ALGEBRAIC COMBINATORICS

James East, Maximilien Gadouleau & James D. Mitchell

Structural aspects of semigroups based on digraphs

Volume 2, issue 5 (2019), p. 711-733.

<http://alco.centre-mersenne.org/item/ALCO_2019__2_5_711_0>

Algebraic Combinatorics
Volume 2, issue 5 (2019), p. 711–733
https://doi.org/10.5802/alco.56

Structural aspects of semigroups based on digraphs

James East, Maximilien Gadouleau & James D. Mitchell

Abstract Given any digraph D without loops or multiple arcs, there is a natural construction of a semigroup $\langle D \rangle$ of transformations. To every arc (a, b) of D is associated the idempotent transformation $(a \to b)$ mapping a to b and fixing all vertices other than a. The semigroup $\langle D \rangle$ is generated by the idempotent transformations $(a \to b)$ for all arcs (a, b) of D.

In this paper, we consider the question of when there is a transformation in $\langle D \rangle$ containing a large cycle, and, for fixed $k \in \mathbb{N}$, we give a linear time algorithm to verify if $\langle D \rangle$ contains a transformation with a cycle of length k. We also classify those digraphs D such that $\langle D \rangle$ has one of the following properties: inverse, completely regular, commutative, simple, 0-simple, a semilattice, a rectangular band, congruence-free, is $\mathcal{K}$-trivial or $\mathcal{K}$-universal where $\mathcal{K}$ is any of Green's $\mathcal{H}$-, $\mathcal{L}$-, $\mathcal{R}$-, or $\mathcal{J}$-relation, and when $\langle D \rangle$ has a left, right, or two-sided zero.

1. Introduction

A *transformation* of degree $n \in \mathbb{N}$ is a function from $\{1, \ldots, n\}$ to itself. A *transformation semigroup* is a semigroup consisting of transformations of equal degree and with the operation of composition of functions. For the sake of brevity we will denote $\{1, \ldots, n\}$ by $[n]$. We define $(a \to b)$ to be the transformation defined by

$$v(a \to b) = \begin{cases} b & \text{if } v = a \\ v & \text{otherwise} \end{cases}$$

where $a, b \in [n]$ and $a \neq b$. A *digraph* is an ordered pair (V, A), where V is a set whose elements are referred to as *vertices*, and $A \subseteq (V \times V) \setminus \{(v, v) : v \in V\}$ is a set of ordered pairs called *arcs*. We identify a transformation $(a \to b)$ with an arc (a, b) in a digraph and we refer to $(a \to b)$ as an arc. If D is a digraph, we denote by $\langle D \rangle$ the semigroup generated by the arcs of D, and we refer to such a semigroup as *arc-generated*.

If A is empty, then D is an empty digraph, and $\langle D \rangle$ is the empty semigroup. Because this is a degenerate case, and in order to simplify the statement of our results, we shall always assume that D is not empty (and hence, that $\langle D \rangle$ is not the empty semigroup).

Arc-generated semigroups were first introduced by John Rhodes in the 1960s [20, Definition 6.51], under the name *semigroups of flows*. In [20], Rhodes was largely concerned with determining the maximal subgroups of an arc-generated semigroup, and he conjectured that every such subgroup was isomorphic to a direct product of

Manuscript received 2nd March 2018, revised 26th November 2018 and 21st December 2018, accepted 21st December 2018.

Keywords. digraphs, flow semigroup of digraph, semigroups, monoids.

ISSN: 2589-5486 http://algebraic-combinatorics.org/

cyclic, alternating, and symmetric groups. This conjecture was recently proved in a remarkable paper [11] by Horváth, Nehaniv, and Podoski.

Arc-generated semigroups are also closely related to the *pebble motion problem* on graphs [16], which we briefly describe here. Let $G = (V, E)$ be a graph on n vertices and $P = \{1, \ldots, k\}$ be a set of pebbles, with $k < n$. An *arrangement of pebbles* is an injective mapping $S : P \to V$, which places every pebble on a different vertex. A *move* then consists of transferring a pebble $p \in P$ from a vertex u to an adjacent vertex v, provided there is no pebble present in v already. The main problem is, given two arrangements of pebbles S and S', to determine whether one can go from S to S' by using a sequence of moves. A canonical example of the pebble motion problem is the *fifteen-puzzle* on graphs [23], which corresponds to the case $k = n - 1$. Clearly, moving a pebble from u to v corresponds to applying the arc $(u \to v)$; as such, we shall use some of the results in [23] later in this paper. Conversely, one of our main results, namely Theorem 3.15, can be applied to the pebble motion problem.

Many famous examples of semigroups are arc-generated. Perhaps the best known example is the semigroup Sing_n of all non-invertible, or singular, transformations on $[n]$, which was shown to be arc-generated by J. M. Howie in [12]. Other examples include the semigroup of singular order-preserving transformations [1], and the so-called *Catalan semigroup* [9, 21], which are generated by the arcs of the digraphs $\{(i, i+1), (i+1, i) : i \in \{1, \ldots, n-1\}\}$ and $\{(i, i+1) : i \in \{1, \ldots, n-1\}\}$, respectively; these digraphs can be seen in Figure 1.

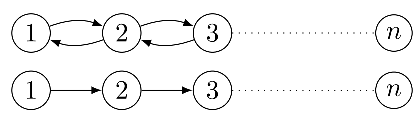


Figure 1. The digraphs D where $\langle D \rangle$ is the semigroup of singular order-preserving transformations (top) or the Catalan semigroup (bottom).

In [13], Howie showed that Sing_n is generated by $\frac{1}{2}n(n-1)$, but no fewer, arcs. In [8] it was shown that $\frac{1}{2}n(n-1)$ is the minimum size of any generating set for Sing_n whether it consists of arcs or not. It was shown in [13] that Sing_n is generated by the arcs of a digraph D if and only if D is strongly connected and D contains a tournament. As a corollary, the minimal-size idempotent generating sets of Sing_n are in one-one correspondence with the strongly connected tournaments on n vertices; these were enumerated by Wright [24]. In [6] it was shown that a digraph D is strongly connected and contains a tournament if and only if it contains a strongly connected tournament. Hence every idempotent generating set for Sing_n contains one of minimum size, something that is not true for generating sets of Sing_n, in general.

Several authors have classified those digraphs D such that $\langle D \rangle$ has a specific semigroup property. For instance, in [25] those digraphs D such that $\langle D \rangle$ is regular are classified; and in [5] those D where $\langle D \rangle$ is a band are classified. In [25, 26], necessary and sufficient conditions on digraphs D and D' are given so that $\langle D \rangle = \langle D' \rangle$ or $\langle D \rangle \cong \langle D' \rangle$, respectively. In this paper, we continue in this direction, by classifying those digraphs D for which the semigroup $\langle D \rangle$ has one of a variety of properties.

The paper is organised as follows. In Section 2 we review some relevant terminology and basic results about digraphs and semigroups. In Section 3 we investigate the presence of transformations with long cycles in arc-generated semigroups and classify those arc-generated semigroups that are $\mathcal{H}$-trivial. It is possible that Proposition 3.6 and the converse of Proposition 3.12 in Section 3 can be proved using

Theorem 1(5)(b) and Lemma 15 from [11]. Our proofs were produced independently of the results in [11], and are relatively concise and self-contained, and so we have included the proofs for the sake of completeness. Arc-generated semigroups that are $\mathcal{L}$-, $\mathcal{R}$- or $\mathcal{J}$-trivial are classified in Section 4. Further classes of arc-generated semigroups (including bands, completely regular semigroups, inverse semigroups, semilattices and commutative semigroups) are classified in Section 5. Finally, properties related to left and right zeros are classified in Section 6, which among other things, allows us to classify those arc-generated semigroups that are rectangular bands, simple, 0-simple, or congruence-free.

Many of the results in Sections 4, 5, and 6 were suggested by initial computational experiments conducted using the Semigroups package [17] for GAP [22].

2. Preliminaries

2.1. Digraphs. In this subsection, we review some terminology and basic results on digraphs. We refer the reader to [2] for an authoritative account of digraphs.

Unless otherwise stated, the vertex set of a digraph will be $[n]$ for some $n \in \mathbb{N}$.

The *in-degree* of a vertex v in a digraph D is the number of arcs of the form (u, v) in D; similarly, the *out-degree* of v is the number of arcs of the form (v, u) in D. A vertex v in a digraph D is called a *sink* if the out-degree of v is 0. A vertex is *isolated* if it has no incoming or outgoing arcs.

If $D = (V, A)$ is a digraph, and U is subset of the vertices V of D, then the *subdigraph of D induced by U* is the digraph with vertices U and arcs $A \cap (U \times U)$. In general, a *subdigraph* of $D = (V, A)$ is any digraph $D' = (V', A')$ with $V' \subseteq V$ and $A' \subseteq A \cap (V' \times V')$.

If $D = (V, A)$ is a digraph, and ε is an equivalence relation on V, then the *quotient digraph D/ε* is defined as follows. The vertex set is the set of all ε-classes of V, and if W, U are ε-classes, then D/ε has the arc (W, U) if and only if $W \neq U$ and D has an arc (w, u) for some $w \in W$ and $u \in U$.

A *walk* in a digraph is a finite sequence $(v_0, v_1, \ldots, v_r)$, $r \geqslant 1$, of vertices such that (v_i, v_{i+1}) is an arc for all $i \in \{0, \ldots, r-1\}$; the *length* of this walk is r. A *path* is a walk where all vertices are distinct. A *cycle* in a digraph is a walk where $v_0 = v_r$ and all other vertices are distinct. A digraph is called *acyclic* if it has no cycles.

A *graph* G is defined to be a digraph where (u, v) is an arc if and only if (v, u) is an arc in G. We refer to the pair of arcs above as the *edge* $\{u, v\}$. Vertices u and v of a graph G are *adjacent* if $\{u, v\}$ is an edge of G.

An induced subdigraph of a graph, is also a graph, which we refer to as an *induced subgraph.* A *spanning subgraph* (as opposed to an induced subgraph) of a graph $G = (V, A)$ is any graph $H = (V, B)$ where $B \subseteq A$.

The *degree* of a vertex in a graph is its in-degree, which equals its out-degree. If u and v are vertices of a graph G, then the *distance* from u to v is the length of a shortest path from u to v, if such a path exists.

If G and H are graphs, then H is a *minor* of G if H can be obtained by successively deleting vertices, deleting edges, or contracting edges of G (where contracting an edge corresponds to deleting it and then identifying its end vertices).

If v is a vertex of a digraph D, then the *strong component* of v is the induced subdigraph of D with vertices v and all u such that there is a path from u to v and from v to u. Every digraph is partitioned by its strong components, and the quotient of a digraph by its strong components is acyclic. If D only has one strong component, then it is *strongly connected.* A *terminal* component of a digraph D is a strong component C such that (u, v) is not an arc in D for all $u \in C$, $v \notin C$.

Alternatively, C is terminal if it is a sink in the quotient of a digraph by its strong components. A strong component of a digraph is *trivial* if it only has one vertex.

The *underlying graph* of a digraph D is the graph with an edge $\{u, v\}$ for each arc (u, v) of D. The *component* of v is the induced subdigraph of D with vertices u such that there is a path from u to v in the underlying graph of D. Every digraph is partitioned by its components; we say the digraph is *connected* if it only has one component.

A graph G is *separable* if it can be decomposed into two connected induced subgraphs G_1 and G_2 with exactly one vertex in common, where all edge of Gs belong to either G_1 or G_2; a graph is *non-separable* if G admits no such decomposition. A *block* of a graph is an induced subgraph that is non-separable and is maximal with respect to this property.

A graph G is *bipartite* if it can be decomposed into two subgraphs G_1 and G_2 such that every edge connects a vertex from G_1 with a vertex from G_2. A graph is *odd bipartite* if it is bipartite and it has an odd number of vertices.

