

GEOMETRIC SEMIGROUP THEORY

JON MCCAMMOND, JOHN RHODES, AND BENJAMIN STEINBERG

ABSTRACT. Geometric semigroup theory is the systematic investigation of finitely-generated semigroups using the topology and geometry of their associated automata. In this article we show how a number of easily-defined expansions on finite semigroups and automata lead to simplifications of the graphs on which the corresponding finite semigroups act. We show in particular that every finite semigroup can be finitely expanded so that the expansion acts on a labeled directed graph which resembles the right Cayley graph of a free Burnside semigroup in many respects.

CONTENTS

1. Introduction	2
2. The topology of directed graphs	2
2.1. Directed graphs	3
2.2. Morphisms of directed graphs	6
2.3. Semigroups and automata	7
2.4. Rooted graphs	9
2.5. The unique simple path property	14
2.6. Cutting sloops	19
2.7. The McCammond cover of a graph	24
2.8. McCammond expansion of automata	27
3. Automata from a universal algebra point-of-view	29
3.1. Presentations and rewriting	31
3.2. The standard Kleene expression	33
4. Semigroup expansions	38
4.1. Semigroups	38
4.2. Straightline automata	39
4.3. Expansions	42
4.4. Mal'cev expansions	46
4.5. Rectangular bands	50
4.6. Improving stabilizers	55
5. The McCammond expansion revisited	58
5.1. Properties of the M^c -expansion	58

Date: April 13, 2011.

The first author was partially supported by the National Science Foundation. The third author gratefully acknowledges the support of NSERC and the DFG.

6. Algebraic rank function	62
References	65

1. INTRODUCTION

Geometric semigroup theory is the systematic investigation of finitely-generated semigroups using the topology and geometry of their associated automata. An early example of this approach is the article [24] where the first author proved that the Burnside semigroups

$$\mathcal{B}(m, n) = \langle A \mid x^m = x^{m+n} \rangle$$

(for fixed $m \geq 6$ and $n \geq 1$) are finite \mathcal{J} -above, have a decidable word problem, and their maximal subgroups are cyclic. In addition, and perhaps most importantly, the Brzozowski conjecture — that the equivalence classes of elements form regular languages — was verified in this range. Independent proofs of these results were obtained by A. de Luca and S. Varricchio [10] at about the same time, and shortly thereafter additional cases were covered by A. do Lago [11] and V. Guba [16]. The techniques used in these other papers, however, were predominantly combinatorial in nature. In the present article we wish to generalize the geometric nature of the arguments used in [24]. In particular, we will prove the following:

Theorem 1.1 (Rough statement). *If S is a finite A -semigroup, then there is a finite expansion of S which acts faithfully on a labeled directed graph which has many of the nice geometric properties possessed by the right Cayley graph of a Burnside semigroup $\mathcal{B}(m, n)$, $m \geq 6$.*

In the course of the article, we will explicate the particular expansions involved and the precise nature of the resemblance.

The proofs in [24] involved a detailed examination of the automata which recognize the equivalence classes of words in A^+ under the relations defining a Burnside semigroup. In particular, for each word $w \in A^+$, the equivalence class $[w]$ of words equal to w in $\mathcal{B}(m, n)$ was described as the language accepted by a non-deterministic finite-state automaton with a fractal-like structure. The deterministic version of this automaton is a full subautomaton of the right Cayley graph of $\mathcal{B}(m, n)$, and its geometric properties were described in some detail by the second author in [30]. The main theorem will follow immediately once we have shown that every finite automata can be finitely expanded, so that the expanded automata closely resembles these “McCammond automata” as described in [24] and [30].

2. THE TOPOLOGY OF DIRECTED GRAPHS

We view here automata from two angles: as labeled directed graphs and as universal algebras. The former viewpoint is geometric, whereas the latter is algebraic. This section establishes the essential properties of the topology

of directed graphs that we shall need. Probably nothing in the first two subsections is original although maybe our slant is different. To some extent it follows [37].

2.1. Directed graphs. We begin with the definition of a directed graph.

Definition 2.1 (Graph). A (directed) *graph* Γ consists of a set $V(\Gamma)$ of vertices, $E(\Gamma)$ of edges and two maps $\iota, \tau: E(\Gamma) \rightarrow V(\Gamma)$ selecting the initial, respectively, terminal vertices of an edge e . Often we write $e: v \rightarrow w$ to indicate $\iota(e) = v$ and $\tau(e) = w$.

Directed and undirected paths in a graph are defined in the usual way. If p is an undirected path, then $\iota(p), \tau(p)$ will denote the initial and terminal vertices of p , respectively and we shall write $p: \iota(p) \rightarrow \tau(p)$. We admit an empty path at each vertex. When we say “path” without any modifier, we mean a directed path, although we may include the word “directed” for emphasis. A directed or undirected path is called (vertex) *simple* if it visits no vertex twice; empty paths are considered simple. By a *circuit* we mean a non-empty closed path p (i.e., $\iota(p) = \tau(p)$). A circuit is called *simple* if the only repetition in the vertices it visits is when it returns to its origin. An undirected path is called *reduced* if it contains no backtracking (i.e., no subpath of length 2 using an edge first in one direction and then in the other). The inverse of an undirected path is defined in the usual way.

A graph is *connected* if there is an undirected path from any vertex to any other. A connected graph is called a *tree* if it contains no reduced undirected circuits. By an *induced* or *full* subgraph we mean a subgraph obtained by considering some subset of vertices and *all* edges between them.

There is a natural preorder on the vertices of any directed graph.

Definition 2.2 (Accessibility order). Let Γ be a graph. Define a preorder on $V(\Gamma)$ by $v \prec w$ if there is a path from w to v . A graph is termed *acyclic* if \prec is a partial order (equivalently, there are no directed circuits in Γ). As usual, an equivalence relation can be obtained from \prec by setting $v \sim w$ if $v \prec w$ and $w \prec v$. A *strong component* of a graph is an induced (or full) subgraph of Γ obtained by considering all vertices in a \sim -equivalence class.

A graph Γ is *strongly connected* if it has a unique strong component. In general, the set of strong components of Γ is partially ordered by putting $C \geq C'$ if there is a path from a vertex in C to a vertex in C' .

It is convenient to divide strong components into two sorts: trivial and non-trivial.

Definition 2.3 (Trivial strong components). A strong component can contain no edge. In this case, it consists of a single vertex and we shall call it *trivial*. Hence by a *non-trivial* strong component, we mean a strong component with at least one edge. In particular, a strong component with a single vertex and one or more loops edges is considered non-trivial.

Remark 2.4. Notice that if $v \prec w$, then there is a simple path from w to v . Indeed, let $p: w \rightarrow v$ be a minimum length path. If p is not simple, we may factor it $p = uts$ with t a directed circuit (in particular t is non-empty). Then $us: w \rightarrow v$ is shorter than p , a contradiction. So p must be simple.

Other authors order vertices via the opposite convention. Our choice was made to be compatible with Green's relations in semigroups. That is, the accessibility order on the right Cayley graph of a semigroup (to be defined shortly) corresponds to the $\leq_{\mathcal{R}}$ ordering on the semigroup.

2.1.1. *Preordered sets.* It is convenient here to introduce some terminology from the theory of preordered sets. If (P, \leq) is a preordered set, then a *downset* is a subset $X \subseteq P$ such that $x \in X$ and $y \leq x$ implies $y \in X$. Downsets are also called *order ideals* by some authors. If $Y \subseteq P$, then Y^\downarrow denotes the downset generated by Y . A downset of the form p^\downarrow with $p \in P$ is called *principal*. One can define *upsets* (called *filters* by some authors) dually. The upset generated by Y will be denote Y^\uparrow . A subset X of a preordered set P is said to be *convex* if $x \prec y \prec z$ and $x, z \in X$ implies $y \in X$. The *equivalence classes* of P are defined by $p \sim q$ if $p \leq q$ and $q \leq p$. A preordered set P is a *chain* if any two elements of P are comparable.

A *principal series* for P is an unrefinable chain

$$P = P_0 \supset P_1 \supset \cdots \supset P_n \tag{2.1}$$

of principal downsets. Every finite preordered set has a principal series. A principal series for a poset amounts to a topological ordering of the poset. In general, one can verify that $P_i \setminus P_{i+1}$ is always an equivalence class of P and that every equivalence class must arise this way.

Suppose now that V is the vertex set of a directed graph Γ and order V by the accessibility order. Then a downset in V is a subset X of vertices with the property that if the initial vertex of an edge belongs X , then so does the terminal vertex. The equivalence classes of the accessibility ordering are the strong components and a principal series amounts to the same thing as a topological ordering on the strong components. Therefore, when we assign indices to the strong components, we use notation consistent with (2.1).

2.1.2. *Transition edges.* An important role in the theory is played by those edges that go between strongly connected components. One facet of geometric semigroup theory is to simplify the structure of these edges.

Definition 2.5 (Transition edge). An edge e of a graph Γ is called a *transition edge* if $\tau(e) \approx \iota(e)$ (or equivalently, there is no directed path from $\tau(e)$ to $\iota(e)$). The *frame* $\text{Fr}(\Gamma)$ of Γ is the graph with vertex set the strong components of Γ and edge set the transition edges of Γ (i.e., we contract each strong component to a point). If e is a transition edge, then e starts at the strong component of its initial vertex and ends at the strong component of its terminal component. Evidently, $\text{Fr}(\Gamma)$ is an acyclic graph.

The edge set E of a directed graph Γ is also a preordered set. One can define $e < f$ if either $e = f$, or there is a directed path in Γ of the form epf where p is some path. Notice that distinct edges e, f are equivalent if and only if they belong to the same strong component. Transition edges are precisely those edges belonging to a singleton equivalence class. In particular, the transition edges form a poset.

It is particularly important in geometric semigroup theory when the strong components form a chain.

Definition 2.6 (Quasilinearity). If Γ is a directed graph we say that Γ is *quasilinear* if the natural partial order on its strong components is a total ordering. Vertices in the top-most strong component of a quasilinear directed graph (if one exists) will be called *top-most vertices*. In a finite graph, when the strong components are linearly ordered in this fashion we will number them starting with 0 and beginning with the top-most component to be consistent with (2.1).

Of course, a quasilinear graph is connected and every strongly connected graph is quasilinear. The class of quasilinear graphs for which the transition edges form a chain plays a salient role in geometric semigroup theory. In this case $\text{Fr}(\Gamma)$ looks like a line, whence the following terminology.

Definition 2.7 (Linearity). A quasilinear graph Γ is said to be *linear* if its transition edges form a chain.

Remark 2.8. If Γ is a finite linear graph with exactly $k+1$ strong components (numbered $0, 1, \dots, k$), then it has exactly k edges which do not belong to strong components, (which then necessarily connect the $(i-1)^{\text{st}}$ strong component to the i^{th} strong component, $i = 1, \dots, k$).

Definition 2.9 (Entry and exit points). If Γ is a finite linear graph, the unique transition edge connecting the $(i-1)^{\text{st}}$ strong component to the i^{th} will be called the *i^{th} transition edge*. Its start point will be denoted q_{i-1} and its end point will be denoted p_i . Notice that the subscripts indicates the strong component which contains the vertex. Since the vertices p_i and q_i are the places where transition edges enter and exit the i^{th} strong component, we will sometimes refer to these vertices as the *entry and exit points* of the i^{th} component. Note that $p_i = q_i$ is possible.

If $p = p_0$ is a specified vertex in the 0^{th} strong component and $q = q_k$ is a specified vertex in the k^{th} strong component, then any simple directed path in Γ from p to q will be called a *quasi-base* for Γ . Notice that every quasi-base for Γ will contain the transition edges plus simple paths in each strong component connecting p_i to q_i , and that conversely, any choice of simple paths connecting p_i to q_i , $i = 0, \dots, k$ can be strung together with the transition edges to form a quasi-base. In the case that there is a unique simple path from p to q (i.e., there is a unique simple path from p_i to q_i for each i) we shall call the corresponding quasi-base a *base*.

Example 2.10. An example of a finite linear graph has been schematically drawn in Figure 2.1. The shaded areas are meant to represent non-trivial strong components. In this example, there are 6 strong components and there are 5 transition edges, and their numbering has been illustrated. The first strong component is trivial (i.e., has no edges). The well-defined entry and exit points for each strong component have also highlighted. Notice that in this example p_3 and q_3 are identical, so that any base must use the trivial path to connect p_3 to q_3 .

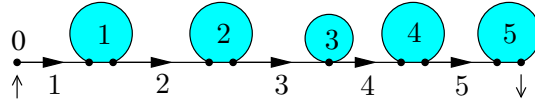


FIGURE 2.1. A directed graph with a quasi-base.

2.2. Morphisms of directed graphs. A *morphism* of graphs $\varphi: \Gamma \rightarrow \Gamma'$ consists of a pair (φ_V, φ_E) of maps $\varphi_V: V(\Gamma) \rightarrow V(\Gamma')$, $\varphi_E: E(\Gamma) \rightarrow E(\Gamma')$ so that $\varphi_V(\iota(e)) = \iota(\varphi_E(e))$ and $\varphi_V(\tau(e)) = \tau(\varphi_E(e))$ for all $e \in E(\Gamma)$. Normally we use φ to denote both maps. There is an obvious way to extend φ from edges to paths.

Two especially important classes of morphisms are *directed coverings* and *directed immersions*. If $v \in V(\Gamma)$, then the *star* of v is $St(v) = \iota^{-1}(v)$.

Definition 2.11 (Directed coverings and immersions). A graph morphism $\varphi: \Gamma \rightarrow \Gamma'$ is called a *directed covering* if it is surjective on vertices and, for each vertex $v \in V(\Gamma)$, the induced map $\varphi: St(v) \rightarrow St(\varphi(v))$ is a bijection. If φ merely injective on stars, the φ is called a *directed immersion*. We do not require directed immersions to be surjective on vertices.

Remark 2.12. Notice that if φ is a directed immersion, then $\varphi\psi$ is a directed immersion if and only if ψ is a directed immersion.

The following topological proposition is proved by straightforward induction on the length of a path.

Proposition 2.13 (Path lifting). *Let $\varphi: \Gamma \rightarrow \Gamma'$ be a graph morphism. Then φ is a directed immersion if and only if, for each (directed) path p at $v' \in V(\Gamma')$ and each $v \in \varphi^{-1}(v')$, there is at most one path q at v with $\varphi(q) = p$. The map φ is a directed covering if and only if it is surjective on vertices and, for each (directed) path p at $v' \in V(\Gamma')$ and each $v \in \varphi^{-1}(v')$, there is a unique path q at v with $\varphi(q) = p$.*

Notice that there is a bijection between sets and graphs with a single vertex. Hence we will frequently identify an alphabet A with the “bouquet” graph \mathcal{B}_A consisting of a single vertex with edge set A . A *labeling* of Γ by an alphabet A is then a graph morphism $\ell: \Gamma \rightarrow \mathcal{B}_A$. We can now define automata using our topological language cf. [37]. See [12, 13, 21] for background on automata theory.

Definition 2.14 (Automaton). A non-deterministic *automaton* over the alphabet A is a pair $\mathcal{A} = (\Gamma, \ell)$ where Γ is a graph and $\ell: \Gamma \rightarrow \mathcal{B}_A$ is a labeling. A morphism of automata $\varphi: (\Gamma_1, \ell_1) \rightarrow (\Gamma_2, \ell_2)$ is a graph morphism $\varphi: \Gamma_1 \rightarrow \Gamma_2$ so that

$$\begin{array}{ccc} \Gamma_1 & \xrightarrow{\varphi} & \Gamma_2 \\ & \searrow \ell_1 & \swarrow \ell_2 \\ & & \mathcal{B}_A \end{array}$$

commutes.

An automaton (Γ, ℓ) is called *deterministic* if ℓ is a directed covering; it is termed a *partial deterministic* automaton if ℓ is a directed immersion. By an automaton or A -automaton, we shall mean a partial deterministic automaton. In the context of automata, vertices are often called states and edges are termed transitions. If $\mathcal{A} = (\Gamma, \ell)$ is an A -automaton with vertex set Q , we sometimes abusively write $\mathcal{A} = (Q, A)$.

By Remark 2.12 any morphism of A -automata is a directed immersion. If $\mathcal{A} = (\Gamma, \ell)$ is an A -automaton, then we can associate to it its transition monoid $M(\mathcal{A})$. Namely, we can define a “monodromy” action of the free monoid A^* generated by A on $V(\Gamma)$ via path lifting. If $q \in V(\Gamma)$ is a vertex and $w \in A^*$ then we can view w as a path p in \mathcal{B}_A . This path has at most one lift \tilde{p} with $\iota(\tilde{p}) = q$ by Proposition 2.13. Define $qw = \tau(\tilde{p})$ if \tilde{p} exists, and leave it undefined otherwise. One easily checks that this defines an action of A^* on $V(\Gamma)$ by partial functions; the associated faithful partial transformation monoid is denoted $M(\mathcal{A})$ and is called the *transition monoid* of A . Note that the action is by total functions if and only if \mathcal{A} is deterministic. We denote by $\eta_{\mathcal{A}}$ the *transition morphism* $\eta_{\mathcal{A}}: A^* \rightarrow M(\mathcal{A})$. Notice that $p \prec q$ if and only if $p \in q \cdot M(\mathcal{A})$. Thus the accessibility order on $V(\Gamma)$ corresponds to the inclusion ordering on cyclic $M(\mathcal{A})$ -invariant subsets. The subsemigroup of $M(\mathcal{A})$ generated by A is called the *transition semigroup* and is denoted $S(\mathcal{A})$.

2.3. Semigroups and automata. An important example of a deterministic automaton is the Cayley graph of an A -semigroup.

Definition 2.15 (A -semigroup). An A -semigroup is a pair (S, φ) where $\varphi: A^+ \rightarrow S$ is a surjective homomorphism, where A^+ denotes the free semigroup on A . To avoid reference to φ , we put $[w]_S = \varphi(w)$ for $w \in A^+$.

If S is a semigroup, then S^I denotes S with an adjoined identity I (even if S was already a monoid). If S is an A -semigroup, it is convenient to consider the empty word as mapping to I .

Definition 2.16 (Cayley graph). Let S be an A -semigroup. The *right Cayley graph* $\text{Cay}(S, A)$ of S with respect to A is the deterministic automaton with vertex set S^I and edge set $S^I \times A$. Here $\iota(s, a) = s$ and $\tau(s, a) = sa$. The edge (s, a) is usually drawn $s \xrightarrow{a} sa$. The advantage of this “monoid”

Cayley graph is that the non-trivial paths from I correspond exactly to the words in A^+ and two such paths have the same endpoints if and only if the words corresponding to these paths represent the same element in S .

The strong components of these graphs are also of interest.

Definition 2.17 (Schützenberger graphs). If S is an A -semigroup then the strong components of $\text{Cay}(S, A)$ are the *Schützenberger graphs* of the \mathcal{R} -classes of S^I . In other words, the vertex set of a strong component contains exactly those vertices which represent the elements in an \mathcal{R} -class, and the strong component itself is the full subgraph of $\text{Cay}(S, A)$ on this vertex set. For each word $w \in A^+$, we will denote the strong component of $\text{Cay}(S, A)$ containing the vertex labeled $[w]_S$ by $\text{Sch}^S(w)$ (where Sch stands for Schützenberger). Suppose $[w]_S = s$. Since $\text{Sch}^S(w)$ only depends on the element s and not on the word w , we sometimes write $\text{Sch}^S(s)$ instead. The study of the way in which S acts on its Schützenberger graphs has been dubbed the *semilocal theory*. See Chapter [20, Chapter 8] and [32, Chapter 4].

Directed coverings correspond to transformation semigroup homomorphisms. In this paper, a *partial transformation semigroup* is a pair (X, S) where S is a semigroup acting faithfully on the right of X by partial transformations. Recall that if (X, S) and (Y, T) are partial transformation semigroups, then a *morphism* is a pair (φ, ψ) where $\varphi: X \rightarrow Y$ is a function and $\psi: S \rightarrow T$ is a homomorphism such that $\varphi(xs) = \varphi(x)\psi(s)$ for all $x \in X$ and $s \in S$, where equality means that either both sides are undefined or both are defined and equal. If S and T are A -generated, we shall call (φ, ψ) a morphism of A -partial transformation semigroups if ψ is a homomorphism of A -semigroups.

The following lemma is standard [13].

Lemma 2.18. *Let (X, S) and (Y, T) be partial transformation semigroups and $\varphi: X \rightarrow Y$ a surjective function so that, for all $s \in S$, there exists $\hat{s} \in T$ such that $\varphi(xs) = \varphi(x)\hat{s}$ for all $x \in X$ and $s \in S$ (interpreting equality as above). Then there is a unique homomorphism $\psi: S \rightarrow T$ so that (φ, ψ) is a morphism.*

Proof. First we show that \hat{s} is unique. Suppose $t, t' \in T$ satisfy $\varphi(xs) = \varphi(x)t = \varphi(x)t'$ for all $x \in X, s \in S$. Let $y \in Y$ with yt defined and choose $x \in \varphi^{-1}(y)$. Then $\varphi(x)t$ is defined and hence $\varphi(xs)$ is defined and so $\varphi(x)t'$ is defined. Moreover, $yt = \varphi(xs) = yt'$. Similarly, yt' defined implies yt is defined and $yt' = yt$. Thus by faithfulness $t = t'$. Hence the element \hat{s} in the hypothesis is unique. Define $\psi: S \rightarrow T$ by $\psi(s) = \hat{s}$. Notice for $s, s' \in S$ that xss' is defined if and only if $\varphi(xs)\psi(s')$ is defined, if and only if $\varphi(x)\psi(s)\psi(s')$ is defined and that $\varphi(xss') = \varphi(x)\psi(s)\psi(s')$. The uniqueness then implies $\psi(ss') = \psi(s)\psi(s')$. Finally the uniqueness of ψ follows from the uniqueness of \hat{s} . \square

We now verify that, for partial deterministic automata, a morphism is a directed covering if and only if it is surjective on vertices and induces a morphism of partial transformation semigroups.

Proposition 2.19. *Let \mathcal{A} and \mathcal{A}' be partial deterministic A -automata. Then the following are equivalent:*

- (1) *A directed covering of A -automata $\varphi: \mathcal{A} \rightarrow \mathcal{A}'$;*
- (2) *A surjective morphism $(\varphi, \psi): (V(\mathcal{A}), S(\mathcal{A})) \rightarrow (V(\mathcal{A}'), S(\mathcal{A}'))$ where ψ is a morphism of A -semigroups.*

Proof. Suppose first that φ is a directed covering. We claim that for all vertices q of \mathcal{A} and $w \in A^+$, one has $\varphi(qw) = \varphi(q)w$ (with the usual meaning). Indeed, the image of a path labeled w must be labeled w . So qw defined means $\varphi(q)w$ is defined and $\varphi(qw) = \varphi(q)w$. Conversely, if $\varphi(q)w$ is defined, then Proposition 2.13 implies that there is a lift of w starting at q , which must also have label w . So qw is defined and $\varphi(qw) = \varphi(q)w$. Hence we can define $[\widehat{w}]_{S(\mathcal{A})} = [w]_{S(\mathcal{A}'')}$ in Lemma 2.18 to obtain ψ .

Conversely, suppose (φ, ψ) is well defined. Define $\varphi: \mathcal{A} \rightarrow \mathcal{A}'$ to agree with φ on vertices. If $e: p \rightarrow q$ is an edge of \mathcal{A} with label a , then $pa = q$ and so $\varphi(q) = \varphi(pa) = \varphi(p)a$. Thus there is a (unique by determinism) edge labeled by a from $\varphi(p)$ to $\varphi(q)$, which we define to be $\varphi(e)$. Clearly φ is a morphism of A -automata. Let us check that it is a directed covering. Let q be a vertex of \mathcal{A} and let $e \in \iota^{-1}(\varphi(q))$ be an edge labeled by a . Then $\varphi(q)a$ is defined, so qa must be defined and $\varphi(qa) = \varphi(q)a$. Since the automaton is deterministic, there is a unique edge labeled by a emanating from q and it must map to e under φ (again by determinism). \square

2.4. Rooted graphs. In this article, we shall mostly be interested in rooted graphs.

Definition 2.20 (Rooted graph). A *rooted graph* is a pair (Γ, v) where Γ is a graph and v is a vertex of Γ such that every vertex of Γ can be reached from v by a directed path (i.e., the strong component of v is the unique maximum component in the ordering on strong components).

There is an analogous definition for automata.

Definition 2.21 (Pointed automaton). By a *pointed* (or *initial*) automaton, we mean an A -automaton $\mathcal{A} = (\Gamma, \ell)$ with a distinguished vertex I so that (Γ, I) is a rooted graph, that is, $I \cdot M(\mathcal{A}) = V(\Gamma)$. We denote the pointed automaton by (\mathcal{A}, I) . More generally, we say that a subset \mathfrak{l} is a *generating* or *initial* set for \mathcal{A} if $\mathfrak{l} \cdot M(\mathcal{A}) = V(\Gamma)$, that is, the downset generated by \mathfrak{l} is $V(\Gamma)$. In this case, we indicate the generating set by writing $(\mathcal{A}, \mathfrak{l})$.

Often it is important to endow an automaton with initial and terminal states in order to accept a language.

Definition 2.22 (Acceptor). A *non-deterministic acceptor* over an alphabet A is a non-deterministic A -automaton \mathcal{A} equipped with a distinguished

initial state I and a set T of *terminal states*. The language of the acceptor consists of all words in A^+ labeling a path from I to an element of T . If we use the word acceptor unmodified, then the underlying automaton is assumed partial deterministic. By a deterministic acceptor, we mean one in which the underlying automaton is deterministic. The languages accepted by finite acceptors are the so-called *regular* or *rational* languages.

An acceptor (\mathcal{A}, I, T) is said to be *trim* if every vertex is contained in a directed path from the initial state to some terminal state. In particular, trim acceptors are rooted at I .

For example, the Schützenberger graph $\text{Sch}^S(s)$ can be turned into an acceptor by specifying the vertex labeled s as both the initial state and its only terminal state. Whenever we refer to $\text{Sch}^S(s)$ as an acceptor without specifying initial and terminal states, this is what we intend.

Definition 2.23 (Reading words). Let \mathcal{A} be an A -automaton. Examples include the Cayley graph $\text{Cay}(S, A)$ and $\text{Sch}^S(s)$. A word $w \in A^+$ is said to be *readable on \mathcal{A}* if there exists a directed path in \mathcal{A} where the concatenation of labels is the word w . Similarly, if v is a vertex \mathcal{A} then being *readable starting at v* or *readable ending at v* has the obvious meaning.

