The Spanning Galaxy Problem*

Daniel Gonçalves † Frédéric Havet‡ Alexandre Pinlou † Stéphan Thomassé †

August 23, 2009

Abstract

In a directed graph, a *star* is an arborescence with at least one arc, in which the root dominates all the other vertices. A *galaxy* is a vertex-disjoint union of stars. In this paper, we consider the SPANNING GALAXY PROBLEM of deciding whether a digraph D has a spanning galaxy or not. We show that although this problem is NP-complete (even when restricted to acyclic digraphs), it becomes polynomial-time solvable when restricted to strongly connected digraphs. We prove indeed that in the strongly connected case, the problem is equivalent to find a strong subgraph with an even number of vertices. As a consequence of this work, we improve some results concerning the notion of directed star arboricity of a digraph D, which is the minimum number of galaxies needed to cover all the arcs of D. We show in particular that $dst(D) \leq \Delta(D) + 1$ for every digraph D and that $dst(D) \leq \Delta(D)$ for every acyclic digraph D.

1 Introduction

All digraphs considered here are finite and loopless. An *arborescence* is a connected digraph in which every vertex has indegree 1 except one, called the *root*, which has indegree 0. A *diforest* is a disjoint union of arborescences. A *star* is an arborescence with at least one arc, in which the root dominates all the other vertices. A *galaxy* is a diforest of stars. A galaxy S in a digraph D is *spanning* if $V(S) = V(D)$.

In this paper, we study the complexity of the following decision problem:

SPANNING GALAXY PROBLEM
Instance: A digraph D.
Question: Does D have a spanning galaxy?

In Section 2, we show that the SPANNING GALAXY PROBLEM is linear-time solvable for arborescences. We also explore the relations between spanning galaxies and winning diforests for the parity game.

In Section 3, we prove that the SPANNING GALAXY PROBLEM is NP-complete for the class of acyclic digraphs.

A digraph $D = (V, A)$ is *strongly connected* or *strong* if for every pair $(u, v) \in V^2$ there is a directed path from u to v. It has been proved that deciding whether a strong digraph contains an even circuit is polynomial-time solvable [9, 10] although the problem is surprisingly NP-complete when restricted to circuits going through a given arc [11]. In Section 4, we show that the SPANNING GALAXY PROBLEM has similar characteristics when restricted to strong digraphs.

*This work was partially supported by the grant INRIA-Color 2008 PAGRO.

†Université Montpellier 2 - CNRS, LIRMM, 161 rue Ada, 34392 Montpellier Cedex 5, France. E-mail: [Daniel.Goncalves, Alexandre.Pinlou, thomasse]@lirmm.fr

‡Projet Mascotte I3S (CNRS & UNSA) and INRIA, INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 Sophia-Antipolis Cedex, France E-mail: Frederic.Havet@sophia.inria.fr. Partially supported by the european project FET-AEOLUS.

In Section 5, we explore some parameterizations of the Spanning Galaxy Problem.

Finally, in Section 6, using the notion of spanning galaxy we improve some results concerning directed star arboricity. The *directed star arboricity* of a digraph D, denoted by $dst(D)$, is the minimum number of galaxies needed to cover $A(D)$. This notion has been introduced in [7] and is an analog of the *star arboricity* defined in [1].

Let us denote by $\mu(G)$, the maximum multiplicity of an edge in a multigraph. By Vizing's theorem, one can colour the edges of a multigraph with $\Delta(G) + \mu(G)$ colours so that two edges have different colours if they are incident. Since the multigraph underlying a digraph has maximum multiplicity at most two, for any digraph D, $dst(D) \leq \Delta + 2$. Amini et al. [3] conjecture the following:

Conjecture 1 *Every digraph D with maximum degree $\Delta \geq 3$ satisfies $dst(D) \leq \Delta$.*

The condition $\Delta \geq 3$ in the above conjecture is necessary since the odd circuits have maximum degree 2 and directed star arboricity 3. This conjecture would be tight since every digraph with $\Delta = \Delta^-$ has directed star arboricity at least Δ. In [3], Amini et al. proved that Conjecture 1 holds when $\Delta = 3$.

A *nice galaxy* in a digraph G is a galaxy spanning all the vertices of maximum degree. To prove Conjecture 1, it suffices to show that every digraph with maximum degree $\Delta \geq 4$ has a nice galaxy.

Conjecture 2 *Every digraph with maximum degree $\Delta \geq 4$ has a nice galaxy.*

Amini et al. [3] showed the conjecture for 2-diregular digraphs. We prove Conjecture 2 for acyclic digraphs, which implies Conjecture 1 for acyclic digraphs. We also prove that every digraph has a galaxy spanning the vertices with indegree at least two and derive that $dst(D) \leq \Delta(D) + 1$ for every digraph D.

2 Spanning Galaxies and Winning Diforests

Proposition 3 *The* Spanning Galaxy Problem *is solvable in linear time when restricted to the class of arborescences.*

Proof. If an arborescence T has no vertices, it clearly admits a spanning galaxy. If T is however restricted to its root, it has none. Now if T has at least two vertices, we consider a furthest leaf v from r and we denote by u the inneighbour of v. By definition, all the outneighbours of u are leaves. Thus, if T admits a spanning galaxy, this galaxy contains the star with root u whose leaves are the outneighbours of u. Hence, T admits a spanning galaxy if and only if $T \setminus (\{u\} \cup N^+(u))$ does. This gives a simple linear-time algorithm for arborescences. □

The proof of the above proposition also implies the following lemma.

