set of vertices of R' adjacent to $N \cup F$, and let A be the set of edges between $N \cup F$ and B.

Claim 2.3.26. *If $|A| \leqslant \frac{9}{10} r(G)$, then G has an ε_k-rich vertex.*

Proof. Deleting A from G decreases the cycle rank by at most $|A|$, hence $r(G-A) \geqslant r(G)/10$. Since $G[C \cup N \cup F]$ and $G[R']$ are unions of connected components of $G - A$, we have

$$r(G - A) = r(G[C \cup N \cup F]) + r(G[R']).$$

Thus either $G[C \cup N \cup F]$ or $G[R']$ has cycle rank at least $r(G)/20$. If it is $G[C \cup N \cup F]$, then we can apply Theorem 2.3.22 to find a $(\delta_k/20)$-rich vertex, and if it is $G[R']$, then we can apply the induction hypothesis to find an $(\varepsilon_{k-1}/20)$-rich vertex. In either case, this gives an ε_k-rich vertex. ⌟

Thus we can now assume that $|A| \geqslant \frac{9}{10} r(G)$.

Let B_1, resp. B_2, be the set of vertices of B incident to exactly one, resp. at least two edges of A, and let $A_1, A_2 \subseteq A$ be the set of edges of A incident to B_1, B_2 respectively. Remark that A_1, A_2 and B_1, B_2 partition A and B respectively, and $|A_1| = |B_1|$.

Claim 2.3.27. *If $|A_2| \geqslant \frac{4}{9} |A|$, then G has an ε_k-rich vertex.*

Proof. Assume that $|A_2| \geqslant \frac{4}{9} |A|$, and thus $|A_2| \geqslant \frac{2}{5} r(G)$.

By Lemma 2.3.13, G is $2k$-degenerate, hence it can be vertex-partitioned into $k+1$ forests. Consider this partition restricted to B_2, and choose $B_3 \subseteq B_2$ which induces a forest and maximizes the set A_3 of edges incident to B_3. Thus $|A_3| \geqslant |A_2| /(k+1) \geqslant \frac{2}{5(k+1)} r(G)$. By Claim 2.3.23, $F \cup B_3$ is a forest, hence Theorem 2.3.22 applies to $G[C \cup N \cup F \cup B_3]$.

By Claim 2.3.25, $G[C \cup N \cup F]$ is connected, thus adding the vertices B_3 and the edges A_3 increases the cycle rank by $|A_3| - |B_3|$. This quantity is at least $|A_3| /2$ since any vertex of B_3 is incident to at least two edges of A_3, and each edge of A_3 is incident to exactly one vertex of B_3. Thus Theorem 2.3.22 yields a vertex of degree at least

$$\delta_k \cdot r(G[C \cup N \cup F \cup B_3]) \geqslant \frac{|A_3|}{2} \delta_k \geqslant \frac{1}{5(k+1)} \delta_k \cdot r(G) \geqslant \varepsilon_k \cdot r(G) \tag{2.8}$$

as desired. ⌟

Thus we can now assume that $|A_1| \geqslant \frac{5}{9} |A|$, and thus $|B| \geqslant \frac{5}{9} |A| \geqslant \frac{1}{2} r(G)$.

Let X, resp. Y, be the set of vertices of B with degree at least 3, resp. exactly 2, in $G[R']$. By Claim 2.3.24, this is a partition of B.

Claim 2.3.28. *If $|X| \geqslant |B| /5$, then G has an ε_k-rich vertex.*

Proof. Assume that $|X| \geqslant |B|/5$, and thus $|X| \geqslant \frac{1}{10} r(G)$.

The cycle rank is lowerbounded by the following sum:

$$r(G[R']) \geqslant |E(G[R'])| - |R'| = \frac{1}{2} \sum_{x \in R'} (d_{G[R']}(x) - 2). \tag{2.9}$$

By Claim 2.3.24, every term in the sum is non-negative, and each $x \in X$ contributes by at least $1/2$ to the sum. Thus $r(G[R']) \geqslant |X|/2 \geqslant \frac{1}{20} r(G)$, and the induction hypothesis applied to $G[R']$ (which is O_{k-1}-free) yields an $(\varepsilon_{k-1}/20)$-rich vertex, which is also ε_k-rich. ⌟

Thus we can now assume that $|Y| \geqslant \frac{4}{5} |B| \geqslant \frac{2}{5} r(G)$.

Let Z be the set of vertices of R' that either are in Y or have degree at least 3 in $G[R']$. Remark that Z is exactly the set of vertices of R' with degree at least 3 in G. In $G[R']$, a *direct path* is a path whose endpoints are in Z, and whose internal vertices are not in Z. In particular, internal vertices of a direct path have degree 2. A direct path need not be induced, as its endpoints may be adjacent. As a degenerate case, we consider a cycle that contains a single vertex of Z to be a direct path whose two endpoints are equal. One can naturally construct a multigraph G_Z with vertex set Z and whose edges correspond to direct paths in $G[R']$. Remark that vertices of Z have the same degree in G_Z and in $G[R']$.

Any $y \in Y$ has two neighbors x_1, x_2 in G_Z. In degenerate cases, it may be that $x_1 = x_2 \neq y$ (multi-edge in G_Z), in which case $G[R']$ contains a banana between y and x_1, or that $x_1 = x_2 = y$ (loop in G_Z), in which case there is a cycle C_y which is a connected component of $G[R']$, and such that y is the only vertex of Z in C_y. We partition Y into Y_i, Y_e as follows: for y, x_1, x_2 as above, if $x_1, x_2 \in Y$, then we place y in Y_i, and otherwise (x_1 or x_2 is in $Z \setminus Y$) we place y in Y_e.

Claim 2.3.29. *If $|Y_e| \geqslant \frac{3}{4} |Y|$, then G has an ε_k-rich vertex.*

Proof. Assume $|Y_e| \geqslant \frac{3}{4} |Y|$, and thus $|Y_e| \geqslant \frac{3}{10} r(G)$.

By definition, any vertex of Y_e is adjacent in G_Z to some vertex of $Z \setminus Y$. Thus, using that $d_{G_Z}(z) = d_{G[R']}(z)$ for any $z \in Z$, we obtain

$$\sum_{z \in Z \setminus Y} d_{G[R']}(z) \geqslant |Y_e| . \tag{2.10}$$

Recall inequality (2.9) on cycle rank:

$$r(G[R']) \geqslant \frac{1}{2} \sum_{x \in R'} (d_{G[R']}(x) - 2). \tag{2.11}$$

By Claim 2.3.24, the terms of this sum are non-negative. Thus, restricting it to $Z \setminus Y$, we have

$$r(G[R']) \geqslant \frac{1}{2} \sum_{z \in Z \setminus Y} (d_{G[R']}(z) - 2). \tag{2.12}$$

By definition of Z, vertices of $Z \setminus Y$ have degree at least 3 in $G[R']$. Thus, each term of the previous sum satisfies $d_{G[R']}(z) - 2 \geqslant d_{G[R']}(z)/3$. It follows using (2.10) that

$$r(G[R']) \geqslant \frac{1}{2} \sum_{z \in Z \setminus Y} \frac{d_{G[R']}(z)}{3} \geqslant \frac{|Y_e|}{6} \geqslant \frac{1}{20} r(G). \tag{2.13}$$

Thus the induction hypothesis applied to $G[R']$ (which is O_{k-1}-free) yields an $(\varepsilon_{k-1}/20)$-rich vertex, which is also ε_k-rich. ⌟

Thus we can now assume that $|Y_i| \geqslant \frac{1}{4} |Y| \geqslant \frac{1}{10} r(G)$.

We now consider the induced subgraph H of $G[R']$ consisting of Y, and direct paths joining vertices of Y. Thus H has maximum degree 2, and since $G[R']$ is O_{k-1}-free, at most $k-2$ components of H are cycles, the rest being paths. Remark that the endpoints of paths in H correspond exactly to Y_e. Also, each connected component of H must contain at least one vertex of Y.

We perform the following cleaning operations, in this order:

- In each cycle of H, pick an arbitrary vertex and delete it, so that all connected components are paths.
- Iteratively delete a vertex of degree 0 or 1 which is not in Y, so that the endpoints of paths are all in Y.
- Delete any isolated vertex.

