A Characterization of Partial Directed Line Graphs*

Nicola Apollonio† Paolo G. Franciosa‡

Abstract

Can a directed graph be completed to a directed line graph? If possible, how many arcs must be added? In this paper we address the above questions characterizing partial directed line (PDL) graphs, i.e., partial subgraph of directed line graphs. We show that for such class of graphs a forbidden configuration criterion and a Krausz's like theorem are equivalent characterizations. Furthermore, the latter leads to an recognition algorithm that requires $O(m)$ worst case time, where m is the number of edges in the graph. Given a partial line digraph, our characterization allows us to find a minimum completion to a directed line graph within the same time bound.

The class of partial directed line graphs properly contains the class of directed line graphs, characterized in [1], hence our results generalizes those already known for directed line graphs. In the undirected case, we show that finding a minimum line graph edge completion is NP-hard, while the problem of deciding whether or not an undirected graph is a partial graph of a simple line graph is trivial.

1 Introduction

Line graphs and adjoint graphs are probably two of the most well known classes of intersection graph models ([2, 4]). Line graphs are defined as the intersection graphs of the set of edges of undirected graphs. They have as vertex set the edges of a given simple undirected graph, and there is an edge between two vertices in the line graph if the corresponding edges are adjacent. Adjoint graphs have as node set the arc set of a given "root" graph and there is an arc between two nodes if and only if the corresponding arcs are "consecutive" in the root graph. More precisely, if xy is an arc of the adjoint graph, then, in the root graph, the head of the arc corresponding to x coincides with the tail of the arc corresponding to y. (Adjoint graphs are the intersection graphs of the dual of the hypergraph having as hyperedges the family of all pairs of consecutive arcs). Adjoint graphs can be even more generally defined for bidirected graphs, and their study has been shown fruitful in connection with Boolean Optimization (see [6]).

Line graphs have been characterized by Krausz (see Theorem 8.1 below), Van Rooij and Wilf and Beineke ([2], pag. 110). In particular, Beineke's characterization is based on nine forbidden induced subgraphs.

Adjoint graphs are those graphs satisfying the so called Duchenne's Condition (see Section 6), after Duchenne's characterization. Equivalently, adjoint graphs of directed graphs are those graphs not containing any subgraph in Figure 3 with the dotted arcs missing.

In this paper we deal with line graphs and directed line graphs.

Directed line graphs are adjoint graphs of directed 1-graphs (i.e., a directed graph with no parallel arcs, in particular at most one loop is allowed at each node). Directed line graphs have been studied in the past decades by several authors [1, 3, 5] in connection e.g., with problems arising in DNA

*Partially supported by the Italian MIUR Project ALGO-NEXT "Algorithms for the Next Generation Internet and Web: Methodologies, Design and Experiments".

†Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università di Roma "La Sapienza", piazzale Aldo Moro 5, I-00185 Roma, Italy. nicola.apollonio@uniroma1.it

‡Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università di Roma "La Sapienza", piazzale Aldo Moro 5, I-00185 Roma, Italy. paolo.franciosa@uniroma1.it

sequencing and the design of interconnection networks. Furthermore, they or their iterations, have nice properties such as high connectivity and small diameter (see [3, 5] and references cited therein). Line graphs and directed line graphs have a number of nice features. A remarkable property of directed line graphs is perhaps that the Hamiltonian Cycle Problem reduces to the Eulerian Cycle Problem (see [1]). As for line graphs, the most famous one is perhaps the polynomial-time solvability of the Maximum Stable Set Problem by a reduction to the Matching Problem. Directed line graphs have been characterized in [1] as those adjoint graphs not containing as subgraphs any of those in Figure 3, or equivalently as those graphs satisfying condition (7), that specializes Duchenne's condition. In [1] it was also shown that directed line graphs can be recognized in $O(n^3)$ time, where n is the number of nodes.

The property of being a directed line graph is preserved under taking induced subgraphs, but it is not inherited by more general subgraphs. For example, removing an edge from an adjoint graph could lead to a graph that is no longer the adjoint of any graph.

In this paper we define the class of partial directed line graphs (PDL graphs for short). A graph G is a PDL graph if G is a partial subgraph of a directed line graph. So the class of PDL graphs is closed under taking *subgraphs*, while the class of directed line graphs is only closed under taking *induced subgraphs*. Clearly, the class of PDL graphs properly contains the class of directed line graphs.

Let us put things in a more precise way. Suppose we are given a directed graph G and we look for a directed 1-graph R with the following property: arcs of R correspond bijectively to nodes of G and whenever two nodes x and y are adjacent in G and x precedes y the head of the arc corresponding to x in R coincides with the tail of arc y. Two arcs can be consecutive in R but the corresponding nodes in G could be not adjacent. If such a graph exists we call it a *weak root* of G (or simply *root*, where no confusion arises). In other words, we have just relaxed the correspondence that associates with a given graph its directed line graph. In our case such a correspondence is not a bijection.

Clearly, if G has a weak-root R, G is a partial graph of the directed line graph of that root. Conversely, if G is a partial graph of some directed line graph G', the root of the subgraph of G' induced by the nodes of G, is a weak-root for G. It follows that deciding if a graph has a weak root is tantamount to decide if the graph is a partial graph of some directed line graph.

Not every 1-graph can be completed to a directed line graph. Take for instance one of the graph in Figure 2 (d)÷(f). Such graphs cannot occur as subgraphs in any directed line graph (see [1, 5]). In particular they cannot occur as a subgraph in any partial directed line graph. So the following recognition problem makes sense.

Problem 1 [Partial directed line graph Recognition]. *Given a 1-graph G decide if it is a partial line graph.*

Once we have decided that a graph $G = (V, E)$ is a partial directed line graph, we are interested in the directed line graph completion problem.

Problem 2 [Minimum directed line graph Completion]. *Given a a partial directed line graph G find a minimum cardinality set of arcs E' such that $G = (V, E \cup E')$ is a directed line graph.*

In this paper we provide some characterizations of the class of PDL graphs. We study the case in which the root graph is allowed to contain self-loops and symmetric pairs, and the case in which the root graph is required to be loopless and antisymmetric. These characterizations lead us to provide a simple algorithms for deciding whether a graph G is a PDL graph, requiring $O(m)$ time, where m is the number of arcs in G. It is immediate to see that our algorithm is asymptotically optimal. The same algorithm also provides a root graph R_G of G. The directed line graph of R_G determines a minimum completion of G to a directed line graph. We also show that any possible completion to a directed line graph must contain our minimum completion.

When the minimum completion is found to be an empty set, our algorithm improves the recognition algorithm in [1] by a factor $\Omega(n)$ (actually, a factor $\Theta(n^3/m)$). Our results strongly rely on the notion of alternating path between a pair of nodes (x, y), that is a path from x to y whose arcs are alternatively oriented forward and backward.

Having recognized PDL graphs and solved the minimum completion problem for directed graphs, it is natural to wonder whether or not similar results hold for undirected graphs. Surprisingly the complexity status of the two problems (recognition and completion) goes in opposite directions; from the one hand the property of being a partial graph of the line graph of a undirected graph is trivial: every simple undirected graph is a partial graph of the complete graph on the same set of vertices, which is the line graph of a star. On the other hand, we show that finding a minimum cardinality set of edges whose addition causes the graph to be a line graph is an NP-hard problem (see Theorem 8.2).

The rest of the paper goes as follows. In Section 2 we briefly recall some graph terminology. Section 3 is devoted to the study of properties of alternating paths. In Section 4 we identify the basic component of PDL graphs, and their role in building root graphs is shown in Section 5, where Problem 1 is solved.

Our characterization is restricted to directed line graphs in Section 6 where we also solve Problem 2. Section 7 shows the recognition algorithm. Finally, in Section 8, we study the minimum line graph completion problem.

