Avoidable paths in graphs

Marthe Bonamy, Oscar Defrain, Meike Hatzel, Jocelyn Thiebaut

Avoidable paths in graphs*

Marthe Bonamy[1], Oscar Defrain[2], Meike Hatzel[3], and Jocelyn Thiebaut[4]

[1] *CNRS, LaBRI, Université de Bordeaux, France.*
[2] *LIMOS, Université Clermont Auvergne, France.*
[3] *LaS, Technische Universität Berlin, Germany.*
[4] *LIRMM, Université de Montpellier, France.*

August 13, 2019

Abstract

We prove a recent conjecture of Beisegel et al. that for every positive integer k, every graph containing an induced P_k also contains an avoidable P_k. Avoidability generalises the notion of simpliciality best known in the context of chordal graphs. The conjecture was only established for $k \in \{1, 2\}$ (Ohtsuki et al. 1976, and Beisegel et al. 2019, respectively). Our result also implies a result of Chvátal et al. 2002, which assumed cycle restrictions. We provide a constructive and elementary proof, relying on a single trick regarding the induction hypothesis. In the line of previous works, we discuss conditions for multiple avoidable paths to exist.

1 Introduction

A graph G is *chordal* if every induced cycle is of length three. A classical result of Dirac [Dir61] states that every chordal graph has a *simplicial* vertex, that is, a vertex which neighbourhood is a clique. However, not all graphs exhibit the nice structure of chordal graphs, and the statement does not extend to general graphs.

1.1 From simplicial vertices to avoidable paths

One way to generalise Dirac's result is through the following more flexible notion.

Definition 1.1 (Avoidable vertex). *A vertex v in a graph G is* avoidable *if every induced path on three vertices with middle vertex v is contained in an induced cycle in G.* ⊣

Note that in a chordal graph, every avoidable vertex is simplicial. The next theorem can be inferred from [OCF76, BB98, ACTV15]; see also [BCG+19] for a nice introduction.

Theorem 1.2. *Every graph has an avoidable vertex.*

*The first author has been supported by the ANR project GrR ANR-18-CE40-0032. The second author has been supported by the ANR project GraphEn ANR-15-CE40-0009. The third author was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC Consolidator Grant DISTRUCT, grant agreement No 648527).

Recently in [BCG+19], the authors considered a generalisation of the concept of avoidable vertices to edges, and extended Theorem 1.2 to that notion.

Definition 1.3 (Avoidable edge). *An edge uv in a graph G is* avoidable *if every induced path on four vertices with middle edge uv is contained in an induced cycle in G.* ⊣

Theorem 1.4 (Beisegel et al. [BCG+19]). *Every graph has an avoidable edge.*

This notion naturally generalises to paths, as follows.

Definition 1.5 (Extension). *Given an induced path P in a graph G, an* extension *of P is an induced path xPy in G for some vertices x, y.* ⊣

Definition 1.6 (Failing). *An induced path P in a graph G is* failing *if there is no induced cycle of G containing P.* ⊣

Definition 1.7 (Avoidable). *A path P in a graph G is* avoidable *if it is induced and has no failing extension. Given a subgraph G' of G, we say that P is an avoidable path of G in G' if it is avoidable in G and $V(P) \subseteq V(G')$.* ⊣

A graph G is P_k-free if it does not contain a P_k, that is, an induced path on k vertices. In [BCG+19] the authors conjecture that for every positive integer k, every graph either is P_k-free or contains an avoidable path on k vertices. This conjecture is motivated by the following result of Chvátal et al. [CRS02], which generalises Dirac's theorem. A $C_{\geqslant p}$-free graph is a graph where every induced cycle has at most $p-1$ vertices. The $C_{\geqslant 4}$-free graphs are exactly the chordal graphs. Unless specified otherwise, we consider cycles to be induced.

