Graphs with no induced P_5 or $\overline{P_5}$

Maria Chudnovsky [*] Louis Esperet [†] Laetitia Lemoine [‡]
Peter Maceli [§] Frédéric Maffray [¶] Irena Penev [‖]

September 30, 2014

Abstract

We prove that a graph G contains no induced 5-vertex path and no induced complement of a 5-vertex path if and only if G is obtained from 5-cycles and split graphs by repeatedly applying the following operations: substitution, split unification, and split unification in the complement (where split unification is a new class-preserving operation that is introduced in this paper).

1 Introduction

All graphs in this paper are finite and simple. For fixed $n \geq 1$, let P_n denote the path on n vertices, and for $n \geq 3$ let C_n denote the cycle on n vertices. The complement of a graph G is denoted by $\overline{G}$. Given graphs G and F, we say that G is F-*free* if G does not contain (an isomorphic copy of) F as an induced subgraph. Given a family $\mathcal{F}$ of graphs, we say that a graph G is $\mathcal{F}$-*free* provided that G is F-free for all $F \in \mathcal{F}$.

There are various instances of the collection $\mathcal{F}$ such that the $\mathcal{F}$-free graphs are highly structured in a way that can be described precisely, which is an interesting fact in itself and also for the purpose of solving some optimization problems on such graphs. Two famous cases are the P_4-free graphs (also known as *cographs*) and the $\{C_4, \overline{C}_4, C_5\}$-free graphs (also known as *split* graphs). Clearly, the class of $\{P_5, \overline{P_5}\}$-free graphs contains all cographs and all split graphs. The goal of this paper is to understand the structure of $\{P_5, \overline{P_5}\}$-free graphs.

This paper results from the merging of the two manuscripts [2] and [4] on the same subject; it combines the proofs and results from these two manuscripts so as to present them in the most succint way.

[*] Columbia University, New York, NY 10027, USA. E-mail: mchudnov@columbia.edu. Partially supported by NSF grants DMS-1001091 and IIS-1117631.

[†] CNRS, Laboratoire G-SCOP, University of Grenoble, France.

[‡] Laboratoire G-SCOP, University of Grenoble, France.

[§] Columbia University, New York, NY 10027, USA. E-mail: plm2109@columbia.edu.

[¶] CNRS, Laboratoire G-SCOP, University of Grenoble, France.

[‖] Université de Lyon, LIP, UMR 5668, ENS de Lyon - CNRS - UCBL - INRIA. E-mail: irena.penev@ens-lyon.fr. This work is supported by the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). Partially supported by ANR project Stint under reference ANR-13-BS02-0007.

2 Definitions

For a graph G, we denote by $V(G)$ its vertex-set and by $E(G)$ its edge-set. Given a set $S \subseteq V(G)$, let $N(S)$ be the set of vertices in $V(G) \setminus S$ that have a neighbor in S. Let $G[S]$ denote the induced subgraph of G with vertex-set S, and let $G \setminus S$ denote the induced subgraph $G[V(G) \setminus S]$. We say that a vertex v in $V(G) \setminus S$ is *complete* to S if v is adjacent to every vertex of S, and v is *anticomplete* to S if v has no neighbor in S. A vertex of $V(G) \setminus S$ that is neither complete nor anticomplete to S is *mixed* on S. Given two disjoint sets $S, T \subseteq V(G)$, we say that S is *complete* to T when every vertex of S is complete to T, and S is *anticomplete* to T when every vertex of S is anticomplete to T.

An *anticomponent* of a set $S \subseteq V(G)$ is any subset of S that induces a component of the graph $\overline{G}[S]$.

A *homogeneous set* is a set $S \subseteq V(G)$ such that every vertex of $V(G) \setminus S$ is either complete or anticomplete to S. A homogeneous set S is *proper* when $|S| \geq 2$ and $S \neq V(G)$. Let G be a graph that admits a proper homogeneous set S, and let s be any vertex in S. We can decompose G into the two graphs $G[S]$ and $G \setminus (S \setminus s)$. Note that, up to isomorphism, the latter graph is the same whatever the choice of s, because S is a homogeneous set. Moreover, both $G[S]$ and $G \setminus (S \setminus s)$ are induced subgraphs of G.
The reverse operation, known as *substitution*, can be defined as follows. Let G and H be two vertex-disjoint graphs and let x be a vertex in G. Make a graph G' with vertex-set $V(G \setminus x) \cup V(H)$, taking the union of the two graphs $G \setminus x$ and H and adding all edges between $V(H)$ and the neighborhood of x in G. Clearly, in G', the set $V(H)$ is a homogeneous set, $H = G'[V(H)]$ and G is isomorphic to an induced subgraph of G'. Moreover $V(H)$ is a proper homogeneous set if both G and H have at least two vertices. Thus, a graph G is obtained by substitution from smaller graphs if and only if G contains a proper homogeneous set.