Let H' be the subgraph of H obtained after these steps.

Claim 2.3.30. *All but* $3(k-2)$ *vertices of* Y_i *are internal vertices of paths of* H'.

Proof. If $y \in Y_i$ belongs to a path of H, then it must be an internal vertex of this path, and the path is unaffected by the cleaning operations. Thus it suffices to prove that in each cycle of H, at most 3 vertices of Y_i are deleted or become endpoints of paths during the clean up.

Let C' be a cycle of H. If C' contains no more than 2 vertices of Y_i, there is nothing to prove. Remark in this case that C' is entirely deleted by the clean up. Otherwise, let x be the vertex deleted from H (which may be in Y_i), and let y_1, y_2 be the first vertices of Y_i strictly before and after x in the cyclic order of C'. Since C' has at least 3 vertices of Y_i, x, y_1, y_2 are all distinct. Then, it is clear that the cleaning operations transform C' into a path with endpoints y_1, y_2, such that any $y \in Y_i \cap C'$ distinct from x, y_1, y_2 is an internal vertex of this path. ⌟

We now add H' to F, which yields a forest by Claim 2.3.23. Recall that vertices of Y are adjacent to $N \cup F$, and all endpoints of paths of H' are in Y. Thus, in $G[C \cup N \cup F \cup H']$, every vertex of H' has degree at least 2, and vertices of Y_i in the interior of paths of H' have degree at least 3. Since $G[C \cup N \cup F]$ is connected by Claim 2.3.25, the addition of H' does not change the number of connected components. Using Claim 2.3.30, this implies that

$$r(G[C \cup N \cup F \cup H']) \geqslant |Y_i| - 3(k-2). \tag{2.14}$$

We finally apply Theorem 2.3.22 to $G[C \cup N \cup F \cup H']$ to obtain a vertex with degree at least

$$\delta_k \cdot (|Y_i| - 3(k-2)).$$

Since G contains vertices of degree at least 2, we can always assume that $\varepsilon_k \cdot r(G) \geqslant 2$, and thus

$$|Y_i| \geqslant \frac{1}{10} \cdot 2\varepsilon_k^{-1} \geqslant \frac{1}{5} \cdot 30(k-2) = 6(k-2). \tag{2.15}$$

It follows that $|Y_i| - 3(k-2) \geqslant |Y_i|/2$, and the previous argument yields a vertex of degree at least $\frac{\delta_k}{2}|Y_i| \geqslant \frac{\delta_k}{20} r(G)$, which is an ε_k-rich vertex. □

2.3.6 Cutting trees into stars and paths

The setting in all this subsection is the following. We consider a graph H which is the disjoint union of a forest F and an independent set N, such that the neighborhood of N in F is a subset $S \subseteq V(F)$ containing all the leaves of F, and each vertex of S has a unique neighbor in N. Moreover H is O_k-free (and it can be deduced from the definition that H is 3-biclique-free).

A path is a special case of subdivided star. The *center* of a subdivided star is the vertex of degree at least 3, if any. If none, the subdivided star is a path, and its *center* is a vertex of degree 2 that belongs to S, if any, and an arbitrary vertex otherwise. We say a forest $F' \subseteq F$ is *S'-clean*, for some $S' \subseteq S$, if $V(F') \cap S' = L(F')$, where $L(F')$ denotes the set of leaves of F'. We define being *quasi-S'-clean* for a subdivided star as intersecting S' at exactly its set of leaves, plus possibly its center. Formally, a subdivided star T is *quasi-S'-clean* if $L(T) \subseteq V(T) \cap S' \subseteq L(T) \cup \{c\}$ where c is the center of T. The *degree* of a subdivided star is the degree of its center. A forest $F' \subseteq F$ of subdivided stars is *quasi-S'-clean*, for some $S' \subseteq S$, if all its connected components are quasi-S'-clean (subdivided stars).

Our approach in this subsection is summarized in Figure 2.6. We start with an arbitrary forest F and a subset S of vertices including all the leaves of F (the vertices of S are depicted in white, while the vertices of $F - S$ are depicted in black). We first extract quasi-S-clean subdivided stars (Lemma 2.3.31). We then extract

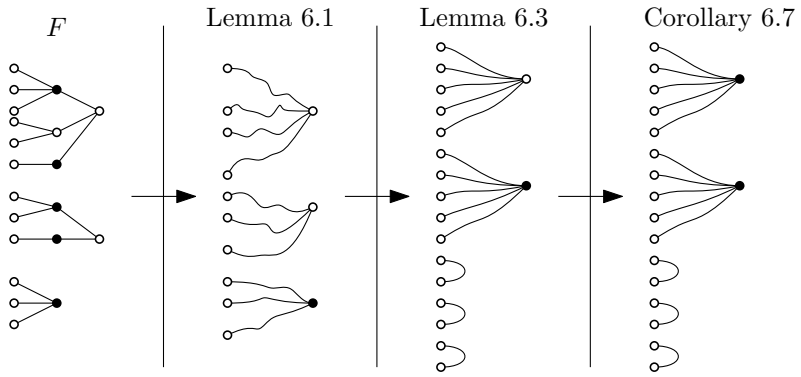


Figure 2.6: A visual summary of subsection 2.3.6.

quasi-S-clean subdivided stars of large degree, or S-clean paths (Lemma 2.3.33). Finally we extract S-clean subdivided stars of large degree or paths (Corollary 2.3.37). At each step the number of vertices of S involved in the induced subgraph of F we consider is linear in $|S|$.

Lemma 2.3.31. *There is a subset $F^* \subseteq F$ containing at least $\frac{1}{2}|S|$ vertices of S such that each connected component of $H[F^*]$ is a quasi-S-clean subdivided star.*

Proof. We first use the following claim.

Claim 2.3.32. *There is a set of edges $X \subseteq E(H)$ such that every connected component of $H[F] \setminus X$ is either a quasi-S-clean subdivided star or a single vertex that does not belong to S.*

Proof. We proceed greedily, starting with $X = \varnothing$. While $H[F] \setminus X$ contains a component T and an edge $e \in T$ such that the two components of $T - e$ each contains either no vertex of S or at least two vertices of S, we add e to X.

Observe that in $H[F]$, every connected component with at least one edge contains at least 2 vertices of S. Throughout the process of defining X, every connected component of $H[F] \setminus X$ with at least one edge contains either 0 or at least 2 vertices of S.

At the end of the process, for any connected component T of $H[F] \setminus X$ with at least one edge, all the leaves of T belong to S. Otherwise, the edge incident to the leaf of T that is not in S can be added to X.

Thus, $H[F] \setminus X$ does not contain any component with more than one vertex of degree at least 3, since otherwise any edge on the path between these two vertices would have been added to X, yielding two components containing at least 2 leaves, and thus at least 2 vertices of S.

Observe also that if $H[F] \setminus X$ contains a component T with a vertex $v \in S$ that has degree 2 in T, then T is a path containing exactly 3 vertices of S, and thus

T is a subdivided star whose center and leaves are in S, and whose other internal vertices are not in S. ⌟

To conclude, we need to select connected components of $H[F] \setminus X$ with more than one vertex and that are pairwise non-adjacent in H. Consider the minor G_F of F obtained by contracting each connected component of $H[F] \setminus X$ into a single vertex and deleting those that are a single vertex not in S. Since F is a forest, the graph G_F is a forest. We weigh each vertex of G_F by the number of elements of S that the corresponding connected component of $H[F] \setminus X$ contains. Since G_F is a forest, there is an independent set $\{u_1, u_2, \ldots, u_p\}$ that contains at least half the total weight. The connected components corresponding to $u_1, u_2, \ldots, u_p$ together form a forest F^* with the required properties. □

We observe that subdivided stars of small degree can be transformed into paths for a low price, as follows. A *subdivided star forest* is a forest whose components are subdivided stars (possibly paths).