Lemma 4 *Every arborescence T contains a galaxy spanning every vertex except possibly the root.*

The *parity game* is a widely studied game. Its restriction to arborescences is played on an arborescence T (with root r) by two players, Player 0 and Player 1, as follows. At the beginning of a play, a token is placed on the root r and is then moved over $V(T)$ following the transitional relation: if the token is placed on a vertex v, then the next position of the token is one of the vertices of $N^+(v)$. The players move the token alternatively (starting with Player 0) until the token reaches a leaf. A player wins if its opponent cannot move anymore. Since our arborescences are finite, one of the two player has a winning strategy. If Player 0 has a winning strategy, we say that T is *winning*; otherwise, T is *loosing*. By convention, an arborescence T with zero vertices is winning.

Lemma 5 *An arborescence T admits a spanning galaxy if and only if T is winning.*

Proof. This directly follows from:

- An arborescence T with one vertex is loosing.
- Given an arborescence T with at least two vertices, where v is any furthest leaf from the root, and where u is the inneighbour of v; T is winning if and only if $T \setminus (\{u\} \cup N^+(u))$ is winning.

□

A diforest is *winning* if all its arborescences are winning, otherwise it is *loosing*. Since stars are winning arborescences we have that:

Lemma 6 *A digraph D admits a spanning galaxy if and only if D contains a winning spanning diforest.*

The directed path $P_l = (r, v_1, v_2, \ldots, v_l)$, admits a spanning galaxy if and only if l is odd (recall that the length of a path is its number of arcs). Given two arborescences T and T' and a vertex v of T, we denote by $T \vee_v T'$ the arborescence obtained by identifying v in T with the root of T'. When v is the root of T, we simply write $T \vee T'$. Observe that $T \vee T'$ is winning if and only if T or T' is winning. Similarly, if T' is loosing, then $T \vee_v T'$ is winning if and only if T is winning.

Thus, we have the two following lemmas, which we will use in Section 6.

Lemma 7 *Given any arborescence T and any odd l, the arborescence $T \vee P_l$ is winning.*

Lemma 8 *Given any arborescence T, any vertex v of T, and any even l, the arborescence $T \vee_v P_l$ is winning if and only if T is winning.*

3 Acyclic Digraphs

A digraph $D = (V, A)$ is *acyclic* if it does not contain any circuit.

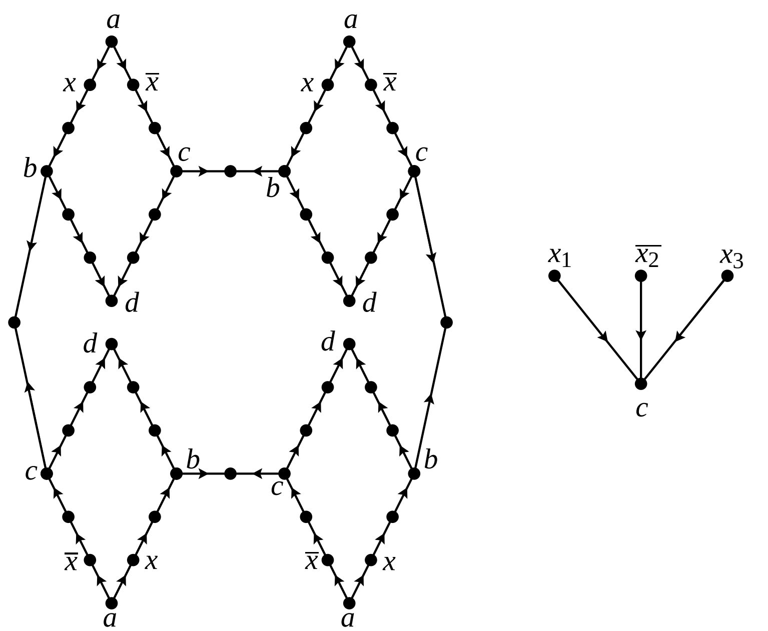
Theorem 9 *The* Spanning Galaxy Problem *is NP-complete, even if restricted to digraphs which are acyclic, planar, bipartite, subcubic, with arbitrary girth, and with maximum outdegree* 2*.*

Proof. This problem is clearly in NP and we prove now that it is NP-hard for this restricted family of digraphs. Kratochvíl proved that Planar $(3, \leq 4)$-Sat is NP-complete [8]. In this restricted version of Sat, the graph of incidence variable-clause of the input formula is planar, every clause is a disjunction of three literals, and every variable occurs in at most four clauses. We reduce Planar $(3, \leq 4)$-Sat to the Spanning Galaxy Problem. Given an instance I of Planar $(3, \leq 4)$-Sat, we construct a planar digraph D_I such that I is a satisfiable instance of Planar $(3, \leq 4)$-Sat if and only if D_I has a spanning galaxy. For this, we take one copy of the graph depicted in Figure 1(a) per variable of I, and one copy of the graph depicted in Figure 1(b) per clause of I. Whenever the literal x (resp. $\overline{x}$) appears in a clause c in I, we identify one vertex labelled x (resp. $\overline{x}$) of the variable gadget of x with a source of the clause gadget of c.

Let us observe that the digraph D_I is acyclic, planar, bipartite, subcubic, with maximum indegree 3 and with maximum outdegree 2.

The variable gadget of x in the graph D_I is connected to the rest of the graph by the vertices labelled by x or $\overline{x}$. The vertices which are not labelled by x or $\overline{x}$ are called *internal vertices* of the variable gadget of x. One can observe that there are only two possible galaxies that span all the internal vertices of a variable gadget. Actually, these two galaxies span all the vertices of the variable gadget. Moreover, in the first galaxy, every vertex x is the root of a star and every vertex $\overline{x}$ is a leaf of a star; in the second one, every vertex $\overline{x}$ is a root of a star and every vertex x is a leaf of a star.

