DECISION PROBLEMS FOR INVERSE MONOIDS PRESENTED BY A SINGLE SPARSE RELATOR

SUSAN HERMILLER, STEVEN LINDBLAD, AND JOHN MEAKIN

Abstract. We study a class of inverse monoids of the form $M = \mathrm{Inv}\langle X \mid w = 1\rangle$, where the single relator w has a combinatorial property that we call **sparse**. For a sparse word w, we prove that the word problem for M is decidable. We also show that the set of words in $(X \cup X^{-1})^*$ that represent the identity in M is a deterministic context free language, and that the set of geodesics in the Schützenberger graph of the identity of M is a regular language.

Dedicated to the memory of Douglas Munn

1. Introduction

In a seminal paper in 1974, Douglas Munn [Mun74] introduced the notion of birooted edge labeled trees (subsequently referred to as "Munn trees") to solve the word problem for the free inverse monoid. Munn's work was extended by Stephen [Step90] who introduced the notion of Schützenberger graphs to study presentations of inverse monoids. The Schützenberger graphs of an inverse monoid presentation are the strongly connected components of the Cayley graph of the presentation (or equivalently the restrictions of the Cayley graph to the $\mathcal{R}$-classes of the monoid). From a Schützenberger graph for an inverse monoid presentation, the corresponding Schützenberger complex can be defined as the 2-complex whose 1-skeleton is the Schützenberger graph and whose faces have boundaries labeled by the sides of relations [Stei03].

One-relator inverse monoids of the form $M = \mathrm{Inv}\langle X \mid w = 1\rangle$, where $w \in (X \cup X^{-1})^*$, have received some attention in the literature. Birget, Margolis, and Meakin [BMM94] proved that the word problem is solvable for inverse monoids of the form $\mathrm{Inv}\langle X \mid e = 1\rangle$, where e is an idempotent in the free inverse monoid (i.e., reduces to 1 in the free group). Stephen [Step93] observed that if the inverse monoid $M = \mathrm{Inv}\langle X \mid w = 1\rangle$ is E-unitary, then the word problem for M is decidable if there is an algorithm to decide, for any word $u \in (X \cup X^{-1})^*$, whether or not $u = 1$ in M. Furthermore, Ivanov, Margolis, and Meakin [IMM01] proved that if w is cyclically reduced, then $M = \mathrm{Inv}\langle X \mid w = 1\rangle$ is E-unitary. Thus the word problem for $M = \mathrm{Inv}\langle X \mid w = 1\rangle$, w cyclically reduced, is reduced to understanding the Schützenberger graph of 1 in M. This has been used to solve the word problem in several special cases (see for example the paper by Margolis, Meakin and Šunić [MMS05]), but the problem remains open in general, even if w is a cyclically reduced word.

The present paper is concerned with a class of one-relator inverse monoids of the form $M = \mathrm{Inv}\langle X \mid w = 1\rangle$ where $w \in (X \cup X^{-1})^*$ satisfies a combinatorial condition that enables us to understand the structure of the Schützenberger complex corresponding to the identity of M.

Let $w = a_0 \cdots a_{n-1}$ with each a_i in $X \cup X^{-1}$. A **cyclic subword** $q = w(i, j, \epsilon)$ of w is a nonempty word in $(X \cup X^{-1})^*$ of length at most $n-1$ of the form $q = a_i a_{i+1} a_{i+2} \cdots a_{j-1}$

if $\epsilon = 1$ and $q = a_{i-1}^{-1}a_{i-2}^{-1}a_{i-3}^{-1}\cdots a_j^{-1}$ if $\epsilon = -1$, where $i, j \in \mathbb{Z}/n\mathbb{Z}$. The **zone** of the cyclic subword $q = w(i, j, \epsilon)$ is the subset of $\mathbb{Z}/n\mathbb{Z}$ given by $\text{zone}(q) := \{i, i + \epsilon, i + 2\epsilon, ..., j\}$.

Definition 1.1. *A word $w \in (X \cup X^{-1})^*$ is* **sparse** *if w is freely reduced, $l(w) > 1$, and whenever $(q_k, q'_k) = (w(i_k, j_k, \epsilon_k), w(i'_k, j'_k, \epsilon'_k))$ are two pairs of cyclic subwords of w satisfying $q_k = q'_k$ in $(X \cup X^{-1})^*$, $\text{zone}(q_k) \neq \text{zone}(q'_k)$ and $0 \in \text{zone}(q'_k)$ for $k = 1, 2$, then*

- **(sparse 1):** $\text{zone}(q_1) \cap \text{zone}(q'_2) = \emptyset = \text{zone}(q'_1) \cap \text{zone}(q_2)$, *and*
- **(sparse 2):** *either* $\text{zone}(q_1) \cap \text{zone}(q_2) = \emptyset$ *or both* $\epsilon_1\epsilon'_1 = \epsilon_2\epsilon'_2$ *and* $i_1 - \epsilon_1\epsilon'_1 i'_1 = i_2 - \epsilon_2\epsilon'_2 i'_2 \mod n$.

For example one may see easily from this definition that the word $w = aba^{-1}b^{-1}cdc^{-1}d^{-1}$ and all of its cyclic conjugates are sparse. However the word $w = aba^{-1}b^{-1}$ is not sparse. To see this, note that if $q_1 = w(3, 2, -1), q'_1 = w(0, 1, 1), q_2 = w(1, 2, 1)$ and $q'_2 = w(0, 3, -1)$, then $q_1 = q'_1 = a$ in $(X \cup X^{-1})^*$ (where $X = \{a, b, c, d\}$) and $q_2 = q'_2 = b$ in $(X \cup X^{-1})^*$, but $1 \in \text{zone}(q'_1) \cap \text{zone}(q_2)$.

Roughly speaking, if w is a sparse word, then distinct occurrences of prefixes and suffixes of w that occur elsewhere as cyclic subwords of w are separated by at least one letter. This enables us to define an appropriate notion of a dual graph in the Schützenberger complex of 1 and to prove that this dual graph is a tree. From this, we can encode the information contained in the Schützenberger complex of 1 in a pushdown automaton. We can also show that the faces of this Schützenberger complex are of finitely many types and use this to analyze geodesics and cone types in the Schützenberger graph of 1. Specifically, we can prove the following theorems.

Theorem 1.2. *If $w \in (X \cup X^{-1})^*$ is sparse, then the word problem for $M = \text{Inv}\langle X \mid w = 1\rangle$ is solvable.*

Theorem 1.3. *Let w be sparse and let $M = \text{Inv}\langle X \mid w = 1\rangle$. Then:*

(1) *The language of words equal to 1 in M is deterministic context-free.*
(2) *The language of words related to 1 by Green's relation $\mathcal{R}$ in M is deterministic context-free.*

Theorem 1.4. *If w is a sparse word and $M = \text{Inv}\langle X \mid w = 1\rangle$, then the language of geodesics in the Schützenberger graph of 1 for M (i.e. the language of words labeling geodesic paths starting at 1 in $S\Gamma(1)$) is a regular language. That is, the Schützenberger graph of 1 has finitely many cone types.*

In Section 2 of the paper we study some properties of sparse words that enable us to understand how n-gons whose boundaries are labeled by a sparse word may fold together. Section 3 provides information about sequences of complexes that are used to approximate the Schützenberger complex of 1 for an inverse monoid with sparse relator. Section 4 introduces a notion of dual graph to the Schützenberger complex of 1 and this is exploited to provide a proof of Theorem 1.2. In Section 5 we introduce a pushdown automaton that encodes the information contained in the Schützenberger complex of 1 for a one-relator monoid corresponding to a sparse word, and we use this to provide a proof of Theorem 1.3. We also make use of these results to construct a finite state automaton that accepts the geodesics in the Schützenberger graph of 1 for our monoid, and thus provide a proof of Theorem 1.4.

We refer the reader to the book of Lawson [Law98] for much of the basic theory of inverse semigroups and to the paper by Stephen [Step90] for foundational ideas and notation about presentations of inverse monoids.

2. Sparse Words

Throughout this section, $w = a_0 \cdots a_{n-1}$ will denote a fixed sparse word in $(X \cup X^{-1})^*$ as defined in Definition 1.1 above.

Lemma 2.1. *Every sparse word in $(X \cup X^{-1})^*$ is cyclically reduced.*

Proof. Let $w = a_0 \cdots a_{n-1}$ be a sparse word and suppose that $a_0 = a_{n-1}^{-1} = a$. If we let $q_1 = w(0,-1,-1), q_1' = w(0,1,1), q_2 = w(0,-1,-1)$ and $q_2' = w(0,1,1)$, then $q_1 = q_1' = q_2 = q_2' = a$, but $0 \in \text{zone}(q_1) \cap \text{zone}(q_2')$. This contradicts condition (**sparse 1**) of Definition 1.1, so w must be cyclically reduced. □

Lemma 2.2. *Every sparse word $w \in (X \cup X^{-1})^*$ is primitive (i.e. w is not a proper power in $(X \cup X^{-1})^*$).*

Proof. Suppose that $w = u^m$ in $(X \cup X^{-1})^*$ for some $m > 1$. The word u has length $l(u) > 0$ since $l(w) > 0$. If we let $q_1 = w(0, l(u), 1) = q_2'$ and $q_1' = w(-l(u), 0, 1) = q_2$, we again immediately obtain a contradiction of (**sparse 1**). □

We will build 2-dimensional CW-complexes using information from the sparse word w to define the attaching maps. To start, let P be a polygon with n sides; that is, P is a CW-complex with n vertices, n edges and a single 2-cell. We designate a distinguished vertex $\sigma(P)$ of P. We orient the edges of P in a clockwise direction, and label the edges of P so that w is read clockwise from $\sigma(P)$ to $\sigma(P)$ on the boundary ∂P. In addition, we label the vertices of P by the elements of $\mathbb{Z}/n\mathbb{Z}$, starting with 0 at $\sigma(P)$ and labeling in order also in the clockwise direction.