A graph is *prime* if it has no proper homogeneous set.

The following result about the structure of $\{P_5, \overline{P_5}\}$-free graphs was proven by Fouquet in [7].

Theorem 2.1 ([7]) *For each $\{P_5, \overline{P_5}\}$-free graph G at least one of the following holds:*

- *G contains a proper homogeneous set;*
- *G is isomorphic to C_5;*
- *G is C_5-free.*

The graph C_5 is also called a *pentagon*. Theorem 2.1 immediately implies that every $\{P_5, \overline{P_5}, C_5\}$-free graph can be obtained by substitution starting from $\{P_5, \overline{P_5}, C_5\}$-free graphs and pentagons. Furthermore, it is easy to check that every graph obtained by substitution starting from $\{P_5, \overline{P_5}, C_5\}$-free graphs and pentagons is $\{P_5, \overline{P_5}\}$-free. We remark that the Strong Perfect Graph Theorem [3] implies that a $\{P_5, \overline{P_5}\}$-free graph is perfect if and only if it is C_5-free. Thus, every $\{P_5, \overline{P_5}\}$-free graph can be obtained by substitution starting from $\{P_5, \overline{P_5}\}$-free perfect graphs and pentagons. In view of this, the bulk of this paper focuses on prime $\{P_5, \overline{P_5}, C_5\}$-free graphs (equivalently: prime $\{P_5, \overline{P_5}\}$-free perfect graphs).

3 Prime $\{P_5, \overline{P_5}, C_5\}$-free graphs

A graph is *split* if its vertex-set can be partitioned into a stable set and a clique. Földes and Hammer [5, 6] gave the following characterization of split graphs (a short proof is given in [8, p. 151]).

Theorem 3.1 ([5, 6]) *A graph is split if and only if it is $\{C_4, \overline{C_4}, C_5\}$-free.*

Lemma 3.2 *In a $\{P_5, \overline{P_5}, C_5\}$-free graph G, let A and B be non-empty and disjoint subsets of $V(G)$, and let t be a vertex in $V(G) \setminus (A \cup B)$ such that:*

- *t is anticomplete to A and complete to B,*
- *every vertex in B has a neighbor in A, and*
- *A is connected.*

Then some vertex of A is complete to B.

Proof. Pick a vertex a in A with the maximum number of neighbors in B. Suppose that a has a non-neighbor y in B. We know that y has a neighbor a' in A. Since A is connected, there is a path $P = a_0\text{-}\cdots\text{-}a_k$ in $G[A]$ with $k \geq 1$, $a_0 = a'$ and $a_k = a$. Choose a' such that k is minimal. So P is chordless and y has no neighbor in $P \setminus \{a_0\}$. Then $k = 1$, for otherwise t, y, a_0, a_1, a_2 induce a P_5. By the choice of a, since y is adjacent to a' and not to a, there is a vertex z in B adjacent to a and not to a'. Then a, z, t, y, a' induce a C_5 or $\overline{P_5}$ (depending on the pair y, z), a contradiction. Thus a is complete to B. $\Box$

We say that a set, or a graph, is *big* if it contains at least two vertices.