In addition, one can observe that the previous remark, stating that the vertices x are roots of stars whenever the vertices $\overline{x}$ are leaves, holds for any odd paths linking a and b (resp. a and c, b and d, c and d). Therefore, the girth of the graph D_I can be made arbitrarily large.



(a) The gadget for the variable x.

(b) The gadget for the clause $c = (x_1 \vee \overline{x_2} \vee x_3)$.

Figure 1: The gadgets for Theorem 9.

Let I be an instance of PLANAR $(3, \leq 4)$-SAT.
Suppose first that I is satisfiable by some truth assignment ϕ.

Let us exhibit then a spanning galaxy of D_I. For every variable x, we span its gadget with a galaxy in such a way that the vertices labelled x are roots of stars if and only if $\phi(x) =$TRUE. In this way, we can span the internal vertices c of the clause gadgets. Indeed, since c is satisfied by ϕ, the vertex c has an inneighbour x_1 which is the root of a star. We then add the arc x_1c to our galaxy to span c.
Suppose now that D_I has a spanning galaxy T. Let ϕ be the truth assignment ϕ defined by $\phi(x) =$TRUE if and only if the vertices labelled x are roots of stars of T. Then ϕ satisfies I since every clause vertex c needs one of its inneighbours to be the root of some star. □

4 Strong Digraphs

Theorem 10 *It is NP-complete to decide, given a strong digraph and one of its arc, whether there exists a spanning galaxy containing this arc.*

Proof. The reduction from the SPANNING GALAXY PROBLEM in the acyclic case is straightforward. Given an acyclic digraph D, we construct D' from D by adding a disjoint directed path (a_1, a_2, a_3, a_4), all possible arcs from a_4 to a source of D, and all possible arcs from a sink of D to a_1. Note that D' is strong. Observe that D' has a spanning galaxy F containing the arc a_1a_2 if and only if D has a spanning galaxy. □

Note that it is also NP-complete to decide if a strong digraph has a spanning galaxy avoiding one prescribed arc. Indeed, in the previous proof, $a_2a_3 \notin A(F)$ if and only if D has a spanning galaxy.

Given a strong digraph D, a *handle* h of D is a directed path $(s, v_1, \ldots, v_\ell, t)$ from s to t (where s and t may be identical, or the handle possibly restricted to the arc st) such that:

- the vertices v_i satisfy $\mathrm{d}^-(v_i) = \mathrm{d}^+(v_i) = 1$, for every $1 \leq i \leq \ell$, and
- the digraph $D \setminus h$ obtained by *suppressing* the arcs and internal vertices of h is strongly connected.

The vertices s and t are the *endvertices* of h while the vertices v_i are its *inner vertices*. The vertex s is the *tail* of h and t its *head*. The *length* of a handle is the number of arcs in the path, here $\ell + 1$. A handle of length one is said to be *trivial*. For any $1 \leq i, j \leq \ell$, we say that v_i *precedes* (resp. *strictly precedes*) v_j on the handle h if $i \leq j$ (resp. $(i < j)$).

Given a strong digraph D, a *handle decomposition* of D starting at $v \in V(D)$ is a triplet $(v, (h_i)_{1\leq i\leq p}, (D_i)_{0\leq i\leq p})$, where $(D_i)_{0\leq i\leq p}$ is a sequence of strong digraphs and $(h_i)_{1\leq i\leq p}$ is a sequence of handles such that:

- $V(D_0) = \{v\}$,
- h_i is a handle of D_i, for $1 \leq i \leq p$ and D_i is the edge disjoint union of D_{i-1} and h_i, and
- $D = D_p$.

A handle decomposition is uniquely determined by v and either $(h_i)_{1\leq i\leq p}$, or $(D_i)_{0\leq i\leq p}$. The number of handles p in any handle decomposition of D is exactly $|A(D)| - |V(D)| + 1$. The value p is also called the *cyclomatic number* of D. Observe that $p = 0$ when D is a singleton and $p = 1$ when D is a circuit. A digraph D with cyclomatic number two is called a *theta*.

The following lemma is straighforward.

Lemma 11 *For every strong subdigraph D' of some strong digraph D, there is a handle decomposition $(v, (h_i)_{1\leq i\leq p}, (D_i)_{0\leq i\leq p})$ of D such that $D' = D_i$ for some i.*

A handle is *even* if its length is even. A handle decomposition is *even* if one of its handles is even. A strong digraph is *even* if it has an even number of vertices. Handles, handle decompositions and strong digraphs are *odd* when they are not even.

Theorem 12 *Given a strong digraph D, the following are equivalent:*

(1) D has a spanning galaxy.

(2) D contains a spanning winning arborescence.

(3) D has an even handle decomposition.

(4) D contains an even circuit or an even theta.

(5) D contains a even strong subgraph.

Proof. We prove the equivalence using six implications. The playful reader is invited to check that the implication digraph of this proof is an odd theta consisting of two circuits of length 3, hence it does not satisfy (4). This provides an illustration of Theorem 12, since this digraph does not satisfy any of the other properties.

(2) ⇒ (1) Consider a digraph D containing a spanning winning arborescence T. Lemma 5 implies that T contains a spanning galaxy, which also spans D.

(3) ⇒ (2) Let $(v, (h_i)_{1\leq i\leq p}, (D_i)_{0\leq i\leq p})$ be an even handle decomposition of D. Let q be the largest integer such that h_q is an even handle. Since D_{q-1} is strong, it contains a spanning arborescence T_{q-1} rooted at s_q, the first vertex of h_q. Now for every $q \leq r \leq p$, we define a spanning arborescence T_r of D_r as follows. For every $h_r = (s_r, v_1, \ldots, v_\ell, t_r)$, we let $T_r = T_{r-1} \vee_{s_r} P_r$ where P_r is the path $(s_r, v_1, \ldots, v_\ell)$, i.e. the handle h_r minus its last arc. By Lemma 7, the arborescence T_q is winning since $T_{q-1} \vee_{s_q} P_q$ is exactly $T_{q-1} \vee P_q$. Therefore, by Lemma 8, T_r is winning, for every $q \leq r \leq p$. Thus T_p is a spanning winning arborescence of D.

