A combinatorial algorithm for immersed loops in surfaces

J.M. Paterson

DPMMS, 16 Mill Lane, Cambridge, England, UK, CB2 1SB
Received 30 March 1999; received in revised form 31 March 2001

Abstract

In this paper, we develop a purely combinatorial algorithm which minimizes the number of double points of an immersed loop in a closed, orientable surface and converts between the ambient isotopy classes of two homotopic loops using an explicit sequence of elementary homotopies. We note that by introducing a curve-shortening flow known as the disc flow, Hass and Scott have shown that given a pair of general position, immersed loops, each with \( k \) double points, then they are homotopic through loops with at most \( k \) double points, and that this homotopy may be assumed to be regular except at finitely many points. We demonstrate this here without recourse to the geometry of the surface, by giving an explicit homotopy, which relies solely on the notion of a spanning disc for an immersed loop, which we define here to be an embedding of the standard 2-disc into the universal covering space of the surface, for which the boundary is mapped into the union of the lifts of the loop. The most important of these spanning discs have a 1-gon, 2-gon or 3-gon structure relative to the family of covering curves for the loop.

AMS classification: 57N05; 57N10

Keywords: Surfaces; Immersions; Loops; Combinatorial

1. Introduction

In this paper, we take an arbitrary general position immersion of an oriented loop into a closed, orientable surface and present a combinatorial algorithm which homotops the loop to one with the minimal number of double points using only local homotopies of an especially simple form. We then generalize this approach by developing an algorithm which, given two homotopic immersed loops, homotops the first loop through a sequence of these elementary moves to produce a loop which is ambient isotopic to the second.
In [4], Hass and Scott gave shown, by means of techniques which relied intensively on the geometry of the surface, that it is possible to move between a pair of homotopic loops each with \( k \) double points for some \( k \in \mathbb{N} \), through maps with at most \( k \) double points. In particular, they constructed a piecewise-linear analogue of curvature flow which shortened a curve by successively replacing segments with geodesics. They were not, however, able to write down an explicit algorithm which achieved this, using information purely relating to the combinatorial arrangement of self-intersections of the loop on the immersed surface.

We develop an algorithm here which does precisely this, with the added benefit of being very simple to implement. For this reason it should be more directly applicable to other combinatorial problems in two and three-dimensional manifold theory. We remark, moreover, that this procedure is intimately related to the “football-reduction” and “football-exchange” procedures, which we develop in a forthcoming paper [5].

Our approach encodes the spanning discs for a general position immersed loop into a graph which we call the state graph for the immersed loop. Its edges capture the relative arrangement of these spanning discs in terms of special sorts of adjacencies and inclusions. The inclusions give us a natural partial ordering, so that the “higher” up a chain a spanning sub-disc is represented, the more spanning sub-discs it has. Our techniques then hinge upon analysing the way in which performing an elementary local homotopy effects this graph.

Throughout this paper, we shall take \( \Sigma \) to be a closed, orientable surface other than \( S^2 \) and \( c: S^1 \to \Sigma \), to be a general position immersion. We follow the convention of [3] and use \( c \) to represent both the map and its image. Writing \( \tilde{\Sigma} \) for the universal covering space of \( \Sigma \), which is, in this case, a copy of \( \mathbb{R}^2 \), we observe that \( c \) lifts to a collection of general position immersed curves in \( \tilde{\Sigma} \). We denote this covering family by \( \Lambda \). If \( c \) is null-homotopic, then \( \Lambda \) consists of simple closed curves in \( \tilde{\Sigma} \). Otherwise, \( \Lambda \) consists of non-compact curves which we refer to as lines. We develop our algorithm on the basis of the following key concept, where \( D^2 \) denotes the closed 2-disk.

**Definition 1.1.** Suppose that \( \Sigma \) is a closed, orientable surface and that \( c: S^1 \to \Sigma \), is a general position immersion. We shall say that an embedding,

\[
e: D^2 \to \tilde{\Sigma},
\]

is a spanning disc for \( c \) if \( \partial e(D^2) \) is contained in \( \bigcup \Lambda \).

In practice, we refer to the image, \( \Delta = e(D^2) \), as a spanning disc for \( c \). Note that the boundary \( \partial e(D^2) \) is an embedded copy of \( S^1 \) in \( \tilde{\Sigma} \). Moreover, each such disc inherits a natural polygonal structure from \( \bigcup \Lambda \). In particular, a vertex of \( \Delta \) is a double point for \( \Lambda \) which is the intersection of two edges of \( \Delta \). An edge of \( \Delta \) is a maximal (connected) segment of \( \partial \Delta \), which is contained within a single curve in the covering family, \( \Lambda \), and has no vertices in its interior. We shall say that a spanning disc, \( \Delta \) is \( \pi_1(\Sigma) \)-equivariant if it is disjoint from \( g\Delta \), for every \( g \in \pi_1(\Sigma) \). A spanning disc, \( \Delta \) is innermost if \( \text{Int}\Delta \) is disjoint from \( \bigcup \Lambda \). We note that an innermost spanning disc need not be \( \pi_1(\Sigma) \)-equivariant and vice versa. In particular, a pair of innermost spanning discs, \( \Delta \) and \( g\Delta \), for some \( g \in \pi_1(\Sigma) \), may share a common vertex. Restricting our attention to innermost, \( \pi_1(\Sigma) \)-
equivariant spanning discs, we define the following three fundamental local homotopies of \( c \).

1. If \( \Xi \) is a \( \pi_1(\Sigma) \)-equivariant spanning 1-gon disc for \( c \), then we define the 1-gon move, \( \xi \), by equivariantly shrinking the disc until we eliminate it.

2. If \( \Gamma \) is a \( \pi_1(\Sigma) \)-equivariant 2-gon disc, we define the 2-gon move, \( \gamma \), by equivariantly deforming one edge across the other to remove the disc.

3. If \( \Delta \) is a 3-gon disc, we define the 3-gon move, \( \delta \), by equivariantly deforming an edge of the disc across the unique vertex outside of that edge, giving rise to a new 3-gon disc. We note that this disc may be thought of as having the opposite orientation to the original one.

We note that these moves bear some similarity to the well-known Reidemeister moves of knot theory. The key difference, however, is that these moves may always be performed across an equivariant spanning 1-gon, 2-gon or 3-gon, respectively. This is not true for Reidemeister moves which are constrained by the 3-dimensionality of the knot problem. We illustrate these in Fig. 1. Using these ideas, we prove:

**Theorem 1.1.** Let \( c : S^1 \to \Sigma \) be a general position immersion and \( \Sigma \), a closed, orientable surface. Then the set of spanning discs for \( c \) defines a sequence of elementary moves which converts \( c \) into an immersion with the minimal number of self-intersection points in its homotopy class. Furthermore, this conversion sequence is monotonically decreasing with respect to self-intersection number.

**Theorem 1.2.** Let \( c : S^1 \to \Sigma \) and \( c' : S^1 \to \Sigma \) be a pair of homotopic, general position immersed loops in a closed, orientable surface, \( \Sigma \). Moreover, suppose that both \( c \) and \( c' \) have precisely \( k \) self-intersection points, this being the minimal number for a general position immersion in the given homotopy class. Then the spanning discs for \( c \) and \( c' \) define a sequence of elementary moves which converts \( c \) to an immersion which is ambient.
isotopic to \( c' \). Moreover, the curve’s self-intersection number remains equal to \( k \) throughout the conversion sequence.

2. Immersed loops in surfaces

Suppose that \( c : S^1 \to \Sigma \) is a general position immersion as above with a covering family, \( \Lambda \), in \( \tilde{\Sigma} \). We shall say that a pair of lines in \( \Lambda \) is linked if they meet in an odd number of points, see [1]. Suppose now that we have two homotopic, general position immersed loops, \( c \) and \( d \). We identify the members of the two covering families of lines in \( \tilde{\Sigma} \) with a natural bijection as follows.

**Definition 2.1.** Suppose that \( c : S^1 \to \Sigma \) and \( c' : S^1 \to \Sigma \) are homotopic essential immersions and \( h \) is a homotopy taking \( c \) to \( c' \) which lifts to a homotopy \( \tilde{h} \), in \( \tilde{\Sigma} \) taking the line \( l \) above \( c \) to the line, \( l' \), above \( c' \). We shall say that each line, \( gl \), above \( c \), where \( g \in \pi_1(\Sigma) \), corresponds to \( gl' \) above \( c' \). Using this, we define a correspondence bijection,

\[
\Phi : \Lambda \to \Lambda',
\]

\[
: gl \to gl',
\]

\[g \in \pi_1(\Sigma)\].

We observe that a pair of linked lines in \( \Lambda \) corresponds under \( \Phi \) to a linked pair in \( \Lambda' \), for any immersion, \( c' : S^1 \to \Sigma \) in the homotopy class of \( c \). It is further easy to see that if a pair of lines intersect infinitely often, then they are stabilized by an infinite cyclic subgroup, \( \langle g \rangle \), of \( \pi_1(\Sigma) \) and the number of \( \langle g \rangle \)-orbits of the points of intersection is finite. We note also that the lines in \( \Lambda \) may have points of transverse self-intersection.

Recall from [3], that a general position immersion, \( c \), satisfies the 1-point intersection property if the covering family, \( \Lambda \), consists of embedded lines, any intersecting pair of which meet transversely in a single point. In [3], however, Hass and Scott constructed examples of immersed loops which cannot be homotoped to have the 1-point property. These are, in fact, curves which carry non-primitive elements of \( \pi_1(\Sigma) \). A weaker property therefore has more practical value.

**Definition 2.2.** A general position immersion, \( c : S^1 \to \Sigma \), has the minimal intersection property if it has the minimal number of self-intersection points of any general position immersed loop in its homotopy class.

