2012 IEEE 53rd Annual Symposium on Foundations of Computer Science

On the Homotopy Test on Surfaces

Francis Lazarus
GIPSA-Lab
CNRS
Grenoble, France
Francis.Lazarus@gipsa-lab.fr

Julien Rivaud
GIPSA-Lab
CNRS
Grenoble, France
Julien.Rivaud@gipsa-lab.fr

***Abstract*—Let G be a graph cellularly embedded in a surface $\mathcal{S}$. Given two closed walks c and d in G, we take advantage of the RAM model to describe linear time algorithms to decide if c and d are homotopic in $\mathcal{S}$, either freely or with fixed basepoint. After $O(|G|)$ time preprocessing independent of c and d, our algorithms answer the homotopy test in $O(|c| + |d|)$ time, where $|G|$, $|c|$ and $|d|$ are the respective numbers of edges of G, c and d. These results were previously announced by Dey and Guha (1999). Their approach was based on small cancellation theory from combinatorial group theory. However, several flaws in their algorithms make their approach fail, leaving the complexity of the homotopy test problem still open. We present a geometric approach, based on previous works by Colin de Verdière and Erickson, that provides optimal homotopy tests.**

***Keywords*-computational topology, curve homotopy, combinatorial surface.**

I. Introduction

Computational topology of surfaces has received much attention in the last two decades. Among the notable results we may mention the test of homotopy between two cycles on a surface [1] or the computation of a shortest cycle homotopic to a given cycle [2], [3], [4]. In their 1999 paper, Dey and Guha announced a linear time algorithm for testing whether two curves on a triangulated surface are freely homotopic. This appeared as a major breakthrough for one of the most basic problem in computational topology. Dey and Guha's approach relies on results by Greendlinger [5], [6] for the conjugacy problem in one relator groups satisfying some small cancellation condition. In the appendix, we show several subtle flaws in the paper of Dey and Guha [1] that invalidate their approach[1]. Inspired by the recent work of Colin de Verdière and Erickson [4], we propose a self-contained approach and confirm the results of Dey and Guha. In addition, our free homotopy test covers the cases of orientable surfaces of genus 2 or non orientable surfaces of genus 3 and 4 which are not addressed in Dey and Guha's approach.

As commonly assumed in computational topology, we analyze the complexity of our algorithms with the uniform cost RAM model of computation [7]. A notable feature of this model is the ability to manipulate arbitrary integers in constant time per operation and to access an arbitrary memory register in constant time.

Let G be a graph cellularly embedded in a surface $\mathcal{S}$ of genus g. In a first part we consider the homotopy test for curves with fixed endpoints drawn in G. This test reduces to decide if a loop is contractible in $\mathcal{S}$, *i.e.*, null-homotopic, since two curves c and d are homotopic with fixed endpoints if and only if the concatenation $c \cdot d^{-1}$ is contractible. The contractibility test was already considered by Dey and Schipper [8] using a partial and implicit construction of the universal cover of $\mathcal{S}$. Indeed, a curve is contractible in $\mathcal{S}$ if and only if its lift is closed in the universal cover of $\mathcal{S}$. Given a closed curve c, Dey and Schipper detect if c is contractible in $O(|c| \log g)$ time. Their implicit construction is relatively complex and does not seem to extend to handle the free homotopy test. Our contractibility test relies on the more explicit construction in [4, Sec. 3.3 and 4] for tightening paths. It amounts to build a convex region of the universal cover of $\mathcal{S}$ large enough to contain a lift of c. An argument *à la Dehn* shows that this region can be chosen to have size $O(|c|)$, leading to our first theorem:

Theorem 1 (Contractibility test)**.** *Let G be a graph of complexity n cellularly embedded in a surface $\mathcal{S}$. We can preprocess G in $O(n)$ time, so that for any loop c on $\mathcal{S}$ represented as a closed walk of k edges in G, we can decide whether c is contractible or not in $O(k)$ time.*

We next study the free homotopy test, that is deciding if two cycles c and d drawn in G can be continuously deformed one to the other on $\mathcal{S}$. By theorem 1, we may assume that none of c and d is contractible. We first build (part of) the cyclic covering $\mathcal{S}_c$ of $\mathcal{S}$ induced by the cyclic subgroup generated by c in the fundamental group of $\mathcal{S}$. Assuming that $\mathcal{S}$ is orientable (see the end of Section IV for the non-orientable case), $\mathcal{S}_c$ is a topological cylinder and we call any of its non-contractible simple cycles a *generator*. Since the generators of $\mathcal{S}_c$ are freely homotopic, their projection on $\mathcal{S}$ are freely homotopic to c. Our next task is to extract from $\mathcal{S}_c$ a *canonical* generator γ_R whose definition only depends on the isomorphism class of $\mathcal{S}_c$. To this end, we lift

[1] In a private communication, one of the authors of [1] claims that their contractibility test can be fixed.

0272-5428/12 $26.00 © 2012 IEEE
DOI 10.1109/FOCS.2012.12

the graph G in $\mathcal{S}_c$ and we endow $\mathcal{S}_c$ with the corresponding cross-metric [4]. The set of generators that are minimal for this metric form a compact annulus in $\mathcal{S}_c$. We define γ_R as the "right" boundary of this annulus. Similarly, we extract a canonical generator δ_R of $\mathcal{S}_d$. If c and d are freely homotopic we know that $\mathcal{S}_c$ and $\mathcal{S}_d$ are isomorphic covering spaces [9, §V.6]. It follows that c and d are freely homotopic if and only if γ_R and δ_R have equal projections on $\mathcal{S}$. Proving that γ_R and δ_R can be constructed in time proportional to $|c|$ and $|d|$ respectively, we finally obtain:

Theorem 2 (Free homotopy test)**.** *Let G be a graph of complexity n cellularly embedded in a surface $\mathcal{S}$. We can preprocess G in $O(n)$ time, so that for any cycles c and d on $\mathcal{S}$ represented as closed walks in G with k edges in total, we can decide if c and d are freely homotopic in $O(k)$ time.*

As an immediate consequence of our two theorems, we can solve the word problem and the conjugacy problem [10, Section 6.1] in surface groups in optimal linear time. The word problem has a long standing history starting a century ago with Dehn's seminal papers [11]. The famous Dehn's algorithm is based on the property that every freely reduced word that is trivial in a surface group contains more than half of a cyclic permutation of the defining relator or its inverse. Replacing this piece of relator by its complementary piece allows to reduce the size of the word until it becomes empty. Dehn's algorithm has been extended to larger classes of groups through the theory of *small cancellation groups* based on van Kampen diagrams (see [6, Chap. V] and [12]). Most of the small cancellation groups appear themselves to be hyperbolic groups [12], to which a generalized Dehn's algorithm apply. Recent developments show that the word problem in such groups admits (linear) *real time* solutions [13], [14], [15]. We emphasize that such developments assume a multi-tape Turing machine as a model of computation where the number of tapes is related to the size of the symmetric set of defining relators. In our case, this would incure a multiplicative factor in the algorithm complexity proportional to the genus of the surface. In contrast, our solutions consider the group itself as part of the input and, after the preprocessing phase, the decision problems have linear time solutions *independent of the genus of the surface*. Although our main arguments are reminiscent of basic properties of van Kampen diagrams, such as Lyndon's curvature formula [6, Chap. V] or linear isoperimetric inequalities [16, Sec. III.2], it is not clear how to establish a precise connection[2]. For complexity purposes we are indeed working with graphs that can not be described as Cayley graphs of groups in an obvious manner, that is considering the group elements as vertices and indexing the edges into a set of generators.

[2]At the time of printing, J. Erickson and K. Whittlesey have succeded to establish this connection to obtain simpler algorithms.

Organization of the paper: We start recalling some necessary terminology and properties of surfaces, coverings and cellular embeddings of graphs in Section II. We solve the contractibility test and prove Theorem 1 in Section III. The proof of Theorem 2 for the free homotopy test is given in Section IV. In the spirit of the first Dehn's papers [11, paper 4] we sometimes appeal to concise arguments from hyperbolic geometry. Purely topological and combinatorial arguments can be found in the full version of our paper.

II. Background

We refer the reader to Massey [9] or Stillwell [10] for further details on topological concepts.