2 Definitions

Notation and terminology used throughout is mostly standard. If G is a (undirected or directed) graph we use the symbol $V(G)$ both for its vertex set (if it is undirected) and for its node set (if it is directed). Similarly, $E(G)$ denotes both the edge set of an undirected graph and the arc set of a directed graph. If G is undirected and $x, y \in V(G)$, the edge having x and y as endpoints is denoted by xy. If x and y are nodes of a directed graph G, let xy denote the arc leaving x and entering y[1]. The symbol yx will denote the arc leaving y and entering x. An arc xx is called a loop, and we say that there is a loop at x. For an arc xy of a directed graph we say that the x is its *tail* and that y is its *head*. Head and tail of arc $e \in E(G)$ are also denoted by $h(e)$ and $t(e)$, respectively. Two arcs e and e' are said to be *consecutive*, if either $h(e) = t(e')$ and in this case we say that e *precedes* e' or $h(e') = t(e)$ and we say that e *follows* e'. If G is a graph and $x \in V(G)$, let $\deg_G(v)$ denote the number of edges (or arcs) incident in v. If G is directed $\deg^+_G(v)$ and $\deg^-_G(v)$ denote the set of arcs leaving x (i.e., having x as tail) and entering x (i.e., having x as head), respectively. Let G be a graph and $U \subseteq V(G)$: the subgraph induced by U is the graph $G[U]$ having $V(G[U]) = V$ and $E(G[U]) = E(G) \cap (U \times U)$. Let x be a node in a directed graph G. We say that x is a *source* in G if no arc of G enters x; x is said to be a *sink* in G if no arc in G leaves G; x is said to be *flowing* in G otherwise. If $U \subseteq V(G)$ we say that $x \in U$ is a source, a sink or flowing in U, if it is a source, a sink or flowing in $G[U]$, respectively. In an undirected graph, two edges are *parallel* if they join the same endpoints. In a directed graph two arcs are *parallel* if they join the same endpoints and have the same head. An undirected graph is *simple*, if it does not contain parallel edges. A directed graph is a *1-graph* if it does not contain parallel arcs. In particular, at most one loop is allowed at each node. Two arcs e and e' in a directed graph G, are said to be *symmetric* if $h(e) = t(e')$ and $t(e) = h(e')$. If $e = xy$ and $e' = yx$, $x, y \in V(G)$, we say that there is a *digon* at x and y or that xy and yx form a *digon*. The *support* of a directed graph G is the simple undirected graph $\tilde{G}$, having as vertex set the node set of G and where two vertices are joined by an edge if the corresponding nodes are joined by at least one arc. An *anti-symmetric* directed graph is a digon-free 1-graph. An *anti-symmetric* directed graph is *simple* if it loopless.

Let R be an undirected graph. The *line graph* of R, denoted by $L(R)$, is the graph having $V(L(R)) = E(R)$ and where two vertices are joined by an edge if and only if the corresponding edges are adjacent in R. A simple undirected graph G is a *line graph* if it is isomorphic to the line graph of some graph R, we call R a root of G.

Let R be a directed graph. The *adjoint* of R, denoted by $L^*(R)$, is the directed graph having $V(L^*(R)) = E(R)$, and where there is an arc xy joining two nodes $x, y \in V(L^*(R))$ if and only if arc x precedes arc y in R. A directed graph G is an adjoint graph if it is isomorphic to the adjoint of some

[1]This notation is not standard. Usually arcs in directed graphs as thought of as ordered pair (x, y). We reserve this symbol for other special ordered pairs

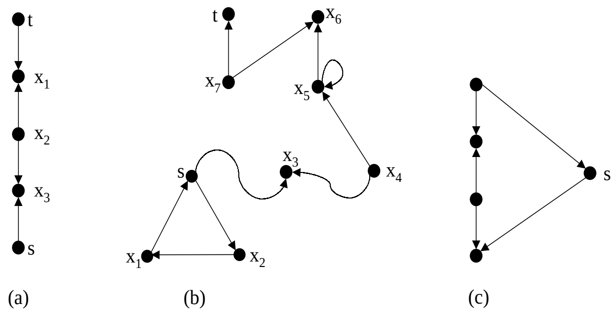

Figure 1: (a). A simple st-alternating path; (b). An st alternating path visiting (in this order) nodes s, x_1, x_2, s, x_3, x_4, x_5 , x_5 (again), x_6, x_7, t; (c). A simple alternating cycle from s to s.

---

graph R, we call R a root of G. If G is an adjoint of a directed 1-graph we say that G is a *directed line graph*. Notice that, by definition, directed line graphs are 1-graphs.

Let G be a 1-graph; *G admits a weak-root* if there is a 1-graph R and a bijection $g : V(G) \to E(R)$, such that $xy \in E(G)$ implies $g(x), g(y)$ are consecutive arcs in R.

Clearly if G has a weak-root R, G is partial graph of $L^*(R)$. Conversely, if G is a partial graph of some directed line graph G', the root of $G'[V(G)]$ is a weak-root of G. It follows that deciding if a graph has a weak root is tantamount to decide if the graph is a partial graph of some directed line graph. Therefore the following definition is well justified. A 1-graph G is a *partial directed line graph*, (*PDL graph*, for shortness) if it is a partial graph of some directed line graph. Given a 1-graph, a set of arcs $E' \subseteq V(G) \times V(G) \setminus E(G)$, is called *directed line graph completion* if $G' = (V(G), E(G) \cup E')$ is a directed line graph. The notion of line graph completion for simple undirected graphs is defined similarly. In the sequel, we will use the term root both for a weak root and for a root. A 1-graph G is a *simple partial directed line graph*, (*SPDL graph*, for shortness) if it is a PDL graph whose root is simple (i.e., loopless ans antysimmetric).

Remark 2.1 *If G is a PDL graph, each of its connected components is itself a PDL graph. Moreover, single nodes, single arcs, single loops or digons are trivially directed line graph. Therefore we may assume w.l.o.g. that the input graph in Problems 1 and 2, is connected and it is neither a single arc nor a digon.*

3 Alternating paths

Let G be a directed graph and let $s, t \in V(G)$. An *$s\,t$-alternating path* is a subgraph P whose arc set $E(P)$ can be ordered as $\{e_1, \ldots, e_p\}$, where e_1 is incident in s, e_p is incident in t, e_i and e_{i+1} are adjacent but not consecutive. Nodes s and t will be referred to as the *endpoints* of the alternating path, while edges e_1 and e_p as its *endarcs*. Figure 1 gives some intuition on how $s\,t$ alternating paths "go". Note that the support of an $s\,t$-alternating path is just a walk (with loop degrees counted twice) between s and t whose edges can be colored red and blue in such a way that reversing the orientation of, say, the blue edges we get an oriented eulerian walk. Therefore, if say e_1 is red and $t(e_1) = s$, red edges can be thought of as forward edges walking from s to t while the blue ones as backward edges walking in the same direction. A single arc $e = s\,t$ is always an alternating path. An *alternating cycle*, is an alternating path whose endpoints coincide (a closed alternating path).

3.1 Composing alternating paths

The set of alternating paths is partitioned into four classes according to how they enter/leave the endpoints. Let G be a directed graph. The node-arc incidence mapping of G is a mapping $\varphi : V(G) \times E(G) \to \{-1, 0, 1\}$, defined by $\varphi(x, e) = -1$ if $h(e) = x$, $\varphi(x, e) = -1$ if $t(e) = x$ and $\varphi(x, e) = 0$ otherwise. Let s and t be two not necessarily distinct nodes of G and let P be an *st*-alternating path P. Let the two endarcs of P be a and b with a incident in s and b incident in t. We say that the *sign* of P is (α, β) (or that P is of *class* (α, β) or simply that P *is* (α, β)) if $\varphi(s, a) = \alpha$ and $\varphi(t, b) = \beta$, where $\alpha, \beta \in \{-1, +1\}$. We also say that the *sign of P in s* is α and the *sign of P in t* is β. Since for most of what follows the only thing that matters dealing with alternating paths is the sign of such paths, it is convenient to have a shorthand notation for pairs of nodes that are endpoints of alternating paths. Let G be any directed graph. For two not necessarily distinct nodes x, y of $V(G)$ and $\alpha, \beta \in \{-, +\}$, we say that pair (x, y) is an *(α, β)-pair in G* (or that (x, y) *is of class* (α, β) in G, or simply that (x, y) is (α, β) *in G*) if there exists an xy-alternating path in G whose sign in x and y is α and β, respectively. We also say that (x, y) is a *signed pair*. With a little abuse of notation, we will say that (x, y) is an (α, β)-pair *in a node set U* instead of *in $G[U]$*.

The following result shows that signed pairs can be combined.

Lemma 3.1 *Let G be a directed graph and x, y, z three (not necessarily distinct) nodes of G. If (x, y) and (y, z) are, respectively, a (α, β)-pair and a (β, γ)-pair in $V(G)$ then (x, z) is an (α, γ)-pair in $V(G)$.*

Proof. We have only to show that given an xy-alternating path P_1 in G of sign (α, β) and a yz-alternating path P_2 in G of sign (β, γ), there exists an xz-alternating path in G of sign (α, γ). If P_1 and P_2 have no common edges, then the required path is just their concatenation $P = P_1 \circ P_2$. Otherwise, let e be the last common edge on P_2 walking from y to z and denote by a and b the edges of P_1 and P_2 that follow e walking from y to x and from y to z, respectively. Let s and t be the endpoints of e and suppose that t follows s walking from y to z. Clearly, t lies on P_1. Let P_2' be the subpath of P_2 joining t and z and P_1' be the subpath of P_1 joining t and x. Now P_1' and P_2' are edge disjoint and both have t has an endpoint. Moreover, on P_1, either t follows s or s follows t. In the former case $\varphi(t, a) = \varphi(t, b)$, because P_1 and P_2 are alternating. In the latter case, P_1' ends in t with e and since P_2 is alternating and e belongs to it $\varphi(t, e) = \varphi(t, b)$. It follows that P_1' and P_2' have the same sign at t and, being edge disjoint, they can be concatenated so as to obtain the required path $P = P_1' \circ P_2'$. $\square$

3.2 Roots of Alternating Paths

Alternating paths play a fundamental role in deciding whether a directed graph G is a PDL graph, since, as we shall see, the only obstruction to the property of being a PDL graph consists in the presence in G of alternating paths of different signs between pairs of nodes.