Theorem 3.3 *Let G be a prime $\{P_5, \overline{P_5}, C_5\}$-free graph that contains a $\overline{C_4}$. Then there are pairwise disjoint subsets $X_0, X_1, \ldots, X_m, Y_0, Y_1, \ldots, Y_m$, with $m \geq 2$, whose union is equal to $V(G)$, such that the following properties hold, where $X = X_0 \cup X_1 \cup \cdots \cup X_m$ and $Y = Y_0 \cup Y_1 \cup \cdots \cup Y_m$:*

(i) *For each $i \in \{1, \ldots, m\}$, X_i is connected, $|X_i| \geq 2$, X_0 is a stable set, and $X_0, X_1, \ldots, X_m$ are pairwise anticomplete to each other.*

(ii) *For each $i \in \{1, \ldots, m\}$, $Y_i \neq \emptyset$, every vertex of Y_i is mixed on X_i and complete to $X \setminus (X_i \cup X_0)$, and Y_0 is complete to $X \setminus X_0$.*

(iii) *$Y_0, Y_1, \ldots, Y_m$ are pairwise complete to each other. (So each anticomponent of Y is included in some Y_i with $i \in \{0, \ldots, m\}$.)*

(iv) *No vertex of $X \setminus X_0$ is mixed on any anticomponent of Y.*

(v) *For each $i \in \{1, \ldots, m\}$, X_i contains a vertex that is complete to Y.*

(vi) *Every vertex of X_0 is mixed on at most one anticomponent of Y.*

(vii) *For every big anticomponent Z of Y, the set X_Z of vertices of X_0 that are mixed on Z is not empty. Moreover, if Z and Z' are any two distinct big anticomponents of Y, then $X_Z \cap X_{Z'} = \emptyset$.*

(viii) *Each big anticomponent Z of Y contains a vertex that is anticomplete to X_Z.*

(ix) If Y is not a clique, there is a big anticomponent Z of Y such that X_Z is anticomplete to all big anticomponents of $Y \setminus Z$.

Proof. Since G contains a $\overline{C_4}$, there is a subset X of $V(G)$ such that $G[X]$ has at least two big components. We choose X maximal with this property. Let $X_1, \ldots, X_m$ ($m \geq 2$) be the vertex-sets of the big components of $G[X]$, and let $X_0 = X \setminus (X_1 \cup \cdots \cup X_m)$. So (i) holds. Let $Y = V(G) \setminus X$. We claim that:

$$\text{For every } y \in Y \text{ and } i \in \{1, \ldots, m\},\ y \text{ has a neighbor in } X_i. \tag{1}$$

Proof. If y has no neighbour in X_i, then $X \cup \{y\}$ induces a subgraph of G with at least two big components (one of which is X_i), which contradicts the maximality of X. Thus (1) holds.

$$\begin{array}{c}\text{For every vertex } y \in Y, \text{ there is at most one integer } i \text{ in} \\ \{1, \ldots, m\} \text{ such that } y \text{ has a non-neighbor in } X_i.\end{array} \tag{2}$$

Proof. Suppose that y has a non-neighbor in two distinct components X_i and X_j of X. For each $h \in \{i, j\}$, y has a neighbor in X_h by (1), and since X_h is connected, there are adjacent vertices $u_h, v_h \in X_h$ such that y is adjacent to u_h and not to v_h. Then v_i, u_i, y, u_j, v_j induce a P_5, a contradiction. Thus (2) holds.

An immediate consequence of Claims (1) and (2) is the following.

$$\begin{array}{l}\text{For every vertex } y \in Y, \text{ either } y \text{ is complete to } X \setminus X_0, \\ \text{or there is a unique integer } i \in \{1, \ldots, m\} \text{ such that } y \text{ is} \\ \text{complete to } X \setminus (X_i \cup X_0) \text{ and } y \text{ is mixed on } X_i.\end{array} \tag{3}$$

For each $i \in \{1, \ldots, m\}$, let $Y_i = \{y \in Y \mid y \text{ is mixed on } X_i\}$, and let $Y_0 = Y \setminus (Y_1 \cup \cdots \cup Y_m)$. By (3), the sets $Y_0, Y_1, \ldots, Y_m$ are pairwise disjoint and their union is Y. For each $i \in \{1, \ldots, m\}$, since G is prime, X_i is not a homogeneous set, so there exists a vertex in $V(G) \setminus X_i$ that is mixed on X_i; by (i), any such vertex is in Y, and so $Y_i \neq \emptyset$. Thus (ii) holds.