Theorem 1.8 (Chvátal et al. [CRS02]). *For every positive integer k, every $C_{\geqslant k+3}$-free graph either is P_k-free or contains an avoidable path on k vertices.*

In fact, Theorem 1.8 originally states the existence of a simplicial path in the class of $C_{\geqslant k+3}$-free graphs. A *simplicial* path is an induced path with no extension: it is avoidable by vacuity. Note that these two definitions coincide in such a class, as no cycle on at most $k+2$ vertices can contain the extension of an induced path on k vertices.

Here, we confirm the aforementioned conjecture [BCG+19, Conjecture 1], as follows.

Theorem 1.9. *For every positive integer k, every graph either is P_k-free or contains an avoidable P_k.*

In fact, we prove Theorem 1.9 using a stronger induction hypothesis, in the exact same flavour as [CRS02], see Theorem 2.4 in Section 2.

1.2 Consequences

We point out that the proof of Theorem 1.9 is self-sufficient, thus this supersedes the arguments for theorems 1.2, 1.4 and 1.8.

By using ingredients of Theorem 2.4 (namely Lemma 2.3), we obtain a way to build more than one avoidable P_k.

Corollary 1.10. *For every positive integer k, graph G and subset $X \subseteq V(G)$ such that $G[X]$ is connected, either $G-N[X]$ is P_k-free or there is an avoidable P_k of G in $G-N[X]$.*

Corollary 1.11. *For every positive integer k and graph G, either G does not contain two non-adjacent P_k, or it contains two non-adjacent avoidable P_k.*

Since Corollary 1.11 is not as straightforward as its predecessor, we include a proof.

Proof. Let Q_1 and Q_2 be two non-adjacent P_k. By Corollary 1.10, either $G - N[Q_1]$ is P_k-free or there is an avoidable P_k of G in $G - N[Q_1]$. The first outcome is ruled out by the existence of Q_2. Let Q_2' be an avoidable P_k of G in $G - N[Q_1]$. We repeat the argument with Q_2' instead of Q_1, and obtain an avoidable P_k of G in $G - N[Q_2']$, call it Q_1'. The two paths Q_1' and Q_2' are two non-adjacent avoidable P_k, as desired. □

We can also wonder:

Question 1.12. *For every positive integer k, does every graph G either not contain two disjoint P_k, or contain two disjoint avoidable P_k?*

We know the answer to be positive in the case $k \in \{1, 2\}$, due to [BCG+19, Theorems 3.3 and 6.4]. The answer turns out to be negative in all other cases, as exhibited in the following counter-example for $k \geqslant 3$, which consists of a cycle on $2k - 1$ vertices with an added vertex adjacent to two consecutive vertices on the cycle (see Figure 1 for the case $k = 3$). This graph contains two disjoint P_k, and it has $2k$ vertices, so any two disjoint P_k are in fact complementary in the graph. Suppose that it contains two disjoint avoidable P_k, and note that each intersects the triangle (otherwise the complement would not be a path). Since there are three vertices in the triangle, there is an avoidable P_k containing a single vertex in the triangle. This P_k has a failing extension, a contradiction.

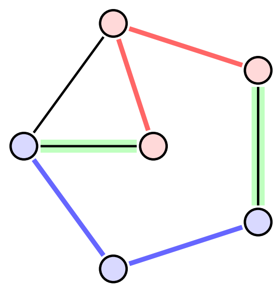

Figure 1: A graph that contains two disjoint P_3 (in blue and in red) but no two disjoint avoidable P_3 (there is a unique partition into two disjoint P_3, up to symmetry). In green, a failing extension of the blue path.

In Section 3, we present a concise algorithm which follows the proof of Theorem 2.4. As discussed there, the algorithm has complexity $O(n^{k+2})$ which, while naive, is the right order of magnitude under ETH.