(1) ⇒ (3) By way of contradiction, suppose that there exists a strong digraph D with no even handle decomposition admitting a spanning galaxy. Observe that in particular, D has no even circuit. Choose such a D with minimum number of arcs. Let F be a spanning galaxy of D. Observe that every trivial handle st of D belongs to F, otherwise deleting the arc st from D leaves a strong digraph with no even handle decomposition, against the minimality of D.

Consider a handle decomposition $(v, (h_i)_{1\leq i\leq p}, (D_i)_{0\leq i\leq p})$ of D which minimizes the number of trivial handles. Let q be the largest integer such that $h_q = (v_0, \ldots, v_{\ell+1})$ is non trivial (here we adopt the notation $s := v_0$ and $t := v_{\ell+1})$). Hence, every handle h_i is trivial for every $q < i \leq p$. Moreover, since h_q is odd and non trivial, we have $\ell \geq 2$. Having chosen a minimal number of trivial handles in this decomposition, we have the following properties.

(i) there is no arc v_iv_j with $j \geq i + 2$, except possibly st;

(ii) for $2 \leq i \leq \ell$, the vertex v_i has no inneighbour in D_{q-1};

(iii) for $1 \leq i \leq \ell - 1$, the vertex v_i has no outneighbour in D_{q-1}.

In addition, the above observation implies that:

(iv) v_1 has no inneighbours in $D_{q-1} \setminus \{v_0\}$.

Indeed if u is such an inneighbour, both arcs uv_1 and v_0v_1 would be trivial handles of D. Hence, according to the previous observation, they both belong to F which is impossible.

Furthermore,

(v) there is no arc v_jv_i with $0 \leq i < j \leq \ell + 1$.

Such an arc v_jv_i is *short* if there is no arc $v_{j'}v_{i'}$ for which $i \leq i' < j' \leq j$ unless $(i, j) = (i', j')$. By way of contradiction, consider a short arc v_jv_i which minimizes i. By (i) and since there is no even circuit, the vertices $\{v_i, v_{i+1}, \ldots, v_j\}$ form an induced odd circuit. Moreover, since deleting the arc v_jv_i leaves D strongly connected, we have $v_jv_i \in F$. Hence there is at least one vertex in $X = \{v_{i+1}, \ldots, v_{j-1}\}$ which has a neighbour in $F \setminus X$. By (i), (ii), (iii) and the choice of v_jv_i, there is an arc $v_{j'}v_{i'}$ such that $i < i' < j < j'$. Let us consider such an arc $v_{j'}v_{i'}$ which minimizes i'. If $i' - i$ is odd then $(v_j, v_i, v_{i+1}, \ldots, v_{i'})$ is an even handle on the circuit $(v_{i'}, v_{i'+1}, \ldots, v_{j'})$, contradicting the fact that D has no even handle decomposition. If $i' - i$ is even then $X' = \{v_{i+1}, \ldots, v_{i'-1}\}$ has odd cardinality, and both arcs $v_{j'}v_{i'}$ and v_jv_i belong to F. Hence there must be a vertex in X' which has a neighbour in $F \setminus X'$, contradicting the definition of i'. This proves (v).

These properties imply that the only arc entering $S = \{v_1, \ldots, v_\ell\}$ is v_0v_1 and the only arcs leaving S are those leaving v_ℓ. Moreover $(v_0, v_1, \ldots, v_\ell)$ is an induced path. If $\{v_1v_2, v_3v_4, \ldots, v_{\ell-1}v_\ell\} \subseteq F$ then the digraph D_{q-1} would also be a counterexample, contradicting the minimality of D. Thus F contains the arcs $v_0v_1, v_2v_3, \ldots, v_{\ell-2}v_{\ell-1}$ and all the arcs leaving v_ℓ (by the earlier observation). Thus the digraph obtained from D by contracting v_0v_1 and v_1v_2 has a spanning galaxy and no even handle decomposition. This contradicts the minimality of D.

(3) ⇒ (4) By way of contradiction, suppose that there are digraphs with an even handle decomposition containing no even circuits or even thetas. Consider such a digraph D with an even handle decomposition $(v,(h_i)_{1\le i\le p},(D_i)_{0\le i\le p})$ minimizing p. It is clear, by minimality of p that the only even handle of this decomposition is h_p. Otherwise D_{p-1} would contradict the minimality of p.

In the remainder, we denote by s and t the tail and the head respectively of the handle h_p.

Claim 1 $p > 2$.

If $p = 1$ then D would be an even circuit. If $p = 2$, then h_1 has odd length and thus D would either be an even theta or contain an even circuit.

By Lemma 11, there is a handle decomposition $(s,(h'_i)_{1\le i\le p},(D'_i)_{0\le i\le p})$ of D starting at s and such that $h'_p = h_p$. For every $1 < i < p-1$, let us denote by s_i the tail of h'_i and by t_i its head.

The vertex t is an inner vertex of h'_{p-1} otherwise the digraph obtained from D by suppressing h'_{p-1} would contradict the minimality of p. Thus, we can divide h'_{p-1} into two subpaths: P with tail s_{p-1} and head t and Q with tail t and head t_{p-1}. Furthermore Q has odd length otherwise the digraph obtained from D by suppressing P would contradict the minimality of p.