Now we prove (iii). Let Z be an anticomponent of Y, and suppose that $Z \not\subseteq Y_0$. So Z contains a vertex y from Y_i for some $i \in \{1, \ldots, m\}$; say $y \in Y_1$. Since X_1 is connected, there are adjacent vertices u_1 and v_1 in X_1 such that y is adjacent to u_1 and not to v_1. Consider any non-neighbor z of y in Z. By (ii), z has a neighbor x_2 in X_2, and y is complete to X_2. If z is anticomplete to $\{u_1, v_1\}$, then z, x_2, y, u_1, v_1 induce a P_5. If z is complete to $\{u_1, v_1\}$, then the same five vertices induce a $\overline{P_5}$. So z is mixed on X_1, i.e., $z \in Y_1$. Since Z is anticonnected, we can repeat this argument along the edges of a spanning tree of $\overline{G}[Z]$, which implies that $Z \subseteq Y_1$. Thus (iii) holds.

Now we prove (iv). Suppose on the contrary, and up to symmetry, that a vertex x in X_1 is mixed on some anticomponent Z of Y. Since Z is anticonnected, there are non-adjacent vertices $y, z \in Z$ such that x is adjacent to y and not to z. By (ii), z has a neighbor u in X_1, so $z \in Y_1$. Since X_1 is connected, there is a path u_0-$\cdots$-u_k in $G[X_1]$ with $u_0 = u$, $u_k = x$ and $k \geq 1$. Choose u such that k is minimal. By (ii), y has a neighbor x_2 in X_2, and since $z \in Y_1$, z is adjacent to x_2. If $k = 1$, then x, y, z, u, x_2 induce a C_5 or $\overline{P_5}$ (depending on the pair y, u). So $k \geq 2$. The minimality of k implies that z is not adjacent to u_1 or u_2, and u is not adjacent to u_2. Then x_2, z, u, u_1, u_2 induce a P_5, a contradiction.

Now we prove (v). Observe that any vertex t from a big component of $X \setminus X_i$ is complete to Y_i and anticomplete to X_i, so we can apply Lemma 3.2 to X_i, Y_i and t. It follows that some vertex a of X_i is complete to Y_i. By (ii), X_i is complete to $Y \setminus Y_i$. Thus a is complete to Y.

Now we prove (vi). Suppose that a vertex x in X_0 is mixed on two anticomponents Z_1 and Z_2 of Y. For each $j \in \{1, 2\}$, since Z_j is anticonnected, there are non-adjacent vertices y_j and z_j in Z_j such that x is adjacent to y_j and not to z_j. Then y_1, z_1, x, z_2, y_2 induce a $\overline{P_5}$, a contradiction.

Now we prove (vii). If Z is any big anticomponent of Y, then, since G is prime, Z is not a homogeneous set, so there exists a vertex of $V(G) \setminus Z$ that is mixed on Z. The definition of Z and (iv) imply that any such vertex is in X_0. So $X_Z \neq \emptyset$. The second sentence of (vii) follows directly from (vi).

Now we prove (viii). Let Z be a big anticomponent of Y. By (iii), Z is included in one of $Y_0, Y_1, \ldots, Y_m$. By (ii), some vertex t of $X \setminus X_0$ is complete to Z and anticomplete to X_Z. Hence we can apply Lemma 3.2 to Z, X_Z and t in the complementary graph $\overline{G}$, and we obtain that some vertex in Z is complete (in $\overline{G}$) to X_Z.

Finally we prove (ix). Suppose that Y is not a clique, and choose a big anticomponent Z of Y that minimizes the number of big anticomponents of Y that are not anticomplete to X_Z. If this number is 1, then Z satisfies the desired property. So suppose that this number is at least 2, that is, there is a vertex $x \in X_Z$ and a big anticomponent Z' of $Y \setminus Z$ that contains a neighbor of x. There are non-adjacent vertices $y, z \in Z$ such that x is adjacent to y and not to z. By (vi), x is complete to Z'. Consider any $t \in X_{Z'}$; there are non-adjacent vertices $y', z' \in Z'$ such that t is adjacent to y' and not to z'. If t has any neighbor in Z, then, by (vi), t is complete to Z, and then z, x, t, z', y' induce a $\overline{P_5}$, a contradiction. Since this holds for any $t \in X_{Z'}$, we obtain that $X_{Z'}$ is anticomplete to Z. Now the choice of Z implies that there is a third big anticomponent Z'' of Y (a big anticomponent of $Y \setminus (Z \cup Z')$) such that some vertex u of $X_{Z'}$ has a neighbor y'' in Z'' and X_Z is anticomplete to Z''. There are non-adjacent vertices $a, b \in Z'$ such that u is adjacent to a and not to b. Then a, b, u, x, y'' induce a $\overline{P_5}$, a contradiction. This completes the proof. $\square$