Surfaces: A compact surface without boundary is homeomorphic to a sphere where either:

- $g \geq 0$ open disks are removed and a perforated torus with one boundary component is attached to each resulting circle, or
- $g \geq 1$ open disks are removed and a Möbius band is attached to each resulting circle.

The surface is called ***orientable*** in the former case and ***non-orientable*** in the latter case. In both cases, g is the ***genus*** of the surface. In the sequel, $\mathcal{S}$ designates a compact surface of genus g.

Homotopy and fundamental group: A ***path*** in $\mathcal{S}$ is a continuous map $p : [0,1] \to \mathcal{S}$. A ***loop*** is a path p whose endpoints $p(0)$ and $p(1)$ coincide. This common endpoint is called the ***basepoint*** of the loop. Two paths p, q in $\mathcal{S}$ with common basepoint are ***homotopic*** if one can be continuously deformed into the other while fixing the basepoint. The homotopy classes of loops with given basepoint $x \in \mathcal{S}$ form the ***fundamental group*** $\pi_1(\mathcal{S}, x)$ of $\mathcal{S}$. The homotopy class of a loop c is denoted by $[c]$. The loop c is ***contractible*** if it is homotopic to the constant loop, *i.e.*, if $[c]$ is the identity of $\pi_1(\mathcal{S}, x)$. A group defined by a set A of generators and a set R of relators is denoted by $\langle A \ ; \ R\rangle$. In particular, the fundamental group of the orientable surface of genus $g \geq 1$ is isomorphic to the presentation $\langle a_1, b_1, \ldots, a_g, b_g \ ; \ a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_g b_g a_g^{-1} b_g^{-1}\rangle$.

Two loops c, d in $\mathcal{S}$ with respective basepoints x and y are ***freely homotopic*** if one can be continuously deformed into the other or, equivalently, if $[c]$ and $[u \cdot d \cdot u^{-1}]$ are conjugate in $\pi_1(\mathcal{S}, x)$ for any path u linking x to y. Recall that two group elements a, b (resp. two subgroups A, B) are ***conjugate*** if $b = gag^{-1}$ (resp. $B = gAg^{-1}$) for some group element g.

Covering spaces: A ***covering space*** of $\mathcal{S}$ is a surface $\mathcal{S}'$ together with a continuous surjective map $\pi : \mathcal{S}' \to \mathcal{S}$ such that every $x \in \mathcal{S}$ lies in an open neighborhood U such that $\pi^{-1}(U)$ is a disjoint union of open sets in $\mathcal{S}'$, each of which is mapped homeomorphically onto U by π. The map π induces a monomorphism $\pi_* : \pi_1(\mathcal{S}', y) \to \pi_1(\mathcal{S}, \pi(y))$, so that $\pi_1(S', y)$ can be considered as a subgroup of $\pi_1(S, \pi(y))$. If p is a path in $\mathcal{S}$ and $y \in \mathcal{S}'$ with

$\pi(y) = p(0)$, there exists a unique path $q : [0,1] \to \mathcal{S}'$, called a ***lift*** of p, such that $\pi \circ q = p$ and $q(0) = y$.

A ***morphism*** between the covering spaces $(\mathcal{S}', \pi)$ and $(\mathcal{S}'', \pi')$ of $\mathcal{S}$ is a continuous map $\varphi : \mathcal{S}' \to \mathcal{S}''$ such that $\pi' \circ \varphi = \pi$. The ***universal cover*** $\tilde{\mathcal{S}}$ of $\mathcal{S}$ is its unique (up to isomorphism) simply connected covering space. Unless $\mathcal{S}$ is a sphere or a projective plane $\tilde{\mathcal{S}}$ has the topology of a plane. More generally, for every subgroup $\mathcal{G}$ of $\pi_1(\mathcal{S}, x)$ there is a covering space of $\mathcal{S}$, unique up to isomorphism, whose fundamental group is conjugate to $\mathcal{G}$ in $\pi_1(\mathcal{S}, x)$. If $\mathcal{G}$ is cyclic and generated by the homotopy class of a non-contractible loop c in $\mathcal{S}$, this covering space is called the c-***cyclic cover*** and denoted by $\mathcal{S}_c$. It can be constructed as follows: take a lift $\tilde{c}$ of c in the universal cover $\tilde{\mathcal{S}}$ and let τ be the unique automorphism of $\tilde{\mathcal{S}}$ sending $\tilde{c}(0)$ to $\tilde{c}(1)$; then $\mathcal{S}_c$ is the quotient of the action of the cyclic group $\langle \tau \rangle$ generated by τ on $\tilde{\mathcal{S}}$. When $\mathcal{S}$ is orientable, $\mathcal{S}_c$ has the topology of a cylinder whose generators project on $\mathcal{S}$ to loops that are freely homotopic to c or its inverse. In Section IV we shall use the following property.

Lemma 3. *[See [17]] A self-intersecting generator of a cylinder has a contractible closed subpath.*

Cellular embeddings of graphs: All the considered graphs may have loop edges and multiple edges. We denote by $V(G)$ and $E(G)$ the set of vertices and edges of a graph G, respectively. A graph G is ***cellularly embedded*** on $\mathcal{S}$ if every open face of the embedding of G on $\mathcal{S}$ is a disk. The embedding of G can be encoded by adjoining to G a ***rotation system*** [18]. This encoding takes linear space in the ***complexity*** of G, that is, in its total number of vertices and edges. A ***facial walk*** of G is then obtained by the face traversal procedure described in [18, p. 93]. Rotation systems can be implemented efficiently [19], [20] so that we can traverse the neighbors of a vertex or the edges of a facial walk in time proportional to their degree. Any graph G cellularly embedded in $\mathcal{S}$ has a ***dual graph*** denoted by G^* whose vertices and edges are in one-to-one correspondence with the faces and edges of G respectively; if two faces of G share an edge $e \in E(G)$ its ***dual edge*** $e^* \in E(G^*)$ links the corresponding vertices of G^*. This dual graph can be cellularly embedded on $\mathcal{S}$ so that each face of G contains the matching vertex of G^* and each edge e^* dual of $e \in E(G)$ crosses only e, only once.

Let $(\mathcal{S}', \pi)$ be a covering space of $\mathcal{S}$. The ***lifted graph*** $G' = \pi^{-1}(G)$ is cellularly embedded in $\mathcal{S}'$. The restriction of π from the star of each vertex $x \in V(G')$ to the star of $\pi(x) \in V(G)$ is an isomorphism. The star of x is the set of oriented edges with origin x, and G does not need to be simple [21, Ch. 10].

Regular paths and crossing weights: For the free homotopy test we use the cross metric surface model [4] defined by a graph G cellularly embedded in $\mathcal{S}$. A path p in $\mathcal{S}$ is ***regular*** for G if every intersection point of p and G is a transverse crossing, *i.e.*, has a neighborhood in which $p \cup G$ is homeomorphic to two perpendicular line segments intersecting at their midpoint. The ***crossing weight*** with respect to G of a regular path p is the number $|p|$ of its transverse crossings and is always finite.

Homotopy encoding in cellular graph embeddings: If H is a subgraph of a cellularly embedded graph G in $\mathcal{S}$, we denote by $\mathcal{S} \setminus\!\!\setminus H$ the surface obtained after cutting $\mathcal{S}$ along H. If $\mathcal{S} \setminus\!\!\setminus H$ is a topological disk, H is called a ***cut graph***; it defines a cellular embedding in $\mathcal{S}$ with a unique face. A cut graph can be computed in linear time [22], [19]. Let T be a spanning tree of a cut graph H and consider the set of edges $A := E(H) \setminus E(T)$. (In the particular case where the cut graph H is also spanning the vertices of G, the tree T and $\mathcal{S} \setminus\!\!\setminus H$ correspond to a *tree-cotree* decomposition [19].) Euler's formula implies that A contains $2g$ edges if $\mathcal{S}$ is orientable and g edges otherwise. For each vertex $s \in H$, we have $\pi_1(\mathcal{S}, s) \cong \langle A\ ;\ R \rangle$, where R denotes the restriction of the facial walk f_H of the unique face of H to the edges in A. Indeed, if we contract T to the vertex s in $\mathcal{S}$, the graph H becomes a bouquet of circles whose complementary set in $\mathcal{S}$ is a disk bounded by the facial walk R. The above group presentation then follows from the Seifert-Van Kampen Theorem [9, Ch. IV].

