TIGHTENING NONSIMPLE PATHS AND CYCLES ON SURFACES*

ÉRIC COLIN DE VERDIÈRE† AND JEFF ERICKSON‡

Abstract. We describe algorithms to compute the shortest path homotopic to a given path, or the shortest cycle freely homotopic to a given cycle, on an orientable combinatorial surface. Unlike earlier results, our algorithms do not require the input path or cycle to be simple. Given a surface with complexity n, genus $g \geq 2$, and no boundary, we construct in $O(gn \log n)$ time a *tight octagonal decomposition* of the surface—a set of simple cycles, each as short as possible in its free homotopy class, that decompose the surface into a complex of octagons meeting four at a vertex. After the surface is preprocessed, we can compute the shortest path homotopic to a given path of complexity k in $O(gnk)$ time, or the shortest cycle homotopic to a given cycle of complexity k in $O(gnk \log(nk))$ time. A similar algorithm computes shortest homotopic curves on surfaces with boundary or with genus 1. We also prove that the recent algorithms of Colin de Verdière and Lazarus for shortening embedded graphs and sets of cycles have running times polynomial in the complexity of the surface and the input curves, regardless of the surface geometry.

Key words. Computational topology, combinatorial surface, embedded graph, homotopy, topological graph theory

AMS subject classifications. 05C85, 55Q05, 57M15, 68R10, 68W05

Introduction. We consider the following topological version of the shortest path problem in geometric spaces: Given a path or cycle γ on an arbitrary surface, find the shortest path or cycle that can be obtained from γ by continuous deformation, keeping the endpoints fixed if γ is a path. Except in very special cases (such as hyperbolic surfaces), local improvement algorithms do not always converge to the true shortest path, but only to a local minimum. A more global approach is required.

Versions of this problem have been studied by several authors during the last decade. Hershberger and Snoeyink [21] find the shortest path or cycle homotopic to a given path or cycle in a triangulated piecewise-linear surface where every vertex lies on the boundary—for example, a triangulated polygon with holes in the plane. Using techniques developed by Cabello et al. [4], Efrat et al. [12] and Bespamyatnikh [1] describe algorithms to find homotopic shortest paths in the plane minus a finite set of points.

Building on earlier work, we formulate the shortest homotopic curve problem in the *combinatorial surface* model. A combinatorial surface is an abstract 2-manifold provided with a weighted embedded graph such that each face of the embedding is a disk; the curves considered are walks on this graph (in the graph-theoretical sense: paths in the graph, possibly with repeated vertices and edges). For example, a polyhedral surface where the curves are drawn on its 1-skeleton falls into this model.

*A preliminary version appeared in *Proc. ACM-SIAM Symp. on Discrete Algorithms, 2006.*

†Laboratoire d'informatique, École normale supérieure, CNRS, Paris, France; Eric.Colin.de.Verdiere@ens.fr; http://www.di.ens.fr/~colin/.

‡Department of Computer Science, University of Illinois at Urbana-Champaign; jeffe@cs.uiuc.edu; http://www.cs.uiuc.edu/~jeffe/. Partially supported by NSF under CAREER award CCR-0093348 and grants DMR-012169, CCR-0219594, and DMS-0528086. Portions of this work were done while this author was visiting École normale supérieure, LORIA/INRIA Lorraine, Polytechnic University, and Freie Universität Berlin.

Many previous works consider topological problems in combinatorial surfaces.[1] Vegter and Yap [28] and Lazarus et al. [24] describe how to build canonical polygonal schemata in this model. Dey and Guha [10] describe an algorithm to determine whether two paths or cycles are homotopic in linear time. Several algorithms have been developed recently for computing shortest families of curves with certain topological properties: Examples include the shortest non-contractible or non-separating cycle (Erickson and Har-Peled [15], Cabello and Mohar [5], Kutz [23], and Cabello and Chambers [2]), the shortest cut graph (Erickson and Har-Peled [15]), the shortest fundamental system of loops (Erickson and Whittlesey [16]). Colin de Verdière and Lazarus [7, 8] describe algorithms to compute the shortest *simple* loop homotopic to a given *simple* loop, or the shortest cycle homotopic to a given *simple* cycle, in time polynomial in the complexity of the surface, the complexity of the input curve, and the ratio between the largest and smallest edge lengths. A variant by Colin de Verdière [6] allows to compute the shortest graph embedding isotopic, with fixed vertices, to a given graph embedding.