The simplest example of a loop which cannot be homotoped to have the 1-point property has two lines, \( l \) and \( hl \), in \( \Lambda \), which meet transversely in an infinite number of points. These lifts cover a loop with a single point of transverse self-intersection in the cyclic covering space, \( \Sigma_h \), of \( \Sigma \) corresponding to the sub-group, \( \langle h \rangle \), of \( \pi_1(\Sigma) \). We may generalize this to a situation involving an \( n \)-tuple of distinct lines, \( \langle l, hl, h^2l, \ldots, h^{n-1}l \rangle \), for which \( |l \cap h^k l| = \infty, k = 1, \ldots, n - 1 \), and \( h^n = 1 \), \( n \geq 2 \). If, in addition, there are two \( \langle h \rangle \)-orbits
of these intersection points for each $k$ where $n \geq 3$ and one if $n = 2$, then standard cut-and-paste arguments tell us that this is the minimal number of transverse intersections. We note that a feature such as this in $\Lambda$ occurs precisely when $c$ winds $n$ times around a primitive loop in $\Sigma$. We refer to a feature of this form as an $n$-strand in $\Lambda$. If an $n$-strand is disjoint from its translates, then the primitive loop around which $c$ winds may be embedded in the surface $\Sigma$. We review next some important definitions from [3].

**Definition 2.3.** If $c : S^1 \to \Sigma$ is a general position immersion, then $c$ has a singular 1-gon if there is a sub-arc, $\alpha$, of $S^1$ such that the map $c$ identifies the endpoints of $\alpha$ and $c|_\alpha$ is a null-homotopic loop in $\Sigma$. Moreover, a disc $D$ is an embedded 1-gon for $c$ if there exists a sub-arc, $\alpha$, of $S^1$ with $c(\alpha) = \partial D$ and the map $c|_\alpha$ injects. Similarly, $c$ has a singular 2-gon if there exist disjoint sub-arcs, $\alpha$ and $\beta$, of $S^1$ such that $c$ cyclically identifies their endpoints and $c|_{\alpha \cup \beta}$ is a null-homotopic loop in $\Sigma$. Moreover, a disc $D$ is an embedded 2-gon for $c$ if there exist disjoint sub-arcs, $\alpha$ and $\beta$, of $S^1$ which embed under $c$, where $c(\alpha) \cup c(\beta) = \partial D$ and $c(\alpha) \cap c(\beta) = c(\partial \alpha) = c(\partial \beta)$. Finally, $c$ has a singular 3-gon if there exist disjoint sub-arcs, $\alpha$, $\beta$ and $\gamma$, of $S^1$ such that $c$ cyclically identifies their endpoints and $c|_{\alpha \cup \beta \cup \gamma}$ is a null-homotopic loop in $\Sigma$. Moreover, a disc $D$ is an embedded 3-gon for $c$ if there exist disjoint sub-arcs, $\alpha$, $\beta$ and $\gamma$ of $S^1$ which embed under $c$ for which $c(\alpha) \cup c(\beta) \cup c(\gamma) = \partial D$ and each of the sets, $c(\alpha) \cap c(\beta)$, $c(\beta) \cap c(\gamma)$ and $c(\alpha) \cap c(\gamma)$ consists of a single point.

**Remark 1.** If $c$ has an embedded $k$-gon, $D$, $k = 1, 2, 3$, then $D$ lifts to a $\pi_1(\Sigma)$-equivariant spanning $k$-gon for $c$ in $\tilde{\Sigma}$.

We next paraphrase Theorems 4.2 and 2.7 of [3] as follows:

**Theorem 2.1** (Hass and Scott). Suppose that $c$ is a general position immersion of $S^1$ into the closed orientable surface $\Sigma$. If $c$ does not have the minimal intersection property, then it has a singular 1-gon or 2-gon. Furthermore, if $c$ is homotopic to an embedded loop, then $c$ has an embedded 1-gon or 2-gon.

It follows from this that any null-homotopic loop in $\Sigma$ has a $\pi_1(\Sigma)$-equivariant spanning 1-gon or 2-gon across which an elementary move may be performed, thereby reducing the number of self-intersection points. Given a singular loop which is homotopic to an embedding, this gives us a natural procedure for achieving an embedded loop via a sequence of 1-gon and 2-gon moves. Our chief interest, therefore, lies with essential loops which are not homotopic to embeddings. We note that in [3], Hass and Scott give some explicit constructions of immersed loops without the minimal intersection property which have no embedded 1-gon or 2-gon discs.

**Lemma 2.1.** Suppose that $c : S^1 \to \Sigma$ is a general position, essential immersion and $\Sigma$ is a closed, orientable surface. If $c$ has a spanning 1-gon, 2-gon or 3-gon, then there exists an innermost spanning $n$-gon for some $n \leq 3$. 
Proof. By a traversing segment of an $n$-gon disc, we mean an immersion of the closed unit interval into the $n$-gon where the endpoints are mapped into the boundary. Using this, we proceed case-wise.

(1) The loop $c$ has a spanning 1-gon: We choose a sub-1-gon which is small. It follows that any traversing segment must be embedded. Then either the 1-gon is innermost and the conclusion is immediate or it has a sub-2-gon. In the latter case we have the situation of 2 below.

(2) The loop $c$ has a spanning 2-gon: We choose a sub-2-gon which is small. If this disc is non-innermost, then it has either a singular or an embedded traversing segment. In the former case, it has a sub-1-gon and we argue as in 1 above. In the latter case, we have a sub-3-gon and we deal with this in 3 below.

(3) The loop $c$ has a spanning 3-gon: We choose a sub-3-gon which is small. If this disc is non-innermost, we consider, as in 2, the possibilities of singular or embedded traversing segments. In the event of the former, we have a sub-1-gon and argue as in case 1. In the event of the latter there must be a sub-2-gon and we argue as in 2.

By compactness of the original $n$-gon, this process must terminate with a spanning disc which is a 1-gon, 2-gon or 3-gon, as claimed. "}

Suppose, in the sequel, that $c$ is an oriented loop. We define the positive side of $c$ to be the side for which the normal to the curve is inward-pointing. We label the closed half-planes defined by a line, $l_s$, in $\Lambda$ by $l^{+1}_s$ and $l^{-1}_s$, where the superscripts are consistent with our orientation on $c$. Using this, the spanning discs for $c$ may be expressed naturally as components of the intersection of closed half-planes in $\tilde{\Sigma}$.

**Definition 2.4.** Suppose that $\Delta$ is a spanning $n$-gon disc for $c$ with boundary segments in the $n$-tuple of lines, $(l_1, l_2, \ldots, l_n)$ in $\Lambda$, where two or more of these components may coincide. We shall say that the ordered $n$-tuple, $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ in $\{-1, +1\}^n$ is the $(l_1, l_2, \ldots, l_n)$-intersection index of $\Delta$ if $\Delta$ is a component of

$$\bigcap_{i=1}^{n} l^{\varepsilon_i}_i.$$  \hfill (2.3)

If $c$ has the 1-point property, then we shall say that $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ is the intersection index for the $n$-tuple $(l_1, l_2, \ldots, l_n)$.

We note that it is often convenient, when the $n$-tuple of lines is clear from the context, to refer to the intersection index of a spanning $n$-gon disc. In particular, we fix an ordering of these $n$ lines for any given disc to ensure uniqueness. We note that intersection index is a $\pi_1(\Sigma)$-equivariant quantity, that is it is invariant under the action of $\pi_1(\Sigma)$.

**Remark 2.** If $c$ has the 1-point property, then the intersection index of a triple of mutually intersecting lines is invariant under transverse homotopy up to negation in each component. This corresponds to the two possible configurations for a spanning 3-gon with boundary segments in these lines.
We note that a 3-gon move produces a 3-gon disc with intersection index the negative of that of the original disc.

Suppose next that \(c\) and \(c'\) are homotopic essential immersions. We would like to compare the spanning discs for \(c\) with those of \(c'\). In order to do this, we attempt to pair spanning discs for \(c\) and \(c'\) in a natural way.

**Definition 2.5.** Suppose that \(\Delta\) is a spanning \(n\)-gon for \(c\) which is a component of \(\bigcap_{i=1}^{n} t_i^{\varepsilon_i}\). We shall say that \(\Delta\) properly corresponds to a spanning \(n\)-gon, \(\Delta'\), for \(c'\) if \(\Delta'\) is a component of \(\bigcap_{i=1}^{n} t_i^{\varepsilon_i}\), where \(t_i'\) corresponds to \(t_i\). If \(\Delta\) is a 3-gon in \(\bigcap_{i=1}^{3} t_i^{\varepsilon_i}\), then we shall say that \(\Delta\) reverse corresponds to \(\Delta'\) if \(\Delta'\) is a component of \(\bigcap_{i=1}^{3} t_i'^{-\varepsilon_i}\).

We note that if \(c\) has the 1-point property, then this relation of correspondence defines a bijection between the spanning discs for \(c\) and those for \(c'\). In general, however, this is not the case and attempts to define an injective correspondence leave us with 1-gon, 2-gon or 3-gon discs in \(\tilde{\Sigma}\) with no counterparts in \(\tilde{\Sigma}'\) and vice versa. Likewise, a 1-gon, 2-gon or 3-gon disc for \(c\) may have a number of corresponding discs in \(\tilde{\Sigma}'\). We note, moreover, that if this relation is bijective on the 1-gons, 2-gons and 3-gons and all the correspondences of 3-gons are proper, then \(c\) and \(c'\) are ambient isotopic. Attempting to construct a bijection in this way to decide whether two immersions are ambient isotopic appears at first to present a rather daunting task. By encoding the information of inclusions and adjacencies of spanning discs into a graph, the problem reduces to recognizing isomorphisms of graphs. This graph has the practical advantage of determining all the explicit sequences of elementary moves which change the ambient isotopy class of an immersed loop. In order to build this graph, we start with the following definitions.

**Definition 2.6.** We shall say that a pair of spanning discs, \(\Delta\) and \(\Gamma\), are **vertex adjacent** if they have opposite angles at some common vertex. Likewise, we shall say that a pair of spanning discs are **edge adjacent** if they locally oppose across some common boundary edge sub-segment. A spanning disc, \(\Delta\), **vertex includes** a spanning disc, \(\Gamma\), if \(\Delta\) contains \(\Gamma\) and the two discs share a common vertex. A disc, \(\Delta\), **edge includes** \(\Gamma\) if the \(\Delta\) contains \(\Gamma\) and they share a common edge sub-segment of their boundaries but \(\Delta\) does not vertex include \(\Gamma\). We shall refer to inclusion which is neither of the vertex or edge type as **interior inclusion**.

We note that where a disk includes another disk it cannot be adjacent to that disk and vice versa. We further note that a pair of 1-gons cannot be vertex adjacent if the loop is essential but that they may be edge adjacent. We illustrate these inclusions and adjacencies with examples in Fig. 2 below.