We denote by K_n the *complete graph* with vertices $[n]$ and edges $\{u, v\}$ for all distinct $u, v \in [n]$; by $K_{k,1}$ the *star graph* with vertices $\{1, \dots, k + 1\}$ and edges $\{i, k + 1\}$ for all $i \in \{1, \dots, k\}$. We denote by P_n the *path graph*, or simply *path* if there is no ambiguity, with vertices $[n]$ and edges $\{i, i+1\}$ for all $i \in \{1, \dots, n-1\}$. We denote by C_n the *cycle graph* with vertices $\{1, \dots, n\}$ and edges $\{1, n\}$ and $\{i, i + 1\}$ for all $i \in \{1, \dots, n - 1\}$.

2.2. Semigroups and monoids.

In this subsection, we review some terminology about semigroups. We refer the reader to [14] and [7] for further background material about semigroups.

A *semigroup* is a set with an associative binary operation. A *monoid* is a semigroup S with an *identity*: i.e. an element $e \in S$ such that $es = se = s$ for all $s \in S$. If S is a semigroup, then $s \in S$ is an *idempotent* if $s^2 = s$. If S is a semigroup without identity, then we denote by S^1 the monoid obtained by adjoining an identity $1_S \notin S$ to S; if S is a monoid, then $S^1 = S$. A semigroup S is *regular* if for all $x \in S$ there exists $y \in S$ such that $xyx = x$. A *subsemigroup* of a semigroup S is a subset T of S that is also a semigroup with the same operation as S; denoted $T \leqslant S$.

A *congruence* on a semigroup S is an equivalence relation ε on S for which $(a, b) \in \varepsilon$ and $(c, d) \in \varepsilon$ imply $(ac, bd) \in \varepsilon$ for all $a, b, c, d \in S$. A semigroup S is *congruence-free* if the only congruences on S are the universal and trivial relations.

Let S be a semigroup and let $x, y \in S$ be arbitrary. We say that x and y are *$\mathcal{L}$-related* if the principal left ideals generated by x and y in S are equal; in other words, if $S^1x = \{sx : s \in S^1\} = S^1y = \{sy : s \in S^1\}$. We write $x\mathcal{L}y$ to denote that x and y are $\mathcal{L}$-related.

Green's $\mathcal{R}$-relation is defined dually to Green's $\mathcal{L}$-relation; Green's $\mathcal{H}$-relation is the meet, in the lattice of equivalence relations on S, of $\mathcal{L}$ and $\mathcal{R}$. Green's $\mathcal{J}$-relation is defined so that $x, y \in S$ are *$\mathcal{J}$-related* if $S^1xS^1 = S^1yS^1$. We will refer to the equivalence classes as $\mathcal{K}$-classes where $\mathcal{K}$ is any of $\mathcal{R}$, $\mathcal{L}$, $\mathcal{H}$, or $\mathcal{J}$. We write $x\mathcal{K}y$ to indicate $(x, y) \in \mathcal{K}$, where $\mathcal{K}$ is any of $\mathcal{R}$, $\mathcal{L}$, $\mathcal{H}$, or $\mathcal{J}$.

We denote by Tran_n the monoid consisting of all of the transformations of degree n where $n \in \mathbb{N}$; called the *full transformation monoid*. This monoid plays the same role in semigroup theory as the symmetric group does in group theory, in that every finite semigroup is isomorphic to a subsemigroup of some Tran_n. Green's relations on Tran_n can be described in terms of the following natural parameters associated to transformations. The *image* of a transformation $\alpha \in \mathrm{Tran}_n$ is the set

$$\mathrm{im}(\alpha) = \{x\alpha : x \in \{1, \dots, n\}\};$$

the *kernel* of α is the equivalence relation

$$\ker(\alpha) = \{(x, y) \in \{1, \ldots, n\} \times \{1, \ldots, n\} : x\alpha = y\alpha\};$$

and the *rank* of α is

$$\operatorname{rk}(\alpha) = |\operatorname{im}(\alpha)|.$$

It is well-known that two elements of Tran_n are $\mathcal{R}$-, $\mathcal{L}$- or $\mathcal{J}$- related if and only if they have the same kernel, image or rank, respectively; see [14, Exercise 2.6.16].

A semigroup is *aperiodic* if all of its subgroups are trivial. A semigroup is *$\mathcal{K}$-trivial* for $\mathcal{K} \in \{\mathcal{R}, \mathcal{L}, \mathcal{H}, \mathcal{J}\}$, if $x\mathcal{K}y$ implies $x = y$.

2.3. Arc-generated semigroups. We now characterise some basic semigroup theoretic properties of $\langle D\rangle$ in terms of digraph theoretic properties of D.

Suppose that D is a digraph with vertex set V, that $v \in V$ is an isolated vertex, and that D' is the subdigraph of D induced by $V \setminus \{v\}$. Then it is clear that the arc-generated semigroups $\langle D\rangle$ and $\langle D'\rangle$ are isomorphic. So, we may assume without loss of generality, where appropriate and if it is convenient, that a digraph D has no isolated vertices.

The following proposition will allow us to only consider connected digraphs in some cases; its proof is trivial and is omitted.

Proposition 2.1. *Let D be a digraph with components $D_1, D_2, \ldots, D_k$ and no isolated vertices. Then $\langle D\rangle$ is isomorphic to $\langle D_1\rangle^1 \times \cdots \times \langle D_k\rangle^1 \setminus \{(1_{\langle D_1\rangle}, \ldots, 1_{\langle D_k\rangle})\}$.*

The next result is also trivial.

Proposition 2.2. *Let D be a digraph. Then the following are equivalent:*

1. *$\langle D\rangle$ is trivial;*
2. *$\langle D\rangle$ is a group;*
3. *$\langle D\rangle$ has a unique $\mathcal{H}$-class;*
4. *D has only one arc.*

The semigroup $\langle D\rangle$ can contain arcs that are not present in D. It was shown in [25, Lemma 2.3] that the set of arcs in $\langle D\rangle$ is

$$\{(a \to b) : (a \to b) \in D \text{ or } (b \to a) \text{ belongs to a cycle of } D\}.$$

The *closure* of D, denoted $\bar{D}$, is the digraph on $[n]$ with the set of arcs as above; it is clear that $\langle \bar{D}\rangle = \langle D\rangle$. By construction, $\bar{D} = D$ if and only if every strong component C of D is a graph. We say that D is *closed* if $\bar{D} = D$.

3. Cyclic properties

A *cycle* of length k in $\alpha \in \mathrm{Tran}_n$ is a sequence of distinct points $a_0, a_1, \ldots, a_{k-1} \in [n]$ such that $a_i\alpha = a_{i+1}$ for all i, where the indices are computed modulo k. This section is concerned with cycles of transformations in an arc-generated semigroup. In particular, we are interested in the presence of *long* cycles. As mentioned in the introduction, the results in this section are related to those in [11], where the authors describe the structure and actions of the maximal subgroups of any arc-generated semigroup $\langle D\rangle$ in terms of properties of D. It is possible that Proposition 3.6 and the converse of Proposition 3.12 could be proved using Theorem 1(5)(b) and Lemma 15 from [11]. However, determining the properties of the specific digraphs in Propositions 3.6 and 3.12 required to apply the results in [11] is non-trivial, and requires the notation and terminology used in [11]. Since the proofs presented in this section are self-contained, and relatively concise, and were found independently of [11], we have opted not to use the results of [11].

One application of the results in this section is a classification of the digraphs D such that $\langle D\rangle$ is $\mathcal{H}$-trivial. In particular, we shall prove the following result.

Proposition 3.1. *Let D be a digraph. Then $\langle D\rangle$ is $\mathcal{H}$-trivial if and only if all the strong components of D are paths or isolated vertices.*

3.1. Preliminary results. The length of a longest cycle of α is denoted as $l(\alpha)$ and for a digraph D we write

$$l(D) = \max\{l(\alpha) : \alpha \in \langle D\rangle\}.$$

Lemma 3.2. *Let D be a digraph. Then the following are equivalent:*

(i) $l(D) = 1$*;*
(ii) $\langle D\rangle$ *is aperiodic;*
(iii) $\langle D\rangle$ *is $\mathcal{H}$-trivial.*

Proof. Conditions (ii) and (iii) are equivalent for any finite semigroup; see [18, Proposition 4.2].

(i) $\Rightarrow$ (ii). We prove the contrapositive. Suppose $\alpha \in \langle D\rangle$ belongs to a non-trivial subgroup and that α is not an idempotent. Then the restriction of α to $\mathrm{im}(\alpha)$ is a non-trivial permutation, so $l(\alpha) \geqslant 2$.

(ii) $\Rightarrow$ (i). Again, we prove the contrapositive. Suppose $\alpha \in \langle D\rangle$ has a cycle of length $k \geqslant 2$, say $a_0, a_1, \ldots, a_{k-1}$. Choose $r \geqslant 1$ such that α^r is an idempotent, and let H be the $\mathcal{H}$-class of α^r. Then H is a group, and $H = \{\alpha^s : s \geqslant r\}$. But α^r and α^{r+1} are distinct elements of H, since $a_0\alpha^r = a_0 \neq a_1 = a_0\alpha^{r+1}$. ☐

Since $\langle D\rangle = \langle \bar{D}\rangle$ for any digraph D, we clearly have $l(D) = l(\bar{D})$. Thus, when studying $l(D)$, we can assume without loss of generality that D is closed. If $\alpha \in \langle D\rangle$, then any cycle of α belongs entirely to a strong component of D. Therefore, if D has strong components $S_1, \ldots, S_r$, then

$$l(D) = \max\{l(S_1), \ldots, l(S_r)\}. \tag{1}$$

Thus, in this section we may assume without loss of generality, if it is convenient, that D is a connected graph.

Lemma 3.3. *For the cycle graph C_n, $n \geqslant 3$, we have $l(C_n) = n - 1$.*

Proof. Clearly, for any digraph G on n vertices, $l(G) \leqslant n - 1$. Conversely,

$$(n-1 \to n)(n-2 \to n-1)\cdots(1 \to 2)(n \to 1) \in \langle C_n\rangle$$

has the cycle $1, 2, \ldots, n-1$. ☐

If G' is a subgraph of G, then $\langle G'\rangle \leqslant \langle G\rangle$, and so $l(G') \leqslant l(G)$, which will allow us to isolate subgraphs of G in order to obtain lower bounds on $l(G)$. In the next lemma, we extend this result to graph minors.

Lemma 3.4. *If G is a graph and H is a minor of G, then $l(H) \leqslant l(G)$.*

Proof. For any $k \leqslant n$ and $S \leqslant \mathrm{Tran}_k$ and any $T \leqslant \mathrm{Tran}_n$, we write $S \preceq T$ if, relabelling the vertices of $\{1, \ldots, k\}$ if necessary, for any $\alpha \in S$, there exists $\beta \in T$ such that $v\beta = v\alpha$ for all $v \in \{1, \ldots, k\}$. It is clear that if $\langle G\rangle \preceq \langle H\rangle$, then $l(G) \leqslant l(H)$. Hence it suffices to show that $\langle H\rangle \preceq \langle G\rangle$.

This clearly holds if H is obtained from G by deleting an edge or a vertex. Suppose that H is obtained from G by contracting the edge $\{n-1, n\}$. Let B be the set of vertices that are adjacent to n but not to $n-1$ in G.