We will build finite 2-complexes iteratively from the n-gon P by successively attaching new copies of P at existing vertices and applying certain edge foldings. More specifically, given a finite collection of copies F_1,F_2,...,F_m of P, first attach the vertex $\sigma(F_2)$ to any vertex of F_1 other than $\sigma(F_1)$. At the glued vertex v, if there are two edges incident to v with either (1) the same orientation and edge label, or (2) opposite orientation and edge labels that are inverse letters in $X \cup X^{-1}$, then we identify those edges to a single 1-cell (and identify the vertices at the other ends to a single vertex). Repeat this successively at all of the vertices of the complex until no further edge identification according to rules (1)–(2) can be done, to obtain a new CW-complex with two 2-cells. Denote the images of F_1 and F_2 in the quotient by $\bar{F}_1$ and $\bar{F}_2$, respectively, and denote the image of $\sigma(F_i)$ by $\bar{\sigma}(F_i)$ for $i = 1, 2$. At the i-th step, we attach F_i to the complex $\bar{F}_1 \cup \cdots \cup \bar{F}_{i-1}$ by identifying $\sigma(F_i)$ with a vertex v' other than one of the $\bar{\sigma}(F_j)$ for $j < i$. We again glue edges according to rules (1)–(2) (where the orientation and label of any edge incident to a face $\bar{F}_j$ can be considered to be that inherited from F_j), to obtain a quotient CW-complex with i faces. (Note that at each step, the complex is finite, so this process must stop.) We say that the face F_i is **folded onto** $\bar{F}_1 \cup \cdots \cup \bar{F}_{i-1}$ **at** v', or that F_i is **attached** at v'.

This process is repeated to create a CW-complex with images $\bar{F}_1$,...,$\bar{F}_m$ of the original polygons as faces. For any index j and vertex v in $\bar{F}_j$, let $i(F_j, v)$ denote the **index** (or the set of indices) of the vertex (resp. vertices) in F_j that is sent to v via the canonical map $F_j \to \bar{F}_1 \cup \cdots \cup \bar{F}_m$.

Note that as a consequence of Lemma 2.1, the two edges of a single face F incident to $\sigma(F)$ cannot be identified to a single edge in this procedure. The definition of sparse also implies restrictions on edge gluings in complexes built from two or three faces, as the following lemmas demonstrate. These lemmas will be applied to determine the structure of the Schützenberger complex of 1 in Section 3.

Lemma 2.3 (The two-face lemma). *Let $\bar{F}_1 \cup \bar{F}_2$ be the CW-complex obtained by folding one face F_2 onto another face F_1 at a vertex $v \neq \sigma(F_1)$. Then $\bar{\sigma}(F_1) \notin \bar{F}_1 \cap \bar{F}_2$.*

Proof. Suppose to the contrary that $\bar{\sigma}(F_1) \in \bar{F}_1 \cap \bar{F}_2$. Since F_2 is folded onto the single face F_1, there must be a path in $\bar{F}_1 \cap \bar{F}_2$ from $\bar{\sigma}(F_1)$ to $v = \bar{\sigma}(F_2)$. The preimage of this path under the map $F_1 \to \bar{F}_1 \cup \bar{F}_2$ is a path in ∂F_1 starting at the vertex $\sigma(F_1)$, and so this path defines a cyclic subword q_1 of w starting at vertex 0 when w is viewed as a word labeling ∂F_1. Similarly, this path defines a cyclic subword q_1' of w ending at vertex 0 when w is viewed as a word labeling ∂F_2. The two pairs of cyclic subwords (q_1, q_1') and $(q_2, q_2') := (q_1', q_1)$ satisfy $0 \in \text{zone}(q_1) \cap \text{zone}(q_2')$, contradicting Definition 1.1. □

Lemma 2.4 (The three-face lemma). *Suppose that the face F_2 is folded onto the face F_1 with at least one pair of edges glued, and suppose that face F_3 is folded onto a vertex $v \in \bar{F}_1 \cap \bar{F}_2$. Then no edges are glued via the folding process for F_3; that is, no edge of F_3 can be glued to an edge of $\bar{F}_1 \cup \bar{F}_2$, and no two edges of $\bar{F}_1 \cup \bar{F}_2$ are identified.*

Proof. By construction, $\bar{F}_1 \cap \bar{F}_2$ is a connected non-empty edge path containing $\bar{\sigma}(F_2)$ and the vertex $v = \bar{\sigma}(F_3)$, so there is a subpath p_1 of $\bar{F}_1 \cap \bar{F}_2$ with endpoints $\bar{\sigma}(F_2)$ and $\bar{\sigma}(F_3)$. When viewed as a path in ∂F_2, p_1 determines a cyclic subword $q_1' = w(i_1', j_1', \epsilon_1')$ such that $\text{zone}(q_1')$ contains both $0 = i(F_2, \bar{\sigma}(F_2))$ and the index $i(F_2, v)$ of the vertex corresponding to v. When viewed as a path in ∂F_1, p_1 determines a cyclic subword $q_1 = w(i_1, j_1, \epsilon_1)$ such that $\text{zone}(q_1)$ contains $i(F_1, \bar{\sigma}(F_2))$ and $i(F_1, v)$.

Suppose that some edge of F_3 is glued onto an edge of $\bar{F}_1 \cup \bar{F}_2$.

Case 1. F_3 folds onto an edge of $\bar{F}_1$. Then there is a non-trivial path p_2 in $\bar{F}_1 \cap \bar{F}_3$ with endpoint $v = \bar{\sigma}(F_3)$. When viewed as a path in ∂F_3, p_2 determines a cyclic subword $q_2' = w(i_2', j_2', \epsilon_2')$ with $0 \in \text{zone}(q_2')$. When viewed as a path in ∂F_1, p_2 determines a cyclic word $q_2 = w(i_2, j_2, \epsilon_2)$ such that $i(F_1, v) \in \text{zone}(q_2)$. Then $i(F_1, v) \in \text{zone}(q_1) \cap \text{zone}(q_2) \neq \emptyset$. But $i_1 - \epsilon_1\epsilon_1' i_1' = i(F_1, \bar{\sigma}(F_2))$ and $i_2 - \epsilon_2\epsilon_2' i_2' = i(F_1, v)$, so $i_1 - \epsilon_1\epsilon_1' i_1' \neq i_2 - \epsilon_2\epsilon_2' i_2'$, contradicting condition (**sparse 2**) of Definition 1.1. Thus Case 1 cannot occur.

Case 2. F_3 folds onto an edge of $\bar{F}_2$. Then there is a non-trivial path p_3 in $\bar{F}_2 \cap \bar{F}_3$ with endpoint $v = \bar{\sigma}(F_3)$. When viewed as a path in ∂F_3, p_3 determines a cyclic subword $q_3' = w(i_3', j_3', \epsilon_3')$ with $0 \in \text{zone}(q_3')$. When viewed as a path in ∂F_2, p_3 determines a cyclic subword $q_3 = w(i_3, j_3, \epsilon_3)$ with $i(F_2, v) \in \text{zone}(q_3)$. In this case, $i(F_2, v) \in \text{zone}(q_1') \cap \text{zone}(q_3) \neq \emptyset$, so condition (**sparse 1**) fails, a contradiction.

Since no edge of F_3 is folded onto any edge of $\bar{F}_1 \cup \bar{F}_2$, no additional edge folding can occur in $\bar{F}_1 \cup \bar{F}_2$. □

3. The Schützenberger complex $SC(1)$

Throughout this section, w will denote a fixed sparse word and $M = \text{Inv}\langle X \mid w = 1 \rangle$. We recall that the Schützenberger graph of 1 for this presentation is the restriction of the Cayley graph of M to the $\mathcal{R}$-class of 1. We denote this graph by $S\Gamma(1)$: its vertices are the elements $s \in M$ such that $ss^{-1} = 1$ in M and there is an edge labeled by $x \in X \cup X^{-1}$ from s to t if $ss^{-1} = tt^{-1} = 1$ and $sx = t$ in M. We denote this edge by (s, x, t). Its inverse edge is the edge (t, x^{-1}, s) in $S\Gamma(1)$, where we interpret $(x^{-1})^{-1} = x$, and this inverse pair is interpreted as a single topological edge. The Schützenberger complex of 1 is the complex $SC(1)$ obtained from $S\Gamma(1)$ by adding a face with boundary label w for each closed path labeled by w in $S\Gamma(1)$. Stephen's iterative construction of a sequence of approximations of the Schützenberger graph $S\Gamma(1)$ may easily be adapted to yield a sequence of approximations of the Schützenberger complex $SC(1)$. In particular, we may construct such a sequence of complexes in the following way.