Lemma 2.3.33. *For every $S' \subseteq S$, quasi-S'-clean subdivided star forest $F' \subseteq F$, and integer $D \geqslant 2$, there is a subdivided star forest $F'' \subseteq F'$ such that every connected component of $H[F'']$ is either an S'-clean path or a quasi-S'-clean subdivided star of degree at least D. Additionally, F'' contains at least $\frac{2|S' \cap F'|}{D}$ vertices of S'.*

Proof. We define F'' from F' as follows. Consider a connected component T of $H[F']$. If the center of T has degree at least D, we add T to F''. Consider now the case where T is a quasi-S'-clean subdivided star whose center c has degree less than D. If $c \in S'$, we select a non-edgeless path $P \subseteq T$ between c and S', and add P to F''. If $c \notin S'$, we select two internally-disjoint paths $P_1, P_2 \subseteq T$ between c and S', and add $P_1 \cup P_2$ to F''. Note that $P_1 \cup P_2$ yields an S'-clean path.

To see that F'' contains at least $\frac{2|S' \cap F'|}{D}$ vertices of S', we simply observe that out of a maximum of $(D-1)+1$ vertices of S' in a component T, we keep at least 2 in F'. This adds up to $\frac{2|S'|}{D}$ vertices of S' since connected components of $H[F']$ are disjoint by definition. □

Lemma 2.3.34. *If each vertex of N has degree less than $\frac{1}{8k}|S|$, one of the following holds.*

- *there is a subset S' of S and a subset F_2 of F such that F_2 contains $\frac{1}{32}|S|$ vertices of S', and each connected component of $H[F_2]$ is an S'-clean subdivided star.*
- *there is a subset F_3 of F such that every connected component of F_3 is a quasi-S-clean subdivided star of degree at most 4 and F_3 contains at least $\frac{1}{8}|S|$ vertices of S.*

Proof. Let F^* be the forest obtained from Lemma 2.3.31, which contains at least $\frac{1}{2}|S|$ vertices of S, and such that each component of $H[F^*]$ is a quasi-S-clean subdivided star or an S-clean path. We define the *label* of a vertex of S to be its only neighbor in N.

Claim 2.3.35. *There is a subset F_1 of F^* containing at least $\frac{1}{4}|S|$ vertices of S, such that no subdivided star of F_1 has its center and one of its endpoints sharing the same label.*

Proof. Let ℓ be the maximum integer such that there exist ℓ subdivided stars $S_1, S_2, \ldots, S_\ell$ in $H[F^*]$ and ℓ different labels $v_1, \ldots, v_\ell \in N$, such that for any $1 \leqslant i \leqslant \ell$, S_i has its center and at least one of its endpoint labeled v_i. Note that in this case G contains ℓ independent cycles, and thus $\ell < k$ by assumption.

For any $1 \leqslant i \leqslant \ell$, remove all the leaves u of F^* that are labeled v_i, and also remove the maximal path of $H[F^*]$ ending in u. By assumption, there are at most $\frac{1}{8k}|S|$ such vertices u for each $1 \leqslant i \leqslant \ell$, and thus we delete at most $k \cdot \frac{1}{8k}|S| \leqslant \frac{1}{8}|S|$ vertices of S from F^*. We also delete the centers that have no leaves left (there are at most $k \cdot \frac{1}{8k}|S| \leqslant \frac{1}{8}|S|$ such deleted centers). Let F_1 be the resulting subset of F^*. Note that F_1 contains at least $|F^* \cap S| - 2 \cdot \frac{1}{8}|S| \geqslant (\frac{1}{2} - \frac{1}{4})|S| = \frac{1}{4}|S|$ vertices of S. ⌟

We can assume that a subset Y of at least $\frac{1}{8}|S|$ vertices of S in the forest F_1 obtained from Claim 2.3.35 are involved in a quasi-S-clean subdivided star of degree at least 5. Indeed, otherwise at least $\frac{1}{8}|S|$ vertices of S in the forest F_1 obtained from Claim 2.3.35 are involved in a quasi-S-clean subdivided star of degree at most 4 (note that an S-clean path is an S-clean subdivided star), and in this case the second outcome of Lemma 2.3.34 holds.

For each label $v \in N$, we choose uniformly at random with probability $\frac{1}{2}$ whether v is a *center label* or a *leaf label*. We then delete all the subdivided stars of F_1 whose center is labeled with a leaf label, and all the leaves whose label is a center label. Moreover, we delete from N all the vertices that are a center label, and let S' be the set of vertices of S whose neighbor in N is not deleted.

Take a vertex u of Y. If u is a center of a subdivided star, then the probability that u is not deleted is at least $\frac{1}{2}$. If u is a leaf, u is kept only if u and the center of the subdivided star it belongs to (which has by construction a different label) are correctly labeled, so u is kept with probability at least $\frac{1}{4}$. Overall, each vertex u of Y has probability at least $\frac{1}{4}$ to be kept. Thus the expectation of the fraction of vertices of Y not deleted is at least $\frac{1}{4}$, thus we can find an assignment of the labels to leaf labels or center labels, such that a subset $Z \subseteq Y$ with $|Z| \geqslant \frac{1}{4}|Y|$ survives.

We then iteratively delete vertices of degree 1 that do not belong to S' and all vertices of degree 0. Let F_2 be the resulting forest. Note that S' contains only the endpoints of stars with a leaf label, thus the forest F_2 is S'-clean. It

remains to argue that F_2 contains a significant fraction of vertices of S. Note that a connected component of F_1 is deleted if and only if it contains at most one element of Z. Every such component has at least 4 elements in $Y \setminus Z$, hence there are at most $\frac{1}{4} \cdot \frac{3}{4}|Y| = \frac{3}{16}|Y|$ such components. It follows that F_2 contains at least $|Z| - \frac{3}{16}|Y| \geqslant \frac{1}{4}|Y| - \frac{3}{16}|Y| \geqslant \frac{1}{16}|S|$ elements of $Z \subseteq S$. □

We now have all the ingredients to obtain the following two corollaries.

Corollary 2.3.36. *For any $D \geqslant 2$, there is a subset $F^* \subseteq F$ containing at least $\frac{1}{2D}|S|$ vertices of S such that each*

1. *F^* induces a quasi-S-clean subdivided star forest whose components all have degree at least D, or*
2. *F^* induces an S-clean path forest.*

Corollary 2.3.36 follows from Lemma 2.3.31 by applying Lemma 2.3.33 and observing that one of the two outcomes contains half the corresponding vertices in S.

Corollary 2.3.37. *Let $D \geqslant 2$. If each vertex of N has degree less than $\frac{1}{8k}|S|$, then there are $F'' \subseteq F$, $S' \subseteq S$ such that F'' contains at least $\frac{1}{32D}|S|$ vertices of S' and one of the following two cases apply.*

1. *F'' induces an S'-clean subdivided star forest whose components all have degree at least D, or*
2. *F'' induces an S'-clean path forest.*

Similarly, Corollary 2.3.37 follows from Lemma 2.3.34 by applying Lemma 2.3.33 and observing that one of the two outcomes contains half the corresponding vertices in S.

2.3.7 Trees, stars, and paths

In the proof of Theorem 2.3.22, we will apply Corollaries 2.3.36 and 2.3.37 several times, and divide our graph into two parts: a union of subdivided stars on one side, and a union of subdivided stars or paths on the other side. We now explain how to find a rich vertex in this context.

We start with the case where subdivided stars appear on both sides.

Lemma 2.3.38 (Star-star lemma)**.** *Let $c > 0$ be the constant of Lemma 2.3.17. Let H be an O_k-free graph whose vertex set is the union of two sets L, R, such that*

- *$S = L \cap R$ is an independent set,*

- *there are no edges between $L \setminus S$ and $R \setminus S$, and*
- *L (resp. R) induces in H a disjoint union of subdivided stars, whose centers have average degree at least $3ck \log k$, and whose set of leaves is precisely S.*

Then H contains a vertex of degree at least $\frac{1}{2f'(3,k)}|S| = \Omega(\frac{1}{k^3}|S|)$, where f' is the function of Lemma 2.3.12.

Proof. Note that H is 3-biclique-free (but might contain some $K_{2,2}$ subgraph) and O_k-free. By Lemma 2.3.12, there is a set X of at most $f'(3,k)$ vertices of H such that all bananas of H intersect X. Since the centers of the subdivided stars are the only vertices of degree larger than 2 in H, we can assume that X is a subset of the centers of the subdivided stars.