Let c be a closed walk in G with basepoint $x \in V(H)$. Denote by $T(s,x)$ the unique simple path in T from s to x. We can express the homotopy class of the closed walk $c' := T(s,x) \cdot c \cdot T(x,s)$ as follows. Let $x = u_0, u_1, \ldots, u_k = x$ be the sequence of vertices that belong to H, while walking along c'. The subpath of c' between u_{i-1} and u_i is homotopic to a subpath w_i of the facial walk f_H. Denote by $w_i|_A$ the restriction of w_i to the edges in A. We have, in the above presentation of $\pi_1(\mathcal{S}, s)$:

$$[c'] = (w_1|_A) \cdot (w_2|_A) \cdots (w_k|_A)$$

We call such a product a ***term product*** representation of $[c']$ of ***height*** k. If we encode each term $w_i|_A$ implicitly by two pointers corresponding to its first and last edges in R, the above representation can be stored as a list of k pairs of pointers.

Lemma 4 (See [1, Sec. 3.1])**.** *Let G be a graph of complexity n cellularly embedded on $\mathcal{S}$. We can preprocess G and its embedding in $O(n)$ time such that the following holds. For any closed walk c in G with k edges, we can compute in $O(k)$ time a term product representation of height at most k of some closed walk freely homotopic to c.*

III. The Contractibility Test

After reduction to a simplified framework, our contractibility test for a closed curve p relies on the construction of a *relevant region* Π_p in a specific tiling of the universal cover $\tilde{\mathcal{S}}$ of $\mathcal{S}$. This relevant region, introduced by

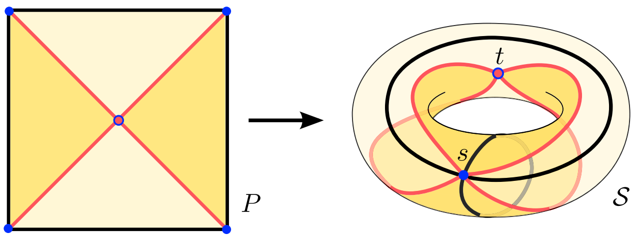

Figure 1. When $\mathcal{S}$ is a torus, P is a 4-gon and $r = 4$. The black edges of ∂P projects to G in $\mathcal{S}$. Here G is composed of two loops with basepoint s. The four radial edges in P projects to the s–t edges of the radial graph which possesses two faces.

Colin de Verdière and Erickson [4], contains a lift of p and we can decide whether p is contractible by checking if its lift is a closed curve in Π_p.

A. A simplified framework

Graphic interpretation of term products: Following Lemma 4, we can assume that G is in reduced form, with a single vertex s and a single face, and that the input closed walks are given by their term product representations in $\langle E(G)\ ;\ f_G\rangle$. We denote by r the size of the facial walk f_G, so that $r = 4g$ if $\mathcal{S}$ is orientable and $r = 2g$ otherwise. We can view $\mathcal{S} \setminus\!\!\setminus G$ as an open regular r-gon P whose boundary sides are labelled by the edges in f_G. The boundary ∂P of P maps to G after gluing back the sides of P, and its vertices all map to s. In order to represent the terms of a product as walks of constant complexity, we introduce an embedded graph H. For this, consider the radial graph in P linking the center of P to each vertex of ∂P along a straight segment; this graph has $r + 1$ vertices and r edges. After gluing back the boundary of P we obtain a bipartite graph H cellularly embedded on $\mathcal{S}$ with two vertices $\{s, t\}$ and r edges. Both vertices of H are r-valent; each of the $r/2$ faces of H is of length 4 and is cut in two "triangles" by the unique edge of G it contains — see figure 1. We call H the ***radial graph*** of G. A rotation system for H can be computed from that of G in $O(r)$ time. Any subpath of ∂P is now homotopic to the 2-walk with the same extremities and passing through t. Consequently, if $[c] = w_1 \cdots w_k$ is a term product representation of height k stored as a list of pointers then in $O(k)$ time we can obtain a closed walk of length $2k$ in H, homotopic to c.

Tiling of the universal cover: Let $(\tilde{\mathcal{S}}, \pi)$ be the universal cover of $\mathcal{S}$. A loop of H is contractible if and only if its lift in $\tilde{\mathcal{S}}$ is a loop. We shall construct a finite part of $\tilde{\mathcal{S}}$ large enough to contain that lift. To this end we rely on a tiling of $\tilde{\mathcal{S}}$ sharing similar properties with the octagonal decomposition of [4]. The lifted graph $\tilde{H} := \pi^{-1}(H)$ is an infinite regular bipartite graph with r-valent vertices and 4-valent faces. Let H^* be the *rectification* of G, *i.e.*, the dual of H on $\mathcal{S}$ chosen so that its vertices lie in the middle of the edges of G. The lifted graph $\tilde{H}^*$ is the dual of $\tilde{H}$; it defines a $\{r, 4\}$-tiling of $\tilde{\mathcal{S}}$, *i.e.*, four r-gon faces meet at every vertex of $\tilde{H}^*$. We say that two edges of H^* or $\tilde{H}^*$ are ***facing*** each other if they share an endpoint x and are not consecutive in the circular order around x. The four edges meeting at any vertex form two pairs of facing edges. The ***line*** induced by $e^* \in E(\tilde{H}^*)$ is the smallest set $\ell_{e^*} \subset E(\tilde{H}^*)$ containing e^* and the facing edge of any of its edges. Every line is an infinite lift of some cycle of facing edges of H^*, but contrary to the tight cycles of the octagonal decomposition in [4] this cycle may self-intersect. However, our set of lines in $\tilde{H}^*$ share similar properties with the lines in [4].

Proposition 5. *Suppose that $\mathcal{S}$ is orientable with genus ≥ 2 or non-orientable with genus ≥ 3. Then lines of $\tilde{H}^*$ do not self-intersect nor are cycles and two distinct lines intersect at most once. Moreover each line is separating in $\tilde{H}^*$.*

Proof: Since $\tilde{H}^*$ is $\{r, 4\}$-regular with $1/r + 1/4 < 1/2$, it can be realized as a regular geometric graph in the hyperbolic plane. In particular, the lines of $\tilde{H}^*$ are realized as hyperbolic lines. ■

Proposition 6. *Moreover, the following properties hold.*

• ***(triangle-free Property)*** *Three pairwise distinct lines cannot pairwise intersect.*

• ***(quad-free Property)*** *Four pairwise distinct lines can neither form a quadrilateral.*

Proof: We again view $\tilde{H}^*$ as a union of hyperbolic lines defining a regular tiling of the hyperbolic plane. In particular lines meet perpendicularly. The area of a hyperbolic k-gon drawn in $\tilde{H}^*$ is [23, Section 5.6]: $(k-2)\pi - \text{angle sum} \leq (k-2)\pi - k\frac{\pi}{2}$. Since the area is positive, this enforces $k > 4$. ■

B. The relevant region

Let p be a path in $\tilde{H}$. Following [4] we denote by Π_p, and call the ***relevant region*** with respect to p, the union of closed faces of $\tilde{H}^*$ reachable from $p(0)$ by crossing only lines crossed by p, in any order. In other words, if for any line ℓ we denote by ℓ^+ the component of $\tilde{\mathcal{S}} \setminus \ell$ that contains $p(0)$, then Π_p is the "convex polygon" of $\tilde{\mathcal{S}}$ formed by intersection of the sets $\ell \cup \ell^+$ for all ℓ *not* crossed by p. See Figure 2.A. This region has interesting properties which make it easy and efficient to build.