Using this language, we can restate the advantage of the Cayley graph $\text{Cay}(S, A)$ as follows. Every word $w \in A^+$ is readable starting at I , and because $\text{Cay}(S, A)$ is deterministic, w is readable starting at I in exactly one way.

An important fact is that the Cayley graph of the transition semigroup of a complete deterministic pointed automaton is always a directed cover of the automaton.

Proposition 2.24. *Let $\mathcal{A} = (\Gamma, v)$ be a pointed deterministic A -automaton. Then there is a directed covering $\rho: \text{Cay}(S(\mathcal{A}), A) \rightarrow \mathcal{A}$ of A -automata given by $s \mapsto vs$ on vertices.*

Proof. The map ρ is clearly surjective and satisfies $\rho(ts) = vts = \rho(t)s$. Proposition 2.19 yields the desired result. \square

A rooted graph (Γ, v) is called a *directed rooted tree* if Γ is a tree. Directed rooted trees are characterized by having a unique directed path from the root to any vertex.

Proposition 2.25. *A rooted graph (Γ, v) is a directed rooted tree if and only if, for each vertex $w \in V(\Gamma)$, there is a unique directed path from v to w . This path is necessarily simple.*

Proof. Suppose first that Γ is a rooted directed tree and that there are two directed paths $p, q: v \rightarrow w$. Then by considering the longest common initial and terminal segments of p, q we can write $p = urs$ and $q = uts$ so that r, t begin and end with different edges. Then rt^{-1} is a reduced undirected circuit in Γ and so Γ is not a tree. Suppose conversely, that Γ is not a tree.

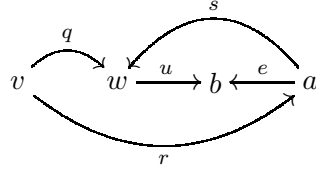


FIGURE 2.2.

Consider a reduced undirected circuit p in Γ . Let $w = \iota(p)$ and let $q: v \rightarrow w$ be a directed path. Replacing p by a cyclic conjugate of its reverse circuit if necessary, we may assume the first edge of p is traversed in the positive direction. If p is a directed circuit, then q, qp are two directed paths from v to w and we are done. Otherwise, we may factor $p = ue^{-1}s$ where u is the longest directed initial segment of p and e^{-1} indicates that e is traversed backwards. Let $e: a \rightarrow b$. Then $b = \tau(u)$. Let $r: v \rightarrow a$ be a directed path. See Figure 2.2. Then re and qu are two directed paths from v to b . We claim they are distinct. Indeed, if the last edge of u were e , then $p = ue^{-1}s$ would not be reduced. This completes the proof.

If (Γ, v) is a directed rooted tree and $p: v \rightarrow w$ is the unique directed path, then obviously p is simple. For if we could write $p = qur$ where u is a non-empty path with $\iota(u) = \tau(u)$, then u is a reduced circuit in Γ , contradicting that Γ is a tree. \square

The unique directed path from v to w is denoted $[v, w]$ and called the *geodesic* from v to w . More generally, if T is a rooted directed tree and $w \leq u$, then there is a unique (directed) simple path from u to w , which we denote $[u, w]$ and call the geodesic from u to w .

Let (Γ, v) be a rooted graph. By a *directed spanning tree* T for (Γ, v) we mean a directed rooted subtree (T, v) containing all the vertices of Γ . The next proposition shows that every rooted graph admits a directed spanning tree. One should think of a directed spanning tree as a collection of normal forms for each vertex of Γ .

Proposition 2.26. *Let (Γ, v) be a rooted graph. Then (Γ, v) admits a directed spanning tree.*

Proof. Let \mathcal{C} be the collection of all directed rooted subtrees (T, v) of (Γ, v) ordered by inclusion. It is non-empty since it contains $(\{v\}, v)$. Next observe that if $\{(T_\alpha, v) \mid \alpha \in A\}$ is a chain in \mathcal{C} , then $(\bigcup_{\alpha \in A} T_\alpha, v) \in \mathcal{C}$ since any reduced undirected circuit in this union must belong to some T_α . Thus by Zorn's Lemma, \mathcal{C} contains a maximal element (T, v) . We claim that this is the desired directed spanning tree. For suppose $V(T) \subsetneq V(\Gamma)$. We claim that there is an edge $e \in E(\Gamma)$ with $\iota(e) \in V(T)$ and $\tau(e) \notin V(T)$. For if this is not the case, then $V(T)$ is a downset in the preorder \prec on $V(\Gamma)$. But $v \in V(T)$ and Γ is rooted at v . Thus $V(T) = V(\Gamma)$. So let e be such an edge. One easily verifies that the graph T' obtained by adjoining e to T ,

so $V(T') = V(T) \cup \{\tau(e)\}$ and $E(T') = E(T) \cup \{e\}$, is a tree. Indeed, if $w \in V(T)$, then the unique directed path in T' from v to w is the geodesic in T , whereas the unique directed path from v to $\tau(e)$ is pe where p is the geodesic $[v, \iota(e)]$ in T . Thus T' is a directed rooted tree by Proposition 2.25. Clearly (T', v) is a larger element of \mathcal{C} than T . This contradiction completes the proof. \square

Remark 2.27. The proof of Proposition 2.26 can easily be adapted to prove that if (T_0, v) is any directed rooted subtree of a rooted graph (Γ, v) , then there is a directed spanning tree (T, v) containing (T_0, v) .

It turns out that every rooted directed graph (Γ, v) has a unique directed cover that is a tree. Moreover, this tree is a directed cover of all directed covers of (Γ, v) and hence is called the *universal directed cover* of (Γ, v) .

Theorem 2.28. *Let (Γ, v) be a rooted graph. Then there is a rooted tree $(\tilde{\Gamma}, \tilde{v})$ and a directed covering $\pi: (\tilde{\Gamma}, \tilde{v}) \rightarrow (\Gamma, v)$. Moreover, given a directed covering $\varphi: (\Gamma', v') \rightarrow (\Gamma, v)$, there is a unique morphism (necessarily a directed covering) $\psi: (\tilde{\Gamma}, \tilde{v}) \rightarrow (\Gamma', v')$ so that*

$$\begin{array}{ccc} (\tilde{\Gamma}, \tilde{v}) & \xrightarrow{\psi} & (\Gamma', v') \\ & \searrow \pi & \swarrow \varphi \\ & (\Gamma, v) & \end{array}$$

commutes.

Proof. The proof is so similar to the classical undirected case that we just give the construction of $\tilde{\Gamma}$ and leave the remaining details to the reader. The vertices of $\tilde{\Gamma}$ are the directed path starting at v . One takes \tilde{v} to be the empty path at v . The edges of $\tilde{\Gamma}$ consist of pairs (p, e) where p is a path from v and $\tau(p) = \iota(e)$. The incidence functions are given by $\iota(p, e) = p$ and $\tau(p, e) = pe$. The directed covering π is given by $\pi(p) = \tau(p)$ on vertices and $\pi(p, e) = e$ on edges. \square

For example, the Cayley graph of A^* is the universal directed cover of the bouquet \mathcal{B}_A , as well as of any other deterministic A -automaton.

2.4.1. The fundamental monoid of a rooted graph. If (Γ, v) is a rooted graph and $v \in V(\Gamma)$, then one can define the *fundamental monoid* $\Gamma^*(v)$ of (Γ, v) to be the monoid of all loops at v with the concatenation product. For instance, the fundamental monoid of the bouquet \mathcal{B}_A at its unique vertex is the free monoid A^* . Notice that the fundamental monoid ignores all of Γ except the strong component of v and moreover, it depends on the root vertex even for strongly connected graphs. Thus one should really work with the free category Γ^* on the graph Γ [23]. Analogously to the case of fundamental groups, the fundamental monoid of a rooted graph is free.

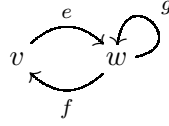


FIGURE 2.3. An infinitely generated fundamental monoid

Proposition 2.29. *Let (Γ, v) be a rooted graph. Then $\Gamma^*(v)$ is free on the set P_v of non-empty paths $p: v \rightarrow v$ that do not visit v except for at the beginning and the end.*

Proof. Indeed, if $q: v \rightarrow v$ is any non-empty loop, then it has a unique factorization $q = p_1 \cdots p_n$ with $p_1, \dots, p_n \in P$ by partitioning q according to each time it visits v . Thus $\Gamma^*(v)$ is free on P . \square

The next example shows that the generating set P_v need not be finite even when the graph Γ is finite and also exhibits the dependence on the generating set.

Example 2.30. Consider the strongly connected graph Γ in Figure 2.3. Then $\Gamma^*(v)$ is freely generated by the infinite set $\{eg^n f \mid n \geq 0\}$, whereas $(\Gamma, w)^*$ is freely generated by f, e, g .

The next result generalizes the situation for free monoids that can be found, for instance, in [3]. A submonoid N of a monoid M is called *right unitary* if $u, uv \in N$ implies $v \in N$ for $u, v \in M$.

Proposition 2.31. *Let $\varphi: (\tilde{\Gamma}, \tilde{v}) \rightarrow (\Gamma, v)$ be a directed covering of rooted graphs. Then the induced map $\varphi: \tilde{\Gamma}^*(\tilde{v}) \rightarrow \Gamma^*(v)$ is injective and the image is right unitary. Conversely, every right unitary submonoid of $\Gamma^*(v)$ is of this form.*

Proof. Since φ is a directed covering, any loop at v has at most one lift starting at \tilde{v} and so φ is an injective homomorphism. Suppose that $p, pq \in \varphi(\tilde{\Gamma}^*(\tilde{v}))$ with $q \in \Gamma^*(v)$. Choose lifts \tilde{p}, \tilde{q} of p and q respectively starting at \tilde{v} . Then \tilde{p} must be a loop at \tilde{v} by uniqueness of lifts. Also, since $\tilde{p}\tilde{q}$ lifts pq , it too must be a loop at \tilde{v} . It follows now that \tilde{q} is a loop at \tilde{v} and so $q = \varphi(\tilde{q}) \in \varphi(\tilde{\Gamma}^*(\tilde{v}))$, as required.

Now suppose that N is a right unitary submonoid of $\Gamma^*(v)$. Let X be the set of directed paths in Γ starting at v . Define an equivalence relation on X by $p \equiv q$ if $\tau(p) = \tau(q)$ and $pu \in N$ if and only if $qu \in N$ for all $u: \tau(p) \rightarrow v$. Let V be the quotient of X by \equiv and let E consist of all pairs $([p], e)$ with $e \in E(\Gamma)$ and $\tau(p) = \iota(e)$. We define $\tilde{\Gamma}$ to have vertex set V and edge set E where $([p], e)$ goes from $[p]$ to $[pe]$. Notice that $[p] = [q]$ implies $peu \in N$ if and only if $qeu \in N$ and so the incidence functions are well defined. The reader can easily verify that the maps $[p] \mapsto \tau(p)$ and $([p], e) \mapsto e$ yield a directed covering. Surjectivity on vertices requires that (Γ, v) is rooted. Let

$\tilde{v} = [1_v]$. One can verify directly that if q is a loop at v , then the lift of q to \tilde{v} ends at $[q]$. Now $[q] = \tilde{v}$, if and only if $q \equiv 1_v$. Let us show that this is equivalent to $q \in N$. If $q \equiv 1_v$, then since $1_v \in N$, it follows that $q = q1_v \in N$. Conversely, if $q \in N$, then the very definition of right unitary implies $qu \in N$ if and only if $u \in N$, if and only if $1_v u \in N$. Thus $q \equiv 1_v$, completing the proof that $\tilde{\Gamma}^*(\tilde{v})$ maps onto N . \square

There can be multiple directed coverings corresponding to a given right unitary submonoid of $\Gamma^*(v)$, but the construction given in Proposition 2.31 is the unique minimal one in the sense that all others cover it. The above proof also shows that directed immersions induce injective maps on fundamental monoids.

Notice that if $\varphi: (\tilde{\Gamma}, \tilde{v}) \rightarrow (\Gamma, v)$ is a directed covering of rooted graphs, then $\Gamma^*(v)$ has a monodromy action on $\varphi^{-1}(v)$ given on $w \in \varphi^{-1}(v)$ by $wp = w'$ where w' is the end point of the unique lift of p starting at w . This generalizes the action of the transition monoid of a deterministic automaton.

2.5. The unique simple path property. Perhaps the most important notion in geometric theory is that of a rooted graph with the unique simple path property. We give several equivalent properties that will define this notion. Recall that path unmodified means directed path.

Proposition 2.32. *Let (Γ, v) be a rooted graph. Then the following are equivalent:*

- (1) *For each vertex w , there is a unique simple path from v to w ;*
- (2) *(Γ, v) admits a unique directed spanning tree;*
- (3) *(Γ, v) admits a directed spanning tree (T, v) such that, for each edge $e \in E(\Gamma) \setminus E(T)$ one has $[v, \tau(e)]$ is an initial segment of $[v, \iota(e)]$ (i.e., $\tau(e)$ is visited by the geodesic $[v, \iota(e)]$).*

Proof. To see that (1) implies (2), suppose that (T, v) and (T', v) are distinct spanning trees. Then there is an edge e that belongs to, say, T and not T' . Set $w = \tau(e)$. Let $p: v \rightarrow \iota(e)$ and $q: v \rightarrow w$ be the geodesics in T and T' respectively. Then $pe: v \rightarrow w$ is a directed path in T and hence simple by Proposition 2.25. Since $e \notin T'$, clearly $q \neq pe$. This contradicts (1).

For (2) implies (3), let T be the unique directed spanning tree for Γ and suppose that $e \in E(\Gamma) \setminus E(T)$. Suppose that $\tau(e) \notin [v, \iota(e)]$. The path $[v, \iota(e)]e$ is then simple and so its support T' is a directed tree rooted from v . By Remark 2.27 we may complete T' to a directed spanning tree that evidently is different from T . This contradicts the uniqueness of T .

Finally, we prove (3) implies (1). Let T be the spanning tree provided by (3). Let $p = [v, w]$ be the geodesic in T . Then p is a simple path from v to w . Suppose that $q: v \rightarrow w$ is another simple path; it cannot be contained in T so let e be the first edge used by q that does not belong to T . Then we can factor $q = ret$ where $e \in E(\Gamma) \setminus T$ and $r = [v, \iota(e)]$. But then by assumption $\tau(e) \in [v, \iota(e)]$. This contradicts that q is simple. We conclude that p is the unique simple path from v to w . \square

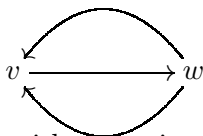


FIGURE 2.4. A graph with the unique simple path property at v but not w .

The third condition says that Γ is a “rooted tree falling back on itself.” Here is the key definition in geometric semigroup theory.

Definition 2.33 (Unique simple path property). A rooted graph (Γ, v) is said to have the *unique simple path property* if it satisfies the equivalent conditions of Proposition 2.32. If (Γ, v) has the unique simple path property, then we denote the unique simple path from v to w by $[v, w]$ and call it the *geodesic*, just as we did for directed rooted trees.

Remark 2.34. Having the unique simple path property depends not just on the graph Γ , but also on the choice of the root. For example the graph in Figure 2.4 has the unique simple path property from v but not from w .

Remark 2.35 (Non-planar). Also, having the unique simple path property does not force the graph to be planar. For example, the complete bipartite graph $K_{3,3}$ is non-planar but as shown in Figure 2.5 it can be oriented to have the unique simple path property from one of its vertices (p_1 in this case).

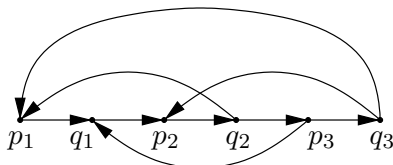


FIGURE 2.5. A non-planar graph with the unique simple path property.

Remark 2.36. Notice that if (Γ, v) has the unique simple path property, then each transition edge of Γ belongs to the directed spanning tree by the third item of Proposition 2.32.

It is easy to see that if (Γ, v) has the unique simple path property and X is a convex set of vertices of Γ containing v , then $(\Delta(X), v)$ has the unique simple path property where $\Delta(X)$ is the subgraph induced by X .

The frame of a rooted graph with the unique simple path property is a directed rooted tree. Moreover, each strong component has a unique transition edge entering it, called its *entry edge*. The endpoint of the entry edge is termed the *entrance* of the strong component. Most importantly, if C is a strong component with entrance w , then (C, w) has the unique simple path property as does (C^\downarrow, w) where C^\downarrow is the subgraph of Γ induced by the

downset in \prec generated by the vertices of C (or equivalently by w). This is summarized in the following proposition.

Proposition 2.37. *Let (Γ, v) have the unique simple path property. Then $(\text{Fr}(\Gamma), C_v)$ is a directed rooted tree where C_w denotes the strong component of a vertex w . Each strong component has a unique transition edge entering it, called its entry edge. If C is a strong component with entry edge e and entrance $w = \tau(e)$, then (C^\downarrow, w) has the unique simple path property and hence (C, w) has the unique simple path property (being convex in (C^\downarrow, w)).*

Proof. We show that there is a unique directed path from C_v to C in $\text{Fr}(\Gamma)$ for any strong component C . Proposition 2.25 then implies that $(\text{Fr}(\Gamma), C_v)$ is a directed rooted tree. It is immediate from the definition of $\text{Fr}(\Gamma)$ that there is a path $C_x \rightarrow C_y$ if and only if there is a path from x to y . Thus there is some path from C_v to C . Suppose that $p = e_1 \cdots e_n$ and $q = f_1 \cdots f_m$ are two distinct such paths; they are necessarily simple since $\text{Fr}(\Gamma)$ is acyclic. The e_i and f_j are transition edges of Γ . Since e_n and f_m end in C , we can find a simple path in $r: \tau(e_n) \rightarrow \tau(f_m)$ in C . Also, for each $1 \leq i \leq n-1$ we can find a simple path p_i contained in the strong component of $\tau(e_i)$ from $\tau(e_i)$ to $\iota(e_{i+1})$. Let $t: v \rightarrow \iota(e_1)$ be a simple path contained in C_v . Then $p' = te_1p_2 \cdots e_{n-1}p_{n-1}e_n r$ is a simple path from v to $\tau(f_m)$ whose transition edges are precisely e_1, \dots, e_n . Similarly, we can construct a simple path $q' = t'f_1q_1 \cdots f_{m-1}q_{m-1}f_m$ from v to $\tau(f_m)$ whose transition edges are precisely f_1, \dots, f_m . Since $p' \neq q'$, this contradicts the unique simple path property. We conclude $(\text{Fr}(\Gamma), C_v)$ is a directed rooted tree.

Suppose that e, e' are distinct transition edges entering C . Say $e: C_1 \rightarrow C$ and $e': C_2 \rightarrow C$ in $\text{Fr}(\Gamma)$. Let p, q be the geodesics in $\text{Fr}(\Gamma)$ from C_v to C_1, C_2 respectively. Then $pe, qe': C_v \rightarrow C$ are distinct directed paths, a contradiction to $\text{Fr}(\Gamma)$ being a rooted tree.

Finally, let C be a strong component with entry edge e and entrance $w = \tau(e)$. Let z be any vertex with $z \prec w$. Suppose $p, q: w \rightarrow z$ are simple paths. Since e is a transition edge, p and q cannot use the edge e . It follows that $[v, \iota(e)]ep$ and $[v, \iota(e)]eq$ are simple paths from v to z . We conclude $p = q$, as required. \square

The following corollary is immediate.

Corollary 2.38. *If (Γ, v) has the unique simple path property and is quasi-linear, then Γ is linear.*

Let us fix for the moment a rooted graph (Γ, I) with the unique simple path property and denote by T its unique directed spanning tree. Let us use \leq_T to denote the accessibility order in T . Notice that $u \leq_T v$ implies $u \leq v$, but not conversely. For example, \leq is trivial on strongly connected components of Γ , but \leq_T is still non-trivial. If v is a vertex of Γ , denote by v^\downarrow the set of vertices $u \in V(\Gamma)$ with $u \leq_T v$. Abusively, we shall also denote the induced subgraph with vertex set v^\downarrow by the same notation. Let us denote by C_{v^\downarrow} the strong component of v in v^\downarrow .

Proposition 2.39. *Let (Γ, I) be a graph with the unique simple path property and suppose $v \in V(\Gamma)$. The rooted graph (v^\downarrow, v) has the unique simple path property. Consequently, (C_{v^\downarrow}, v) has the unique simple path property.*

Proof. If $u \in v^\downarrow$, then there is a simple path in T from v to u by definition, which we denote $[v, u]$. Then $[I, u] = [I, v][v, u]$. Suppose that $p: v \rightarrow u$ is a simple path in v^\downarrow and consider $[I, v]p$. If this path is simple, $[v, u] = p$ and we are done. If not, then there is a vertex r visited twice by $[I, v]p$. Since $[I, v]$ and p are simple, it follows that $r \neq v$ and is visited by both $[I, v]$ and p . But this is impossible since this means that $v <_T r$ and $r \leq_T v$. \square

Sometimes one wants to extract a quasilinear subgraph from a graph with the unique simple path property. If w is a vertex of Γ , then by abuse of notation, w^\uparrow will also denote the full subgraph of Γ with vertex set $w^\uparrow = \{q \in V(\Gamma) \mid q \succ w\}$.

Proposition 2.40. *Suppose that (Γ, v) has the unique simple path property and $w \in V(\Gamma)$. Then (w^\uparrow, v) has the unique simple path property and is quasilinear. In particular, w^\uparrow is linear.*

Proof. Since w^\uparrow is convex, it has the unique simple path property from v . To see that it is quasilinear, let C, C' be strong components of w^\uparrow . Then since $\text{Fr}(\Gamma)$ is a tree and C, C' are between the strong component of v and the strong component of w , it follows that C and C' are comparable in the accessibility order. Thus w^\uparrow is quasilinear and hence is linear by Corollary 2.38. \square

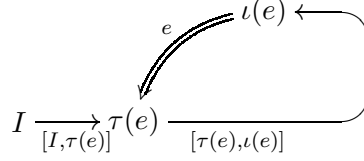
We now define the “bold” arrows of a rooted graph with the unique simple path property.

Definition 2.41 (Bold arrows). Let (Γ, I) have the unique simple path property with directed spanning tree T . Then $\mathbf{E}(\Gamma) = E(\Gamma) \setminus E(T)$ is called the set of *bold arrows* of Γ . (Notice the bold font is used to denote the set of bold arrows.)

The set $\mathbf{E}(\Gamma)$ of bold arrows is in bijection with a generating set for $\pi_1(\Gamma, I)$ (or equivalent $H_1(\Gamma)$) and so the number of bold arrows is just the first Betti number of Γ . Notice that if Γ' is any subgraph of Γ containing the spanning tree T , then (Γ', I) still has the unique simple path property. Thus removing bold arrows does not cause one to lose the unique simple path property. Induction on the number of bold arrows is a key idea in geometric semigroup theory.

An important notion in geometric semigroup theory is that of a sloop.

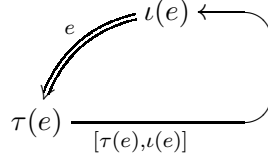
Definition 2.42 (Sloop). If $e \in \mathbf{E}(\Gamma)$ is a bold arrow, then the sloop (think “simple loop”) is the path $\mathbf{slp}(e) = [I, \iota(e)]e$. In pictures, we have



where the bold arrow is drawn doubled.

If e is a bold arrow, then the corresponding generator of $\pi_1(\Gamma, I)$ is $\mathbf{slp}(e)[I, \tau(e)]^{-1}$. Next we define the loop of a sloop.

Definition 2.43 (Loop of a sloop). If $e \in \mathbf{E}(\Gamma)$ is a bold arrow, define the loop $\mathbf{lp}(e)$ of the sloop $\mathbf{slp}(e)$ to be the simple circuit $[\tau(e), \iota(e)]e$. So the loop $\mathbf{lp}(e)$ of the sloop $\mathbf{slp}(e)$ in the picture from Definition 2.42 is



Note that $\mathbf{slp}(e) = [I, \tau(e)]\mathbf{lp}(e)$.

In a graph with the unique simple path property, each non-trivial strong component is a union of loops of sloops.

Proposition 2.44. *Let (Γ, v) have the unique simple path property and let C be a strong component of Γ containing at least one edge. Then $C = \bigcup_{e \in \mathbf{E}(\Gamma) \cap C} \mathbf{lp}(e)$.*

Proof. Obviously, $\mathbf{lp}(e) \subseteq C$ for any $e \in \mathbf{E}(\Gamma) \cap C$ since $\mathbf{lp}(e)$ is strongly connected. For the converse, it suffices to show that each edge of C belongs to $\mathbf{lp}(e)$ for some $e \in \mathbf{E}(\Gamma)$ since every vertex of C must be on some edge. Let u be the entrance of C and suppose $f \in E(C)$. If $f \in \mathbf{E}(\Gamma)$, there is nothing to prove. So suppose that $f \in E(T)$; say $f: v \rightarrow w$. Choose a simple path $p: w \rightarrow u$. Then $q = [u, v]fp$ is not simple and so we can factor $q = [u, \iota(e)]eq'$ where e is the first bold arrow used by q . Evidently, $[u, v]f$ is an initial segment of $[u, \iota(e)]$ and so $\iota(e) \leq_T w$. On the other hand, since p is simple, we must have $v \leq_T \tau(e)$ and so $\iota(e) \leq_T w \leq_T v \leq_T \tau(e)$. See Figure 2.6. Thus f is an edge of $\mathbf{lp}(e) = [\tau(e), \iota(e)]e$. \square

Next we wish to define the notion of a geometric rank function. Later on, when we deal with automata, we shall impose some algebraic restrictions on our rank functions, but here we define things in complete generality. From now on we assume that all graphs have finite out-degree.

Definition 2.45 (Geometric rank function). Let (Γ, I) have the unique simple path property. Then a *geometric rank function* is a mapping $r: \mathbf{E}(\Gamma) \rightarrow$

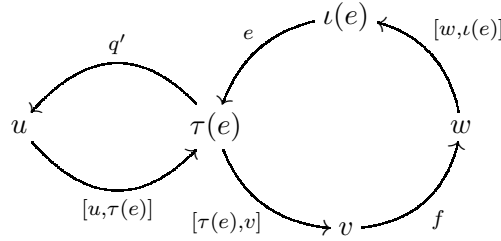


FIGURE 2.6.