Start with a trivial complex S_0 consisting of one vertex v_0 and no edges or faces. Take a copy F_1 of the n-gon P, identify its start vertex $\sigma(P)$ with v_0, and denote this complex by S_1. As in Section 2, we build a sequence of complexes $S_1 = \bar{F}_1$, $S_2 = \bar{F}_1 \cup \bar{F}_2$, $S_3 = \bar{F}_1 \cup \bar{F}_2 \cup \bar{F}_3$, ... by successively folding faces F_i onto $\bar{F}_1 \cup \bar{F}_2 \cup \cdots \cup \bar{F}_{i-1}$ at vertices $v_{i-1} \in S_{i-1}$ at which no face has yet been attached, in such a way that $d(v_0, v_{i-1})$ is as small as possible, where d is the path metric in S_{i-1}. Lemma 2.2 guarantees that such a v_{i-1} exists. To see this, note that if no such v_{i-1} exists, then $S\Gamma(1) = S_{i-1}$, so $S\Gamma(1)$ is finite. Thus if x is the first letter in w, since x^j labels a path in $S\Gamma(1)$ for each $j > 0$ we see that x is a torsion element in M (i.e. $x^j = x^k$ for some $k \neq j$). It follows that x must be a torsion element of $G = Gp\langle X | w = 1\rangle$, but G is torsion free if w is primitive.

A sequence of complexes obtained in the above manner is referred to as a **Schützenberger approximation sequence**. Since $v_i = \bar{\sigma}(F_{i+1})$ is chosen so as to minimize the distance from v_0, we can see that every vertex of S_i is the start vertex of some face in S_{i+j} for some j. From the results of Stephen [Step90], the corresponding sequence of 1-skeleta of a Schützenberger approximation sequence has a direct limit that is independent of the choice of the vertices v_i, and this direct limit is $S\Gamma(1)$. By an argument similar to the formal category theoretical argument in [Step90] used to show this, it follows that the Schützenberger approximation sequence of complexes has a direct limit, and since the approximation sequence attaches faces whenever a closed path labeled by w is attached, the limit of the Schützenberger approximation sequence is the Schützenberger complex $SC(1)$.

Theorem 3.1. *Let $S_0, S_1, S_2, \ldots$ be any Schützenberger approximation sequence for $SC(1)$ corresponding to a sparse word w. Then for all $m \geq 0$ and for all distinct faces $\bar{F}_i, \bar{F}_j, \bar{F}_k, \bar{F}_l$ in S_m*

(1) *The natural map $F_i \to \bar{F}_i$ is an embedding of F_i into S_m.*
(2) *If $\bar{F}_i \cap \bar{F}_j \neq \emptyset$, then $\bar{F}_i \cap \bar{F}_j$ is a connected path such that either $\bar{\sigma}(F_i) \in \bar{F}_j$ with $\bar{\sigma}(F_j) \notin \bar{F}_i$, or $\bar{\sigma}(F_j) \in \bar{F}_i$ with $\bar{\sigma}(F_i) \notin \bar{F}_j$.*
(3) *If $\bar{F}_i \cap \bar{F}_j \cap \bar{F}_k \neq \emptyset$, then there exists $r \in \{i, j, k\}$ with $\bar{F}_i \cap \bar{F}_j \cap \bar{F}_k = \bar{\sigma}(F_r)$ and $\bar{F}_r$ shares no other vertices with the other two faces.*
(4) *$\bar{F}_i \cap \bar{F}_j \cap \bar{F}_k \cap \bar{F}_l = \emptyset$.*
(5) *The natural map from S_{m-1} to S_m is an embedding.*

Proof. The proof proceeds by induction on m. The result is clear if m is 0 or 1. Suppose that the result is true for approximation sequences of length $m - 1$. Let v be the vertex of S_{m-1} at which F_m is attached to S_{m-1}. From part (4) of the induction assumption, at most three faces contain the point v.

Case 1. Suppose that v is on the boundary of three faces in S_{m-1}. Then by part (3) of the induction assumption, one of these faces $\bar{F}$ satisfies $v = \bar{\sigma}(F)$. But then the algorithm for constructing the Schützenberger approximation sequence would not attach F_m at v also. Hence Case 1 cannot occur.

Case 2. Suppose that v is on the boundary of exactly two faces $\bar{F}_i$ and $\bar{F}_j$ in S_{i-1}. By part (2) of the induction hypothesis, we may assume without loss of generality that

$\bar{\sigma}(F_i) \in \bar{F}_j$ and again by this induction hypothesis there is a non-trivial path in $\bar{F}_i \cap \bar{F}_j$ from v to $\bar{\sigma}(F_i)$. Then by the three-face lemma, no edge of F_m is glued onto any edge of $\bar{F}_i \cup \bar{F}_j$ at v, and hence no edge of F_m is glued onto any edge of S_{m-1} at all. Hence properties (1)–(5) of the statement of the theorem hold for S_m.

Case 3. Suppose that v is on the boundary of exactly one face $\bar{F}_i$ of S_{m-1}. Consider the complex $\hat{S_m}$ obtained from S_{m-1} and F_m by just gluing edges of F_m and $\bar{F}_i$ starting from v, and no additional edge foldings. Then $\bar{F}_i \cap \hat{F_m}$ is a connected path. If there exists a vertex v' in $\bar{F}_i \cap \hat{F_m}$ with $v' \neq v = \hat{\sigma}(F_m)$, the two-face lemma says that $v' \neq \bar{\sigma}(F_i)$ also. In this case the three-face lemma then says that any other face incident to v' cannot contain an edge that can be identified with an edge of either $\bar{F}_i$ or $\hat{F_m}$ in a further folding process. Thus in any case no further edges can be glued, and $\hat{S_m} = S_m$. Hence properties (1)–(5) of the statement of the theorem hold for S_m. □

Using part (5) of Theorem 3.1, we may consider $S_0 \subset S_1 \subset S_2 \subset S_3 \subset \cdots$, and so $SC(1) = \cup_{m=0}^{\infty} S_m$ for any Schützenberger approximation sequence constructed as above. Hence the corollary below follows immediately.

Corollary 3.2. *Properties (1)–(5) of Theorem 3.1 hold with S_m replaced by $SC(1)$.*

For every Schützenberger approximation sequence, there is a vertex v_0 which is the unique vertex incident to only one face in the direct limit, and so there is a unique vertex in $SC(1)$, which we will also call v_0, that is incident to only one face, which we will refer to throughout as the face F_1. For any face A of $SC(1)$, the sparse property of w implies that there is only one vertex in ∂A that can be the start vertex $\bar{\sigma}(A)$, and only one possible orientation starting from this vertex in which the word w labels the boundary path.

For distinct faces A and B of $SC(1)$, we define $A < B$ if the face A must be attached before the face B in every Schützenberger approximation sequence. The corresponding partial ordering $\leq$ is the **face ordering** on the faces of $SC(1)$. This partial ordering is well-founded, and the face F_1 is a minimal element.

Corollary 3.3 (**Order Corollary**)**.** *Let v be a vertex of $SC(1)$, and let B be the face with $v = \bar{\sigma}(B)$.*

(1) *If v is incident to exactly one other face A, then $A < B$.*
(2) *If v is incident to two other faces A and C with $\bar{\sigma}(C) \in A$, then $A < B$ and $A < C$.*
(3) *If v is incident to a face A and $A \cap B$ contains at least one edge, then $A < B$.*

Proof. Let $S_0, S_1, S_2, \ldots$ be any Schützenberger approximation sequence for $SC(1)$ corresponding to a sparse word w, with face F_i attached to S_{i-1} in the construction of S_i, as above. In the case that v is incident only to faces $A = \bar{F}_j$ and $B = \bar{F}_k$, the vertex v must exist in a complex S_i before B can be attached, and so we must have $j < k$.

In the case that v is also incident to a third face $C = \bar{F}_l$ with $\bar{\sigma}(C) \in A$, then Theorem 3.1 says that $A \cap C$ contains a connected non-empty edge path from $\bar{\sigma}(C)$ to v, and so at the vertex $\bar{\sigma}(C)$, an edge of C is glued to an edge of A. Again applying Theorem 3.1, no face other than A and C can be incident to $\bar{\sigma}(C)$ in any of the S_i. Then as in the paragraph above, we have $j < l$. Now the face F_k can be attached at v only after v has been built in the sequence, and hence only after at least one of F_j, F_l has been attached. Therefore $j < k$ also.

Finally, if $v \in A$ and $A \cap B$ contains at least one edge, then Theorem 3.1 says that no other face can be incident to v, and so the first paragraph of this proof applies. □

4. The dual graph and the word problem

In this section we define a notion of a dual graph of the Schützenberger complex $SC(1)$ for an inverse monoid $M = \mathrm{Inv}\langle X \mid w = 1\rangle$ corresponding to a sparse word w. We show that this dual graph is a tree and we make use of this to provide a solution to the word problem for M.