Theorem 3.1 *Let P be an alternating path in a 1-graph G. Any root of G contains a subgraph R_P whose arcs corresponds bijectively to nodes of P and such that:*

1. *all arcs in R_P share a common node w;*
2. *arcs of R_P corresponding to sources in P have the same head and enter w while arcs corresponding to sinks in P have the same tail and leave w;*
3. *if there is a loop ww in R it correspond to a flowing node x in P such that (x, x) is of class $(+, -)$ in G (possibly because of a loop xx in G).*

Proof. Let P be an an alternating *st*-path of sign (α, β) in G. We will prove the claim only for $\alpha = +$ and $\beta = -$ the other cases being similar. Let $e_1 e_2 \dots e_p$ be the sequence of edges visited in P walking from s to t and let $x_1 x_2 \cdots x_{p+1}$, $x_1 = s$, $x_{p+1} = t$ be the corresponding sequence of (not

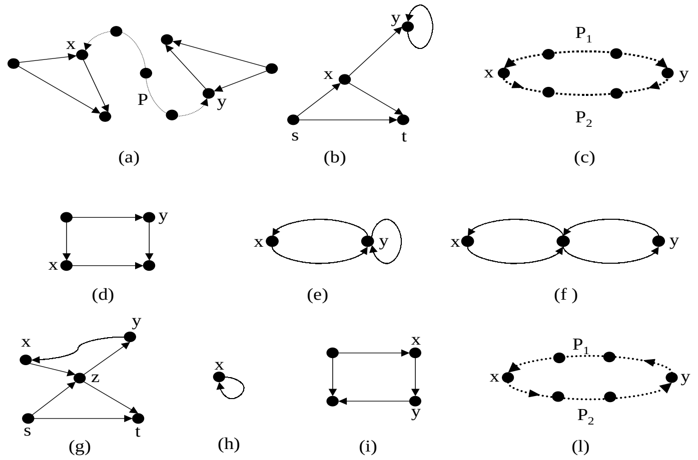

Figure 2: various patterns of parallel and symmetric pairs; pairs (x, y) in (a)÷(f) are parallel pairs. Pairs (x, y) in (g)÷(l) are symmetric pairs.

---

necessarily distinct) nodes. It follows, by definition of alternating path, that e_i leaves x_i, if i is odd and e_i enters x_i if i is even, $i = 1, \ldots, p+1$. Hence, p is odd. Therefore, in any root of G, if any, one must have $h(x_i) = t(x_{i+1}) = h(x_{i+2})$, if i is odd, and $t(x_i) = h(x_{i+1}) = t(x_{i+2})$ if i is even. Let $S_P = \{x \in V(P) \mid x = x_i, i \text{ odd}\}$ and $T_P = \{x \in V(P) \mid x = x_i, i \text{ even}\}$. It follows that in any root of G, if any, all edges in S_P have the same head as x_1 and all edges in T_P have the same tail as x_2; since P contains the edge $e_1 = x_1x_2$ one has $h(x_1) = t(x_2)$. Therefore, in any root of G, if any, all arcs corresponding to nodes of P, share a common node w, and those in S_P all have the same head and enter w, while those in T_P all have the same tail and leave w. If $x \in (S_P \cup T_P) \setminus (S_P \cap T_P)$ then x is either a source of a sink in P. It follows that $S_P \cap T_P$ is the set of flowing nodes in P. Moreover, $x \in S_P \cap T_P$ if and only if x is a loop on P or x is contained in some odd sub-cycle of P. In any case if $x \in S_P \cap T_P$ then $t(x) = h(x)$, hence x must be a loop at w. Consequently any root of G, if any, contains a subgraph R_P as stated. □

Alternating paths are the simplest non trivial subgraphs that impose local constraints on the structure of the root. Further and more global conditions arise from the way alternating paths interact. Suppose for instance that $x, y \in V(G)$ and that P_1 and P_2 are xy-alternating paths of sign $(+,+)$ and $(-,-)$, respectively. These paths are not required to be edge-disjoint. By Theorem 3.1 applied to P_1, it follows that $t(x) = t(y)$ in any root R of G. On the other hand, applying Theorem 3.1 to P_2, we get $h(x) = h(y)$ whence, x and y must be parallel arcs in any root of G. Therefore, the above configuration is forbidden in any PDL graph G since its root must be a 1-graph. Similarly, given an SPDL graph G and $x, y \in V(G)$ the pair (x, y) cannot be both a $(-,+)$-pair and a $(+,-)$-pair. Indeed such a pair would be represented as a digon in any root of G contradicting that roots of G are required to be antisymmetric. The above discussion motivates the following notions.

Definition 1 *Let H be a directed graph. Two nodes x, y form a* parallel pair *in H (or in $V(H)$) if (x, y) is both a $(-,-)$-pair and $(+,+)$-pair. Nodes x, y are a* symmetric pair *in H if (x, y) is both a $(-,+)$-pair and $(+,-)$-pair. If (x, x) is a $(+,-)$-pair (and obviously it also is a $(-,+)$-pair) we say that node x is* odd *in H. We say a graph is $\rightrightarrows$-free if does not contain parallel pairs, and $\rightleftarrows$-free if it does not contain either parallel or symmetric pairs.*

For instance, in Figure 2.g, the xy-alternating path traversing nodes x,z,s, t, z and y, in this order, defines (x, y) as a $(+,-)$-pair. On the other hand arc yx defines (x, y) as a $(-,+)$-pair. In particular, node z is odd, because of the cycle through z, s, t and z. Note that the two forbidden paths in the definition are not necessarily edge disjoint. As an example consider Figure 2.b where there is an xy-alternating path P' traversing (in this order) nodes x, s, t, x and y and an xy-alternating path P'' traversing (in this order) nodes x and y (twice).

The above discussion shows that parallel pairs must be represented in the root graph as pair of parallel edges, symmetric pairs by digons and odd nodes by loops. Therefore, the following fact is a direct consequence of the definition of PDL graph.

Fact 3.1 *Let G be a directed graph. If G is a PDL graph then G is $\rightrightarrows$-free. If G is a SPDL graph then G is $\rightleftarrows$-free.*

In Theorem 5.1 we will see that this conditions turns out to be also sufficient for G being a PDL graph or a SPDL graph.

4 Kernels

In this section we use the concept of alternating paths to identify larger portions of a PDL graph, called kernels. We show that the subgraph induced by a kernel essentially is a directed bipartite graph, with some exceptions due to the presence of odd nodes, and to some "badly oriented" arcs, that contrast with the orientation of a $(+,-)$-pair.

Given a node x, we define the two sets of all nodes of G reachable from x by alternating paths:

$$B_+(x) = \{x\} \cup \{y \in V(G) \mid (x,y) \text{ is a } (+,\beta)\text{-pair}\} \tag{1}$$

$$B_-(x) = \{x\} \cup \{y \in V(G) \mid (x,y) \text{ is a } (-,\beta)\text{-pair}\}, \tag{2}$$

Fact 4.1 *Given any node $y \in B_\alpha(x)$, assuming x, y is a (α,β)-pair, sets $B_\alpha(x)$ and $B_\beta(y)$ coincide.*

Proof. In fact, assume $z \in B_\alpha(x)$, and (x,z) is an (α,γ)-pair. Since (x,y) is an (α,β)-pair, then (y,z) is a (β,γ)-pair due to Lemma 3.1. By the same argument, given any $z \in B_\beta(y)$, we have $z \in B_\alpha(x)$. $\square$

Fact 4.1 shows that any node in $B_\alpha(x)$ could be used to generate the same set of reachable nodes, using alternating paths starting by the convenient sign.

Definition 2 *A* kernel *is a non-singleton set of nodes of the form $B_\alpha(x)$ for some node x and some $\alpha \in \{-,+\}$.*

Every kernel induces a connected graph.

After Fact 4.1, we see that every non-odd node can be thought of as signed with respect to the kernel it belongs to. Actually, we will see later that a kernel in a PDL graph could contain a single special node (an odd node), having both signs.