In this paper, a curve is *tight* if it is as short as possible in its homotopy class.[2] Our main contribution is to provide efficient algorithms to tighten *possibly non-simple* paths and cycles in polynomial time, regardless of the surface geometry:

Main Theorem. *Let $\mathcal{M}$ be an orientable combinatorial surface with complexity n, genus g, and b boundaries. Let γ be a (possibly non-simple) path or cycle on $\mathcal{M}$, represented as a closed walk of complexity k in the vertex-edge graph of the combinatorial surface.*

After a preprocessing of the surface that takes $O((g+b)n)$ space, we can compute a shortest path or cycle homotopic to γ with complexity $k' = O((g+b)nk)$; the running times of the preprocessing step and of the shortening step depend on g and b and are indicated in the table below, where $\bar{k} = \min\{k, k'\}$.

	preprocessing step	path tightening	cycle tightening
$g \geq 2$, $b = 0$	$O(gn \log n)$	$O(g(k + n\bar{k}))$	$O(g(k + n\bar{k} \log(n\bar{k})))$
$g = 1$, $b = 0$	$O(n \log n)$	$O(k + n\bar{k}^2)$	$O(k + n\bar{k}^2 \log(n\bar{k}))$
$b \geq 1$	$O(n \log n + (g+b)n)$	$O((g+b)(k + n\bar{k}))$	$O((g+b)(k + n\bar{k} \log(n\bar{k})))$

For the case of genus at least two and without boundary, the $O(gn \log n)$ preprocessing step makes use of an independent result by Cabello et al. [3]. Without this result, the preprocessing step would require $O(n^2 \log n)$ time.

For the preprocessing step, we decompose the surface with a set of tight arcs or cycles C, such that the way a curve γ crosses C determines (a) its homotopy class and (b) the way some shortest curve homotopic to γ crosses C. Decompositions introduced in earlier papers [10, 7, 8, 6] share, at least partly, these properties. Using this decomposition, we prove that a shortest curve homotopic to a given curve lifts to a (small) portion of the universal cover of the surface; solving the shortest homotopic path problem then essentially amounts to computing shortest paths in this region.

In the case of a surface with boundary, such a decomposition is provided by a so-called *tight system of arcs*, generalizing the greedy system of loops constructed

[1] In some of these papers, an appropriate notion of disjointness of curves is necessary: Two curves, even sharing edges and vertices of the graph, are said disjoint provided they can be spread out infinitesimally so as to become really disjoint on the surface. This is equivalent to what we do in the dual "cross-metric surface", see §1.2.

[2] Erickson and Whittlesey [16] define a cycle to be tight if it contains the global shortest path between any two of its points. Our notion of tightness is different.

by Erickson and Whittlesey [16]. Surfaces without boundary are considerably more difficult (and interesting), because we cannot decompose them with pairwise disjoint curves. For this case, we introduce the notion of *tight regular decomposition* of the surface: an arrangement of tight cycles where every vertex of the arrangement has degree four and every face is a disk with the same number of sides; the efficiency of our algorithm follows from classical results in combinatorial group theory and hyperbolic geometry.

The existence of the decompositions described above also implies that the algorithms of Colin de Verdière and Lazarus [7, 8, 6] for shortening simple curves run in polynomial time.

This paper is organized as follows. We first introduce the background in topology and dualize the problem from combinatorial surfaces to *cross-metric surfaces*, which enable to keep track of crossings between curves. We introduce a series of tools on tight curves in §2. Then we describe our algorithms for tightening curves, for surfaces without boundary and with genus at least two (§3 and §4), for the torus (§5), and for the case of surfaces with boundary (§6). We finally give our improved analysis of the algorithms by Colin de Verdière and Lazarus [7, 8, 6] in §7.

1. Background.

1.1. Topology.

We begin by recalling several standard definitions from manifold topology. Further background can be found in textbooks by Hatcher [19] and Stillwell [27].

A *surface* (or 2-manifold possibly with boundary) $\mathcal{M}$ is a topological Hausdorff space where each point has a neighborhood homeomorphic to either the plane or the closed half-plane. The points without neighborhood homeomorphic to the plane comprise the *boundary* of $\mathcal{M}$. A (g, b)-*surface* is any surface homeomorphic to a sphere with g handles attached and b open disks removed. Every compact, connected, orientable surface $\mathcal{M}$ is a (g, b)-surface for unique integers g (its *genus*) and b (its *number of boundaries*). A *sphere* is a $(0, 0)$-surface; a *disk* is a $(0, 1)$-surface; an *annulus* (or *cylinder*) is a $(0, 2)$-surface; a *pair of pants* is a $(0, 3)$-surface; a *torus* is a $(1, 0)$-surface.

We distinguish between four different types of *curves*. A *path* on a surface $\mathcal{M}$ is (the image of) a continuous map $p : [0, 1] \to \mathcal{M}$; its *endpoints* are $p(0)$ and $p(1)$. A *loop* is a path p whose endpoints coincide. An *arc* is a path intersecting the boundary of a surface exactly at its endpoints. A *cycle* is (the image of) a continuous map $\gamma : S^1 \to \mathcal{M}$ where $S^1 = \mathbb{R}/\mathbb{Z}$ is the standard circle. A curve is *simple* if does not self-intersect (except, for a loop, at its basepoint).