Claim 2 *For every $1 < i < p$, the endvertices of $h'_i = (s_i,\dots,t_i)$ are inner vertices of h'_{i-1}.*

Suppose for a contradiction that the claim does not hold. Let q be the largest integer such that one of the two endvertices of h'_q is not an inner vertex of h'_{q-1}. One of the endvertices of h'_q is an inner vertex of h'_{q-1}. Otherwise h'_{q-1} would be a handle of D and the digraph obtained from D by suppressing h'_{q-1} would contradict the minimality of p. By directional duality, we may assume that s_q is an inner vertex of h'_{q-1} and t_q is not. Let us divide h'_{q-1} into two paths, the path R with tail s_{q-1} and head s_q and the path S with tail s_q and head t_{q-1}. Then S is a handle of D and the digraph obtained from D by suppressing h'_{q-1} contradicts the minimality of p (the handles h'_{q-1} and h'_q are replaced by a single handle with tail s_{q-1} and head t_q). This proves Claim 2.

Claim 3 *For every $1 < i < p$, the vertex t_i precedes s_i on h'_{i-1}.*

Suppose not. Then s_i stricly precedes t_i on h'_{i-1}. Let R be the subpath of h'_{i-1} with tail s_i and head t_i. Then R is a handle of D and the digraph obtained form D by suppressing R contradicts the minimality of p (the handles h'_{i-1} and h'_i are replaced by a single handle going from s_{i-1} to t_{i-1} containing h'_i). This proves Claim 3.

According to the Claim 3, the circuit h'_1 can be divided into two paths: P_1 with tail s_2 and head t_2 and P_2 with tail t_2 and head s_2. If s_2 and t_2 are identical, we assume that P_2 has no arc. Observe that P_1 is a handle of D which suppression leaves a digraph with an even handle decomposition and no even circuit or theta. The contradiction follows.

(4) ⇒ (5) Trivial since even circuits and thetas are strong digraphs with an even number of vertices.

(5) ⇒ (3) By Lemma 11 consider a handle decomposition $(v,(h_i)_{1\le i\le p},(D_i)_{0\le i\le p})$ of D such that some digraphs D_i have an even number of vertices. Let q be the smallest integer such that D_q has an even number of vertices. This implies that the handle h_q has an odd number of inner vertices, thus has even length. □

Theorem 13 SPANNING GALAXY *is polynomial-time solvable when restricted to strong digraphs.*

Proof. Actually, there exists a polynomial-time algorithm to decide whether a strong digraph contains an even strong subdigraph (ESS for short), which is by Theorem 12 equivalent to Spanning Galaxy for this class of graphs. The algorithm performs as follows. We first find a handle decomposition $(v,(h_i)_{1\le i\le p},(D_i)_{0\le i\le p})$ where $h_q=(x_0,x_1,\dots,x_\ell)$ is the last non-trivial handle. If there exists an arc with tail in $V(D_{q-1})$ and head in $\{x_2,\dots,x_{\ell-1}\}$, or with tail in $\{x_1,\dots,x_{\ell-2}\}$ and head in $V(D_{q-1})$, or an arc x_ix_j with $j>i+1$, then this arc is a trivial handle $h_{q'}$ with $q'>q$; one can easily find a new handle decomposition with less trivial handles. Since this operation is done in constant time and since the initial number of trivial handles is polynomial, one can compute in polynomial time a handle decomposition where there is no such trivial handles. If the decomposition has an even handle then return "YES" thanks to Theorem 12. We can then suppose in the remainder that ℓ is odd. Let D' be the digraph obtained from D_{q-1} by adding all the arcs between $N^-_D(x_1)$ and $N^+_D(x_{\ell-1})$. Let $S=\{x_1,x_2,\dots,x_{\ell-1}\}$ be the set of inner vertices of h_q.

Claim 4 *D has an ESS if and only if $D[S]$ has an ESS or D' has a ESS.*

Since every subdigraph of $D[S]$ is a subdigraph of D, if $D[S]$ has an ESS, then D has an ESS. Therefore, suppose that $D[S]$ does not have an ESS. For every ESS E of D, the digraph $E'=D'[V(E)\setminus S]$ is an ESS of D'. Indeed, if $|V(E)|\neq|V(E')|$, the handle h_q is a subdigraph of E; however h_q is odd so $|V(h_q)|$ is even, and thus $|V(E')|=|V(E)|-|V(h_q)|$ is also even. Furthermore, since the paths of E from $N^-_D(x_1)$ to $N^+_D(x_{\ell-1})$ are replaced by single arcs in E', E' is strong. Finally, it is also clear that given any ESS F' of D' one of the graphs $D[V(F')]$ or $D[V(F)\cup S]$ is an ESS of D – according whether or not there is an arc uv in $A(F)\cap\left(N^-_D(x_1)\times N^+_D(x_{\ell-1})\right)$.

Checking if $D[S]$ has an ESS can be done in polynomial time. We first check if there exists a backward arc (i.e. an arc x_bx_a such that $a<b$) such that a and b have distinct parity. If there is such an arc, the graph $D[\{x_a,x_{a+1},\dots,x_b\}]$ is an ESS.

If there exists no such arc, we distinguish two types of backward arcs x_bx_a of $D[S]$: the arcs where a and b are both even, called *e-arcs*, and those where a and b are both odd, called *o-arcs*. Observe that the vertex set of an ESS F of $D[S]$ is of the form $\{x_i,x_{i+1},\dots,x_j\}$. Indeed, since there is no arc x_ax_b with $a+1<b$ in $D[S]$ and since there is a path from the vertex with smaller index in $V(F)$ (here x_i) to the one with higher index (here x_j), all the vertices between x_i and x_j are in F.