Definition 4.1. *Let w be a sparse word. The* dual graph *of $SC(1)$ for $M = \mathrm{Inv}\langle X \mid w = 1\rangle$ is the directed graph $\mathcal{D}$ with*

- *vertex set $V(\mathcal{D})$ given by the set of faces of $SC(1)$, and*
- *set $E(\mathcal{D})$ of directed edges (A, B) (oriented from A to B) for $A, B \in V(\mathcal{D})$ satisfying $A < B$ in the face ordering and $A \cap B \neq \emptyset$ in $SC(1)$.*

As a consequence of Corollaries 3.2 and 3.3, the definition of $E(\mathcal{D})$ can also be phrased purely in terms of the combinatorial properties of $SC(1)$, namely (A, B) is a directed edge in $\mathcal{D}$ if and only if $A \neq B$, $\bar{\sigma}(B) \in A$, and whenever $C \in V(\mathcal{D})$ with $\bar{\sigma}(B) \in C$ then $\bar{\sigma}(C) \in A$.

Proposition 4.2. *Let w be a sparse word and $M = \mathrm{Inv}\langle X \mid w = 1\rangle$. Then the dual graph $\mathcal{D}$ of $SC(1)$ is a directed, rooted, infinite tree (with root F_1) in which each vertex has at most $l(w) - 1$ children.*

Proof. Recall that the face F_1 is the only face of $SC(1)$ containing the unique vertex v_0 of $SC(1)$ incident to only one face. Let $A \neq F_1$ be any other face in $SC(1)$, and assume by Noetherian induction that for all faces $B < A$ with respect to the well-founded face ordering, there is a directed edge path in $\mathcal{D}$ from F_1 to B. From Corollary 3.2, there are either 2 or 3 faces incident to the vertex $\bar{\sigma}(A)$ in $SC(1)$, including A.

If there is only one other face B incident to $\bar{\sigma}(A)$, then the Order Corollary 3.3 implies that $B < A$. Since $\bar{\sigma}(B) \in A \cap B \neq \emptyset$, then $(B, A) \in E(\mathcal{D})$. The concatenation of the path from F_1 to B from the induction assumption with this edge (B, A) then gives a directed edge path in $\mathcal{D}$ from F_1 to A.

On the other hand, if there are two other faces B and C incident to $\bar{\sigma}(A)$, then Corollary 3.2 says that one of these faces contains the $\bar{\sigma}$ vertex of the other; without loss of generality, suppose that $\bar{\sigma}(C) \in B$. Then the Order Corollary 3.3 again implies that $B < A$, and as in the previous paragraph we obtain a directed path in $\mathcal{D}$ from F_1 to A. Hence $\mathcal{D}$ is connected.

Suppose that $\mathcal{D}$ is not a tree. Then there is an undirected circuit in this graph.

Suppose that two edges of this circuit have a common target; that is, suppose that there are edges $(A, C), (B, C) \in E(\mathcal{D})$ with $A \neq B$. Using the combinatorial description of $E(\mathcal{D})$ above, then $\bar{\sigma}(C) \in A \cap B$. From Corollary 3.2 part (2), $A \cap B$ is a path containing one of $\bar{\sigma}(A)$ or $\bar{\sigma}(B)$ but not both. This contradicts the existence of one of the edges $(A, C), (B, C)$, and so the circuit must also be a directed circuit.

The consecutive vertices $A_1, A_2, ..., A_k$ following the directed edges in this circuit must then satisfy $A_1 < A_2 < \cdots A_k < A_1$ in the face ordering, which is again a contradiction. Hence $\mathcal{D}$ is a directed tree with root F_1.

Since each face A of $SC(1)$ has $l(w) - 1$ vertices other than its vertex $\bar{\sigma}(A)$, there are at most $l(w) - 1$ directed edges in $\mathcal{D}$ with source vertex A. In addition, as remarked in

Section 3, the fact that w is primitive guarantees $SC(1)$ is infinite. Therefore the tree $\mathcal{D}$ must be infinite. □

To simplify notation later, it will be helpful to consider a slight modification of $\mathcal{D}$. The **augmented dual graph** $\mathcal{D}'$ is obtained from $\mathcal{D}$ by adding an additional vertex v_0 to $\mathcal{D}$ and an additional directed edge from v_0 to F_1. Then $\mathcal{D}'$ is a directed rooted tree with root v_0. Using standard language for rooted trees, if (A, B) is a directed edge in $\mathcal{D}'$, we call A the parent of B, and B a child of A.

Define a map $\Omega : V(SC(1)) \to V(\mathcal{D}')$ as follows. For each vertex $v \neq v_0$ in $SC(1)$, let $\Omega(v)$ be the unique face of $SC(1)$ that is closest to v_0 in $\mathcal{D}'$ from among the faces that are incident to v, and let $\Omega(v_0) := v_0$.

For any face A of $SC(1)$, then $\Omega(\bar{\sigma}(A))$ is the parent of A, and so $\Omega(\bar{\sigma}(A)) < A$; i.e., $\Omega(\bar{\sigma}(A))$ must be attached before A in any Schützenberger approximation. By Corollaries 3.2 and 3.3, in the folding process edges of A can be glued to edges of $\Omega(\bar{\sigma}(A))$ but not to edges of any other face, and the glued edges are a connected path. Recall that the boundary ∂A of the polygon A is labeled by the word w, when read starting at the vertex $\sigma(A)$ in the clockwise direction. The connected set $\gamma(A) := A \cap \Omega(\bar{\sigma}(A))$, then, can be regarded as the image of a ("gluing") path (which we will also call $\gamma(A)$) going clockwise around ∂A from the ("reverse") vertex $\rho(A)$ to the ("forward") vertex $\phi(A)$. Note that if no edges are glued when A is attached to its parent $\Omega(\bar{\sigma}(A))$, then $\rho(A) = \bar{\sigma}(A) = \phi(A)$ and $\gamma(A)$ is this point.

Lemma 4.3. *Let A be a face of the complex $SC(1)$ for a sparse word w.*

1. *The lengths $l(w)$ and $l(\gamma(A))$ satisfy $l(\gamma(A)) \leq \frac{1}{2}l(w) - 1$.*
2. *If v is a vertex in $\gamma(A)$, then $\Omega(v) = \Omega(\bar{\sigma}(A))$.*
3. *If v is a vertex in $\partial A \setminus \gamma(A)$, then $\Omega(v) = A$.*

Proof. The path $\gamma(A)$ determines a cyclic subword q' of w when viewed as a path in ∂A, and determines a cyclic subword q when viewed as a path in the parent $\partial\Omega(\bar{\sigma}(A))$ of A. Since w is sparse we must have $\text{zone}(q') \cap \text{zone}(q) = \emptyset$, (take $(q_1, q_1') = (q_2, q_2') = (q, q')$ in Definition 1.1), and so there must also be at least one edge between the endpoints of these cyclic subwords on both sides. Then $l(w) \geq 2l(\gamma(A)) + 2$.

If v is a vertex in $\gamma(A)$, then by definition of the set γ, the point v is also in the parent $\Omega(\bar{\sigma}(A))$ of A. If there is a third face C incident to v in $SC(1)$, then by the order corollary and the definition of $\mathcal{D}'$, the face $\Omega(\bar{\sigma}(A))$ is also the parent of C.

For a vertex v in $\partial A \setminus \gamma(A)$, then $v = \bar{\sigma}(B)$ for another face B of $SC(1)$. If there is no other face incident to v, the order corollary then says $A < B$. If C is a third face incident to v, then Corollary 3.2 says that A and C must share at least one edge in common, and either $\bar{\sigma}(A)$ is in C, or $\bar{\sigma}(C)$ is in A. The order corollary then says that the face among A and C that contains the start vertex $\bar{\sigma}$ of the other is the parent of the pair. However, since $v \notin \Omega(\bar{\sigma}(A))$, we must have $C \neq \Omega(\bar{\sigma}(A))$, and hence $\bar{\sigma}(C) \in A$ and A is the parent of both B and C. □

Let the 1-skeleton $S\Gamma(1)$ of the 2-complex $SC(1)$ have the path metric $d_{S\Gamma}$, and let the augmented dual graph have path metric $d_{\mathcal{D}'}$. The following theorem shows that geodesics in these metric spaces are closely related.

Theorem 4.4 (Geodesic Theorem). *Let p be any geodesic edge path in $S\Gamma(1)$ from v_0 to a vertex v. Let $v_0, v_1, ..., v_k = v$ be the successive vertices in the path p. Then for all i, either $\Omega(v_i) = \Omega(v_{i+1})$ or $\Omega(v_i)$ is the parent of $\Omega(v_{i+1})$ in $\mathcal{D}'$, and the edge* *from* v_i *to* v_{i+1} *is contained in* $\Omega(v_{i+1})$. *Moreover, whenever* $\Omega(v_i) < \Omega(v_{i+1})$, *then* $v_i \in \{\rho(\Omega(v_{i+1})), \phi(\Omega(v_{i+1}))\}$.

Proof. We prove this by induction on the length k of the edge path p. If $k = 0$, then p is the constant path at v_0 in $SC(1)$, and there is no other vertex. If $k = 1$, then p follows a single edge from v_0 to $v_1 \neq v_0$. Then $\Omega(v_0) = v_0$ is the parent of $\Omega(v_1) = F_1$ in $\mathcal{D}'$.