Let $\alpha \in \langle H\rangle$ be arbitrary. Then there exist arcs $\beta_1, \ldots, \beta_k \in H$ such that $\alpha = \beta_1\beta_2\cdots\beta_k$. If $\beta_i = (b \to n-1)$ for some $b \in B$, then we replace β_i in the product for α by $(b \to n)(n \to n-1)$. Similarly, we replace any arc $(n-1 \to b)$, $b \in B$, by $(n-1 \to n)(n \to b)$. If $\beta \in \langle G \rangle$ denotes this modified product, then $v\beta = v\alpha$ for all $v \in \{1, \ldots, n-1\}$, and so $\langle H \rangle \preceq \langle G \rangle$. □

LEMMA 3.5. *Let G be a graph. Then the following hold:*

(i) *if G has a vertex of degree k, then $l(G) \geqslant k-1$;*
(ii) *if G contains a subgraph that is a tree with k leaves, then $l(G) \geqslant k-1$;*
(iii) *if G is connected and t is the number of vertices of degree not equal to 2, then*

$$l(G) \geqslant \frac{1}{4}(t-2)+1.$$

Proof. (i). For distinct vertices u and v of the star graph $K_{k,1}$ such that $u, v \neq k+1$, we write $(u \rightsquigarrow v) = (u \to k+1)(k+1 \to v)$. Then

$$(k-1 \rightsquigarrow k)(k-2 \rightsquigarrow k-1) \cdots (1 \rightsquigarrow 2)(k \rightsquigarrow 1) \in \langle K_{k,1} \rangle$$

has the cycle $1, 2, \ldots, k-1$. The result now follows from Lemma 3.4.

(ii). If T is a tree with k leaves, then $K_{k,1}$ is a minor of T, so the result follows from Lemma 3.4 and part (i).

(iii). By [3], any graph with t vertices of degree not equal to 2 contains a spanning tree with at least $\frac{1}{4}(t-2)+2$ leaves. □

The next result concerns connected graphs that can be decomposed into two connected induced subgraphs with a path connecting them. With this in mind, we introduce a construction based on paths. Let L and R be two connected graphs on m and s vertices, respectively, where $m \leqslant s$, and let P be a path with q vertices. Let $L \oplus_q R$ denote the graph obtained by adding an edge between an endpoint of P and a vertex of L of degree not equal to 1, and an edge between the other endpoint of P to a vertex of R of degree not equal to 1. Even though this definition depends on the choice of attachment vertices, we will omit them in the notation, for our purpose is to derive results that do not depend on them, apart from the fact that they do not have degree 1 in L and R. We remark that $m, s \neq 2$, since the only connected graph on two vertices is K_2, whose vertices both have degree 1. However, it is possible to have $m = 1$ or $s = 1$.

The vertices of R are denoted as $r_1, \ldots, r_s$, where they are sorted in weakly increasing order of distance to the path P. In particular, r_1 is attached to P, and r_2 and r_3 are neighbours of r_1 if $s \neq 1$. A similar notation is used for L; in particular l_1 is attached to P. Write $L^* = L \smallsetminus \{l_1\}$, $R^* = R \smallsetminus \{r_1\}$, $P^* = P \cup \{l_1, r_1\}$, and order the elements of the path as $p_1, \ldots, p_q$ so that p_1 is adjacent to l_1, and p_q is adjacent to r_1; finally, we also write $l_1 = p_0$ and $r_1 = p_{q+1}$. For instance, the graph $K_{3,1} \oplus_4 C_4$ is illustrated in Figure 2.

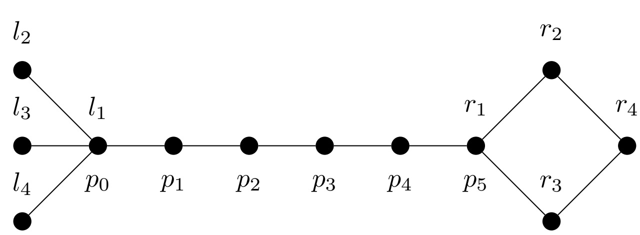


FIGURE 2. The graph $K_{3,1} \oplus_4 C_4$.

Proposition 3.6 (cf. Theorem 1(5)(b) in [11]). *With the above notation, if $q \geqslant s$, then*

$$l(L \oplus_q R) = \begin{cases} 1 & \text{if } m = s = 1 \\ s-1 & \text{if } m = 1, s \geqslant 3 \\ m+s-3 & \text{otherwise.} \end{cases}$$

Proof. We write $G = L \oplus_q R$. If $m = s = 1$, then $G = P_n$ where $n = q+2$, and hence $\langle P_n \rangle$ is the semigroup of order-preserving transformations [1], which is aperiodic. Otherwise, we have $s \geqslant 3$ and we view the result as two matching upper and lower bounds on $l(G)$.

Lower bound. Throughout this part of the proof, if $u, w_1, \dots, w_t, v$ is a path in G, then we define

$$(u \rightsquigarrow v) = (u \to w_1)(w_1 \to w_2) \cdots (w_{t-1} \to w_t)(w_t \to v) \in \langle G \rangle.$$

Since this transformation depends on the choice of path, we will always specify the path.

Case 1: $m = 1$. We will show that there exists $\alpha \in \langle G \rangle$ containing the cycle $r_2, \dots, r_s$.

For each $3 \leqslant i \leqslant s$, we choose $r'_i \in \{r_2, r_3\}$ such that there is a shortest-length path from r_1 to r_i that avoids r'_i. We also define

- $(r_2 \rightsquigarrow p_2)$ to follow the path $r_2, r_1, p_q, \dots, p_2$,
- $(p_{j-1} \rightsquigarrow r_j)$ to follow a shortest-length path avoiding r'_j and $(r_j \rightsquigarrow p_j)$ to follow the reverse of such a path (but omitting the last edge) for all $3 \leqslant j \leqslant s$,
- $(r'_j \rightsquigarrow r'_{j-1}) = (r'_j \to r_1)(r_1 \to r'_{j-1})$ for all $4 \leqslant j \leqslant s$, even if $r'_j = r'_{j-1}$,
- $(r_s \rightsquigarrow r'_s)$ to follow any path avoiding vertices from P.

It is straightforward to verify that

$$\alpha = \left[\prod_{i=2}^{s-1} (r_i \rightsquigarrow p_i) \right] \cdot (r_s \rightsquigarrow r'_s) \cdot \left[\prod_{j=s}^{4} (p_{j-1} \rightsquigarrow r_j)(r'_j \rightsquigarrow r'_{j-1}) \right] \cdot (p_2 \rightsquigarrow r_3) \in \langle G \rangle$$

contains the cycle $r_2, \dots, r_s$, as required (the second product is computed in descending order of the indices). We also note that $l_1 \alpha = l_1$.

Case 2: $m \geqslant 3$. As in Case 1, we may use $K_1 \oplus_q R \subseteq G$ to create $\alpha \in \langle G \rangle$ containing the cycle $r_2, \dots, r_s$ and such that $l_i \alpha = l_i$ for all i. Similarly, we may use $L \oplus_q K_1 \subseteq G$ to create $\beta \in \langle G \rangle$ containing the cycle $l_2, \dots, l_m$ and such that $r_i \beta = r_i$ for all i. If $(r_2 \rightsquigarrow l_3)$ and $(l_2 \rightsquigarrow r_2)$ follow the unique shortest paths, then $\gamma = \alpha\beta(r_2 \rightsquigarrow l_3)(l_2 \rightsquigarrow r_2) \in \langle G \rangle$ contains the cycle $l_3, \dots, l_m, r_2, \dots, r_s$.

Upper bound. Since $l(G_1) \leqslant l(G_2)$ if G_1 is a subgraph of G_2, we assume without loss of generality that $G = K_m \oplus_q K_s$. We define a pre-order $\preceq$ on the vertices of G such that $a \preceq b$ if $a \in L^*$, or $b \in R^*$, or $a = p_i$ and $b = p_j$ for some $0 \leqslant i \leqslant j \leqslant q+1$. If $a \preceq b$ and $b \not\preceq a$, then we write $a \prec b$ or $b \succ a$. We note that $b \not\preceq a$ implies $a \preceq b$. We define the sets

$$a^+ = \{b \in [n] : b \succ a\} \cup \{a\} \quad \text{and} \quad a^- = \{b \in [n] : a \succ b\} \cup \{a\}.$$

For the remainder of the proof, we fix some $\gamma \in \langle G \rangle$, and we write $\gamma = \epsilon_1 \cdots \epsilon_k$, where $\epsilon_1, \dots, \epsilon_k$ are arcs in G. We also define $\gamma_0 = \mathrm{id}$ and $\gamma_i = \epsilon_1 \cdots \epsilon_i$ for all $1 \leqslant i \leqslant k$.

Case 1: $m = 1$. We require the following claim.

Claim 3.7. *Suppose that $m = 1$. If there are vertices u and v in G such that $u \preceq v$ and $u\gamma \succ v\gamma$, then $\gamma \in \alpha\langle G \rangle$ for some $\alpha \in \langle G \rangle \cup \{\mathrm{id}\}$ such that $u^+\alpha \subseteq R^*$.*

Proof. If $u \in R^*$, then $u^+ = \{u\}$ and so $\alpha = \mathrm{id}$ has the required properties. Suppose that $u \notin R^*$. Since $u\gamma \succ v\gamma$, there exists i such that $u\gamma_i \in R^*$. If j is the least such value, then $u\gamma_{j-1} = r_1$, $u^+\gamma_{j-1} \subseteq R$, and $\epsilon_j = (r_1 \to r_a)$ for some $r_a \in R^*$. If we set $\alpha = \gamma_j$, then $u^+\alpha = (u^+\gamma_{j-1})\epsilon_j \subseteq R^*$. □

Seeking a contradiction, suppose that γ has a cycle of length c, where $c \geqslant s$, and let $C = \{u_1, \dots, u_c\}$ be such a cycle, where $u_1 \preceq \cdots \preceq u_c$. It is not necessarily the case that $u_i\gamma = u_{i+1}$. Since $|C| = c > s-1 = |R^*|$, C is not contained in R^* and so $u_1 \prec u_i$ for all $i \neq 1$. This gives $C \subseteq u_1^+$. Let j be such that $u_j\gamma = u_1$. Then $u_1 \prec u_j$ but $u_1\gamma \succ u_j\gamma$. So by Claim 3.7, $c = |C\gamma| \leqslant |u_1^+\gamma| \leqslant |R^*| = s-1$, which is the desired contradiction.

Case 2: $m \geqslant 3$. For the sake of obtaining a contradiction, suppose that γ has a cycle of length at least $m+s-2$. Let $C = \{u_1, \dots, u_c\}$ be such a cycle, sorted so that $u_1 \preceq \cdots \preceq u_c$. Again it is not necessarily the case that $u_i\gamma = u_{i+1}$. We note that $u_c \notin L$ (since $c \geqslant m+1$) and $u_1 \notin R$ (since $c \geqslant s+1$), whence $u_c \succ u_1$.

We say that a vertex v of G is of *L-type* if there is $\beta \in \langle G\rangle \cup \{\mathrm{id}\}$ such that $\gamma \in \beta\langle G\rangle$ and $v^-\beta \subseteq L^*$. Similarly, we say that v is of *R-type* if there is $\alpha \in \langle G\rangle \cup \{\mathrm{id}\}$ such that $\gamma \in \alpha\langle G\rangle$ and $v^+\alpha \subseteq R^*$.

Claim 3.8. *Suppose that $m \geqslant 3$. If there are vertices u and v in G such that $u \preceq v$ and $u\gamma \succ v\gamma$, then either u is of R-type, or v is of L-type.*

Proof. As in the proof of Claim 3.7, if $u \in R^*$, then $\alpha = \mathrm{id}$ witnesses that u is of R-type. Similarly, if $v \in L^*$ then $\beta = \mathrm{id}$ shows that v is of L-type.

Suppose that $u \notin R^*$ and $v \notin L^*$. As before, for some $i \in \{1, \dots, k\}$, $u\gamma_i$ and $v\gamma_i$ both belong to L^* or R^*. Suppose that both $u\gamma_i$ and $v\gamma_i$ belong to R^* before they both belong to L^*; the case when they both first belong to L^* is symmetric. Let i be the least value such that $u\gamma_i, v\gamma_i \in R^*$. Then for all $j < i$, $u\gamma_j \preceq v\gamma_j$ and either $u\gamma_j \notin R^*$ or $v\gamma_j \notin R^*$. If $u\gamma_j \in R^*$ for some $j < i$, then since $u\gamma_j \preceq v\gamma_j$, it follows that $v\gamma_j \in R^*$. Hence i is the least value such that $u\gamma_i \in R^*$.