Assume first that less than $\frac{1}{2}|S|$ vertices of S are leaves of subdivided stars centered in an element of X. Let $S' \subseteq S$ be the leaves of the subdivided stars whose center is not in X (note that $|S'| \geqslant \frac{1}{2}|S|$), and remove from the subdivided stars of $H[L]$ and $H[R]$ all branches whose endpoint is not in S' to get new sets of vertices L', R'. The centers of the resulting S'-clean subdivided stars now have average degree at least $\frac{1}{2} \cdot 3ck \log k > ck \log k$. We denote the resulting S'-clean subdivided stars of $H[L']$ by S_1, S_2, etc. and their centers by s_1, s_2, etc. Similarly, we denote the resulting S'-clean subdivided stars of $H[R']$ by S'_1, S'_2, etc. and their centers by s'_1, s'_2, etc. Observe that by the definition of X, for any two centers s_i, s'_j, there is at most one vertex $u \in S'$ which is a common leaf of S_i and S'_j.

Let B be the bipartite graph with partite set $s_1, s_2, \ldots$ and $s'_1, s'_2, \ldots$, with an edge between s_i and s'_j if and only if some vertex of S' is a common leaf of S_i and S'_j. Note that B has average degree more than $ck \log k$, and some induced subgraph of H (which is O_k-free) contains a strict subdivision of B. This contradicts Lemma 2.3.17.

So we can assume that at least $\frac{1}{2}|S|$ vertices of S are leaves of subdivided stars centered in an element of X. Then some vertex of X has degree at least $\frac{1}{2f'(3,k)}|S|$, as desired.

□

We now consider the case where subdivided stars appear on one side, and paths on the other.

Lemma 2.3.39 (Star-path lemma)**.** *Let $c > 0$ be the constant of Lemma 2.3.17. Let H be an O_k-free graph whose vertex set is the union of two sets L, R, such that*

- *$S = L \cap R$ is an independent set,*
- *there are no edges between $L \setminus S$ and $R \setminus S$,*

- *L induces in H a disjoint union of paths, whose set of endpoints is precisely S, and*
- *R induces in H a disjoint union of subdivided stars, whose centers have average degree at least $4ck\log k$, and whose set of leaves is precisely S.*

Then H contains a vertex of degree at least $\frac{1}{3f'(2,k)}|S| = \Omega(\frac{1}{k^2}|S|)$, where f' is the function of Lemma 2.3.12.

Proof. Note that H is 2-biclique-free and O_k-free. By Lemma 2.3.12, there is a set X of at most $f'(2,k)$ vertices of H such that all bananas of H intersect X. Since the centers of the subdivided stars are the only vertices of degree more than 2 in H, we can assume that X is a subset of the centers of the subdivided stars.

Assume first that less than $\frac{1}{3}|S|$ vertices of S are leaves of subdivided stars centered in an element of X. Then there are at least $\frac{1}{6}|S|$ paths in $H[L]$ whose endpoints are not leaves of stars centered in X. Let $S' \subseteq S$ be the endpoints of these paths (note that $|S'| \geqslant \frac{1}{3}|S|$), and remove from the subdivided stars of $H[R]$ all branches whose endpoint is not in S' to get R'. The centers of the resulting S'-clean subdivided stars in $H[R']$ now have average degree at least $\frac{1}{3}\cdot 4ck\log k > ck\log k$. We denote these subdivided stars by $S_1,\dots,S_t$, and their centers by $s_1,\dots,s_t$.

Given two centers s_i, s_j, we say that a pair $u_i, u_j \in S'$ is an *$\{i,j\}$-route* if u_i is a leaf of S_i, u_j is a leaf of S_j, and there is a path with endpoints u_i, u_j in $H[L]$. Observe that by the definition of X, for every pair s_i, s_j, there is at most one $\{i,j\}$-route.

Let G be the graph with vertex set $s_1,\dots,s_t$, with an edge between s_i and s_j if and only if there is an $\{i,j\}$-route. Note that G has average degree more than $ck\log k$, and some induced subgraph of H (which is O_k-free) contains a strict subdivision of G. This contradicts Lemma 2.3.17.

So we can assume that at least $\frac{1}{3}|S|$ vertices of S are leaves of subdivided stars centered in an element of X. Then some vertex of X has degree at least $\frac{1}{3f'(2,k)}|S|$, as desired. □

From the two previous lemmas and Lemma 2.3.31 we deduce the following.

Lemma 2.3.40 (Star-tree lemma)**.** *There is a constant $c > 0$ such that the following holds. Let H be an O_k-free and t-biclique-free graph whose vertex set is the union of two sets L, R, such that*

- *$S = L \cap R$ is an independent set partitioned into S_P, S_T,*
- *there are no edges between $L \setminus S$ and $R \setminus S$,*

- *L induces in H a disjoint union of subdivided stars, whose centers have average degree at least $(8ck\log k)^2$, and whose set of leaves is equal to S, and*
- *R induces in H the disjoint union of*
  - *paths on a vertex set R_P, whose set of endpoints is equal to S_P, and*
  - *a tree T on a vertex set R_T such that S_T is a subset of leaves of T.*

Then H contains a vertex of degree at least $\Omega(\frac{1}{k^4\log k}|S|)$.

Proof. Let $c > 0$ be the constant of Lemma 2.3.17. Assume first that $|S_T| \leqslant 1$. Then since the subdivided stars of L have average degree at least $(8ck\log k)^2$, we have $|S_P| = |S| - |S_T| \geqslant (8ck\log k)^2 - 1 \geqslant 1$ and thus $|S_P| \geqslant \frac{1}{2}|S|$. By removing the branch of a subdivided star of L that has an endpoint in S_T (if any), we obtain a set of S_P-clean subdivided stars of average degree at least $\frac{1}{2}\cdot(8ck\log k)^2 \geqslant 4ck\log k$. By Lemma 2.3.39, we get a vertex of degree at least $\Omega(\frac{1}{k^2}|S_P|) = \Omega(\frac{1}{k^2}(|S|))$, as desired. So in the remainder we can assume that $|S_T| \geqslant 2$.

Let T' be the subtree of T obtained by repeatedly removing leaves that are not in S_T. Since $|S_T| \geqslant 2$, $L(T') = S_T$. Observe that $F' = T' \cup R_P$ is an S-clean forest (with $L(F') = S$), thus any S-quasi-clean subforest of F' is S-clean. It follows from Corollary 2.3.36 (applied to S, F', and $D = 4ck\log k$) that F' contains a subset F^* containing at least $\frac{1}{2\cdot 4ck\log k}|S|$ vertices of S, such that $H[F^*]$ induces either (1) an S-clean forest of path, or (2) an S-clean forest of subdivided stars of degree at least $4ck\log k$.

We denote this intersubsection of S and F^* by S^*, and we remove in the subdivided stars of $H[L]$ all branches whose endpoint is not in S^* to get a new set of vertices $L^* \subset L$. By assumption, the average degree of the subdivided stars in L^* is at least $\frac{(8ck\log k)^2}{8ck\log k} = 8ck\log k \geqslant 4ck\log k$.

In case (1) above we can now apply Lemma 2.3.39, and in case (2) we can apply Lemma 2.3.38. In both cases we obtain a vertex of degree at least $\Omega(\frac{1}{k^3}|S^*|) = \Omega(\frac{1}{k^4\log k}|S|)$, as desired. □

2.3.8 Proof of Theorem 2.3.22

We start by recalling the setting of Theorem 2.3.22. The graph G is a connected O_k-free graph of girth at least 11, and C is a shortest cycle in G. The neighborhood of C is denoted by N, and the vertex set $V(G) \setminus (C \cup N)$ is denoted by R. The subset of R consisting of the vertices adjacent to N is denoted by S. Since C is a shortest cycle, of size at least 11, each vertex of S has a unique neighbor in N, and a unique vertex at distance 2 in C. Moreover N and S are independent sets. In the setting of Theorem 2.3.22, R is a forest.

Our goal is to prove that there is a vertex whose degree is linear in the cycle rank $r(G)$. To this end, we assume that G has maximum degree at most $\delta \cdot r(G)$, for some $\delta > 0$, and prove that this yields a contradiction if δ is a small enough function of k.

By Lemma 2.3.18, we can assume that G is reduced, i.e., contains no vertex of degree 0 or 1. If G consists only of the cycle C, then $r(G) = 1$ and the theorem is immediate. Thus we can assume that N is non-empty, which in turn implies that S is non-empty since G is reduced.