Lemma 7 ([4, lemma 4.1])**.** *For any path $p \subset \tilde{H}$ and any line ℓ the intersection $\ell \cap \Pi_p$ is either empty or a segment of ℓ whose relative interior is included in either the interior or the boundary of Π_p.*

Lemma 8 ([4, lemma 4.3])**.** *Π_p contains at most $\max(5\,|p|\,, 1)$ faces of $\tilde{H}$.*

Lemma 9 ([4, lemma 4.2])**.** *Let p be a path of $\tilde{H}$ and let $\tilde{e}$ be an edge with $p(1)$ as an endpoint. Suppose e crosses a line ℓ not already crossed by p. Then $\Pi_p \cap \ell$ is a segment of*

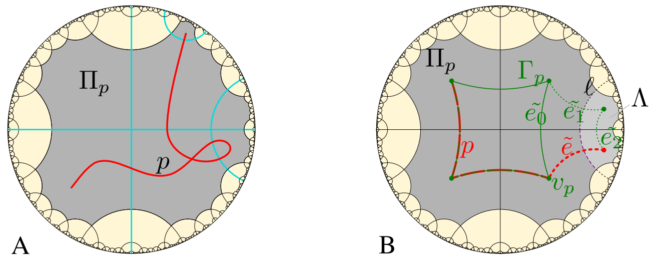

Figure 2. (A) The Poincaré disk model of $\tilde{\mathcal{S}}$ with lines represented as hyperbolic geodesics. Remark that the union of (light blue) lines crossed by p is not necessarily connected. (B) The path p is composed of two edges. The graph Γ_p corresponding to its relevant region Π_p has four edges bounding a face of $\tilde{H}$.

the boundary of Π_p *along a connected set of faces* $V \subset \Pi_p$. *Moreover* $\Pi_{p\cdot\tilde{e}} = \Pi_p \cup \Lambda$ *where* Λ *is the reflection of* V *across* ℓ.

Following lemma 9 we build the relevant region of the lift $\tilde{c}$ of a loop c incrementally as we lift its edges one at a time. Let p be the subpath of $\tilde{c}$ already traversed, and let $\tilde{e}$ be the edge following p in $\tilde{c}$. Although Π_p contains $O(|p|)$ faces by Lemma 8, a naive representation of Π_p with its faces and edges would require $O(r\,|p|)$ space. We rather store the interior of Π_p by its dual graph Γ_p, *i.e.*, by the subgraph of $\tilde{H}$ induced by the vertices dual to the faces of Π_p.

Suppose Γ_p is already constructed and let v_p be the endpoint of p in Γ_p. If $\tilde{e} \in \Gamma_p$, we have $\Gamma_{p\cdot\tilde{e}} = \Gamma_p$. Otherwise, $\tilde{e}$ is exiting Π_p and we need to enlarge Γ_p, performing a *mirror operation* as suggested by Lemma 9. We repeat this procedure until we reach the end of c, *i.e.*, when $\Gamma_{\tilde{c}}$ is constructed. We can then check that c is contractible by comparing $v_{\tilde{c}}$ with the first created vertex in $\Gamma_{\tilde{c}}$.

The mirror operation: We say that two edges of $\tilde{H}$ are ***siblings*** if their dual edges are facing each other in $\tilde{H}^*$. The four edges bounding any face of $\tilde{H}$ thus form two pairs of siblings. Let $\tilde{e_1}$ be one of the two siblings of $\tilde{e}$. Denote by $\tilde{e_0}$ and $\tilde{e_2}$ the other sibling pair bounding the same face as $\tilde{e}$ and $\tilde{e_1}$, so that v_p is an endpoint of $\tilde{e_0}$ — see Figure 2.B. If $\tilde{e_1}$ has an endpoint in Π_p, or equivalently if $\tilde{e_0} \in E(\Gamma_p)$, then by lemma 9 the edges $\tilde{e_1}$ and $\tilde{e_2}$ need to be added to Γ_p as well as their common endpoint. If on the contrary $\tilde{e_0} \notin E(\Gamma_p)$ then v_p lies in a face of $\tilde{H}^*$ that is extremal in the chain V of lemma 9, and the sibling $\tilde{e_1}$ should not be added to Γ_p. We handle similarly the sibling of $\tilde{e_1}$ which is not $\tilde{e}$, and so on until we reach the end of V. There remains to do the mirror in the other direction, starting from the still unprocessed face bounded by $\tilde{e}$. Eventually, every vertex dual to a face in the chain Λ of Lemma 9 has been created, as well as its links to existing vertices in Γ_p, and we have constructed $\Gamma_{p\cdot\tilde{e}}$.

Complexity: Any of the above operations, such as adding a vertex or an edge to Γ_p, takes constant time if we represent this graph with indexed tables for the adjacency lists. However the degree of a vertex in Γ_p may go up to r. To avoid the initialization of these tables, we use a technique inspired from [7, exercise 2.12 p. 71] taking advantage of the RAM model to allocate in $O(1)$ time an r-sized segment of memory without initializing it. It is clear from the description of the mirror subprocedure that the construction of $\Gamma_{\tilde{c}}$ takes constant time per added vertex. Lemma 8 then implies that our algorithm takes $O(|c|)$ time. Taking into account the precomputation of Lemma 4 we have proved Theorem 1 when $r \geq 6$. In the remaining cases, that is when $\mathcal{S}$ is either a torus, a Klein bottle or a projective plane, we can expand the term product representations in $O(|c|)$ time to obtain a word in the computed presentation of $\pi_1(\mathcal{S}, s)$. Testing if a word represents the unity in $\pi_1(\mathcal{S}, s)$ has trivial linear time solutions in those cases.

IV. The Free Homotopy Test

We now tackle the free homotopy test. We restrict to the case where $\mathcal{S}$ is an orientable surface of genus at least two. After fixing an orientation of $\mathcal{S}$, we can associate with every oriented edge e of H the dual edge of H^* oriented from the left face to the right face of e. This correspondence between the embedded graph and its dual will be implicit in the sequel.

We want to decide if two cycles c and d on $\mathcal{S}$ are freely homotopic. After running our contractibility test on c and d, we can assume that they are not contractible. From Section III-A, we can also assume that c and d are given as closed walks in the radial graph H. Let $(\mathcal{S}_c, \pi_c)$ be the c-cyclic cover of $\mathcal{S}$. We view $\mathcal{S}_c$ as the orbit space of the action of $\langle\tau\rangle$ where τ is the automorphism of $(\tilde{\mathcal{S}}, \pi)$ sending $\tilde{c}(0)$ on $\tilde{c}(1)$ for a given lift $\tilde{c}$ of c. We refer to τ as a *translation* of $\tilde{\mathcal{S}}$, as it can indeed be realized as a translation of the hyperbolic plane. Notice that $\mathcal{S}$ being orientable, τ is orientation preserving. The projection φ_c sending a point of $\tilde{\mathcal{S}}$ to its orbit makes $(\tilde{\mathcal{S}}, \varphi_c)$ a covering space of $\mathcal{S}_c$ with $\pi = \pi_c \circ \varphi_c$. We denote by H_c and H_c^* the respective lifts of H and H^* in $(\mathcal{S}_c, \pi_c)$. Regularity of paths in $\mathcal{S}$, $\tilde{\mathcal{S}}$ or $\mathcal{S}_c$ is considered with respect to H^*, $\tilde{H}^*$ and H_c^* respectively, and so are their crossing weights.

A. Structure of the cyclic cover

Recall that a line in $\tilde{\mathcal{S}}$ is an infinite sequence of facing edges in $\tilde{H}^*$. We start by stating some structural properties of lines.

τ-transversal and τ-invariant lines: Let ℓ be a line such that $\ell \cap \tau(\ell) = \varnothing$ and denote by $\mathring{B}_\ell$ the open band of $\tilde{\mathcal{S}}$ bounded by ℓ and $\tau(\ell)$. The line ℓ is said τ-***transversal*** if $B_\ell := \mathring{B}_\ell \cup \ell$ is a fundamental domain[3] for the action of $\langle\tau\rangle$ over $\tilde{\mathcal{S}}$. In such a case we can obtain $\mathcal{S}_c$ by point-wise identification of the boundaries ℓ and $\tau(\ell)$ of B_ℓ. The

[3] *i.e.*, every orbit of $\langle\tau\rangle$ has a unique representative in B_ℓ.

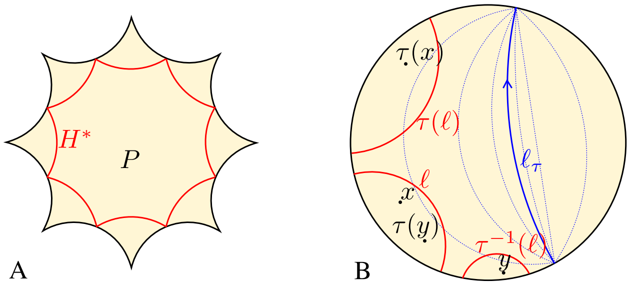

Figure 3. (A) The regular hyperbolic r-gon P. (B) The non τ-transversal line ℓ and its translates. For $z \in \{x, y\}$ the line ℓ separates z from $\tau(z)$. These two points are also separated by either $\tau(\ell)$ or $\tau^{-1}(\ell)$.

following proposition gives a characterization of τ-transversal lines whose existence are stated in Proposition 15.