We will show that $u\gamma_j \notin L^*$ for all $0 \leqslant j < i$. Seeking a contradiction, suppose that $u\gamma_j \in L^*$ for some $0 \leqslant j < i$, and let $b = \max\{j : j < i,\ u\gamma_j \in L^*\}$. The vertices $[n]$ of G can be partitioned into two parts:

$$A = (\{u\gamma_b\} \cup P^* \cup R^*)\gamma_b^{-1} \qquad \text{and} \qquad B = (L^* \setminus \{u\gamma_b\})\gamma_b^{-1}.$$

Let $x \in A$. If $x\gamma_b = u\gamma_b$, then $x\gamma_i = u\gamma_i \in R^*$. Otherwise, $u\gamma_b \prec x\gamma_b$. By maximality of b and minimality of i, we have $\epsilon_{b+1} = (u\gamma_b \to l_1)$, $\epsilon_i = (r_1 \to u\gamma_i)$, $u\gamma_j \in P^*$ for all $b < j < i$, and $u\gamma_j \preceq x\gamma_j$ for all $b < j < i$. It follows that $x\gamma_i \in R^*$. Therefore, $A\gamma_i \subseteq R^*$, and so

$$c \leqslant \mathrm{rk}(\gamma) \leqslant |B\gamma_b| + |A\gamma_i| \leqslant (m-2) + (s-1) = m+s-3,$$

which contradicts the fact that $c > m+s-3$.

We conclude that $u\gamma_j \in P^*$ for all $j < i$, and by the argument concluding the proof of Claim 3.7, we obtain $u^+\gamma_i \subseteq R^*$, so that u is of R-type. □

Claim 3.9. *There exist $u, v \in C$ such that $u \preceq v$, $u_c \preceq v$, and $u\gamma \succ v\gamma$. There also exist $u', v' \in C$ such that $u' \preceq v'$, $u' \preceq u_1$, and $u'\gamma \succ v'\gamma$.*

Proof. If $u_c \in R^*$, then since C intersects R^* but is not contained in R^*, there exist $u, v \in C$ such that $u \notin R^*$, $u\gamma \in R^*$, $v \in R^*$, and $v\gamma \notin R^*$. Then u and v have the required properties.

If $u_c \notin R^*$, then $u_c \succ u_i$ for all $1 \leqslant i \leqslant c-1$. In particular, if $u_c = u_j\gamma$, then $u = u_j$ and $v = u_c$ have the required properties.

The proof of the existence of u' and v' is symmetrical. □

We now write $X = \{u_i \in C : \exists j,\ u_i \preceq u_j,\ u_i\gamma \succ u_j\gamma\}$, and note that $X \neq \varnothing$ by Claim 3.9. We enumerate $X = \{x_1, \dots, x_d\}$ such that $x_1 \preceq \cdots \preceq x_d$. For each $1 \leqslant i \leqslant d$, let y_i be an element of C such that $x_i \preceq y_i$ and $x_i\gamma \succ y_i\gamma$; we also assume that y_i is maximal with respect to this property: that is, if $v \in C$ is such that $x_i \preceq v$ and $x_i\gamma \succ v\gamma$, then $v \preceq y_i$. Note that $\{y_1, \dots, y_d\}$ is not necessarily sorted according to the pre-order. If y_M is a maximal element of $\{y_1, \dots, y_d\}$ with respect to $\preceq$, then Claim 3.9 indicates that $u_c \preceq y_M$. We also have $x_1 \preceq u_1$, since $x_1 \preceq u'$, where u' is as in Claim 3.9.

Claim 3.10. *There exists $1 \leqslant a < d$ such that $y_1, \dots, y_a$ are all of L-type and $x_{a+1}, \dots, x_d$ are all of R-type. Moreover, for all $i > a \geqslant j$, $x_i \succ y_j$.*

Proof. We shall prove a sequence of facts about the set X, the last two of which give the claim.

(a) If x_i is of R-type, and if $x_i \preceq y_j$, then y_j is not of L-type.
Suppose to the contrary that we have the following: $x_i \preceq y_j$; $\alpha, \beta \in \langle G\rangle \cup \{\mathrm{id}\}$; $\gamma \in \alpha\langle G\rangle$ and $\gamma \in \beta\langle G\rangle$; $x_i^+\alpha \subseteq R^*$ and $y_j^-\beta \subseteq L^*$. We then have $x_i \notin L^*$, for otherwise $[n] = L^* \cup x_i^+$ and

$$c \leqslant \mathrm{rk}(\gamma) \leqslant |L^* \setminus \{x_i\}| + |x_i^+\alpha| \leqslant (m-2) + (s-1) = m+s-3,$$

a contradiction. Similarly, we have $y_j \notin R^*$. Thus

$$x_i, y_j \in P^*, \quad [n] = x_i^+ \cup y_j^-, \quad x_i \in x_i^+ \cap y_j^-.$$

Denoting $S = \{w \in y_j^- : w\gamma = x_i\gamma\}$, we have

$$c \leqslant \mathrm{rk}(\gamma) \leqslant |(x_i^+ \cup S)\gamma| + |(y_j^- \setminus S)\gamma| = |x_i^+\gamma| + |y_j^-\gamma| - 1 \leqslant (s-1)+(m-1)-1 = m+s-3,$$

a contradiction.

(b) For every i, x_i is of R-type if and only if y_i is not of L-type.
Apply (a) with $i = j$ and combine with Claim 3.8.

(c) If x_i is of R-type, then so too are $x_{i+1}, \dots, x_d$.
If x_i is of R-type and $i < j$, then because $x_i \preceq x_j \preceq y_j$, (a) says that y_j is not of L-type and (b) in turn says that x_j is of R-type.

(d) y_1 is of L-type and x_M is of R-type.
We prove that x_1 is not of R-type, which by (b) implies that y_1 is of L-type. Suppose that $x_1^+\alpha \subseteq R^*$ for some $\alpha \in \langle G\rangle \cup \{\mathrm{id}\}$ with $\gamma \in \alpha\langle G\rangle$. If $x_1 \in L^*$, then $x_1^+ = \{x_1\} \cup P^* \cup R^*$ and hence

$$c \leqslant \mathrm{rk}(\gamma) \leqslant |L^* \setminus \{x_1\}| + |(\{x_1\} \cup P^* \cup R^*)\alpha| \leqslant (m-2) + (s-1) = m+s-3,$$

a contradiction. If $x_1 \notin L^*$, then $x_1 = u_1$ (since $x_1 \preceq u_1$ and $u_1 \notin R^*$) and hence $C \subseteq x_1^+$, so that $c = |C\gamma| \leqslant |x_1^+\alpha| \leqslant s-1$, a contradiction. The proof for x_M is symmetrical.

(e) There exists $1 \leqslant a < d$ such that $y_1, \dots, y_a$ are all of L-type and $x_{a+1}, \dots, x_d$ are all of R-type.
This follows from combining (b), (c) and (d), with

$$a = \max\{i : y_1, \dots, y_i \text{ are of } L\text{-type}\}.$$

(f) For all $i > a \geqslant j$, $x_i \succ y_j$.
By (e), x_i is of R-type, and y_j of L-type. It follows from (a) that

$$x_i \succ y_j.$$ □

We now partition C into two parts A and B defined by

$$A = \{u \in C : x_{a+1} \succ u\} \qquad \text{and} \qquad B = \{v \in C : x_{a+1} \preceq v\}.$$

Note that A and B are both non-empty: for example, $x_{a+1} \in B$ and $y_a \in A$. Since C is a cycle of γ, $S\gamma \neq S$ for any non-empty proper subset S of C. In particular, $A\gamma \neq A$ and $B\gamma \neq B$, and so there exist $u \in A$ and $v \in B$ such that $u\gamma \in B$ and $v\gamma \in A$. It follows that $u \preceq v$ and $u\gamma \succ v\gamma$, and hence $u = x_j$ and $v \preceq y_j$ for some $j \leqslant a$. But then $x_{a+1} \preceq v \preceq y_j$, which contradicts Claim 3.10. ☐

3.2. Classification results. In this subsection, we give a classification of the connected graphs G for which $l(G)$ is equal to 1, 2 or $n-1$. From this, and in light of equation (1), it is easy to deduce such classifications for arbitrary graphs G. We also consider the computational complexity of determining whether a given graph G satisfies $l(G) \leqslant k$.

The classification of graphs with $l(G) = 1$ or $l(G) = 2$ is based on the following family of graphs. The graph Q_n for $n \geqslant 3$ is obtained by adding the edge $\{n-2, n\}$ to the path P_n, so that the three last vertices form a triangle (and indeed $Q_3 = K_3$). The graph R_n for $n \geqslant 4$ is obtained by removing the edge $\{n-1, n\}$ from Q_n, so that the last four vertices form the star graph $K_{3,1}$ (and indeed $R_4 = K_{3,1}$). The graphs Q_6 and R_6 are illustrated in Figure 3.

Figure 3. The graphs Q_6 (left) and R_6 (right).

A number of other graphs, pictured in Figure 4, will feature in the proofs. It can be shown, using GAP [22] for instance, that if G is the bull graph or the E-graph, then $l(G) = 3$ and that $l(\theta_0) = 6$.

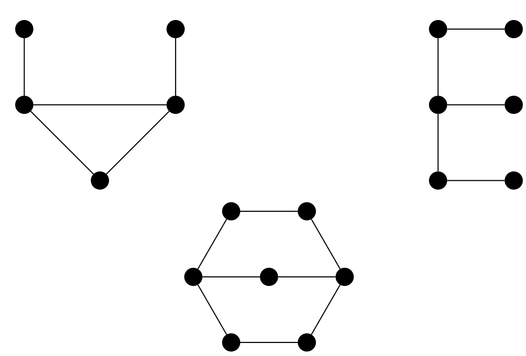


Figure 4. The bull graph (left), E-graph (right) and θ_0 graph (below).

Proposition 3.11. *Let G be a connected graph. Then $l(G) = 1$ if and only if G is a path.*

Proof. Since $\langle P_n \rangle$ is the semigroup of order-preserving transformations of $[n]$, it is aperiodic. Conversely, suppose that $l(G) = 1$. By Lemmas 3.3, 3.4 and 3.5, G is a tree with maximum degree 2, in other words, G is a path. ☐

Proposition 3.1 easily follows from Proposition 3.11 and equation (1).

Proposition 3.12. *Let G be a connected graph. Then $l(G) = 2$ if and only if G is Q_n $(n \geqslant 3)$ or R_n $(n \geqslant 4)$.*

Proof. Note that $Q_n = K_1 \oplus_{n-4} K_3$ and $R_n = K_1 \oplus_{n-4} K_{2,1}$. It follows from Proposition 3.6 that $l(Q_n) = l(R_n) = 2$ for $n \geqslant 7$; this can also be verified for $n \leqslant 6$, using GAP [22]. This part of the proof also follows from [11, Lemma 15].

Conversely, suppose that $l(G) = 2$; the case $n \leqslant 3$ is easy so let us assume $n \geqslant 4$. By Lemmas 3.3 and 3.4, G does not have any cycle of length 4 or more. By Lemma 3.5, G has no vertices of degree greater than 3.