**Definition 2.7.** Let \(c: S^1 \to \Sigma\) be an oriented immersion. We define a graph, \(\tilde{S}(c)\), of \(c\) the vertices of which are the 1-gon, 2-gon and 3-gon spanning discs for \(c\) in \(\tilde{\Sigma}\) and the
edges of which are defined as follows. If $\Delta_1$ and $\Delta_2$ are vertex adjacent spanning discs for $c$, then we write
\[ \Delta_1 \leftrightarrow \Delta_2. \] (H1)
If $\Delta_1$ and $\Delta_2$ are edge adjacent and $\Delta = \Delta_1 \cup \Delta_2$, then we write
\[ \Delta_1 \iff \Delta_2. \] (H2)
We say, moreover, that $\Delta_1$ and $\Delta_2$ are paired in $\Delta$, and that $\Delta_1$ is the partner of $\Delta_2$ in $\Delta$.
If $\Delta_1$ vertex includes $\Delta_2$, then we write
\[ \Delta_1 \downarrow \Delta_2. \] (V1)
If $\Delta_2$ is a 2-gon which is edge included by $\Delta_1$, then we also define an edge of the form (V1) above. Similarly, if $\Delta_2$ is a 1-gon and the inclusion is interior inclusion. Finally, if $\Delta_1$ is a 2-gon, $\Delta_2$ is a 1-gon and the pair show the special case of vertex inclusion shown in Fig. 3, then we define an edge
\[ \Delta_1 \Downarrow \Delta_2. \] (V2)
We then define the state graph, $S(c)$, to be the quotient graph of $\widetilde{S}(c)$ under the action of $\pi_1(\Sigma)$. Hence the vertices of $S(c)$ are the $\pi_1(\Sigma)$-orbits of the spanning 1-gon, 2-gon and 3-gon discs for $c$. Representing these vertices by a collection, $D(c)$, of spanning discs, we may assign the superscript $(g)$, $g \in \pi_1(\Sigma)$, to a vertical edge of $S(c)$ if the edge results from an inclusion of $g$. $\Delta_2$ by $\Delta_1$ in $\widetilde{S}(c)$, where $\Delta_1$ and $\Delta_2$ are discs in $D(c)$. Similarly, we may add a superscript $(g) \rightarrow$ to a horizontal edge resulting from the adjacency of $\Delta_1$ and $g \Delta_2$ in $\widetilde{S}(c)$, $\Delta_1, \Delta_2 \in D(c)$. We refer to the edges of resulting from type (H1) and (H2) edges in $\widetilde{S}(c)$ as horizontal edges in $S(c)$ and those of type (V1) and (V2) as vertical edges in $S(c)$. If the group element on an edge is the identity for our choice of $D(c)$, we omit it.

Fig. 2. Inclusions and adjacencies.
We note that inclusions other than those of type (V1) and (V2) above are omitted from $S(c)$. In general, when representing sub-graphs of $S(c)$, we choose $D(c)$ so that some vertical or horizontal edge carries the identity element in $\pi_1(\Sigma)$. Suppose that we have constructed $S(c)$ in the above way. We observe that a type (H1) edge may connect a vertex, $v_i$, to itself, although orientation arguments prevent similar behaviour with a type (H2) edge. We denote this feature by $v_i^{(g)}$, where the superscript $(g)$ refers to an edge of the form $\rightarrow \leftarrow$. We refer to $v_i^{(g)}$ as a unit circuit on $v_i$. We note that a unit circuit on a 2-gon type vertex occurs precisely when we have a pair of lifts, $l$ and $gl$, of $c$ which meet transversely in an infinite number of points, all of which project to a single point in the covering space corresponding to the infinite cyclic group, $(g)$. We further note that a unit circuit on a 3-gon type vertex has a representative disc, $\Delta$, with boundary segments in lines of the form $l, gl, g^2l$, where $g \in \pi_1(\Sigma)$ and the $(l, gl, g^2l)$-intersection index of $\Delta$ is either $(1, -1, 1)$ or $(-1, 1, -1)$. It is easy to see that a 1-gon vertex cannot support a unit circuit. We note, moreover, that a 3-gon unit circuit gives us the sole example of a non-equivariant, yet innermost, spanning 3-gon.

**Remark 3.** The graph $S(c)$ is uniquely determined by the ambient isotopy class of $c$ up to the $\pi_1(\Sigma)$ superscripts which vary according to the choice of $D(c)$. An alternative viewpoint considers $S(c)$ as the quotient of a graph whose vertices are all the spanning discs for $c$ by the action of $\pi_1(\Sigma)$.

**Lemma 2.2.** The state graph, $S(c)$, of a general position, essential immersion, $c : S^1 \rightarrow \Sigma$ is a finite graph.

**Proof.** Since $c$ has only finitely many double points, there must exist maximal spanning 1-gon and 2-gon discs with respect to inclusion. Furthermore, there must exist innermost 1-gon, 2-gon and 3-gon discs as $\Sigma$ is compact, see [3]. It remains, however, to exclude the possibility of an infinite sequence of nested 3-gons. We start by considering the case where the genus of $\Sigma$ is at least two. Then we may equip $\tilde{\Sigma}$ with a hyperbolic metric. Given the finite number of double points of $c$, the infinite collection of lines must partition into three finite collections, together with their translates by three infinite cyclic groups. These cyclic groups are the stabilizer sub-groups of the three lines which bound the smallest 3-gon in the
chain of inclusions. These groups, however, must be one and the same in order to allow this infinite nesting. To see this, we observe that the members $\pi_1(\Sigma)$ act as loxodromic actions in the hyperbolic plane from which it follows easily that the configuration is impossible. It therefore remains to consider the case where $\Sigma$ is a torus. Even then, however, the configuration in question cannot result from lifts of single curve but requires instead three curves where each has lifts in precisely one of the three families defined above. \hfill $\blacksquare$

We next describe some of the important features of $S(c)$.

**Definition 2.8.** A vertical chain in $S(c)$ headed by a vertex, $v_i$, in $S(c)$, is a sub-graph consisting of vertices, $v_{i_0} = v_i, v_{i_1}, \ldots, v_{i_n}$ and vertical edges, such that each $v_{i_k}$ lies beneath $v_{i_{k-1}}, k \in \mathbb{N}$. The shadow, $s_i$, of a vertex, $v_i$, in $S(c)$ is the full sub-graph of $S(c)$ on the union of all vertical chains headed by $v_i$.

A vertex $v_i$ in $S(c)$ is minimal if there is no vertex $v_j$ such that a vertical edge runs from $v_i$ to $v_j$. For example, given a 3-gon $\Delta$, any segment which crosses $\Delta$ parallel to an edge of $\Delta$ produces a new 3-gon, $\Delta_i$, which is $(V1)$-included by $\Delta$. Hence the vertex $v$ corresponding to $\Delta$ cannot be minimal in $S(c)$. The vertical edges give us a natural partial ordering on the vertices of $S(c)$. In particular, a minimal vertex in $S(c)$ with respect to this ordering is the $\pi_1(\Sigma)$-orbit of an innermost spanning disc for $c$. We note, moreover, that the vertices in a vertical chain in $S(c)$ may be represented by nested spanning discs for $c$.

**Example 2.1.** We illustrate a possible local arrangement of spanning discs in $\tilde{\Sigma}$, in Fig. 4, together with the associated state sub-graph. The spanning discs may be described by the following counter-clockwise vertex triples: $\Delta_{i1}: (P_1, P_2, P_3), \Delta_{i2}: (P_1, P_3, P_6), \Delta_{i3}: (P_1, P_4, P_7), \Delta_{i4}: (g.P_1, P_4, P_4), \Delta_{i5}: (g.P_1, P_7, P_6)$.
Definition 2.9. We shall say that a sub-graph of $S(c)$ with vertex set \{v_1, \ldots, v_n\} is a \textit{length $n$ vertical circuit} if there exists a choice for $D(c)$ which produces a sub-graph with one of the following $n$ configurations:

\[
\begin{align*}
&v_{i_1} \downarrow \\
&(g) \quad \vdots \quad \downarrow \\
&v_{i_n},
\end{align*}
\]  

(2.4)

which we call a \textit{standard form vertical circuit}, or one of the form

\[
\begin{align*}
v_{i_1} \downarrow \quad (g) \quad \vdots \quad \downarrow \\
&(v_{i_n} \quad \vdots \\
v_{i_{n-k}} \quad \vdots \\
v_{i_{n-k+1}},
\end{align*}
\]

where $k \in \{1, \ldots, n-1\}$. We denote the standard vertical circuit, 2.4, by $v_{i_1/n}^k$. In each of these situations, we shall say that the vertex $v_{i_1}$ \textit{heads} the vertical circuit.

An example of a vertical circuit in which the turning vertex is a 2-gon is shown in Fig. 5, where the 2-gon $\Gamma$ has vertices $gP_2$ and $Q$, the 3-gon $\Delta$ has vertices $(P_1, P_2, P_3)$ and the 3-gon $\Delta_1$ has vertices $(P_4, P_2, P_5)$.

We leave it as an easy exercise for the reader to check that if a spanning disc, $\Delta$, meets a non-trivial translate, $g\Delta$, then the corresponding vertex for $\Delta$ in $S(c)$ heads a vertical circuit.

Another important class of sub-graphs of $S(c)$ is described below.

Definition 2.10. We shall say that a sub-graph of $S(c)$ is a lateral escape sub-graph with \textit{turning vertex} $v_i$ if it has one of the following forms:

\[
\begin{align*}
\text{Fig. 5.}
\end{align*}
\]
where each of the vertices in this sub-graph are 3-gon type vertices and the following conditions are satisfied:
(1) \( v_s \) is (H1)-related to the partner of \( v_i \) with respect to \( v_u \),
(2) \( v_i \) is (H1)-related to the partner of \( v_s \) with respect to \( v_r \).