Proposition 10. *Let ℓ be a line such that $\ell \cap \tau(\ell) = \varnothing$ and let $x \in \tilde{\mathcal{S}}$ be separated from $\tau(x)$ by ℓ. Then, ℓ is τ-transversal if and only if neither $\tau^{-1}(\ell)$ nor $\tau(\ell)$ separates x from $\tau(x)$.*

Proof: As described in Section III-A, we view $\mathcal{S}$ as an r-gon P whose sides are identified. We next endow $\mathcal{S}$ with a hyperbolic metric by taking for P a regular hyperbolic r-gon with angle sum 2π (see [23, Section 5.6]). We eventually join the midpoints of consecutive sides of P with r geodesic arcs (see Figure 3.A). After gluing P these arcs form an embedding of H^* in $\mathcal{S}$ and by symmetry of P, facing arcs of H^* make an angle of π. The lift $\tilde{H}^*$ of H^* in the hyperbolic plane $\tilde{\mathcal{S}}$ is thus made of hyperbolic lines. The translation τ is now a hyperbolic translation whose *axis*, ℓ_τ, is its unique globally invariant line. The action of τ on the pencil determined by its axis [25, §7.34] shows that a line is τ-transversal if and only if the line crosses ℓ_τ. Such a line clearly satisfies the conditions in the lemma. On the contrary, a line ℓ that is *not* τ-transversal must be contained in one of the two halfplanes bounded by ℓ_τ. Furthermore, if ℓ does not intersect its translate $\tau(\ell)$, it is easily seen that for any point x separated from $\tau(x)$ by ℓ, either $\tau(\ell)$ or $\tau^{-1}(\ell)$ also separates x from $\tau(x)$ (see Figure 3.B). ∎

A line ℓ such that $\tau(\ell) = \ell$ is said ***τ-invariant***. Note that this equality does not hold pointwise, but globally. The axis of τ may or may not coincide with some line of $\tilde{H}^*$. In any case:

Lemma 11. *There is at most one τ-invariant line.*

Lemma 12. *A line ℓ that intersects three consecutive translates of a τ-transversal λ is τ-invariant.*

Proof: Otherwise, ℓ, $\tau(\ell)$, $\tau(\lambda)$ and $\tau^2(\lambda)$ would contradict the quad-free Property. ∎

Generators of the cyclic cover: A loop of $\mathcal{S}_c$ that is regular (for H_c^*) and freely homotopic to $\varphi_c(\tilde{c})$ is called a ***generator*** of $\mathcal{S}_c$. A generator is said ***minimal*** if its crossing weight is minimal among generators; it projects to a regular loop of minimal crossing weight in the free homotopy class of c. By Lemma 3, we may consider that all minimal generators are simple curves. The following property is easily obtained by switching the sides of the bigons of two intersecting minimal generators.

Lemma 13. *Let μ and ν be two minimal generators. There exist two disjoint minimal generators γ and σ such that $\gamma \cup \sigma$ crosses the same set of edges as $\mu \cup \nu$.*

Consider a minimal generator γ and a lift $\tilde{\gamma} \subset \tilde{\mathcal{S}}$. The reciprocal image $\ell_\gamma := \varphi_c^{-1}(\gamma)$ is the simple curve obtained by concatenation of all the translates $\tau^i(\tilde{\gamma})$, $i \in \mathbb{Z}$. The curve ℓ_γ behaves like a line in $\tilde{\mathcal{S}}$:

Lemma 14. *The curve ℓ_γ is separating in $\tilde{\mathcal{S}}$ and each line of $\tilde{H}^*$ intersect ℓ_γ at most once.*

Proposition 15. *Let γ be a minimal generator whose basepoint is not on any line. Let $\tilde{\gamma}$ be a lift of γ in $\tilde{\mathcal{S}}$. Any line ℓ crossed by $\tilde{\gamma}$ is τ-transversal. In particular, there exists a τ-transversal line that separates the endpoints of $\tilde{c}$ from each other.*

B. The canonical generator

Since $\mathcal{S}$ is oriented we can speak of the left or right side of a minimal generator. Our aim is to prove that the set of minimal generators of $\mathcal{S}_c$ covers a bounded cylinder allowing us to define its right boundary as a canonical representative of the free homotopy class of c. By definition, a τ-transversal line projects in $\mathcal{S}_c$ to a simple curve. We call this projection a ***c-transversal***. Lemma 14 easily implies that each c-transversal ℓ crosses exactly once every minimal generator γ. Moreover, by Proposition 15, any minimal generator is crossed by c-transversals only. The number of c-transversals in $\mathcal{S}_c$ is thus equal to the length of the minimal generators which is in turn no larger than $|c|$. Notice that the orientation of $\mathcal{S}$ and of the minimal generators induce a left-to-right orientation of the c-transversals.

We now consider two disjoint minimal generators γ and σ. They bound an annulus $\mathcal{A}$ in $\mathcal{S}_c$. Since γ and σ are crossed by c-transversals only, a line ℓ of $\tilde{\mathcal{S}}$ whose projection $\varphi_c(\ell)$ intersects $\mathcal{A}$ is either τ-transversal or τ-invariant. Indeed, if ℓ is not τ-transversal, $\varphi_c(\ell)$ must stay in the finite subgraph of H_c^* interior to $\mathcal{A}$; it follows that $\varphi_c(\ell)$ uses some edge twice, which can only happen if ℓ is τ-invariant by Proposition 5. In this latter case, $\varphi_c(\ell)$ is a generator crossed once by every c-transversal and composed of $|\gamma|$ edges. By analogy with the linear isoperimetric inequality for annular diagrams in word hyperbolic groups [16, Prop. III.2.14], we first bound the complexity of $\mathcal{A}$.

Lemma 16. *Let V_I, E_I and F be the respective numbers of vertices, edges and faces of H_c^* intersected by $\mathcal{A}$. Then $V_I \leq |\gamma|$, $E_I \leq 3|\gamma|$ and $F \leq 2|\gamma|$.*

Proof: We only treat the case where there is no τ-invariant line. In $\mathcal{A}$, the c-transversals form an arrangement

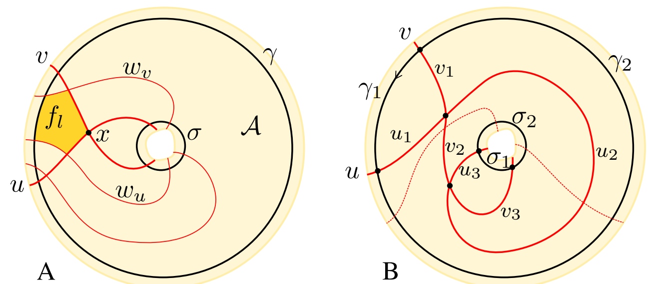

Figure 4. (A) f_l is the left face of x in the annulus $\mathcal{A}$. (B) The lines u and v cut $\mathcal{A}$ into simply connected (open) faces.

of $|\gamma|$ curves that pairwise cross at most twice by Lemma 12. Together with γ and σ, this arrangement defines a subdivision of $\mathcal{A}$ whose number of boundary vertices is $2|\gamma|$. We distinguish two cases according to whether c-transversals pairwise intersect at most once or twice.

Case 1: any two c-transversals intersect at most once: Suppose $\mathcal{A}$ has an interior vertex x; it is the intersection of two c-transversals u and v (refer to Figure 4.A). Call $t_\gamma(x)$ the triangle formed by u, v and γ. No c-transversal can join u and v inside $t_\gamma(x)$. Otherwise u, v and this c-transversal would contradict the triangle-free Property. Moreover, a c-transversal w that crosses the u-side of $t_\gamma(x)$ cannot be crossed inside $t_\gamma(x)$ by any w', as w, w', u and v would then form a quadrilateral. If the u-side of $t_\gamma(x)$ is indeed crossed, we let w_u be the crossing curve closer to x along the u-side. We define w_v similarly for the v-side of $t_\gamma(x)$. The c-transversal curves w_u and w_v, if any, together with u, v and γ bound a face of the subdivision of $\mathcal{A}$. This is the only face incident to x in $t_\gamma(x)$. We call it the *left face* of x; it has one side along γ and no side along σ. Conversely, we can prove that every face f of the subdivision of $\mathcal{A}$ with no side along σ is the left-face of a unique interior vertex. We define *right faces* analogously and remark that a face that is neither a left nor a right face must have one side along γ and one side along σ.