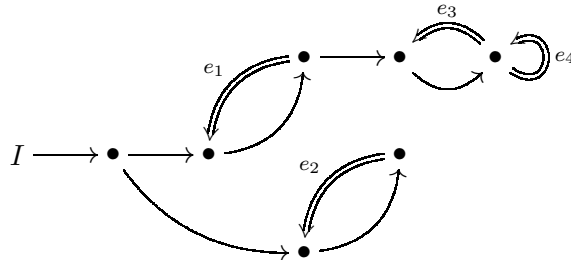
\mathbb{N} so that, for each vertex v , r maps $\tau^{-1}(v) \cap \mathbf{E}(\Gamma)$ bijectively to the interval $[0, |\tau^{-1}(v) \cap \mathbf{E}(\Gamma)| - 1]$. We then totally order $\tau^{-1}(v) \cap \mathbf{E}(\Gamma)$ by putting $e \leq_r f$ if and only if $r(e) \leq r(f)$.

Said differently, a geometric rank function is a way of totally ordering the bold arrows at each vertex. Given a geometric rank function, we can now define a partial order on the bold arrows as follows. Recall that C_v denotes the strong component of a vertex v .

Definition 2.46 (Order on bold arrows). Fix a geometric rank function r . Then define a partial order on $\mathbf{E}(\Gamma)$ by putting $e \ll f$ for $e \neq f \in \mathbf{E}(\Gamma)$ if:

- (1) $C_{\tau(f)} < C_{\tau(e)}$; or
- (2) $C_{\tau(e)} = C_{\tau(f)}$ and $\tau(e) <_T \tau(f)$; or
- (3) $\tau(e) = \tau(f)$ and $r(e) < r(f)$.

Example 2.47. Consider the graph



The bold arrows are e_1, e_2, e_3, e_4 . One has $e_1 \ll e_4 \ll e_3$, whereas e_2 is incomparable to the other arrows.

2.6. Cutting sloops. A major idea in geometric semigroup is to work by induction on the number of bold arrows. To do this, we need to cut sloops. First we talk about cutting arbitrary rooted graphs.

Definition 2.48 (Cutting a graph). Let (Δ, v) be a rooted graph. Define $\text{cut}(\Delta, v)$ to be the graph with vertex set $V(\Delta) \cup (\{v\} \times \tau^{-1}(v))$ and edge set $E(\Delta)$. The initial vertex function ι is as before. The new terminal vertex

function τ' is given by

$$\tau'(e) = \begin{cases} \tau(e) & \tau(e) \neq v \\ (v, e) & \tau(e) = v. \end{cases}$$

Geometrically, $\text{cut}(\Delta, v)$ is obtained from Δ by taking each edge ending at v and moving its end away from v . Notice the map $\text{cut}(\Delta, v) \rightarrow \Delta$ which is the inclusion on $V(\Delta) \cup E(\Delta)$ and which sends $\{v\} \times \tau^{-1}(v)$ to v is a directed graph immersion.

Example 2.49. Suppose (Δ, v) is the rooted graph in Figure 2.7. Then

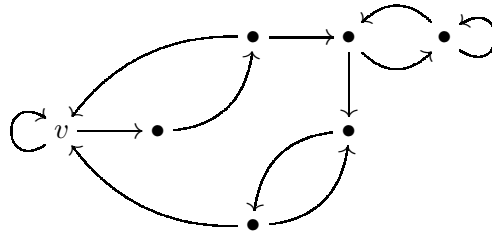


FIGURE 2.7. The graph (Δ, v) be

$\text{cut}(\Delta, v)$ is the graph in Figure 2.8.

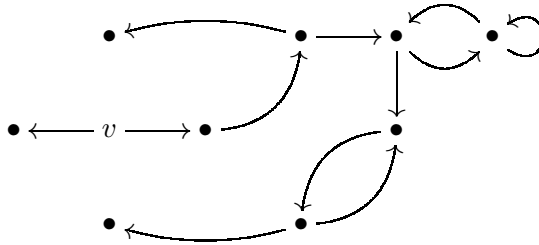


FIGURE 2.8. The graph $\text{cut}(\Delta, v)$

The next proposition establishes some basic properties of cut graphs.

Proposition 2.50. *The graph $\text{cut}(\Delta, v)$ is rooted at v . Moreover, the rooted graph $(\text{cut}(\Delta, v), v)$ has the unique simple path property if and only if (Δ, v) does. If (Δ, v) has the unique simple path property, then the bold arrows of $(\text{cut}(\Delta, v), v)$ are the bold arrows of Δ that do not end at v .*

Proof. We shall repeatedly use the following fundamental observation: if $w \in V(\Delta)$ with $w \neq v$, then the simple paths between v and w in Δ and $\text{cut}(\Delta, v)$ are one in the same. This is because on the one hand, a simple path from v never uses an edge of $\tau^{-1}(v)$; on the other hand, a path in $\text{cut}(\Delta, v)$ using an edge $e \in \tau^{-1}(v)$ can only use that edge as its last edge and therefore does not end in $V(\Delta)$.

It follows that if $w \in V(\Delta)$, then there is a simple path from v to w in $\text{cut}(\Delta, v)$. On the other hand, if $e \in \tau^{-1}(v)$ and p is a simple path from v to $\iota(e)$ in Δ , then the fundamental observation implies pe is a simple path in $\text{cut}(\Delta, v)$ from v to (v, e) .

From the fundamental observation, it is immediate that if $(\text{cut}(\Delta, v), v)$ has the unique simple path property, then so does (Δ, v) . Suppose now that (Δ, v) has the unique simple path property. If $w \in V(\Delta)$, then the fundamental observation shows that there is a unique simple path from v to w in $\text{cut}(\Delta, v)$. Now any simple path from v to (v, e) for $e \in \tau^{-1}(v)$ must end in the edge e since it is the only edge ending at (v, e) . Thus the unique simple path in $\text{cut}(\Delta, v)$ from v to (v, e) is pe where p is the unique simple path in Δ from v to $\iota(e)$. It follows from this discussion that the bold arrows of $(\text{cut}(\Delta, v), v)$ are those bold arrows of Δ that do not end at v . \square

Next we need to define a notion of a closed subgraph with respect to a base point.

Definition 2.51 (Closed subgraph with respect to a point). Let (Γ, I) have the unique simple path property and let T be its unique directed spanning tree. Fix $u \in V(\Gamma)$. A subgraph Δ of Γ is said to be *closed with respect to u* , or the pair (Δ, u) is said to be *closed* if:

- (1) $u \in \Delta \subseteq u^\downarrow$;
- (2) For all $v \in V(\Delta)$, one has $[u, v] \subseteq \Delta$;
- (3) If $e \in \mathbf{E}(\Gamma)$ and $u \neq \tau(e) \in \Delta$, then $e \in \Delta$.

Let us show that the set of all closed subgraphs with respect to u is a complete lattice. Throughout (Γ, I) is a fixed graph with the unique simple path property.

Proposition 2.52. *The collection L_u of all closed subgraphs of Γ with respect to u has maximum element u^\downarrow and is closed under non-empty intersections. Thus it is a complete lattice. The bottom of L_u is $\{u\}$ and its join is determined by the meet.*

Proof. First note that the graph u^\downarrow is closed with respect to u . The first two axioms are obvious. The last one follows since in a graph with the unique simple path property $\iota(e) \leq_T \tau(e)$ for any bold arrow e . Clearly it is the largest element of L_u by (1). It is trivial that the set of elements satisfying (1)–(3) is closed under non-empty intersections. \square

As a consequence we can define the closure of a subgraph.

Definition 2.53 (Closure of a pointed subgraph). If $\Delta \subseteq u^\downarrow$, define $\overline{(\Delta, u)} = (\overline{\Delta}, u)$ where $\overline{\Delta}$ is the meet of all elements of L_u containing Δ . This definition makes sense since the top of L_u is u^\downarrow .

Let us establish some basic properties of $\overline{(\Delta, u)}$.

Proposition 2.54. *Let (Δ, u) be closed. Then (Δ, u) has the unique simple path property.*

Proof. Proposition 2.39 shows that (u^\downarrow, u) has the unique simple path property. Since $u \in \Delta \subseteq u^\downarrow$, it suffices to show that (Δ, u) is rooted at u . But this is immediate from Definition 2.51(2). \square

Consequently, $\overline{(\Delta, u)}$ has the unique simple path property for any $\Delta \subseteq u^\downarrow$.

Proposition 2.55. *Suppose that (Δ, u) is closed and $e \in \mathbf{E}(\Gamma)$ with $\tau(e) \in \Delta$ and $u \neq \tau(e)$. Then $\text{lp}(e) \in \Delta$.*

Proof. It follows from the third axiom in the definition of closed that $e \in \Delta$. Therefore, $[u, \iota(e)] \subseteq \Delta$ by the second axiom and hence $\text{lp}(e) = [\tau(e), \iota(e)]e \subseteq \Delta$ as $\iota(e) \leq_T \tau(e) \leq_T u$. \square

Next we show that taking the closure preserves strong connectivity.

Proposition 2.56. *Suppose that $\Delta \subseteq u^\downarrow$ is strongly connected and $u \in \Delta$. Then $\overline{(\Delta, u)}$ is strongly connected.*

Proof. Let C be the strong component of u in $\overline{(\Delta, u)}$. Then $\Delta \subseteq C$ since $u \in \Delta$ and Δ is strongly connected. Thus it suffices to show that (C, u) is closed. The first axiom is clear. If $v \in C$, choose a path $p: v \rightarrow u$ in C . Then $[u, v] \subseteq \overline{(\Delta, u)}$ and so the existence of the circuit $[u, v]p$ shows that $[u, v] \subseteq C$. Suppose that $e \in \mathbf{E}(\Gamma)$ with $\tau(e) \in C$ and $u \neq \tau(e)$. Then $e \in \overline{(\Delta, u)}$. Therefore, $\text{lp}(e) \subseteq \overline{(\Delta, u)}$ by Proposition 2.55 and hence $\text{lp}(e) \subseteq C$. We conclude $e \in C$. This completes the proof that (C, u) is closed. It now follows that $\overline{\Delta} = C$. \square

In order to show that certain automata are trim, we need to show that closure preserves the property that some vertex is reachable from all vertices.

Proposition 2.57. *Suppose that $(\Delta, u) \subseteq u^\downarrow$ and w is a vertex of Δ which is accessible from every vertex of Δ by a path in Δ . Then there is a directed path in $\overline{\Delta}$ from every vertex of $\overline{\Delta}$ to w .*

Proof. Let Λ be the full subgraph of $\overline{\Delta}$ containing all vertices v of $\overline{\Delta}$ such that there is a path from v to w in $\overline{\Delta}$. We claim that (Λ, u) is a closed subgraph containing Δ . It will then follow that $\Lambda = \overline{\Delta}$, as required. Indeed, by hypothesis $u \in \Delta \subseteq \Lambda \subseteq \overline{\Delta} \subseteq u^\downarrow$. In particular, the first condition in the definition of a closed subgraph is satisfied. If $v \in \Lambda$, then there is a directed path from v to w in $\overline{\Delta}$. Since $[u, v] \subseteq \overline{\Delta}$, it follows immediately that $[u, v] \subseteq \Lambda$. Finally if $e \in \mathbf{E}(\Gamma)$ and $u \neq \tau(e) \in \Lambda$, then $e \in \overline{\Delta}$ since $\overline{\Delta}$ is closed. Hence if p is a path in $\overline{\Delta}$ from $\tau(e)$ to w , then ep is a directed path in $\overline{\Delta}$ from $\iota(e)$ to w . Thus $\iota(e) \in \Lambda$ and so $e \in \Lambda$. This concludes the proof that (Λ, u) is closed. \square

We can now define the cut sloop of a sloop.

Definition 2.58 (Cut sloop). If $e \in \mathbf{E}(\Gamma)$, then we define the *cut sloop* of e , denoted $\text{cut}(e)$, to be $\text{cut}(\text{lp}(e), \tau(e))$.

This definition makes sense because $\text{lp}(e)$ is a subgraph of $\tau(e)^\downarrow$. Putting together Propositions 2.50 and 2.54, we conclude that $(\text{cut}(e), \tau(e))$ has the unique simple path property.

Proposition 2.59. *Let $e \in \mathbf{E}(\Gamma)$. Then $(\text{cut}(e), \tau(e))$ has the unique simple path property.*

It turns out that $\text{cut}(e)$ can be described more explicitly.

Proposition 2.60. *Let $e \in \mathbf{E}(\Gamma)$ and set $u = \tau(e)$. Let K_u be the strong component of u in the graph Γ' obtained from u^\downarrow by removing the bold arrows ending at u other than e (so $K_u \subseteq C_{u^\downarrow}$). Then:*

- (1) $\overline{(\text{lp}(e), u)} \subseteq K_u$;
- (2) e is the unique edge of $\overline{(\text{lp}(e), u)}$ ending at u ;
- (3) If $f \in \mathbf{E}(\Gamma)$ and f belongs to $\overline{(\text{lp}(e), u)}$, then $f \ll e$;
- (4) If $f \in \mathbf{E}(\Gamma)$ and $\tau(f) \in \overline{(\text{lp}(e), u)} \setminus \{u\}$, then $\text{lp}(f) \subseteq \overline{(\text{lp}(e), u)}$;
- (5) Suppose that $f \in \mathbf{E}(\Gamma) \cap K_u$. Then $f \in \overline{(\text{lp}(e), u)}$;
- (6) $K_u = \overline{(\text{lp}(e), u)}$.

Proof. To prove (1), clearly $\text{lp}(e) \subseteq K_u$. It therefore suffices to show that (K_u, u) is closed. The first two axioms are clear. Suppose that $f \in \mathbf{E}(\Gamma)$ and $u \neq \tau(f) \in K_u$. Then $\iota(f) \leq_T \tau(f) <_T u$ where the first inequality holds since (Γ, v) has the unique simple path property, whereas the second follows by definition of u^\downarrow . Thus $f \in \Gamma'$. Also $f \in K_u$ since bold arrows are never transition edges in Γ' and $\tau(f) \in K_u$. This completes the proof of (1).

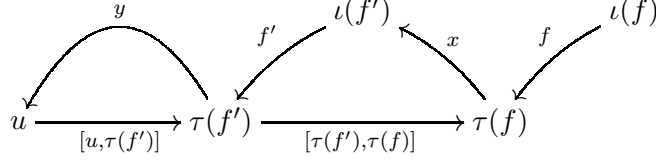
We immediately deduce (2) from (1) since Γ' contains only the edge e coming into u as any edge of u^\downarrow going into u must necessarily be a bold arrow.

By (1) and (2), necessary conditions for a bold arrow f to belong to $\overline{(\text{lp}(e), u)}$ are that $C_{\tau(f)} = C_u$ and $\tau(f) \neq u$. Also by definition of a closed set, we must have that $\tau(f) <_T u$. But this implies $f \ll e$.

Statement (4) is just a restatement of Proposition 2.55 in this special case.

For (5), if $f = e$, there is nothing to prove. So suppose $f \neq e$. By definition of K_u we must then have $\tau(f) <_T u$. We prove that $f \in \overline{(\text{lp}(e), u)}$ by induction on the number of vertices of the form $\tau(f')$ with $f' \in \mathbf{E}(\Gamma)$ and $\tau(f) <_T \tau(f') <_T u$. Suppose first that there is no vertex w with $\tau(f) <_T w <_T u$ that is the tip of a bold arrow. We claim that $\tau(f) \in \text{lp}(e)$. Indeed, choose a simple path p from $\tau(f)$ to $\iota(e)$ in Γ' (we can do this since $f \in K_u$). Consider the path $[u, \tau(f)]p$. If this path is contained in T , then it equals $[u, \iota(e)]$ and so $\tau(f) \in \text{lp}(e)$. Otherwise, there is a first bold arrow f' used by p . Since p is simple, we must have $\tau(f) <_T \tau(f') \leq_T u$. See Figure 2.9. Now $u \neq \tau(f')$ by definition of Γ' . This contradicts our assumption on f , so $\tau(f) \in \text{lp}(e)$ and hence $f \in \overline{(\text{lp}(e), u)}$ by definition of a closed subgraph with respect to u .

Now we proceed by induction. Choose a simple path p from $\tau(f)$ to $\iota(e)$ in Γ' . If the path $[u, \tau(f)]p$ is contained in T , then it equals $[u, \iota(e)]$ and so

FIGURE 2.9. Factoring $p = xf'y$

$\tau(f) \in \text{lp}(e)$ and so $f \in \overline{(\text{lp}(e), u)}$ by definition of a closed subgraph with respect to u . Otherwise, there is a first bold arrow f' used by p . Again, because p is simple, we must have $\tau(f) <_T \tau(f') \leq_T u$; see Figure 2.9. Also the definition of Γ' implies $u \neq \tau(f')$. Thus $\tau(f) <_T \tau(f') <_T u$ and so by induction $f' \in \overline{(\text{lp}(e), u)}$. Therefore $\text{lp}(f') \subseteq \overline{(\text{lp}(e), u)}$ by (4). Suppose that $p = xf'y$. Then $\text{lp}(f') = [\tau(f'), \tau(f)]xf'$ and so $\tau(f) \in \text{lp}(f') \subseteq \overline{(\text{lp}(e), u)}$. We conclude that $f \in \overline{(\text{lp}(e), u)}$ by the definition of a closed subgraph with respect to u .

To establish the final statement, (1) implies $\overline{(\text{lp}(e), u)} \subseteq K_u$. On the other hand, Proposition 2.44 shows that $K_u = \bigcup_{f \in \mathbf{E}(\Gamma') \cap K_u} \text{lp}(f)$. But (4) and (5) imply that $\text{lp}(f) \subseteq \overline{(\text{lp}(e), u)}$ for all $f \in \mathbf{E}(\Gamma') \cap K_u$. This completes the proof. \square

It follows from the proposition that when forming $\text{cut}(\overline{(\text{lp}(e), u)})$ we adjoin a single new vertex (u, e) , which we normally denote simply by u' .

Proposition 2.61. *Every edge of $\text{cut}(e)$ is on a path from u to u' . Consequently, $(\text{cut}(e), u)$ is linear.*

Proof. The previous proposition implies $\overline{(\text{lp}(e), u)}$ is strongly connected and hence every edge f of $\overline{(\text{lp}(e), u)}$ is on a circuit q starting at u . Since e is the only edge entering u , the last edge of q must be e . We may assume without loss of generality that q uses the edge e only once. Then q corresponds to a path in $\text{cut}(\overline{(\text{lp}(e), u)})$ from u to u' using f .

By Proposition 2.59 $(\text{cut}(e), \tau(e))$ has the unique simple path property. It now follows from the first statement and Proposition 2.40 that $\text{cut}(e)$ is linear. \square

A key property of $(\text{cut}(e), u)$ is that it has strictly fewer bold arrows than (Γ, I) since e no longer is a bold arrow. This is important for inductive constructions.

If u, v are vertices with $v \leq_T u$, then $[u, v] \subseteq u^\downarrow$ and so we can define $\overline{([u, v], u)}$. This will be used later.

2.7. The McCammond cover of a graph. In this subsection, we present a construction, due to the first author, which the second author terms the ‘‘McCammond’’ expansion in the context of semigroups. In the setting of graphs, the term ‘‘cover’’ seems more appropriate and so we will adhere to

this terminology in this section, whereas in the context of semigroups and automata we shall use the term “expansion.”

If (Γ, I) is a rooted graph, let $\text{Simple}(\Gamma, I)$ be the set of simple paths of Γ starting at I (including the empty path). If p is a simple path in a graph and v is a vertex visited by p , denote by $p[v]$ the unique initial segment of p ending at v . Suppose that $\varphi: (\Gamma', I') \rightarrow (\Gamma, I)$ is a directed covering of rooted graphs. Then we obtain an injective map $\tilde{\varphi}: \text{Simple}(\Gamma, I) \rightarrow \text{Simple}(\Gamma', I')$ by $\tilde{\varphi}(p) = \tilde{p}$ where \tilde{p} is the unique lift of p starting at I' . Indeed, the lift of a simple path must be simple. It is then natural to focus our attention on directed covers φ where $\tilde{\varphi}$ is bijection.

Proposition 2.62. *Let $\varphi: (\Gamma', I') \rightarrow (\Gamma, I)$ be a directed cover. Then $\tilde{\varphi}: \text{Simple}(\Gamma, I) \rightarrow \text{Simple}(\Gamma', I')$ is a bijection if and only if φ takes simple paths at I' to simple paths at I .*

Proof. Suppose first that $\tilde{\varphi}$ is a bijection and let q be a simple path at I' . Then $q = \tilde{\varphi}(p)$ for some path $p \in \text{Simple}(\Gamma, I)$. But then $\varphi(q) = p$ is simple. Conversely, suppose φ takes simple paths at I' to simple paths at I . Let $q \in \text{Simple}(\Gamma', I')$; then $\varphi(q) \in \text{Simple}(\Gamma, I)$ and $q = \tilde{\varphi}(\varphi(q))$. This completes the proof. \square

Definition 2.63 (Simple covering). A morphism $\varphi: (\Gamma', I') \rightarrow (\Gamma, I)$ of rooted graphs will be called a *simple covering* if it is a directed covering satisfying the equivalent conditions of Proposition 2.62.

It turns out that each rooted graph (Γ, I) has a universal simple cover, which is characterized by having the unique simple path property. This cover is called the *McCammond cover* of (Γ, I) . First we establish its universal property (and hence uniqueness). Then we provide its construction.

Proposition 2.64. *Let $\alpha: (\tilde{\Gamma}, \tilde{I}) \rightarrow (\Gamma, I)$ be a simple covering such that $(\tilde{\Gamma}, \tilde{I})$ has the unique simple path property. Then given any simple covering $\varphi: (\Gamma', I') \rightarrow (\Gamma, I)$, there is a unique morphism (necessarily a simple covering) $\psi: (\tilde{\Gamma}, \tilde{I}) \rightarrow (\Gamma', I')$ so that the diagram*

$$\begin{array}{ccc} (\tilde{\Gamma}, \tilde{I}) & \xrightarrow{\psi} & (\Gamma', I') \\ & \searrow \alpha & \swarrow \varphi \\ & (\Gamma, I) & \end{array}$$

commutes.

Proof. Define ψ as follows. On vertices, set $\psi(v) = \tau(\tilde{\varphi}(\alpha([\tilde{I}, v])))$; this makes sense since α is a simple covering. For $e \in E(\tilde{\Gamma})$ with $e: v \rightarrow w$, define $\psi(e)$ to be the unique lift \tilde{e} of $\alpha(e)$ starting at $\psi(v)$. The uniqueness and existence of such an edge follows since φ is a directed cover and $\varphi(\tau(\tilde{\varphi}(\alpha([\tilde{I}, v]))) = \tau(\alpha([\tilde{I}, v])) = \alpha(v)$. We need to check that φ is a

graph morphism. By construction $\psi(v) = \iota(\psi(e))$. The issue is to prove $\psi(w) = \tau(\psi(e)) = \tau(\tilde{e})$.

The proof divides into two cases. Suppose first that $[\tilde{I}, v]e$ is a simple path (and hence is $[\tilde{I}, w]$). Then $\psi(w) = \tau(\tilde{\varphi}(\alpha([\tilde{I}, w]))) = \tau(\tilde{\varphi}(\alpha([\tilde{I}, v])\tilde{e})) = \tau(\psi(e))$ since $\tilde{\varphi}(\alpha([\tilde{I}, v])\tilde{e})$ lifts $\alpha([\tilde{I}, v]e) = \alpha([\tilde{I}, w])$.

Next suppose that $[\tilde{I}, v]e$ is not simple. Then $[\tilde{I}, w]$ is an initial segment of $[\tilde{I}, v]$. It follows that $\tilde{\varphi}(\alpha([\tilde{I}, w]))$ is an initial segment of $\tilde{\varphi}(\alpha([\tilde{I}, v]))$ since path lifting preserves initial segments. Next observe that $\tilde{\varphi}(\alpha([\tilde{I}, v])\tilde{e})$ is not simple as φ is a simple covering and $\varphi(\tilde{\varphi}(\alpha([\tilde{I}, v])\tilde{e})) = \alpha([\tilde{I}, v]e)$, which is not simple. Hence $\psi(w) = \tau(\tilde{e})$ is a vertex w' of $\tilde{\varphi}(\alpha([\tilde{I}, v]))$ with $\varphi(w') = \alpha(w)$. But since $\tilde{\varphi}(\alpha([\tilde{I}, v]))$ is simple and φ is a simple covering, there is exactly one vertex w' of $\tilde{\varphi}(\alpha([\tilde{I}, v]))$ lying over $\alpha(w)$. Since $\tilde{\varphi}(\alpha([\tilde{I}, w]))$ is an initial segment of $\tilde{\varphi}(\alpha([\tilde{I}, v]))$, this w' must be $\tau(\tilde{\varphi}(\alpha([\tilde{I}, w]))) = \psi(w)$. This completes the proof that ψ is a graph morphism. Since the above diagram commutes, ψ must automatically be a simple covering.

If β is another morphism making the diagram commute, then it must be a simple covering. Thus $\beta([\tilde{I}, v])$ is a simple path from I' to $\beta(v)$ and its image under φ is the simple path $\alpha([\tilde{I}, v])$. Thus $\tilde{\varphi}(\alpha([\tilde{I}, v])) = \beta([\tilde{I}, v])$ and so $\beta(v) = \psi(v)$. If e is an edge at v , then $\beta(e)$ must be the unique lift of $\alpha(e)$ at $\beta(v) = \psi(v)$. This proves uniqueness. \square

The existence of the McCammond cover (or universal simple cover) is established via the same schema as the construction of the universal covering space in algebraic topology. Abusively, in what follows we shall also use I to denote the empty path at I .

Definition 2.65 (McCammond cover). Let (Γ, I) be a rooted graph. Define $(\Gamma, I)^{M^c} = (\Gamma^{M^c}, I)$ to be the graph given by:

$$\begin{aligned} V(\Gamma^{M^c}) &= \text{Simple}(\Gamma, I); \\ E(\Gamma^{M^c}) &= \{(p, e) \mid p \in \text{Simple}(\Gamma, I), e \in E(\Gamma), \tau(p) = \iota(e)\} \end{aligned}$$

where the incidence functions are defined by $\iota(p, e) = p$ and

$$\tau(p, e) = \begin{cases} pe & pe \in \text{Simple}(\Gamma, I) \\ p[\tau(e)] & \text{else.} \end{cases}$$

Define $\eta: (\Gamma^{M^c}, I) \rightarrow (\Gamma, I)$ by $\eta(p) = \tau(p)$ on vertices and $\eta(p, e) = e$ on edges. The map η is called the *McCammond covering* of Γ . The graph Γ^{M^c} is termed the *McCammond expansion* or *McCammond cover* of Γ .