More formally, for a kernel K we define the following sets:

$$S(K) = \{x \in K \mid (x,y) \text{ is a } (+,\beta)\text{-pair}, \forall y \in K, \beta \in \{-,+\}\}, \tag{3}$$

$$T(K) = \{x \in K \mid (x,y) \text{ is a } (-,\beta)\text{-pair}, \forall y \in K, \beta \in \{-,+\}\}. \tag{4}$$

Note that $S(K) = \{x \in V(G) \mid B_+(x) = K\}$ and that $T(K) = \{x \in V(G) \mid B_-(x) = K\}$. Sets $S(K)$ and $T(K)$ are called *shores* of the kernel. By definition of shores, (x,y) is a $(+,+)$-pair in K if and only if $x, y \in S(K)$; (x,y) is a $(+,-)$-pair in K if and only if $x \in S(K)$ and $y \in T(K)$; (x,y) is a $(-,-)$-pair in K if and only if $x, y \in T(K)$.

Lemma 4.1 *$S(K) \cap T(K)$ is the set of odd nodes in K.*

Proof. If node x is both in $S(K)$ and in $T(K)$, there exists a node y such that (x, y) is both a $(+, \beta)$-pair and a $(-, \beta)$-pair for some sign β; hence (x, x) is a $(+, -)$-pair and x is odd.

Conversely, let x be odd. For any $y \in K$, y is reachable from x both by a $(-, \beta)$-path and a $(+, \beta)$-path, where $\beta = +$ if $y \in S(K)$ and $\beta = -$ if $y \in T(K)$. Hence x is in $S(K) \cap T(K)$. □

4.1 Flowing nodes

Lemma 4.2 *Let G be a directed graph, and let K be a kernel in G. If G is $\rightrightarrows$-free at most one node is odd in K. If G is $\rightleftarrows$-free then no node is odd in K.*

Proof. The proof proceeds by contradiction. Let us assume that x and y are both odd in K. Both (x, x) and (y, y) are $(+, -)$-pairs, and since K is a kernel, (x, y) is an (α, β)-pair. Either $\alpha = -\beta$ or $\alpha = \beta$.

If $\alpha = -\beta$, since (x, x) and (y, y) are both $(\beta, -\beta)$, combining (x, x) with (x, y) gives (x, y) as (β, β), while combining (x, y) with (y, y) gives (x, y) as $(-\beta, -\beta)$ and (x, y) would be a parallel pair (see Figure 2.b).

If $\alpha = \beta$, (x, y) is (β, β). Since (x, x) is a $(-\beta, \beta)$-pair, and (y, y) is a $(\beta, -\beta)$-pair, combining (x, x), (x, y) and (y, y) in this order, defines (x, y) as $(-\beta, -\beta)$. Again a parallel pair would arise. (see Figure 2.b).

This contradiction proves the first part of the statement. The second part follows by definition of $\rightleftarrows$-free and odd node. □

Let K be a kernel of G. Let $a = ts$ be an arc of $G[K]$. Arc a is said to be a *backward* arc (or simply *backward*) in K if $s \in S(K)$ and $t \in T(K)$. Clearly, if ts is backward in K, then (s, t) is a symmetric pair. Indeed, since K is connected and $s \in S(K)$ and $t \in T(K)$, (s, t) is a $(+, -)$-pair. On the other hand, (s, t) is a $(-, +)$-pair, because of arc ts. Backwards arcs in a given kernel can be thought of as badly oriented chords that shortcut rightly oriented alternating paths.

Lemma 4.3 *Let K be a kernel in G, and x be flowing and non-odd in K.*

(i) *If G is $\rightrightarrows$-free then at least one neighbor of x in $G[K]$ is flowing and non-odd in K. Moreover, among such neighbors there exists one such that the arc joining it to x is backward in K and no other backward arc of K is incident in x.*

(ii) *If $x \in S(K)$ then $\deg^-_H(x) = 1$, analogously, If $x \in T(K)$ then $\deg^+_H(x) = 1$, where $H = G[K]$. Moreover, the only alternating path of H that uses a backward arc is the backward arc itself.*

Proof.

(i) W.l.o.g. we assume $x \in S(K)$; the proof in the case $x \in T(K)$ being symmetric. Since x is flowing there is an arc yx in $G[K]$. Since $x \in S(K)$, and $y \in K$, pair (x, y) is a $(+, \beta)$-pair with $\beta = -$. Otherwise, if $\beta = +$, combining pair (x, y) and arc yx, we would obtain that x is odd. Moreover, y is non odd, otherwise, as above, combining pair (x, y), pair (y, y) with the convenient sign and arc yx as in Figure 2.(e) we obtain again that x is odd, a contradiction. Clearly, y is flowing, and arc yx is backward.

We show now, that there cannot be two backward arcs yx, $y'x$, in $G[K]$. In fact, both (x, y) and (x, y') are $(+, -)$-pairs. Combining pairs (y', x) and (x, y) defines (y', y) as a $(-, -)$-pair; combining arcs $y'x$ and xy defines (y', y) as a $(+, +)$-pair; Therefore (y', y) would be a parallel pair. (Figure 2.(f))

(ii) The property concerning the degrees of x directly follows from by part (i).

For completing the proof we only need to show that the only alternating path of $G[K]$ that uses a backward arc yx is the backward arc itself. If there existed an alternating path with at least two arcs traversing yx, it would exist an arc d either leaving y or entering x. But then $\deg^+_H(y) \geq 2$

or $\deg_H^-(x) \geq 2$. In any case a contradiction would arise, since $y \in T(K)$ and the first part of the thesis also applies to node y.

$\square$

4.2 Structure of kernels

The structure of a kernel is described in the following theorem.

Theorem 4.1 *Let K be a kernel in G. Let X denote the set of odd nodes of K (X is either empty or a singleton) and denote by $F(K)$ the set of non odd flowing nodes of K. If G is $\rightrightarrows$-free, then the following hold.*

1. *$|S(K) \cap T(K)| \leq 1$,*
2. *The set of arcs of $G[K]$ can be partitioned into three disjoint sets:*

   forward arcs: *all arcs xy such that $x \in S(K) \setminus X$ and $y \in T(K) \setminus X$;*

   backward arcs: *backwards arcs xy such that $x \in T(K) \cap F(K)$ and $y \in S(K) \cap F(K)$; these arcs form a perfect matching on nodes $S(K) \cap F(K)$ and $T(K) \cap F(K)$;*

   odd arcs: *the set of all arcs of $G[K]$ incident in the unique odd node of K (if any).*

Proof.

1. By Lemma 4.1, $S(K) \cap T(K)$ is the set of odd nodes, and by Lemma 4.2 there is at most one odd node in each kernel.
2. Observe first that no backward arc is incident in the odd node. Let ts be any such arc and suppose that one among s and t is odd. Then the other one is flowing and non odd. By Lemma 4.3, this latter node is incident in exactly one backward arc, the other endpoint of which being flowing and non odd. Therefore either two backward arcs are incident in the same non odd flowing node or none among s and t is odd. In any case a contradiction arises. It follows that the sets defined in part 2 of the statement are disjoint. Let H' be the graph obtained from $G[K]$ by removing all backward arcs. By Lemma 4.3 (ii), all nodes except the odd node are either sources or sinks in H'. Hence, all arcs in H', except possibly those incident in the odd node, go from $S(K)$ to $T(K)$. By definition, backward arcs go from a node in $T(K)$ to a node in $S(K)$. Therefore the set of arcs of $G[K]$, can be partitioned as stated in part 2. We only have to show that backward arcs form a perfect matching on $S(K) \cap F(K)$ and $T(K) \cap F(K)$, that follows from Lemma 4.3 (ii).

$\square$

The following is a direct consequence of Theorem 4.1.

Corollary 4.1 *Let K be a kernel in a $\rightleftarrows$-free directed graph G. Then $G[K]$ is a connected graph whose support is bipartite.*

Proof. Directly by Theorem 4.1, observing that there are neither backwards arcs nor odd nodes. $\square$

4.3 Roots of kernels

Importance of kernels is easily recognized looking at the following result that generalizes Theorem 3.1.

Theorem 4.2 *Let K be a kernel in a $\rightrightarrows$-free graph G. Let R be a 1-graph whose arcs correspond bijectively to nodes of $G[K]$. Suppose that R is such that:*

1. *all arcs of R share a common node w;*

2. *arcs of R corresponding to nodes in $S(K)$ enter w and arcs corresponding to nodes in $T(K)$ leave w; in particular an odd node, if any, corresponds to a loop at w;*

3. *for any backward arc yx, $x \in S(K)$, $y \in T(K)$, the arcs of R corresponding to x and y form a digon.*

4. *any two arcs corresponding to non flowing nodes, have exactly one node in common (this node being w).*

Then $G[K]$ is a PDL graph and R is one of its root. Any other root R' of G must satisfy (1)÷(3). Moreover, among all such roots R', those minimizing $|E(L^*(R'))|$, satisfy (4) as well.