Furthermore since F is even, i and j have distinct parity. Consider a set A of backward arcs such that the union of the directed path $\{x_i,x_{i+1},\dots,x_j\}$ and A is strongly connected, and moreover such that A is minimum with respect to inclusion. Such a set A can be obtained from the set of backward arcs of F by deleting greedily some arcs. The arcs of A, when ordered increasingly according to the index of their tail, are such that two consecutive arcs x_cx_a and x_dx_b satisfy $a<b<c<d$. Note that since i and j have distinct parity, there exists two consecutive backward arcs of distinct types (one is an e-arc and the other one is an o-arc). Thus at this stage, $D[S]$ contains an ESS if and only if it contains an ESS with exactly two backward arcs. Hence this can be checked in polynomial time.

In the case of D', we check whether it contains an ESS or not by applying the original algorithm recursively. □

5 Parameterizations of galaxy problems

The spanning galaxy problem being hard in the general case, it is natural to ask if some parameterized version is tractable. A first attempt could be to ask for a fixed parameter tractable algorithm on parameter k (i.e. admitting an algorithm in time $O(f(k)n^c)$ for some constant c) deciding if a digraph admits a spanning galaxy with at most k stars. Unfortunately, the problem k-Domination (which is

W[2]-complete [6]) admits a straightforward reduction to this problem. Indeed, every minimal dominating set A of a graph G (with no isolated vertex) corresponds to the set of roots of a spanning galaxy of the digraph D obtained from G by replacing each edge $ab \in E(G)$ by the arcs ab and ba. Hence this galaxy problem is at least as hard as k-DOMINATION, thus it is $W[2]$-hard.

However, the following problem is easier to handle:

Problem 14 (k-Galaxy)
INSTANCE: A digraph D.
QUESTION: Does D have a galaxy spanning at least k vertices?

This problem is very easily fixed parameter tractable, but we will show a much stronger result. Indeed, there is a polynomial algorithm (in size of D) which tramsforms every instance (D, k) of k-GALAXY into an instance (D', k') which is equivalent to (D, k) and such that D' has at most $2k - 2$ vertices. This algorithm is called a *kernelization algorithm*, and the output D' is called a *kernel*. Observe that applying a brute force algorithm on D' to check if it admits a galaxy spanning at least k' vertices takes $O(f(k))$ time. Hence the existence of the kernelization algorithm gives an FPT algorithm for k-GALAXY running in $O(f(k) + n^c)$ time.

A galaxy F of D is *locally maximal* if it satisfies the following conditions:

(a) The vertices of $V(D) \setminus V(F)$ form a stable set.

(b) If uv is an arc of F and uw is an arc of D, we have $w \in V(F)$.

(c) If $u \in V(F)$ and $uv, uw \in A(D)$, at least one of v and w belong to $V(F)$.

(d) If $uv, uw \in F$ and $wx \in A(D)$, then $x \in V(F)$.

If a galaxy F does not satisfy one of the previous conditions, one can easily find a galxy spanning more vertices than F. Hence one can compute a locally maximal galaxy G in polynomial time. If G spans at least k vertices, then we are done.

The conditions (b), (c) and (d) imply that the set $N^+_G = \{v \in V(D) \setminus V(G) \mid \exists u \in V(G), uv \in A(D)\}$ has cardinality at most $\frac{|V(G)|}{2}$. Now let $N^-_G = V(D) \setminus (V(G) \cup N^+_G)$. Note that (a) implies that N^-_G is a stable set of D. Note also that N^+_G may contain in-neighbors of $V(G)$ while N^-_G does not contain out-neighbors of $V(G)$. Consider now a maximum matching M (polynomially computable) in the bipartite subdigraph of D induced by the sets $V(G)$ and N^-_G. Observe that this matching M has at most $\frac{|V(G)|}{2}$ arcs otherwise we would choose M as a galaxy spanning more vertices than G.

Lemma 15 *The digraph D has a galaxy of maximum size contained in the subgraph D' induced by $V(G) \cup N^+_G \cup V(M)$.*

Proof. Consider for contradiction a galaxy G^* of D such that D' does not contain a galaxy spanning $|V(G^*)|$ vertices. Among the possible choices of G^*, select one which minimizes its number of vertices in $N^-_G \setminus M$, and then which minimizes its number of arcs between $N^-_G \setminus M$ and $V(G)$. Since $G^* \not\subset D'$, G^* has a vertex $u \in N^-_G \setminus M$, and thus G^* has an arc uv_1 with $v_1 \in V(G)$. Since $uv_1 \notin M$ there is an arc u_1v_1 in M. We inductively define the vertices u_i and v_i, for $i \geq 2$, as follows. If u_{i-1} does not belong to G^* then u_j and v_j are not defined for $j \geq i$. Otherwise, let v_i be any vertex such that $u_{i-1}v_i$ is an arc of G^*. Note that $v_i \in V(M)$, otherwise the path $(v_i, u_{i-1}, v_{i-1}, \dots, u_1, v_1, u)$ would be an augmenting path with respect to M, contradicting the maximality of M. Thus let u_i be the vertex such that $u_iv_i \in M$. Let t be the greater index such that the vertices u_t and v_t are defined. Since $u_t \notin V(G^*)$ we can replace the arcs uv_1 and $u_{i-1}v_i$, for $2 \leq i \leq t$, by the arcs u_iv_i, for $1 \leq i \leq t$. Note that since u_t was not previously spanned, the obtained galaxy spans at least as many vertices as G^* but covers more arcs of M. Thus the lemma holds. □

Since the digraph D' has at most $2.|V(G)|$ vertices, we have our kernel of size at most $2k-2$.