LE(B): a sub-graph of the form (2.5), where \( v_u \) and \( v_i \) are 2-gon type vertices and \( v_s \) and \( v_r \) are 3-gon type vertices and we have conditions 1 and 2 above, or a sub-graph of the form

\[
\begin{array}{c}
v_u \\
\downarrow \quad (g) \\
\quad \downarrow \quad (f) \\
v_s \\
v_i \\
\end{array}
\]

where \( v_u \) and \( v_s \) are 3-gons and \( v_u \) is a 2-gon vertex. Moreover \( v_i \) is edge adjacent to \( v_u \). A special sort of LE(B) lateral escape sub-graph is known as an LE(B)* lateral escape. In this case each of the vertices of the 2-gon \( \Gamma_i \) supports a vertex adjacency with a type LE(B) lateral escape 3-gon.

LE(C): a sub-graph of the form (2.5), where \( v_u \) and \( v_i \) are 2-gon type vertices and \( v_s \) and \( v_r \) are 3-gon type vertices and we have conditions 1 and 2 above.

LE(D):

\[
\begin{array}{c}
v_{s_1} \\
\leftrightarrow v_{s_2} \\
\ldots \\
\leftrightarrow v_i \\
\ldots \\
\leftrightarrow v_{s_{2k+1}},
\end{array}
\]

where \( k \in \mathbb{N} \) and each of the vertices \( v_{s_1}, \ldots, v_{s_{2k+1}} \), is a 2-gon vertex and neither \( v_{s_1} \) nor \( v_{s_k} \) is (H1)-related to another 2-gon outside this sub-graph.

LE(E):

\[
\begin{array}{c}
v_u \\
\downarrow \quad (g) \\
\quad \downarrow \quad (f) \\
v_{s_1}
\end{array}
\]

where either \( v_u \) is a 2-gon type vertex and \( v_i \) and \( v_u \) are 3-gon type vertices or \( v_u \) is a 1-gon vertex split into a 2-gon and a 3-gon vertex.

We refer to each of the vertices, \( v_s \), and each of the 2-gon vertices, \( v_{s_1}, \ldots, v_{s_{2k+1}} \), in LE(D) above as lateral escape vertices.

In the argument that follows, these lateral escape sub-graphs play the significant role of allowing us to perform a local homotopy for some 1-gon, 2-gon or 3-gon spanning disc in the initial presence of a non-equivariant sub-2-gon or sub-3-gon. In order to explain better what is meant by this, we first discuss the geometric pictures described by the lateral escape sub-graphs, LE(A)–LE(E), above. In particular, we claim that these configurations are precisely those shown in Figs. 6, 7 and 8. That these are the configurations determined by the sub-graphs follows immediately in cases LE(D) and LE(E). In cases LE(A), LE(B) and LE(C), it is easy to see that the conditions 1 and 2 exclude all other configurations. Inspecting the LE(A) picture it is clear that we may reverse the 3-gon \( \Delta_s \) without reversing \( \Delta_i \), by first reversing the 3-gon \( \Delta_s \) followed by \( \Delta_r \) and finally \( \Delta_u \). In the LE(B) case,
we can similarly remove the large 2-gon, $\Delta_u$, without first removing $\Delta_i$. In the LE(C) situation, we can remove the 2-gon $\Delta_u$, without first reversing the 3-gon $\Delta_i$, and in the LE(D) case, performing $k+1$ 2-gon moves removes the 2-gon disc, $\Delta_i$. In the LE(E) case, we are able to remove $\Delta_u$ by acting first on $\Delta_s$ instead of $\Delta_i$, should $\Delta_i$ prove intractable.

One instance when we might look for lateral escape sub-graphs is when the vertex $v_i$ either supports a unit circuit or heads a vertical circuit in $S(c)$. We also use them to determine how great an obstacle a vertical circuit poses.

**Definition 2.11.** We shall say that a vertical circuit, $v_i^{(g)}$, is *genuine* if none of the vertices $\text{inv}\{v_i, \ldots, v_k\}$ is the turning vertex for either a type LE(D) sub-graph or a lateral escape sub-graph of type LE(A), LE(B), LE(C) or LE(E) in $S(c)$, itself containing a sub-graph of the form

$$
\begin{array}{c}
\downarrow \\
v_{i-1}
\end{array}
\begin{array}{c}
\downarrow \\
v_i
\end{array}
$$

(2.9)
or

\[ \begin{array}{c}
\text{\(v_{i_1}\)} \\
\downarrow \\
\text{\(v_{i_k}\).}
\end{array} \]

We note that if a vertex, \(v_i\), either supports a unit circuit or heads a genuine vertical circuit, then it is impossible to act on this vertex by the obvious transverse local move. In particular, if \(c\) has the 1-point property and \(I_1\) is a 3-gon, then any general position immersion, \(c' : S^1 \rightarrow \Sigma\) in the same homotopy class as \(c\) must have a corresponding unit or vertical circuit. A non-genuine vertical circuit, however, may be “broken” by reversing the lateral escape vertex first. We illustrate an example of this in Fig. 9. Bearing all these observations in mind, we attempt to define, given a vertex of \(S(c)\), a sub-graph which incorporates all these features and defines a natural sequence of elementary moves which allow subsequent action on \(v_i\). We draw the readers attention to the fact that this construction is the natural ancestor of the 3-dimensional namesake, introduced in [5] for immersed \(\pi_1\)-injective surfaces in 3-manifolds.

**Definition 2.12.** Let \(v_i\) be a vertex in \(S(c)\). A lateral escape structure, \(S(v_i)\), for \(v_i\) is a finite a sub-graph of \(S(c)\), which we define by induction on a quantity which we call the order of \(S(v_i)\).

We start by defining the structure sub-graphs of order one to be the minimal vertices in \(S(c)\) which do not support unit circuits.

Suppose that we have defined structure sub-graphs of all orders less than \(k\). Then \(T(v_i)\) is a structure sub-graph of order \(k\) if it has the following properties:

If \(v_i\) is the turning vertex for a type LE(D) lateral escape sub-graph in \(S(c)\) with \(2s + 1\) lateral escape 2-gons, \(s \geq 0\), then \(T(v_i)\) contains this lateral escape sub-graph, together with structure sub-graphs of order less than \(k\) for precisely \(s + 1\) of the lateral escape
2-gons, no two of these being (H1) related. If, on the other hand, \( v_i \) is not a turning vertex of this sort, we consider the set of all sub-graphs of the form:

\[
\begin{array}{c}
  v_i \\
  \downarrow (g) \\
  v_{i1} \\
  \downarrow (\rho) \\
  v_{i2}.
\end{array}
\]  

(2.11)

We deal with the possibilities as follows:

(1) If \( v_i \) is a 2-gon we have one of the following:

(a) If \( v_{i1} \) is a 2-gon, then \( T(v_i) \) contains one of

\[
\begin{array}{c}
  v_i \\
  \downarrow \\
  v_{i1}.
\end{array}
\]  

(2.12)

joined with a structure sub-graph, \( T(v_{i1}) \), for \( v_{i1} \) of order less than \( k \) or

(ii) there exists a type LE(B) lateral escape sub-graph

\[
\begin{array}{c}
  v_{i0} \\
  \downarrow (g) \\
  v_{i1} \\
  \downarrow (h) \\
  v_{i2}.
\end{array}
\]  

(2.13)

joined with a structure sub-graph, \( T(v_{i0}) \), of order less than \( k \) for \( v_{i0} \) containing the vertex \( v_{i2} \).

(b) If \( v_{i1} \) is a 3-gon, then \( T(v_i) \) contains one of

(i) the sub-graph (2.12), together with a structure sub-graph for \( v_{i1} \) of order less than \( k \),

(ii) the sub-graph

\[
\begin{array}{c}
  v_i \\
  \downarrow (g) \\
  v_{i2}.
\end{array}
\]  

(2.14)

together with a structure sub-graph for \( v_{i2} \) of order less than \( k \),

(iii) a type LE(C) lateral escape sub-graph, (2.13), joined with a structure sub-graph, \( T(v_{i0}) \), for \( v_{i0} \) containing the vertex \( v_{i1} \) of order less than \( k \),

(iv) a type LE(A) lateral escape sub-graph,

\[
\begin{array}{c}
  v_i \\
  \downarrow \\
  v_{i0} \\
  \downarrow (g) \\
  v_{i1} \\
  \downarrow (h) \\
  v_{i2}.
\end{array}
\]  

(2.15)

joined with structure sub-graphs, \( T(v_{i0}) \) and \( T(v_{i0}) \), of orders less than \( k \).

(v) If \( v_{i1} \) is a 1-gon, then \( T(v_i) \) contains (2.12), together with a structure sub-graph for \( v_{i1} \) of order less than \( k \).
(c) If $v_i$ is a 1-gon, we have the following possibilities:
   (i) $v_{i_1}$ is a 1-gon and $T(v_i)$ contains the sub-graph (2.12), together with a structure sub-graph for $v_{i_1}$ of order less than $k$,
   (ii) $v_{i_2}$ is a 2-gon and $T(v_i)$ contains the sub-graph (2.12), together with a structure sub-graph for $v_{i_2}$ of order less than $k$,
   (iii) $v_{i_3}$ is a 3-gon and $T(v_i)$ contains the sub-graph (2.12), together with a structure sub-graph for $v_{i_3}$ of order less than $k$,
   (iv) any of (i), (ii) or (iii) above, replacing $v_{i_1}$ by $v_{i_2}$.

(d) If $v_i$ is a 3-gon, we have the following possibilities:
   (i) If $v_{i_1}$ is a 2-gon, then either $T(v_i)$ contains (2.12), together with a structure sub-graph for $v_{i_1}$ of order less than $k$ or $T(v_i)$ contains a type LE(B) lateral escape sub-graph and lateral escape structures as in 1(a)(ii) above.
   (ii) If $v_{i_2}$ is a 3-gon, then either $T(v_i)$ contains (2.12), together with a structure sub-graph for $v_{i_2}$ of order less than $k$ or $T(v_i)$ contains a type LE(A) lateral escape sub-graph, (2.13) and lateral escape structures as in 1(b)(iii).

(2) If $v_i$ is not a minimal vertex in $T(v_i)$ and either heads a vertical circuit or supports a unit circuit, then we have one of the following:
   (a) The vertex $v_i$ heads a non-genuine vertical circuit, $v_{i_{\text{g}}}$ in which case the vertices in $v^{(g)}_{i_{\text{L-L}}}$ and the lateral escape sub-graph are contained in $T(v_i)$, together with structure sub-graphs of order less than $k$ as in 1 above,
   (b) $v_i$ heads a genuine vertical circuit, $v^{(g)}_{i_{\text{L-L}}}$ and $v_i$ is the turning vertex in a type LE(D) lateral escape sub-graph which is contained in $T(v_i)$ together with structure sub-graphs for $2s - 1$ of the $2s + 1$ lateral escape 2-gon vertices of order less than $k$.