Proof. Let R be a graph fulfilling (1)÷(4), and let g be the bijection that associates arcs in R with nodes in G. Conditions (1)÷(3) guarantee that $xy \in E(G[K]) \Rightarrow g(x)$ and $g(y)$ are consecutive arcs in R. Therefore, R is a root of $G[K]$ and $G[K]$ is a PDL graph.

Let us prove that any other root R' of $G[K]$ satisfies (1)÷(3). By definition of kernel, for any two nodes $x, y \in K$, there exists and alternating path xy-path P such that P is of sign $(+,+)$ if $x, y \in S(K)$, P is of sign $(+,-)$ if $x \in S(K)$ and $y \in T(K)$ and P is of sign $(-,-)$ if $x, y \in T(K)$. By Theorem 3.1, all arcs in R' corresponding to nodes of $S(K)$ have the same head; all those corresponding to nodes of $T(K)$ have the same tail; the head of any arc corresponding to a node in $S(K)$ coincides with the tail of any arc corresponding to a node of $T(K)$. Moreover, since $|S(K) \cap T(K)| \leq 1$, at most one node in K corresponds to a loop in R'. It follows that R' must satisfy (1) and (2). Let xy be a backward arc in K and denote by $g(x)$ and $g(y)$ be the arcs of R' corresponding to x and y, respectively. Since R' satisfies (1) and (2), the arc $g(x)$ enters w and $g(y)$ leaves w. On the other hand, by definition of root, $yx \in E(G[K]) \Rightarrow g(y)$ and $g(x)$ are consecutive arcs in R. Therefore, $h(g(x)) = t(g(y)) = w$ and $h(g(y)) = t(g(x))$ both holds. Hence $g(x)$ and $g(y)$ form a digon and R' satisfies also (3).

Let now R' be a root of $G[K]$ satisfying (1)÷(3). After Theorem 4.1. (2), any root R' of $G[K]$ contains at least $|F(K)|$ digons. Since a root is a 1-graph, no two different digons can have the same endpoints. Therefore if there are more than $|F(K)|$ digons, they must be sought among arcs corresponding to non flowing nodes. Clearly $|E(L^*(R'))|$ is strictly increasing in $|F(K)|$. Therefore, any root minimizing $|E(L^*(R'))|$, must contain as few digon as possible. Consequently, a root R' of $G[K]$ minimizing $|E(L^*(R'))|$, satisfies (4) as well. □

For a kernel K in a $\rightrightarrows$-free graph, call the four-tuple $(S(K), T(K), X, F(K))$ the *skeleton* of K and call a root R of $G[K]$ *canonical* if it satisfies the conditions of Theorem 4.2. In view of this conditions, such a canonical root depends only on the skeleton of K in the following sense.

Fact 4.2 *Let G be a $\rightrightarrows$-free graph and K be one of its kernels. Let $x \in S(K)$ and $y \in T(K)$ be such that $xy \notin E(G)$. Then K is a kernel in $G' = (V(G), E(G) \cup \{xy\})$ and it has the same skeleton both in G and G'. In particular $G[K]$ and $G'[K]$ have the same canonical root.*

Proof. Directly from the definition of kernel and from Theorem 4.2. □

5 Main result

Having investigated the structure of kernels of $\rightrightarrows$-free and $\rightleftarrows$-free directed graphs and their roots, let us see how kernels interact and how a root of a PDL graph can be built.

A first important consequence of Theorem 4.1 is the following.

Lemma 5.1 *Given any directed graph G, a node lies in exactly one kernel if and only if it is either a source or a sink or it is an odd node.*

Proof. If node x is a source, then it only belongs to $B_+(x)$ ($B_-(x) \setminus \{x\}$ being empty), and similarly for sinks (($B_+(x) \setminus \{x\}$ being empty). If x is odd we have $B_+(x) = B_-(x)$. Conversely, if x belongs to only one kernel then either $B_+(x) \setminus \{x\} = \emptyset$, or $B_-(x) \setminus \{x\} = \emptyset$, or $B_+(x) = B_-(x)$ (x is odd). □

By definition of kernel, starting from a node x we can define at most two kernels $B_+(x)$ and $B_-(x)$. It follows that:

Fact 5.1 *Let G be a directed graph. A node of G lies in at most two distinct kernels. If a node of G belongs to two distinct kernels K and L then either $x \in S(K) \cap T(L)$ or $x \in S(L) \cap T(K)$.*

Lemma 5.2 *Let G be a directed $\rightrightarrows$-free graph and $K, L \subseteq V(G)$ be two distinct kernels of G. Then $|K \cap L| \leq 2$. In particular, if $|K \cap L| = 2$ then $|S(K) \cap T(L)| = 1$ and $|S(L) \cap T(K)| = 1$. If G is $\rightleftarrows$-free, then $|K \cap L| \leq 1$.*

Proof. Notice first that, by Lemma 5.1, nodes in $K \cap L$ cannot be odd since odd nodes belong to exactly one kernel. Suppose that $|K \cap L| \geq 3$. There are at least two common nodes in the same shore of K or of L. Let x, y be such two nodes and suppose w.l.o.g. that $x, y \in S(K)$. By Fact 5.1, x and y must belong to $T(L)$, but then (x, y) would be $(+,+)$-pair in K and $(-,-)$-pair in L contradicting the fact that G is $\rightrightarrows$-free. Since any two nodes in the intersection of distinct kernels would form a symmetric pair, the second part of the statement follows as well. $\square$

Let $\mathcal{K} = \{K_1, K_2, \ldots, K_q\}$ be the family of kernels in a $\rightrightarrows$-free graph G. For $1 \leq i \leq q$, let (S_i, T_i, X_i, F_i) be the skeleton of K_i where, S_i and T_i are the shores of K_i, X_i is the set of odd nodes of K_i, and F_i is the set of non-odd flowing nodes of K_i. Recall that by Theorem 4.1, X_i is either empty or a singleton and that $|S_i \cap F_i| = |T_i \cap F_i|$. For a node $x \in V(G)$, call a node $y \in V(G)$ a *mate* of x if $\{x, y\} = K_i \cap K_j$, for some $i \neq j$. Let

$$\tilde{F} = \{x \in V(G) \mid \{x, y\} = K_i \cap K_j \text{ for some } y \in V(G) \text{ and } i \neq j\} \tag{5}$$

be the set of nodes of G having a mate.

Fact 5.2 *Let G be $\rightrightarrows$-free. Any node of G has at most one mate.*

Proof. Let $x \in V(G)$ have two distinct mates y and z. Then $K_i \cap K_j = \{x, y\}$ and $K_r \cap K_t = \{x, z\}$ both hold for some $i \neq j$ and $r \neq t$. Necessarily $\{i, j\} = \{r, t\}$ must hold (for otherwise x would belong to more than two kernels contradicting Fact 5.1). But this is impossible because $K_i \cap K_j$ would contain $\{x, y, z\}$. $\square$

Let G_i be the subgraph induced by K_i and denote by $G_{0,i}$ the subgraph of G_i obtained removing all backward arcs of G_i. It is worth stating the following:

Lemma 5.3 *Let G be $\rightrightarrows$-free. Then $|E(G_i) \cap E(G_j)| \leq 2$. In particular, $|E(G_i) \cap E(G_j)| = 2$ implies that these edges form a digon in G. In any case one of the common edges is a backward and $F_i \subseteq \tilde{F} \cap K_i$, for $i = 1, \ldots, q$. Moreover, the $G_{0,i}$'s are arc disjoint and so are the G_i's when G is $\rightleftarrows$-free.*

Proof. Clearly $|E(G_i) \cap E(G_j)| \geq 1 \Rightarrow |K_i \cap K_j| = \{x, y\}$ for some $x, y \in V(G)$ (by Lemma 5.2). Hence $x, y \in \tilde{F}$ and x and y are mates of each other. Now either $xy, yx \in E(G)$ or exactly one among xy and yx is in $E(G)$. Assume w.l.o.g. that in the latter case $xy \in E(G)$. In both cases, due to Lemma 5.2, (x, y) is a $(+,-)$-pair in one kernel and a $(-,+)$-pair in the other kernel, hence xy is backward in one kernel and, if $yx \in E(G)$, yx is backward in the other kernel. It follows that the endpoints of any backward arc are mates of each other. Since such endpoints belongs to some F_i, one has $F_i \subseteq \tilde{F} \cap K_i$. Removing all backward arcs from the G_i's leads to the arc disjoint subgraphs $G_{0,i}$'s. If G is $\rightleftarrows$-free there are no backward arcs, whence $G_{0,i} \cong G_i,\ i \in I$. $\square$

We are now in position to state and prove our main theorem.