Using that G is connected, remark that

$$r(G) = |E(G)| - |V(G)| + 1 = 1 + \frac{1}{2} \sum_{v \in V(G)} (d(v) - 2) \tag{2.16}$$

We start by proving that the cardinality of S is at least the cycle rank $r(G)$.

Claim 2.3.41. *$|S| \geqslant r(G)$, and thus G has maximum degree at most $\delta|S|$.*

Proof. Observe that $\frac{1}{2}\sum_{v \in C \cup N}(d(v) - 2) = \frac{1}{2}|S|$. Furthermore $\frac{1}{2}\sum_{v \in R}(d(v) - 2)$ is equal to $\frac{1}{2}|S|$ minus the number of connected components of $G[R]$, as R induces a forest and each vertex of S has a unique neighbor outside of R. Since R is non-empty, it follows from (2.16) that $r(G) \leqslant |S|$. We assumed that G has maximum degree at most $\delta \cdot r(G)$ which is at most $\delta|S|$, as desired. ⌟

In the remainder of the proof, we let $c > 0$ be a sufficiently large constant such that Lemmas 2.3.16 and 2.3.40 both hold for this constant.

We consider $\delta < \frac{1}{8k}$, and use Claim 2.3.41 to apply Corollary 2.3.37 to the subgraph H of G induced by N and $F = R$ (which is O_k-free), with $D = 2 \cdot (8ck \log k)^2$. We obtain subsets $N' \subseteq N$, $R'' \subseteq R$ such that if we define S' as the subset of $S \cap R''$ with a neighbor in N', we have $|S'| \geqslant \frac{1}{32D}|S|$ and at least one of the following two cases apply.

1. Each connected component of $H[R'']$ is an S'-clean subdivided star of degree at least D, or
2. Each connected component of $H[R'']$ is an S'-clean path.

We first argue that the second scenario holds.

Claim 2.3.42. *Each connected component of $H[R'']$ is an S'-clean path.*

Proof. Assume for a contradiction that Case 2 does not apply, hence Case 1 applies.

Let G_1 be the subgraph of G induced by $C \cup N' \cup R''$ (see Figure 2.7, left). Since $|C| \geqslant 11$ and vertices of C have disjoint second neighborhoods in S', there exists

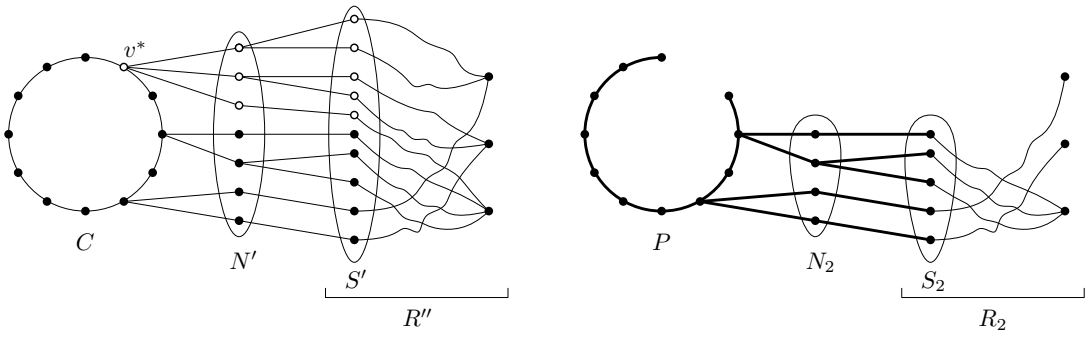


Figure 2.7: The graphs G_1 (left) and G_2 (right) in the proof of Claim 2.3.42.

a vertex $v^* \in C$ that sees at most $\frac{1}{11}|S'|$ vertices of S' in its second neighborhood. If we remove from G_1 the vertex v^*, its neighborhood $N(v^*) \subseteq N'$, its second neighborhood $N^2(v^*) \subseteq S'$, and the corresponding branches of the subdivided stars of R'', we obtain a graph G_2 whose vertex set is partitioned into a path $P = C - v^*$, its neighborhood $N_2 = N' - N(v^*)$, and the rest of the vertices R_2 (which includes the set $S_2 = S' - N^2(v^*)$), with the property that each component of $G_2[R_2]$ is an S_2-clean subdivided star (see Figure 2.7, right). More importantly,

$$|S_2| \geqslant \tfrac{10}{11}|S'| \geqslant \tfrac{10}{11} \cdot \tfrac{1}{32D}|S| \geqslant \tfrac{1}{36D}|S|,$$

and the average degree of the centers of the subdivided stars is at least $\frac{10}{11}D \geqslant (8ck \log k)^2$.

Observe that $P \cup N_2 \cup S_2$ induces a tree in G_2, such that all leaves of $G_2[P \cup N_2 \cup S_2]$ except at most two (the two neighbors of v^* on C) lie in S_2, and non leaves of the tree are not in S_2. We can now apply Lemma 2.3.40 with $R = P \cup N_2 \cup S_2$ and $L = R_2$. It follows that G_2 contains a vertex of degree at least $\Omega(\frac{1}{k^4 \log k}|S_2|) = \Omega(\frac{1}{k^6 \log^3 k}|S|) > \delta|S|$. Since G_2 is an induced subgraph of G, this contradicts Claim 2.3.41. ⌟

We denote the connected components of $H[R'']$ by $P_1, \ldots, P_\ell$, with $\ell \geqslant \frac{1}{64D}|S|$.

Claim 2.3.43. *There is a vertex u^* in C which has at least $\frac{1}{16(8ck \log k)^3}|S|$ endpoints of the paths $P_1, \ldots, P_\ell$ in its second neighborhood, where $c > 0$ is the constant of Lemma 2.3.16.*

Proof. Assume for the sake of contradiction that each vertex of C has less than $\frac{1}{16(8ck \log k)^3}|S|$ endpoints of the paths $P_1, \ldots, P_\ell$ in its second neighborhood.

Let G_3 be subgraph of G induced by $C \cup N'$ and $\bigcup_{i=1}^{\ell} V(P_i)$ (see Figure 2.8, left), and let G_4 be the graph obtained from G_3 by contracting each vertex of

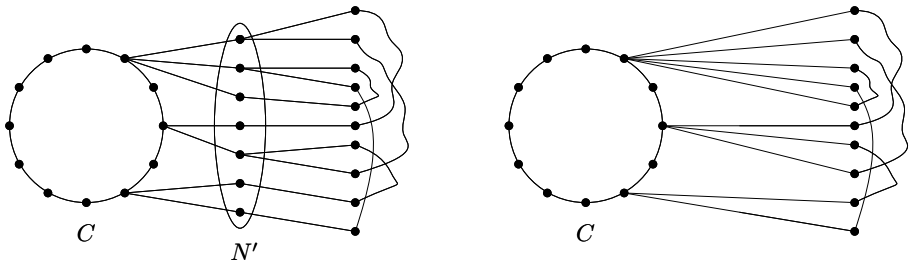


Figure 2.8: The graphs G_3 (left) and G_4 (right) in the proof of Claim 2.3.43.

N' with its unique neighbor in C (i.e., G_4 is obtained from G_3 by contracting disjoint stars into single vertices), see Figure 2.8, right. Note that since G is O_k-free, G_3 and G_4 are also O_k-free (from the structural properties of C, N, and S, each cycle in G_4 can be canonically associated to a cycle in G_3, and for any set of independent cycles in G_4, the corresponding cycles in G_3 are also independent). By our assumption, each vertex of C in G_4 has degree at most $\frac{1}{16(8ck\log k)^3}|S|+2$, and G_4 consists of the cycle C together with $\ell \geqslant \frac{1}{64D}|S|$ paths whose endpoints are in C and whose internal vertices are pairwise disjoint and non-adjacent. By Lemma 2.3.16, it follows that

$$\tfrac{1}{64D}|S| < \ell \leqslant c \cdot \tfrac{1}{16(8ck\log k)^3}|S| \cdot k\log k,$$

and thus $D > 2(8ck\log k)^2$, which contradicts the definition of $D = 2(8ck\log k)^2$. ⌟

Claim 2.3.44. *If the vertices in $N[u^*]$ have average degree at least $(8ck\log k)^2$ in S', then G contains a vertex of degree at least $\delta|S|$.*

Proof. The key idea of the proof of the claim is to consider the neighbors of u^* as the centers of stars (L) in Claim 2.3.40. In order to do that, we consider the subgraph G_5 of G induced by

- the path $C - u^*$,
- $N(u^*)$ and the paths P_i $(1 \leqslant i \leqslant \ell)$ with at least one endpoint in the second neighborhood $N^2(u^*)$ of u^* (call these paths $P'_1, \ldots, P'_t$), and
- the neighbors of the endpoints of the paths $P'_1, \ldots, P'_t$ in N.