Since any directed cover of an A -automaton is an A -automaton, it follows that if (\mathcal{A}, I) is a pointed A -automaton, then $(\mathcal{A}, I)^{M^c}$ is a pointed A -automaton.

Proposition 2.66. *The map $\eta: (\Gamma^{M^c}, I) \rightarrow (\Gamma, I)$ is a simple covering and (Γ^{M^c}, I) has the unique simple path property. Hence it is the universal simple covering of (Γ, I) .*

Proof. Let us verify that the map η is a graph morphism. First we check that $\iota(\eta(p, e)) = \iota(e) = \tau(p) = \eta(\iota(p, e))$. Second, observe that $\tau(\tau(p, e))$ is a simple path ending at $\tau(e)$ in either of the two cases in the definition of τ . Thus $\eta(\tau(p, e)) = \tau(\tau(p, e)) = \tau(e) = \tau\eta(p, e)$. Thus η is a graph morphism. It is a directed covering since if $e \in E(\Gamma)$ is an edge with $\iota(e) = \eta(p) = \tau(p)$, then (p, e) is the unique edge at p mapping to e under η .

The unique lift of the empty path at I to Γ^{M^c} is the empty path at I . If $p \in \text{Simple}(\Gamma, I)$ is non-empty, say $p = e_1 \cdots e_n$, then

$$\tilde{\eta}(p) = (I, e_1)(e_1, e_1e_2) \cdots (e_1 \cdots e_{n-1}, e_n),$$

which is a simple path in (Γ^{M^c}, I) from I to p . Hence (Γ^{M^c}, I) is rooted at I . To prove that η is simple and (Γ^{M^c}, I) has the unique simple path property it therefore suffices to prove that the paths $\tilde{\eta}(p)$ with $p \in \text{Simple}(\Gamma, I)$ are the only simple paths at I in Γ^{M^c} .

Let $q \in \text{Simple}(\Gamma^{M^c}, I)$. We prove by induction on $|q|$ that $q = \tilde{\eta}(p)$ for some $p \in \text{Simple}(\Gamma, I)$. If $|q| = 0$, then $q = \tilde{\eta}(I)$. Suppose now that it is true for paths of smaller length than $|q|$. Then $q = \tilde{\eta}(p, e)$ for some simple path $p \in \text{Simple}(\Gamma, I)$ by induction. Suppose that $pe \notin \text{Simple}(\Gamma, I)$. Then $\tau(p, e) = p[\tau(e)]$. But every initial segment of p is visited by $\tilde{\eta}(p)$, contradicting that q is simple. We conclude that $\tau(p, e) = pe \in \text{Simple}(\Gamma, I)$ and so $q = \tilde{\eta}(pe)$. \square

As immediate consequences of the above result and the universal property we obtain the following corollaries.

Corollary 2.67. *(Γ^{M^c}, I) is the unique simple cover of (Γ, I) with the unique simple path property.*

Corollary 2.68. *If $\varphi: (\Gamma, I) \rightarrow (\Gamma', I')$ is a simple covering, then $(\Gamma, I)^{M^c} \cong (\Gamma', I')^{M^c}$.*

2.8. McCammond expansion of automata. As mentioned earlier, the McCammond cover of an A -automaton is again an A -automaton. In this context we call it the *McCammond expansion* of the automaton. However, the reader is cautioned that the construction is not functorial. Let's make the McCammond expansion explicit from an automaton point-of-view.

Remark 2.69 (Words and simple paths). Notice that when \mathcal{A} is a pointed A -automaton, the words readable starting at a vertex p are in one-to-one correspondence with the paths starting at p . Thus, if we let $\text{Simple}(\mathcal{A})$ be the words readable along simple paths starting at the root of \mathcal{A} , then \mathcal{A} has the unique simple path property from its root if and only if the paths corresponding to these words all end at distinct vertices of \mathcal{A} .

Next, we interpret the action of A on the simple paths starting at a vertex.

Definition 2.70 (Acting on simple paths). Let \mathcal{A} be a pointed A -automaton with root p . If u is a word readable as a simple path in \mathcal{A} from p to q (i.e., $u \in \text{Simple}(\mathcal{A})$) and a is the label on an edge in \mathcal{A} which starts at q , then there is a well-defined word in $\text{Simple}(\mathcal{A})$ that we denote $u \cdot a$. There are two possibilities. If the word ua is readable as a simple path starting at p , then $u \cdot a$ is defined to be ua ; otherwise, the path starting at p reading ua ends at a vertex q' already traversed by the path reading u . In that case, $u \cdot a$ is defined to be the initial portion of u which is read from p to q' . The two possibilities are illustrated in Figure 2.10. When $u \cdot a = ua$ we call this an *extension*, and when $u \cdot a$ is an initial subsegment of u we call it a *reduction*. Combining these results gives a partial function from $\text{Simple}(\mathcal{A}) \times A$ to $\text{Simple}(\mathcal{A})$. Notice that this partial function is a function if and only if \mathcal{A} is a complete automaton.

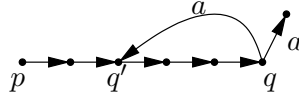


FIGURE 2.10. Multiplying a simple path by an edge.

We are now ready to rephrase the McCammond cover for automata.

Definition 2.71 (Expanding automata). Let \mathcal{A} be an A -automaton rooted at I . Then the A -automaton \mathcal{A}^{M^c} is called the *McCammond expansion* of \mathcal{A} .

Here is an example of the McCammond expansion.

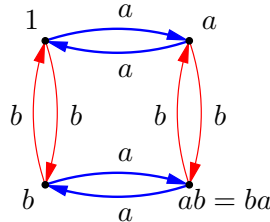


FIGURE 2.11. The Cayley graph of the Klein four-group.

Example 2.72 (Klein four-group). Let $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ be the Klein four-group with the standard presentation $\langle a, b \mid a^2 = b^2 = [a, b] = 1 \rangle$. The Cayley graph of G with respect to $A = \{a, b\}$ is shown in Figure 2.11. In Figure 2.12, the expanded graph $\text{Cay}(G, A)^{M^c}$ is shown on the left and its directed spanning tree is shown on the right. Notice that the expansion has the unique simple path property from 1 but not from the state aba since $aba \cdot b = aba \cdot aba$. Notice that all edges not in the spanning tree connect a vertex p to some vertex on the unique simple path in the tree from 1 to p .

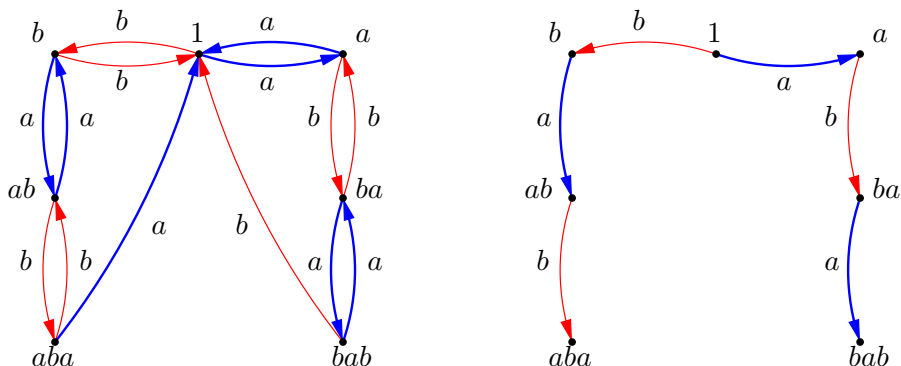


FIGURE 2.12. The expansion of the graph in Figure 2.11 and its underlying tree.

Our next example shows that the McCammond expansion is not functorial.

Example 2.73 (The McCammond expansion is not functorial). We need to include an example.

We conclude this section with a conjecture that has a long history dating back to the 1940s. We mention it here since its assertion can be reformulated in terms of depth of the tree underlying the M^c expansion to the Cayley graph of a finite group.

Conjecture 1 (Hamiltonian path). *If G is a finite group with at least 3 elements and it is generated as a monoid by A , then there exists a simple path in the Cayley graph of G with respect to A which passes through every vertex. In other words, there is a word $w \in A^+$ so that every element of G is represented uniquely by an initial segment of w .*

3. AUTOMATA FROM A UNIVERSAL ALGEBRA POINT-OF-VIEW

There is another viewpoint on A -automata that takes them to be universal algebras [9]. In what follows we write unary operations on the right of their arguments. First let's consider deterministic A -automata. Deterministic A -automata are exactly universal algebras with signature (or type) consisting of a collection of unary operations indexed by A . Given an A -automaton (Γ, ℓ) we can create an algebra with underlying set $V(\Gamma)$; the unary operation associated to $a \in A$ is $q \mapsto qa$. Conversely, if Q is a universal algebra of this type, we can form its *Cayley graph* with vertex set Q and set of edges $Q \times A$. Here $\iota(q, a) = q$ and $\tau(q, a) = qa$. The labeling sends $(q, a) \rightarrow a$. Therefore, one has universal algebraic notions of generators, relations and homomorphisms for automata. We shall see shortly how they relate to our topological notions.

One can similarly deal with partial deterministic automata by formally adjoining a sink. Let us now consider universal algebras with unary operations indexed by A and a distinguished constant \square . We also demand that they satisfy the identity $\square a = \square$ for all $a \in A$. Let \mathfrak{V} be the variety of algebras of this type satisfying these identities. Then, given an A -automaton (Γ, ℓ) we can define an algebra in this type by taking $V(\Gamma) \cup \{\square\}$ as the underlying set, defining $\square a = \square$ for all $a \in A$ and setting

$$qa = \begin{cases} qa & qa \text{ defined} \\ \square & \text{otherwise.} \end{cases}$$

Conversely, given an algebra $Q \in \mathfrak{V}$, we can define its *Cayley graph* by taking the vertex set to be $Q \setminus \{\square\}$ and the edge set to be $\{(q, a) \in Q \times A \mid qa \neq \square\}$. Of course, $\iota(q, a) = q$, $\tau(q, a) = qa$ and $\ell(q, a) = a$. Thus we can also talk about generators, relations and homomorphisms for partial automata.

Let us reinterpret these notions topologically. We say that a morphism φ of \mathfrak{V} -algebras is \square -restricted if $\varphi^{-1}(\square) = \{\square\}$.

Proposition 3.1.

- (1) *The category of algebras in \mathfrak{V} with surjective \square -restricted morphisms is equivalent to the category of A -automata with directed covers as morphisms.*
- (2) *An A -automaton (Γ, ℓ) is generated by l in the sense of universal algebra if and only if l is a generating set in the sense of Definition 2.21. In particular, rooted automata correspond to one-generated \mathfrak{V} -algebras.*
- (3) *The free algebra on a set l is $(\mathsf{l} \times A^*) \cup \{\square\}$ with the obvious \mathfrak{V} -algebra structure.*

Proof. To prove (1), suppose first that $\varphi: \mathcal{A}_1 \rightarrow \mathcal{A}_2$ is a directed covering of A -automata and define $\psi: V(\mathcal{A}_1) \cup \{\square\} \rightarrow V(\mathcal{A}_2) \cup \{\square\}$ by $\psi(\square) = \square$ and $\psi(v) = \varphi(v)$ for $v \in V(\mathcal{A}_1)$. Clearly, ψ is surjective and \square -restricted. Let us check that ψ is a homomorphism. Clearly $\psi(\square \cdot a) = \psi(\square) = \square = \square \cdot a$ for all $a \in A$. Next suppose $q \in V(\mathcal{A}_1)$ and $a \in A$. Then by definition of a directed covering, we have that $\varphi(qa) = \varphi(q)a$ in the sense that either both sides are undefined or both are defined and agreed. If both sides are undefined, then $\psi(qa) = \square = \psi(q)a$. Otherwise, $\psi(qa) = \psi(q)a$.

Conversely, suppose that $\varphi: Q \cup \{\square\} \rightarrow Q' \cup \{\square\}$ is a surjective \square -restricted morphism of \mathfrak{V} -algebras. Then we claim that φ induces a directed cover defined by φ on vertices and by $(q, a) \mapsto (\varphi(q), a)$, for $qa \neq \square$, on edges. Notice this is well defined since $qa \neq \square$ implies $\varphi(q)a = \varphi(qa) \neq \square$ as φ is \square -restricted. Evidently, φ gives a morphism of A -automata. Surjectivity on vertices is immediate. Suppose $\varphi(q)a \neq \square$. Then $\varphi(qa) = \varphi(q)a \neq \square$ and hence $qa \neq \square$ since φ is a homomorphism. Thus the edge $(\varphi(q), a)$ is a lift (q, a) (and is the unique one).

The second item is trivial. To verify (3), if Q is a \mathfrak{V} -algebra generated by l , then the map defined by $(p, w) \mapsto pw$ for $(p, w) \in \mathsf{l} \times A^*$ and $\square \mapsto \square$

is the unique morphism from $I \times A^*$ to Q extending the map $(p, 1) \mapsto p$ for $p \in I$. \square

We leave to the reader to formulate and prove the analogous result for deterministic automata.

3.1. Presentations and rewriting. It now makes sense to talk about presentations of A -automata. Since we are only interested in pointed automata, we restrict to the case of one-generated algebras from \mathfrak{A} and so we need only give the relations. It follows now that if (\mathcal{A}, I) is a pointed automaton, then it is one-generated and can be defined by relations of the form $u = v$ and $u = \square$ where $u, v \in A^*$. The relation $u = v$ forces $Iu = Iv$ in \mathcal{A} , while the relation $u = \square$ forces Iu to be undefined. If $\langle R \rangle$ is a presentation of \mathcal{A} , the underlying \mathfrak{A} -algebra of \mathcal{A} is the quotient of $A^* \cup \{\square\}$ (viewed as a right A^* -set) by the right congruence generated by the relations in R and the relations $\square a = a$. Formally, let us define the one-step derivation relation \Rightarrow by putting $\square a \Rightarrow \square$ for all $a \in A$ and $uw \Rightarrow vw$ if $u = v \in R$ where $u, v \in A^* \cup \{\square\}$, $w \in A^*$. As usual, \Rightarrow^* denotes the reflexive transitive closure of the relation \Rightarrow . We write $u \sim v$ for $u, v \in A^* \cup \{\square\}$ if there is a sequence $u = u_0, u_1 \cdots, u_n = v$ such that, for each $i = 0, \dots, n-1$, either $u_i \Rightarrow u_{i+1}$ or $u_{i+1} \Rightarrow u_i$. Then the state set of \mathcal{A} is $(A^* \cup \{\square\})/\sim$. The root is the class I of the empty word. We remark that we are essentially dealing with what is known as prefix rewriting systems [7]: relations can only be applied to prefixes of words.

We state an obvious result.

Proposition 3.2. *Let $S = \langle A \mid R \rangle$ be a presentation of an A -semigroup. Then the rooted Cayley graph $(\text{Cay}(S, A), I)$ is presented as an A -automaton by $\langle \{ur = ur' \mid u \in A^*, r = r' \in R\} \rangle$.*

Definition 3.3 (Loop automaton). A pointed automaton (\mathcal{A}, I) is said to be a *loop automaton* if it has a presentation of the form $\langle L \cup D \rangle$ where all relations in L are so-called *loop relations* of the form $uv = u$ and all relations in D are *dead-end* relations $w = \square$ where $u, v, w \in A^*$.

We now prove that automata with the unique simple path property are loop automata.

Proposition 3.4. *Let (\mathcal{A}, I) be a partial A -automaton with the unique simple path property. Then*

$$\begin{aligned} \mathcal{A} = & \langle \{ \ell(\text{slp}(e)) = \ell([I, \tau(e)]) \mid e \in \mathbf{E}(\Gamma) \} \cup \\ & \{ \ell([I, v])a = \square \mid va \text{ is undefined}, v \in V(\Gamma), a \in A \} \rangle \end{aligned}$$

where $\ell: E(\Gamma) \rightarrow A$ is the labeling function. Thus (\mathcal{A}, I) is a loop automaton.

Proof. Clearly \mathcal{A} satisfies all these relations. We show all other relations are consequences of these. Let \sim be the equivalence relation on $A^* \cup \{\square\}$

associated to the presentation above. First we claim that if Iu is defined for $u \in A^*$, then $u \sim \ell([I, Iu])$. The proof goes by induction on the number of bold arrows in the path p_u read by u from I to Iu . If there are no bold arrows, then $p_u = [I, Iu]$ and there is nothing to prove. Else we can factor $p_u = vew$ where e is the first bold arrow used by p_u . If $q = \tau(v)$, then $v = [I, q]$ and $\text{slp}(e) = [I, q]e = ve$. See Figure 3.1. Thus $\ell(vew) \sim \ell([I, \tau(e)])$

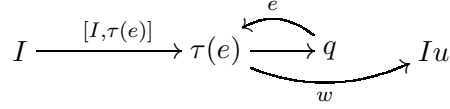


FIGURE 3.1. Reducing to the base

and so $u = \ell(vew) \sim \ell([I, \tau(e)])\ell(w) = \ell([I, \tau(e)]w)$. But $[I, \tau(e)]w$ contains strictly fewer bold arrows than p_u and hence by induction $u \sim \ell([I, \tau(e)]w) \sim \ell([I, \tau(w)]) = \ell([I, Iu])$.

Suppose now that $u, v \in A^*$ and $Iu = Iv$ with both sides defined. Then by the claim $u \sim \ell([I, Iu]) = \ell([I, Iv]) \sim v$. If $u \in A^*$ and Iu is undefined, then we can factor $u = vaw$ uniquely with Iv defined, $a \in A$ and Iva undefined. Then by the claim $v \sim \ell([I, Iv])$. Hence $u \sim \ell([I, Iv])aw \sim \square$ using the relation $\ell([I, Iv])a = \square$. This completes the proof that \mathcal{A} is given by the above presentation. \square

The above proof provides a Church-Rosser type property [7] for pointed automata with the unique simple path property. Namely, we have shown that $u \sim \square$ or $u \sim \ell([I, Iu])$ can be derived using only \Rightarrow^* . Hence if we form a rewriting system with rules

$$\begin{aligned} \ell(\text{slp}(e)) &\rightarrow \ell([I, \tau(e)]) && \text{for } e \in \mathbf{E}(\Gamma), \\ \ell([I, v])a &\rightarrow \square && \text{if } va \text{ is undefined,} \\ \square a &\rightarrow \square \end{aligned}$$

then we obtain a confluent terminating prefix rewriting system (where derivations are defined via \Rightarrow). The unique irreducible form derivable from $w \in A^*$ is \square if Iw is not defined and otherwise is $\ell([I, Iw])$, which from now on we denote $\text{Red}(w)$. Here we consider \square to be irreducible, whereas a word $w \in A^*$ is irreducible if it has no prefix which is the left hand side of a rule.

Let us formalize this.

Definition 3.5 (Elementary loop automaton). A loop automaton \mathcal{A} is said to be an *elementary loop automaton* if it can be given a loop automaton presentation $\langle L \cup D \rangle$ with the property that distinct irreducible elements for

the prefix rewriting system

$$\begin{aligned} uv &\rightarrow u && \text{such that } uv = u \in L, \\ u &\rightarrow \square && \text{such that } u = \square \in D, \\ \square a &\rightarrow \square && a \in A \end{aligned}$$

represent distinct classes of \sim .

It turns out that elementary loop automata are exactly the automata with unique simple path property.

Proposition 3.6. *A pointed A -automaton (\mathcal{A}, I) has the unique simple path property if and only if it is an elementary loop automaton.*

Proof. The proof of Proposition 3.4 shows that if (\mathcal{A}, I) has the unique simple path property, then it is an elementary loop automaton with the presentation in that proposition. We prove the converse. Suppose (\mathcal{A}, I) is an elementary loop automaton with loop presentation $\langle L \cup D \rangle$ and let N be the set of irreducible elements of A^* ; notice that it is prefix-closed. Since each element of N gives a unique state of \mathcal{A} and N is prefix-closed, it follows that subgraph of \mathcal{A} spanned by the paths labeled by the elements of N from I is a directed rooted tree isomorphic to the induced subgraph of the Cayley graph of A^* with vertices the elements of N . Moreover, it must be a spanning tree since if q is a state of \mathcal{A} , then we can find a word $w \in A^*$ with $Iw = q$. Replacing w by the unique irreducible element to which it can be reduced shows that T is a spanning tree. We claim that each edge e of \mathcal{A} that does not belong to T satisfies $\tau(e) \in [I, \iota(e)]$. Proposition 2.32 will then imply that (\mathcal{A}, I) has the unique simple path property. So suppose $u \in N$, $ua \notin N$ with $a \in A$ and Iua defined. Since ua is not irreducible and cannot be reduced to \square , we can write $ua = xyw$ where $xy = x \in L$. Since u is irreducible, we must have that w is empty and y ends in a . That is, $u = xy'$ where $y = y'a$ and $ua = xy$. But then $ua = xy \sim x$ and x is a prefix of u . Thus if e is the edge labeled by a at Iu , then $\tau(e) = Ix$, which is a vertex of $[I, Iu]$. This completes the proof. \square

3.2. The standard Kleene expression. We now aim to define a unionless Kleene expression for the language accepted by a finite trim pointed acceptor $(\mathcal{A}, I, \{q\})$ with the unique simple path property and a single terminal state q . This expression turns out to be good for pumping arguments and has the advantage of being unionless. The intuition is that for applications to complexity, one wants to associate to the automaton a unionless Kleene expression and then replace each occurrence of the Kleene star $*$ by ω in order to obtain an element of the free aperiodic ω -algebra [25].

Let $\mathcal{A} = (\Gamma, \ell)$. We assume that a geometric rank function r on \mathcal{A} has been fixed. It follows from Proposition 2.40 and $(\mathcal{A}, I, \{q\})$ being trim that \mathcal{A} is linear. Clearly the vertex q belongs to the bottommost strong component and I belongs to the topmost one. The Kleene expression $\mathcal{K}(\mathcal{A}, I, q, r)$ is defined via the *Principal of Induction* (or PI), which is based on induction

on the pair $(|\mathbf{E}(\Gamma)|, |V(\Gamma)|)$ where $\mathbb{N} \times \mathbb{N}$ is ordered lexicographically. Let $C_0 > C_1 > \dots > C_n$ be the chain of strong components of Γ and let p_i and q_i be the respective entry and exit points of C_i where we take $p_0 = I$. Let $e_i: q_{i-1} \rightarrow p_i$, for $i = 1, \dots, n$, be the i^{th} transition edge. See Figure 3.2. Notice that each (C_i, p_i) has the unique simple path property. It is possible

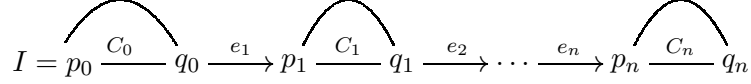


FIGURE 3.2. Trim automaton with the unique simple path property

that $p_i = q_i$ and some of the C_i may be trivial.

3.2.1. *Base of the induction.* If (\mathcal{A}, I) has no bold arrows, then it consists simply of the geodesic $[I, q]$ and we define

$$\mathcal{K}(\mathcal{A}, I, q, r) = \ell([I, q]) \quad (3.1)$$

where ℓ is the function labeling the edges of \mathcal{A} .

3.2.2. *Elementary induction.* If $n \geq 1$, then we define

$$\mathcal{K}(\mathcal{A}, I, q, r) = \mathcal{K}(\mathcal{A}(q_{n-1}), I, q_{n-1}, r|_{\mathcal{A}(q_{n-1})}) \ell(e_n) \mathcal{K}(C_n, p_n, q, r|_{C_n}) \quad (3.2)$$

where $\mathcal{A}(q_{n-1})$ is the full subautomaton of \mathcal{A} whose vertices are greater than or equal to q_{n-1} in the accessibility order (cf. Proposition 2.40). This make sense since $(\mathcal{A}(q_{n-1}), I)$ and (C_n, p_n) have the unique simple path property and each have strictly fewer vertices than \mathcal{A} and no more bold arrows.

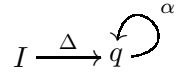
3.2.3. *Loop case.* Suppose that $n = 0$ (so there is only one strong component) and that \mathcal{A} has at least one bold arrow. Let e be the bold arrow ending at I with largest geometric rank. Let \mathcal{B} be the automaton obtained from \mathcal{A} by removing the edge e and taking the strong component of I . Then (\mathcal{B}, I) has the unique simple path property and has fewer bold arrows than \mathcal{A} . Also $\text{cut}(e)$ has the unique simple path property from I by Propositions 2.50 and is linear by Proposition 2.61. Finally, we observe that $(\overline{[I, q], I})$ has the unique simple path property. Moreover, every vertex of $(\overline{[I, q], I})$ is on a path from I to q by Proposition 2.57. Thus $(\overline{[I, q], I})$ is trim. It also has fewer bold arrows than \mathcal{A} since the graph obtained by removing the bold arrows ending at I is closed with respect to I and contains $[I, q]$. Thus we may define

$$\begin{aligned} \mathcal{K}(\mathcal{A}, I, q, r) = & \mathcal{K}(\mathcal{B}, I, I, r|_{\mathcal{B}}) [\mathcal{K}(\text{cut}(e), I, I', r|_{\text{cut}(e)}) \mathcal{K}(\mathcal{B}, I, I, r|_{\mathcal{B}})]^* \\ & \cdot \mathcal{K}(\overline{([I, q], I)}, I, q, r|_{\overline{([I, q], I)}}) \end{aligned} \quad (3.3)$$

where $r|_{\text{cut}(e)}$ makes sense as each bold arrow of $\text{cut}(e)$ is a bold arrow of \mathcal{A} .

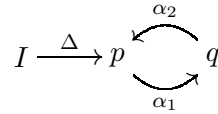
Before verifying that the Kleene expression yields the correct language, we compute some examples. In the first few examples there will be no multiple bold arrows ending at the same vertex and so we do not need to give the geometric rank function.