Two paths p and p' are *homotopic* if there is a continuous map $h : [0, 1] \times [0, 1] \to \mathcal{M}$ such that $h(0, t) = p(t)$ and $h(1, t) = p'(t)$ for all t, and $h(\cdot, 0)$ and $h(\cdot, 1)$ are constant maps. Two cycles γ and γ' are *(freely) homotopic* if there is a continuous map $h : [0, 1] \times S^1 \to \mathcal{M}$ such that $h(0, t) = \gamma(t)$ and $h(1, t) = \gamma'(t)$ for all t. A loop or cycle is *contractible* if it is homotopic to a constant loop or cycle; an arc is contractible if it is homotopic to a path on the boundary. A simple loop, arc, or cycle is *separating* if $\mathcal{M}$ minus (the image of) this curve is disconnected. In particular, every simple contractible curve is separating. Any cycle homotopic to the boundaries of an annulus is called a *generating cycle*.

A map $\pi : \mathcal{M}' \to \mathcal{M}$ between two surfaces is called a *covering map* if each point $x \in \mathcal{M}$ lies in an open neighborhood U such that (1) $\pi^{-1}(U)$ is a countable union of disjoint open sets $U_1 \cup U_2 \cup \cdots$ and (2) for each i, the restriction $\pi|_{U_i} : U_i \to U$ is a homeomorphism. If there is a covering map π from $\mathcal{M}'$ to $\mathcal{M}$, we call $\mathcal{M}'$ a *covering space* of $\mathcal{M}$.

If p is a path in $\mathcal{M}$ and $\pi(x') = p(0)$ for some point $x' \in \mathcal{M}'$, there is a unique path p' in $\mathcal{M}'$ such that $p'(0) = x'$ and $\pi \circ p' = p$. This path p' is called a *lift* of p. Two paths are homotopic in $\mathcal{M}$ if and only if they have homotopic lifts in $\mathcal{M}'$. Similarly, a *lift* of a cycle γ is either a cycle γ' on $\mathcal{M}'$ such that $\pi \circ \gamma' = \gamma$, or a continuous 'open arc' $p' : \mathbb{R} \to \mathcal{M}'$ such that $\pi(p'(t)) = \gamma(t \bmod 1)$ for all t. Any lift of a contractible cycle is itself a contractible cycle.

Every surface $\mathcal{M}$ has a unique covering space $\widetilde{\mathcal{M}}$ in which every cycle is contractible, called the *universal cover* of $\mathcal{M}$.

Let γ be a non-contractible cycle on a surface $\mathcal{M}$. An *annular cover* of $\mathcal{M}$ with respect to γ is a covering space in which every simple cycle is either contractible or homotopic to a lift of γ or to its reverse [14, Lemma 2.5], [19, Proposition 1.36]. Topologically, the annular cover is an annulus with some points of the boundary removed.

All surfaces considered in this paper are connected, compact, and orientable, although their covering spaces are of course not necessarily compact.

1.2. Combinatorial and Cross-Metric Surfaces. In this paper, a *combinatorial surface* is a surface $\mathcal{M}$ together with a weighted undirected graph $G(\mathcal{M})$, embedded on $\mathcal{M}$ so that each open face is a disk, and the boundary of $\mathcal{M}$ is the union of some edges in $G(\mathcal{M})$. (We will simply write G if the surface $\mathcal{M}$ is clear from context.) In this model, the only allowed curves are *walks* in G: as usual, a *walk* is an alternating sequence of vertices and edges of G, starting and ending with a vertex, such that two contiguous elements in this sequence are incident. A walk is *closed* if the first and last vertex are identical. The *length* of a curve is the sum of the weights of the edges traversed by the curve, counted with multiplicity. The *complexity* of a curve is the number of edges of G it uses, counted with multiplicity. The *complexity* of a combinatorial surface is the total number of vertices, edges, and faces of G.

Most of our results are developed in a dual formulation of this model, which allows to define *crossings* between curves. A *cross-metric* surface is an abstract surface $\mathcal{M}$ together with an undirected weighted graph $G^* = G^*(\mathcal{M})$, embedded so that every open face is a disk, and the boundary of $\mathcal{M}$ is the union of some edges in G^*. We consider only *regular* paths and cycles on $\mathcal{M}$, which intersect the edges of G^* only transversely and away from the vertices. The *length* of a regular curve p is defined to be the sum of the weights of the dual edges that p *crosses*, counted with multiplicity. To emphasize this usage, we sometimes refer to the weight of a dual edge as its *crossing weight*.