6 Directed Star Arboricity

6.1 Acyclic digraphs

In this subsection, we settle Conjecture 2 for acyclic digraphs and derive that Conjecture 1 holds for acyclic digraphs. To do so, we need the following lemma on mixed graphs. A *mixed graph* is the disjoint union of odd circuits and a matching.

Lemma 16 *Every graph has a mixed subgraph spanning all the vertices of maximum degree.*

Proof. Let G be a graph of maximum degree Δ and V_Δ be the set of vertices of degree Δ. The result holds trivially if $\Delta = 1$ so we may assume that $\Delta \geq 2$. Let H be a mixed subgraph that spans the maximum number of vertices of V_Δ. Let $C_1, \ldots, C_p$ be the odd circuits of H and M its matching. Suppose by way of contradiction that there is a vertex v in $V_\Delta \setminus V(H)$. An *alternating v-path* is a path starting at v such that every even edge is in M (and so every odd edge is not in M). Let A_0 (resp. A_1) be the set of vertices u such that there exists a v-alternating path of even (resp. odd) length ending at u. Note that $v \in A_0$ as (v) is an alternating v-path of length 0.

Claim 5 $A_0 \subset V_\Delta$.

Suppose that $A_0 \not\subset V_\Delta$. Then there is a vertex $x \in A_0 \setminus V_\Delta$. Let P be the even alternating v-path ending at x. Then the mixed subgraph obtained from H by replacing the matching M by $M' = M \triangle P$ spans one more vertex of V_Δ, namely v, than H. This is a contradiction.

Claim 6 $A_1 \subseteq V(H)$.

Suppose by way of contradiction that a vertex $x \in A_1$ is in $V(G) \setminus V(H)$. Let P be an odd alternating v-path ending at x. Then the mixed subgraph obtained form H by replacing the matching M by $M' = M \triangle P$ spans one more vertex of V_Δ, namely v, than H. This is a contradiction.

Claim 7 $A_1 \subseteq V(M)$.

Suppose by way of contradiction that a vertex $x \in A_1$ is in $\bigcup_{i=1}^{p} C_i$, say in C_p. Then $C_p - x$ has a matching M_1. Let P be an odd alternating v-path ending at x. This path of odd length has a perfect matching $M_2 = P \setminus M$. Thus the disjoint union of $C_1, \ldots, C_{p-1}$ and $(M \setminus P) \cup M_1 \cup M_2$ is a mixed subgraph spanning more vertices of V_Δ than M. This is a contradiction.

Claim 8 $|A_0| = |A_1| + 1$.

Indeed, M matches every vertex of A_0, except v, with a vertex of A_1, and vice versa.

Claim 9 A_0 *is a stable set.*

Suppose to the contrary that there exist two adjacent vertices x and y in A_0. Let P_x and P_y be two even alternating v-path ending at x and y, respectively. We choose x, y, P_x and P_y in such a way that $|V(P_x) \cup V(P_y)|$ is minimum. Note that P_x P_y may share common vertices and arcs at the beginning. If $xy \in M$, then x is the predecessor of y in P_y and vice-versa. In this case let $Q_y = P_x - y$ and $Q_x = P_y - x$.

Otherwise let $Q_x = P_x$ and $Q_y = P_y$. In both cases, Q_x and Q_y are alternating v-paths of same parity. Note that by minimality of $|V(P_x) \cup V(P_y)|$ there exists only one vertex $z \in V(Q_x) \cap V(Q_y)$ (possibly $z = v$) and three paths Q_{v-z}, Q_{z-x} and Q_{z-y}, going respectively from v to z, from z to x and from z to y such that $Q_x = Q_{v-z} \cup Q_{z-x}$, $Q_y = Q_{v-z} \cup Q_{z-y}$, and $V(Q_{z-x}) \cap V(Q_{z-y}) = \{z\}$. Note that we necessarily have $z \in A_0$ since every odd vertex in Q_x and Q_y is followed by its neighbour in M. Let C_{p+1} be the odd circuit formed by the paths Q_{z-x} and Q_{z-y}, and by the edge xy. Then the mixed subgraph obtained from H by replacing the matching M by $M' = M \triangle Q_{v-z}$ and adding the odd circuit C_{p+1} spans one more vertex of V_Δ than H. This is a contradiction.

By Claim 9, all the edges with an end in A_0 have the other end in A_1 and thus, by Claims 5 and 8, there are $|A_0| \times \Delta = (|A_1| + 1) \times \Delta$ edges between A_0 and A_1. This is impossible because the vertices in A_1 have maximum degree Δ. □

Theorem 17 *Every acyclic digraph has a nice galaxy.*

Proof. Let D be an acyclic digraph and G its underlying undirected graph D. By Lemma 16, G has a mixed subgraph H spanning all the vertices of maximum degree. The subdigraph D' of D which is an orientation of H is the union of oriented odd circuits and a matching. Each oriented circuit is not directed because D is acyclic and thus has a spanning galaxy. Thus D' has a spanning galaxy, which is a nice galaxy of D. □

Corollary 18 *If D is an acyclic digraph then $dst(D) \le \Delta(D)$.*

Proof. We prove the result by induction on $\Delta(D)$, the result holding trivially when $\Delta(D) = 1$. Suppose now $\Delta(D) = k > 1$. By Theorem 17, D has a nice galaxy F_k. Hence $D' = D \setminus E(F_k)$ has maximum degree at most $k-1$. By induction, D' has an arc-partition into $k-1$ galaxies $F_1, \dots, F_{k-1}$. Thus $(F_1, \dots, F_k)$ is an arc-partition of D into k galaxies. □