2 A stronger induction hypothesis

All graphs considered in this paper are finite, simple and loopless. Given a graph G, we denote by $V(G)$ its set of vertices, and by $E(G) \subseteq \{\{x, y\} \mid x, y \in V(G),\ x \neq y\}$ its set of edges. Edges are denoted by uv (or vu) instead of $\{u, v\}$. If uv is an edge, then we say that u and v are *adjacent*. Given a vertex u, the neighbourhood $N(u)$ of u is the set of vertices of G that are adjacent to u. The closed neighbourhood $N[u]$ of u is the set $N(u) \cup \{u\}$. If $X \subseteq V(G)$, then we define $N[X] \coloneqq \bigcup_{x \in X} N[x]$ and $N(X) \coloneqq N[X] \setminus X$. The subgraph of

G *induced* by X, denoted by $G[X]$, is the graph $(X, E(G) \cap \{\{x,y\} \mid x,y \in X,\ x \neq y\})$, and $G - X$ is the graph $G[V(G) \setminus X]$. Given two adjacent vertices u_1 and u_2 of G, the graph obtained by *merging* u_1 and u_2 is the graph obtained from G by replacing u_1 and u_2 with a new vertex u such that $N(u) = N(\{u_1, u_2\})$. Given a graph G and two subsets X and Y of $V(G)$, we say that X *dominates* Y if every vertex of $Y \setminus X$ has a neighbour in X (equivalently, if $Y \subseteq N[X]$).

We first define two useful properties.

Definition 2.1 (Basic property H_B)**.** *Given a positive integer k and a graph G, the property $H_B(G,k)$ holds if either G is P_k-free or there is an avoidable P_k in G.* $\dashv$

Definition 2.2 (Refined property H_R)**.** *Given a positive integer k, a graph G and a vertex $u \in V(G)$, the property $H_R(G,k,u)$ holds if either $G - N[u]$ is P_k-free or there is an avoidable P_k of G in $G - N[u]$.*

Given a positive integer k and a graph G, the property $H_R(G,k)$ holds if $H_R(G,k,u)$ holds for every $u \in V(G)$. $\dashv$

Note that property H_R does not directly imply property H_B. We also emphasise the fact that an avoidable path in a subgraph is not necessarily an avoidable path in the whole graph.

We now prove a form of heredity in H_R.

Lemma 2.3. *Let k be a positive integer, G a graph and u_1u_2 an edge of G. Let G' be the graph obtained from G by merging the two vertices u_1 and u_2 into one vertex u. If $G' - N[u]$ contains a P_k, then $H_R(G',k,u)$ implies $H_R(G,k,u_1)$.*

Proof. Suppose $G' - N[u]$ contains a P_k, and that $H_R(G',k,u)$ holds but not $H_R(G,k,u_1)$. Since $G' - N[u]$ is not P_k-free, there is an avoidable P_k of G' in $G' - N[u]$. Call it Q. The path Q is contained in $G' - N[u] = G - N[\{u_1,u_2\}]$, so in particular in $G - N[u_1]$. Since $H_R(G,k,u_1)$ does not hold, Q is not an avoidable P_k of G. Thus, there is a failing extension xQy of Q in G. Note that x, y, u_1, and u_2 are all pairwise distinct.

Hence, xQy is an extension of Q in G', and there is an induced cycle C in G' containing the path xQy. If $u \notin C$, then the cycle C is also an induced cycle in G containing xQy, a contradiction. Therefore, $u \in C$. By replacing u with either u_1, u_2 or the edge u_1u_2 as appropriate, we obtain an induced cycle in G containing xQy, a contradiction. $\square$

We are now ready to prove the main technical result of this paper.

Theorem 2.4. *For every positive integer k and every graph G, both properties $H_B(G,k)$ and $H_R(G,k)$ hold.*

Proof. Suppose the statement is false and consider a counter-example G which is minimal with respect to the number of vertices.

Lemma 2.5. *The property $H_R(G,k)$ holds for every k.*

Proof. We proceed by contradiction. Suppose that $H_R(G,k,u)$ does not hold for some k and some vertex $u \in V(G)$, that is, there exists a P_k in $G - N[u]$, and every P_k in $G - N[u]$ has a failing extension. We prove the following.