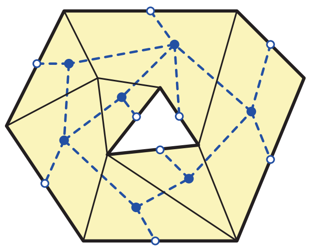

FIG. 1.1. *Primal (solid) and dual (dashed) graphs on a combinatorial annulus.*

To any combinatorial surface $(\mathcal{M}, G)$, we associate by duality a cross-metric surface $(\mathcal{M}, G^*)$, where G^* is the *dual graph* G^* of G built as follows (see Figure 1.1). G^* has a vertex f^* in the interior of each face f of G, and a vertex $\bar{e}^*$ in the relative interior of each boundary edge e of G. There are three types of dual edges. First, for each non-boundary edge e separating faces f_1 and f_2, there is a dual edge e^* between dual vertices f_1^* and f_2^*. Second, for each boundary edge e incident to face f, there is a dual edge e^* between dual vertices f^* and $\bar{e}^*$. Finally, for each boundary vertex v incident to boundary edges e_1 and e_2, there is a dual edge $\bar{v}^*$ between dual vertices $\bar{e}_1^*$ and $\bar{e}_2^*$. Each dual edge e^* in G^* intersects only its corresponding edge e in G and has the same weight as that edge. Each dual edge $\bar{v}^*$ lies entirely on the boundary of the surface and has infinite weight. Each face of G^* corresponds to a vertex of G. To any curve on a combinatorial surface, traversing edges $e_1, \ldots, e_p$, we can associate a curve in the corresponding cross-metric surface, crossing edges $e_1^*, \ldots, e_p^*$, and conversely. This transformation preserves the lengths and homotopy classes of the curves. So far, the notions of combinatorial and of cross-metric surfaces are thus essentially the same, up to duality.

A path or cycle is *tight* if it is as short as possible among all homotopic regular paths or cycles. A set of curves is *tight* if all its curves are tight. From the discussion above, to be able to tighten curves in a combinatorial surface, it suffices to be able to tighten curves in the dual cross-metric surface. We can easily construct shortest paths on a cross-metric surface by restating the usual algorithms (for example, Dijkstra's algorithm) on G in terms of the dual graph G^*.

We can represent an arbitrary arrangement of possibly (self-)intersecting curves on a cross-metric surface $\mathcal{M}$ by maintaining the *arrangement* of G^* and of the curves. Contrary to combinatorial surfaces, this data structure also encodes the crossings between curves. The initial arrangement is just the graph G^*, without any additional curve. We embed each new curve *regularly*: every crossing point of the new curve and the existing arrangement, and every self-crossing of the new curve, creates a vertex of degree four.

Whenever we split an edge e^* of G^* to insert a new curve, we give both sub-edges the same crossing weight as e^*. Each segment of the curve between two intersection points becomes a new edge, which is, unless noted otherwise, assigned weight zero. However, it is sometimes desirable to assign a non-zero weight to the edges of a curve. For example, the cross-metric surface $\mathcal{M}\backslash\alpha$ obtained by cutting $\mathcal{M}$ along an embedded curve α can be represented simply by assigning infinite crossing weights to the edges that comprise α, indicating that these edges cannot be crossed by other curves. Also it is sometimes very useful to assign to all edges of a curve a crossing weight that is a fixed *formal infinitesimal* $\varepsilon > 0$.[3] The number of curve crossings becomes thus a tie-breaking measure for computing shortest paths when two curves have the same length with respect to G^*: this ensures that later shortest-path computations always prefer paths with fewer crossings when the lengths (with respect to the original G^*) are equal. These modifications change the length of the regular curves in $\mathcal{M}$ by at most a multiple of ε; in particular, any path that is tight with respect to the refined graph is tight with respect to the original graph G^*.

We sometimes need to glue surfaces together along a common boundary; again, our representation easily supports this operation. Finally, for any (even infinite) covering space $\mathcal{M}'$ of $\mathcal{M}$, the graph $G^*(\mathcal{M}')$ is simply a lift of $G^*(\mathcal{M})$.

[3]Equivalently, we can take the crossing weight of any edge in G^* to be a vector $(\ell, 0)$ for some non-negative real number ℓ, and the crossing weight of any edge of the curve to be $(0, 1)$. Crossing weights are now vectors, which are added normally and compared lexicographically.

We emphasize that combinatorial and cross-metric surfaces do not have any 'geometry' in the usual sense; only the combinatorial structure is important. Eppstein's *gem representation* [13] is a convenient data structure for maintaining this information.

The *multiplicity* of a set of curves at some edge e of $\mathcal{M}$ is the number of times e^* is crossed by the curves. The *multiplicity* of a set of curves is the maximal multiplicity of the curves at any edge e of $\mathcal{M}$.