We can now determine the complexity of the subdivision of $\mathcal{A}$. Denote by V and E its respective numbers of vertices and edges. With the notations in the lemma, we have $V = V_I+2|\gamma|$ and $E = E_I+2|\gamma|$. By the preceding remark, every face has a side on either γ or σ. It ensues that $F \leq 2|\gamma|$. Euler's formula then implies $0 = V-E+F = V_I-E_I+F$, whence $E_I \leq V_I+2|\gamma|$. Since interior and boundary vertices have respective degree 4 and 3, we get $E_I = 2V_I + |\gamma|$ by double-counting of the vertex-edge incidences. Combining with the previous inequality we obtain $V_I \leq |\gamma|$, and finally conclude that $E_I \leq 3|\gamma|$.

Case 2: at least two c-transversals intersect twice: Let u and v be two c-transversals intersecting twice in $\mathcal{A}$. The curves γ, σ, u and v induce a subdivision of $\mathcal{A}$ where u and v are each cut into three pieces, say u_1, u_2 and u_3 for u and v_1, v_2 and v_3 for v, and the two generators are each cut into two pieces, say γ_1, γ_2 for γ and σ_1, σ_2 for σ (see Figure 4.B). The triangle and quad-free Properties imply that any c-transversal distinct from u and v must extend between γ_2 and σ_2 and cut either u_2 or v_2. Moreover, any two such c-transversals cannot cross without creating a triangle or a quadrilateral bounded by c-transversals, which is again forbidden. It follows that apart from u and v all c-transversals are pairwise disjoint in $\mathcal{A}$. We deduce that every face has one side along γ or σ (but not both), whence $F = 2|\gamma|$. Since any c-transversal distinct from u and v is cut into two pieces we also get $V_I = |\gamma|$ and $E_I = 3|\gamma|$. ■

An edge of H_c^* whose relative interior is crossed by a minimal generator is said ***short***.

Lemma 17. *If $\mathcal{A}$ is crossed by c-transversal curves only, then every edge of H_c^* in $\mathcal{A}$ is short.*

Proof: If $\mathcal{A}$ contains no vertex of H_c^* then $\mathcal{A}$ crosses the same edges as γ and σ and the lemma is trivial. Otherwise, we show how to sweep the entire arrangement inside $\mathcal{A}$ with a minimal generator from γ to σ. A simple induction on $|\gamma|$ shows that the subdivision of $\mathcal{A}$ induced by γ, σ and H_c^* contains a triangle face t with one side along γ. This triangle t has one vertex x interior to $\mathcal{A}$ and incident to four edges u_1, v_1, u_2, v_2 where u_1, v_1 bound t. We can now sweep x with γ by crossing v_2, u_2 instead of u_1, v_1 to obtain a new minimal generator. It bounds with σ a new annulus $\mathcal{A}' \subset \mathcal{A}$. Note that $\mathcal{A}'$ crosses the same set of edges as $\mathcal{A}$ except for u_1, v_1 that were crossed by γ. Moreover, the number of interior vertices is one less in $\mathcal{A}'$ than in $\mathcal{A}$. We conclude the proof with a simple recursion on this number. ■

Given two minimal generators μ and ν, there exists a minimal generator that crosses the rightmost of the short edges crossed by μ and ν along each c-transversal. Indeed, the two disjoint minimal generators returned by Lemma 13 cannot invert their order of crossings along c-transversals. Hence one of them uses all the rightmost short edges. By a simple induction on the number of c-transversals, this implies in turn that there exists a minimal generator γ_R that crosses the rightmost short edge of each c-transversal. We define the ***canonical generator*** with respect to c as the cycle in H_c dual to the sequence of short edges crossed by γ_R.

The canonical belt: As for γ_R, there exists a minimal generator γ_L crossing the leftmost short edges. We define the ***canonical belt*** $\mathcal{B}_c$ as the union of the vertices, edges and faces crossed by the annulus bounded by γ_L and γ_R. By Lemma 17, the edges in $\mathcal{B}_c$ are the short edges and the edges of the projection of the τ-invariant line, if any. All the minimal generators are included in the canonical belt.

We consider the subgraph $\tilde{K}^*$ of $\tilde{H}^*$ induced by the lines in $\tilde{H}^*$ that are neither τ-transversal nor τ-invariant. The projection $\varphi_c(\tilde{K}^*)$ of $\tilde{K}^*$ in $\mathcal{S}_c$ is denoted by K_c^*. The following lemma gives a simple characterization of the canonical belt.

Lemma 18. *$\mathcal{B}_c$ is the unique component of $\mathcal{S}_c \setminus K_c^*$ that contains a generator.*

C. Computing the canonical generator

We now explain how to compute the canonical generator associated with the loop c in time proportional to $|c|$. Let Π be the relevant region of the loop c^6 obtained by six concatenations of c. According to Section III-B, we can build the adjacency graph Γ of the faces of Π in $O(|c|)$ time. The edges dual to the edges of Γ are the edges of $\tilde{H}^*$ interior to Π. They induce a subgraph of $\tilde{H}^*$ which we denote by Γ^*. The graph Γ^* may have multiple components (see Figure 2.A), and its vertices may have degree one or four. Using the projection on H^* we can easily compute in constant time the circular list of the (one or four) edges sharing a same vertex of Γ^*, or the facing edge of any edge of Γ^*.

Identifying lines in Γ^*: From the preceding discussion we can traverse Γ^* to give a distinct tag to each maximal component of facing edges in constant time per edge. Lemma 7 ensures that each such component is supported by a distinct line. We denote by $\ell(\tilde{e})$ the identifying tag of the line supporting the edge dual to $\tilde{e}$. With a little abuse of notation we will identify a line with its tag.

Let $c_1, \ldots, c_6$ be the successive lifts of c in the lift of c^6 and let $x_0, x_1, \ldots, x_6$ be the successive lifts of $c(0)$. Let $\tilde{e}_{i,j}$ be the j-th edge of c_i. Since $\tau(c_i) = c_{i+1}$, we have $\tau(\ell(\tilde{e}_{i,j})) = \ell(\tilde{e}_{i+1,j})$. This allows us to compute the translate of any line crossing one of $c_1, \ldots, c_5$ in constant time per line. Notice that the interior of Π is crossed by a τ-invariant line if and only if $\tau(\ell) = \ell$ for some line ℓ crossing one (thus any) of $c_1, \ldots, c_5$. We can now fill in $O(|c|)$ time a table $C[\ell]$ whose Boolean value is true if ℓ intersects $\tau(\ell)$ in Π and false otherwise. We first identify the τ-transversals separating x_2 from x_3. We start by filling a table $P[\ell, i]$ counting the parity of the number of intersections of each line ℓ with c_i for $i \in \{2, 3, 4\}$. This can be done in $O(|c|)$ time: we initialize all the entries of the table P to 0 and, for each $i \in \{2, 3, 4\}$ and each edge $\tilde{e}$ of c_i, we invert the current parity of $P[\ell(\tilde{e}), i]$. By Proposition 10, the τ-transversals separating x_2 from x_3 are exactly those ℓ for which $C[\ell]$ is false, $P[\ell, 3]$ is odd, and $P[\ell, 2]$ and $P[\ell, 4]$ are even. We can now identify all the τ-transversals separating x_i from x_{i+1}, for $i \in [0, 5]$, by translation of those separating x_2 from x_3. Proposition 15 ensures the existence of at least one transversal separating x_2 from x_3; we choose one and denote it by ℓ in the sequel. We shall concentrate on the part of Π contained in the closure $\bar{B}_\ell$ of the fundamental domain of $\langle \tau \rangle$ comprised between ℓ and $\tau(\ell)$. We put $\mathcal{C} := \Pi \cap \bar{B}_\ell$.