Claim 3.13. *G has exactly one vertex of degree 3.*

Proof. If G has no vertex of degree 3, then it is a path and $l(G) = 1$, or it is a cycle and $l(G) = n - 1 \geqslant 3$. Thus, G has a vertex of degree 3, say x_1, with neighbours x_2, x_3 and x_4. First, suppose that x_2 also has degree 3. If x_3 and x_4 are both neighbours of x_2, then G has the cycle x_1, x_3, x_2, x_4. If x_2 is adjacent to x_3 and to another vertex, say x_5, then G contains a bull. Thus, x_2 is not adjacent to either x_3 or x_4, and instead is adjacent to x_5 and x_6, say, in which case, G contains a tree with leaves x_3, x_4, x_5, and x_6, so Lemma 3.5(ii) applies. So vertex x_2 does not have degree x_3 and, similarly, neither do vertices x_3 and x_4. Second, suppose that G contains another vertex of degree 3, say u, that is not a neighbour of x_1. There is a path from u to x_1, and we may assume this goes through vertex x_2; then by contracting the path from x_2 to u, and applying Lemma 3.4, we get back to the first case. □

We now split the rest of the proof into two cases. First, if G is a tree, then $G = R_n$ or G has the E-graph as a subgraph. The latter case would yield $l(G) \geqslant 3$, hence $G = R_n$. Second, if G is not a tree, then G has a triangle, say induced by the vertices a, b, c. One of them must be the vertex of degree 3, say a, and the other two have degree 2. Then $G = Q_n$. □

On the other extreme, we have the following classification. Recall that a graph G is *non-separable* if for every pair of vertices $u, v \in [n]$, there are at least two vertex-disjoint paths from u to v.

Proposition 3.14. *Let G be a connected graph. Then $l(G) = n - 1$ if and only if $G = K_2$ or G is non-separable and not odd bipartite.*

Proof. The case $n \leqslant 3$ being easily checked, we assume $n \geqslant 4$ throughout the proof.

Let G be non-separable. Recall the puzzle group $\Gamma_G(v)$ from [23], obtained as follows. First of all, create a hole at any vertex v. Then repeatedly slide a vertex a into the hole at vertex b, where a is adjacent to b; this moves the hole to a. Whenever the hole goes back to v, this yields a permutation of $[n] \setminus \{v\}$. The (abstract) group does not actually depend on v. Clearly, creating the hole at v can be done by using any arc $(v \to u)$ where u is a neighbour of v, and then sliding a vertex a to the hole in b is equivalent to using the arc $(a \to b)$. Therefore, for any initial hole v and any $g \in \Gamma_G(v)$ acting on $[n] \setminus \{v\}$, there exists $\alpha \in \langle G \rangle$ such that $ug = u\alpha$ for all $u \neq v$.

($\Leftarrow$) We have already noted that $l(\theta_0) = 6$, and that $l(C_n) = n - 1$. Let G be non-separable and neither a cycle nor the graph θ_0. According to [23, Theorem 2], $\Gamma_G(v) = \mathrm{Alt}_{n-1}$ if G is bipartite and $\Gamma_G(v) = \mathrm{Sym}_{n-1}$ otherwise. Therefore $l(G) = n-1$ if G is non-separable and not odd bipartite, or $l(G) \geqslant n-2$ if G is non-separable and odd bipartite.

($\Rightarrow$) We prove the contrapositive. Suppose first that G is non-separable and odd bipartite. To obtain a contradiction, suppose that $\beta \in \langle G \rangle$ has $l(\beta) = n - 1$. Due to the form of β, there exist u and v such that $u\beta = v\beta$, $u\beta^{-1} \neq \varnothing$ and $v\beta^{-1} = \varnothing$ (note that β acts as a cyclic permutation π on $[n] \setminus \{v\}$). Since $(v \to u)\beta = \beta$, we can assume that the first arc in any word expressing β is $(v \to u)$. This corresponds to creating a hole in v, and then expressing π as a member of $\Gamma_G(v)$, which is impossible since π is an odd permutation while $\Gamma_G(v) = \mathrm{Alt}_{n-1}$.

Now suppose that G is separable. So there exist $L, R \subseteq [n]$ and $v \in [n]$ such that $2 \leqslant |L| \leqslant |R|$, $L \cap R = \{v\}$, and for any edge $\{l, r\}$ of G with $l \in L$ and $r \in R$ we have $v \in \{l, r\}$; see for example [4, Theorems 5.1 and 5.2]. Then G is a minor of $L \oplus_n R$, which is itself a minor of $K_m \oplus_n K_s$, where $m = |L|$ and $s = |R|$. Let us abuse notation slightly, and define $K_2 \oplus_n K_s := K_1 \oplus_{n+1} K_s$. By Lemma 3.4 and Proposition 3.6,

$$l(G) \leqslant l(K_m \oplus_n K_s) = \begin{cases} 1 & \text{if } m = s = 2 \\ s-1 & \text{if } m = 2 < s \\ m+s-3 & \text{if } m, s > 2. \end{cases}$$

Thus, in all cases, $l(G) \leqslant n-2$. □

We remark that the proof of Proposition 3.14 (in conjunction with Proposition 3.14 itself) indicates that if G is non-separable and odd bipartite, then $l(G) = n-2$.

A classification of graphs G such that $l(G) \leqslant k$ for arbitrary k seems beyond reach at the moment. However, since these graphs form a minor-closed class, we can determine whether $l(G) \leqslant k$ in time $O(n^2)$ [15]. We show that in fact this can be done in linear time. Here, we consider a computational model where the atomic operations are integer operations. This is not only the consensus in the analysis of graph algorithms, but is also similar to the analysis of sorting algorithms, where typically the running time is given as the number of comparisons made (e.g. Merge-Sort makes $O(n \log n)$ comparisons; clearly each comparison could take $O(\log n)$ bitwise operations).

Theorem 3.15. *For any fixed k, deciding whether a connected graph G, given as an adjacency list, satisfies $l(G) \leqslant k$ can be done in $O(n)$ time.*

Proof. Let us refer to a maximal path in G consisting of vertices of degree 2 as a branch. If a branch does not belong to a non-separable block, then $G = L \oplus_q R$, where the branch is the path in the middle. We say that a branch is terminal if $L = K_1$ and non-terminal otherwise. We shall use the same notation as for Proposition 3.6.

The result is clear for $k = 1$ (Proposition 3.11), so suppose $k \geqslant 2$. The algorithm goes as follows.

(1) If $n \leqslant (k+2)(k+1)(2k-1)$, solve by brute force, i.e. by enumerating all elements of $\langle G \rangle$.
(2) If G is a path, then return *Yes*.
(3) If G has a vertex of degree at least $k+2$, then return *No* (Lemma 3.5).
(4) If G has at least $4k$–1 vertices of degree not 2, then return *No* (Lemma 3.5(iii)).
(5) If G has a non-separable block of size at least $k+3$, then return *No* (Proposition 3.14 and the remark after its proof).
(6) Let P be the longest branch of G. If P is terminal and has length at most $n-k-3$, or if P is non-terminal and has length at most $n-k-4$, then return *No*. Otherwise, return *Yes*.

If the first five properties are not satisfied, then the number of vertices of degree 2 is at least

$$n - t \geqslant (k+2)(k+1)(2k-1) + 1 - (4k-2) > (k+1)^2(2k-1).$$

On the other hand, the number of branches is at most $(k+1)t/2 \leqslant (k+1)(2k-1)$. Thus the longest branch P of G has length $q \geqslant k+2$.

First, suppose that P is terminal, i.e. $G = L \oplus_q R$ with $m = 1$ and $s \geqslant 3$. If $q \leqslant n-k-3$, then $s - 1 = n - q - 2 \geqslant k+1$ and $l(G) \geqslant l(G') = k+1$, where G'

is the subgraph of G induced by $L \cup P \cup \{r_1, \ldots, r_{k+2}\}$. Otherwise, $s \leqslant k+1$ hence $q \geqslant s$ and $l(G) = s-1 \leqslant k$.

Second, suppose that P is non-terminal: i.e. $G = L \oplus_q R$ with $s \geqslant m \geqslant 3$. If $q \leqslant n-k-4$, then $m+s-3 = n-q-3 \geqslant k+1$. Let

$$\mu = \min\left\{m, \left\lfloor \frac{k+4}{2} \right\rfloor\right\} \qquad \text{and} \qquad \sigma = k+4-\mu.$$

We then have

$$\sigma + \mu - 3 = k+1, \quad q \geqslant k+2 \geqslant \sigma \geqslant \mu \geqslant 3, \quad m \geqslant \mu, \quad s \geqslant \sigma$$

and $l(G) \geqslant l(G') = k+1$, where G' is the subgraph of G induced by $\{l_1, \ldots, l_\mu\} \cup P \cup \{r_1, \ldots, r_\sigma\}$. Otherwise, $m+s-3 \leqslant k$ hence $q \geqslant s$ and $l(G) = m+s-3 \leqslant k$.

Step 1 runs in $O(1)$ time; properties 2 to 4 are decidable in time $O(n)$. If the first four properties are not satisfied, the number m of edges of G is at most $\frac{1}{2}(k+1)t + n - t \leqslant 2n$. Then the following steps, which run in $O(n+m)$ (an algorithm to find the non-separable blocks in linear time is given in [10]), actually run in $O(n)$ time. □

We now give an interpretation of Theorem 3.15 in terms of the pebble motion problem, reviewed in the introduction. In general, the problem of determining whether S' can be obtained from S via a sequence of moves is NP-hard [19]. However, we can easily extend any pebble arrangement $T : P \to [n]$ into a singular transformation τ of $[n]$, by letting $\tau(p) = T(p)$ for all $p \in P$ and $\tau(q) = T(1)$ otherwise. Moreover, any sequence of moves corresponds to an element $\alpha \in \langle D \rangle$. Thus, S' can be obtained from S via a sequence of moves if and only if the corresponding transformations satisfy $\sigma' = \sigma\alpha$ for some $\alpha \in \langle D \rangle^1$ – in particular, if D is strongly connected, then this is equivalent to $\sigma' \mathcal{R} \sigma$. Theorem 3.15 then gives a linear-time algorithm to show that some arrangements cannot be obtained from S, solely based on how much they "shuffle" the pebbles. Indeed, for the sake of simplicity, let us assume that σ and σ' have the same image, say I. The restriction of α to I is then a permutation π, since $(p\sigma)\alpha = (p\sigma')$ for all $p \in P$. If π has a cycle of length $k > l(D)$, then the algorithm in the proof of Theorem 3.15 will return *No*.

4. Properties related to Green's relations

In this section we characterise some semigroup theoretic properties of $\langle D \rangle$ in terms of certain digraph theoretic properties of D. In Proposition 3.1, we classified the digraphs D for which $\langle D \rangle$ is $\mathcal{H}$-trivial. The purpose of this section is to give analogous classifications for Green's $\mathcal{R}$-, $\mathcal{L}$- and $\mathcal{J}$-relations in Propositions 4.3, 4.4 and 4.5, respectively.

The proof of [14, Proposition 2.4.2] gives the following.

Lemma 4.1. *Let T be a subsemigroup of a semigroup S, let $a, b \in T$, and suppose a, b are regular in T. Then the following hold:*

(i) *a, b are $\mathcal{R}$-related in T if and only if they are $\mathcal{R}$-related in S;*
(ii) *a, b are $\mathcal{L}$-related in T if and only if they are $\mathcal{L}$-related in S.*

Recall that two elements of Tran_n are $\mathcal{R}$-, $\mathcal{L}$-, or $\mathcal{J}$- related if and only if they have the same kernel, image, or rank, respectively.

Lemma 4.2. *Let D be a digraph. If D contains a cycle and $(a \to b)$ is an arc in that cycle, then $(b \to a) \in \langle D \rangle$ and $(a \to b)\mathcal{R}(b \to a)$.*

Proof. As noted earlier, it follows from [25, Lemma 2.3] that $(b \to a)$ belongs to $\langle D \rangle$. Since $(a \to b)$ and $(b \to a)$ are idempotents, and hence regular, and they have equal kernels, it follows follows from Lemma 4.1 that $(a \to b)\mathcal{R}(b \to a)$. □

Proposition 4.3. *Let D be a digraph. Then $\langle D\rangle$ is $\mathcal{R}$-trivial if and only if D is acyclic.*

Proof. ($\Rightarrow$) It follows immediately from Lemma 4.2 that if D contains a cycle, then it is not $\mathcal{R}$-trivial, and so the contrapositive of this implication holds.