Example 3.7. Suppose the automaton \mathcal{A} is given by



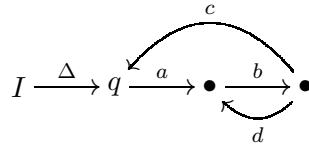
Then $\mathcal{K}(\mathcal{A}, I, q, r) = \Delta\alpha^*$.

Example 3.8. Suppose the automaton \mathcal{A} is given by

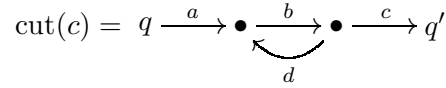


Then $\mathcal{K}(\mathcal{A}, I, q, r) = \Delta(\alpha_1\alpha_2)^*\alpha_1$.

Example 3.9. Let the automaton \mathcal{A} be given by



Then



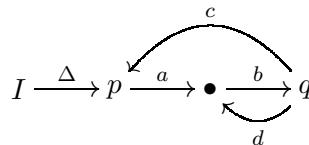
A computation similar to the previous example shows that

$$\mathcal{K}(\text{cut}(c), q, q', r) = a(bd)^*bc.$$

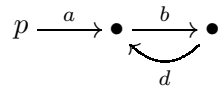
Thus $\mathcal{K}(\mathcal{A}, I, q, r) = \Delta(a(bd)^*bc)^*$.

Our next example modifies the previous one by changing the terminal state q .

Example 3.10. Suppose the automaton \mathcal{A} is given by



The only new ingredient is to compute $\overline{([p, q], p)}$, which is



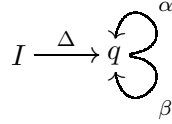
and so has Kleene expression

$$\mathcal{K}\left(\left(\overline{[p, q]}, p\right), p, q, r\right) = a(bd)^*b.$$

Thus $\mathcal{K}(\mathcal{A}, I, q, r) = \Delta(a(bd)^*bc)^*a(bd)^*b$.

In our next example, we will make use of the geometric rank function.

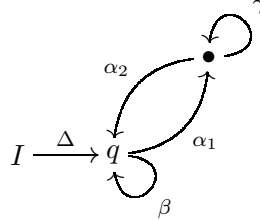
Example 3.11. Let the automaton \mathcal{A} be given by



and let us suppose that β has greater geometric rank than α . Then the Kleene expression is $K(\mathcal{A}, I, q, r) = \Delta\alpha^*(\beta\alpha^*)^*$.

A slight variation on the previous example is the following.

Example 3.12. Let the automaton \mathcal{A} be given by



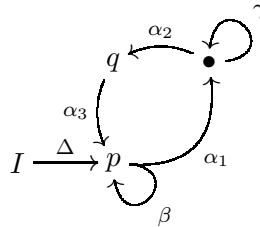
and let us suppose that β has greater geometric rank than α_2 . Then

$$\text{cut}(\alpha_2) = q \xrightarrow{\alpha_1} \bullet \xrightarrow{\alpha_2} q'$$

and so $K(\mathcal{A}, I, q, r) = \Delta(\alpha_1\gamma^*\alpha_2)^*[\beta(\alpha_1\gamma^*\alpha_2)^*]^*$.

Our final example mixes several ingredients.

Example 3.13. Suppose the automaton \mathcal{A} is given by



and let us suppose that α_3 has greater geometric rank than β . Then

$$\text{cut}(\alpha_3) = p \xrightarrow{\alpha_1} \bullet \xrightarrow{\alpha_2} q \xrightarrow{\alpha_3} p'$$

and

$$\overline{([p, q], p)} = p \xrightarrow{\alpha_1} \bullet \xrightarrow{\alpha_2} q$$

Thus $K(\mathcal{A}, I, q, r) = \Delta\beta^* [(\alpha_1\gamma^*\alpha_2\alpha_3)\beta^*]^* \alpha_1\gamma^*\alpha_2$.

We now prove that the Kleene expression accepts the language recognized by the automaton.

Theorem 3.14. *Let $(\mathcal{A}, I, \{q\})$ be a finite trim pointed A -acceptor with the unique simple path property and with a single terminal state q where $\mathcal{A} = (\Gamma, \ell)$. Fix a geometric rank function r on (\mathcal{A}, I) . Then $\mathcal{K}(\mathcal{A}, I, q, r)$ is a Kleene expression for the language of words accepted by \mathcal{A} with initial state I and terminal state q that does not use union.*

Proof. The proof goes by induction on the pair $(|\mathbf{E}(\Gamma)|, |V(\Gamma)|)$ where we order $\mathbb{N} \times \mathbb{N}$ lexicographically. It is clear that if \mathcal{A} has no bold arrows, then $\mathcal{K}(\mathcal{A}, I, q, r)$ is unionless and accepts the same language as \mathcal{A} . Similarly, in the case of the elementary induction 3.2.2, one easily verifies that $\mathcal{K}(\mathcal{A}, I, q, r)$ is a unionless Kleene expression accepting the language of \mathcal{A} using the induction hypothesis. The only difficult case is the loop case 3.2.3.

Let us retain the notation of 3.2.3. By induction we have \mathcal{B} with initial and terminal state both set to I accepts precisely $\mathcal{K}(\mathcal{B}, I, I, r|_{\mathcal{B}})$, whereas $\mathcal{K}(\text{cut}(e), I, I', r|_{\text{cut}(e)})$ accepts all strings reading from I to I' in $\text{cut}(e)$ and $\mathcal{K}(\overline{([I, q], I)}, I, q, r|_{\overline{([I, q], I)}})$ accepts all words reading in $\overline{([I, q], I)}$ from I to q . Then trivially, $\mathcal{K}(\mathcal{A}, I, q, r)$ is accepted by \mathcal{A} with initial state I and final state q and is a unionless Kleene expression using the induction and (3.3). The converse direction requires some proof.

Suppose w labels a path from I to q . We can factor $w = uv$ with $u, v \in A^*$ so that $Iu = I$ and the path labeled by v from I to q does not revisit I . First we claim that v is accepted by $\overline{([I, q], I)}$ with initial state I and terminal state q . If $q = I$, then v is empty and there is nothing to prove. Otherwise, we proceed by induction on the number of applications of \Rightarrow needed to reduce v to $\ell([I, q])$ in the rewriting system appearing after Proposition 3.4. If $v = \ell([I, q])$, then there is nothing to prove. Suppose $v \Rightarrow x$. Then we can write $v = \ell(\text{slp}(f))z$ for some bold arrow f with $x = \ell([I, \tau(f)])z$. By choice of v , the edge f does not end at I . By induction, x is accepted by $\overline{([I, q], I)}$. Since f is a bold arrow with $\tau(f)$ visited by x on its run from I , it follows from the definition of being closed with respect to I that $f \in \overline{([I, q], I)}$. Thus $\text{lp}(f) \subseteq \overline{([I, q], I)}$ by Proposition 2.55. Therefore, $\text{slp}(f) = [I, \tau(f)]\text{lp}(f) \subseteq \overline{([I, q], I)}$ and so v is accepted by $\overline{([I, q], I)}$ with terminal state q . Hence $v \in \mathcal{K}(\overline{([I, q], I)}, I, q, r|_{\overline{([I, q], I)}})$.

It thus remains to show that

$$u \in \mathcal{K}(\mathcal{B}, I, I, r|_{\mathcal{B}}) [\mathcal{K}(\text{cut}(e), I, I', r|_{\text{cut}(e)}) \mathcal{K}(\mathcal{B}, I, I, r|_{\mathcal{B}})]^*.$$

Let s be the path read by u from I to I and factor $p = s_0 e s_1 e \cdots s_{m-1} e s_m$ where no s_i uses the edge e and where we interpret the case $m = 0$ as $p = s_0$. Then $\ell(s_m)$ is accepted by \mathcal{B} with initial and terminal state I . We may further factor each s_i with $0 \leq i \leq m-1$ as $s_i = t_i u_i$ so that $t_i: I \rightarrow I$ and so that $u_i: I \rightarrow \iota(e)$ does not revisit I . Then $\ell(t_i)$ is accepted by \mathcal{B} with initial and terminal state I for all $0 \leq i \leq m-1$. We claim that $\ell(u_i)a$ is accepted by $\text{cut}(e)$ with initial state I and terminal state I' . If $I = \iota(e)$ (that is, e is a loop edge), then each u_i is empty and $\text{cut}(e) = I \xrightarrow{a} I'$ and so there is nothing to prove. So assume $I \neq \iota(e)$. We show by induction on the number of applications of \Rightarrow needed to reduce $\ell(u_i)$ to $\ell([I, \iota(e)])$ in the rewriting system after Proposition 3.4 that $\ell(u_i)$ reads in $\text{cut}(e)$ from I to $\iota(e)$. Since a labels an edge from $\iota(e)$ to I' in $\text{cut}(e)$, this will establish the claim. Recall that $\text{cut}(e)$ is obtained by cutting $\overline{(\text{lp}(e), I)}$ at I . If $u_i = [I, \iota(e)]$ there is nothing to prove since $[I, \iota(e)] \subseteq \overline{(\text{lp}(e), I)}$ and doesn't use the edge e , and hence "is" the geodesic from I to $\iota(e)$ in $\text{cut}(e)$. Thus $\ell(u_i)$ reads from I to $\iota(e)$ in $\text{cut}(e)$. Next suppose that $\ell(u_i) \Rightarrow x$ where x reads a path from I to $\iota(e)$ in $\text{cut}(e)$. By definition of \Rightarrow , we can write $\ell(u_i) = \ell(\text{slp}(f))z$ for some bold arrow f so that $x = \ell([I, \tau(f)])z$. Since u_i does not revisit I , it follows that f does not end at I . Then since x labels a path in $\overline{(\text{lp}(e), I)}$ from I to $\iota(e)$, it follows by the definition of a closed subgraph at I that $f \in \overline{(\text{lp}(e), I)}$ and hence, by Proposition 2.55, we must have $\text{lp}(f) \subseteq \overline{(\text{lp}(e), I)}$. Consequently, $\text{slp}(f) = [I, \tau(f)]\text{lp}(f)$ is contained in $\overline{(\text{lp}(e), I)}$. Because u_i does not revisit I , it follows that $\ell(u_i) = \ell(\text{slp}(f))z$ is readable from I to $\iota(e)$ in $\text{cut}(e)$.

Thus we now have (using the inductive hypothesis)

$$\begin{aligned} u &= [\ell(t_0)(\ell(u_0)a)] \cdots [\ell(t_{m-1})(\ell(u_{m-1})a)]\ell(s_m) \\ &\in [\mathcal{K}(\mathcal{B}, I, I, r|_{\mathcal{B}}) \mathcal{K}(\text{cut}(e), I, I', r|_{\text{cut}(e)})]^{m-1} \mathcal{K}(\mathcal{B}, I, I, r|_{\mathcal{B}}) \\ &\subseteq \mathcal{K}(\mathcal{B}, I, I, r|_{\mathcal{B}}) [\mathcal{K}(\text{cut}(e), I, I', r|_{\text{cut}(e)}) \mathcal{K}(\mathcal{B}, I, I, r|_{\mathcal{B}})]^* \end{aligned}$$

as required. This completes the proof of the theorem. \square

4. SEMIGROUP EXPANSIONS

4.1. Semigroups. Once again we fix an alphabet A (usually assumed finite). In later sections we often require semigroups to have certain finiteness properties such as being finite, or at least finite \mathcal{J} -above. A semigroup S is *finite \mathcal{J} -above* if there are only finitely many elements in S which are \mathcal{J} -above (i.e., $\geq_{\mathcal{J}}$) any particular element of S . We also need the following equivalent characterization. In this paper we allow ideals to be empty.

Lemma 4.1. *A semigroup S is finite \mathcal{J} -above if and only if there exists a family of co-finite ideals whose intersection is empty.*

Proof. If such a family of ideals exists, then for each $s \in S$ there exists an ideal I in the family with $s \notin I$. Since $S \setminus I$ is finite and contains all

Notation	Description
Sgp	semigroups
FSgp	finite semigroups
FJSgp	finite \mathcal{J} -above semigroups
Sgp_A	A -semigroups
FSgp_A	finite A -semigroups
FJSgp_A	finite \mathcal{J} -above A -semigroups

TABLE 1. Commonly used categories

the elements of S that are \mathcal{J} -above s , it follows that S is finite \mathcal{J} -above. Conversely, if S is finite then the empty ideal is co-finite and the result is immediate. If S is infinite and finite \mathcal{J} -above, then for each \mathcal{J} -class J let I_J denote the elements in S which are not \mathcal{J} -above J . Clearly these ideals are co-finite and we will show that their intersection is empty. Since S is infinite, for each $s \in S$ there is an element t which is not \mathcal{J} -above s . Because s is not in the ideal corresponding to the \mathcal{J} -class of st , evidently s is not in the intersection of all such ideals and thus the intersection of these ideals is empty. \square

Definition 4.2 (Categories). In Table 1 we list the categories commonly used throughout the article and the notations used to describe them. When the objects are A -semigroups, the morphisms are restricted to A -morphisms.

4.2. Straightline automata. In this subsection we define two finite-state automata for each finite \mathcal{J} -above A -semigroup S and each word $w \in A^+$, one which depends on the word w (the straightline automaton), and a second which only depends on the element in S which w represents (the Cayley automaton). These two types of automata will be crucial to all of the results that follow.

Straightline automata play a critical role in the solution to the word problem for Burnside semigroups [24].

Definition 4.3 (Straightline automata). Let (\mathcal{A}, I) be a pointed A -automaton and w a word in A^+ . The *straightline automaton of w with respect to \mathcal{A}* , denoted $\text{Str}^{\mathcal{A}}(w)$, has an underlying directed graph which is a subgraph of \mathcal{A} . The directed graph underlying the automaton $\text{Str}^{\mathcal{A}}(w)$ is the union of the path from I labeled by w , and the strong components of each of the vertices that this path passes through. This labeled directed graph is made into an acceptor by specifying the vertex I as its initial state and Iw as its terminal state.

If S is an A -semigroup and w is a word in A^+ , then the *straightline automaton of w with respect to S* is by definition $\text{Str}^S(w) = \text{Str}^{\text{Cay}(S,A)}(w)$.

Observe that its vertex set consists of all of the vertices that represent elements of S^I which are \mathcal{R} -equivalent to elements of S^I represented by initial segments of w .

Remark 4.4. If (\mathcal{A}, I) is a finite pointed linear A -automaton and $w \in A^+$ reads from I to a vertex in the bottommost strong component of \mathcal{A} , then $\mathcal{A} = \text{Str}^{\mathcal{A}}(w)$.

The following properties of $\text{Str}^S(w)$ follow easily from the definition.

Lemma 4.5 ($\text{Str}^S(w)$). *If S is an arbitrary finite \mathcal{J} -above A -semigroup and w is word in A^+ , then $\text{Str}^S(w)$ (with initial state I and terminal state $[w]_S$) is a trim, deterministic finite-state acceptor which is linear and its strong components are Schützenberger graphs of $\text{Cay}(S, A)$. In addition, a word $x \in A^+$ is readable on $\text{Str}^S(w)$ starting at the initial state if and only if x is \mathcal{R} -equivalent to an initial segment of w .*

The transition edges of the linear graph $\text{Str}^S(w)$ are exactly those edges whose endpoints represent elements of S which are not \mathcal{R} -equivalent to each other, in other words, edges which represent descents in the \mathcal{R} -order. Thus every edge in $\text{Str}^S(w)$ either is a transition edge or it lies in a unique Schützenberger graph. Let the vertices p_i and q_i be defined as in Definition 2.9 and let s_i and t_i denote the elements of S which corresponding to these vertices. In addition, if $\text{Str}^S(w)$ has exactly $k + 1$ strong components, we define p_0 to be the initial state of $\text{Str}^S(w)$ and q_k to be its terminal state. For $i = 0, 1, \dots, k$, the vertex p_i will be the first vertex in the path w which lies in the i^{th} strong component and q_i will be the last such vertex. Thus, p_i and q_i can also be identified with specific initial segments of w . Under this scheme $p_0 = q_0$ corresponds to the empty prefix and q_k corresponds to word w itself. Notice that the vertices p_0 and q_0 will always correspond since the Cayley graph used to define $\text{Str}^S(w)$ is that of S^I .

Example 4.6. Figure 4.1 illustrates the structure of a typical straightline automaton. In this example there are $k + 1$ strong components. The 0^{th} strong component is trivial, as it always is, and the entry and exit points are equal in the third strong component.

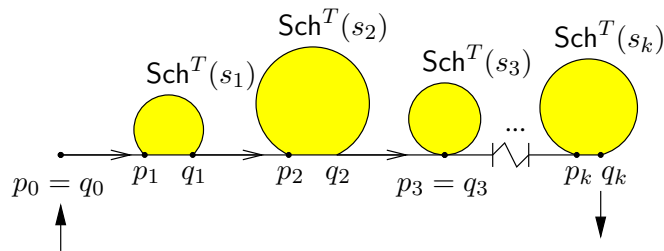


FIGURE 4.1. A straightline automaton

Of course, there is an obvious analogue of Lemma 4.5 for straightline automata associated to finite pointed automata.

Definition 4.7 (Cayley automata). Let S be an A -semigroup and let w be a word in A^+ . The *Cayley automaton of w with respect to S* is the full subgraph of $\text{Cay}(S, A)$ on the set of vertices which represent elements of S^I \mathcal{R} -above $[w]_S$. The initial state and terminal state are defined as the vertices I and $[w]_S$, as before. The Cayley automaton of w will be denoted $\text{Cay}^S(w)$. Notice that $\text{Cay}^S(w)$ actually only depends on the element of S that w represents and not on the word itself. Thus, following the same convention as for Schützenberger graphs, if $[w]_S = s$, we might write $\text{Cay}^S(s)$ instead.

The following lemmas records some elementary properties of $\text{Cay}^S(w)$.

Lemma 4.8 ($\text{Cay}^S(w)$). *If S is a finite \mathcal{J} -above A -semigroup and w is a word in A^+ , then $\text{Cay}^S(w)$ (with initial state I and terminal state $[w]_S$) is a trim, deterministic, finite-state automaton which accepts the language of words equivalent to w in S . In addition, the finite-state automaton $\text{Str}^S(w)$ is a subautomaton of $\text{Cay}^S(w)$. Moreover, a word $x \in A^+$ is readable on $\text{Cay}^S(w)$ if and only if $[x]_S \geq_{\mathcal{J}} [w]_S$, it is readable on $\text{Cay}^S(w)$ starting at the initial state if and only if $[x]_S \geq_{\mathcal{R}} [w]_S$, and it is readable on $\text{Cay}^S(w)$ ending at $[w]_S$ if and only if $[x]_S \geq_{\mathcal{L}} [w]_S$.*

Example 4.9. Figure 4.2 illustrates the structure of a possible Cayley automaton. Notice that the induced order on strong components in this example is not linear, and that transition edges between two particular strong components are not always unique. On the other hand, the automaton $\text{Str}^S(w)$ shown in Figure 4.1 is visible as a subautomaton of $\text{Cay}^S(w)$.

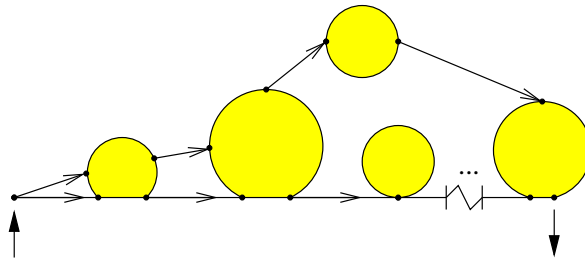


FIGURE 4.2. A Cayley automaton

The situation where $\text{Str}^S(w)$ and $\text{Cay}^S(w)$ coincide is particularly nice.

Lemma 4.10 (Equal). *Let S be a finite \mathcal{J} -above A -semigroup and let w be a word in A^+ . If $\text{Str}^S(w) = \text{Cay}^S(w)$, then this trim, deterministic finite state acceptor — which depends only on the element w represents — both is linear and accepts the language of words equivalent to w in S . In addition, every element \mathcal{R} -above $[w]_S$ is \mathcal{R} -equivalent to an initial segment of w .*

Similarly, if the left-handed versions of $\text{Str}^S(w)$ and $\text{Cay}^S(w)$ are identical, then every element \mathcal{L} -above $[w]_S$ is \mathcal{L} -equivalent to a final segment of w .

Finally, we record a fact about Schützenberger graphs for later use. Recall that an element s of a semigroup S is *regular* if $s = sts$ for some $t \in S$. For instance, idempotents are regular.

Lemma 4.11 ($\text{Sch}^S(w)$). *Suppose S is an A -semigroup, w and x are words in A^+ , and w represents a regular element of S . Then x is readable on $\text{Sch}^S(w)$ if and only if $[x]_S \geq_{\mathcal{J}} [w]_S$.*

Proof. Consider the automaton $\text{Cay}^S(w)$ and recall that $\text{Sch}^S(w)$ is the strong component of the vertex $[w]_S$ in $\text{Cay}^S(w)$. If x is readable on $\text{Sch}^S(w)$, then since the graph is strongly connected, there exist words $y, z \in A^*$ so that yxz is readable as a loop starting and ending at $[w]_S$. But this means that $wyxz$ is accepted by $\text{Cay}^S(w)$ and hence $[wyxz]_S = [w]_S$. In particular, $[x]_S \geq_{\mathcal{J}} [w]_S$. Conversely, if $[x]_S$ is \mathcal{J} -above $[w]_S$ then there exist $y, z \in A^*$ such that $[yxz]_S = [w]_S$. Suppose that $[wrw]_S = [w]_S$ with $r \in A^+$, using regularity of $[w]_S$. Then this means that $[wryxz]_S = [wrw]_S = [w]_S$. Since $\text{Cay}^S(w)$ is partial deterministic this forces $ryxz$ to be readable as a loop in $\text{Sch}^S(w)$ based at $[w]_S$ and, in particular, x is readable in $\text{Sch}^S(w)$, as $[wry]_S \mathcal{R} [w]_S$. \square

4.3. Expansions. In this subsection we review the general notion of an expansion and record some of their chief properties.

Definition 4.12 (Expansions). Let \mathbf{C} be a subcategory of \mathbf{Sgp} . An *expansion on \mathbf{C}* is a functor $E: \mathbf{C} \rightarrow \mathbf{C}$ together with a surjective natural transformation η from E to the identity functor $1_{\mathbf{C}}$ on \mathbf{C} ; see [23] for more on functors and natural transformations. More explicitly, an expansion assigns to every semigroup S in \mathbf{C} a semigroup S^E and a surjective morphism $\eta_S: S^E \twoheadrightarrow S$, and to every morphism $f: S \rightarrow T$ in \mathbf{C} it assigns a morphism $f^E: S^E \rightarrow T^E$ in \mathbf{C} so that the diagram

$$\begin{array}{ccc} S^E & \xrightarrow{f^E} & T^E \\ \eta_S \downarrow & & \downarrow \eta_T \\ S & \xrightarrow{f} & T \end{array}$$

commutes. The semigroups and morphisms assigned must remain in \mathbf{C} . One must assign identity morphisms to identity morphisms and composition must be preserved. If S^E remains finite whenever S is finite, we say that E *preserves finiteness*. Suppose that E and F are expansions. The semigroup $(S^E)^F$ is denoted $S^{E.F}$.

Most of the expansions defined in this article are expansions on a category of A -semigroups. Notice that such an expansion is uniquely determined by the semigroups assigned since A -morphisms such as $f^E: S^E \rightarrow T^E$ and $\eta_S: S^E \twoheadrightarrow S$ are uniquely determined by their domain and range whenever

they exist. Thus only the existence of morphisms such as f^E and η_S need to be checked: all of the remaining functoriality requirements follow automatically. We also encounter situations where the objects can be expanded but the morphisms cannot. In other words, for every S , there is an object S^E and a map $S^E \rightarrow S$, but for maps $S \rightarrow T$ there may or may not exist a map $S^E \rightarrow T^E$. In this situation we call $S \mapsto S^E$ an *object expansion*.

The next lemma essentially shows that an expansion which preserves finiteness also preserves the property of being finite \mathcal{J} -above.

Lemma 4.13 (Preserving finiteness). *If E is an expansion on \mathbf{Sgp}_A which preserves finiteness and for all semigroups $S \in \mathbf{Sgp}_A$ and for all ideals $I \subseteq S$, there is a map $(S/I)^E \rightarrow (S^E/J)$ where $J = \eta_S^{-1}(I)$, then E applied to a finite \mathcal{J} -above semigroup remains finite \mathcal{J} -above.*

Proof. Let S be a finite \mathcal{J} -above A -semigroup and let \mathcal{F} be a family of co-finite ideals in S whose intersection is empty (Lemma 4.1). Consider the inverse images of these ideals under $\eta_S: S^E \rightarrow S$. The intersection of these ideals in S^E is clearly empty, and they are also co-finite since for each $I \in \mathcal{F}$, S^E/J is the image of the semigroup $(S/I)^E$ which is finite by hypothesis. \square

For later use we briefly review some of the most common expansions.

Definition 4.14 (Common expansions). Let S be an A -semigroup. Roughly speaking, the *reverse or right Rhodes expansion* remembers the exit points as the path w travels through $\text{Cay}^S(w)$, while the *reverse or right Karnofsky-Rhodes* variation remembers the actual transition edges used (i.e., exit point, entry point and particular edge traversed). Note that the reverse Karnofsky-Rhodes expansion of S clearly maps onto the reverse Rhodes expansion of S . The reverse Rhodes expansion is thus a fairly small expansion in which $\text{Cay}^T(w)$ is quasilinear for all $w \in A^+$, and the *reverse Karnofsky-Rhodes expansion*, is the smallest expansion such that $\text{Cay}^T(w)$ is linear for all $w \in A^+$. We denote the reverse Rhodes expansion by $\wedge \mathcal{R}$ and the reverse Karnofsky-Rhodes by $\wedge \ell 1$.

The original versions of these two expansions were left-handed versions which we denote $\wedge \mathcal{L}$ and $\wedge \ell 1^{op}$. In other words, the statements above hold when we define left-handed versions of $\text{Str}^S(w)$ and $\text{Cay}^S(w)$ using the left Cayley graph and we read words from right-to-left instead of left-to-right. The *Birget-Rhodes expansion*, denoted BR , is the limit A -semigroup obtained by iterating the Rhodes and the reverse Rhodes expansions until the result stabilizes. All five of these expansions preserve finiteness. See [6] for a precise definition of the Birget-Rhodes expansion, and [13], [14] or [34] for precise definitions of the Rhodes expansion, the Karnofsky-Rhodes expansion, and their reverses.