Theorem 5.1 *Let G be a directed graph. The following statements are equivalent.*

(1) *G is a PDL graph.*

(2) *G is $\rightrightarrows$-free.*

(3) *The family $\mathcal{K} = \{K_1, K_2, \dots, K_q\}$ of kernels of G is a covering of $V(G)$ such that $|S_i \cap T_j| \leq 1$ for each pair of kernels.*

Proof.

(1)⇒(2) This has been already observed in Section 3.2 by Fact 3.1.

(2)⇒(3) It directly follows by the above Lemma 5.1, Fact 5.1, Lemma 5.2.

(3)⇒(1) Assuming that kernels intersect as in (3), we exhibit a root R_G of G. By Theorem 4.2, each $G[K_i]$ admits a canonical root R_i. Let i denote the node shared by all arcs of R_i and call it the center of R_i.

We build the root graph R_G of G by pasting the q graphs $R_1, R_2, \dots, R_q$ along with their arcs. Initially they are arc disjoint. Let x be a node in $V(G)$ covered by two kernels; due to Fact 5.1, we know that $x \in S_i \cap T_j$ for some i, j; in this case we make R_i and R_j share the arc corresponding to x as follows: x corresponds to an arc $t(x)i$ in R_i and to an arc $jh(x)$ in R_j; we identify $t(x)$ and $h(x)$. Notice that this procedure leaves unchanged the centers of the R_i's. In particular, no loop arises in this way and the graph R_G we get eventually has at most one loop at each center.

Since each arc $e \in E(G)$ is contained in the subgraph induced by a kernel K_i, all arcs of G are mapped into consecutive arcs in R_G.

We only have to show that R_G is a 1-graph. Assume by contradiction that there are two parallel arcs $a, b \in E(R_G)$ corresponding to nodes $x, y \in V(G)$. Arcs a and b cannot be loops because there is at most one loop at each center. Moreover, since each R_i is a 1-graph, a and b should come from two different stars R_i, R_j. Therefore a and b are of the form $a = (i, j)$ and $b = (i, j)$, $i \neq j$, meaning that both a and b belong to $S_i \cap T_j$, that contrasts with hypothesis (3).

□

A root R_G of G built as in Theorem 5.1 (3), will be referred to as *canonical*. Theorem 5.1 (3), through Theorem 4.2, shows that canonical roots of PDL graphs are determined by the skeleton of any of its kernels and by the interactions among such skeletons. In particular, R_G may contain more digons than those occurring in the roots of its kernels. By the construction in Theorem 5.1 (3), such digons arise whenever the corresponding nodes belong to exactly two kernels of G, i.e., when such nodes are mates of each other but they are not flowing in some kernel (recall Fact 5.2 and Lemma 5.3). It follows that the set of digons of R_G is determined by the set $\tilde{F}$ defined in (5). Let $\{(S_i, T_i, X_i, F_i), i = 1, \dots q\}$ be the family of the skeletons of the kernels of G. By Fact 5.2 there are exactly $|\tilde{F}|/2$ digons, and among them exactly $\sum_{i=1}^{q} |F_i|/2$ are already determined by the kernels. Recall that, by Lemma 5.3, $F_i \subseteq \tilde{F} \cap K_i$ so that digons corresponding to pairs of mates in $\tilde{F} \setminus \bigcup_i F_i$ are due to the interactions among kernels.

Let us call the family $\{(S_i, T_i, X_i, \tilde{F}_i), i = 1, \dots q\}$ the *skeleton* of G, where, for $i = 1, \dots q$, $\tilde{F}_i = \tilde{F} \cap K_i$.

Corollary 5.1 *Let G and G' be PDL graphs on the same set of nodes. If G and G' have the same skeleton they have isomorphic canonical roots. In particular, if for some i, $x \in S_i$ and $y \in T_i$ are such that $xy \notin E(G)$, then $G' = (V(G), E(G) \cup \{xy\}))$ is still a PDL graph whose root is the root of G.*

If we are looking for simple roots the following corollary solves the problem.

Corollary 5.2 *Let G be a directed graph. The following statements are equivalent.*

(1') *G is a SPDL graph.*

(2') *G is $\rightleftarrows$-free.*

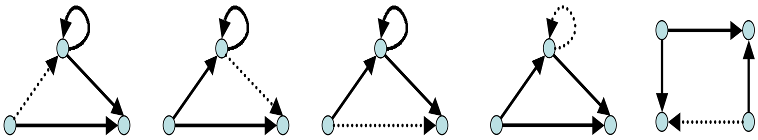

Figure 3: forbidden alternating paths in directed line graphs.

(3') *The family $\mathcal{K} = \{K_1, K_2, \ldots, K_q\}$ of kernels of G is a covering of $V(G)$ such that $|K_i \cap K_j| \leq 1$ for each pair of kernels, and the subgraph $G[K_i]$ is bipartite, for $1 \leq i \leq q$.*

Proof. Directly from Theorem 5.1, via Lemmata 5.2 and 5.3, after observing that loops and digons are special symmetric-pairs. □

If G is simple we may re-state the necessary and sufficient condition for G to be a SPDL graph as follows.

Corollary 5.3 *A simple directed graph G is an SPDL graph if and only if it does not contain any of the following subgraphs:*

- *odd simple cycles with exactly one flowing node;*
- *even simple cycles with exactly two flowing nodes.*

6 Directed line graphs

In Theorem 5.1 we showed how to build a root graph R_G of a PDL graph G. Each pair of consecutive nodes in G is mapped to a pair of consecutive arcs in R_G, but there can be some consecutive arcs in R_G that do not correspond to consecutive nodes in G.

A graph G is a directed line graph (DL graph) if it exists a root R_G such that each pair of consecutive arcs in R_G corresponds to a pair of consecutive nodes in G.

It can be seen that all "extra" consecutive pairs are due to pairs of nodes x, y in G such that $x \in S_i$, $y \in T_i$ for some kernel K_i of G and $xy \notin E(G)$. We call a kernel *complete* if $S_i \times T_i \subseteq E(G)$, i.e., the induced subgraph $G[K_i]$ is a complete directed bipartite graph, plus possibly some backward arcs going from T_i to S_i and arcs going from each node of S_i to the odd node and from the odd node to each node of T_i. By Theorem 4.2 and Fact 4.2, $(S_i \times T_i) \setminus E(G)$ is a minimum directed line completion of $G[K_i]$.

A directed graph is *semi-functional* if

$$N^+(x) \cap N^+(y) \neq \emptyset \Rightarrow N^+(x) = N^+(y). \tag{6}$$

Duchenne characterized adjoint graphs of directed graphs as those graphs satisfying (6), such condition is sometimes referred to as Duchenne condition. He actually showed that (6) is equivalent to excluding subgraphs in Figure 3 with the dotted arcs missing. Directed line graphs, i.e., adjoint of 1-graphs, have been characterized by Blazewich et al. in [1], by means of the following condition

$$N^+(x) \cap N^+(y) \neq \emptyset \Rightarrow N^+(x) = N^+(y) \text{ and } N^-(x) \cap N^-(y) = \emptyset, \tag{7}$$

that, as the authors showed, turns out to be equivalent to excluding from an adjoint the graphs in (d),(e), (f) of Figure 2. It follows that a graph G is a directed line graph if and only if it does not contain any subgraph in Figure 3 with the dotted arcs missing, and no subgraphs in (d),(e), (f) of

Figure 2. Notice that, by Lemma 5.2, after Theorem 5.1, a graph is a PDL graph if and only if no two nodes x, y lie in the same shore of two distinct kernels. This latter condition, as Theorem 6.1 shows, is the same as condition (7) when a PDL graph does not contain induced subgraphs in Figure 3 with the dotted arcs missing.

Lemma 6.1 *Let G be a 1-graph not containing as subgraphs any subgraphs in Figure 3 with the dotted arcs missing. Then if G contains a xy-alternating path of sign (α, β) it contains an xy-alternating path of sign (α, β) and with at most two arcs (depending on the parity).*

Proof. Assume G contains an xy alternating path P of length at least 3. This path must contain one of the solid paths in Figure 3 as subpath. Since those subpaths are allowed only if the dotted arc is in G as well, it follows that an xy alternating path with the same sign and shorter than P can be obtained by using the dotted arc. □

Theorem 6.1 *Let G be a directed graph. The following statements are equivalent.*

(1") *G is a directed line graph.*

(2") *G is a PDL graph satisfying Duchenne's Condition (6).*

(3") *G satisfies Condition (7).*

(4") *G does not contain any of the subgraphs in (d), (e), (f) of Figure 2 and any of those in Figure 3 with the dotted arc missing.*

(5") *G is a PDL graph not containing any of the subgraphs 3 with the dotted arcs missing.*

(6") *The family $\mathcal{K} = \{K_1, K_2, \dots, K_q\}$ of kernels of G, where S_i, T_i are the shores of K_i, is a covering of $V(G)$ such that $|S_i \cap T_j| \leq 1$ for each pair of kernels and all kernels are complete.*

Proof.