4 The split divide

A *split divide* of a graph G is a partition (A, B, C, L, T) of $V(G)$ such that:

- $|A| \geq 2$, A is complete to B and anticomplete to $C \cup T$, and some vertex of A is complete to L;
- L is a non-empty clique, every vertex of L is mixed on A, and L is complete to $B \cup C$;
- $|C| \geq 2$, some vertex of C is complete to B, and no vertex of C is mixed on any anticomponent of B;
- T is a stable set and is anticomplete to C.

Note that the sets B and T may be empty. The split divide can be thought of as a relaxation of the homogeneous set decomposition: a set $X \subseteq V(G)$ is a homogeneous set in G if no vertex in $V(G)\setminus X$ is mixed on X; in the case of the split divide, the set A is not homogeneous, but all the vertices that are mixed on A lie in the clique L, and adjacency between L and the rest of the graph is heavily restricted.

Theorem 4.1 *Let G be a prime $\{P_5, \overline{P_5}, C_5\}$-free graph. Then either G is a split graph or G or $\overline{G}$ admits a split divide.*

Proof. By Theorem 3.1 and up to complementation, we may assume that G contains a $\overline{C_4}$. Consequently G admits the structure described in Theorem 3.3, and we use it with the same notation. All items (i) to (ix) refer to Theorem 3.3.

Suppose that Y is a clique. Let $A = X_1$, $L = Y_1$, $B = Y \setminus Y_1$, $C = X_2 \cup \cdots \cup X_m$ and $T = X_0$. Then (A, B, C, L, T) is a split divide of G; this follows immediately from the definition of the partition $X_0, X_1, \ldots, X_m, Y_0, Y_1, \ldots, Y_m$, the fact that Y is a clique, and items (i)–(v).

Now suppose that Y is not a clique. By (ix), there is a big anticomponent Z of Y such that X_Z is anticomplete to all big anticomponents of $Y \setminus Z$. By (vii), $X_Z \neq \emptyset$. By (iii), and up to relabeling, we may assume that $Z \subseteq Y_0 \cup Y_1$. Hence Z is complete to $X_2 \cup \cdots \cup X_m$, and every vertex of $X_1 \cup (X_0 \setminus X_Z)$ is either complete or anticomplete to Z. Let K be the union of all anticomponents of Y of size 1. So K is a clique and is complete to $Y \setminus K$. Let:

$$\begin{aligned}
A &= Z; \\
L &= X_Z; \\
B &= \{x \in X_1 \cup (X_0 \setminus X_Z) \mid x \text{ is anticomplete to } Z\}; \\
C' &= \{x \in X_1 \cup (X_0 \setminus X_Z) \mid x \text{ is complete to } Z\}; \\
T &= \{k \in K \mid k \text{ has a neighbor in } X_Z\}; \\
C &= X_2 \cup \cdots \cup X_m \cup (Y \setminus (Z \cup T)) \cup C'.
\end{aligned}$$

We claim that:

$$\text{No vertex of } C \text{ is mixed on any component of } B. \tag{4}$$

For suppose that there is a vertex $c \in C$ and adjacent vertices $u, v \in B$ such that c is adjacent to u and not to v. Since X_0 is a stable set, we have $u, v \in \{x \in X_1 \mid x \text{ is anticomplete to } Z\}$. Since c is adjacent to u, we have $c \in (Y\setminus(Z\cup T))\cup\{x \in X_1 \mid x \text{ is complete to } Z\}$. Pick any $x \in X_Z$ and any vertex $z \in Z$ adjacent to x. Then x, z, c, u, v induce a P_5, a contradiction. Thus (4) holds.