Finding a lift of the canonical generator: We can show that $\mathcal{C}$ contains either a whole lift of the canonical belt or half of it. In this latter case, Π is bounded by a τ-invariant line.

Lemma 19. *Exactly one of the two following situations occurs*

1) *Π contains the intersection $\bar{B}_\ell \cap \varphi_c^{-1}(\mathcal{B}_c)$.*
2) *There exists a τ-invariant line λ, whose projection $\varphi_c(\lambda)$ cuts the canonical belt into two open parts $\mathcal{B}_L$ and $\mathcal{B}_R$, each one containing a generator and intersecting exactly one edge of each c-transversal. The relevant region Π contains either (i) the intersection $\bar{B}_\ell \cap \varphi_c^{-1}(\mathcal{B}_L)$ or (ii) the intersection $\bar{B}_\ell \cap \varphi_c^{-1}(\mathcal{B}_R)$, and exclude the other one.*

We now explain how to identify the lift of the canonical belt, or half of it, contained in $\mathcal{C}$.

Lemma 20. *Let Σ^* be the subgraph of $\Gamma^* \cap \mathcal{C}$ projecting to the canonical belt. We can identify the edges of Σ^* in $O(|c|)$ time.*

Proof: From the preceding lemma, $\varphi_c(\mathcal{C}) \cap \mathcal{B}_c$ is connected and contains a generator. Recall that $\tilde{K}^*$ is the union of the lines that are neither τ-transversal nor τ-invariant, and that K_c^* is its projection into $\mathcal{S}_c$. Lemma 18 ensures that $\varphi_c(\mathcal{C}) \cap \mathcal{B}_c$ is the only component of $\varphi_c(\mathcal{C}) \setminus K_c^*$ that contains a generator. Equivalently, $\mathcal{C} \cap \varphi_c^{-1}(\mathcal{B}_c)$ is the only component of $\mathcal{C} \setminus \tilde{K}^*$ that contains an edge $\tilde{e}^* \in \ell$ together with its translate $\tau(\tilde{e}^*)$.

Thanks to Lemma 12 and following the paragraph on line identification, we can detect all the τ-transversal and τ-invariant lines crossing $\mathcal{C}$. By complementarity, we identify the edges of $\tilde{K}^*$ in $\mathcal{C}$. We also identify by a simple traversal the subgraph $\Gamma_{\mathcal{C}}$ of Γ whose dual edges are contained in $\mathcal{C}$. We eventually select in time proportional to $|c|$ the component Σ of $\Gamma_{\mathcal{C}} \setminus K$ that includes an edge $\tilde{e}$ together with its translate $\tau(\tilde{e})$. We finally remark from the initial discussion that the dual of the edges in Σ are the edges of Σ^*. ∎

Proposition 21. *We can compute the canonical generator in $O(|c|)$ time.*

Proof: We first compute Σ^* as in Lemma 20. In situation 2*(i)* of Lemma 19 the canonical generator is composed of the projection on $\mathcal{S}_c$ of the dual of the edges facing the edges of Σ^* to their right. In the other situations 1 and 2*(ii)*, the edges crossed by the lift of the canonical generator in $\bar{B}_\ell$ are the edges of Σ^* that are supported by τ-transversal lines and whose right endpoint is not a crossing with any other τ-transversal or τ-invariant line. In other words, these are the rightmost edges in Σ^* of the pieces of τ-transversals crossing Σ^*, unless they abut on ℓ or $\tau(\ell)$. In either case, we can clearly determine the sequence of edges of the canonical generator in $O(|c|)$ time. ∎

D. End of the proof of Theorem 2

Let c and d be two non-contractible cycles represented as closed walks in H. Assuming that $\mathcal{S}$ is orientable with genus at least two, we compute the canonical generators γ_R and δ_R corresponding to c and d respectively. This takes $O(|c|+|d|)$ time according to Proposition 21. Following the discussion in the Introduction, c and d are freely homotopic if and only if the projections $\pi_c(\gamma_R)$ and $\pi_c(\delta_R)$ in $\mathcal{S}$ are equal as cycles of H. This can be determined, under the obvious constraint that these two projections have the same length, in $O(|c| + |d|)$ time using the Knuth-Morris-Pratt algorithm [24] to check whether $\pi_c(\gamma_R)$ is a substring of the concatenation $\pi_c(\delta_R) \cdot \pi_c(\delta_R)$. It remains to consider the case of $\mathcal{S}$ being a torus. The fundamental group of $\mathcal{S}$ is then commutative and the test reduces to the trivial contractibility test.

We have thus solved the free homotopy test for closed orientable surfaces. We finally consider the free homotopy test when $\mathcal{S}$ is non-orientable. A possibly self-crossing cycle c on $\mathcal{S}$ is ***two-sided*** if a consistent orientation of $\mathcal{S}$ can be propagated all along c. The cycle is otherwise ***one-sided***, which can easily be decided in $O(|c|)$ time with the edge signature of the combinatorial map encoding $\mathcal{S}$ [18, p. 101]. Since the square c^2 of c is two-sided, we may assume that the two given cycles c and d are two-sided. Indeed, for non-orientable surfaces of genus ≥ 3, two one-sided cycles are (freely) homotopic if and only if their square are (freely) homotopic[4]. The automorphisms of $\tilde{\mathcal{S}}$ associated to c and d are thus orientation preserving and the corresponding covering spaces $\mathcal{S}_c$ and $\mathcal{S}_d$ are again cylinders. Since $\mathcal{S}$ is non-orientable there is no a priori way of orienting those cylinders. We can nonetheless orient $\mathcal{S}_d$ arbitrarily and define γ_R as either one of the boundaries of the canonical belt of $\mathcal{S}_c$. We then carry out the whole algorithm as described for orientable surfaces with both choices. This just multiply the complexity of the free homotopy test by two.

We finally note that if $\mathcal{S}$ is a projective plane, its fundamental group is again commutative and the test is trivial. If $\mathcal{S}$ is a Klein bottle, the test was already resolved by Max Dehn [11, p.153] in linear time.

Acknowledgment

The authors would like to thank Éric colin de Verdière for interesting discussions and advices and François Dahmani for pointing out the square argument for one-sided curves in non-orientable surfaces.

References

[1] T. K. Dey and S. Guha, "Transforming Curves on Surfaces," *Journal of Computer and System Sciences*, vol. 58, no. 2, pp. 297–325, 1999.

[2] É. Colin de Verdière and F. Lazarus, "Optimal System of Loops on an Orientable Surface," *Discrete & Computational Geometry*, vol. 33, no. 3, pp. 507 – 534, 2005.

[3] ——, "Optimal Pants Decomposition and Shortest Cycles on an Orientable Surface," *Journal of the ACM*, vol. 54, no. 4, p. art. 18, 2007.

[4] É. Colin de Verdière and J. Erickson, "Tightening non-simple paths and cycles on surfaces," *SIAM Journal on Computing*, vol. 39, no. 8, pp. 3784–3813, 2010.

[5] M. D. Greendlinger, "On Dehn's algorithm for the conjugacy and word problems with applications," *Comm. Pure Appl. Math.*, vol. 13, pp. 641–677, 1960.

[6] R. Lyndon and P. Shupp, *Combinatorial Group Theory*, ser. A Series of Modern Surveys in Mathematics. Springer-Verlag, 1977, vol. 89.

[7] A. V. Aho, J. E. Hopcroft, and J. D. Ullman, *The Design and Analysis of Computer Algorithms*, ser. Series in computer science and information processing. Addison-Wesley, 1974.

[8] T. Dey and H. Schipper, "A new technique to compute polygonal schema for 2-manifolds with application to null-homotopy detection," *Discrete & Computational Geometry*, vol. 14, pp. 93–110, 1995.

[9] W. S. Massey, *A Basic Course in Algebraic Topology*, ser. Graduate Texts in Mathematics. Springer Verlag, 1991, vol. 127.

[10] J. Stillwell, *Classical topology and combinatorial group theory*. New York: Springer-Verlag, 1993.

[11] ——, *Papers on group theory and topology*. Springer-Verlag, 1987.

[12] R. Strebel, "Small cancellation groups," in *Sur les groupes hyperboliques d'après Mikhael Gromov*, E. Ghys and P. de la Harpe, Eds. Birkhäuser, 1990.