In this paper, $\mathcal{M}$ is a combinatorial or a cross-metric surface with genus g and b boundaries; n denotes its complexity. Since each face of $G(\mathcal{M})$ (or $G^*(\mathcal{M})$) is a disk, Euler's formula implies that $g = O(n)$ and $b = O(n)$.

2. Toolbox on Tight Curves. We will need several tools on tight curves on cross-metric surfaces, described in this section.

2.1. Basic Results. A few trivial facts will be used repeatedly without notice but are worth mentioning:

- A shortest path is tight.
- Any subpath of a tight path or cycle is tight.
- Let $\mathcal{M}$ and $\mathcal{N}$ be two surfaces with $\mathcal{N} \subset \mathcal{M}$. If a curve in $\mathcal{N}$ is tight in $\mathcal{M}$, then it is tight in $\mathcal{N}$.
- Let $\mathcal{M}'$ be a covering space of $\mathcal{M}$. Let c be a path (resp. a cycle) on $\mathcal{M}$, lifting to a path (resp. a cycle) c' on $\mathcal{M}'$. Then c is tight on $\mathcal{M}$ if and only if c' is tight on $\mathcal{M}'$.

We also note that tightening curves on a sphere or a disk is trivial, because every loop or cycle is contractible and any two paths with the same endpoints are homotopic: to tighten a loop or a cycle, simply return a constant loop or cycle; to tighten a path, simply return the shortest path between its endpoints.

We will need the following lemma, also noted in [7, Lemma 3].

LEMMA 2.1. *Let c be either a non-contractible simple cycle in the interior of $\mathcal{M}$ or a simple arc in $\mathcal{M}$. Then each lift of c separates $\widetilde{\mathcal{M}}$ into two connected components.*

We will also use results by Hass and Scott [18], which we restate here for completeness. Let C be a set of curves, each being an arc or a cycle. We assume that the curves in C are in general position: they pairwise (self-)intersect at finitely many points, where exactly one crossing between exactly two curves occurs, or exactly one self-crossing occurs. A *monogon* in C is a contractible subpath of some curve in C. A *bigon* in C is a pair of homotopic subpaths of curves in C. A monogon is *embedded* if it is simple; a bigon is *embedded* if both subpaths are simple and disjoint, except at their endpoints; hence an embedded monogon or bigon bounds a disk. We say that two curves have *excess intersection* if they can be homotoped to intersect in less points.

PROPOSITION 2.2. *The following holds.*

(a) Let γ be a non-simple arc or cycle in general position that is homotopic to a simple arc or cycle. Then γ has an embedded monogon or bigon [18, Theorems 2.1 and 2.7].

(b) Let γ and δ each be a simple arc or cycle such that they are in general position. If γ and δ have excess intersection, then there is an embedded bigon between γ and δ whose interior meets neither γ nor δ [18, Lemma 3.1].

In the rest of this section, $\mathcal{M}$ is a cross-metric surface.

2.2. Tight Curves Cross Minimally. An easy consequence of the previous proposition is the following result.[4]

LEMMA 2.3. *Let C be a set of pairwise disjoint, simple, tight curves on $\mathcal{M}$, each being an arc or a cycle. Let γ be a simple arc or cycle that is tight, in the cross-metric surface $\mathcal{M}$ where we put infinitesimal crossing weights for the curves in C. Then γ has excess intersection with no curve in C.*

Proof. Assume γ has excess intersection with a curve c in C. Applying Proposition 2.2(b), we obtain an embedded bigon between γ and c. Let x and y be the endpoints of the associated subpaths γ_1 and c_1 of γ and c. In γ, we replace the subpath γ_1 by a path running along c_1, obtaining a curve γ' that does not cross c at x and y.

The paths γ_1 and c_1 are homotopic and tight. This implies that γ and γ' are homotopic and have the same length with respect to the original graph G^*. Furthermore, the interior of c_1 does not intersect any other curve in C because the curves in C are disjoint; so γ' has at least two crossings fewer with C than γ has. This contradicts the fact that γ is tight with respect to the graph G^* with infinitesimal weights for the curves in C. ☐

2.3. Tightness and Cut. The following proposition is crucial.

PROPOSITION 2.4. *Let α and β be each a simple arc or cycle in $\mathcal{M}$ such that α and β are disjoint. Assume that α is tight in $\mathcal{M}$, and, if it is a cycle, assume it is non-contractible. Then β is tight in $\mathcal{M} \setminus \alpha$ if and only if β is tight in $\mathcal{M}$.*

Proof. If β is tight in $\mathcal{M}$, then it is obviously tight in $\mathcal{M} \setminus \alpha$. We assume β is tight in $\mathcal{M} \setminus \alpha$ and prove that it is tight in $\mathcal{M}$. The result is obvious if β is a contractible cycle in $\mathcal{M}$: in that case, it bounds a disk; this disk cannot contain α, because α is either an arc or a non-contractible cycle; so β is contractible in $\mathcal{M} \setminus \alpha$ and thus has length zero; hence it is tight in $\mathcal{M}$.