For the reader's convenience we include here the definition of the right Karnofsky-Rhodes expansion (cut to the generating set A) so that the reader gets the idea.

Definition 4.15 (Karnofsky-Rhodes expansion). If S is an A -semigroup, define a congruence on A^+ by putting $u \equiv v$ if:

- (1) $[u]_S = [v]_S$;
- (2) The paths read by u and v from I to $[u]_S = [v]_S$ in $\text{Cay}(S, A)$ use the same transition edges.

The quotient semigroup A^+/\equiv is denoted $S^{\wedge \ell 1}$. It is immediate from (1) that there is a surjective morphism $\eta_S: S^{\wedge \ell 1} \rightarrow S$. If $f: S \rightarrow T$ is a homomorphism of A -semigroups, then $f(s) \xrightarrow{a} f(s)a$ a transition edge of $\text{Cay}(T, A)$ forces $s \xrightarrow{a} sa$ to be a transition edge of $\text{Cay}(S, A)$. It follows that there exists a map $f^{\wedge \ell 1}: S^{\wedge \ell 1} \rightarrow T^{\wedge \ell 1}$ and so we have defined an expansion on the category of A -semigroup. It is easy to see from (1) and (2) that $\wedge \ell 1$ preserves finiteness.

Remark 4.16. It is known that the projection $\eta_S: S^{\wedge \ell 1} \rightarrow S$ is universal amongst maps from A -semigroups to S with derived semigroupoid dividing a locally trivial category [14, 31].

When $S = S^E$ we say that S is *stable under the expansion E* . The following result is an immediate consequence of Definition 4.14.

Lemma 4.17. *Let S be a finite \mathcal{J} -above A -semigroup. If S is stable under the reverse Rhodes expansion $\wedge \mathcal{R}$, then $\text{Str}^S(w)$ and $\text{Cay}^S(w)$ have the same vertex set for all words $w \in A^+$ and if S is stable under the reverse Karnofsky-Rhodes expansion, $\wedge \ell 1$, then $\text{Str}^S(w) = \text{Cay}^S(w)$ for all words $w \in A^+$ and all of the properties listed in Lemma 4.10 must hold. Similar results hold for the left-handed versions when S is stable under $\wedge \mathcal{L}$ or $\wedge \ell 1^{op}$.*

Proof. We just handle the case of the reverse Karnofsky-Rhodes expansion as the other case is similar. Since $\text{Str}^S(w)$ is always a subautomaton of $\text{Cay}^S(w)$, it suffices to show the reverse inclusion. To do this, it suffices by the definitions to show that each \mathcal{R} -class of elements \mathcal{R} -above $[w]_S$ contains an initial segment of w . Suppose that $[u]_S \geq_{\mathcal{R}} [w]_S$. Then $[w]_S = [uv]_S$ for some $v \in A^*$. We can factor $u = u'u''$ so that reading u'' from $[u']_S$ stays in a Schützenberger graph, i.e., the last letter of u' labels the last transition edge e read by u from I in $\text{Cay}(S, A)$. The since $S = S^{\wedge \ell 1}$, it follows that w uses the same transition edges as uv and so, in particular, uses e . Thus $w = w'w''$ where the last edge read by w' from I is e . Then $[w']_S = [u']_S$ \mathcal{R} $[u'u'']_S = [u]_S$. This establishes that each \mathcal{R} -class of elements \mathcal{R} -above $[w]_S$ contains an initial segment of w , as was required. \square

For expansions closed under iteration, there is a general stability result.

Lemma 4.18 (Stability). *Let \mathbf{C} be a subcategory of \mathbf{Sgp}_A and let E and F be expansions on \mathbf{C} . If, for every A -semigroup S in \mathbf{C} , $S^{E.E} = S^E$ and there exists a map $S^E \rightarrow S^F$, then for every S in \mathbf{C} , $S^E = S^{E.F} = S^{F.E}$. In other words, S^E is stable under F . Conversely, if $S^E = S^{E.F}$ or $S^E = S^{F.E}$, then there is a map from $S^E \rightarrow S^F$.*

Proof. If we apply the E -expansion to the maps $S^E \rightarrow S^F \rightarrow S$ we get maps in each direction between S^E and $S^{F.E}$ as a result of the hypothesis that $S^{E.E} = S^E$. Within the category of A -semigroups this implies they are isomorphic. Similarly there is a map $S^{E.F} \rightarrow S^E$ since F is an expansion and a map $S^{E.E} \rightarrow S^{E.F}$ since the hypothesis of the theorem can be applied to the S^E instead of S . Once again, maps in both directions implies they are isomorphic. The final assertion is easy since $S^{E.F} \rightarrow S^F$ is the F -expansion applied to map $S^E \rightarrow S$ and $S^{F.E} \rightarrow S^F$ is the E -expansion of semigroup S^F projected back to S^F . \square

We conclude this section by reviewing the notion of a relational morphism, its relationship with expansions, and a remark on the most common ways to construct expansions.

Definition 4.19 (Relational morphisms). Let S and T be semigroups and recall that a *relational morphism from S to T* is a subsemigroup $R \subseteq S \times T$ such that the natural projection from R to S is onto. If $\varphi: R \rightarrow S$ and $\psi: R \rightarrow T$ denote the natural projections restricted to R , then the “relation” alluded to in the name relational morphism is the relation $\psi \circ \varphi^{-1}$. A relation-based definition of a relational morphism also exists, but the equivalent definition given above is easier to work with in practice. We thus refer to the subsemigroup $R \subseteq S \times T$ as the relational morphism from S to T even though technically, it merely encodes the corresponding relation.

In the category of A -semigroups, the concept is further simplified since there is a *canonical relational morphism* between any two A -semigroups, which automatically projects onto both the domain and the range. It is defined as follows. Let S and T be A -semigroups with functions $f: A \rightarrow S$ and $g: A \rightarrow T$. Their *product* in the category \mathbf{Sgp}_A (sometimes called their *product cut-to-generators*) is the subsemigroup of $S \times T$ generated by the ordered pairs $(f(a), g(a))$, $a \in A$. Call this A -subsemigroup R and let $\varphi: R \rightarrow S$ and $\psi: R \rightarrow T$ be the A -morphisms derived from the projection maps. The canonical relational morphism between S and T is the relation $\psi \circ \varphi^{-1}$. The reader should be aware that the composition of two canonical relational morphisms need not be canonical.

The concept of a relational morphism is critically important in finite semigroup theory (see, for example, [13, 26, 33, 36, 38, 39]). Moreover, given a semigroup variety \mathcal{V} and a finite A -semigroup S , it is also important to have detailed information about the relational morphisms from S to the A -semigroups $V \in \mathcal{V}$, particularly when $V \in \mathcal{V}$ is itself finite.

Remark 4.20 (Relational morphisms and expansions). The notion of a relational morphism is very closely tied to that of an expansion and the problem of constructing relational morphisms can often be reduced to the problem of constructing finite expansions of finite A -semigroups. Let E be an expansion on \mathbf{FSgp}_A . Starting from an A -semigroup S and a variety \mathcal{V} , the semigroup S^E can be sent to its maximal image $V \in \mathcal{V}$ (which exists since

\mathcal{V} is a variety). Moreover, since S^E is finite, V is finite. Then,

$$\begin{array}{ccc} S^E & \longrightarrow & V \\ \downarrow & & \\ S & & \end{array}$$

is a relational morphism. Conversely, given a systematic method for constructing relational morphisms

$$\begin{array}{ccc} R & \longrightarrow & V \\ \downarrow & & \\ S & & \end{array}$$

where S is an arbitrary finite semigroup, V is a finite member of \mathcal{V} (depending on S), and all semigroups and morphisms are A -semigroups and A -morphisms, the assignment $S \mapsto R$ is often an expansion on \mathbf{Sgp}_A . See [6] and [13] for a more detailed discussion.

Remark 4.21 (Constructing expansions). Four of the most common ways to construct expansions on \mathbf{FSgp}_A are:

- (1) Ramsey Theory;
- (2) Zimin words and uniformly recurrent sequences;
- (3) Semidirect product expansions;
- (4) Mal'cev expansions.

The first method was applied brilliantly by Ash in [2] to solve the Type-II Conjecture (see [15] for a good general reference on Ramsey theory and see [1, 2, 5, 18, 30] for other applications to semigroup theory). The second method relies heavily on classical universal algebra and semigroup theory, mixed with combinatorics (the article [19] contains an excellent survey of this approach). The third method, which includes the Karnofsky-Rhodes expansion, is expounded in [14] (see also [31]). The fourth method is the one used in this paper and is described in the next section; see also [14, 31].

The last two methods of constructing expansions are encompassed by the notion of relatively free relational morphisms with respect to a variety of relational morphisms [32, Chapter 3].

4.4. Mal'cev expansions. In this subsection we define the Mal'cev expansion of a semigroup with respect to a variety. The results in this section are in general well-known but they provide a language and a context for the newer results contained in the later sections.

Definition 4.22 (Mal'cev kernels). The *Mal'cev kernel* of a semigroup morphism $\varphi: S \rightarrow T$ is the collection of inverse images of idempotents. In other words, it is the set of subsemigroups $\{\varphi^{-1}(e)\}$ where e is an idempotent in T . If each semigroup in the Mal'cev kernel is contained in a variety \mathcal{V} we say the *Mal'cev kernel lies in \mathcal{V}* .

The main result we need involving Mal'cev kernels is Brown's theorem. Recall that a variety is called *locally finite* if its finitely generated free objects remain finite (or equivalently its finitely generated members are all finite).

Theorem 4.23 (Brown [8]). *Let $\varphi: S \rightarrow T$ be a map between two A -semigroups where A is finite. If T is finite and the Mal'cev kernel of φ lies in a locally finite variety \mathcal{V} , then S is also finite.*

Algebraic proofs of Brown's theorem can be found in [22] and [32, Chapter 4]; see also [35]. Mal'cev expansions were considered in [14] (a profinite analogue for pseudovarieties was introduced in [31]).

Definition 4.24 (Mal'cev expansions). Let S be an A -semigroup and let \mathcal{V} be a variety. The *Mal'cev expansion of S by \mathcal{V}* is the largest A -semigroup which maps to S with Mal'cev kernel in \mathcal{V} . More precisely, for each A -morphism $T \rightarrow S$ whose Mal'cev kernel lies in \mathcal{V} , define \sim_T as the congruence on the free semigroup A^+ which produces T , i.e., the congruence which corresponds to the A -morphism $A^+ \rightarrow T$. The Mal'cev expansion of S by \mathcal{V} is then defined as the semigroup A^+/\sim where \sim is the intersection of all of these congruences \sim_T on A^+ . The Mal'cev expansion of S by \mathcal{V} will be denoted $S^\mathcal{V}$. See [6,14] for a more detailed discussion of Mal'cev expansions.

Some properties of Mal'cev expansions are immediate from the definition.

Lemma 4.25. *Let S be an A -semigroup and let \mathcal{V} be a variety. The semigroup $S^\mathcal{V}$ is the largest A -semigroup which maps to S with Mal'cev kernel in \mathcal{V} . More precisely the canonical A -morphism $\eta_S: S^\mathcal{V} \rightarrow S$ has Mal'cev kernel lying in \mathcal{V} and if $\psi: T \rightarrow S$ is another such A -morphism, then there is a (unique) A -morphism $\Psi: S^\mathcal{V} \rightarrow T$ so that*

$$\begin{array}{ccc}
 S^\mathcal{V} & \xrightarrow{\Psi} & T \\
 \searrow \eta_S & & \swarrow \psi \\
 & S &
 \end{array}$$

commutes.

Remark 4.26 (A presentation of $S^\mathcal{V}$). One can alternatively construct $S^\mathcal{V}$ as follows. Let E be a basis of identities for the variety \mathcal{V} . Define an A -semigroup S' with presentation consisting of all relations of the form $u(w_1, \dots, w_n) = v(w_1, \dots, w_n)$ where $u = v$ is an identity from E in n -variables and $w_1, \dots, w_n \in A^+$ all map to the same idempotent of S . By construction, $S' \rightarrow S$ is a well-defined A -morphism with Mal'cev kernel in \mathcal{V} . To verify the universal property of Lemma 4.25, notice that if T is any A -semigroup with $T \rightarrow S$ having Mal'cev kernel in \mathcal{V} and u, v, w_1, \dots, w_n are as above, then since $[w_1]_T, \dots, [w_n]_T$ generate a semigroup in \mathcal{V} they satisfy $[u(w_1, \dots, w_n)]_T = [v(w_1, \dots, w_n)]_T$ and so there is a morphism $S' \rightarrow T$ of A -semigroups. Thus $S' = S^\mathcal{V}$.

Notice that the functoriality of the assignment $S \mapsto S^\mathcal{V}$ is clear from this presentation. Indeed, if $f: S \rightarrow T$ is a homomorphism of A -semigroups and $w_1, \dots, w_n \in A^+$ map to the same idempotent of S , then also w_1, \dots, w_n map to the same idempotent of T . Thus $T^\mathcal{V}$ satisfies all the defining relations of $S^\mathcal{V}$ and so there is a morphism $f^\mathcal{V}: S^\mathcal{V} \rightarrow T^\mathcal{V}$.

One of the main goals of geometric semigroup theory is to determine the structure of $S^\mathcal{V}$ for various locally finite varieties \mathcal{V} . Note, however, that even taking a Mal'cev expansion with respect to the trivial variety is a very non-trivial operation. If S is a finite semigroup and A is the set of all elements in S , then Chris Ash studied, in essence, the Mal'cev expansion $S^{\{1\}}$ with respect to the generating set A in his solution to the Type II conjecture. See [1, 2, 5] for details and [6] for an early version of this approach.

The following theorem establishes the key properties of Mal'cev expansions, namely that they are indeed expansion in the sense of Definition 4.12, and that when the variety \mathcal{V} is locally finite, they preserve finiteness properties.

Theorem 4.27. *For each variety \mathcal{V} , the assignment $S \mapsto S^\mathcal{V}$ defines an expansion on the category \mathbf{Sgp}_A . Moreover, if \mathcal{V} is locally finite, then it preserves finiteness. In particular, it restricts to an expansion on \mathbf{FSgp}_A or \mathbf{FJSgp}_A .*

Proof. Remark 4.26 shows that $S \mapsto S^\mathcal{V}$ is an expansion.

To show the second assertion we need to show that if S is finite, then $S^\mathcal{V}$ is finite, and if S is finite \mathcal{J} -above, then $S^\mathcal{V}$ is finite \mathcal{J} -above. When S is finite, the assertion follows immediately from Brown's theorem (Theorem 4.23). When S is finite \mathcal{J} -above we will use Lemma 4.13. Let I be an ideal in S and let J be the inverse image of I under the A -morphism $S^\mathcal{V} \rightarrow S$. Notice that the Mal'cev kernel of $S^\mathcal{V}/J \rightarrow S/I$ is a subfamily of the Mal'cev kernel of $S^\mathcal{V} \rightarrow S$ with the possible addition of the trivial semigroup, and thus lies in \mathcal{V} . It now follows from Lemma 4.25 that there is an A -morphism from $(S/I)^\mathcal{V}$ to $S^\mathcal{V}/J$ and by Lemma 4.13 the proof is complete. \square

A list of some locally finite varieties is given in Table 2.

The lattice of locally finite varieties is closed under taking joins, subvarieties and semidirect products. The concept of a Mal'cev product can also be extended to form a (nonassociative) product on the set of varieties.

Definition 4.28 (Mal'cev products of varieties). If \mathcal{U} and \mathcal{V} are varieties of semigroups, the *Mal'cev product* of \mathcal{U} and \mathcal{V} (denoted $\mathcal{U} \boxtimes \mathcal{V}$) is the variety generated by all semigroups S admitting a homomorphism $\varphi: S \rightarrow V$ with $V \in \mathcal{V}$ so that the Mal'cev kernel of φ belongs to \mathcal{U} . It follows from Brown's theorem that the Mal'cev product of two locally finite varieties is also a locally finite variety, and, as a result, this table can be extended by taking Mal'cev products. Some care must be exercised with parentheses, since

Notation	Description	Equations
$\{1\}$	trivial	$x = 1$
\mathcal{SL}	semilattices	$xy = yx, x^2 = x$
\mathcal{RZ}	right-zero semigroups	$xy = y$
\mathcal{LZ}	left-zero semigroups	$xy = x$
\mathcal{B}	bands	$x^2 = x$
\mathcal{RB}	rectangular bands	$x^2 = x, xzy = xwy$
\mathcal{LC}	left constants	$xy = xz$
\mathcal{RC}	right constants	$yx = zx$
\mathcal{C}	2-sided constants	$xzy = xwy$
\mathcal{D}_k	k -delay	$x_1 \cdots x_k = x_0 x_1 \cdots x_k$
\mathcal{NULC}	null semigroups	$xy = 0$
\mathcal{N}_k	nilpotent of class k	$x_1 \cdots x_k = 0$
$\mathcal{COM}_{(m,n)}$	commutative semigroups satisfying $x^m = x^{m+n}$	$xy = yx, x^m = x^{m+n}$
$\langle S \rangle$	variety generated by a finite semigroup S	$\text{Eqs}(S)$
$\langle \mathbb{Z}_p \rangle$	vector spaces over \mathbb{Z}_p , p prime	$xy = yx, x^p = 1$

TABLE 2. Locally finite varieties

Mal'cev products are not necessarily associative. The inclusion

$$\mathcal{U} \textcircled{\text{m}} (\mathcal{V} \textcircled{\text{m}} \mathcal{W}) \subseteq (\mathcal{U} \textcircled{\text{m}} \mathcal{V}) \textcircled{\text{m}} \mathcal{W}$$

always holds, but a strict inclusion is possible.

Numerous relations exist between the locally finite varieties listed in the table with respect to the Mal'cev product. For example $\mathcal{B} = \mathcal{B} \textcircled{\text{m}} \mathcal{B}$, as is immediate from the definition. It is a well-known consequence of Green-Rees structure theory that $\mathcal{B} = \mathcal{RB} \textcircled{\text{m}} \mathcal{SL}$. In contrast, if we replace \mathcal{B} by $\bar{\mathcal{B}} = \mathcal{C} \textcircled{\text{m}} \mathcal{B}$, then finite iterated Mal'cev products of $\bar{\mathcal{B}}$, bracketed in the larger way, contain all finite aperiodic semigroups by the two-sided Prime Decomposition Theorem [32, 33].

As an illustration of the power of Lemma 4.25 we prove that the Mal'cev expansion with respect to the variety of rectangular bands applied to the free semilattice over A is the free band over A . Actually, with additional work, one could prove that the Birget-Rhodes expansion of the free semilattice is already equal to free band, but that fact is harder to derive using these techniques.

Lemma 4.29. *Let \mathcal{V} and \mathcal{W} be varieties. Then the free semigroup on A in $\mathcal{V} \textcircled{\text{m}} \mathcal{W}$ is $F_{\mathcal{W}}(A)^{\mathcal{V}}$ where $F_{\mathcal{W}}(A)$ is the free semigroup on A in \mathcal{W} . In particular, the rectangular band expansion of the free semilattice is the free band.*

Proof. Suppose that $\sigma: A \rightarrow T$ is a map with $T \in \mathcal{V} \textcircled{m} \mathcal{W}$. Then we can write T as a quotient $\pi: S \rightarrow T$ where S admits a morphism $\theta: S \rightarrow W$ with the Mal'cev kernel of θ in \mathcal{V} . Let $\tau: A \rightarrow S$ be a map with $\sigma = \pi\tau$. Clearly to extend σ to $F_{\mathcal{W}}(A)^{\mathcal{V}}$ it suffices to extend τ . Without loss of generality we may assume that θ is onto and that $\tau(A)$ generates S . Then there is an A -morphism $W^{\mathcal{V}} \rightarrow S$ by the universal property in Lemma 4.25. But there is an A -morphism $F_{\mathcal{W}}(A) \rightarrow W$ since $W \in \mathcal{W}$, and so by functoriality we have an A -morphism $F_{\mathcal{W}}(A)^{\mathcal{V}} \rightarrow W^{\mathcal{V}} \rightarrow S$, completing the proof. \square

4.5. Rectangular bands. In this subsection we investigate the Mal'cev expansion with respect to the variety of rectangular bands and the closely related varieties of left zero/right zero semigroups. Let S be an A -semigroup and consider the semigroups $S^{\mathcal{R}\mathcal{B}}$, $S^{\mathcal{L}\mathcal{Z}}$, and $S^{\mathcal{R}\mathcal{Z}}$. Since left and right zero semigroups are examples of rectangular bands, it is immediate from the definition that there are maps $S^{\mathcal{R}\mathcal{B}} \rightarrow S^{\mathcal{L}\mathcal{Z}}$ and $S^{\mathcal{R}\mathcal{B}} \rightarrow S^{\mathcal{R}\mathcal{Z}}$. We begin by establishing properties of the projection $S^{\mathcal{R}\mathcal{B}} \rightarrow S$. Recall that if \mathcal{K} is any of Green's relations, then a morphism $\varphi: S \rightarrow T$ is a \mathcal{K}' -map if $s, s' \in S$ regular and $\varphi(s) = \varphi(s')$ implies that $s \mathcal{K} s'$. See [20, Chapter 8] or [32, Chapter 4].

Lemma 4.30. *If S is an A -semigroup then the map $\varphi: S^{\mathcal{R}\mathcal{B}} \rightarrow S$ is one-to-one on each subgroup of $S^{\mathcal{R}\mathcal{B}}$. Moreover, if S is finite \mathcal{J} -above, then the inverse image of each regular \mathcal{J} -class of S contains a single regular \mathcal{J} -class in $S^{\mathcal{R}\mathcal{B}}$. In particular, φ is a \mathcal{J}' -map.*

Proof. Let G be a subgroup of $S^{\mathcal{R}\mathcal{B}}$ and let e be its identity. Since φ restricted to G is a group homomorphism, it is injective on G if and only if $\ker \varphi|_G = \{e\}$. On the other hand, e and $\varphi(e)$ are idempotents, so the inverse image $\varphi^{-1}(\varphi(e))$ is a rectangular band by definition. Since the only idempotent in a group is its identity, we can conclude the restriction is indeed one-to-one. Next, let J be a regular \mathcal{J} -class of S and let x and y be idempotents of $\varphi^{-1}(J)$. Since $\varphi(x)$ and $\varphi(y)$ are \mathcal{J} -equivalent (and φ is onto) there exists u and v in $S^{\mathcal{R}\mathcal{B}}$ such that $\varphi(uxv) = \varphi(y) = e$, where e is an idempotent of J . On the other hand, $\varphi^{-1}(e)$ is a rectangular band by definition, so uxv and y are \mathcal{J} -equivalent in $S^{\mathcal{R}\mathcal{B}}$. Thus $x \geq_{\mathcal{J}} y$ and a similar argument shows $y \geq_{\mathcal{J}} x$. It follows that $\varphi^{-1}(J)$ contains only a single regular \mathcal{J} -class. \square

Remark 4.31. One can similarly show that $S^{\mathcal{L}\mathcal{Z}} \rightarrow S$ is an \mathcal{L}' -map and $S^{\mathcal{R}\mathcal{Z}} \rightarrow S$ is an \mathcal{R}' -map [32, Chapter 4].

Because $S^{\mathcal{R}\mathcal{B}} \rightarrow S$ is a \mathcal{J}' -map which is injective on subgroups, when S is finite the techniques and results from [32, 34] can be applied. In particular, $S^{\mathcal{R}\mathcal{B}}$ and S have the same complexity [13, 32]. In the next lemma, we show that all three expansions are stable under iteration.

Lemma 4.32 (Stable under iteration). *A map $\theta: S \rightarrow T$ has Mal'cev kernel in $\mathcal{L}\mathcal{Z}$ ($\mathcal{R}\mathcal{Z}$) if and only if the inverse image of each left zero (right zero)*

subsemigroup in T is a left zero (right zero) subsemigroup in S . Similarly, θ has Mal'cev kernel in \mathcal{RB} if and only if the inverse image of each rectangular band in T is a rectangular band in S . As a consequence, $S^{\mathcal{LZ}.\mathcal{LZ}} = S^{\mathcal{LZ}}$, $S^{\mathcal{RZ}.\mathcal{RZ}} = S^{\mathcal{RZ}}$, and $S^{\mathcal{RB}.\mathcal{RB}} = S^{\mathcal{RB}}$.

Proof. We just handle the case of rectangular bands, as the other cases are similar (and this result is any event well known). Clearly, it suffices to handle the case of a surjective morphism $\theta: S \rightarrow T$ of bands with T a rectangular band and θ a morphism with Mal'cev kernel in \mathcal{RB} . By Lemma 4.30, S has a single \mathcal{J} -class. It follows that S is a rectangular band. Finally, to see the last assertion, let E denote either \mathcal{LZ} , \mathcal{RZ} , or \mathcal{RB} and consider the composition of maps $S^{E.E} \rightarrow S^E \rightarrow S$. An idempotent in S pulls back to a left zero semigroup/right zero semigroup/rectangular band in S^E which we now know pulls back to a semigroup of the same type in $S^{E.E}$. Thus, by Lemma 4.25 there exists a map $S^E \rightarrow S^{E.E}$. As usual, maps in both directions implies they are isomorphic. \square

One consequence of being stable under the \mathcal{RB} -expansion is that the semigroup remembers the first and last letters of the words representing an element in the following sense.

Lemma 4.33 (Vertex labels). *If S is a non-trivial finite \mathcal{J} -above A -semigroup which is stable under the \mathcal{RB} -expansion and u and v are words in A^+ which represent the same element of S then they have the same first letter and the same last letter. As a consequence, there is a unique edge starting at the initial state of $\text{Str}^S(v)$ and for every vertex v in $\text{Cay}(S, A)$, all of the edges ending at v have identical labels.*

Proof. Let $S \rightarrow \{1\}$ be the collapsing morphism. Viewing $\{1\}$ as an A -generated semigroup via this map, functoriality yields a surjective morphism $S^{\mathcal{RB}} \rightarrow \{1\}^{\mathcal{RB}}$. Since $\{1\}^{\mathcal{RB}}$ is the free rectangular band on A by Lemma 4.29, and hence two words are equal in it if and only if they have the same first and last letter, the result follows. \square

A similar statement and proof shows that $S^{\mathcal{LZ}}$ remembers the first letter and $S^{\mathcal{RZ}}$ remembers the last letter. The \mathcal{LZ} , \mathcal{RZ} and \mathcal{RB} -expansions are also stable under many of the expansions defined in Definition 4.14.