6.2 Galaxy spanning the vertices with indegree at least two

The *outsection* of a vertex x is the set $S^+(x)$ of vertices y to which there exists a directed path from x. An *outgenerator* of D is a vertex $x \in V(D)$ such that $S^+(x) = V(D)$. Note that if D is strong, every vertex is an outgenerator. Every outgenerator is the root of a spanning arborescence, so by Lemma 4 we get the following:

Corollary 19 *Let v be an outgenerator of a digraph D. Then D contains a galaxy F spanning all the vertices of $D - v$.*

Theorem 20 *Every digraph D has a galaxy spanning all the vertices with indegree at least 2.*

Proof. We prove the result by induction on the number of vertices. Free to remove arcs entering vertices of indegree 1 or more than 2, we may assume that every vertex of D has indegree 2 or 0. Suppose first that D contains a vertex v of indegree 0. Set $D^+ = D[S^+(v)]$ and $D' = D - D^+$. By definition of outsection, there are no arcs leaving D^+. So the vertices of D' have the same indegree in D' and D. By the induction hypothesis, there is a galaxy F' spanning all the vertices of D' with indegree 2 and by Corollary 19, there is a galaxy F^+ spanning all the vertices of D^+ with indegree 2. The union of F' and F^+ is the desired galaxy.

Suppose now that all the vertices of D have indegree 2. Consider an initial strong component C of D. Let us recall that every strong digraph with minimum indegree two has a vertex which deletion leaves the digraph strong. Consider for this a handle decomposition minimizing the number of trivial handles. Let $x_0, \ldots, x_l$ be the last non-trivial handle. The vertex x_{l-1} has indegree at least two, hence the other in-arcs entering x_{l-1} are trivial handles. If l is greater than 2, any of these trivial handle, together with $x_0, \ldots, x_l$ would result in two non-trivial handles - which is impossible by assumption. Thus $l = 2$, and then the vertex x_1 can be deleted.

In particular, there exists a vertex v of C such that $C - v$ is strong. Let S^+ be the outsection of v in $D - (C \setminus \{v\})$ and $T = S^+_D(v) \setminus S^+$ and $D' = D - S^+_D(v)$. Note that v is an outgenerator of $D[S^+]$ and $D_1 = D[T \cup \{v\}]$. Moreover since $C - v$ is strong, every vertex of $C - v$ is an outgenerator of $D_2 = D[T]$.

By the induction hypothesis, there is a galaxy F' spanning all the vertices of D' with indegree 2. By Corollary 19, there is a galaxy F^+ of $D[S^+]$ spanning all the vertices of $S^+ \setminus \{v\}$ in which v is either not spanned or a root. If v is a root of F^+ then, by Corollary 19, there is a galaxy F_1 of D_1 spanning all the vertices of T in which v is either not spanned or a root. The union of F', F^+ and F_1 is a spanning galaxy of D. If v is not a root of F^+, let u be an inneighbour of v. By Corollary 19, there is a galaxy F_2 of D_2 spanning all the vertices of $T \setminus \{u\}$ in which u is either not spanned or a root. The union of F', F^+, F_2 and the arc uv is a spanning galaxy of D. □

Note that Theorem 20 implies Amini et al. result [3] showing that a 2-diregular digraph has a spanning galaxy.

Theorem 21 *Let D be a digraph with maximum degree $\Delta \geq 2$. Then $dst(D) \leq \Delta + 1$.*

Proof. Set $D_0 = D$ and for every i from 1 to $\Delta - 2$, let F_i be a galaxy spanning all the vertices of indegree at least 2 in D_{i-1} and $D_i = D_{i-1} \setminus E(F_i)$. Observe that a vertex of $D' = D_{\Delta-2}$ has either indegree at most one or indegree 2 and outdegree 0. Now we just have to prove that $dst(D') \leq 3$. For this, choose one in-arc for each vertex with indegree two and denote the set of these arcs by F. The graph $D' - F$ is a loopless functional digraph (i.e. every vertex has indegree exactly 1). Consider a 3-colouring of the arcs of $D' - F$ such that two incident arcs get different colours. The crucial fact is that every arc xy of F can get three colours. One is forbidden by the other in-arc of x and another by the in-arc of y. Hence there is a colour left to extend the 3-colouring into three galaxies. □

References

[1] I. Algor and N. Alon, The star arboricity of graphs, *Discrete Mathematics* **75** (1989), 11–22.

[2] N. Alon, C. McDiarmid, B. Reed, Star arboricity, *Combinatorica* **12** (1992), 375–380.

[3] O. Amini, F. Havet, F. Huc and S. Thomassé, WDM and directed star arboricity, *manuscript*.

[4] J. Bang-Jensen and G. Gutin, *Digraphs: theory and applications*.

[5] R. Brandt, *Multicasting using WDM in Multifiber Optical Star Networks*, Thesis, UCSB, September 2003.

[6] R. Downey and M. Fellows, Fixed parameter tractability and completeness I: basic results, *SIAM J. Comput.*, **24** (1995), 873–921.

[7] B. Guiduli, On incidence coloring and star arboricity of graphs, *Discrete Mathematics* **163** (1997), 275–278.

[8] J. Kratochvíl, A special planar satisfiability problem and a consequence of its NP-completeness, *Discrete Applied Mathematics* **52(3)** (1994), 233–252.

[9] W. McCuaig, Pólya's permanent problem, *Electron. J. Combin.* **11(1)** (2004), Research Paper 79, 83 pp. (electronic).

[10] N. Robertson, P. D. Seymour and R. Thomas, Permanents, Pfaffian orientations, and even directed circuits, *Ann. of Math. (2)* **150(3)** (1999), 929–975.

[11] C. Thomassen, Even cycles in directed graphs, *European J. Combin.*, **6(1)** (1985), 85–89.