(3) The vertices of $T(v_i)$ are compatible, that is, no pair of non-turning 3-gon vertices in the above construction are horizontally related in $S(c)$.

We refer to the set of all turning vertices in the above constructions as the turning vertex set for $T(v_i)$ and the set of all lateral escape vertices as the lateral escape vertex set for $T(v_i)$.

If $v_i$ has a lateral escape structure of order $k$ for some $k \in \mathbb{N}$, then we shall say that $v_i$ is weak. If this is not the case, then $v_i$ is strong in $S(c)$.

Remark 4. We note that the 2-gons which result when a curve winds several times around a primitive loop with the minimal number of transverse intersections are all strong.

Occasionally, we need a notion of compatibility for a pair of lateral escape structures for different vertices in $S(c)$. This situation arises, for instance, when we attempt to eliminate two different 2-gons using sequences of moves based on lateral escape structures as in Section 3 below. To this end, we say that a lateral escape structure for $v_i$ in $S(c)$ is compatible with a lateral escape structure for a different vertex, $v_j$, in $S(c)$ if there exists no lateral escape, non-turning 3-gon vertex in the former which is horizontally related to a lateral escape, non-turning vertex in the latter.
3. Action on lateral escape structures

We next use lateral escape structures to define sequences of 1-gon, 2-gon and 3-gon moves which either remove a 1-gon or weak 2-gon or equivariantly reverse a weak 3-gon. We start by defining a complexity, $C(v_i)$, for the vertex $v_i$ to be the number of distinct vertices in the lateral escape structure $S(v_i)$. Our sequence is constructed so as to decrease $C(v_i)$, noting that $C(v_i) = 0$ occurs only if $v_i$ is minimal in $S(c)$. In order to do this, we first examine the local effects of these generic moves on $S(c)$. To start with, suppose that $v_j$ is a 3-gon vertex which is minimal and does not support a unit circuit in $S(c)$. Writing $v_j^{-1}$ for the 3-gon vertex replacing $v_j$ after performing the 3-gon move, $\delta_j$, we have

$$\begin{align*}
\downarrow (g) \quad & \quad \delta_j \quad \uparrow (g) \\
v_j \quad & \quad v_i \quad \quad v_j^{-1}
\end{align*}$$

and conversely,

$$\begin{align*}
\downarrow (g) \quad & \quad \delta_j \quad \uparrow (g) \\
v_i \quad & \quad v_j \quad \quad v_j^{-1}
\end{align*}$$

In one very special instance, namely that of a (V2)-inclusion, a 2-gon is converted to a 1-gon by performing the 1-gon move for the smaller vertex, see Fig. 10.

We next examine the effect of these moves on lateral escape sub-graphs. In particular, if $v_j$ is the lateral escape vertex in a type LE(A) sub-graph,

$$\begin{align*}
v_i \quad & \quad \downarrow (h) \\
v_j \quad & \quad (g)
\end{align*}$$

Fig. 10.
and \( v_j \) is minimal in \( S(c) \), and does not support a unit circuit, then we have

\[
\begin{array}{cccc}
 v_a & v_r & v_j & v_u \\
 \downarrow & \downarrow & \downarrow & \downarrow \\
 v_l & \delta_j & v_j^{-1} & v_r \\
 \end{array}
\]

(3.4)

If \( v_r \) is minimal in \( S(c) \) and \( v_r \) does not support a unit circuit, then we repeat this process, performing the 3-gon move \( \delta_r \). If after this we obtain \( v_u \) minimal (not supporting a unit circuit), we perform \( \delta_u \). In this way, we are able to reverse \( v_u \) without first reversing \( v_i \). Similarly, if (3.3) is a type LE(B) lateral escape, we have (3.4) above. If \( v_r \) is minimal, we perform \( \delta_r \). If we then obtain \( v_u \) minimal, we perform the 2-gon move \( \gamma_u \) and the 2-gon vertex \( v_i \) persists. The type LE(C) case is similar and is left as an exercise for the reader. In the type LE(D) case, it is clear that performing a 2-gon move for one of the \( 2k + 1 \) lateral escape 2-gon vertices reduces the number of such vertices by two, if \( k > 0 \). Hence \( 2k \) of these moves removes all these 2-gons and also the turning vertex of the sub-graph. Finally, if

\[
\begin{array}{c}
 v_l \\
 \downarrow \,(g) \rightarrow \, v_j \\
 v_j^{-1} \\
 \end{array}
\]

(3.5)

is a type LE(E) sub-graph and \( v_j \) is minimal in \( S(c) \), then \( \delta_j \) removes both \( v_j \) and \( v_i \) from the shadow, \( s_u \), of \( v_u \) in \( S(c) \).

Using these observations, we are now in a position to prove the following key lemma:

**Lemma 3.1.** Suppose that \( v_i \) is a 1-gon, weak 2-gon or a weak 3-gon vertex in \( S(c) \) with a lateral escape structure, \( S(v_i) \). Then \( S(v_i) \) defines a sequence of 1-gon, 2-gon and 3-gon moves, \( T(v_i) \), which remove \( v_i \) if \( v_i \) is a 2-gon vertex or reverse \( v_i \) if it is a 3-gon vertex. Moreover, the graph which results is the state graph, \( S(c') \), of a general position immersion, \( c': S^1 \rightarrow \Sigma \), in the homotopy class of \( c \).

**Proof.** Let \( V_i^0 \) denote the set of all vertices in \( S(v_i) \) which are minimal in \( S(c) \) and are not turning vertices in \( S(v_i) \). If \( v_{km} \in V_i^0 \), then we perform the elementary move associated with \( v_{km} \). If \( v_{km} \) is a 1-gon or a 2-gon vertex, then we have clearly decreased the quantity \( N(v_i) \) by one. If \( v_{km} \) is a 3-gon vertex, we claim that we decrease \( N(v_i) \) by at least one. To see this, we note that if \( v_{km} \) is not the lateral escape vertex of a type LE(A), LE(B) or LE(C) sub-graph, then we simply remove \( v_{km}^{-1} \) from the modified \( S(v_i) \) to obtain a lateral
escape structure with one less vertex. If, on the other hand, \( v_{kn} \) is such a lateral escape vertex, for instance we have

\[
\begin{align*}
&
\vdots & \vdots \\
& v_{kn-3} & v_{kn-1} \\
& v_{kn-2} \xrightarrow{(b)} v_{kn}, \\
\end{align*}
\tag{3.6}
\]

and \( \delta_{kn} \) acts to produce

\[
\begin{align*}
&
\vdots & \vdots \\
& v_{kn-2} & v_{kn-3} \\
& v_{kn-1} \xrightarrow{(g)} v_{kn-2}, \\
& v_{kn} \xrightarrow{(g^{-1}h)} v_{kn-1} \\
& \vdots \\
\end{align*}
\tag{3.7}
\]

then we remove the vertices \( v_{kn-1} \) and \( v_{kn-2} \). The careful inductive definition of \( S(v_i) \) means that we retain the properties of a lateral escape structure throughout this process and \( C(v_i) \) for the modified structure is two less than that for the original one. Similarly if \( v_{km} \) is the lateral escape vertex in a type LE(E) sub-graph. In the case that \( v_{km} \) is a lateral escape vertex in a type LE(D) sub-graph, the modified lateral escape structure also has complexity \( C(v_i) \) two less than the original.

We repeat this process for each member of the set \( V_0^i \), at the end of which we denote the resulting lateral escape structure by \( S^{(1)}(v_i) \). We then define the set \( V_1^i \) in the analogous fashion to \( V_0^i \) and continue in the same way until we have \( C(v_i) = 0 \). We finish by performing the elementary move for the vertex \( v_i \) and denote the overall sequence of moves by \( T(v_i) \). In the sequel we shall refer to \( T(v_i) \) as the sequence defined by action on \( S(v_i) \).

We note that at no stage in this procedure do we increase the number of double points of the immersed loop.

Suppose next that \( \Gamma \) is a spanning 2-gon for \( c \). If \( \Gamma \) is \( \pi_1(\Sigma) \)-equivariant, then we may cancel it by equivariantly deforming one of its edges across the other. If \( \Delta \) is not \( \pi_1(\Sigma) \)-equivariant, then we shall say that \( \Delta \) is cancellable if we may equivariantly deform the lines in \( \Lambda \) until the 2-gon \( \Delta \) is \( \pi_1(\Sigma) \)-equivariant without introducing any new 1-gons or 2-gons at any time. Once this is achieved we may cancel it as above. By \([2]\), the cancellable 2-gons are precisely those which may be removed by cut-and-paste arguments. In particular, if \([c]\) is primitive in \( \pi_1(\Sigma) \), then all 2-gons are cancellable. If \([c]\) is non-primitive, then we have an \( n \)-strand in \( \Lambda \) and \( n-1 \) non-cancellable 2-gons up to translation in \( \pi_1(\Sigma) \). We may define cancellability for spanning 1-gons in a similar vein, noting, moreover, that all 1-gons are cancellable by \([2]\). When examining 3-gons, we have an analogous concept of
reversibility. That is, we define a spanning 3-gon, $\Delta$, for $c$ to be reversible if we are able to equivariantly deform the members of $\Lambda$ until $\Delta$ is $\pi_1(\Sigma)$-equivariant, after which we may reverse it in the standard fashion.

**Lemma 3.2.** If a spanning 1-gon or 2-gon $\Delta_j$, is cancellable, then we are able to transform $\Delta_j$ into an innermost 2-gon by a homotopy of $c$. Similarly, if $\Gamma_j$ is a reversible spanning 3-gon, then we may transform it into an innermost 3-gon by a homotopy of $c$.