Suppose that $k \geq 2$. The prefix $\hat{p}$ of the path p with vertices $v_0, ..., v_{k-1}$ is also a geodesic path in $S\Gamma(1)$, and so by induction the conditions on the pair $\Omega(v_i), \Omega(v_{i+1})$ in the theorem hold for all $0 \leq i \leq k-2$. The vertex $v_{k-1} \neq v_0$, so Corollary 3.2 says that there are at least two faces $A := \Omega(v_{k-1})$ and B with $\bar{\sigma}(B) = v_{k-1}$, and possibly a third face C, incident to the vertex v in $SC(1)$. By definition of Ω and the Order Corollary, we have $A < B$ and $A < C$. The edge e from v_{k-1} to v_k must be contained in at least one of these faces.

Case 1. Suppose that e is contained in A. If v_k is in the path $\gamma(A)$, then Lemma 4.3 implies that $\Omega(v_k) = \Omega(\bar{\sigma}(A))$, but since $\Omega(v_{k-1}) = A$, the same lemma implies that v_{k-1} is not in $\gamma(A)$. Then v_k must be one of the endpoints $\rho(A)$, $\phi(A)$ of $\gamma(A)$. By induction, the prefix $\hat{p}$ of p traversed one of these endpoints, and since p is a geodesic, $\hat{p}$ must have traversed the endpoint v' of $\gamma(A)$ that is not v_k. However, this implies that a suffix of p is a geodesic in ∂A from v' to v_k that goes through the point v_{k-1} not in $\gamma(A)$. This contradicts Lemma 4.3(1), and so v_k must lie in $\partial A \setminus \gamma(A)$. Lemma 4.3(3) then implies that $\Omega(v_k) = A = \Omega(v_{k-1})$.

Case 2. Suppose that e is contained in a child E of the face A, but not in A. That is, E is one of the faces B or C. In this case, since v_k is not contained in $A \cap E$, then Lemma 4.3(3) says that $\Omega(v_k)$ is E, and we have that $\Omega(v_{k-1}) = A$ is the parent of $\Omega(v_k)$. Moreover, since v_{k-1} is in $A \cap E$ but v_k is not, we have that v_{k-1} is one of the endpoints $\rho(E)$, $\phi(E)$ of $\gamma(E)$. □

We can now provide a solution to the word problem for M.

Proof of Theorem 1.2. As noted in Section 1, it is sufficient to prove that there is an algorithm that takes a word $u \in (X \cup X^{-1})^*$ as input, and outputs whether or not $u = 1$ in M. Given a sparse word w, the following procedure is such an algorithm for $M = \text{Inv}\langle X \mid w = 1\rangle$.

Let $L := l(u)$ be the length of the word u. The algorithm follows the construction of a Schützenberger approximation sequence as described at the beginning of Section 3, attaching a face at each step to a vertex whose distance to v_0 in the approximation complex is minimal from among all of those vertices that are not yet the start vertex ($\bar{\sigma}$) of a face. Continue this process until the next vertex at which a face is to be attached has distance $L \cdot l(w) + 1$ from v_0; the process stops at this time, with an approximation complex S. Since each complex in this sequence is locally finite, this process is finite.

From Theorem 3.1 we know that S embeds in $SC(1)$. From the Geodesic Theorem 4.4, we have that for each vertex v in S, any geodesic path p in $S\Gamma(1)$ from v_0 to v is contained in the union of the the faces labeling vertices of the geodesic in $\mathcal{D}'$ from v_0 to $\Omega(v)$. By the definition of the map Ω, these are the faces that must be constructed in the Schützenberger approximation sequence before the face $\Omega(v)$, together with the face $\Omega(v)$ which must be the first face containing v constructed in the sequence. Hence all of these faces are also in S, as is the path p. Therefore the path metric d_S in the 1-skeleton of S is the same as the metric inherited from $S\Gamma(1)$.

We claim that every face A of $SC(1)$ with $d_{\mathcal{D}'}(v_0, A) \leq L$ lies in S. Suppose not; that is, suppose that there is a face A with $d_{\mathcal{D}'}(v_0, A) \leq L$ and A not in S, and choose A to have minimal distance from v_0 in $\mathcal{D}'$ among all such faces. Then the parent $\Omega(\bar{\sigma}(A))$ of A satisfies $d_{\mathcal{D}'}(v_0, \Omega(\bar{\sigma}(A))) = d_{\mathcal{D}'}(v_0, A) - 1$, and so $\Omega(\bar{\sigma}(A))$ lies in S. But the previous paragraph and Theorem 4.4 imply that $d_S(v_0, \bar{\sigma}(A)) = d_{S\Gamma}(v_0, \bar{\sigma}(A)) < L \cdot l(w)$, and so S has a vertex $\bar{\sigma}(A)$ within $L \cdot l(w)$ of v_0 that is not the start vertex of an attached face, giving the required contradiction.

Now let v' be any vertex of $SC(1)$ with $d_{S\Gamma}(v_0, v') \leq L$. It follows from Theorem 4.4 that $d_{\mathcal{D}'}(v_0, \Omega(v')) \leq L$ also, and so by the previous paragraph, $\Omega(v')$, and hence also v', is in the finite complex S. Putting these results together, we have that u labels a path from v_0 to v_0 in $SC(1)$ if and only if u labels a path from v_0 to v_0 in S. Since $u = 1$ in M if and only if u labels a path in $SC(1)$ from v_0 to v_0, the algorithm outputs $u = 1$ if u labels a path from v_0 to v_0 in S, and outputs $u \neq 1$ in M otherwise. □

5. Languages of geodesics and words representing 1

Throughout this section, w is a sparse word and $M = \mathrm{Inv}\langle X \mid w = 1 \rangle$.

Lemma 5.1. *Let A be any face in $SC(1)$. There is a unique point x_A in $\partial A \setminus \gamma(A)$ satisfying $d_{S\Gamma}(v_0, x_A) \geq d_{S\Gamma}(v_0, y)$ for all $y \in \partial A$.*

Proof. First consider points in the set $T := \partial A \setminus \gamma(A)$. Lemma 4.3(3) and the Geodesic Theorem 4.4 imply that every geodesic from v_0 to a point y in T must traverse one of the points $\rho(A), \phi(A)$, and then follow edges in the path along T to y. Let $a := d_{S\Gamma}(v_0, \rho(A))$, $b := d_{S\Gamma}(v_0, \phi(A))$, and $q := l(\gamma(A))$, and let p be the length of the edge path in T from $\rho(A)$ to $\phi(A)$. The triangle inequality together with Lemma 4.3(1) give $|b-a| \leq q \leq p-1$. Let x be the point in T that is a distance $\frac{1}{2}(p + (b-a)) < p$ from the endpoint $\rho(A)$; then x is a distance $\frac{1}{2}(p + (a-b))$ along T from $\phi(A)$. Now the concatenation of a geodesic path from v_0 to $\rho(A)$ followed by the geodesic in T from $\rho(A)$ to x has the same length $\frac{1}{2}(p + a + b)$ as the concatenation of a geodesic path from v_0 to $\phi(A)$ followed by the geodesic in T from $\phi(A)$ to x, and hence both of these concatenations are geodesics from v_0 to x. Since every other point $y \in T$ lies on one of these paths, we have $d_{S\Gamma}(v_0, x) > d_{S\Gamma}(v_0, y)$.

Similarly, let z be the point in $\gamma(A)$ that is a distance $\frac{1}{2}(q + (b-a)) \leq q$ from the endpoint $\rho(A)$ along the path $\gamma(A)$, and hence a distance $\frac{1}{2}(q + (a-b))$ from $\phi(A)$. The concatenation of a geodesic from v_0 to either $\rho(A)$ or $\phi(A)$, together with the geodesic along $\gamma(A)$ from that endpoint to z, has length $\frac{1}{2}(q + a + b)$, and every point y in $\gamma(A)$ lies on one of these path concatenations. Hence for all $y \in \gamma(A)$, we also have $d_{S\Gamma}(v_0, x) = \frac{1}{2}(p + a + b) > \frac{1}{2}(q + a + b) \geq d_{S\Gamma}(v_0, y)$. □

For a face A of $SC(1)$, choose $\mathbb{Z}$ representatives $\hat{i}(A, \rho(A))$ and $\hat{i}(A, \phi(A))$ of the indices $i(A, \rho(A))$ and $i(A, \phi(A))$ from $\mathbb{Z}/n\mathbb{Z}$, respectively, satisfying $0 \leq \hat{i}(A, \phi(A)) < \hat{i}(A, \rho(A)) \leq n = l(w)$. Similarly, for each vertex v in $\partial A \setminus \gamma(A)$, let $\hat{i}(A, v)$ be the representative of $i(A, v)$ satisfying $0 < \hat{i}(A, v) < n$; from Lemma 4.3, then $\hat{i}(A, \phi(A)) < \hat{i}(A, v) < \hat{i}(A, \rho(A))$.

Define $k_A := \frac{1}{2}[\hat{i}(A, \rho(A)) + \hat{i}(A, \phi(A)) + (d_{S\Gamma}(v_0, \rho(A)) - d_{S\Gamma}(v_0, \phi(A)))]$. The proof above shows that the point x_A lies at the index $i(A, x_A) = k_A \pmod{n\mathbb{Z}}$ if x_A is a vertex, otherwise x_A lies at the midpoint of the edge whose endpoints y, z are the vertices with indices $i(A, y), i(A, z)$ given by $k_A \pm \frac{1}{2} \pmod{n\mathbb{Z}}$.