(1")⇒(2") Trivial.

(2")⇒(3") If G does not satisfy condition (7) then $N^+(x) \cap N^+(y) \neq \emptyset$ and $N^-(x) \cap N^-(y) \neq \emptyset$ for some $x, y \in V(G)$. Then G contains one of the subgraphs in (d),(e), (f) of Figure 2 contradicting that G is a PDL graph.[2]

(3")⇒(4") Just check that each subgraph in the statement violates the condition.

(4")⇒(5") G cannot contain any parallel pair (x, y) otherwise, by Lemma 6.1, the (even) alternating paths connecting x and y could be chosen of length two. Therefore, G would contain a subgraph in (d),(e), (f) of Figure 2. By Theorem 5.1 G is a PDL graph.

(5")⇒(6") By Theorem 5.1, $\mathcal{K}$ is a covering by kernels. Since G does not contain any of the subgraphs in Figure 3 with the dotted arc missing neither each $G[K_i]$ does. It follows that the each $G[K_i]$ must be complete. Indeed, suppose indirectly that for $x \in S_i$, $y \in T_i$ arc xy is missing in $G[K_i]$ and hence is missing in G. As $G[K_i]$ is connected an alternating xy-path exists in $G[K_i]$. In particular there is one containing a minimum number of arcs. This path has length at least three, contradicting (5").

(6")⇒(1") By Theorem (5.1) $\mathcal{K} = \{K_i, i \in I\}$ gives a root R_G of the PDL graph G where stars of R_G correspond to kernels of G; since kernels are complete every pair of consecutive arcs in any star of R_G correspond to arcs in kernels of G.

□

[2]The same contradiction, even if from a different result, is derived in the necessary implication of Corollary 2 in [1].

Theorem 6.1 can be stated as: *DL graphs are precisely those PDL graphs not containing any induced alternating P_3.*

Theorem 6.1, via Corollary 5.1, has the following interesting consequence, that solves Problem 2 posed in Section 1.

Corollary 6.1 *Let G be a PDL graph and let $[G]$ be the graph obtained from G by completing each of its kernels. Then $[G]$ is a directed line graph whose skeleton coincides with the skeleton of G. Moreover, if A_i is the directed line graph completion of the i-th kernel , $i = 1, \ldots, q$, then $\bigcup_i A_i$ is a minimum directed line graph completion of G.*

Proof. The first part of the statement follows directly from Theorem 6.1.(6") and by definition of skeleton. Let us prove the second part. Let A be any directed line graph completion of G, and let $G' = (V(G), E(G) \cup A)$ be the the resulting directed line graph. Let K be a kernel of G. Then $G'[K]$, is a directed line graph and $E(G'[K]) \setminus E(G[K])$ is a directed line graph completion of the PDL graph $G[K]$. Therefore, by Theorem 4.2 and Fact 4.2, $|E(G'[K]) \setminus E(G[K])| \geq |(S(K) \times T(K)) \setminus E(G[K])|$, the r.h.s. being a minimum directed line completion of $G[K]$. It follows that A must contain at least as many arcs as those occurring in the completion of any of its kernels. Since this latter number coincides with $|E([G])| - |E(G)|$, the thesis follows by Theorem 6.1.(6"). □

The proof of Corollary 6.1 also shows that any directed line graph completion of a $\rightleftarrows$-free graph G must contain the completion determined by the canonical root of G.

7 Recognition of partial directed line graphs

The characterization of PLD graphs based on condition (3) in Theorem 5.1 allows us to solve Problem 1, deciding whether a graph G is a PDL graph in $O(m)$ worst case time, where $m = |E(G)|$. A root graph of G, if it exists, can be built within the same time bound. The same algorithm can be adapted to recognize directed line graphs in $O(m)$ worst case time; this improves over the $O(n^3)$ algorithm proposed in [1], where $n = |V(G)|$, based on condition (7). Our algorithm builds the skeleton of a partial directed line graph; starting from the skeleton of a PDL graph, we also solve Problem 2, finding a description of a minimum completion to a directed line graph in $O(m)$ worst case time.

In order to decide whether a graph G is a PDL graph, we build the family $\mathcal{K} = \{K_i\}_{i=1}^q$ of kernels of G and the corresponding skeleton $\{S_i, T_i, X_i, \tilde{F}_i)\}_{i=1}^q$. Any time a node is found to belong to shore S_i (resp., T_i), we check whether $|S_i \cap T_j| \leq 1$ (resp., $|T_i \cap S_j| \leq 1$), for any $j < i$.

Kernels are built by finding all (α, β)-pairs (v, w), for all $v \in V(G)$, and for $\alpha, \beta \in \{+, -\}$. To this aim, we perform an "alternating" breadth first traversal, i.e., a breadth first traversal in which arcs are traversed in an alternating fashion: when a node x is visited by traversing an incoming arc zx (resp., an outgoing arc xz), we continue the visit by traversing all incoming arcs $yx \in E(G)$ (resp., outgoing arcs $xy \in E(G)$). Node x is marked as $(-)$-visited (resp., $(+)$-visited).

The algorithm works as follows. For each node x that is not a sink and has not been $(+)$-visited yet we start an alternating breadth first traversal at x, starting by outgoing arcs, and build a new kernel K_i. Node x and all nodes $(+)$-visited during this traversal are put in shore S_i, while all nodes $(-)$-visited during this traversal are put in T_i. Node x is marked as $(+)$-visited as well. Analogously, for each node x that is not a source and has not been $(-)$-visited yet, we start an alternating breadth first traversal at x, starting by incoming arcs, and build a new kernel K_i; x and all nodes $(-)$-visited during this traversal are put in shore T_i, while all nodes $(+)$-visited during this traversal are put in S_i. Node x is marked as $(-)$-visited.

The algorithm stops when all nodes with outgoing arcs have been $(+)$-visited and all nodes with incoming arcs have been $(-)$-visited, or when a violation of condition (3) in Theorem 5.1 is found. Sources (resp., sinks) will be only $(+)$-visited (resp., $(-)$-visited). Graph G is recognized as a PDL graph if and only if no violation is found.

In order to efficiently test whether condition (3) in Theorem 5.1 is satisfied, we maintain the following information:

- for each node, the shores it belongs to; by construction, any node belongs to at most one shore S_i and one shore T_j (possibly, $i = j$ for odd nodes);

- a global array $\mathbf{S}$ that represents the intersection between the currently built shore S_i and shores T_j's, $j < i$. At most n kernels are built, hence, array $\mathbf{S}$ has at most n entries. More precisely, the j-th entry of $\mathbf{S}$ contains a reference to S_i if and only if a node belonging both to S_i and T_j has been found, with $j < i$.

- a global array $\mathbf{T}$ that represents the intersection between the currently built shore T_i and shores S_j's, $j < i$. As above, array $\mathbf{T}$ has at most n entries. More precisely, the j-th entry of $\mathbf{T}$ contains a reference to T_i if and only if a node belonging both to T_i and S_j has been found, with $j < i$.

While traversing G, any time a node w is found to be in the current shore S_i, we store this information within node w. Moreover, we check whether w was already contained in some shore T_j, $j < i$; this information is stored within node w, as well. If so, we look at the j-th entry of array $\mathbf{S}$: if it already contains a reference to S_i then we know that w is the second node in the intersection between shores S_i and T_j, violating condition (3) in Theorem 5.1, otherwise we store a reference to S_i in the j-th entry of array $\mathbf{S}$. Nodes found to be in a shore T_i are dealt with analogously, replacing S's by T's and viceversa.

In case G is PDL graph, in order to build a root graph of G, we must also explicitly find odd nodes and backward arcs. This can be done during the same visit within the same worst case time. An odd node x is detected if x is both in S_i and in T_i for some kernel K_i. A little more work is required to find backward arcs; assume node x is being $(-)$-visited while building kernel K_i: for each arc $xy \in G$ we check whether $y \in S_i$, in this case arc xy is a backward in K_i. An analogous test is performed when x is $(+)$-visited, looking at arcs yx with $y \in T_i$.