Let β' be a shortest arc or cycle homotopic, in $\mathcal{M}$, to β. It suffices to prove that β and β' have the same length. We may assume that α, β, and β' are in general position by slightly perturbing them if necessary (and moving slightly the endpoints of β' so that they are disjoint with the endpoints of β, if β is an arc). If β' is not simple, then, by Proposition 2.2(a), it must have an embedded monogon or bigon. Removing the monogon or flipping the bigon neither increases the length of β' nor changes its homotopy class, and removes at least one self-intersection. So, by induction, we may assume that β' is simple.

If β' intersects α, then, by Proposition 2.2(b), there is an embedded bigon between them, so we can push the part of β' that is on the boundary of the bigon across the bigon. This does not change the homotopy class of β'. The two subpaths bounding the bigon must have the same length, since α and β' are tight, so this operation does not change the length of β'. This does not introduce self-intersections on β' and removes exactly two crossings with α, since the interior of the bigon meets neither α nor β'. We can continue by induction. So we may assume that α and β' are disjoint and that β' is simple.

If β' intersects β, then, again by Proposition 2.2(b), we can move β' across an embedded bigon bounded by β and β'. As above, this creates no self-intersection on β' and removes exactly two crossings between β and β'. Furthermore, α does not meet the bigon, since neither β nor β' intersects α and since α is not a contractible

[4]We will use this result only in special cases where the existence of a bigon, provided by Proposition 2.2, could be easily proved by hand, but we prefer to mention it in whole generality.

cycle. In particular, this change introduces no crossing between β' and α. It also follows that both subpaths are homotopic in $\mathcal{M} \setminus \alpha$; since they are tight in $\mathcal{M} \setminus \alpha$, this does not modify the length of β'. We can continue by induction. Hence, we may assume that α, β, and β' are simple and pairwise disjoint.

Assume first that β and β' are two arcs; then they bound a disk, which cannot meet α, since α does not meet the boundary of the disk and is not a contractible cycle. The arcs β and β' must thus have the same length, since they are homotopic in $\mathcal{M} \setminus \alpha$ and both tight in $\mathcal{M} \setminus \alpha$. The result follows.

Finally, assume that β and β' are non-contractible cycles: they bound an annulus A in $\mathcal{M}$. If α does not meet A, then β and β', being homotopic in $\mathcal{M} \setminus \alpha$ and both tight in $\mathcal{M} \setminus \alpha$, have the same length. Otherwise, since α does not meet the boundaries β and β' of A, it is a cycle; since it is simple and non-contractible, it is homotopic to β and β' or to their reverses. So A is the union of two annuli, one bounded by β and α, and the other one bounded by α and β'. Hence the length of β equals the length of α, which in turn equals the length of β'. This concludes the proof. ◻

2.4. Curves Wrapping Around Cycles. A path $p : [0, 1] \to \mathcal{M}$ *wraps around* a cycle γ if $p(t) = \gamma((ut + v) \bmod 1)$ for some real numbers u and v; in particular, if p is a subpath of γ, then p wraps around γ. A cycle $\delta : S^1 \to \mathcal{M}$ *wraps around* γ if $\delta(t) = \gamma((m \cdot t) \bmod 1)$, for some integer m; one also says that δ is the mth *power* of γ. We will need the following result:

Proposition 2.5. *Any path or cycle on $\mathcal{M}$ that wraps around a tight cycle is tight.*

Proof. We first observe that it suffices to prove the assertion for cycles. Indeed, if p is a path wrapping around a tight cycle γ, then p is necessarily a subpath of some power of γ; hence, if we can prove that every power of γ is tight, then p will be also tight.

So we prove the result for a cycle wrapping around a tight cycle γ. As a special subcase, we assume that $\mathcal{M}$ is an annulus with generating cycle γ. We first introduce some notations. For each cycle δ, we let $c(\delta)$ be the number of self-crossings of δ. Let $k(\delta)$ be the unique integer such that δ is homotopic to the $(k(\delta))$th power of γ. Finally, let $|\delta|$ be the length of δ. With these notations, we need to prove that $|\delta| \geq k(\delta)|\gamma|$, and we do this by induction on $c(\delta)$.