Definition 5.2. *For any face A of $SC(1)$, we define the associated triple $ft(A) := (\hat{i}(A,\rho(A)), \hat{i}(A,\phi(A)), k_A)$. We define an equivalence relation $\sim_{ft}$ on the set of faces of $SC(1)$ by $A \sim_{ft} B$ if and only $ft(A) = ft(B)$; in this case, we say that A and B have the same* **face type**. *Define an equivalence relation $\sim_{ft}$ on the set of vertices of $SC(1)$ by $u \sim_{ft} v$ if and only if $\Omega(u) \sim_{ft} \Omega(v)$ and $i(\Omega(u),u) = i(\Omega(v),v)$. Denote the equivalence class of a vertex or face z relative to $\sim_{ft}$ by $[z]$.*

Note that there are only finitely many face types, and similarly only finitely many $\sim_{ft}$-equivalence classes of vertices. For example, it follows from this definition that if A is a face of $SC(1)$ that is attached to $\Omega(\bar{\sigma}(A))$ at the vertex $\bar{\sigma}(A)$ in such a way that no edge of A folds onto $\Omega(\bar{\sigma}(A))$, then the triple for A is $(n, 0, n/2)$, and $A \sim_{ft} F_1$. Since $\Omega(v_0)$ is not a face of $SC(1)$, the $\sim_{ft}$-equivalence class $[v_0]$ contains only the vertex v_0.

The following lemma will be used in the constructions of a push-down automaton and a finite state automaton later in this section.

Lemma 5.3. *In $SC(1)$ let u_1, u_2 be vertices with $u_1 \sim_{ft} u_2$ and let $e_1 = (u_1, x, v_1)$ be an edge. Suppose that either*
(i) $\Omega(u_1) = \Omega(v_1)$,
(ii) $(\Omega(u_1), \Omega(v_1)) \in E(\mathcal{D}')$, or
(iii) $(\Omega(v_1), \Omega(u_1)) \in E(\mathcal{D}')$ and $\bar{\sigma}(\Omega(u_1)) \sim_{ft} \bar{\sigma}(\Omega(u_2))$.
Then there is an edge $e_2 = (u_2, x, v_2)$ in $SC(1)$ with $v_1 \sim_{ft} v_2$ satisfying, respectively,
(i) $\Omega(u_2) = \Omega(v_2)$ and $\hat{i}(\Omega(u_1), v_1)$ lies between $\hat{i}(\Omega(u_1), u_1)$ and $k_{\Omega(u_1)}$ (inclusive) if and only if $\hat{i}(\Omega(u_2), v_2)$ lies between $\hat{i}(\Omega(u_2), u_2)$ and $k_{\Omega(u_2)}$.
(ii) $(\Omega(u_2), \Omega(v_2)) \in E(\mathcal{D}')$ and $\bar{\sigma}(\Omega(v_1)) \sim_{ft} \bar{\sigma}(\Omega(v_2))$, or
(iii) $(\Omega(v_2), \Omega(u_2)) \in E(\mathcal{D}')$.

Proof. Suppose first that $u_1 = v_0$. Then $u_1 \sim_{ft} u_2$ implies that $u_2 = v_0 = u_1$, and the result of the lemma follows. For the remainder of the proof, we assume that $u_1 \neq v_0$, and as a consequence $u_2 \neq v_0$. Let A_i be the face $\Omega(u_i)$ for $i = 1, 2$. By definition of $u_1 \sim_{ft} u_2$, then $A_1 \sim_{ft} A_2$ and $i(A_1, u_1) = i(A_2, u_2)$.

Suppose that (i) $\Omega(u_1) = \Omega(v_1)$ holds. Then the edge e_1 lies in the face A_1. The faces A_1 and A_2 are copies of the same polygon with the same boundary label word w, and we have $i(A_1, u_1) = i(A_2, u_2)$, hence there is an edge $e_2 = (u_2, x, v_2)$ in the boundary of A_2 with $i(A_1, v_1) = i(A_2, v_2)$. From the definition of $A_1 \sim_{ft} A_2$, we have $\hat{i}(A_1, \phi(A_1)) = \hat{i}(A_2, \phi(A_2))$ and $\hat{i}(A_1, \rho(A_1)) = \hat{i}(A_2, \rho(A_2))$. From Lemma 4.3, the edge e_1 lies in $\partial A_1 \setminus \gamma(A_1)$, and so we have $\hat{i}(A_1, \phi(A_1)) < \hat{i}(A_1, v_1) < \hat{i}(A_1, \rho(A_1))$. Then $\hat{i}(A_2, \phi(A_2)) < \hat{i}(A_2, v_2) < \hat{i}(A_2, \rho(A_2))$, and so v_2 lies in $\partial A_2 \setminus \gamma(A_2)$. Applying the same lemma again gives $\Omega(v_2) = A_2$. Then both $v_1 \sim_{ft} v_2$ and the betweenness condition follow directly.

Next suppose that (ii) $(\Omega(u_1), \Omega(v_1)) \in E(\mathcal{D}')$. In this case, $B_1 := \Omega(v_1)$ is a face of $SC(1)$. Since $A_1 < B_1$, the edge e_1 lies in B_1, the vertices u_1 and $\bar{\sigma}(B_1)$ (which may or may not be the same point) both lie in $A_1 \cap B_1$, and v_1 lies in $B_1 \setminus \gamma(B_1)$. Let B_2 be the face of $SC(1)$ whose vertex $\bar{\sigma}(B_2)$ lies at the vertex of ∂A_2 satisfying $i(A_2, \bar{\sigma}(B_2)) = i(A_1, \bar{\sigma}(B_1))$. Again using the fact that the pairs of polygons A_1, B_1 and A_2, B_2 have the same boundary labels, the gluings of B_2 onto A_2 correspond to the gluings of B_1 onto A_1. Hence $u_2 \in A_2 \cap B_2$, and there is an edge (u_2, x, v_2) in B_2 with $v_2 \notin A_2$. Since $\Omega(u_2) = A_2$, we have $A_2 < B_2$, and so $(A_2, B_2) \in E(\mathcal{D}')$. In addition, we have $\hat{i}(C_1, \rho(B_1)) = \hat{i}(C_2, \rho(B_2))$ and $\hat{i}(C_1, \phi(B_1)) = \hat{i}(C_2, \phi(B_2))$ for $C_i \in \{A_i, B_i\}$ and $\hat{i}(B_1, v_1) = \hat{i}(B_2, v_2)$. Since $B_1 = \Omega(v_1)$, then $\hat{i}(B_1, v_1)$ lies strictly between $\hat{i}(B_1, \phi(B_1))$ and $\hat{i}(B_1, \rho(B_1))$, and hence $\hat{i}(B_2, v_2)$ lies strictly between $\hat{i}(B_2, \phi(B_2))$ and $\hat{i}(B_2, \rho(B_2))$, giving $B_2 = \Omega(v_2)$. The property $\bar{\sigma}(\Omega(v_1)) \sim_{ft} \bar{\sigma}(\Omega(v_2))$ follows immediately. Now $A_1 \sim_{ft} A_2$ implies that $k_{A_1} = k_{A_2}$. Lemma 5.1 shows that $d_{S\Gamma}(v_0, \rho(B_i)) = d_{S\Gamma}(v_0, x_{A_i}) - d_{S\Gamma}(x_{A_i}, \rho(B_i))$ for $i = 1, 2$, and similarly for $\phi(B_i)$. Then $d_{S\Gamma}(v_0, \rho(B_1)) - d_{S\Gamma}(v_0, \phi(B_1)) = d_{S\Gamma}(x_{A_1}, \phi(B_1)) - d_{S\Gamma}(x_{A_1}, \rho(B_1)) = |k_{A_1} - \hat{i}(A_1, \phi(B_1))| - |k_{A_1} - \hat{i}(A_1, \rho(B_1))|$. Since all of these numbers are the same if the subscript 1 is replaced by 2 everywhere, then we have $d_{S\Gamma}(v_0, \rho(B_1)) - d_{S\Gamma}(v_0, \phi(B_1)) = d_{S\Gamma}(v_0, \rho(B_2)) - d_{S\Gamma}(v_0, \phi(B_2))$. This shows that $k_{B_1} = k_{B_2}$, which is the last item needed to show that $B_1 \sim_{ft} B_2$. Therefore $v_1 \sim_{ft} v_2$.

Finally, suppose that (iii) $(\Omega(v_1), \Omega(u_1)) \in E(\mathcal{D}')$ and $\bar{\sigma}(\Omega(u_1)) \sim_{ft} \bar{\sigma}(\Omega(u_2))$. Suppose further that $\bar{\sigma}(A_1) = v_0$. Then $v_1 = v_0$ and $A_1 = \Omega(u_1) = F_1$. In this case $\bar{\sigma}(\Omega(u_2)) = v_0$, and so $A_1 = A_2$, $u_1 = u_2$, and the lemma holds.