All the components of $G_5 - N(u^*)$ are either paths P'_i with both endpoints in $N^2(u^*)$, or a tree whose leaves are all in $N^2(u^*)$ (except at most two leaves, which are the two neighbors of u^* in C). See Figure 2.9, right, for an illustration, where the vertices of $N^2(u^*)$ are depicted with squares and the components of $G_5 - N(u^*)$ are depicted with bold edges.

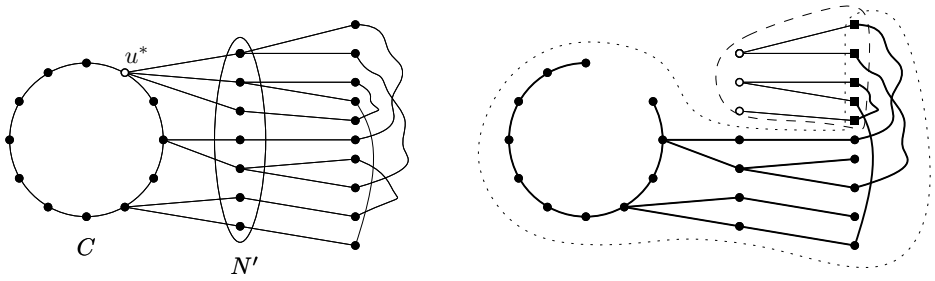


Figure 2.9: The graphs G_3 with the vertex u^* (left) and the graph G_5 (right) in the proof of Claim 2.3.44.

By considering the vertices of $N(u^*)$ and their neighbors in S' as stars (whose centers, depicted in white in Figure 2.9, right, have average degree at least $(8ck\log k)^2$) we can apply Lemma 2.3.40, and obtain a vertex of degree at least $\Omega(\frac{1}{k^4\log k}|S'|) \geqslant \Omega(\frac{1}{k^6\log^3 k}|S|) \geqslant \delta|S|$ in G_5 (and thus in G), which contradicts Claim 2.3.41. ⌟

Observe that if the vertices of $N(u^*)$ have average degree at most $(8ck\log k)^2$ in S', then u^* has degree at least $\frac{1}{16(8ck\log k)^5}|S| \geqslant \delta|S|$. If not, by Claim 2.3.44, G also contains a vertex of degree at least $\delta|S|$. Both cases contradict Claim 2.3.41, and this concludes the proof of Theorem 2.3.22. □

Chapter 3

Conclusion and Perspectives

In this thesis, we studied various structural problems, distributed into two topics: finding small minor-universal graphs and finding large induced trees.

In Chapter 1 we studied the minor-universal graph problem for planar and bounded genus graphs. For planar graphs, since the smallest planar minor-universal graph known is a grid, we compared the area of the smallest grid-major of a planar graph to the area of the smallest grid on which it can be drawn [GH23a]. We showed that is NP-complete to find the smallest grid-major of a planar graph, even given a smallest grid-drawing of the graph. The question of Dieng and Gavoille [DG20] whether a grid-drawing of area A implies a grid-major of area $O(A)$ remains open. However, we show that, when the maximum degree is bounded by a constant, any lower bound for the grid-drawing implies a lower bound for the area of a grid-major. Recall that Robertson, Seymour, and Thomas [RST94] proved that a grid minor-universal for the planar n-vertex graphs has area $O(n^2)$, and the quadratic dependency for the grid area was proved necessary. When restricted to k-outerplanar n-vertex graphs, we show that they are minors of a grid of area $O(kn)$, and that this is asymptotically optimal, generalizing and improving the $O(n \log n)$ previous upper bound for outerplanar graphs [DG20]. Then, we generalized the theorem of Robertson, Seymour, and Thomas to graph of bounded genus, showing that for any surface of genus g, there is a graph on $O(g^2(n+g)^2)$ vertices embedded on the surface and minor-universal for all the n graphs embeddable on this surface [GH23b]. Our construction is inspired from the planar case, hence the quadratic dependency in n. Note that, in contrast to the previous results, our minor-universal graph is not a grid since it cannot be planar for $g > 0$.

This raised the question whether we could find a smaller planar graph minor-universal for planar graphs, if we drop the constraint of being a grid. More precisely,

Question 1.

Is there a planar graph on $o(n^2)$ vertices minor-universal for all planar graphs on at most n vertices?

If such minor-universal exists, it would be interesting to see if the construction can also be generalized to bounded genus graphs.

Moreover, thanks to the Structure Theorem of Robertson and Seymour [RS03], several results on bounded genus graph can be generalized to the classes of graphs excluding a fixed graph as minor. Roughly speaking, any graph excluding some fixed graph has some kind of tree-decomposition where each bag induces a graph "almost embeddable" on a surface of bounded genus. Therefore, it would be very interesting to check if our construction can be generalized to these "almost embeddable" graphs first, and then maybe to those minor-free graphs. This leads to the following question:

Question 2.
For every fixed graph H, is there a H-minor free graph on $n^{O(1)}$ vertices minor-universal for all H-minor free graphs on at most n vertices?

In Chapter 2, we present our result related to tree structures in graphs and more precisely to induced forests in graphs. Note that in this chapter, we worked on both tree decomposition and induced trees, so an interesting related question, asked by Dvořák, was the following: can the tree of a tree decomposition appear (for some containment relation) inside the graph, without having to increase the width much? We answered in the negative [BCH+23] even for the minor relation, for which we need to increase the width of the decomposition by a function of n.

In the first part, we studied the case where the forest is a path and focused on a conjecture of Esperet et al [ELM17]. Is it true that k-degenerate graphs with a path of size n as subgraph also have a path of size $(\log n)^{\Omega(1)}$ as induced subgraph? We proved that in the case of graphs of pathwidth less than k, there is always an induced path on $\frac{1}{3}n^{\frac{1}{k}}$ vertices, but some have no induced path on more than $n^{\frac{2}{k}}$ vertices [HR23]. From this result, we generalized their $(\log n)^{\Theta(1)}$ lower bound for bounded genus graph to the topological-minor-free graphs, going along with their conjecture. However, the conjecture has recently been disproved [DR23] by Defrain and Raymond, who found a family of 2-degenerate graphs whose induced paths have order $O((\log\log n))^2$. This however exhibits three main behaviours for the order of the induced path in a k-degenerate class: $n^{\Theta(1)}$, $(\log n)^{\Theta(1)}$ and $(\log\log n)^{\Theta(1)}$. This raises the question:

Question 3.
Can we categorize the k-degenerate graph classes into three categories, those of longest induced path of size $n^{\Theta(1)}$, $(\log n)^{\Theta(1)}$ or $(\log\log n)^{\Theta(1)}$?

Moreover, for several classes, we have an upper and a lower bound on the order of a longest induced path.

Question 4.
How close can we get for the upper and lower bounds on the order of a longest induced path, for the graph classes we studied?

Also, instead of looking for path-like graphs as substructure, we can also look for them as super structure. More precisely in [DHK+21], we looked for the smallest number of edges to add to a graph (from various chordal classes), to turn it into a special case of interval graph, called a proper interval graph.

The second part of this chapter is a generalization of the Erdős-Pósa Theorem [EP65], showing that sparse n-vertex graphs with a small number of cycles that are pairwise not adjacent admit a set of logarithmically many vertices whose removal yields a forest [BBD+23].