Claim 2.6. *Every P_k in $G - N[u]$ dominates $N(u)$.*

Proof. Assume towards a contradiction that there is a P_k in $G - N[u]$, call it Q, which is not adjacent to some vertex $v \in N(u)$. Then $G - N[\{u, v\}]$ contains a P_k. Let G' be the graph obtained from G by merging u and v into a vertex u'. Since G' has fewer vertices than G, the property $H_R(G', k, u')$ holds by minimality of G. By Lemma 2.3, the property $H_R(G, k, u)$ holds, a contradiction. ⌟

Let $G' := G - N[u]$. Then G' contains a P_k. As G' contains fewer vertices than G, the property $H_B(G', k)$ holds. Let Q be an avoidable P_k of G'. By assumption, Q is not an avoidable P_k of G. So there is a failing extension xQy of Q in G. Since Q has no failing extension in G', we can assume without loss of generality that $y \in N(u)$. It follows that $x \notin N(u)$: otherwise the cycle $xQyu$ contradicts the fact that xQy is failing. By definition of an extension, xQy is an induced path. Let z be the only neighbour of y in Q, and let us now consider the path $xQ - z$. It is a P_k, and it does not intersect $N[u]$. However, no vertex in it is adjacent to y which lies in $N(u)$, contradicting Claim 2.6. ■

Lemma 2.7. *The property $H_B(G, k)$ holds for every k.*

Proof. Assume towards a contradiction that for some k, the property $H_B(G, k)$ does not hold. By Lemma 2.5, the property $H_R(G, k, u)$ holds for every vertex $u \in V(G)$. In other words, the graph G contains a P_k but no avoidable P_k, and for every vertex $u \in V(G)$, either $G - N[u]$ is P_k-free or there is an avoidable P_k of G in $G - N[u]$.

We derive the following claim.

Claim 2.8. *Every P_k in G dominates $V(G)$.*

Proof. Suppose there is a P_k, call it Q, that does not dominate some vertex u of G. Since $H_R(G, k)$ holds, either $G - N[u]$ is P_k-free or there is an avoidable P_k of G in $G - N[u]$. The first case contradicts the existence of Q, and the second contradicts the fact that $H_B(G, k)$ does not hold. ⌟

Since $H_B(G, k)$ does not hold, G contains a P_k, say Q, that is not avoidable. So it has a failing extension xQy. Let z be the only neighbour of y in Q, and consider the path $xQ - z$. It is an induced P_k and none of its vertices is adjacent to y. This contradicts Claim 2.8. ■

Finally, lemmas 2.5 and 2.7 together contradict G being a counter-example. □

Theorem 1.9 directly follows from Theorem 2.4.

3 An algorithm for Theorem 2.4

By going through the proof and extracting the key ingredients, we obtain a straightforward algorithm verifying both properties (see Algorithm 1).

The algorithm uses the subprocedure InducedPath that, given a graph G and a positive integer k, decides whether G contains a P_k. If it does, the procedure returns one, otherwise it returns `null`. The naive algorithm for that (testing all subsets of size k) has complexity $O(n^k)$. However, this is nearly optimal. Indeed, the problem of finding a P_k in a given graph is W[1]-hard[1] when parametrised by k (see [CFK+15, Ex. 13.16, p. 460]).

[1] see e.g. [CFK+15] for definitions around complexity

Algorithm 1 finds an avoidable path of given length in a given graph, if any.

procedure FindAvoidablePathRefined(G, k, u)
    for all $v \in N(u)$ **do**
        if InducedPath$(G - N[\{u, v\}], k) \neq$ `null` **then**
            $G' \leftarrow G$ with u and v merged into u'
            return FindAvoidablePathRefined(G', k, u')
    return FindAvoidablePath$(G - N[u], k)$

procedure FindAvoidablePath(G, k)
    for all $u \in V(G)$ **do**
        if InducedPath$(G - N[u], k) \neq$ `null` **then**
            return FindAvoidablePathRefined(G, k, u)
    return InducedPath(G, k)

In fact, the hinted reduction has a linear blow-up, so it follows that there is no $f(k) \cdot n^{o(k)}$ algorithm under ETH.