On the other hand, suppose that $\bar{\sigma}(A_1) = \bar{\sigma}(\Omega(u_1)) \neq v_0$. Then $E_i := \Omega(\bar{\sigma}(A_i)$ is a face of $SC(1)$ for $i = 1, 2$, and we also have $(E_2, A_2) \in \mathcal{D}'$ and $E_1 = \Omega(v_1)$. The definition of $\bar{\sigma}(A_1) \sim_{ft} \bar{\sigma}(A_2)$ implies that $E_1 \sim_{ft} E_2$ and $i(E_1, \bar{\sigma}(A_1)) = i(E_2, \bar{\sigma}(A_2))$. Now the edge gluings in the folding of A_1 onto its parent face E_1 and in the folding of A_2 onto E_2 must be the same. The edge $e_1 = (u_1, x, v_1)$ lies in A_1 with u_1 in $\partial A_1 \setminus \gamma(A_1)$ and v_1 in $\gamma(A_1)$, and there must be a corresponding edge $e_2 = (u_2, x, v_2)$ in the face A_2. Then $i(A_1, v_1) = i(A_2, v_2)$, and so v_2 lies in $\gamma(A_2)$. Hence $E_2 = \Omega(v_2)$. Finally the correspondence in edge gluings together with $i(E_1, \bar{\sigma}(A_1)) = i(E_2, \bar{\sigma}(A_2))$ imply that $i(E_1, v_1) = i(E_2, v_2)$, and so $v_1 \sim_{ft} v_2$. □

Next we use the face type classes of vertices in $SC(1)$ to build a deterministic pushdown automaton, following the notation for a PDA in [HU79, p.110].

Definition 5.4. *Let $\mathcal{P} = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ be the deterministic pushdown automaton with state set $Q = \{[v] \mid v \in V(SC(1))\}$, input alphabet $\Sigma = X \cup X^{-1}$, stack alphabet $\Gamma = \{[v] \mid v \in V(SC(1))\}$, initial state $q_0 = [v_0]$, initial stack symbol $Z_0 = [v_0]$, final (accept) state $F = \{[v_0]\}$, and transition function the partial function $\delta \colon Q \times \Sigma \times \Gamma \to Q \times \Gamma^*$ for which $\delta([u], x, [t])$ is defined only if there is an edge (u, x, v) for some vertex v in $SC(1)$, by*

$$\delta([u], x, [t]) := \begin{cases} ([v], [t]) & \text{if } \Omega(u) = \Omega(v) \\ ([v], [\sigma(\Omega(v))][t]) & \text{if } (\Omega(u), \Omega(v)) \in E(\mathcal{D}') \\ ([v], \epsilon) & \text{if } (\Omega(v), \Omega(u)) \in E(\mathcal{D}'),\ [t] = [\bar{\sigma}(\Omega(u))] \end{cases}$$

The undefined transitions for δ are viewed as going to a fail state. Note that for an edge (u, x, v) in $SC(1)$ satisfying $v_0 \neq v = \bar{\sigma}(\Omega(u)) = \rho(\Omega(u)) = \phi(\Omega(u))$, so that the face $\Omega(u)$ is attached at v but no edges are glued, the last case of the definition of δ can be split into two subcases. In this situation we have $\Omega(u) \sim_{ft} F_1$, and there is an edge (u_1, x, v_0) in F_1 with $u_1 \sim_{ft} u$. If $[t] = [v] \neq [v_0]$, then $\delta([u], x, [t]) := ([v], \epsilon)$, but if $[t] = [v_0]$, then $\delta([u], x, [t]) := ([v_0], \epsilon)$. The fact that δ is well-defined follows directly from Lemma 5.3.

An instantaneous description (α, z, β) for the PDA $\mathcal{P}$ consists of the current state $\alpha \in Q$ of the machine, the word $z \in (X \cup X^{-1})^*$ that remains to be read, and the current contents $\beta \in \Gamma^*$ of the stack, where the first letter of β is the "top" of the stack. We write $(\alpha, yz, \beta) \vdash^* (\alpha', z, \beta')$ if, when y is read in starting from (α, yz, β), the PDA reaches (α', z, β'), and write $\vdash$ when a single letter $y \in X \cup X^{-1}$ is read.

Define a function $\beta : V(SC(1)) \to \Gamma^*$ as follows. Given any vertex v in $SC(1)$, let $v_0, F_1, ..., F_m = \Omega(v)$ be the labels of the vertices along the geodesic path in the tree

$\mathcal{D}'$ from v_0 to $\Omega(v)$. Then $\beta(v)$ is the associated word over the stack alphabet given by $\beta(v) := [\bar{\sigma}(F_m)] \cdots [\bar{\sigma}(F_1)][v_0]$.

Proposition 5.5. *Let w be sparse, and let $SC(1)$ be the Schützenberger complex of 1 for $M = \mathrm{Inv}\langle X \mid w = 1\rangle$. Let $\alpha \in Q$, $y, z \in (X \cup X^{-1})^*$, and $\beta \in \Gamma^*$. Then $([v_0], yz, [v_0]) \vdash^* (\alpha, z, \beta)$ if and only if y labels an edge path in $SC(1)$ starting at v_0 and $\alpha = [v]$ and $\beta = \beta(v)$ where v is the end vertex of this path.*

Proof. First we prove the forward implication, by induction on the length of y. If $l(y) = 0$, then $y = \epsilon$ and $([v_0], \epsilon z, [v_0]) \vdash^* (\alpha, z, \beta)$ implies that $\alpha = [v_0]$, and $\beta = [v_0] = \beta(v_0)$. The path starting at v_0 labeled by $y = \epsilon$ ends at $v = v_0$, as required.

Now, suppose that the forward implication holds for any word $\tilde{y}$ with $0 \le l(\tilde{y}) < l(y)$, and write $y = y'x$ with $x \in X \cup X^{-1}$. Suppose that $([v_0], yz, [v_0]) \vdash^* (\alpha, z, \beta)$. Then we have $([v_0], y'xz, [v_0]) \vdash^* (\alpha', xz, \beta') \vdash (\alpha, z, \beta)$ for some $\alpha' \in Q$ and $\beta' \in \Gamma^*$. By induction, the word y' labels a path π' in $SC(1)$ starting at v_0, and $\alpha' = [u]$ and $\beta' = \beta(u)$ where u is the ending vertex of the path π'.

Since $(\alpha', xz, \beta') \vdash (\alpha, z, \beta)$, the transition function δ is defined on the triple (α', x, γ), where γ is the first letter of the word $\beta(u) \in \Gamma^*$. This means that there is a representative $\tilde{u}$ of the $\sim_{ft}$-class α' such that there is an edge of the form $e = (\tilde{u}, x, v)$ in $SC(1)$ for some vertex v, and either (i) $\Omega(\tilde{u}) = \Omega(v)$, (ii) $(\Omega(\tilde{u}), \Omega(v)) \in E(\mathcal{D}')$, or (iii) $\gamma = [\bar{\sigma}(\Omega(\tilde{u}))]$ and $(\Omega(v), \Omega(\tilde{u})) \in E(\mathcal{D}')$. In cases (i) and (ii), Lemma 5.3 shows that we may take $\tilde{u} = u$. In case (iii), notice that the first letter γ of $\beta(u)$ satisfies $\gamma = [\bar{\sigma}(\Omega(u))]$ if $\Omega(u) \neq v_0$, and $\gamma = [v_0]$ if $\Omega(u) = v_0$. However, if $\Omega(u) = v_0$, then $u = v_0$, and since $[\tilde{u}] = \alpha' = [u]$, then $\tilde{u} = v_0$, contradicting the existence of the edge $(\Omega(v), \Omega(\tilde{u}))$ in $\mathcal{D}'$. Then $\Omega(u) \neq v_0$, and so we also may take $\tilde{u} = u$ in this case.

Then in all three cases, the path π' followed by the edge e is a path in $SC(1)$ labeled by the word y starting at v_0 and ending at the vertex v. Moreover, we have $\alpha = [v]$.

In case (i), $\delta(\alpha', x, \gamma) = ([v], \gamma)$, and the stack word $\beta = \beta' = \beta(u)$ is unchanged by this transition. Since $\Omega(u) = \Omega(v)$, then $\beta = \beta(v)$.

In case (ii), $\delta(\alpha', x, \gamma) = ([v], [\bar{\sigma}(\Omega(v))]\gamma)$, and we have $\beta = [\bar{\sigma}(\Omega(v))]\beta(u)$. Since $(\Omega(u), \Omega(v)) \in E(\mathcal{D}')$, we again have $\beta = \beta(v)$.

In case (iii), $\delta(\alpha', x, \gamma) = ([v], \epsilon)$. Now $(\Omega(v), \Omega(u)) \in E(\mathcal{D}')$ implies that $\beta(u) = [\bar{\sigma}(\Omega(u))]\beta(v)$, and we have $\beta = \beta(v)$ in this case as well.

This completes the proof of the forward implication.

For the reverse implication, we again induct on the length $l(y)$. If $l(y) = 0$, the as before $y = \epsilon$ labels a path from v_0 to v_0, and so $([v_0], yz, [v_0]) \vdash^* (\alpha, z, \beta)$ where $\alpha = [v_0]$ and $\beta = [v_0] = \beta(v_0)$.