$$T \text{ is complete to } C. \tag{5}$$

For suppose that there are non-adjacent vertices $t \in T$ and $c \in C$. Since K is complete to $Y \setminus K$ and $T \subseteq K$, we have that $c \notin Y \setminus (Z \cup T)$. Thus, $c \in X_2 \cup \cdots \cup X_m \cup C'$. By (ii), Y_0 and Y_1 are complete to $X_2 \cup \cdots \cup X_m$; since $Z \subset Y_0 \cup Y_1$, it follows that Z is complete to $X_2 \cup \cdots \cup X_m$. Thus, $X_2 \cup \cdots \cup X_m \cup C'$ is complete to Z, and so c is complete to Z. Further, since $X_2 \cup \cdots \cup X_m \cup C' \subseteq X \setminus X_Z$ and X_Z is anticomplete to $X \setminus X_Z$ (because $X_Z \subseteq X_0$), we know that c is anticomplete to X_Z. By the definition of T, t has a neighbor x in X_Z. There are non-adjacent vertices $y, z \in Z$ such that

x is adjacent to y and not to z. Since t and c are complete to Z, we see that t, c, y, z, x induce a $\overline{P_5}$, a contradiction. Thus (5) holds.

Now we observe that:

- $|A| \geq 2$ (because Z is big), A is anticomplete to B (by the definition of B), A is complete to $C \cup T$ (by (ii)), and some vertex of A is anticomplete to L (by (viii)).
- L is a non-empty stable set, every vertex of L is mixed on A, and L is anticomplete to $B \cup C$ (by the definition of L, with $L \subseteq X_0$).
- $|C| \geq 2$ (because $X_2 \subseteq C$), some vertex of C is anticomplete to B (every vertex of X_2 has this property), and no vertex of C is mixed on any component of B (by (4)).
- T is a clique and is complete to C (by (5)).

These observations mean that (A, B, C, L, T) is a split divide in $\overline{G}$. This completes the proof. $\square$

Let G be a graph that admits a split divide (A, B, C, L, T) as above, let a_0 be a vertex of A that is complete to L, and let c_0 be a vertex of C that is complete to B. Let $G_1 = G[A \cup B \cup \{c_0\} \cup L \cup T]$ and $G_2 = G[\{a_0\} \cup B \cup C \cup L \cup T]$. Then we consider that G is decomposed into the two graphs G_1 and G_2. Note that G_1 and G_2 are induced subgraphs of G and each of them has strictly fewer vertices than G since $|A| \geq 2$ and $|C| \geq 2$.

Split unification We can define a composition operation that "reverses" the split divide decomposition. Let A, B, C, L, T be pairwise disjoint sets, and assume that A and C are non-empty. Let a^*, c^* be distinct vertices such that $a^*, c^* \notin A \cup B \cup C \cup L \cup T$.

Let G_1 be a graph with vertex-set $A \cup B \cup L \cup T \cup \{c^*\}$ and adjacency as follows:

- L is a (possibly empty) clique;
- T is a (possibly empty) stable set;
- A is complete to B and anticomplete to T;
- Some vertex a_0 of A is complete to L;
- c^* is complete to $B \cup L$ and anticomplete to $A \cup T$.

Let G_2 be a graph with vertex-set $B \cup C \cup L \cup T \cup \{a^*\}$ and adjacency as follows:

- $G_2[B \cup L \cup T] = G_1[B \cup L \cup T]$;
- T is anticomplete to C;
- L is complete to $B \cup C$;
- a^* is complete to $B \cup L$ and anticomplete to $C \cup T$;

- Some vertex c_0 of C is complete to B, and no vertex of C is mixed on any anticomponent of B.

Under these circumstances, we say that (G_1, G_2) is a *composable pair*. The *split unification* of a composable pair (G_1, G_2) is the graph G with vertex-set $A \cup B \cup C \cup L \cup T$ such that:

- $G[A \cup B \cup L \cup T] = G_1 \setminus c^*$;
- $G[B \cup C \cup L \cup T] = G_2 \setminus a^*$;
- A is anticomplete to C in G.

Thus to obtain G from G_1 and G_2, we "glue" G_1 and G_2 along their common induced subgraph $G_1[B \cup L \cup T] = G_2[B \cup L \cup T]$, where $L \cup T$ induces a split graph (hence the name of the operation).

We say that a graph G is *obtained by split unification* provided that there exists a composable pair (G_1, G_2) such that G is the split unification of (G_1, G_2). We say that G is *obtained by split unification in the complement* provided that $\overline{G}$ is obtained by split unification. We now prove that every graph that admits a split divide is obtained by split unification from smaller graphs.