**Proof.** We deal firstly with the case where $\Delta_j$ is a 2-gon, observing that if $\Delta_j$ is cancellable, then we may assume that it is $\pi_1(\Sigma)$-equivariant. We proceed with an explicit construction of a lateral escape structure for the vertex $v_j$ associated to the disc $\Delta_j$, which is contained in the shadow $s_j$. We start by examining the set of sub-graphs of $S(c)$ which are of one of the forms

\[
\begin{align*}
(v_j, v_{j_1}, v_{j_2}) & \quad (3.8) \\
v_j & \downarrow \\
v_{j_1} & \quad (3.9)
\end{align*}
\]

or

\[
\begin{align*}
v_j & \downarrow \\
v_{j_1} & \quad (3.10)
\end{align*}
\]

The second of these, (3.9), refers to the special situation when $v_{j_i}$ is a 1-gon and the third of these, (3.10), occurs precisely when we have neither (3.8) nor (3.9). It follows, in the case of (3.8), that $v_{j_1}$ and $v_{j_2}$ are 3-gon vertices and in the case of (3.10), the vertex $v_{j_1}$ is either a 2-gon or a 1-gon vertex. In the former instance, we choose precisely one of the vertices $v_{j_1}$ and $v_{j_2}$ to be contained in a level one sub-structure, $S^1(v_j)$, for $v_j$. Moreover, we make this choice, where necessary, subject to one important constraint, which we describe next. In particular, if in addition to a sub-graph (3.8), we have a sub-graph

\[
\begin{align*}
(v_j, v_{j_3}, v_{j_4}) & \quad (3.11) \\
v_{j_1} & \downarrow (h) \\
v_{j_4}
\end{align*}
\]

where $v_{j_3}$ is neither $v_{j_1}$ nor $v_{j_2}$ and suppose, moreover, that

\[
\begin{align*}
v_{j_1} & \downarrow (h) \\
v_{j_4} & \quad (3.12) \\
v_{j_1} & \downarrow (g) \\
v_{j_2} & \quad (3.13)
\end{align*}
\]
Then if $S^1(v_j)$ contains the sub-graph
\[
v_j \\
\downarrow
\]
then it must also contain the sub-graph
\[
v_j \\
\downarrow
j_1, j_2.
\]

(3.15)

In the event that we have the sub-graph (3.9) where $v_{j_1}$ is either a 1-gon and $v_{j_2}$ is a 2-gon, we define $S^1(v_j)$ to contain the vertex $v_{j_1}$ and the vertical edge connecting it to $v_j$. Finally, if we have the third situation, (3.10), then we define $S^1(v_j)$ to contain this sub-graph.

We note that according to this construction, no two vertices in $S^1(v_j)$ are horizontally related in $S^1(c)$ and hence this sub-graph is compatible in the sense of Definition 2.12.

Suppose next that we have constructed a sub-structure for each level less than or equal to some $k \in \mathbb{N}$ for the vertex $v_j$. Then we define a level $k + 1$ sub-structure, $S^{k+1}(v_j)$, for $v_j$ by repeating the above process for each vertex which is minimal in $S^k(v_j)$. In this more general situation, we must examine sub-graphs of the form
\[
\begin{array}{c}
v_s \\
\downarrow
\end{array}

\text{or}
\]
\[
\begin{array}{c}
v_{g_1} \\
\downarrow
\end{array}
\]
where $v_s$ may be a 1-gon or a 3-gon vertex. If $v_s$ is a 1-gon, then either $v_{g_1}$ is a 1-gon vertex and $v_{g_2}$ is a 2-gon vertex in (3.16) or $v_{g_1}$ is a 1-gon vertex in (3.17). We note that the latter describes the situation in which a representative 1-gon disc for $v_s$ interior includes a representative 1-gon disc for $v_{g_1}$. Given a sub-graph of the form (3.16), we define $S^{k+1}(v_j)$ to contain 1-gon vertex $v_{j_1}$. Similarly, given the sub-graph (3.17), we define $S^{k+1}(v_j)$ to contain it. If $v_s$ is a 3-gon vertex, then $v_{g_1}$ may be any of a 1-gon, 2-gon or a 3-gon vertex and we have a sub-graph of the form (3.17) in $S^{k+1}(v_j)$.

It is easy to check that this process terminates at some finite level, $n$, with a lateral escape structure, $S^n(v_j)$, for the vertex $v_j$ which lies in the shadow $s_j$. \square

In the second key lemma, we show that all cancellable 2-gons and 1-gons (the latter are invariably cancellable) are weak according to our definition, and so can be removed using an appropriate sequence of the elementary homotopies. To this end, suppose that $c_1 : S^1 \to \Sigma$ is a general position immersion which is homotopic to $c$ and that $c_1$ has the minimal intersection property. We then have the following fairly basic combinatorial measure of distance between the curves $c$ and $c_1$, given by
\[
d(c, c_1) = (D, E)(c, c_1),
\]
(3.18)
where $D$ measures the number of cancellable 2-gon vertices in $S(c)$ and $E$ measures the number of (cancellable) 1-gon vertices in $S(c)$. We note that the condition $d(c, c_1) = (0, 0)$ is insufficient to ensure that $c$ is ambient isotopic to $c_1$, since we may still have very different configurations of spanning 3-gons for the loops $c$ and $c_1$. We therefore refine this distance measure later on.

**Lemma 3.3.** Suppose that $c$ and $c_1$ are immersed loops as above for which $d(c, c_1) \neq (0, 0)$. Then given a 1-gon or cancellable 2-gon vertex, $v_i$, in $S(c)$, there exists a lateral escape structure for $v_i$, action on which produces an immersion, $c': S^1 \to \Sigma$, having $d(c', c_1) = (0, 0)$.

**Proof.** Suppose firstly that $A^{(k)}$ is a finite subset of lines in $\Lambda$. We shall define a configuration disc, $D(A^{(k)})$, for $A^{(k)}$ to be an embedded 2-disc in $\widetilde{\Sigma}$ which has the following properties.

1. If $l$ and $m$ are distinct lines in $A^{(k)}$ which do not share an infinite cyclic stabilizer sub-group in $\pi_1(\Sigma)$, then the points of $l \cap m$ are all contained in $\text{Int} D(A^{(k)})$.
2. If two distinct intersecting lines, $l$ and $m$, in $A^{(k)}$ share an infinite cyclic sub-group, $(h)$, in $\pi_1(\Sigma)$, then $\text{Int} D(A^{(k)})$ contains precisely one representative of each of the $(h)$-orbits in $l \cap m$.
3. Each line $\lambda$ in $A^{(k)}$ meets $D(A^{(k)})$ in a single line segment.
4. The circle $\partial D(A^{(k)})$ is transverse to the members of $A^{(k)}$.

Given a configuration disc, $D(A^{(k)})$, for $A^{(k)}$, we define the associated configuration circle, $C(A^{(k)})$, to be the boundary, $\partial D(A^{(k)})$, of the configuration disc. Moreover, we define a configuration sequence, $S(A^{(k)})$, for $A^{(k)}$ to be a sequence whose members are the points of the set

$$C(A^{(k)}) \cap \left( \bigcup A^{(k)} \right).$$

(3.19)

in a clockwise ordering on the configuration circle, $C(A^{(k)})$. We say that a point, $p$, in a configuration sequence, $S(A^{(k)})$, is derived from the line $\lambda \in A^{(k)}$ if it lies in the set $C(A^{(k)}) \cap \lambda$. Clearly, we have two derived points in $S(A^{(k)})$ for each line $\lambda \in A^{(k)}$. In order to distinguish these two points, we shall say that a point, $p$, in $C(A^{(k)})$ is positive if in moving clockwise in $S(A^{(k)})$ through $x$ we pass from the positive to the negative side of the line $\lambda$. Otherwise, we shall say that $p$ is negative. We see immediately that the two points of $\lambda \cap C(A^{(k)})$ form a pair, one member of which is positive and the other of which is negative. We may then describe $S(A^{(k)})$ unambiguously by a sequence of lines in $A^{(k)}$ in which each member of $A^{(k)}$ appears twice, once marked positive and once marked negative, and the ordering of the underlying points in $C(A^{(k)})$ is that of a configuration sequence for $A^{(k)}$. We shall call this sequence a configuration pattern for $A^{(k)}$. The key feature to observe here is that if no two lines in $A^{(k)}$ share an infinite cyclic stabilizer, then a configuration pattern for $A^{(k)}$ corresponds under the bijection $\Phi: \Lambda \to \Lambda'$ up to cyclic permutation to a configuration pattern for the set, $A^{(k)}'$, of $\Phi$-corresponding lines in $\Lambda'$. On the other hand, if two distinct lines, $l$ and $m$, in $A^{(k)}$ share an infinite cyclic...
stabilizer sub-group in \( \pi_1(\Sigma) \), then \( m = hl \) for some \( h \in \pi_1(\Sigma) \), since these lines are lifts of the same immersed loop, \( c \), in \( \Sigma \). It follows that \( l \) and \( m \) lie in an \( n \)-strand for some \( n \in \mathbb{N} \), allowing us a choice of \( n - 1 \) configuration discs with distinct configuration patterns up to cyclic permutation. It is nonetheless possible in this situation to choose a pair of configuration discs, one for \( \Lambda^{(k)} \) and one for \( \Lambda'^{(k)} \), for which the configuration patterns correspond under \( \Phi \). In the sequel, we assume that such a choice has been made. We note moreover, that this form of rigidity, albeit rather weak, allows us to list the possible 1-complexes, \( \Lambda^{(k)} \cap D(\Lambda'^{(k)}) \), given a configuration pattern for \( \Lambda^{(k)} \). We start by constructing a lateral escape structure for a single cancellable 2-gon vertex, \( \nu \), in \( S(c) \). We note, moreover, that since \( \nu \) is cancellable, then by Lemma 3.2, we may homotop \( c \) to an immersed, general position loop, \( d : S^1 \to \Sigma \), whilst converting \( \nu \) to a minimal 2-gon vertex in \( S(c) \). We go on to show that the situation is similar if \( \nu \) is a 1-gon or a reversible 3-gon. Prior to making our construction, we introduce one more concept. In particular, let us denote by \( H \) the homotopy which converts \( c \) to the map \( d \) above, where \( H(c, 0) = c \) and \( H(c, 1) = d \). Suppose, moreover, that \( H(c, t) \) is a general position immersion for all but finitely many values of \( t \) in \((0, 1)\) and \( H \) introduces no 1-gons or 2-gons. Suppose, in addition, that none of the maps, \( H(c, t), \ t \in (0, 1), \) has a non-transverse triple point, that is, a triple point where two of the lifts are tangential. Then we shall say that \( H \) is a simple homotopy. We note that the last condition is equivalent to the requirement that a 2-gon may be cancelled in the course of the homotopy only if it is first made innermost. We claim that given an arbitrary homotopy, \( H \), carrying the map \( c \) to the map \( d \), there is always a simple homotopy which performs the same function. To see this, we note that we may achieve the first condition by perturbation. The second condition follows from the definition of cancellability. In particular, any new 2-gon or 1-gon would get removed later on in the homotopy. Hence we may continuously deform the homotopy so as to prevent its introduction. It remains only to examine the third condition. In particular, we are concerned with the ways in which non-transverse triple points might arise. We note indeed that the case of three tangential curve segments may be eliminated by perturbation, leaving us with the case where two of the curve segments are tangential and the third is transverse to both of these at a point of intersection. We illustrate this in Fig. 12. The only instance where
this is not easily removable by perturbation is when we are cancelling a 2-gon which is
non-innermost. Then a traversing arc will give this sort of intersection. We have, however,
that a cancellable 2-gon may be homotoped to be innermost. Hence we need only compose
the two homotopies. Repeating this as many times as is necessary gives us the desired
conclusion. Hence we work from now on with simple homotopies.