If $c(\delta) = 0$, then δ is either contractible or homotopic to γ or to its reverse (equivalently, $k(\delta) \in \{-1, 0, +1\}$), so the result follows from the tightness of γ. Assume now that $c(\delta) \geq 1$ and that the result is true for any cycle with smaller value of c. Let p be a simple closed subpath of δ, and let δ' be the cycle δ where the loop p has been removed; we have $c(\delta') = c(\delta) - 1$, so by the induction hypothesis $|\delta'| \geq k(\delta')|\gamma|$.

Henceforth, we consider p as a cycle. Since it is simple, we have $k(p) \in \{-1, 0, +1\}$. We distinguish two cases according to the value of $k(p)$. If $k(p) = 0$, then $|\delta| \geq |\delta'| \geq k(\delta')|\gamma|$ (as shown above) which is in turn equal to $k(\delta)|\gamma|$; this implies the desired result. In the other case, we have $|\delta| = |\delta'| + |p|$. But $|\delta'| \geq k(\delta')|\gamma|$, and $|p| \geq |\gamma|$ by tightness of γ and since p is homotopic to γ or to its reverse, so we have $|\delta| \geq (k(\delta') + 1)|\gamma|$. Since $k(\delta)$ and $k(\delta')$ differ by at most one, this is at least $k(\delta)|\gamma|$. This concludes the proof.

To prove the general case, we may clearly assume that γ is non-contractible. We consider the *annular* cover $\widehat{\mathcal{M}}$ of $\mathcal{M}$ generated by a lift $\widehat{\gamma}$ of γ. Since γ is tight in $\mathcal{M}$, its lift $\widehat{\gamma}$ is tight in $\widehat{\mathcal{M}}$; so each power of $\widehat{\gamma}$ is tight in $\widehat{\mathcal{M}}$, by the special case proved above; by projection, each power of γ is tight in $\mathcal{M}$. ◻

2.5. Multiplicity of Some Tight Cycles. The tool we introduce now bounds the multiplicity of certain tight cycles.

PROPOSITION 2.6. *Any tight simple cycle γ homotopic to a boundary δ of $\mathcal{M}$ has multiplicity at most two.*

Proof. Let $\mathcal{N}$ be the surface $\mathcal{M}$ with a disk glued to δ. Let $\widetilde{\mathcal{N}}$ be its universal cover, with covering map $\pi : \widetilde{\mathcal{N}} \to \mathcal{N}$. Every lift $\tilde{\gamma}$ of γ is a simple cycle enclosing exactly one lift $\tilde{\delta}$ of δ; conversely, every lift of δ is enclosed by exactly one lift of γ.

First, we claim that every lift $\tilde{\gamma}$ of γ has multiplicity one in $\widetilde{\mathcal{N}}$. Suppose to the contrary $\tilde{\gamma}$ crosses some edge $\tilde{e}$ of $G^*(\widetilde{\mathcal{N}})$ at points x and y, and let $\overline{xy}$ denote the open segment of $\tilde{e}$ between these two intersection points. If some lift of γ crosses $\overline{xy}$, the Jordan curve theorem implies that it has to cross $\overline{xy}$ again in the other direction. Thus, by choosing the lift $\tilde{\gamma}$ and points x and y appropriately, we can assume that no lift of γ crosses $\overline{xy}$.

The intersection points x and y split $\tilde{\gamma}$ into two paths. Gluing $\overline{xy}$ to these paths gives us two disjoint simple cycles $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$, each of which is shorter than $\tilde{\gamma}$ (because they do not cross $\tilde{e}$). Because no other lift of γ crosses $\overline{xy}$, the cycle γ does not intersect the segment $\pi(\overline{xy})$ in $\mathcal{N}$. Thus, $\gamma_1 = \pi(\tilde{\gamma}_1)$ and $\gamma_2 = \pi(\tilde{\gamma}_2)$ are disjoint simple cycles in $\mathcal{N}$, both strictly shorter than γ.

Since γ_1 is simple, all its lifts to $\widetilde{\mathcal{N}}$ are pairwise disjoint, and therefore each lift of γ_1 encloses at most one lift of δ. (Otherwise, two lifts of γ_1 would enclose two sets of lifts A and B of δ such that $A \not\subseteq B$ and $B \not\subseteq A$, which is excluded by the Jordan curve theorem.) The same holds for γ_2. Let $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ be lifts of γ_1 and γ_2. If $\tilde{\gamma}_1$ encloses a lift of δ, then γ_1 is homotopic to γ in $\mathcal{M}$, which is impossible since γ is tight. On the other hand, if $\tilde{\gamma}_1$ does not enclose a lift of δ, then γ_1 is contractible in $\mathcal{M}$, which implies that γ_2 is homotopic to γ in $\mathcal{M}$, which is also impossible. This completes the proof of our claim.