A classical generalization of Erdős-Pósa theorem is the study the Erdős-Pósa property of a class. A graph class $\mathcal{H}$ has the Erdős-Pósa property if there is a function f such that for every graph G and integer k, the graph G contains either k vertex-disjoint subgraph isomorphic to a graph of $\mathcal{H}$, or a vertex set of size $f(k)$, whose removal yields a graph with no subgraph isomorphic to a graph in $\mathcal{H}$. Thus Erdős-Pósa theorem means that the cycles have the Erdős-Pósa property with $f(k) = O(k \log k)$. With our new result in mind, it would be interesting to look into a variant of the Erdős-Pósa property for sparse graphs, where instead of looking for k vertex-disjoint copies of graphs from $\mathcal{H}$, we look for independent copies of them, up to authorising the function to depend logarithmically in the order. We thus can explore this new problem:

Question 5.
What graph classes have the "independent" Erdős-Pósa property?

References

[ACP+18] T. Abrishami, M. Chudnovsky, M. Pilipczuk, P. Rzążewski, and P. Seymour, *Induced subgraphs of bounded treewidth and the container method*, in Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA), 2018, pp. 1948–1964. doi: 10.1137/1.9781611976465.116.

[AV00] J. Arocha and P. Valencia, *Long induced paths in 3-connected planar graphs*, Discussiones Mathematicae Graph Theory, 20 (2000), pp. 105–107. doi: 10.1016/j.tcs.2014.04.032.

[BBCR14] J. Babu, M. Basavaraju, L. S. Chandran, and D. Rajendraprasad, *2-connecting outerplanar graphs without blowing up the pathwidth*, Theoretical Computer Science, 554 (2014), pp. 119–134. doi: 10.1016/j.tcs.2014.04.032.

[BCE+82] L. Babai, F. Chung, P. Erdös, R. Graham, and J. Spencer, *On graphs which contain all sparse graphs*, in Theory and Practice of Combinatorics, P. L. Hammer, A. Rosa, G. Sabidussi, and J. Turgeon, eds., vol. 60 of North-Holland Mathematics Studies, North-Holland, 1982, pp. 21–26. doi: 10.1016/S0304-0208(08)73486-8.

[BCE+19] T. C. Biedl, E. W. Chambers, D. Eppstein, A. de Mesmay, and T. Ophelders, *Homotopy height, grid-major height and graph-drawing height*, in 27th International Symposium on Graph Drawing and Network Visualization (GD), vol. 11904 of Lecture Notes in Computer Science, Springer, September 2019, pp. 468–481. doi: 10.1007/978-3-030-35802-0_36.

[BEGS21] M. Bonamy, L. Esperet, C. Groenland, and A. Scott, *Optimal labelling schemes for adjacency, comparability, and reachability*, in Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2021, New York, NY, USA,

2021, Association for Computing Machinery, p. 1109–1117. DOI: 10.1145/3406325.3451102.

[Ber84] A. A. Bertossi, *Dominating sets for split and bipartite graphs*, Information processing letters, 19 (1984), pp. 37–40. DOI: 10.1016/0020-0190(84)90126-1.

[Bie11] T. C. Biedl, *Small drawings of outerplanar graphs, series-parallel graphs, and other planar graphs*, Discrete & Computational Geometry, 45 (2011), pp. 141–160. DOI: 10.1007/s00454-010-9310-z.

[Bie12] T. C. Biedl, *A 4-approximation for the height of drawing 2-connected outer-planar graphs*, in 10th International Workshop on Approximation and Online Algorithms (WAOA), T. Erlebach and G. Persiano, eds., vol. 7846 of Lecture Notes in Computer Science, Springer, September 2012, pp. 272–285. DOI: 10.1007/978-3-642-38016-7_22.

[Bod98] H. L. Bodlaender, *A partial k-arboretum of graphs with bounded treewidth*, Theoretical Computer Science, 209 (1998), pp. 1–45. DOI: 10.1016/S0304-3975(97)00228-4.

[Bod02] O. Bodini, *On the minimum size of a contraction-universal tree*, in 28th International Workshop on Graph-Theoretic Concepts in Computer Science (WG), vol. 2573 of Lecture Notes in Computer Science, Springer, June 2002, pp. 25–34. DOI: 10.1007/3-540-36379-3_3.

[CdV21] É. Colin de Verdière, *Algorithms for embedded graphs*, October 2021. monge.univ-mlv.fr/~colinde/cours/all-algo-embedded-graphs.pdf. Lectures Note.

[CFK+15] M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh, *Parameterized Algorithms*, Springer, 2015. ISBN: 978-3-319-21274-6. DOI: 10.1007/978-3-319-21275-3.

[Cou90] B. Courcelle, *The monadic second-order logic of graphs. I. Recognizable sets of finite graphs*, Information and Computation, 85 (1990), pp. 12–75. DOI: 10.1016/0890-5401(90)90043-H.

[CP84] D. G. Corneil and Y. Perl, *Clustering and domination in perfect graphs*, Discrete Applied Mathematics, 9 (1984), pp. 27–39. DOI: 10.1016/0166-218X(84)90088-X.

[DEG+20] V. Dujmović, L. Esperet, C. Gavoille, G. Joret, P. Micek, and P. Morin, *Adjacency labelling for planar graphs (and beyond)*, in 2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS), 2020, pp. 577–588. doi: 10.1109/FOCS46700.2020.00060.

[DEJ+20] V. Dujmović, L. Esperet, G. Joret, B. Walczak, and D. R. Wood, *Planar graphs have bounded nonrepetitive chromatic number*, Advances in Combinatorics, 5 (2020). doi: 10.19086/aic.12100.

[dFPP88] H. de Fraysseix, J. Pach, and R. Pollack, *Small sets supporting fary embeddings of planar graphs*, in 20th Annual ACM Symposium on Theory of Computing (STOC), ACM Press, 1988, pp. 426–433. doi: 10.1145/62212.62254.

[DG20] Y. Dieng and C. Gavoille, *On the treewidth of planar minor free graphs*, in 4th EAI International Conference on Innovations and Interdisciplinary Solutions for Underserved Areas (InterSol), J. P. Thorn, A. Gueye, and A. P. Hejnowicz, eds., vol. 321 of Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Springer, Cham, March 2020, pp. 238–250. doi: 10.1007/978-3-030-51051-0_17.

[DJM+20] V. Dujmović, G. Joret, P. Micek, P. Morin, T. Ueckerdt, and D. R. Wood, *Planar graphs have bounded queue-number*, Journal of the ACM, 67 (2020), pp. Article No. 22, pp. 1–38. doi: 10.1145/3385731.

[DK99] R. Diestel and D. Kühn, *A universal planar graph under the minor relation*, Graph Theory, 32 (1999), pp. 191–206. doi: 10.1002/(SICI)1097-0118(199910)32:2<191::AID-JGT8>3.0.CO;2-I.

[DLT94] D. Dolev, F. T. Leighton, and H. W. Trickey, *Planar embedding of planar graphs*, Advances in Computer Research, 2 (1994), pp. 147–161.

[DMW17] V. Dujmović, P. Morin, and D. R. Wood, *Layered separators in minor-closed graph classes with applications*, Journal of Combinatorial Theory, Series B, 127 (2017), pp. 111–147. doi: https://doi.org/10.1016/j.jctb.2017.05.006.

[DR23] O. Defrain and J.-F. Raymond, *Sparse graphs without long induced paths*, Tech. Rep. 2304.09679, arXiv, 2023.

[Dvo18] Z. Dvořák, *Induced subdivisions and bounded expansion*, European Journal of Combinatorics, 69 (2018), pp. 143–148. doi: 10.1016/j.ejc.2017.10.004.

[EJM22] L. Esperet, G. Joret, and P. Morin, *Sparse universal graphs for planarity*, Tech. Rep. 2010.05779, arXiv, 2022.

[ELM17] L. Esperet, L. Lemoine, and F. Maffray, *Long induced paths in graphs*, European Journal of Combinatorics, 62 (2017), pp. 1–14. doi: 10.1016/j.ejc.2016.11.011.

[EP65] P. Erdős and L. Pósa, *On independent circuits contained in a graph*, Canadian Journal of Mathematics, 17 (1965), pp. 347–352. doi: 10.4153/CJM-1965-035-8.

[FGL+96] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy, *Interactive proofs and the hardness of approximating cliques*, J. ACM, 43 (1996), p. 268–292. doi: 10.1145/226643.226652.