Suppose again that $l(y) > 0$ and write $y = y'x$ with $x \in X \cup X^{-1}$. By hypothesis, y labels a path in $SC(1)$ from v_0; let v be the vertex at the end of this path, and let u be the penultimate vertex; that is, u is at the end of the path labeled by y'. By induction we have $([v_0], yz, [v_0]) \vdash^* ([u], xz, \beta(u))$. The definition of δ then shows that $([v_0], yz, [v_0]) \vdash^* ([v], z, \beta(v))$. □

We can now prove Theorems 1.3 and 1.4.

Proof of Theorem 1.3

For a word $y \in (X \cup X^{-1})^*$, we have $y = 1$ in M if and only if y labels an edge path from v_0 to v_0 in $S\Gamma(1)$. Proposition 5.5 shows that the latter holds if and only if $([v_0], y, [v_0]) \vdash^* ([v_0], \epsilon, \beta)$ for some β; that is, exactly when the PDA $\mathcal{P}$ finishes in the accept state $[v_0]$. Thus, the set of words representing the identity element in M is a deterministic context-free language.

The word y is in the language of words related to 1 in M by Green's relation $\mathcal{R}$ if and only if y labels a path starting at v_0 in $S\Gamma(1)$, which holds if and only if $([v_0], y, [v_0]) \vdash^* (\alpha, \epsilon, \beta)$ for some $\alpha \in Q$ and $\beta \in \Gamma^*$. Let $\mathcal{P}'$ be the PDA $\mathcal{P}$ with the the set of final (accept) states changed to $F = Q$. Then we have y is accepted by $\mathcal{P}'$ if and only if y is in the $\mathcal{R}$-equivalence class $\mathcal{R}_1$ of 1. Hence the set of words representing an element of $\mathcal{R}_1$ in M is also a deterministic context-free language. □

Proof of Theorem 1.4

Let $(Q, \Sigma, \delta, q_0, F)$ be the finite state automaton with state set $Q = \{[v] \mid v \in V(SC(1))\}$, input alphabet $\Sigma = X \cup X^{-1}$, initial state $q_0 = [v_0]$, final (accept) states $F = Q$, and transition function the partial function $\delta \colon Q \times \Sigma \to Q$, defined by $\delta([u], x) := [v]$ if there is an edge (u, x, v) in $SC(1)$ and either
(i) $\Omega(u) = \Omega(v)$ and either $\hat{i}(\Omega(u), u) < \hat{i}(\Omega(u), v) \leq k_{\Omega(u)}$ or $\hat{i}(\Omega(u), u) > \hat{i}(\Omega(u), v) \geq k_{\Omega(u)}$, or
(ii) $(\Omega(u), \Omega(v)) \in E(\mathcal{D}')$
Lemma 5.3 shows that this transition function is well-defined.

Let p be an arbitrary path in $SC(1)$ starting at v_0. Let $v_0, v_1, ..., v_m$ be the sequence of consecutive vertices traversed by p, and let $A_i := \Omega(v_i)$. Note that the path p is geodesic if and only if $d_{S\Gamma}(v_0, v_{i-1}) > d_{S\Gamma}(v_0, v_i)$ for all i.

If $(A_i, A_{i-1}) \in E(\mathcal{D}')$, then the Geodesic Theorem 4.4 says that p is not a geodesic. If $(A_{i-1}, A_i) \in E(\mathcal{D}')$, then $v_i \in \partial A_i \setminus \gamma(A_i) = \partial A_i \setminus A_{i-1}$, and the vertex v_{i-1} must be one of the endpoints $\rho(A_i), \phi(A_i)$ of the gluing path of A_i onto A_{i-1}; let u_i be the other. The Geodesic Theorem 4.4 says that any geodesic from v_0 to v_i must pass through one of the points v_{i-1}, u_i. Since $d_{S\Gamma}(v_{i-1}, v_i) = 1$, then Lemma 4.3(1) shows that such a geodesic must also pass through v_{i-1}. Hence $d_{S\Gamma}(v_0, v_i) > d_{S\Gamma}(v_0, v_{i-1})$. Finally, if $A_{i-1} = A_i$, then v_{i-1} and v_i are both vertices in $\partial A_i \setminus \gamma(A_i)$. By Lemma 5.1, it follows that $d_{S\Gamma}(v_0, v_i) > d_{S\Gamma}(v_0, v_{i-1})$ if and only if either $i(A_i, v_{i-1}) < i(A_i, v_i) \leq k_{A_i}$ or $i(A_i, v_{i-1}) > i(A_i, v_i) \geq k_{A_i}$.

In the proof of Theorem 1.3, we showed that a word y labels a path starting at v_0 in $SC(1)$ if and only if it is accepted by the PDA $\mathcal{P}'$, which is the PDA in Definition 5.4 but for which all states in Q are final (accept) states. Note that the only transitions of this PDA which utilize the stack in determining the next state are those associated with edges from u to v with $(\Omega(v), \Omega(u)) \in E(\mathcal{D}')$. Combining this with the previous paragraph, then, the finite state automaton defined above is precisely the underlying finite state automaton of the PDA $\mathcal{P}'$ consisting only of transitions associated with edges (u, x, v) such that $d(v_0, v) > d(v_0, u)$. Thus this finite state automaton accepts precisely the words which label geodesic paths in $SC(1)$. □

Remark 1. The minimized form of the finite state automaton defined in the proof of Theorem 1.4 is the automaton of cone types of $S\Gamma(1)$. As an example, S. Haataja showed that the automaton of cone types for $S\Gamma(1)$ for the sparse word $w = aba^{-1}b^{-1}cdc^{-1}d^{-1}$ corresponding to the surface group of genus 2 has 19 cone types (unpublished manuscript). A description of Haataja's example may be found in Meakin's survey article [Me07].

Remark 2. Descriptions of an iterative construction of the PDA in Definition 5.4 and an implementation of the algorithm for solving the word problem is provided in S. Lindblad's PhD thesis [Lin03]. The software is available from http://www.math.unl.edu/ shermiller2/lindblad/ .

Remark 3. In their paper [IMM01], Ivanov, Margolis and Meakin show that the word problem for the inverse monoid $M = \mathrm{Inv}\langle X \mid w = 1\rangle$ corresponding to a cyclically reduced word w is solvable if the membership problem for the submonoid of the corresponding one-relator group $G = Gp\langle X|w = 1\rangle$ generated by the prefixes of w is solvable. However as far as we are aware, it is not known whether the prefix membership problem for this submonoid of G is equivalent to the word problem for M in general. In particular, it is not known whether this prefix membership problem for G is solvable if w is a sparse word.

References

[BMM94] Birget, J.-C., Margolis, S. W. and Meakin, J. C., *The word problem for inverse monoids presented by one idempotent relator*, Theoret. Comput. Sci. **123** (1994), no. 2, 273–289.

[CP67] Clifford, A. H., Preston, G. P. *The Algebraic Theory of Semigroups*, Math. Surveys 7, Amer. Math. Soc., Providence 1961 (Vol I) and 1967 (Vol. II).

[HU79] Hopcroft, J. E. and Ullman, J. D., *Introduction to automata theory, languages, and computation*, Addison-Wesley Publishing Co., Reading, Mass., 1979, Addison-Wesley Series in Computer Science.

[IMM01] Ivanov,S. V., Margolis, S. W. and Meakin, J. C., *On one-relator inverse monoids and one-relator groups*, J. Pure Appl. Algebra **159** (2001), no. 1, 83–111.

[Law98] Lawson, M. V., *Inverse semigroups*, World Scientific Publishing Co. Inc., River Edge, NJ, 1998.

[Lin03] Lindblad, S P., *Inverse monoids presented by a single sparse relator*, PhD Thesis, Univ. of Nebraska-Lincoln, 2003.

[Me07] Meakin, J C., *Groups and semigroups: connections and contrasts*, Proceedings, Groups St Andrews 2005, London Math. Soc. Lecture Note Series 340, Vol 2, 2007, 357-400.

[MMS05] Margolis, S., Meakin, J. and Šuniḱ, Z., Distortion functions and the membership problem for submonoids of groups and monoids, *Contemporary Mathematics* **372** (2005) 109-129.

[Mun74] Munn, W. D., *Free inverse semigroups*, Proc. London Math. Soc. **29** (1974), no. 3, 385–404.

[Step90] Stephen, J. B., *Presentations of inverse monoids*, J. Pure Appl. Algebra **63** (1990), no. 1, 81–112.

[Step93] Stephen, J. B., *Inverse monoids and rational subsets of related groups*, Semigroup Forum **46** (1993), no. 1, 98–108.

[Stei03] Steinberg, B.,*A topological approach to inverse and regular semigroups*, Pacific J. Math. **208** (2003), no. 2, 367–396.

Susan Hermiller, Department of Mathematics, University of Nebraska, Lincoln NE 68588-0130, USA
E-mail address: smh@math.unl.edu

Steve Lindblad, Hewitt Associates LLC, 45 South 7th Street, Suite 2100, Minneapolis MN 55402, USA
E-mail address: splindblad@gmail.com

John Meakin, Department of Mathematics, University of Nebraska, Lincoln NE 68588-0130, USA
E-mail address: jmeakin@math.unl.edu