We work using a series of local observations, starting by examining the sub-graphs of
$S(c)$ whose forms are amongst the following:

\begin{equation}
\begin{array}{c}
\text{(3.20)} \\
\begin{array}{c}
\text{vi} \quad \xrightarrow{\gamma} \\
\begin{array}{c}
\text{vi}_1 \\
\text{vi}_2
\end{array}
\end{array}
\end{array}
\end{equation}

\begin{equation}
\begin{array}{c}
\text{(3.21)} \\
\begin{array}{c}
\text{vi} \quad \xrightarrow{\gamma} \\
\text{vi}_1, \\
\text{vi}_2
\end{array}
\end{array}
\end{equation}

and, in the absence of the above, given vertices $v_i$ and $v_{i1}$,

\begin{equation}
\begin{array}{c}
\text{(3.22)} \\
\begin{array}{c}
\text{vi} \quad \xrightarrow{\gamma} \\
\text{vi}_1, \\
\text{vi}_2
\end{array}
\end{array}
\end{equation}

In the event of (3.20), $v_i$ and $v_{i2}$ are necessarily both 3-gons. In the case of (3.21), $v_i$, is a
1-gon and $v_{i2}$ is a 2-gon and in (3.22), $v_i$, may be either a 1-gon or a 2-gon. Suppose then
that we have the graph (3.20) where $v_i$, and $v_{i2}$, are both 3-gons. We consider representative
discs, $\Gamma_i$, $\Delta_i$ and $\Delta_{i2}$ for these vertices, where $\Delta_i \cup \Delta_{i2} = \Gamma_i$. These discs have edges
lying in a total of three distinct lines, $\lambda_1$, $\lambda_2$ and $\lambda_3$ covering $c$. We write $A^{(3)}$ for the set
$\{\lambda_1, \lambda_2, \lambda_3\}$ and focus on the configuration disc $D(A^{(3)})$. Now homotoping the map $c$
to the map $d$ by a simple homotopy, $H$, converts the 2-gon $\Gamma_i$ to an innermost 2-gon and
leaves the configuration pattern for the collection $A^{(3)}$ fixed. Using this information, we
can draw some conclusions about the effect of $H$ on $\Delta_{i1}$ and $\Delta_{i2}$. In particular, we must
have one of the following effects under the homotopy $H$.

1. $H$ reverses $\Delta_{i1}$,
2. $H$ reverses $\Delta_{i2}$,
(3) $H$ reverses neither $\Delta_{i1}$ nor $\Delta_{i2}$, but $\Delta_{i2}$ is the partner of some $\Delta_{j1}$ in a 2-gon $\Gamma_j$, with edge segments in the lines $\lambda_2$ and $\lambda_3$. Moreover, $H$ reverses $\Delta_{j1}$ and then later cancels $\Gamma_j$. We note that since $H$ is simple, there may be a number of moves occurring in between the reversal of $\Delta_{j1}$ and the cancellation of $\Gamma_j$.

In the event of situation 1, we define the level one sub-structure, $S^1(v_i)$, to contain the subgraph (3.22). Symmetrically, if we have the situation 2, then we define $S^1(v_i)$ to contain the sub-graph

$$
\begin{array}{c}
v_i \\
\downarrow \sigma \\
v_{i2},
\end{array}
$$

(3.23)

Finally, in the event of situation 3 we define $S^1(v_i)$ to contain the type LE(C) lateral escape sub-graph,

$$
\begin{array}{c}
v_j \\
\downarrow \nu \gamma \rightarrow \nu \\
v_{i1} \leftarrow \nu \gamma \rightarrow \nu
\end{array}
$$

(3.24)

Suppose next that we have the sub-graph (3.22) and that $v_{i1}$ is also a 2-gon vertex. The simple homotopy $H$ must have one of the following effects.

(1) $H$ cancels the 2-gon $\Gamma_{i1}$,

(2) $H$ does not cancel $\Gamma_{i1}$ and we have a type LE(B) lateral escape sub-graph,

$$
\begin{array}{c}
v_i \\
\downarrow \nu \gamma \rightarrow \nu \\
v_{i1} \leftarrow \nu \gamma \rightarrow \nu
\end{array}
$$

(3.25)

where $v_{i3}$ and $v_{i4}$ are 3-gon vertices. Moreover, $H$ acts to reverse the 3-gon $\Delta_{i3}$ and then later reverses either the 3-gon $\Delta_{i4}$, or its partner 3-gon in $\Gamma_j$.

In the event of 1, we define the level one sub-structure, $S^1(v_i)$, to contain the sub-graph (3.22). If we have the situation 2, we define $S^1(v_i)$ to contain the type LE(B) sub-graph (3.25).

Suppose next that we have the situation (3.22) where $v_{i1}$ is a 1-gon vertex. We then have two possibilities for the homotopy $H$.

(1) $H$ removes the 1-gon $Z_{i1}$,

(2) $H$ reverses the 3-gon, $\Delta_j$, which is vertex adjacent to $Z_{i1}$ and edge included by $\Gamma_j$, thereby converting $Z_{i1}$ into a 2-gon. The homotopy $H$ then cancels this 2-gon.

If we have the situation 1, then we define $S^1(v_i)$ to contain the sub-graph (3.22). If, on the other hand, we have the situation 2, then we define $S^1(v_i)$ to contain the sub-graph

$$
\begin{array}{c}
v_i \\
\downarrow \sigma \rightarrow \nu \\
v_{i1} \leftarrow \sigma \rightarrow \nu
\end{array}
$$

(3.26)
Finally, we may have the situation (3.21), where \(v_i\) is a 1-gon. In this very special case, the homotopy \(H\) must act to eliminate the 1-gon \(v_i\), and we define \(S^1(v_i)\) to contain the sub-graph (3.21).

We next turn to the case where the vertex \(v_i\) is a 1-gon. If \(v_i\) is not minimal in \(S(c)\), then it must split, either according to (3.20) into a 2-gon and a 3-gon, or it must lie in a sub-graph of the form (3.22), where \(v_i\) is a 1-gon vertex. In the former case, we have two possibilities for \(H\).

1. \(H\) cancels the sub-2-gon, \(\Gamma_i\).
2. \(H\) reverses the sub-3-gon, \(\Delta_i\).

If we have the situation of 1, then we define \(S^1(v_i)\) to contain the sub-graph (3.22). If we have that of 2, then we define \(S^1(v_i)\) to contain the sub-graph (3.23).

In the latter case, where we have the sub-graph (3.22), the homotopy \(H\) must act to remove the 1-gon \(Z_i\), and we define \(S^1(v_i)\) to contain the sub-graph (3.22).

Hence, given a 1-gon or a cancellable 2-gon vertex, \(v_i\), in \(S(c)\), we build up a level one sub-structure, \(S^1(v_i)\). We note, moreover, that this sub-graph will satisfy the compatibility criteria for a lateral escape structure. We extend this sub-graph inductively to obtain the sought-after lateral escape structure. In particular, suppose henceforth that we have defined a level \(k\) sub-structure, \(S^k(v_i)\), for \(v_i\) for some \(k \in \mathbb{N}\). We construct a level \(k+1\) sub-structure, \(S^{k+1}(v_i)\), from \(S^k(v_i)\) by examining the set of vertices which are minimal in \(S^k(v_i)\) and are not turning vertices of any lateral escape sub-graphs in \(S^k(v_i)\). Indeed, if \(v_{ik}\) is some such vertex which is either a 1-gon or a 2-gon, then we extend \(S^k(v_i)\) by adding a level one sub-structure for \(v_{ik}\) in the same manner as before. It remains then to consider the case where \(v_{ik}\) is a 3-gon vertex, where we note that the simple homotopy \(H\) acts, by definition, to reverse the underlying 3-gon disc, \(\Delta_{ik}\). We proceed by examining sub-graphs of the form

\[
v_{ik} \downarrow \quad v_{ik+1},
\]

noting that here we have no analogue of the sub-graph (3.20). The vertex \(v_{ik+1}\) may be any of a 1-gon, 2-gon or a 3-gon vertex. In the case that \(v_{ik+1}\) is a 1-gon, the homotopy \(H\) must eliminate the underlying 1-gon disc, \(Z_{ik+1}\), and we define \(S^{k+1}(v_i)\) to contain the sub-graph (3.27). If \(v_{ik+1}\) is a 2-gon vertex, then we have the following possibilities.

1. \(H\) cancels the 2-gon disc, \(\Delta_{ik+1}\).
2. \(v_{ik}\) and \(v_{ik+1}\) lie in a type LE(B) lateral escape sub-graph,

\[
v_{ik} \downarrow \quad v_{ik+1} \quad \xrightarrow{(g)} \quad v_{ik+2},
\]

in \(S(c)\). Moreover, \(H\) reverses the lateral escape 3-gon \(\Delta_{ik+2}\) and subsequently reverses the resulting sub-3-gon, \(\Delta_{ik+1}\), which replaces the 2-gon \(\Gamma_{ik+1}\).