Theorem 4.34 (Stable under other expansions). *For each A -semigroup S , there exist A -morphisms as shown in Figure 4.3. As a consequence, $S^{\mathcal{LZ}}$ is stable under $\wedge \ell 1$ and $\wedge \mathcal{R}$, $S^{\mathcal{RZ}}$ is stable under $\wedge \ell 1^{op}$ and $\wedge \mathcal{L}$, and $S^{\mathcal{RB}}$ is stable under all of the expansions shown.*

Proof. The maps that are not obvious are consequences of well-known properties of the Rhodes, Karnofsky-Rhodes and Birget-Rhodes expansions [6, 14, 31]. For instance, we verify the projection $\eta_S: S^{\wedge \ell 1} \rightarrow S$ has Mal'cev kernel in \mathcal{LZ} . Suppose $v, w \in A^+$ map to an idempotent e in S . Then since $[vw]_S = e^2 = e = [v]_S$, it follows that v and vw use the same transition

$$\begin{array}{ccccc}
S^{\mathcal{LZ}} & \longleftarrow & S^{\mathcal{RB}} & \longrightarrow & S^{\mathcal{RZ}} \\
\downarrow & & \downarrow & & \downarrow \\
S^{\wedge \ell 1} & & & & S^{\wedge \ell 1^{op}} \\
\downarrow & & \downarrow & & \downarrow \\
S^{\wedge \mathcal{R}} & \longleftarrow & S^{BR} & \longrightarrow & S^{\wedge \mathcal{L}}
\end{array}$$

FIGURE 4.3. Relations between various expansions.

edges from I in $\text{Cay}(S, A)$ in the paths they read to e . Thus $vw = v$ in $S^{\wedge \ell 1}$. It follows that $\eta_S^{-1}(e)$ is a left zero semigroup.

By Lemma 4.32 the expansions \mathcal{LZ} , \mathcal{RZ} and \mathcal{RB} are closed under iteration, so the stability properties of $S^{\mathcal{LZ}}$ and $S^{\mathcal{RZ}}$ follow by Lemma 4.18. Similarly, $S^{\mathcal{RB}}$ is stable under all of these with the possible exception of BR . However, once $S^{\mathcal{RB}}$ is stable under $\wedge \mathcal{L}$ and $\wedge \mathcal{R}$, it is also stable under BR and by Lemma 4.18, there is a map $S^{\mathcal{RB}} \rightarrow S^{BR}$. \square

Combining Lemma 4.17 and Lemma 4.32 proves the following.

Corollary 4.35. *If S is a finite \mathcal{J} -above A -semigroup and T is either $S^{\mathcal{LZ}}$ or $S^{\mathcal{RB}}$, then $\text{Str}^T(w) = \text{Cay}^T(w)$ is a trim, deterministic finite state acceptor which depends only on the element w represents, it is linear and it accepts the language of words equivalent to w in T . In addition, every element \mathcal{R} -above $[w]_T$ is \mathcal{R} -equivalent to an initial segment of w . Similar properties hold for the left-handed versions when T is either $S^{\mathcal{RZ}}$ or $S^{\mathcal{RB}}$.*

We thus have the following result, which shows that we can use straight-line automata (and in particular the factors of its base) to investigate Green's relations in $S^{\mathcal{RB}}$. This is a first small step toward converting algebraic information into geometric form.

Corollary 4.36 (Green's relations). *Let S be an A -semigroup, let $T = S^{\mathcal{RB}}$ and let u and v be words in A^+ . Then $[u]_T \geq_{\mathcal{J}} [v]_T$ if and only if u is readable in $\text{Str}^T(v)$. Moreover, $[u]_T \geq_{\mathcal{R}} [v]_T$ if and only if u is readable in $\text{Str}^T(v)$ starting at the initial state if and only if $[u]_T$ is \mathcal{R} -equivalent to an element represented by an initial segment of v , and $[u]_T \geq_{\mathcal{L}} [v]_T$ if and only if u is readable in $\text{Str}^T(v)$ ending at the terminal state if and only if $[u]_T$ is \mathcal{L} -equivalent to an element represented by a final segment of v .*

Another immediate corollary of these results is that loops in straightline automata can be “read backwards” in the following sense.

Corollary 4.37 (Reading backwards). *Let S be an A -semigroup stable under the \mathcal{RB} -expansion and let w and u be words in A^+ . If wu and w represent the same element of S , then w is readable in $\text{Str}^S(w)$ starting at I and u is readable as a loop starting and ending at the vertex $[w]_S$. Moreover, there*

is a partition of $w = w_1w_2$, with $w_i \in A^*$ such that w_2 is \mathcal{L} -equivalent to u . In particular, there exists words v and v' such that $[v'u]_S = [w_2]_S$ and $[vw_2]_S = [u]_S$ and the entire configuration of paths shown in Figure 4.4 can be read in $\text{Str}^S(w)$ starting at I .

Proof. The first statement is clear. By Corollary 4.36, since $[u]_S \geq_{\mathcal{L}} [w]_S$, we can write $w = w_1w_2$ with $[u]_S \mathcal{L} [w_2]_S$. The remaining statements of the corollary are straightforward. \square

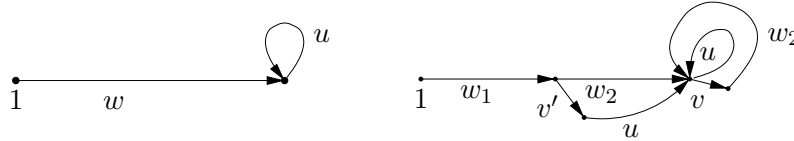


FIGURE 4.4. A standard loop before and after factoring.

Corollary 4.37 takes on added significance once we prove that $S^{\mathcal{RB}}$ can be presented using only relations of the form $uv = u$.

Definition 4.38 (Loop relations). A *loop relation* is one of the form $uv = u$ for words $u, v \in A^*$. An A -semigroup has a *loop presentation* if it has presentation in which each relation is a loop relation. When focusing on the \mathcal{L} -order instead of the \mathcal{R} -order, we consider *reverse loop relations* $uv = v$, which lead to *reverse loop presentations*. Notice that Proposition 3.2 immediately implies that the Cayley graph of a semigroup with a loop presentation is a loop automaton.

Lemma 4.39. For each A -semigroup S , the semigroup $S^{\mathcal{RB}}$ can be given a loop presentation. In particular, $S^{\mathcal{RB}}$ can be defined by adding the relation $uv = u$ whenever the words $u, v \in A^+$ represent the same idempotent e in $S^{\mathcal{RB}}$. The analogous reverse loop presentation is another presentation of $S^{\mathcal{RB}}$.

Proof. This is immediate from Remark 4.26. \square

The fact that $S^{\mathcal{RB}}$ has a loop presentation has important consequences for its straightline automata. In particular, they are all loop automata.

Lemma 4.40. If S is a finite \mathcal{J} -above A -semigroup with a loop presentation, then, for all words $w \in A^+$, $\text{Cay}^S(w)$ is a loop automaton.

Proof. We already observed that $\text{Cay}(S, A)$ is a loop automaton. Clearly a loop presentation of $\text{Cay}^S(w)$ can be obtained by using the loop relations from the presentation of $\text{Cay}(S, A)$ and by adding the dead-end relations $ua = \square$ where $[u]_S \geq_{\mathcal{R}} [w]_S$, $a \in A$ and $[ua]_S \not\leq_{\mathcal{R}} [w]_S$. \square

Actually, more is true. For each specific word w only a finite number of the loop relations are needed in order to build its straightline automaton.

Lemma 4.41. *If T is a finite \mathcal{J} -above A -semigroup with $T = T^{\mathcal{RB}}$, and w is a word in A^+ , then the automaton $\text{Str}^T(w)$ is a loop automaton defined by a finite number of loop relations.*

Proof. By Corollary 4.35, Lemma 4.39 and Lemma 4.40, $\text{Str}^T(w)$ is a loop automaton, but we must prove that it requires only a finite number of loop relations. There is a cofinite ideal I of T so that $\text{Str}^T(w) = \text{Str}^{T/I}(w)$ so we may assume without loss of generality that T is finite.

By Ramsey theory [15], if a word $W \in A^+$ is sufficiently long relative to the size of T , then W can be partitioned into factors so that $W = w_1 u v w_2$ with $w_i \in A^*$ and where u , v and uv all represent the same element of T (which is necessarily an idempotent). In the notation from [15] we need $|W| > [3]_m^2$ where $m = |T|$; call this lower bound n . Indeed, if $W = a_1 \cdots a_n$ with the a_i letters, then consider the complete graph Γ on $\{0, \dots, n\}$. We color the edges of Γ by elements of T . More precisely, if $i < j$ are elements of $\{0, \dots, n\}$, then the edge between i and j is colored by $[a_{i+1} \cdots a_j]_T$. Ramsey's theorem [15] provides a monochromatic triangle. Such a triangle corresponds to $i < j < k$ with

$$[a_{i+1} \cdots a_j]_T = [a_{j+1} \cdots a_k]_T = [a_{i+1} \cdots a_k]_T$$

and so taking $u = a_{i+1} \cdots a_j$, $v = a_{j+1} \cdots a_k$ gives the required partition.

Thus by applying a loop relation of the form $uv = u$, with $|u|, |v| \leq n$, and u and v representing the same idempotent in T , all words are equivalent to words of length at most n . For this finite number of short words, a finite number of loop relations of the form $u_i v_i = u_i$ with u_i, v_i equal to the same idempotent in T are needed to derive whatever further equalities are needed. This gives a *finite* list of defining loop relations for $\text{Str}^T(w)$. \square

Next, we show that $S^{\mathcal{RB}}$ is almost never a monoid, but before doing so, we note that when S is the trivial semigroup generated by $A = \{a\}$, then $S^{\mathcal{RB}}$ is still trivial. This is because the entire semigroup must be a rectangular band and a one-generated rectangular band is trivial. On the other hand, this turns out to be the only situation where $S^{\mathcal{RB}}$ is a monoid.

Lemma 4.42 (Almost never a monoid). *If S is an A -semigroup then $S^{\mathcal{RB}}$ is not a monoid unless $|A| = 1$ and $S = \{1\}$.*

Proof. If $|A| \geq 2$, then the collapsing morphism $S \rightarrow \{1\}$ yields a surjective morphism $S^{\mathcal{RB}} \rightarrow \{1\}^{\mathcal{RB}}$ (where $\{1\}$ is viewed as A -generated). Since $\{1\}^{\mathcal{RB}}$ is the free rectangular band on A , and hence not a monoid, it follows that $S^{\mathcal{RB}}$ is not a monoid. So we may suppose that $|A| = 1$, i.e., S is cyclic. If S is not a monoid, then clearly $S^{\mathcal{RB}}$ is also not a monoid. So we may assume that S is a finite cyclic group of order $n \geq 2$ generated by a . Let T be the semigroup given by the presentation $\langle a \mid a^2 = a^{2+n} \rangle$. Then the canonical morphism $T \rightarrow S$ has trivial Mal'cev kernel. Thus T is a non-monoid homomorphic image of $S^{\mathcal{RB}}$ and so $S^{\mathcal{RB}}$ is not a monoid. \square

As a consequence of Lemma 4.42, our decision to use S^I rather than only adding an identity when one is not already present is irrelevant once S is stable under the \mathcal{RB} -expansion.

4.6. Improving stabilizers. In this short subsection we first consider the effect of the rectangular band expansion on stabilizers, and then we show how an expansion due to Le Saec, Pin and Weil can further simplify their structure.

Definition 4.43 (Stabilizers). Let S be a semigroup. The *right stabilizer* of $s \in S$, is the subsemigroup $S_s = \{t \mid st = s\}$. Similarly, there are *left stabilizers* ${}_sS = \{t \mid ts = s\}$ and *double stabilizers*

$${}_{s'}S_s = {}_{s'}S \cap S_s = \{t \mid ts = s \text{ and } s't = s'\}.$$

The next lemma allows us to equate right stabilizers with loops accepted by Schützenberger graphs, thus turning the study of right stabilizers into an investigation into the properties of various automata.

Lemma 4.44 (Loops in $\text{Sch}^S(s)$). *Let S be an A -semigroup and let s be an element of S . The elements of S represented by words accepted by $\text{Sch}^S(s)$ are precisely the right stabilizers of s , i.e., the elements in S_s .*

Proof. If u is a word representing s and v is a word accepted by $\text{Sch}^S(s)$, then uv is accepted by $\text{Cay}^S(u)$ so that uv and u both represent s , whence $s \cdot [v]_S = [uv]_S = [u]_S = s$, and $[v]_S$ is a right stabilizer of s . Conversely, if t is a right stabilizer of s and u and v represent s and t respectively, then $[uv]_S = [u]_S$ and uv is accepted by $\text{Cay}^S(u)$. Moreover, since $\text{Cay}^S(u)$ is deterministic, v is read as a loop based at $[u]_S$ in the strong component of $[u]_S$, but this is the Schützenberger graph $\text{Sch}^S(s)$ and v is reading an accepted loop. \square

In a left cancellative semigroup, a semigroup S acts on the left of its Cayley graph by injective functions. In the next proposition, we show that if S is stable under the \mathcal{RZ} -expansion, then the left action of S is injective when restricted to Schützenberger graphs. Recall the a partial order is said to be *unambiguous* if $x \leq y, z$ implies y, z are comparable.

Theorem 4.45. *Let S be a finite \mathcal{J} -above A -semigroup stable under the \mathcal{RZ} -expansion. Then the \mathcal{L} -order on S is unambiguous. Moreover, if $r, s, t \in S$ with $r \mathcal{R} t$ and $sr = st$, then $r = t$. Consequently, if $s, t \in S$, then the natural map $\text{Sch}^S(t) \rightarrow \text{Sch}^S(st)$ induced by left multiplication by s is an injective morphism of A -automata.*

Proof. It is conceptually easier to establish the dual result when S is stable under the \mathcal{LZ} -expansion, which we proceed to do. We use that $S = S^{\wedge \ell 1}$. Suppose that r, t have a common lower bound in the \mathcal{R} -order on S . Let $u, v, w, z \in A^+$ satisfy $[u]_S = r$, $[v]_S = t$ and $[uw]_S = [vz]_S$. We prove that r and t are comparable in the \mathcal{R} -order. By hypothesis, the paths read by uw and vz from I in $\text{Cay}(S, A)$ use the same set $\{e_1, \dots, e_m\}$ of transition

edges. Since $\text{Cay}(S, A)$ is linear, we may assume the indices are chosen so that $e_1 \succ e_2 \cdots \succ e_m$. Hence the set of transition edges used by u in its run from I is of the form $\{e_1, \dots, e_k\}$ for some $k \leq m$ and similarly the set used by v is of the form $\{e_1, \dots, e_r\}$ for some $r \leq m$. Without loss of generality we may assume that $k \leq r$. It then follows that $[u]_S \geq_{\mathcal{R}} [v]_S$.

Next suppose that $r, t \in S$ satisfy $r \mathcal{L} t$ and $rs = ts$. Then since $rs = ts$ is a common lower bound to r and t in the \mathcal{R} -order, they are comparable by the previous paragraph. It follows that $[u]_S \mathcal{H} [v]_S$ in S . Choose $u, v, w \in A^+$ such that $[u]_S = r$, $[v]_S = t$ and $[w]_S = s$. There are two cases. Suppose first $[uw]_S \mathcal{R} [u]_S$. Then since $[uw]_S = [vw]_S$, Green's Lemma implies $[u]_S = [v]_S$, as required. Next suppose that $[u]_S >_{\mathcal{R}} [uw]_S$. As $[uw]_S = [vw]_S$, both uw and vw use the same transition edge e when exiting the Schützenberger graph of $[u]_S$. Also note that if w' is an initial segment of w , then $[uw']_S \mathcal{L} [vw']_S$ and hence

$$[uw']_S \mathcal{R} [u]_S \iff [uw']_S \mathcal{J} [u]_S \iff [vw']_S \mathcal{J} [u]_S \iff [vw']_S \mathcal{R} [u]_S$$

as $[v]_S \mathcal{H} [u]_S$. It now follows that if w' is the longest initial segment of w so that $[uw']_S \mathcal{R} [u]_S$, then it is also the longest initial segment so that $[vw']_S \mathcal{R} [u]_S$ and moreover $[uw']_S = [vw']_S$ is the initial vertex of the transition edge e (and hence belongs to the Schützenberger graph of $[u]_S$). Then again Green's Lemma implies $[u]_S = [v]_S$. \square

As a corollary, we obtain a structural result due to the second author [28].

Theorem 4.46 (Right stabilizers). *If S is a finite \mathcal{J} -above A -semigroup stable under the \mathcal{RZ} -expansion, then each right stabilizer S_s is finite \mathcal{R} -trivial and an \mathcal{L} -chain in S , but not necessarily within the subsemigroup S_s .*

Proof. Since each element of S_s is \mathcal{L} -above s , it is finite by assumption on S and is an \mathcal{L} -chain in S by unambiguity of the \mathcal{L} -order (Theorem 4.45). Moreover, if r, t are right stabilizers of s with $r \mathcal{R} t$, then $sr = s = st$ implies $r = t$ by Theorem 4.45. Thus ${}_s S$ is \mathcal{R} -trivial, completing the proof. \square

If $S = S^{\mathcal{RB}}$, then S has unambiguous \mathcal{L} - and \mathcal{R} -orders, a property already achieved by S^{BR} [4].

Corollary 4.47. *If S is a finite \mathcal{J} -above A -semigroup stable under the \mathcal{RB} -expansion, then the \mathcal{L} - and \mathcal{R} -orders on S are unambiguous.*

Although it is not needed later, we record a two-sided version of Theorem 4.46 (also from [28]) which is available whenever S is stable under the \mathcal{RB} -expansion.

Theorem 4.48 (Double stabilizers). *If S is a finite \mathcal{J} -above A -semigroup stable under the \mathcal{RB} -expansion, then each double stabilizer $U = {}_s S_{s'}$ is finite \mathcal{J} -trivial and an \mathcal{H} -chain in S but not necessarily within U .*

Proof. Since S is stable under \mathcal{RB} , it is also stable under \mathcal{LZ} and \mathcal{RZ} (Theorem 4.34). Furthermore, ${}_sS_{s'}$ is a subsemigroup of ${}_{ss'}S_{ss'}$ so without loss of generality we can take $s = s'$ and apply Theorem 4.46 and its dual. \square

Remark 4.49 (Fixing stabilizers). If S is a finite A -semigroup stable under the \mathcal{RZ} -expansion and the right stabilizers are not \mathcal{L} -chains within themselves, then one way to remedy this is by repeatedly applying the Henckell-Schützenberger expansion to S . In fact, Henckell, in his thesis, rediscovered the Schützenberger product in order to prove this result. The precise statement goes as follows. If S is a finite A -semigroup, then the Henckell-Schützenberger expansion $S^{\wedge k}$ of S with $k = 2^{|S|}$ has the property that right stabilizers of $S^{\wedge k}$ mapped into S are \mathcal{L} -chains within themselves. This was exploited by the second and third authors to show that stabilizers in certain relatively free profinite semigroups are \mathcal{L} -chains within themselves. See [6, 17] for details.

Another way to improve the properties of the right stabilizers is to use the Mal'cev expansion with respect to the locally finite variety $\langle \mathbb{Z}_p \rangle$, first studied by Le Saec, Pin and Weil in [22]. See also [14].

Theorem 4.50 (Idempotent stabilizers). *Let S be a finite A -semigroup stable under the \mathcal{RZ} -expansion, let p be a prime, and let $T = S^{\langle \mathbb{Z}_p \rangle}$. If p is sufficiently large relative to the size of S , then each right stabilizer T_t is an \mathcal{R} -trivial band. In other words, each subsemigroup T_t satisfies the identities $yx = xy$ and $x^2 = x$.*

Combining Theorem 4.50 with Theorem 4.46, we have the following, the major improvements being that all of the right stabilizers are now idempotents, the stabilizers form an \mathcal{L} -chain within themselves, and they satisfy an additional equation which is useful later in the article.

Theorem 4.51 (Improved stabilizers). *If S is a finite A -semigroup, p is a sufficiently large prime relative to the size of $S^{\mathcal{RB}}$ and $T = S^{\mathcal{RB}.\langle \mathbb{Z}_p \rangle.\mathcal{RB}}$, then each right stabilizer subsemigroup T_t is an \mathcal{R} -trivial (also called left regular) band, i.e., it satisfies the identities $yx = xy$ and $x^2 = x$. Moreover, T_t forms an \mathcal{L} -chain within itself so that for any two elements $r, s \in T_t$, one is \mathcal{L} -above the other, say $r \geq_{\mathcal{L}} s$, where $\geq_{\mathcal{L}}$ now represents the \mathcal{L} -order in T_t rather than T . In addition, whenever r and s are right t stabilizers and $r \geq_{\mathcal{L}} s$, the equation $sr = s$ also holds. As a consequence, the \mathcal{L} -classes of T_t are left zero semigroups.*

Proof. Since $S^{\mathcal{RB}}$ is stable under the \mathcal{RZ} expansion, the right stabilizer subsemigroups of $S^{\mathcal{RB}.\langle \mathbb{Z}_p \rangle}$ are \mathcal{R} -trivial bands. Let $T' = S^{\mathcal{RB}.\langle \mathbb{Z}_p \rangle}$ and let t' denote the image of $t \in T$ under the natural map $T \rightarrow T'$. Once it is noticed that T_t is contained in the inverse image of $T'_{t'}$ under this map and that the Mal'cev kernel of this map lies in \mathcal{RB} by definition, it is clear that T_t consists solely of idempotents. Moreover, it is \mathcal{R} -trivial and an \mathcal{L} -chain in T by Theorem 4.46. Since the \mathcal{L} -relation restricted to idempotents is independent of the ambient semigroup, this completes the proof. \square

5. THE MCCAMMOND EXPANSION REVISITED

Let us investigate what happens when we apply the McCammond expansion to straightline automata and Cayley graphs.

Theorem 5.1. *Let (\mathcal{A}, I) be a pointed A -automaton. Then, for $w \in A^+$, the pointed A -automaton $(\text{Str}^{\mathcal{A}^{M^c}}(w), I)$ has the unique simple path property and is linear.*

Proof. This follows immediately from Proposition 2.40. \square

Example 5.2 (Expanding straightline automata). Suppose that S is an A -semigroup and w is a word in A^+ . If $\text{Str}^S(w)$ is the automaton on the left in Figure 5.1 (with the edge labels suppressed) and w is the path which passes through vertices $1, p, q, r, s, t$, and u in that order, then $\text{Str}^{(\text{Str}^S(w))^{M^c}}(w)$ is the automaton on the right, where we have simplified the vertex labels as well as leaving out the edge labels. In particular, the labels on the right simply indicate the vertex in $\text{Str}^S(w)$ to which it is sent under the projection map. Similarly, if w' is the path which passes through vertices $1, p, q, r', s, t$, and u in that order, then $\text{Str}^S(w') = \text{Str}^S(w)$, but $\text{Str}^{(\text{Str}^S(w'))^{M^c}}(w') \neq \text{Str}^{(\text{Str}^S(w))^{M^c}}(w)$. Specifically, the automaton $\text{Str}^{(\text{Str}^S(w))^{M^c}}(w)$ would be altered by having the edge connecting t to u starting instead at the other vertex labeled t to obtain $\text{Str}^{(\text{Str}^S(w'))^{M^c}}(w')$.

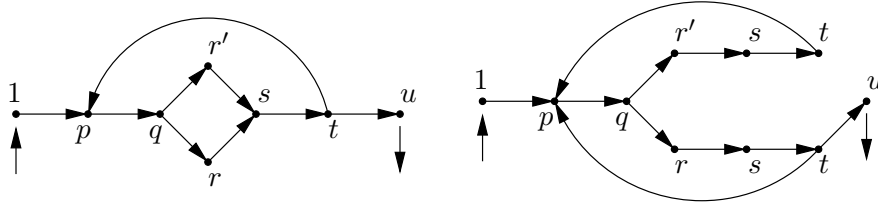


FIGURE 5.1. Expansion of a straightline automaton.

Definition 5.3 (The M^c -expansion of a semigroup). If S is a A -semigroup, then S^{M^c} will denote the transition semigroup of $\text{Cay}(S, A)^{M^c}$. This is an object expansion on A -generated semigroups which preserves finiteness. Note that $\text{Cay}(S^{M^c}, A)$ is not $\text{Cay}(S, A)^{M^c}$. In general, we will never be interested in $\text{Cay}(S^{M^c}, A)$ but rather in the action of S^{M^c} on $\text{Cay}(S, A)^{M^c}$.

Example 5.4 (Not a Cayley graph). Add example that $\text{Cay}(S^{M^c}, A)$ is not $\text{Cay}(S, A)^{M^c}$.

5.1. Properties of the M^c -expansion. In this section we prove that for each A -semigroup S , the M^c -expansion of S has Mal'cev kernel contained in a locally finite variety of aperiodic semigroups and so in particular preserves being finite \mathcal{J} -above (Theorem 5.12). The key step will be to show that the M^c -expansion of the right zero A -semigroup is a finite band.

Definition 5.5 (Right zero semigroup). For each finite set A , let A^r denote the right zero semigroup on A (i.e., $a \cdot b = b$ for all $a, b \in A$) and let $\Gamma_{A^r} = \text{Cay}(A^r, A)$. Recall that by convention $\text{Cay}(S, A)$ is the Cayley graph for S^I rather than S itself, so that Γ_{A^r} has $|A| + 1$ vertices. If we use 1 to denote the root vertex of Γ_V and define $A^1 = A \cup \{1\}$, then the edges of Γ_{A^r} can be described as follows. For each $a \in A^1$ and for each $b \in A$ there is an edge labeled b from a to b . Viewed as a function, $\cdot b$ is the constant function from A^1 to A^1 sending everything to b . See Figure 5.2.

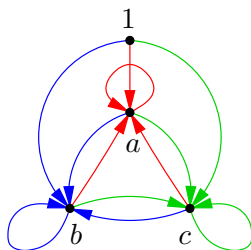


FIGURE 5.2. The automaton $\Gamma_{A^r} = \text{Cay}(A^r, A)$ when $A = \{a, b, c\}$.