Theorem 4.2 *If a graph G admits a split divide, then it is obtained from a composable pair of smaller graphs (each of them isomorphic to an induced subgraph of G) by split unification.*

Proof. Let G be a graph that admits a split divide. Let (A, B, C, L, T) be a split divide of G, let a_0 be a vertex of A that is complete to L, and let c_0 be a vertex of C that is complete to B. Let $G_1 = G[A \cup B \cup L \cup T \cup \{c_0\}]$. Since $|C| \geq 2$, we have $|V(G_1)| < |V(G)|$. Let $G_2 = G[B \cup C \cup L \cup T \cup \{a_0\}]$. Since $|A| \geq 2$, we have $|V(G_2)| < |V(G)|$. Now (G_1, G_2) is a composable pair, and G is obtained from it by split unification. $\square$

The split unification can be thought of as generalized substitution. Indeed, we obtain the graph G from G_1 and G_2 by first substituting $G_1[A]$ for a^* in G_2, and then reconstructing the adjacency between A and L in G using the adjacency between A and L in G_1. We include T and c^* in G_1 in order to ensure that split unification preserves the property of being $\{P_5, \overline{P_5}, C_5\}$-free. In fact, we prove now something stronger than this: split unification preserves the (individual) properties of being P_5-free, $\overline{P_5}$-free, and C_5-free.

Theorem 4.3 *Let (G_1, G_2) be a composable pair and let G be the split unification of (G_1, G_2). Then, for each $H \in \{P_5, \overline{P_5}, C_5\}$, G is H-free if and only if both G_1 and G_2 are H-free.*

Proof. We use the same notation as in the definition of the split unification above. First suppose that G is H-free. Observe that G_1 is isomorphic to the induced subgraph $G[A \cup B \cup L \cup T \cup \{c_0\}]$, and G_2 is isomorphic to the induced subgraph $G[B \cup C \cup L \cup T \cup \{a_0\}]$. Hence G_1 and G_2 are H-free. Now suppose that G_1 and G_2 are H-free and that G contains an induced copy of H. Let W be a five-vertex subset of $V(G)$ such that $G[W] \simeq H$. We claim that W must

contain two non-adjacent vertices b and c with $b \in W \cap B$ and $c \in W \cap C$. For suppose the contrary. Then $W \cap C$ is complete to $W \cap (L \cup B)$ and anticomplete to $W \cap (A \cup T)$. If $|W \cap C| \geq 2$, then either $|W \cap C| \leq 4$, so $W \cap C$ is a proper homogeneous set in $G[W]$ (a contradiction since H is prime), or $W \subseteq C$, so W is isomorphic to an induced subgraph of G_2 (a contradiction since G_2 is H-free). So $|W \cap C| \leq 1$, and then W is isomorphic to an induced subgraph of G_1 (where c^* plays the role of the vertex in $W \cap C$ if there is such a vertex), a contradiction since G_1 is H-free. Therefore the claim holds. By a similar argument, W must contain two non-adjacent vertices a and ℓ with $a \in W \cap A$ and $\ell \in W \cap L$. Let w be the fifth vertex in W, so that $W = \{a, b, c, \ell, w\}$. By the definition of the split unification, a, b, ℓ, c induce a P_4 with edges $ab, b\ell, \ell c$. Consequently we must have one of the following two cases:
(i) W induces a P_5 or C_5. So w is anticomplete to $\{b, \ell\}$ and has a neighbor in $\{a, c\}$. Since w is anticomplete to $\{b, \ell\}$, it cannot be in A, B, L or C, so it is in T. But then w should be anticomplete to $\{a, c\}$.
(ii) W induces a $\overline{P_5}$. So w is adjacent to a and c and has exactly one neighbor in $\{b, \ell\}$. Since w is adjacent to a, it is not in $C \cup T$, and since it is adjacent to c, it is not in A. Moreover, since w is adjacent to exactly one of b and ℓ, it is not in L. So $w \in B$, and so it is adjacent to ℓ and, consequently, not to b. Hence b and w lie in the same anticomponent of B, and c is adjacent to exactly one of them, a contradiction (to the last axiom in the definition of a split unification).
□

5 The main theorem

In this section, we use Theorem 2.1 and the results of the preceding sections to prove Theorem 5.1, the main theorem of this paper.