In the first instance, we define \(S^{k+1}(v_i)\) to contain the sub-graph (3.27). If we have the situation of 2 on the other hand, then we define \(S^{k+1}(v_i)\) to contain the sub-graph (3.28).
Suppose next that \( v_{i_k+1} \) is a 3-gon. The homotopy \( H \) must act in one of the following ways.

1. \( H \) reverses the 3-gon \( \Delta_{i_k+1} \).
2. \( v_i \) and \( v_{i_k+1} \) lie in a type LE(A) lateral escape sub-graph,

\[
\begin{align*}
& v_{i_k} \\
\xrightarrow{(g)} & v_{i_k+1} \\
\rightarrow & v_{i_k+2} \\
& v_{i_k+4},
\end{align*}
\]  

(3.29)

where \( H \) reverses the 3-gon \( \Delta_{i_k+1} \) and then subsequently reverses the 3-gon \( \Delta_{i_k+3} \), followed by \( \Delta_{i_k} \).

In the first case, we define \( S_{i_k+1}^{(vi)} \) to contain the sub-graph (3.27) and in the second, we define \( S_{i_k+1}^{(vi)} \) to contain the lateral escape sub-graph (3.29).

We note that the level \( k + 1 \) sub-structure, \( S_{i_k+1}^{(vi)} \), defined in this way satisfies the compatibility criteria for a lateral escape structure. Moreover, there must exist some \( n \in \mathbb{N} \) for which \( S^n(v_i) \) is a lateral escape structure, since there exists a simple homotopy \( H \).

We let \( T(c) \) denote the sequence of 1-gon, 2-gon and 3-gon moves derived from action on this lateral escape structure. Applying this sequence of moves to the curve \( c \) removes the vertex \( v_i \), giving us a new immersed loop, \( c' \). We continue by locating a 1-gon or cancellable 2-gon for \( c' \) and applying the same procedure. Eventually we are left with a loop which has no 1-gons or cancellable 2-gons and we have proved Theorem 1.1.

Suppose, in the sequel, that \( c \) and \( c' \) are homotopic, general position immersions with the minimal intersection property. As mentioned earlier, these need not be ambient isotopic. We therefore develop our existing techniques further by designing a combinatorial algorithm for homotoping \( c \) until it is ambient isotopic to \( c' \). We do this starting with the “easy” case where \( c \) (and therefore also \( c' \)) carries a primitive element of \( \pi_1(\Sigma) \). In this circumstance we have no 1-gons or 2-gons. In particular, we may bijectively pair a 3-gon vertex in \( S(c) \) with a 3-gon vertex in \( S(c') \), where they both have corresponding underlying spanning 3-gon discs. We then focus on those vertices which reverse correspond to their counterparts under this pairing. Using this, we define a new notion of distance, \( \Delta(c, c') \), between \( c \) and \( c' \), which is the number of pairs of reverse corresponding 3-gon vertices for \( S(c) \) and \( S(c') \). Clearly,

\[
\Delta(c, c') = 0 \tag{3.30}
\]

if and only if \( c \) is ambient isotopic to \( c' \). Suppose then that \( \Delta(c, c') > 0 \). It follows, since \( c \) has no spanning 1-gons or 2-gons that there must exist some minimal 3-gon vertex, \( v_i \), in \( S(c) \) which reverse corresponds to its counterpart, \( v'_i \), in \( S(c') \). To see this, we recall the argument of Lemma 3.3 above. In particular, since \( \Delta(c, c') \) is non-zero, there must exist some 3-gon vertex, \( v_j \), in \( S(c) \) which reverse corresponds to some 3-gon vertex, \( v'_j \), in \( S(c') \). Hence there must exist a simple homotopy, \( H \), which makes \( v_j \) minimal in \( S(c) \) and this homotopy must reverse some innermost 3-gon for \( c \), as required. We are therefore able to reduce \( \Delta(c, c') \) by one through performing this 3-gon move. We repeat this process until
\[ \Delta(c, c') \] is zero. We note that since \( v_j \) is reversible, it may be made minimal by action on a lateral escape structure, \( S(v_j) \), determined as in Lemma 3.3 above. This action reduces \( \Delta(c, c') \) by at least one, since we define our lateral escape structure in terms of the simple homotopy \( H \).

The problem becomes more intricate if we allow \( c \) to carry a non-primitive element of \( \pi_1(\Sigma) \). Indeed, we must take into account non-cancellable 2-gons in \( n \)-strands. In particular, suppose that \( c \) lifts to an \( n \)-strand consisting of lines, \( \lambda, g\lambda, \ldots, g^{n-1}\lambda \), where \( n \in \mathbb{N} \), \( g^n \lambda = \lambda \), and that \( h\lambda \) is some other lift of \( c \) which crosses one of the 2-gons for this \( n \)-strand and is not a member of the \( n \)-strand itself. We denote the collection of lines, \( \{\lambda, g\lambda, h\lambda\} \), by \( \Lambda(g, h) \) and work in a configuration disc, \( D(\Lambda(g, h)) \), for \( \Lambda(g, h) \).

As mentioned earlier, we select a configuration disc, \( D(\Lambda'(g, h)) \), for the set of \( \Phi \)-corresponding lines, \( \Lambda'(g, h) = \{\lambda', g\lambda', h\lambda'\} \), with a corresponding configuration pattern. That is, the clock-wise sequence of lines and associated signs encountered on the boundary of \( D(\Lambda'(g, h)) \) corresponds to a cyclic permutation of that for \( D(\Lambda(g, h)) \). We note that the line \( h\lambda \) may only meet a single 2-gon with boundary segments in \( \lambda \) and \( g\lambda \) and that it may only split this 2-gon into a pair of sub-3-gons. Otherwise it is easy to see that it would have to belong to the \( n \)-strand itself, contrary to our assumption. As in the case of the primitive loop, we can establish a bijection between the 3-gon vertices of \( S(c) \) and those of \( S(c') \) based on the relation of correspondence. It may be the case, however, that we need to perform a succession of 3-gon moves, deforming \( h\lambda \) over more than one double point for the same pair of lines to make the picture in \( D(\Lambda(g, h)) \) look like that in \( D(\Lambda'(g, h)) \). We therefore need to refine our measure of distance. In particular, we define \( \Delta^*(c, c') \), to be given by the lexicographically ordered pair,

\[
(D, \Delta)(c, c'),
\]

where \( \Delta \) is as above. In order to define the quantity \( S \), we return to examine the discs \( D(\Lambda(g, h)) \) and \( D(\Lambda'(g, h)) \) and suppose that the 1-complexes for \( \Lambda(g, h) \) and \( \Lambda'(g, h) \) in their respective configuration discs are non-isomorphic. We then define \( D_{g, h} \) to be the maximum number of whole 2-gons with boundary segments in \( \lambda \) and \( g\lambda \) across which \( h\lambda \) must be deformed in \( D(\Lambda(g, h)) \), for the 1-complex associated with \( \Lambda(g, h) \) in \( D(\Lambda(g, h)) \) to be isomorphic with that for \( \Lambda'(g, h) \) in \( D(\Lambda'(g, h)) \). This maximum will be formed over at most two numbers. To see this, we refer the reader to Fig. 13, in which we illustrate a situation where this quantity is two, whilst a shorter “zero-length” path still exists. We define these quantities for each set of the form \( \Lambda(g, h) \), where the line \( \lambda \) is fixed and we examine precisely one representative per \( \pi_1(\Sigma) \)-orbit of ordered pair \( (g, h) \). We then define \( D(c, c') \) by summing the terms \( D_{g,h} \).

**Lemma 3.4.** Suppose that \( c \) and \( c' \) are homotopic immersions with the minimal intersection property and that \( \Delta^*(c, c') \) is non-zero. Then there exists a reversible 3-gon, action on which decreases \( \Delta^*(c, c') \) by at least one.

**Proof.** If \( D(c, c') = 0 \), then we proceed precisely as in the case of a primitive loop described above. If, on the other hand, \( D(c, c') \) is non-zero, then some term, \( D_{g,h} \), is
non-zero and from examining the configuration disc for these three lines, it is easy to see that there must exist a reversible 3-gon, \( v_j \). Moreover, the simple homotopy \( H \) may be described in the vicinity of \( v_j \) by action on some lateral escape structure, \( S(v_j) \), determined as in Lemma 3.3 above. Action on this lateral escape structure must, by its construction, reduce either \( D(c, c') \) or \( \Delta(c, c') \) by at least one.

\[ \Box \]

We continue this process until \( \Delta^*(c, c') = (0, 0) \) and the modified loop \( c \) is ambient isotopic to \( c' \). Writing \( T(c, c') \) to denote a sequence of 3-gon moves defined as above, and combining this with Theorem 1.1, we have proved Theorem 1.2.

We may summarize the achievements of this paper as follows. In particular, given an immersed loop, \( c \), in a closed, orientable surface, \( \Sigma \), we can construct the state graph, \( S(c) \), of \( c \), from examining the arrangement of its spanning discs in the universal covering plane \( \tilde{\Sigma} \). We then examine the 1-gon and 2-gon vertices of \( S(c) \) and construct a sequence of lateral escape structures, successive action on whose terms eliminates all the weak 2-gons and 1-gons in \( S(c) \). Moreover, the existence of this sequence is guaranteed and we use it to replace \( c \) by an immersion, \( c' \), which is homotopic to \( c \) and has the minimal number of self-intersections of any general position loop in its homotopy class. We extend this technology to produce an algorithm which converts a minimally self-intersecting general position immersed loop, \( c \), to one which is ambient isotopic to any other specified homotopic general position immersed loop, \( d \). This amounts to locating a sequence of lateral escape structures, successive action on which reverses the “right” 3-gons and so moves the immersion closer to \( d \). We note, moreover, that the construction of the state graph and the search for lateral escape structures should be achievable by hand in many cases and relatively straightforward to implement by computer.

Finally, we note that this technology should transfer to the situation where we have two immersed curves, \( c \) and \( d \), in \( \Sigma \) in general position, where we wish to minimize the number of points in \( c \cap d \). To do this, we would need to set up a state graph, \( S(c, d) \), the vertices of which are derived from spanning 2-gons and 3-gons with edges in both lifts of \( c \) and lifts of \( d \). By making the analogous constructions, we should then have an algorithm which minimizes \( |c \cap d| \) by a sequence of 2-gon and 3-gon moves.
Acknowledgements

I would like to thank Joel Hass and Peter Scott for helpful conversations.

References