Let k be a positive integer, and let $B(n)$ (resp. $R(n)$) be the worst case complexity of FindAvoidablePath (resp. FindAvoidablePathRefined) on an n-vertex graph with parameter k. We have $B(n) \leqslant n \cdot n^k + \max(R(n), n^k)$, and $R(n) \leqslant n \cdot n^k + \max(R(n-1), B(n-2))$. We obtain $R(n) \leqslant n^{k+2}$ and $B(n) \leqslant n^{k+2} + n^{k+1}$. While this may well be improved, the known limitations for finding an induced path on k vertices also apply for an induced avoidable path on k vertices (by Theorem 1.9, if the first exists, then so does the second). Therefore, the order of magnitude of this naive algorithm is correct.

Note that there is a yet more naive algorithm blindly checking for every subset of size k if it corresponds to an avoidable path. That algorithm has comparable complexity to ours (though slightly worse, at least at first sight). However, we wanted to emphasise that our proof of Theorem 2.4 is constructive and yields an elementary algorithm. Also, we believe that it provides an outline of the proof which might be helpful to the reader.

4 Conclusion

Given the discussions in Section 1.2, it is tempting to ask when a graph admits three (or more) disjoint (resp. pairwise non-adjacent) avoidable paths. Note that though Corollary 1.10 arms us with sufficient conditions for there to be more than two avoidable P_k, we do not believe that the corresponding sufficient conditions are necessary. However, it seems the picture is murky already for chordal graphs.

It is tempting to wonder whether we can obtain another avoidable structure. Though in some cases the very notion of extension becomes unclear (what should an extension of a clique be?), it does not seem like any other structure survives the test of chordal graphs or simple ad hoc constructions—even when allowing a family of graphs instead of fixing a single pattern (like a path on k vertices). This motivates us to formulate the following question.

Question 4.1. *Does there exist a family $\mathcal{H}$ of connected graphs, not containing any path, such that any graph is either $\mathcal{H}$-free or contains an avoidable element of $\mathcal{H}$?*

The notion of avoidability in this context is deliberately left up to interpretation.

Acknowledgements

We gratefully acknowledge support from Nicolas Bonichon and the Simon family for the organisation of the 4th Pessac Graph Workshop, where this research was done. We are indebted to Michał Pilipczuk for providing helpful references regarding the complexity of finding an induced path of given length. Last but not least, we thank Peppie for her unwavering support during the work sessions.

References

[ACTV15] Pierre Aboulker, Pierre Charbit, Nicolas Trotignon, and Kristina Vušković. Vertex elimination orderings for hereditary graph classes. *Discrete Mathematics*, 338(5):825–834, 2015.

[BB98] Anne Berry and Jean-Paul Bordat. Separability generalizes Dirac's theorem. *Discrete Applied Mathematics*, 84(1-3):43–53, 1998.

[BCG+19] Jesse Beisegel, Maria Chudnovsky, Vladimir Gurvich, Martin Milanič, and Mary Servatius. Avoidable vertices and edges in graphs. In Zachary Friggstad, Jörg-Rüdiger Sack, and Mohammad R Salavatipour, editors, *Algorithms and Data Structures*, pages 126–139, Cham, 2019. Springer International Publishing.

[CFK+15] Marek Cygan, Fedor V Fomin, Łukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michał Pilipczuk, and Saket Saurabh. *Parameterized algorithms.* Springer, 2015.

[CRS02] Vašek Chvátal, Irena Rusu, and R. Sritharan. Dirac-type characterizations of graphs without long chordless cycles. *Discrete Mathematics*, 256(1-2):445–448, 2002.

[Dir61] Gabriel Andrew Dirac. On rigid circuit graphs. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 25(1):71–76, 1961.

[OCF76] Tatsuo Ohtsuki, Lap Kit Cheung, and Toshio Fujisawa. Minimal triangulation of a graph and optimal pivoting order in a sparse matrix. *Journal of Mathematical Analysis and Applications*, 54(3):622–633, 1976.