($\Leftarrow$) Again we prove the contrapositive. Suppose that $\alpha, \beta \in \langle D\rangle$ are such that $\alpha \neq \beta$ and $\alpha\mathcal{R}\beta$. Then there exist $\gamma, \delta \in \langle D\rangle$ such that $\alpha\gamma = \beta$ and $\beta\delta = \alpha$, and there is $i \in [n]$ with $i\alpha \neq i\beta$. Hence $i\alpha\gamma = i\beta \neq i\alpha$ and $i\alpha = i\alpha\gamma\delta$. The former implies that D contains a non-trivial path from $i\alpha$ to $i\alpha\gamma$, and the latter that D contains a path from $i\alpha\gamma$ to $i\alpha$. Thus D contains a cycle. ☐

Proposition 4.4. *Let D be a digraph. Then $\langle D\rangle$ is $\mathcal{L}$-trivial if and only if the following hold:*

(i) *the out-degree of every vertex in D is at most* 1*; and*
(ii) *D contains no cycles of length greater than* 2*.*

Proof. ($\Rightarrow$) We prove the contrapositive (i.e. that if either (i) or (ii) is not true, then $\langle D\rangle$ is not $\mathcal{L}$-trivial). If there are distinct arcs $\alpha = (a \to b)$ and $\beta = (a \to c)$ in D, then α and β are regular, and have the same image, and so $\alpha\mathcal{L}\beta$. If D contains a cycle of length greater than 2, then $\langle D\rangle$ is not $\mathcal{H}$-trivial, by Proposition 3.1, and hence it is not $\mathcal{L}$-trivial.

($\Leftarrow$) Suppose that both (i) and (ii) both hold. We begin by making some observations about the elements of $\langle D\rangle$ and their action on the vertices of D.

Suppose that $x_0 \in [n]$ is an arbitrary vertex of D with out-degree 1. By the assumptions on the structure of D, there is a unique path

$$x_0 \to x_1 \to \cdots \to x_{k-1} \to x_k \tag{2}$$

in D starting at x_0, and where x_k has out-degree 0 or $(x_k \to x_{k-1})$ is an arc in D. Since there are no vertices in D with out-degree exceeding 1, it follows that $(x_i \to x_{i+1})$ is the only arc in D starting at x_i for every i. So, if $\gamma \in \langle D\rangle$, then

$$x_s\gamma = x_t \quad \text{and} \quad s \leqslant t \quad \text{for all} \quad s \leqslant k-1. \tag{3}$$

First, we will show that

(4)
$$\text{if } \gamma = \gamma_0(x_0 \to x_1)\gamma_1(x_1 \to x_2)\gamma_2 \cdots \gamma_{r-1}(x_{r-1} \to x_r)\gamma_r \text{ where } \gamma_0, \gamma_1, \dots, \gamma_r \in \langle D\rangle,$$
$$\text{then } x_0\gamma = x_s \text{ where either } s \geqslant r \text{ or } s = k-1 \text{ and } r = k$$

for all $0 \leqslant r \leqslant k$.

We proceed by induction on r. If $r = 0$, then $\gamma = \gamma_0$ and $x_0\gamma = x_l$ for some $l \geqslant 0$ by (3). If $r > 0$, then by induction there exists $l \geqslant r-1$ such that

$$x_0\gamma_0(x_0 \to x_1)\gamma_1(x_1 \to x_2)\gamma_2 \cdots \gamma_{r-1} = x_l.$$

Suppose first that $l \geqslant r$. Then $x_0\gamma = x_l(x_{r-1} \to x_r)\gamma_r = x_l\gamma_r$. If $l \leqslant k-1$, then (3) gives $x_l\gamma_r = x_s$ for some $s \geqslant l \geqslant r$. If $l = k$, then by the form of D, $x_l\gamma_r = x_k\gamma_r$ can only be one of x_k or x_{k-1}. On the other hand, if $l = r-1$, then $x_0\gamma = x_{r-1}(x_{r-1} \to x_r)\gamma_r = x_r\gamma_r$, and $x_r\gamma_r = x_m$ where $m \geqslant r$, again by (3), unless $r = k$ in which case it is possible that $x_r\gamma_r = x_{k-1}$. This completes the proof of (4).

Second, suppose that $\gamma \in \langle D\rangle$ and $x_0\gamma = x_r$ for some r. Since there is a unique path from x_0 to x_r in D, it follows that any factorisation of γ in the arcs of D must contain each of

$$(x_0 \to x_1), (x_1 \to x_2), \dots, (x_{r-1} \to x_r)$$

in this order. In other words,
(5)
$$\text{if } x_0\gamma = x_r \text{ for some } r, \text{ then } \gamma = \gamma_0(x_0 \to x_1)\gamma_1(x_1 \to x_2)\gamma_2 \cdots \gamma_{r-1}(x_{r-1} \to x_r)\gamma_r$$
for some $\gamma_0, \gamma_1, \dots, \gamma_r \in \langle D \rangle$.

We will now begin the proof of this implication in earnest. Suppose that there are $\alpha, \beta \in \langle D \rangle$ such that $\alpha \neq \beta$. We will show that α and β are not $\mathcal{L}$-related. Since $\alpha \neq \beta$, there exists $x_0 \in [n]$ such that $x_0\alpha \neq x_0\beta$. Since at least one of α, β does not fix x_0, it follows that the out-degree of x_0 is equal to 1. Suppose that $x_1, \dots, x_k$ are as in (2). From (3), $x_0\alpha = x_r$ and $x_0\beta = x_s$ for some $r, s \geqslant 0$. We may assume without loss of generality that $r > s$. We consider two cases separately.

Case 1: $s \leqslant k-2$ or $(x_k \to x_{k-1})$ is not an arc in D. By (5),
$$\alpha = \alpha_0(x_0 \to x_1)\alpha_1(x_1 \to x_2)\alpha_2 \cdots \alpha_{r-1}(x_{r-1} \to x_r)\alpha_r \quad \text{for some } \alpha_0, \alpha_1, \dots, \alpha_r \in \langle D \rangle.$$
It follows that if $\gamma \in \langle D \rangle$ is such that $\beta = \gamma\alpha$, then
$$\beta = \gamma\alpha = (\gamma\alpha_0)(x_0 \to x_1)\alpha_1(x_1 \to x_2)\alpha_2 \cdots \alpha_{r-1}(x_{r-1} \to x_r)\alpha_r.$$
But $x_0\beta = x_s$ and so (4) implies that $s \geqslant r$, contradicting the assumption that $s < r$. Hence $(\alpha, \beta) \notin \mathcal{L}$.

Case 2: $s = k-1$ and $(x_k \to x_{k-1})$ is an arc in D. Since $r > s$, it follows that $x_0\alpha = x_k$ and $x_0\beta = x_{k-1}$. By (5), there exist $\alpha_0, \alpha_1, \dots \alpha_k \in \langle D \rangle$ such that
$$\alpha = \alpha_0(x_0 \to x_1)\alpha_1(x_1 \to x_2)\alpha_2 \cdots \alpha_{k-1}(x_{k-1} \to x_k)\alpha_k,$$
and we may assume without loss of generality that α_k does not have $(x_k \to x_{k-1})$ as a factor. Suppose that $\gamma \in \langle D \rangle$ is arbitrary. By (4),
$$x_0\gamma\alpha_0(x_0 \to x_1)\alpha_1(x_1 \to x_2)\alpha_2 \cdots \alpha_{k-2}(x_{k-2} \to x_{k-1})\alpha_{k-1} \in \{x_{k-1}, x_k\}.$$
In either case, $x_0\gamma\alpha = x_k \neq x_{k-1} = x_0\beta$ and so $\beta \neq \gamma\alpha$ for any $\gamma \in \langle D \rangle$, which implies $(\alpha, \beta) \notin \mathcal{L}$. ☐

Proposition 4.5. *Let D be a digraph. Then the following are equivalent:*

(i) *$\langle D \rangle$ has at most one idempotent in every $\mathcal{L}$-class and every $\mathcal{R}$-class;*
(ii) *$\langle D \rangle$ is $\mathcal{J}$-trivial;*
(iii) *D is acyclic and the out-degree of every vertex in D is at most* 1.

Proof. (i) ⇒ (iii). If D had a vertex of out-degree greater than 1, then, as in the proof of Proposition 4.4, $\langle D \rangle$ would contain two distinct $\mathcal{L}$-related idempotents.

If D contained a cycle, then, by Lemma 4.2, $\langle D \rangle$ would contain two distinct $\mathcal{R}$-related idempotents.

(iii) ⇒ (ii). If D is acyclic and the out-degree of every vertex in D is at most 1, then, by Propositions 4.3 and 4.4, $\langle D \rangle$ is both $\mathcal{R}$- and $\mathcal{L}$-trivial. Hence $\langle D \rangle$ is $\mathcal{J}$-trivial.

(ii) ⇒ (i). Since $\langle D \rangle$ is $\mathcal{J}$-trivial, it is both $\mathcal{L}$- and $\mathcal{R}$-trivial. Hence every $\mathcal{L}$-class and every $\mathcal{R}$-class contains exactly one element, and, in particular, at most one idempotent. ☐

5. Other classical semigroup properties

A semigroup S is called *completely regular* if every element belongs to subgroup. Equivalently, a semigroup is completely regular if and only if every element is $\mathcal{H}$-related to an idempotent. A finite semigroup S is completely regular if and only if $x\mathcal{J}x^2$ for all $x \in S$.

If D is any digraph with at most 2 vertices, then $\langle D \rangle$ is a band. Hence in the next two results we will assume that the number of vertices in D is at least 3.

We say a digraph D is *directed-bipartite* if there is a partition of the vertices $[n]$ of D into two parts V_1 and V_2 such that every arc $(a \to b)$ of D satisfies $a \in V_1$ and $b \in V_2$.

Proposition 5.1. *Let D be a connected digraph with at least 3 vertices. Then the following are equivalent:*

(i) *$\langle D \rangle$ is a band;*
(ii) *$\langle D \rangle$ is completely regular;*
(iii) *D is directed-bipartite.*

Proof. (i) and (iii) are shown to be equivalent in [5, Theorem 2.12].

(i) $\Rightarrow$ (ii). This implication follows immediately, since every band is completely regular.

(ii) $\Rightarrow$ (iii). We prove the contrapositive. Suppose that D contains the arcs $(a \to b)$ and $(b \to c)$, where $a, b, c \in [n]$ are distinct, and consider $\alpha = (b \to c)(a \to b)$. Then $\mathrm{rk}(\alpha) = n - 1$ and $\mathrm{rk}(\alpha^2) = n - 2$, and so α and α^2 are not $\mathcal{J}$-related in Tran_n. It follows that α and α^2 are not $\mathcal{J}$-related in $\langle D \rangle$, and so $\langle D \rangle$ is not completely regular. □

Corollary 5.2. *Let D be a connected acyclic digraph with at least 3 vertices. Then the following are equivalent:*

(i) *$\langle D \rangle$ is a band;*
(ii) *$\langle D \rangle$ is completely regular;*
(iii) *$\langle D \rangle$ is regular;*
(iv) *D is directed-bipartite.*

Proof. It suffices to prove that (iii) implies (i), so we suppose that $\langle D \rangle$ is regular. Since D is acyclic, $\langle D \rangle$ is $\mathcal{R}$-trivial by Lemma 4.3. Since $\langle D \rangle$ is regular, it follows that every $\mathcal{R}$-class contains an idempotent. But every $\mathcal{R}$-class is of size 1, and so every element of $\langle D \rangle$ is an idempotent. In other words, $\langle D \rangle$ is a band. □

There are non-acyclic digraphs D such that $\langle D \rangle$ is regular but not a band. For example, if D is any strong tournament with $n \geqslant 3$ vertices, then $\langle D \rangle = \mathrm{Sing}_n$, and Sing_n is regular but not a band.