[13] B. Domanski and M. Anshel, "The complexity of Dehn's algorithm for word problems in groups," *Journal of Algorithms*, vol. 6, pp. 543–549, 1985.

[14] D. Holt, "Word-hyperbolic groups have real-time word problem," *International Journal of Algebra and Computation*, vol. 10, no. 2, pp. 221–227, 2000.

[15] D. Holt and S. Rees, "Solving the word problem in real time," *Journal London Math Soc.*, vol. 63, pp. 623–639, 2001.

[16] M. Bridson and A. Häfliger, *Metric Spaces of Non-Positive Curvature*. Springer Verlag, 1999.

[17] J. Hass and P. Scott, "Intersections of curves on surfaces," *Israel Journal of Mathematics*, vol. 51, no. 1-2, pp. 90–120, 1985.

[18] B. Mohar and C. Thomassen, *Graphs on Surfaces*. John Hopkins University Press, 2001.

[19] D. Eppstein, "Dynamic generators of topologically embedded graphs," in *Proceedings of the fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA-03)*. ACM Press, 2003, pp. 599–608.

[20] S. Lins, "Graph-encoded maps," *Journal of Combinatorial Theory, Series B*, vol. 32, pp. 171–181, 1982.

[21] L. W. Beineke and R. J. Wilson, Eds., *Topics in Topological Graph Theory*, ser. Encyclopedia of Mathematics and its Applications. Cambrigde university Press, 2009, vol. 128.

[22] S. Cabello, É. Colin de Verdière, and F. Lazarus, "Algorithms for the edge-width of an embedded graph," *Comput. Geom. Theory Appl.*, vol. 45, no. 5-6, pp. 215–224, 2012.

[4]For a quick argument, the automorphisms of the Poincaré disk $\tilde{\mathcal{S}}$ induced by a loop and its square have the same axis. If the square of two orientation reversing automorphisms are equal then their restriction to their common axis are equal, implying that the automorphisms are themselves equal.

[23] J. Stillwell, *Geometry of surfaces*. New York: Springer-Verlag, 1992.

[24] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. MIT Press, 2009.

[25] A. F. Beardon, *The Geometry of discrete groups*. New York: Springer-Verlag, 1992.

APPENDIX
COUNTER-EXAMPLES TO DEY AND GUHA'S APPROACH

We refer to the paragraph *Homotopy encoding in cellular graph embeddings* preceding Lemma 4 for the notations. In a first stage, Dey and Guha [1] obtain a term product representation of c and d as in the present Lemma 4. Suppose $R = a_1 a_2 \cdots a_{4g}$, then a term $a_i a_{i+1} \cdots a_j$ is denoted (i, j). This term is equivalent in $\langle A\,;\,R\rangle$ to the complementary term $a_{i-1}^{-1} a_{i-2}^{-1} \cdots a_{j+1}^{-1}$ *going backward* along R. This complementary term is denoted $\overline{(i-1, j+1)}$. The ***length*** $|(i,j)|$ of a term (i,j) is the length of the sequence $a_i a_{i+1} \cdots a_j$. The length of a complementary term is defined analogously, so that $|(i,j)| + |\overline{(i-1,j+1)}| = 4g$. The ***length*** of a product of (possibly complementary) terms is the sum of the lengths of its terms. Let us rename the above term and complementary term as respectively a ***forward term*** and a ***backward term***. A term will now designate either a forward or backward term. Note that a term being equivalent to its complementary term, we may use a forward or backward term in place of each term. By convention, we will write a term in backward form only if it is strictly shorter than its complementary forward term. This convention will be implicitly assumed in this section and corresponds to the notion of *rectified* term in [1].

Let us say that a product $t_1 t_2$ of two terms

- 0-***reacts*** if $t_1 t_2 = 1$, the unit element in the group $\langle A\,;\,R\rangle$,
- 1-***reacts*** if $t_1 t_2 = t$ in $\langle A\,;\,R\rangle$, for a term t such that $|t| \leq |t_1| + |t_2|$, and
- 2-***reacts*** if $t_1 t_2 = t_1' t_2'$ in $\langle A\,;\,R\rangle$, for two terms t_1', t_2' such that $|t_1'| + |t_2'| < |t_1| + |t_2|$.

The aim of Dey and Guha is to apply reactions to a given term product in order to reach a canonical form where no two consecutive terms react in that form. For this, they define a function `apply` that recursively applies reductions to a product of terms. This function is in turn called by another function `canonical`, supposed to produce a canonical form.

The following claim appears as points 2 and 3 in Lemma 4 of [1] and aims at showing that the function `apply` does terminate.

Let u, v, w be 3 terms such that uv does not react. If vw 1-reacts or 2-reacts with $vw = v'$ or $vw = v'w'$ (and $v'w'$ does not react), then uv' does not 1-react.

The non-existence of such 1-reactions is essential in the proof that the function `canonical` indeed returns a canonical form [1, Prop. 7]. However, this claim is false as demonstrated by the following examples. Consider a genus 2 surface with $R = abcda^{-1}b^{-1}c^{-1}d^{-1}$. Put $u = (2,4)$, $v = \overline{(1,7)}$, and $w = (7,8)$. Then $uv = bcd \cdot a^{-1}dc$ does not react and $vw = a^{-1}dc \cdot c^{-1}d^{-1}$ 1-reacts, yielding $v' = a^{-1}$. But uv' 1-reacts, in contradiction with the claim, since $uv' = bcd \cdot a^{-1} = (2,5)$. Likewise, if we now set $u = (2,4)$, $v = \overline{(1,8)}$ and $w = \overline{(4,2)}$, we have: uv does not react, vw 2-reacts, yielding $v'w' = (5,5) \cdot \overline{(3,2)}$, and uv' 1-reacts, in contradiction with the claim, since $uv' = bcd \cdot a^{-1} = (2,5)$.

Define the ***expanded word*** of a term product as the word in the elements of A (and their inverses) obtained by replacing each term in the product with the corresponding factor of R or R^{-1}. Again, R and R^{-1} should be considered cyclically. Call a product of terms ***stable*** if no two consecutive terms react. Another important claim [1, Lem. 8] states that

The expanded word of a stable product of terms does not contain a factor of length $2g+1$ that is also a factor of R or R^{-1}.

This claim is used to prove that the (supposed) canonical form of a product is equivalent to 1 if and only if it is the empty product [1, Prop. 6]. However this claim is again false as demonstrated by the following example. Consider the same genus 2 surface as in the previous examples. Then the product $\overline{(1,7)} \cdot (2,4) \cdot \overline{(1,7)} = cba \cdot bcd \cdot a^{-1}dc$ is stable and contains the factor $a \cdot bcda^{-1}$ of length $2g+1 = 5$ that is also a factor of R.

Finally, the canonical form defined by Dey and Guha is not canonical. By definition of the `canonical` function in [1, p. 314], a stable (rectified) product w is canonical, *i.e.*,
`canonical`$(w) = w$. Using the same genus 2 surface as before, consider the products $w_1 = \overline{(8,6)} \cdot \overline{(8,6)} = dcb \cdot dcb$ and $w_2 = (1,4) \cdot (2,5) = abcd \cdot bcda^{-1}$. It is easily seen that none of these products react. It follows that `canonical`$(w_i) = w_i$, $i = 1, 2$. However $w_1 = w_2$ in $\langle A\,;\,R\rangle$. Indeed, since $\overline{(8,6)} = (1,5)$ in $\langle A\,;\,R\rangle$, we have

$$w_1 = abcda^{-1} \cdot abcda^{-1} = abcd \cdot bcda^{-1} = w_2$$

This contradicts the fact that an element of $\langle A\,;\,R\rangle$ can be expressed as a unique canonical product of terms. In particular, Proposition 7 in [1] is wrong.

The counterexamples easily generalize to genus $g > 2$ orientable surfaces with $R = a_1 a_2 \cdots a_{2g} a_1^{-1} a_2^{-1} \cdots a_{2g}^{-1}$. Similar counterexamples for non-orientable surfaces can also be found starting with the product of squares as a (canonical) relator. In a private communication, one of the authors of [1] suggests to add an extra reaction rule that would take care of the present counter-examples. It is plausible that this extra rule leads to a linear time algorithm for the contractibility test.