Theorem 7.1 *It is possible to decide whether a graph G is a PDL graph in $O(m)$ worst case time, where $m = |E(G)|$. If G is a PDL graph it is also possible to build a canonical root of G within the same worst case time.*

Proof. The algorithm described above performs some alternating breadth first traversals of G, each starting from unvisited nodes. Each node is visited at most two times, and each arc is traversed a constant number of times. More precisely, each arc xy is traversed a first time while visiting one of its endpoints (thus visiting the opposite endpoint as well); one or two more traversals of xy may occur if x and/or y are flowing nodes, checking for the existence of a backward arc. $\square$

Once a graph G is found to be a PDL graph, we can check whether G is a directed line graph just by verifying that the directed graph induced by shores S_i and T_i is complete (in the sense of Theorem 6.1), for each kernel K_i. To this aim, it is sufficient to check that each node $x \in S_i$ has exactly $|T_i|$ outgoing edges in G. The size of each shore can be easily maintained during the alternating breadth first traversals, while building shores.

Concerning Problem 2, we showed in Corollary 6.1 that the canonical root of a $\rightrightarrows$-free graph G identifies a minimum directed line graph completion of G. Hence, Theorem 7.1 has the following consequence.

Corollary 7.1 *The minimum directed line graph completion of a PDL graph can be found in $O(m)$ worst case time.*

Actually, the size of the completion can be larger than m, since $\Theta(n^2)$ arcs can be needed to complete the bipartite graphs induced by kernels. The skeleton of the kernels, plus backward arcs and arcs incident on odd nodes, give a compact representation of the minimum completion that has size $O(n+m)$, and allows to enumerate the m_c arcs needed to complete the graph in $O(m_c)$ time, and also allows to easily check whether a given arc is contained or not in the minimum completion.

8 The undirected case: minimum line graph completion

Having seen that Problems 1 and 2 are solvable in strongly polynomial time, it is natural to ask whether the same holds for undirected graphs. We have already observed that Problem 1 is trivial for undirected simple graph: *every simple undirected graph is a partial graph of some simple line graph*. Indeed, every simple undirected graph is a partial graph of the complete graph on the same set of vertices and every complete graph is a line graph (its root is a star). Quite surprisingly, when undirected graphs are considered, Problem 2 is NP-hard. In Theorem 8.2 we show that the problem of finding a minimum bisection of a graph can be reduced in polynomial time to the problem of a finding minimum line graph completion.

A *minimum bisection* of a simple undirected graph G, with $|V(G)|$ even, is a partition of $V(G)$ into two sets A, B of equal size such that the number of edges having one endpoint in A and the other endpoint in B is minimum. This problem has been shown to be NP-hard in [7]. Recall that line graphs are characterized by Krausz's Theorem (see [2]).

Theorem 8.1 (see [2]) *A simple undirected graph G is a line graph if and only if $E(G)$ can be partitioned into complete subgraphs such that no node of G lies in more than two such subgraphs.*

Theorem 8.2 *The problem of finding a minimum line graph completion is NP-hard.*

Proof. Given an undirected graph G with order $n = |V(G)|$, n even, and size $m = |E(G)|$, we show how to derive a graph H, such that a minimum completion of H to a line graph determines a minimum bisection of G. Graph H is derived by substituting each vertex $v_i \in V(G)$ by a complete graph on $\kappa = n+1$ vertices $C_i = \{v_{i,1}, v_{i,2}, \ldots, v_{i,\kappa}\}$, and adding a single dummy vertex α that is adjacent to all vertices in each C_i. Edges in $E(G)$ are represented in H by edge set X: for each edge $(v_i, v_j) \in E(G)$ there is an edge in X from a vertex in C_i to a vertex in C_j; we do not want any two edges in X share any endpoint: this can be obtained by representing each pair $(v_i, v_j) \in E(G)$ by edge $(v_{i,j}, v_{j,i})$.

More precisely, H is built on the vertex set

$$V(H) = \{\alpha\} \cup C_1 \cup C_2 \cup \ldots \cup C_n \ ,$$

where $C_i = \{v_{i,1}, v_{i,2}, \ldots, v_{i,\kappa}\}$, $1 \le i \le n$, and its edge set is the union of three disjoint sets $E(H) = K \cup \Gamma \cup X$, where

$$K = \bigcup_{\substack{1 \le i \le n \\ 1 \le j < k \le \kappa}} \{(v_{i,j}, v_{i,k})\}$$

$$\Gamma = \bigcup_{\substack{i=1,2,\ldots,n \\ j=1,2,\ldots,\kappa}} \{(\alpha, v_{i,j})\}$$

$$X = \bigcup_{(v_i,v_j) \in E(G)} \{v_{i,j}, v_{j,i}\}$$

Assume we have a completion E' of H to a line graph and let H' be the resulting line graph. By Theorem 8.1, there exists a family $\mathcal{K}$ of cliques of H' such that any pair of cliques in the family intersects in at most one vertex, and each vertex is in at most two cliques. The only violation in H is due to vertex α, which is contained in n maximal cliques (each of the form $C_i \cup \{\alpha\}$, $1 \le i \le n$), while α should be in at most two cliques of $\mathcal{K}$. Let V_1 and V_2 be the (possibly empty) sets of vertices of the two cliques of $\mathcal{K}$ containing α. Again, by Theorem 8.1, each C_i must be contained either in V_1 or in V_2. To show this, assume $\kappa > 2$, $|V_1 \cap C_i| \ge 1$ and $|V_2 \cap C_i| \ge 1$ both hold and at least one inequality is strict. It follows that C_i contains a triangle T with at least one vertex in V_1 and at least one vertex in V_2. Suppose w.l.o.g. that $V(T) \cap V_1 = \{x\}$. Then the edges of T incident in x must be contained in a same clique C of $\mathcal{K}$, otherwise, x would be contained in more than two cliques. But C must contain also the edge of T not incident in x meaning that this edge is contained in two cliques of $\mathcal{K}$.

If we choose a balanced solution, where both V_1 and V_2 contain exactly $n/2$ sets C_i, we need to add no more than

$$M_{\text{bal}} = \frac{n}{2}\left(\frac{n}{2} - 1\right)\kappa^2 = \left(\frac{n^2}{4} - \frac{n}{2}\right)\kappa^2$$

edges to get the completion H', assuming no edge in X helps in getting the completion. On the other hand, if V_1 contains $n/2 + \beta$ sets C_i, $\beta \geq 1$, and V_2 contains $n/2 - \beta$ sets C_i, the number of edges to add is at least

$$M_{\text{unbal}} = \left(\frac{n^2}{4} - \frac{n}{2}\right)\kappa^2 + \beta^2\kappa^2 - m\ ,$$

where term m is an upper bound on the number of edges in X that could help in getting the completion. Since we have set $\kappa = n + 1 > \sqrt{m}$, we have $M_{\text{unbal}} > M_{\text{bal}}$ for any $\beta \geq 1$, hence any minimum completion of H is determined by a balanced partition of sets C_i, and it determines a (not necessarily minimum) bisection of G.

Any balanced partition actually needs exactly

$$M_{\text{bal}} - \left| X \setminus \left(\binom{V_1}{2} \cup \binom{V_2}{2} \right) \right|$$

edges for completing H to a line graph, where the negative term takes into account edges in X with both endpoints in V_1 or both endpoints in V_2.

Hence, among all the balanced partitions of sets C_i, the best one is the partition that maximizes the number of edges in X with both endpoints in V_1 or both endpoints in V_2, thus giving a minimum bisection of G. □

Acknowledgment: We are thankful to Bruno Simeone for the illuminating discussions on the subject.

References

[1] J.Blazewicz, A. Hertz, D. Kobler, D. de Werra, On some Properties of DNA Graphs. Discrete Applied Mathematics 98(1-2): 1-19 (1999).

[2] A. Brandstdt , V. Bang Le , J. P. Spinrad, Graph classes: a survey, Society for Industrial and Applied Mathematics, Philadelphia, PA, (1999).

[3] M. A. Fiol, A. S. Lladó, The Partial Line Digraph Technique in the Design of Large Interconnection Networks, IEEE Transactions on Computers 41(7):848 - 857 (1992).

[4] E. Prisner, A Journey through Intersection Graph County. http://www.math.uni-hamburg.de /spag/gd/mitarbeiter/prisner/Pris/Rahmen.html.

[5] J. L. Villar, The underlying graph of a line digraph, Discrete Applied Mathematics 37-38: 525-538 (1992).

[6] C. Benzaken, S. C. Boyd, P.L. Hammer, B. Simeone, Adjoints of Pure Bidirected Graphs, Proc. Fourteenth Southeastern Conf. on Com- binatorics, Graph Theory and Computing (Boca Raton, Fla., 1983). Congressus Numer. 39 (1983), 123-144.

[7] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete problems. Theoret. Comput. Sci., 1(3):237-267, (1976).