For $n \geqslant 1$, an *n-fan* is a connected acyclic digraph isomorphic to the digraph with arcs (i, n) for all $i \in \{1, \ldots, n-1\}$. A 1-fan is just a one-vertex digraph with no arcs. A picture of an n-fan can be seen in Figure 5.

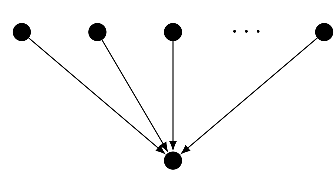


Figure 5. An n-fan.

A semigroup S is called *inverse* if for all $x \in S$ there exists a unique $y \in S$ such that $xyx = x$ and $yxy = y$. It is well-known (see, for example, [14, Theorem 5.1.1]) that a semigroup S is inverse if and only if it is regular and its idempotents commute. The same theorem from [14] also says that S is inverse if and only if each $\mathcal{R}$-class and each $\mathcal{L}$-class of S contains exactly one idempotent. A *semilattice* is a semigroup of commuting idempotents. For any set X, the *power set* $2^X = \{A : A \subseteq X\}$ of X is a semilattice under $\cup$; the subsemigroup $2^X \smallsetminus \{\varnothing\}$ is called the *free semilattice of degree* $|X|$. If $X_1, \ldots, X_k$ are disjoint finite sets, then $(2^{X_1} \times \cdots \times 2^{X_k}) \smallsetminus \{(\varnothing, \ldots, \varnothing)\}$ is a free semilattice of degree $|X_1| + \cdots + |X_k|$, isomorphic to $2^{X_1 \cup \cdots \cup X_k} \smallsetminus \{\varnothing\}$.

Proposition 5.3. *Let D be a connected digraph on n vertices. Then the following are equivalent:*

(i) $\langle D\rangle$ *is a free semilattice of degree* $n-1$*;*
(ii) $\langle D\rangle$ *is inverse;*
(iii) $\langle D\rangle$ *is commutative;*
(iv) D *is a fan.*

If any of the above conditions holds, then $|\langle D\rangle| = 2^{n-1}-1$.

Proof. (i) $\Rightarrow$ (ii). Every semilattice is an inverse semigroup.

(ii) $\Rightarrow$ (iii). Since $\langle D\rangle$ is inverse, it has exactly one idempotent in every $\mathscr{L}$- and $\mathscr{R}$-class, and hence by Proposition 4.5, $\langle D\rangle$ is $\mathscr{J}$-trivial. Thus $\langle D\rangle$ is a semilattice, and hence it is commutative.

(iii) $\Rightarrow$ (iv). Assume that $\langle D\rangle$ is commutative. If D contains distinct arcs $\alpha = (a\to b)$ and $\beta = (b\to c)$, then $\alpha\beta \neq \beta\alpha$, a contradiction. If D contains distinct arcs $\gamma = (d\to e)$ and $\delta = (d\to f)$, then $\gamma\delta\neq\delta\gamma$, a contradiction. Since D is connected, it follows that D is a fan.

(iv) $\Rightarrow$ (i). We may assume that the unique sink in D is the vertex n. If S is any subset of $\{1,\dots,n-1\}$, then we define $\alpha_S\in\mathrm{Tran}_n$ by

$$v\alpha_S = \begin{cases} n & \text{if } v\in S,\\ v & \text{if } v\notin S.\end{cases}$$

If $S=\{s_1,\dots,s_k\}$ is not empty, then $\alpha_S = (s_1\to n)\cdots(s_k\to n)\in\langle D\rangle$. Conversely, the arcs of D commute and so any transformation in $\langle D\rangle$ is of the form α_S for some non-empty subset S of $\{1,\dots,n-1\}$. If S and T are non-empty subsets of $\{1,\dots,n-1\}$, then it is routine to verify that $\alpha_S\alpha_T = \alpha_{S\cup T}$. It follows that the map $\phi : 2^{\{1,\dots,n-1\}}\setminus\{\varnothing\}\to\langle D\rangle$ defined by $S\phi = \alpha_S$ is an isomorphism. □

If D is a digraph with connected components $D_1,\dots,D_r$, then it follows from Proposition 2.1 that $\langle D\rangle$ is inverse if and only if each $\langle D_i\rangle$ is inverse. From this we obtain the following corollary to Proposition 5.3.

Corollary 5.4. *The number of digraphs (up to isomorphism) with n vertices such that $\langle D\rangle$ is an inverse semigroup is equal to $n-1$.*

Proof. Suppose that $\langle D\rangle$ is inverse, that the connected components of D are $D_1,\dots,D_r$, and write $d_i = |D_i|$ for each i. It follows from Proposition 5.3 that each D_i is a fan, and each $\langle D_i\rangle$ is a free semilattice of degree d_i-1. It follows that $\langle D\rangle$ is a free semilattice of degree $(d_1-1)+\cdots+(d_r-1) = n-r$. So the isomorphism class of $\langle D\rangle$ is completely determined by r, the number of connected (fan) components of D. Since r can take any value from 1 to $n-1$, the proof is complete. □

6. Zeros

An element a of a semigroup S is a *left zero* if $ab=a$ for all $b\in S$. Right zeros are defined analogously. An element is a *zero* if it is both a left and right zero. If a semigroup has a left zero and a right zero, then it has a unique zero. In this section, we obtain necessary and sufficient conditions on a digraph D so that $\langle D\rangle$ has various properties associated to left or right zeros. Some of our results also classify the digraphs D for which $\langle D\rangle$ consists of a single $\mathscr{R}$-, $\mathscr{L}$-, or $\mathscr{J}$-class (see Proposition 2.2 for the analogous result for the $\mathscr{H}$-relation). We begin with two technical lemmas.

Lemma 6.1. *If D is strongly connected, then $\langle D\rangle$ contains every constant map.*

Proof. Since D is strongly connected, it suffices to show that $\langle D\rangle$ contains any constant map. Suppose that there exists $\alpha \in \langle D\rangle$ with rank r, where $2 \leqslant r \leqslant n-1$. We prove that there exists $\beta \in \langle D\rangle$ with $\mathrm{rk}(\beta) < \mathrm{rk}(\alpha)$, from which the lemma follows. Since $\mathrm{rk}(\alpha) \geqslant 2$, there exist distinct $i, j \in \mathrm{im}(\alpha)$, and since D is strongly connected, there is a path in D from i to j. If $(i \rightsquigarrow j)$ denotes the product of the arcs in a path from i to j, then it is routine to check that $\mathrm{rk}(\alpha(i \rightsquigarrow j)) < \mathrm{rk}(\alpha)$, as required. ☐

Lemma 6.2. *If D is connected, then there exists $\alpha \in \langle D\rangle$ such that $i\alpha$ belongs to a terminal component of D for each $i \in [n]$.*

Proof. Suppose without loss of generality that $\{1, \dots, k\}$ are the vertices of D that do not belong to a terminal component of D. Then for every $i \in \{1, \dots, k\}$ there exists a vertex t_i in a terminal component of D such that there is a path in D from i to t_i. The product $\alpha = \prod_{i=1}^{k}(i \rightsquigarrow t_i)$ has the required property. ☐

Proposition 6.3. *Let D be a connected digraph. Then the following hold:*

(i) *$\langle D\rangle$ has a left zero if and only if all terminal components of D are trivial. If this is the case, then α is a left zero of $\langle D\rangle$ if and only if $v\alpha$ belongs to a terminal component for all vertices v.*
(ii) *$\langle D\rangle$ has a right zero if and only if it has exactly one terminal component. If this is the case, then α is a right zero of $\langle D\rangle$ if and only if it is a constant map.*
(iii) *$\langle D\rangle$ has a zero if and only if it has exactly one terminal component T, which is trivial. If this is the case, then the zero of $\langle D\rangle$ is the constant map with image T.*

Proof. (i). Suppose that $\alpha \in \langle D\rangle$ is a left zero, let T be an arbitrary terminal component, and let $a, b \in T$. By Lemma 6.1, there exist $\beta, \gamma \in \langle D\rangle$ such that $t\beta = a$ and $t\gamma = b$ for all $t \in T$. Let $t \in T$ be arbitrary. Since T is terminal, $t\alpha \in T$. But then, since α is a left zero, $a = (t\alpha)\beta = t(\alpha\beta) = t\alpha = t(\alpha\gamma) = (t\alpha)\gamma = b$. This shows that $|T| = 1$.

Conversely, suppose that all terminal components of D are trivial: $\{t_1\}, \dots, \{t_k\}$. By Lemma 6.2, there exists $\alpha \in \langle D\rangle$ such that $v\alpha$ belongs to a terminal component for each vertex $v \in [n]$. Now let $\beta \in \langle D\rangle$ be arbitrary. Let $v \in [n]$ and put $v\alpha = t_j$. Then $v\alpha\beta = t_j\beta = t_j = v\alpha$, so that $\alpha\beta = \alpha$, and α is a left zero.

On the other hand, suppose that all of the terminal components of D are trivial, and that $\alpha \in \langle D\rangle$ is such that $v\alpha$ does not belong to a terminal component for some $v \in [n]$. Then D contains some arc $(v\alpha \to j)$. But then $\alpha \neq \alpha(v\alpha \to j)$, so that α is not a left zero.

(ii). Suppose that D has at least two terminal components. Then there exist distinct terminal components T_1 and T_2 and a vertex v such that there is a path from v to a vertex from T_1 and a path from v to a vertex from T_2. For $i = 1, 2$, let $\beta_i \in \langle D\rangle$ be such that $v\beta_i \in T_i$. If $\alpha \in \langle D\rangle$ is a right zero, then

$$v\alpha = v\beta_1\alpha \in T_1 \quad \text{and} \quad v\alpha = v\beta_2\alpha \in T_2,$$

which is the desired contradiction.

Conversely, suppose that D has only one terminal component T and fix some $t \in T$. By Lemmas 6.1 and 6.2, there exist $\alpha_1, \alpha_2 \in \langle D\rangle$ such that $\mathrm{im}(\alpha_1) \subseteq T$, and $T\alpha_2 = \{t\}$. Then $\alpha_1\alpha_2 \in \langle D\rangle$ is a constant map (with image $\{t\}$), and hence a right zero.

On the other hand, suppose that D only has one terminal component T, and that $\alpha \in \langle D\rangle$ is not a constant map. So $\mathrm{rk}(\alpha) \geqslant 2$. We know that $\langle D\rangle$ contains some constant map β. But then $\mathrm{rk}(\beta\alpha) = \mathrm{rk}(\beta) = 1 \neq \mathrm{rk}(\alpha)$, so that $\beta\alpha \neq \alpha$, whence α is not a right zero. ☐

A semigroup S is called a *left zero semigroup* if every element of S is a left zero; *right zero semigroups* are defined analogously.

Proposition 6.4. *Let D be a digraph. Then the following are equivalent:*

(i) *$\langle D\rangle$ is a left zero semigroup;*
(ii) *$\langle D\rangle$ has a unique $\mathcal{L}$-class;*
(iii) *there is a unique non-trivial connected component K of D, and*

$$K^{-1} = \{(y \to x) : (x \to y) \in K\}$$

is a fan.

Proof. (i) ⇒ (ii). Every left zero semigroup has a unique $\mathcal{L}$-class.

(iii) ⇒ (iii). Since $\langle D\rangle$ has a unique $\mathcal{L}$-class, it follows that all of the arcs in D belong to the same $\mathcal{L}$-class. Hence the arcs in D have the same image: say, $[n] \setminus \{x\}$ for some fixed $x \in [n]$. In other words, the arcs in D are all of the form $(x \to y)$, $y \neq x$.