Theorem 5.1 *A graph G is $\{P_5, \overline{P_5}\}$-free if and only if at least one of the following holds:*

- *G is a split graph;*
- *G is a pentagon;*
- *G is obtained by substitution from smaller $\{P_5, \overline{P_5}\}$-free graphs;*
- *G or $\overline{G}$ is obtained by split unification from smaller $\{P_5, \overline{P_5}\}$-free graphs.*

Proof. We first prove the "if" part. If G is a split graph or a pentagon, then it is clear that G is $\{P_5, \overline{P_5}\}$-free. Since both P_5 and $\overline{P_5}$ are prime, we know that the class of $\{P_5, \overline{P_5}\}$-free graphs is closed under substitution, and consequently, any graph obtained by substitution from smaller $\{P_5, \overline{P_5}\}$-free graphs is $\{P_5, \overline{P_5}\}$-free. Finally, if G or $\overline{G}$ is obtained by split unification from smaller $\{P_5, \overline{P_5}\}$-free graphs, then the fact that G is $\{P_5, \overline{P_5}\}$-free follows from Theorem 4.3 and from the fact that the complement of a $\{P_5, \overline{P_5}\}$-free graph is again $\{P_5, \overline{P_5}\}$-free.

For the "only if" part, suppose that G is a $\{P_5, \overline{P_5}\}$-free graph. We may assume that G is prime, for otherwise, G is obtained by substitution from smaller $\{P_5, \overline{P_5}\}$-free graphs, and we are done. If some induced subgraph of G is isomorphic to the pentagon, then by Theorem 2.1, G is a pentagon, and again we are done. Thus we may assume that G is $\{P_5, \overline{P_5}, C_5\}$-free. By Theorem 4.1, we know that either G is a split graph, or one of G and $\overline{G}$ admits a split divide. In the former case, we are done. In the latter case, Theorem 4.2 implies that G or $\overline{G}$ is the split unification of a composable pair of smaller $\{P_5, \overline{P_5}, C_5\}$-free graphs, and again we are done. $\square$

As an immediate corollary of Theorem 5.1, we have the following.

Theorem 5.2 *A graph is $\{P_5, \overline{P_5}\}$-free if and only if it is obtained from pentagons and split graphs by repeated substitutions, split unifications, and split unifications in the complement.*

Finally, a proof analogous to the proof of Theorem 5.1 (but without the use of Theorem 2.1) yields the following result for $\{P_5, \overline{P_5}, C_5\}$-free graphs.

Theorem 5.3 *A graph G is $\{P_5, \overline{P_5}, C_5\}$-free if and only if at least one of the following holds:*

- *G is a split graph;*
- *G is obtained by substitution from smaller $\{P_5, \overline{P_5}, C_5\}$-free graphs;*
- *G or $\overline{G}$ is obtained by split unification from smaller $\{P_5, \overline{P_5}, C_5\}$-free graphs.*

Acknowledgment

We would like to thank Ryan Hayward, James Nastos, Paul Seymour, and Yori Zwols for many useful discussions.

References

[1] M. Chudnovsky, P. Maceli. Ups and downs of the 4-edge path. Submitted for publication.

[2] M. Chudnovsky, P. Maceli, I. Penev. Excluding four-edge paths and their complements. Manuscript, 2013. ArXiv: http://arxiv.org/abs/1302.0405.

[3] M. Chudnovsky, N. Robertson, P. Seymour and R. Thomas. The strong perfect graph theorem. *Annals of Math.* 164 (2006) 51–229.

[4] L. Esperet, L. Lemoine, F. Maffray. On $(P_5, \overline{P_5})$-free graphs. Manuscript, 2013.

[5] S. Földes, P.L. Hammer. Split graphs. *Proc. 8th Southeastern Conf. on Combinatorics, Graph Theory and Computing, Congressus Numerantium* XIX, Utilitas Math. (1977) 311–315.

[6] S. Földes, P.L. Hammer. Split graphs having Dilworth number two. *Canadian Journal of Mathematics* 29 (1977) 666–672.

[7] J.-L. Fouquet. A decomposition for a class of $(P_5, \overline{P_5})$-free graphs. *Discrete Mathematics* 121 (1993) 75–83.

[8] M.C. Golumbic. *Algorithmic Graph Theory and Perfect Graphs.* 2nd edition, Annals of Discrete Mathematics 57, Elsevier, 2004.