The simple paths starting at 1 in Γ_{A^r} have a particularly easy description: a word $w \in A^*$ can be read as a simple path starting at 1 if and only if no letter occurs more than once in w . We call these words *distinct letter words*. Thus the vertices in $(\Gamma_{A^r})^{M^c}$ are in one-to-one correspondence with the distinct letter words.

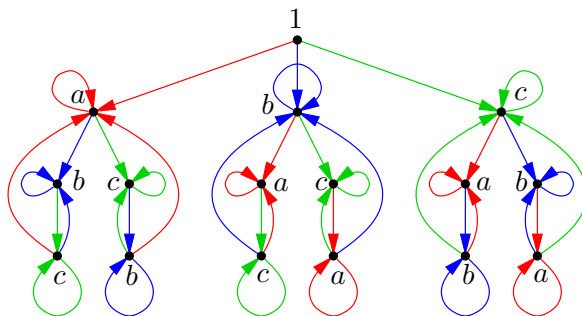


FIGURE 5.3. The automaton $(\Gamma_{A^r})^{M^c}$ when $A = \{a, b, c\}$.

Definition 5.6 (Reduction). Every word $w \in A^*$ can be read in $(\Gamma_{A^r})^{M^c}$ starting at 1 and ending at some vertex, say v . By construction there is a unique simple path in this graph from 1 to v . The word read by this simple path is called the *reduction of w* and denoted $\text{Red}(w)$.

For example, if $A = \{a, b, c, d, e\}$ then $\text{Red}(aba) = a$, $\text{Red}(abc) = abc$ and $\text{Red}(abcdce) = abce$. Here are two easy observations about reductions.

Lemma 5.7 (Recurrence). *Suppose $w \in A^+$ is a word with first letter a . If a recurs in w , that is, $w = auav$ for possibly empty words $u, v \in A^*$, then $\text{Red}(w) = \text{Red}(av)$.*

Proof. Reading one letter at a time it is easy to see that the path corresponding to w only passes through vertices whose label starts with a . Moreover, when a recurs in w , the path w returns to the vertex labeled a . Thus the paths corresponding to w and to av end at the same vertex and hence have the same reduction. \square

The most efficient use of Lemma 5.7 would, of course, focus on the last occurrence of a in w so that a does not occur in v .

Lemma 5.8 (No recurrence). *Suppose $w \in A^+$ is a word with first letter a . If a does not recur in w , then $w = av$ for some $v \in (A \setminus \{a\})^*$ and $\text{Red}(w) = a\text{Red}(v)$.*

Proof. Let B denote $A \setminus \{a\}$ and notice that there is an label-preserving embedding of $(\Gamma_{B^r})^{M^c}$ into $(\Gamma_{A^r})^{M^c}$ which sends the root in the domain to the vertex labeled a in the range. The result is an immediate consequence of this observation. \square

Example 5.9. Using Lemma 5.7 and Lemma 5.8 it is easy to calculate the reduction of any word. For example, if $w = abacadabdbcccebgfdf$ then

$$\begin{aligned}
\text{Red}(abacadabdbcccebgfdf) &= \text{Red}(abdbcccebgfdf) && \text{Lemma 5.7} \\
&= a\text{Red}(bdbcccebgfdf) && \text{Lemma 5.8} \\
&= a\text{Red}(bgfdf) && \text{Lemma 5.7} \\
&= ab\text{Red}(gfdf) && \text{Lemma 5.8} \\
&= abg\text{Red}(fdf) && \text{Lemma 5.8} \\
&= abg\text{Red}(f) && \text{Lemma 5.7} \\
&= abgf && \text{Lemma 5.8}
\end{aligned}$$

We can now show that $(A^r)^{M^c}$ is a band.

Lemma 5.10 (Band). *If A is a finite set, then the M^c -expansion of the right zero semigroup A^r is a finite band.*

Proof. Note that it suffices to check that for each distinct letter word $w \in A^*$ and for each word $u \in A^+$, $\text{Red}(wu) = \text{Red}(wuu)$ since this shows that the functions $\cdot u$ and $\cdot uu$ agrees on each vertex of $(\Gamma_{A^r})^{M^c}$. If w is non-empty and the first letter of w does not occur in u , then applying Lemma 5.8 to both wu and wuu reduces the problem to a similar problem with a shorter word w . Continuing in this way, we may assume that either w is the empty word or that the first letter of w occurs in u . When w is empty, a single application of Lemma 5.7 shows that $\text{Red}(uu) = \text{Red}(u)$ and we are done. Similarly, when w is not empty and the first letter of w , say a , occurs in u then applying Lemma 5.7 to the last occurrence of a in wu and in wuu produces the same result in each case. Thus they have the same reduction, which completes the proof. \square

The following example shows that, in contrast with the Cayley graphs of right zero semigroups, the Mal'cev kernel of $S(\mathcal{A}^{M^c}) \rightarrow S(\mathcal{A})$, does not, in general, consist of bands.

Example 5.11 (Not always a band). Let \mathcal{A} denote the A -automaton shown on the left-hand side of Figure 5.4 and root the graph at the vertex labeled 1. It is easy to check that abc and $(abc)^2$ act the same on each vertex. In particular, they both send 1 and r to r and they both fail when starting at p or q . Thus abc is an idempotent in $S(\mathcal{A})$. On the other hand, in \mathcal{A}^{M^c} , shown on the right-hand side of Figure 5.4, abc and $(abc)^2$ act differently since $1 \cdot abc = r'$ and $1 \cdot (abc)^2 = r$. Thus abc no longer represents an idempotent in $S(\mathcal{A}^{M^c})$.

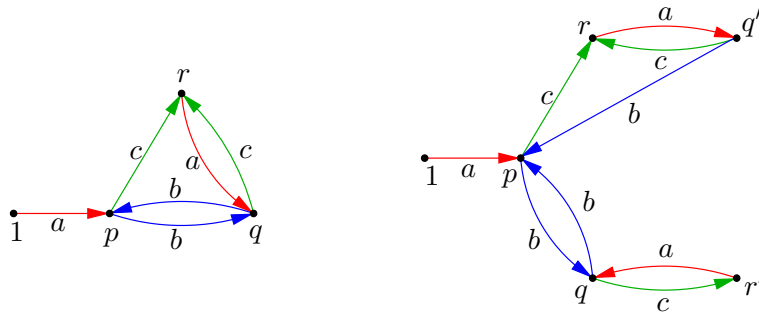


FIGURE 5.4. An \mathcal{A} -automaton \mathcal{A} and its expansion \mathcal{A}^{M^c} .

Despite the fact that the Mal'cev kernel does not always lie in \mathcal{B} , it is quite restricted. Since we only intend to apply this expansion to A -semigroups stable under the \mathcal{RB} -expansion, nothing is lost if we restrict our attention to automata in which the label on an edge is a function of its terminal vertex (Lemma 4.33).

Theorem 5.12 (Bounding the Mal'cev kernel). *If $\mathcal{A} = (V, A)$ is a pointed A -automaton with root I in which the label on an edge is a function of its terminal vertex, then there is a finite band B acting faithfully on a set X such that $S(\mathcal{A}^{M^c})$ divides $(X, B) \wr (V, S(\mathcal{A}))$ where \wr denotes the wreath product of partial transformation semigroups.*

Proof. Let V denote the vertex set of \mathcal{A} and let B be the M^c -expansion of the V -semigroup V^r . By Lemma 5.10, B is a finite band. Put $X = \text{Simple}(V^r)$. At this point, we can define a function $f: \text{Simple}(\mathcal{A}) \rightarrow X \times V$ as follows. Associate to each word $u = a_1 \cdots a_n$ of $\text{Simple}(\mathcal{A})$ the pair $f(u) = ((I, Ia_1, \dots, Ia_1 \cdots a_{n-1}), Iu)$ where elements of V^* are written as strings. The map f is injective precisely because the edge labels are determined by their terminal vertices, so that the word u can be completely recovered from the sequence of vertices it visits. To each $a \in A$, we associate the pair $\hat{a} = (g_a, [a]_{S(\mathcal{A})}) \in (X, B) \wr (V, S(\mathcal{A}))$ where $g_a: Q \rightarrow B$ is given by

$qg_a = \cdot(qa)$ if qa is defined and is arbitrary otherwise. If $u \in \text{Simple}(\mathcal{A})$ is as above and $u \cdot a$ is defined in \mathcal{A}^{M^c} , then there are two cases. If ua is simple, then

$$\begin{aligned} f(u)\widehat{a} &= ((I, Ia_1, \dots, Ia_1 \cdots a_{n-1}), Iu)\widehat{a} \\ &= ((I, Ia_1, \dots, Ia_1 \cdots a_{n-1}, Iu), Iua) = f(u \cdot a). \end{aligned}$$

Otherwise, $u \cdot a = I \cdot a_1 \cdots a_j$ for the unique $1 \leq j \leq n-1$ so that one has $Iua = I \cdot a_1 \cdots a_j$. But then $\text{Red}(I, Ia_1, \dots, Ia_1 \cdots a_{n-1}, Iu) = (I, Ia_1, \dots, Ia_1 \cdots a_j)$. Thus

$$\begin{aligned} f(u)\widehat{a} &= ((I, Ia_1, \dots, Ia_1 \cdots a_{n-1}), Iu)\widehat{a} \\ &= (\text{Red}(I, Ia_1, \dots, Ia_1 \cdots a_{n-1}, Iu), Iua) \\ &= ((I, Ia_1, \dots, Ia_1 \cdots a_j), Iua) = f(u \cdot a). \end{aligned}$$

It now follows that $S(\mathcal{A}^{M^c}) \prec (X, B) \wr (V, S(\mathcal{A}))$ (c.f. [13]). \square

Recall that if \mathcal{V} and \mathcal{W} are varieties, then the semidirect products $\mathcal{V} * \mathcal{W}$ is contained in the Mal'cev product $(\mathcal{L}\mathcal{C} \textcircled{\text{m}} \mathcal{V}) \textcircled{\text{m}} \mathcal{W}$ where we recall that $\mathcal{L}\mathcal{C}$ is the locally finite variety defined by the identity $xy = xz$ [32, Chapter 2]. Thus by Theorem 5.12, the projection $S^{M^c} \rightarrow S$ has a Mal'cev kernel which lies in the locally finite aperiodic variety $\mathcal{L}\mathcal{C} \textcircled{\text{m}} \mathcal{B}$ if S is stable under the $\mathcal{R}\mathcal{B}$ -expansion. Finally, Lemma 4.25 and Lemma 4.33 complete the proof of the following corollary of Theorem 5.12.

Corollary 5.13 (Bounding S^{M^c}). *If S is an A -semigroup stable under the $\mathcal{R}\mathcal{B}$ -expansion, then the Mal'cev kernel of $S^{M^c} \rightarrow S$ lies in the locally finite aperiodic variety $\mathcal{L}\mathcal{C} \textcircled{\text{m}} \mathcal{B}$. As a result, there is an A -morphism $S^{\mathcal{L}\mathcal{C} \textcircled{\text{m}} \mathcal{B}} \rightarrow S^{M^c}$.*

Finally, we note that even though the inverse images of idempotents is not always a band (Example 5.11), the torsion in the Mal'cev kernel is quite controlled.

Lemma 5.14 (Bounding torsion). *If S is an A -semigroup stable under the $\mathcal{R}\mathcal{B}$ -expansion, then the semigroups in the Mal'cev kernel of $S^{M^c} \rightarrow S$ satisfy the identity $x^2 = x^3$. As a result, if S satisfies the identity $x^m = x^{m+n}$ for constants m and n , then S^{M^c} satisfies the identity $x^{m+1} = x^{(m+1)+n}$ for the same constants m and n .*

Proof. It suffices to observe that the variety $\mathcal{L}\mathcal{C} \textcircled{\text{m}} \mathcal{B}$ clearly satisfies the identity $x^2 = x^3$ since in any band x and x^2 map to the same idempotent and so by definition of $\mathcal{L}\mathcal{C}$ we have $xx = xx^2$. \square

6. ALGEBRAIC RANK FUNCTION

In this section we begin by defining an algebraic rank function. For Kleene expressions the rank is the nested star-height, for loop automata it is the maximum number of loops within loops. In the case of the Burnside semigroups, $\mathcal{B}(m, n)$ for $m \geq 6$ and $n \geq 1$, the first author proved that the

automata, Kleene and algebraic definition all agree for $\text{Str}(w)$ in [24]. He also proved that these semigroups are finite \mathcal{J} -above. One main idea of geometric semigroup theory is to run this procedure backwards. That is, to start with a finite \mathcal{J} -above semigroup, define the rank algebraically and then prove it is the automata rank in many cases. This also has strong ties with the holonomy theorem for semigroups [29].

Definition 6.1 (Algebraic rank function). Let S be a finite \mathcal{J} -above A -semigroup and let s be one of its element. Since S is torsion, there is a unique idempotent in the cyclic subsemigroup generated by s which we denote s^ω . Moreover, since S is finite \mathcal{J} -above, there are only a finite number of idempotents in S which are \mathcal{J} -above s^ω . We define the *algebraic rank of s* to be the length of the longest strictly increasing \mathcal{J} -chain of idempotents starting at s^ω . This defines a map $\text{Rank}_S: S \rightarrow \{0, 1, 2, \dots\} = \mathbb{N}$. More specifically,

$$\text{Rank}_S(s) = \max\{r \mid s^\omega = e_0 <_{\mathcal{J}} e_1 <_{\mathcal{J}} \dots <_{\mathcal{J}} e_r\}$$

where all e_i are idempotents in S .

When defining a geometric rank function on $\text{Cay}(S, A)^{M^c}$, we will always assume that the ordering on the loops at a vertex refines the algebraic rank of the images in S of the words labeling the loops.

The following lemma records some elementary properties of the algebraic rank function.

Lemma 6.2 (Elementary properties). *Let S be a finite \mathcal{J} -above A -semigroup and let $\text{Rank}_S: S \rightarrow \mathbb{N}$ denote its algebraic rank function. If s and t are elements of S and e and f are idempotents in S , then:*

- $\text{Rank}_S(s) = \text{Rank}_S(s^k)$ for all $k \geq 1$;
- $\text{Rank}_S(e) \geq \text{Rank}_S(f)$ whenever e is \mathcal{J} -above f ;
- $\text{Rank}_S(e) = \text{Rank}_S(f)$ whenever e and f are \mathcal{J} -equivalent;
- $\text{Rank}_S(e) = \text{Rank}_S(f)$ and $e \geq_{\mathcal{J}} f$ implies $e =_{\mathcal{J}} f$;
- $\text{Rank}_S(st) = \text{Rank}_S(ts)$ for all $s, t \in S$.

Proof. The only part which is not immediate is the last one, but this follows from the first statement once it is noticed that $(st)^\omega$ is always \mathcal{J} -equivalent to $(ts)^\omega$. □

The next lemma studies the effect of \mathcal{J}' -maps on the rank.

Lemma 6.3 (Rank and \mathcal{J}' -maps). *If $S \rightarrow T$ is an onto \mathcal{J}' -map between finite semigroups, then $\text{Rank}_S(s) = \text{Rank}_T(\varphi(s))$ for all $s \in S$.*

Proof. By assumption the map φ has the property that φ^{-1} of a regular \mathcal{J} -class consists of one regular \mathcal{J} -class plus maybe some null elements above it. Hence the restriction of the \mathcal{J} -order to the regular \mathcal{J} -classes results in the same poset for both S and T . The fact that $\text{Rank}_S = \text{Rank}_T$ now follows immediately. □

On the other hand, the rank function can fall an arbitrarily large amount under other types of maximal proper surmorphisms [32–34] that are not \mathcal{J}' -maps.

Remark 6.4. For example, consider the semilattice $S = \{0, 1, \dots, n\}$ with max. Then $\text{Rank}_S(i) = i$. Now put $S' = S \cup \{n'\}$ where $n'j = jn' = n$ for $0 \leq j \leq n$ and $n'n' = n'$; so S' is obtained from S by inflating n to $\{n, n'\}$. Then $\text{Rank}_{S'}(n') = 0$, but its image under the natural map $S' \rightarrow S$ that identifies n with n' has rank n . This map is a maximal proper surmorphism of class $III_{R>R}$ in the terminology of [34] and the Mal'cev kernel of the surmorphism is in \mathcal{SL} .

Lemma 6.5 (Lifting Lemma). *Let S be finite \mathcal{J} -above and let $\varphi: S \rightarrow T$ be an onto map. For each $t \in T$ there exists an $s \in S$ such that $\varphi(s^\omega) = t^\omega$ and $\text{Rank}_S(s) \geq \text{Rank}_T(t)$.*

Proof. Suppose $t^\omega = e_0 <_{\mathcal{J}} e_1 <_{\mathcal{J}} \dots <_{\mathcal{J}} e_r$, where $<_{\mathcal{J}}$ denotes the \mathcal{J} -order in T . First choose an idempotent f_r in S which maps to e_r , then in Sf_rS choose an idempotent f_{r-1} which maps to e_{r-1} , etc. This yields idempotents $f_0 < \dots < f_r$ in S with $\varphi(f_i) = e_i$. Setting $s = s^\omega = f_0$ establishes the result. \square

Recall that idempotents in a semigroup are partially ordered by putting $e \leq f$ if $ef = e = fe$, or equivalently, $e \leq_{\mathcal{H}} f$.

Lemma 6.6 (Alternate definition). *Let S be a finite \mathcal{J} -above A -semigroup. The function $f: S \rightarrow \mathbb{N}$ defined by*

$$f(t) = \max\{r \mid t^\omega \mathcal{J} e_0 < e_1 < \dots < e_r\},$$

where each e_i is an idempotent, is the same as the rank function $\text{Rank}_S(t)$.

Proof. By an elementary lemma from [27], if J and J' are regular \mathcal{J} -classes in a finite semigroup with $J > J'$, then for each idempotent $e \in J$, there exists an idempotent $f \in J'$ so that $e > f$. Indeed, if $e' = uev \in J'$ is an idempotent set $f = eve'ue$. Then $f^2 = evuevueevuev = eve'e'e'ue = eve'ue = f$ and $ufv = ueve'uev = e'e'e' = e'$; so $f \in J'$ is an idempotent with $f < e$. This plus the fact that \mathcal{J} -equivalent idempotents have the same rank proves the result. \square

The following corollary is immediate.

Corollary 6.7 (Subsemigroups). *If S is finite \mathcal{J} -above and T is a subsemigroup of S , then T is finite \mathcal{J} -above and, for all $t \in T$, the inequality $\text{Rank}_T(t) \leq \text{Rank}_S(t)$ holds.*

The importance of the algebraic rank function is as follows. Suppose that S is an A -semigroup and let $\eta: \text{Cay}(S, A)^{M^c} \rightarrow \text{Cay}(S, A)$ be the corresponding simple directed cover. Let v be a vertex of $\text{Cay}(S, A)^{M^c}$. Then $[w]_S$ stabilizes $\eta(v)$. Often S will already have been expanded to have idempotent stabilizers. Thus the loops at v , when mapped into S , will be an

\mathcal{L} -chain of idempotents and we want to choose a geometric rank function accordingly. More formally, we shall always choose a geometric rank function r for $\text{Cay}(S, A)^{M^c}$ so that the following occurs. If e, f are bold arrows ending at v and y, z are the labels of $\text{lp}(e), \text{lp}(f)$, respectively, then we shall assume that if $\text{Rank}_S(y) < \text{Rank}_S(z)$ then $r(e) < r(f)$. In other words, the order on bold arrows at v given by the geometric rank function will be a topological sorting of the algebraic rank of the labels of the loops (taken in S).

REFERENCES

- [1] C. J. Ash. Finite semigroups with commuting idempotents. *J. Austral. Math. Soc. Ser. A*, 43(1):81–90, 1987.
- [2] C. J. Ash. Inevitable graphs: a proof of the type II conjecture and some related decision procedures. *Internat. J. Algebra Comput.*, 1(1):127–146, 1991.
- [3] J. Berstel, D. Perrin, and C. Reutenauer. *Codes and automata*, volume 129 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2010.
- [4] J.-C. Birget. Iteration of expansions—unambiguous semigroups. *J. Pure Appl. Algebra*, 34(1):1–55, 1984.
- [5] J.-C. Birget, S. Margolis, and J. Rhodes. Semigroups whose idempotents form a subsemigroup. *Bull. Austral. Math. Soc.*, 41(2):161–184, 1990.
- [6] J.-C. Birget and J. Rhodes. Almost finite expansions of arbitrary semigroups. *J. Pure Appl. Algebra*, 32(3):239–287, 1984.
- [7] R. V. Book and F. Otto. *String-rewriting systems*. Texts and Monographs in Computer Science. Springer-Verlag, New York, 1993.
- [8] T. C. Brown. An interesting combinatorial method in the theory of locally finite semigroups. *Pacific J. Math.*, 36:285–289, 1971.
- [9] S. Burris and H. P. Sankappanavar. *A course in universal algebra*, volume 78 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1981.
- [10] A. de Luca and S. Varricchio. On noncounting regular classes. *Theoret. Comput. Sci.*, 100(1):67–104, 1992.
- [11] A. P. do Lago. On the Burnside semigroups $x^n = x^{n+m}$. *Internat. J. Algebra Comput.*, 6(2):179–227, 1996.
- [12] S. Eilenberg. *Automata, languages, and machines. Vol. A*. Academic Press, New York, 1974. Pure and Applied Mathematics, Vol. 58.
- [13] S. Eilenberg. *Automata, languages, and machines. Vol. B*. Academic Press, New York, 1976. With two chapters (“Depth decomposition theorem” and “Complexity of semigroups and morphisms”) by Bret Tilson, Pure and Applied Mathematics, Vol. 59.
- [14] G. Z. Elston. Semigroup expansions using the derived category, kernel, and Malcev products. *J. Pure Appl. Algebra*, 136(3):231–265, 1999.
- [15] R. L. Graham, B. L. Rothschild, and J. H. Spencer. *Ramsey theory*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., New York, second edition, 1990. A Wiley-Interscience Publication.
- [16] V. S. Guba. The word problem for the relatively free semigroup satisfying $T^m = T^{m+n}$ with $m \geq 3$. *Internat. J. Algebra Comput.*, 3(3):335–347, 1993.
- [17] K. Henckell. Stable pairs. *Internat. J. Algebra Comput.*, to appear.
- [18] K. Henckell, S. W. Margolis, J.-E. Pin, and J. Rhodes. Ash’s type II theorem, profinite topology and Mal’cev products. I. *Internat. J. Algebra Comput.*, 1(4):411–436, 1991.
- [19] O. G. Kharlampovich and M. V. Sapir. Algorithmic problems in varieties. *Internat. J. Algebra Comput.*, 5(4-5):379–602, 1995.

- [20] K. Krohn, J. Rhodes, and B. Tilson. *Algebraic theory of machines, languages, and semigroups*. Edited by Michael A. Arbib. With a major contribution by Kenneth Krohn and John L. Rhodes. Academic Press, New York, 1968. Chapters 1, 5–9.
- [21] M. V. Lawson. *Finite automata*. Chapman & Hall/CRC, Boca Raton, FL, 2004.
- [22] B. Le Saëc, J.-E. Pin, and P. Weil. Semigroups with idempotent stabilizers and applications to automata theory. *Internat. J. Algebra Comput.*, 1(3):291–314, 1991.
- [23] S. Mac Lane. *Categories for the working mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1998.
- [24] J. McCammond. The solution to the word problem for the relatively free semigroups satisfying $T^a = T^{a+b}$ with $a \geq 6$. *Internat. J. Algebra Comput.*, 1(1):1–32, 1991.
- [25] J. P. McCammond. Normal forms for free aperiodic semigroups. *Internat. J. Algebra Comput.*, 11(5):581–625, 2001.
- [26] J.-E. Pin. *Varieties of formal languages*. Foundations of Computer Science. Plenum Publishing Corp., New York, 1986. With a preface by M.-P. Schützenberger, Translated from the French by A. Howie.
- [27] J. Rhodes. Some results on finite semigroups. *J. Algebra*, 4:471–504, 1966.
- [28] J. Rhodes. Infinite iteration of matrix semigroups. II. Structure theorem for arbitrary semigroups up to aperiodic morphism. *J. Algebra*, 100(1):25–137, 1986. With an appendix by Jerrold R. Goodwin.
- [29] J. Rhodes. Monoids acting on trees: elliptic and wreath products and the holonomy theorem for arbitrary monoids. In *Monoids and semigroups with applications (Berkeley, CA, 1989)*, pages 371–389. World Sci. Publ., River Edge, NJ, 1991.
- [30] J. Rhodes. Flows on automata. Technical report, Center for Pure & Applied Mathematics, University of California, Berkeley, CA 94720-3840, 1995.
- [31] J. Rhodes and B. Steinberg. Profinite semigroups, varieties, expansions and the structure of relatively free profinite semigroups. *Internat. J. Algebra Comput.*, 11(6):627–672, 2001.
- [32] J. Rhodes and B. Steinberg. *The q -theory of finite semigroups*. Springer Monographs in Mathematics. Springer, New York, 2009.
- [33] J. Rhodes and B. Tilson. The kernel of monoid morphisms. *J. Pure Appl. Algebra*, 62(3):227–268, 1989.
- [34] J. Rhodes and P. Weil. Decomposition techniques for finite semigroups, using categories. I, II. *J. Pure Appl. Algebra*, 62(3):269–284, 285–312, 1989.
- [35] I. Simon. Locally finite semigroups and limited subsets of a free monoid. Unpublished manuscript, 1978.
- [36] B. Steinberg. On pointlike sets and joins of pseudovarieties. *Internat. J. Algebra Comput.*, 8(2):203–234, 1998. With an addendum by the author.
- [37] B. Steinberg. Finite state automata: a geometric approach. *Trans. Amer. Math. Soc.*, 353(9):3409–3464 (electronic), 2001.
- [38] B. Steinberg and B. Tilson. Categories as algebra. II. *Internat. J. Algebra Comput.*, 13(6):627–703, 2003.
- [39] B. Tilson. Categories as algebra: an essential ingredient in the theory of monoids. *J. Pure Appl. Algebra*, 48(1-2):83–198, 1987.

DEPT. OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SANTA BARBARA, CA 93106
E-mail address: `mccammon@math.ucsb.edu`

DEPT. OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720
E-mail address: `rhodes@math.berkeley.edu`

SCHOOL OF MATHEMATICS AND STATISTICS, CARLETON UNIVERSITY, OTTAWA, ON
 K1S 5B6, CANADA
E-mail address: `bsteinbg@math.carleton.ca`