(iii) ⇒ (i). If the only arcs in D are of the form $(x \to y)$ for some fixed x, then $(x \to y)(x \to z) = (x \to y)$ for all y, z, and so $\langle D\rangle = D$ is a left zero semigroup. □

Proposition 6.5. *Let D be a digraph. Then the following are equivalent:*

(i) *$\langle D\rangle$ is a right zero semigroup;*
(ii) *$\langle D\rangle$ has a unique $\mathcal{R}$-class;*
(iii) *there is a unique non-trivial connected component K of D, and K has* 2 *vertices.*

Proof. (i) ⇒ (ii). Every right zero semigroup has a unique $\mathcal{R}$-class.

(ii) ⇒ (iii). Since $\langle D\rangle$ has a unique $\mathcal{R}$-class, all elements of D are $\mathcal{R}$-related. But $(i \to j)\mathcal{R}(k \to l)$ if and only if $\{i, j\} = \{k, l\}$ and so all of the arcs of D involve the same two vertices.

(iii) ⇒ (i). In this case, $\langle D\rangle$ is isomorphic to a subsemigroup of Sing_2, which is a right zero semigroup, and hence $\langle D\rangle$ is a right zero semigroup also. □

Recall that a semigroup is *simple* if it contains a single $\mathcal{J}$-class.

Proposition 6.6. *Let D be a connected digraph. Then the following are equivalent:*

(i) *$\langle D\rangle$ is a rectangular band;*
(ii) *$\langle D\rangle$ is simple;*
(iii) *$\langle D\rangle$ is a left or right zero semigroup.*

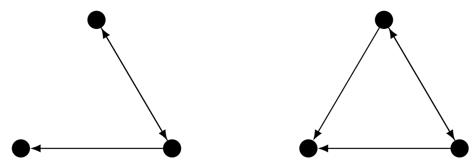
Proof. (iii) ⇒ (ii). Every left or right zero semigroup is simple.

(ii) ⇒ (i). If $\langle D\rangle$ is simple, then it is completely regular. If $n \geqslant 3$, then we conclude from Proposition 5.1 that $\langle D\rangle$ is a band; this is also true if $n \leqslant 2$. Every simple band is a rectangular band.

(i) ⇒ (iii). Suppose that $\langle D\rangle$ is a rectangular band. Since $\langle D\rangle$ has a single $\mathcal{J}$-class, every element of $\langle D\rangle$ has rank $n-1$. It follows that $\langle D\rangle$ consists entirely of arcs. If $\langle D\rangle$ is not a left or right zero semigroup, then there exist distinct arcs $\alpha, \beta, \gamma \in \langle D\rangle$ such that $\alpha\mathcal{R}\beta$ and $\beta\mathcal{L}\gamma$. The former implies that $\alpha = (a \to b)$ and $\beta = (b \to a)$ for some a, b, and the latter that $\gamma = (b \to c)$ for some c, where $a, b, c \in [n]$ are distinct. But then $\gamma\alpha \in \langle D\rangle$ is not an idempotent, a contradiction. □

Recall that a semigroup S with a zero element 0 is *0-simple* if $S^2 \neq \{0\}$ and its $\mathcal{J}$-classes are $\{0\}$ and $S \setminus \{0\}$.

Proposition 6.7. *Let D be a digraph. Then $\langle D\rangle$ is 0-simple if and only if the only non-trivial connected component of D is one of the following:*



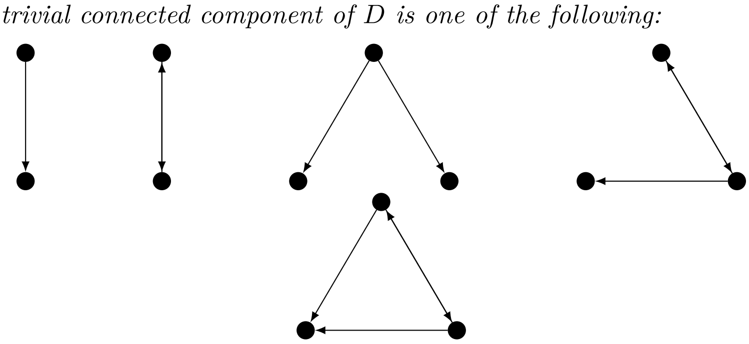
Proof. ($\Leftarrow$) It is straightforward to verify that $\langle D\rangle$ is 0-simple if D is either of the given digraphs, using GAP [22] for instance.

($\Rightarrow$) Suppose that $\langle D\rangle$ is 0-simple and, without loss of generality, that D has no isolated vertices. Since $\langle D\rangle$ is 0-simple, it has two $\mathscr{J}$-classes, and the minimum one contains only the zero element. In particular, $\langle D\rangle$ contains at most one element of rank smaller than $n-1$ and no elements of rank smaller than $n-2$.

Suppose that D has two connected components D_1 and D_2. If there is an arc α in D_1, and distinct arcs β,γ in D_2, then $\alpha\beta,\alpha\gamma\in\langle D\rangle$ are distinct elements of rank $n-2$, a contradiction. Hence, if D has more than one connected component, then every connected component has exactly one arc. In this case, $\langle D\rangle$ contains at least 2 arcs and the zero element and so $|\langle D\rangle|\geqslant 3$. But Proposition 4.5 implies that $\langle D\rangle$ is $\mathscr{J}$-trivial and so $\langle D\rangle$ has at least three $\mathscr{J}$-classes, a contradiction. Thus D is connected.

If $\alpha\in\langle D\rangle$ is the zero element, then by Proposition 6.3, α is constant which implies that $1=\operatorname{rk}(\alpha)\geqslant n-2$, and so $n\leqslant 3$. It is possible to check that if D' is any digraph with at most 3 vertices such that $\langle D'\rangle$ is 0-simple, then D' is isomorphic to one of the two given digraphs. □

Proposition 6.8. *Let D be a digraph. Then $\langle D\rangle$ is congruence-free if and only if the only non-trivial connected component of D is one of the following:*



Proof. Let D_1, D_2, D_3, D_4, D_5 (left to right) be the digraphs in the statement of the proposition.

($\Leftarrow$) The semigroups $\langle D_1\rangle, \langle D_2\rangle, \langle D_3\rangle$ have size at most 2, and so are congruence-free. It is straightforward to verify that $\langle D_4\rangle$ and $\langle D_5\rangle$ are both congruence-free (using GAP [22] for instance).

($\Rightarrow$) If $\langle D\rangle$ is congruence-free, then either $|\langle D\rangle|\leqslant 2$, $\langle D\rangle$ is 0-simple, or $\langle D\rangle$ is a simple group; see [14, Theorems 3.7.1 and 3.7.2]. So, by Propositions 2.2 and 6.7, it suffices to note that the only digraphs D so that $|\langle D\rangle|=2$ are the digraphs D_2 and D_3. □

References

[1] A. Ya. Aĭzenshtat, *The defining relations of the endomorphism semigroup of a finite linearly ordered set*, Sib. Mat. Zh. **3** (1962), 161–169.

[2] Jørgen Bang-Jensen and Gregory Z. Gutin, *Digraphs*, Springer Monographs in Mathematics, Springer London, 2009.

[3] A. V. Bankevich and Dmitriy V. Karpov, *Bounds of the number of leaves of spanning trees*, Journal of Mathematical Sciences **184** (2012), no. 5, 564–572.

[4] Adrian Bondy, *Graph Theory*, Graduate Texts in Mathematics, Springer, 2008.

[5] P. J. Cameron, A. Castillo-Ramirez, M. Gadouleau, and J. D. Mitchell, *Lengths of words in transformation semigroups generated by digraphs*, Journal of Algebraic Combinatorics **45** (2017), no. 1, 149–170.

[6] Igor Dolinka and James East, *Idempotent generation in the endomorphism monoid of a uniform partition*, Communications in Algebra **44** (2016), no. 12, 5179–5198.

[7] Olexandr Ganyushkin and Volodymyr Mazorchuk, *Classical finite transformation semigroups*, Algebra and Applications, vol. 9, Springer London, 2009.

[8] Gracinda M. S. Gomes and John M. Howie, *On the ranks of certain finite semigroups of transformations*, Mathematical Proceedings of the Cambridge Philosophical Society **101** (1987), no. 3, 395–403.

[9] Peter M. Higgins, *Combinatorial results for semigroups of order-preserving mappings*, Mathematical Proceedings of the Cambridge Philosophical Society **113** (1993), no. 2, 281–296.

[10] John Hopcroft and Robert Tarjan, *Algorithm 447: efficient algorithms for graph manipulation*, Communications of the ACM **16** (1973), no. 6, 372–378.

[11] Gábor Horváth, Chrystopher L. Nehaniv, and Károly Podoski, *The maximal subgroups and the complexity of the flow semigroup of finite (di)graphs*, International Journal of Algebra and Computation **27** (2017), no. 7, 863–886.

[12] John M. Howie, *The subsemigroup generated by the idempotents of a full transformation semigroup*, J. London Math. Soc. **41** (1966), no. 1, 707–716.

[13] ———, *Idempotent generators in finite full transformation semigroups*, Proceedings of the Royal Society of Edinburgh: Section A Mathematics **81** (1978), no. 3-4, 317–323.

[14] ———, *Fundamentals of semigroup theory*, London Mathematical Society Monographs. New Series, vol. 12, The Clarendon Press Oxford University Press, New York, 1995, Oxford Science Publications.

[15] Ken-ichi Kawarabayashi, Yusuke Kobayashi, and Bruce Reed, *The disjoint paths problem in quadratic time*, Journal of Combinatorial Theory, Series B **102** (2012), no. 2, 424–435.

[16] Daniel Kornhauser, Gary Miller, and Paul Spirakis, *Coordinating pebble motion on graphs, the diameter of permutation groups, and applications*, 25th Annual Symposium on Foundations of Computer Science, 1984., IEEE, 1984, pp. 241–250.

[17] James D. Mitchell et al., *Semigroups - GAP package, Version 3.1.1*, 2019.

[18] Jean Eric Pin, *Varieties of Formal Languages*, Foundations of Computer Science, Springer, 1986.

[19] Daniel Ratner and Manfred Warmuth, *The ($n^2 - 1$)-puzzle and related relocation problems*, Journal of Symbolic Computation **10** (1990), no. 2, 111–137.

[20] John Rhodes, *Applications of automata theory and algebra: Via the mathematical theory of complexity to biology, physics, psychology, philosophy, and games*, World Scientific Pub Co Inc, 2009.

[21] Andrew Solomon, *Catalan monoids, monoids of local endomorphisms, and their presentations*, Semigroup Forum **53** (1996), no. 1, 351–368.

[22] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.10.1*, 2019, `https://www.gap-system.org`.

[23] Richard M. Wilson, *Graph puzzles, homotopy, and the alternating group*, Journal of Combinatorial Theory, Series B **16** (1974), no. 1, 86–96.

[24] E. M. Wright, *The number of irreducible tournaments*, Glasgow Mathematical Journal **11** (1970), no. 2, 97–101.

[25] Xiuliang Yang and Haobo Yang, *Maximal regular subsemibands of* Sing_n, Semigroup Forum **72** (2005), no. 1, 75–93.

[26] ———, *Isomorphisms of transformation semigroups associated with simple digraphs*, Asian-European Journal of Mathematics **2** (2009), no. 4, 727–737.

JAMES EAST, Centre for Research in Mathematics, School of Computing, Engineering and Mathematics, Western Sydney University, Locked Bag 1797, Penrith NSW 2751, Australia.
E-mail : `j.east@westernsydney.edu.au`

MAXIMILIEN GADOULEAU, Department of Computer Science, Durham University, South Road, Durham, DH1 3LE, UK.
E-mail : `m.r.gadouleau@durham.ac.uk`

JAMES D. MITCHELL, School of Mathematics and Statistics, University of St Andrews, St Andrews, Fife KY16 9SS, UK.
E-mail : `jdm3@st-andrews.ac.uk`