[FHdM22] N. Fuladi, A. Hubard, and A. de Mesmay, *Short topological decompositions of non-orientable surfaces*, Tech. Rep. 2203.06659v1 [cs.CG], arXiv, March 2022.

[FJWY21] S. Fiorini, G. Joret, S. Weltge, and Y. Yuditsky, *Integer programs with bounded subdeterminants and two nonzeros per row*, in 62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022, IEEE, 2021, pp. 13–24. doi: 10.1109/FOCS52979.2021.00011.

[Gav74] F. Gavril, *The intersection graphs of subtrees in trees are exactly the chordal graphs*, Journal of Combinatorial Theory, Series B, 16 (1974), pp. 47–56. doi: 10.1016/0095-8956(74)90094-X.

[GGL95] R. Graham, M. Grötschel, and L. Lovász, *Handbook of Combinatorics*, vol. 1, Elsevier, 1995.

[GL20] P. Gartland and D. Lokshtanov, *Independent set on P_k-free graphs in quasi-polynomial time*, in 2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS), 2020, pp. 613–624. doi: 10.1109/FOCS46700.2020.00063.

[GM15] M. Grohe and D. Marx, *Structure theorem and isomorphism test for graphs with excluded topological subgraphs*, SIAM Journal on Computing, 44 (2015), pp. 114–159. doi: 10.1137/120892234.

[Gol04] M. C. Golumbic, *Algorithmic graph theory and perfect graphs*, Elsevier, 2004.

[GRS82] F. Galvin, I. Rival, and B. Sands, *A Ramsey-type theorem for traceable graphs*, Journal of Combinatorial Theory, Series B, 33 (1982), pp. 7–16. doi: 10.1016/0095-8956(82)90053-3.

[GT12] Q.-P. Gu and H. Tamaki, *Improved bounds on the planar branchwidth with respect to the largest grid minor size*, Algorithmica, 64 (2012), pp. 416–453. doi: 10.1007/s00453-012-9627-5.

[Har82] J. Hartmanis, *Computers and intractability: A guide to the theory of NP-completeness (Michael R. Garey and David S. Johnson)*, SIAM Rev., 24 (1982), p. 90–91. doi: 10.1137/1024022.

[HH22] H. Hatami and P. Hatami, *The implicit graph conjecture is false*, in 2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS), 2022, pp. 1134–1137. doi: 10.1109/FOCS54457.2022.00109.

[Jae74] F. Jaeger, *On vertex-induced forests in cubic graphs*, in Proceedings of the 5th Southeastern Conference on Combinatorics, Graph Theory and Computing, 1974, pp. 501–512.

[KKTW01] D. Král, J. Kratochvíl, Z. Tuza, and G. J. Woeginger, *Complexity of coloring graphs without forbidden induced subgraphs*, in International Workshop on Graph-Theoretic Concepts in Computer Science, Springer, 2001, pp. 254–262. doi: 10.1007/3-540-45477-2_23.

[KO04] D. Kühn and D. Osthus, *Induced subdivisions in $K_{s,s}$-free graphs of large average degree*, Combinatorica, 24 (2004), pp. 287–304. doi: 10.1007/s00493-004-0017-8.

[KST54] T. Kövári, V. T. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloquium Mathematicum, 3 (1954), pp. 50–57. doi: 10.4064/cm-3-1-50-57.

[Lei80] C. E. Leiserson, *Area-efficient graph layouts (for VLSI)*, in 21st Annual IEEE Symposium on Foundations of Computer Science (FOCS), IEEE Computer Society Press, October 1980, pp. 270–280. doi: 10.1109/SFCS.1980.13.

[LPVV01] F. Lazarus, M. Pocchiola, G. Vegter, and A. Verroust, *Computing a canonical polygonal schema of an orientable triangulated surface*, in 17th Annual ACM Symposium on Computational Geometry (SoCG), ACM-SIAM, June 2001, pp. 80–89. doi: 10.1145/378583.378630.

[NDM12] J. Nešetřil and P. O. De Mendez, *Sparsity: graphs, structures, and algorithms*, vol. 28, Springer Science & Business Media, 2012.

[NSS22] T. Nguyen, A. Scott, and P. Seymour, *Induced paths in graphs without anticomplete cycles*, Tech. Rep. 2212.01089, arXiv, 2022.

[Pil11] M. Pilipczuk, *Problems parameterized by treewidth tractable in single exponential time: A logical approach*, in Mathematical Foundations of Computer Science 2011 - 36th International Symposium, MFCS 2011, Warsaw, Poland, August 22-26, 2011. Proceedings, F. Murlak and P. Sankowski, eds., vol. 6907 of Lecture Notes in Computer Science, Springer, 2011, pp. 520–531. doi: 10.1007/978-3-642-22993-0_47.

[Pol74] S. Poljak, *A note on stable sets and colorings of graphs*, Commentationes Mathematicae Universitatis Carolinae, 15 (1974), pp. 307–309.

[PT99] A. Proskurowski and J. A. Telle, *Classes of graphs with restricted interval models*, Discrete Mathematics & Theoretical Computer Science, Vol. 3 no. 4 (1999). doi: 10.46298/dmtcs.263.

[RS86] N. Robertson and P. D. Seymour, *Graph minors. V. Excluding a planar graph*, Journal of Combinatorial Theory, Series B, 41 (1986), pp. 92–114. doi: 10.1016/0095-8956(86)90030-4.

[RS95] N. Robertson and P. D. Seymour, *Graph minors. XIII. The disjoint paths problem*, Journal of Combinatorial Theory, Series B, 63 (1995), pp. 65–110. doi: 10.1006/jctb.1995.1006.

[RS03] N. Robertson and P. D. Seymour, *Graph minors. XVI. excluding a non-planar graph*, Journal of Combinatorial Theory, Series B, 89 (2003), pp. 43–76. doi: 10.1016/S0095-8956(03)00042-X.

[RS04] N. Robertson and P. D. Seymour, *Graph minors. XX. Wagner's conjecture*, Journal of Combinatorial Theory, Series B, 92 (2004), pp. 325–357. doi: 10.1016/j.jctb.2004.08.001.

[RST94] N. Robertson, P. D. Seymour, and R. Thomas, *Quickly excluding a planar graph*, Journal of Combinatorial Theory, Series B, 62 (1994), pp. 323–348. doi: 10.1006/jctb.1994.1073.

[Sud03] M. Suderman, *Pathwidth and layered drawings of trees*, International Journal of Computational Geometry and Applications, 14 (2003). doi: 10.1142/S0218195904001433.

[Tur41] P. Turán, *On an external problem in graph theory*, Matematikai és Fizikai Lapok, 48 (1941), pp. 436–452.

[Whi31] H. Whitney, *A theorem on graphs*, Annals of Mathematics, 32 (1931), pp. 378–390. doi: 10.2307/1968197.

Publications

Article in Peer-reviewed Conference

[BBD+23] M. Bonamy, E. Bonnet, H. Déprés, L. Esperet, C. Geniet, C. Hilaire, S. Thomassé, and A. Wesolek, *Sparse graphs with bounded induced cycle packing number have logarithmic treewidth*, in Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), Society for Industrial and Applied Mathematics, 2023, pp. 3006–3028. doi: 10.1137/1.9781611977554.ch116.

Article in Peer-reviewed Journal

[HR23] C. Hilaire and J.-F. Raymond, *Long induced paths in minor-closed graph classes and beyond*, The Electronic Journal of Combinatorics, 30 (2023), pp. 1–18. doi: 10.37236/11029.

Preprints

[BCH+23] P. Blanco, L. Cook, M. Hatzel, C. Hilaire, F. Illingworth, and R. McCarty, *On tree decompositions whose trees are minors*, Tech. Rep. 2302.12106v1, arXiv, 2023.

[DHK+21] F. Dross, C. Hilaire, I. Koch, V. Leoni, N. Pardal, M. I. L. Pujato, and V. F. dos Santos, *On the proper interval completion problem within some chordal subclasses*, Tech. Rep. 2110.07706, arXiv, 2021.

[GH23a] C. Gavoille and C. Hilaire, *Grid-major of poly-line drawings*, tech. rep., manuscript, 2023.

[GH23b] C. Gavoille and C. Hilaire, *Minor-universal graph for graphs on surfaces*, Tech. Rep. 2305.06673, arXiv, 2023.