Now suppose γ crosses some edge of $G^*(\mathcal{M})$ three or more times. Then there is an edge $\tilde{e}$ of $G^*(\widetilde{\mathcal{N}})$ that is crossed at least three times by lifts of γ, at some points x^1, x^2, and x^3, labeled in order along $\tilde{e}$. By our earlier claim, these intersection points lie on three different lifts $\tilde{\gamma}^1$, $\tilde{\gamma}^2$, and $\tilde{\gamma}^3$ of γ. Each lift $\tilde{\gamma}^i$ encloses exactly one lift $\tilde{\delta}^i$ of δ. Since $\tilde{\gamma}^2$ is a simple cycle that crosses $\tilde{e}$ exactly once, the Jordan curve theorem implies that it encloses either $\tilde{\gamma}^1$ or $\tilde{\gamma}^3$. But then $\tilde{\gamma}^2$ must also enclose $\tilde{\delta}^1$ or $\tilde{\delta}^3$, which is impossible. We conclude that the multiplicity of γ is at most two. □

2.6. Construction of Tight Curves. We introduce six elementary constructions of tight arcs and cycles on a cross-metric surface. Our algorithms will combine these elementary constructions by computing a tight simple arc or cycle on the input surface $\mathcal{M}$, cutting $\mathcal{M}$ along it, and iterating on the resulting surface. Proposition 2.4 implies that the arcs and cycles computed in this fashion are always tight *on the original surface $\mathcal{M}$*.

Let $\mathcal{M}$ be a cross-metric surface with boundary δ. We say that an edge e of $\mathcal{M}$ is *adjacent* to δ if e shares with δ a vertex v such that e is the only non-boundary edge incident to v.

PROPOSITION 2.7.

(a) *If x and y are points on the boundary of $\mathcal{M}$, we can compute in $O(n \log n)$ time a shortest arc with endpoints x and y. The computed arc is simple and has multiplicity one.*

(b) *If $b \geq 2$, we can compute in $O(n \log n)$ time a shortest arc joining two specified boundaries. The computed arc is simple, has multiplicity one, and has* *multiplicity zero at each edge of G^* adjacent to one of the two specified boundaries.*

(c) If $b = 1$ and $g \neq 0$, we can compute in $O(n \log n)$ time a shortest non-contractible, or non-separating, arc. The computed arc is simple, has multiplicity at most two, and has multiplicity zero at each edge of G^* adjacent to the boundary.

(d) If $b = 0$, we can compute in $O(n \log n)$ time a tight non-separating cycle that is simple and has multiplicity at most six.

(e) If $g = 0$ and $b = 2$, we can compute in $O(n \log n)$ time a shortest generating cycle. The computed cycle is simple and has multiplicity one.

(f) If $g = 0$ and $b = 3$, we can compute in $O(n \log n)$ time a shortest cycle homotopic to a chosen boundary cycle. The computed cycle is simple and has multiplicity at most two.

Proof.

(a) Compute a shortest path between x and y using Dijkstra's algorithm.

(b) Temporarily fill the two specified boundaries with a disk, and assign infinitesimal crossing weights to the edges of the boundary of each of these disks. Pick points x and y inside each of these disks. Compute a shortest path from x to y using Dijkstra's algorithm. This path is simple; it crosses each of the boundaries of the aforementioned disks exactly once, so it corresponds to a shortest arc in $\mathcal{M}$. This arc cannot cross an edge e of G^* adjacent to one of the two prescribed boundaries, because otherwise it could be shortened by removing the crossing with that edge.

(c) Temporarily fill the boundary with a disk, and assign infinitesimal crossing weights to the edges of the boundary of this disk. Pick a point p inside this disk, and compute a shortest non-contractible or non-separating loop with basepoint p, using an algorithm of Erickson and Har-Peled [15, Lemmas 5.2 and 5.4]. This loop is simple and corresponds to the desired arc on $\mathcal{M}$.

(d) This follows from a result by Cabello et al. [3, Theorem 4.3]. More precisely, they prove that the following algorithm provides a tight non-separating simple cycle: cut the surface along a shortest simple non-separating loop with an arbitrary basepoint; for each of the two boundaries of the resulting surface, compute a shortest cycle homotopic to that boundary; return the shortest of these two cycles. The resulting cycle has multiplicity at most six, since the initial loop has multiplicity at most two, and by Proposition 2.6.[5]

(e) It is known [15, Lemma 5.2] that some shortest non-contractible cycle γ in the annulus is simple, hence generating. We claim that γ passes at most once through a given face of $G^*(\mathcal{M})$. Indeed, otherwise, there would be two points x and y on two different pieces of γ in a given face f^* of $G^*(\mathcal{M})$, such that some path $\overline{xy}$ between x and y belongs to f^* and does not intersect γ except at x and y. These points x and y split γ into two paths, γ_1 and γ_2. The cycle γ is generating, hence non-contractible, hence, for $i = 1$ or $i = 2$, the concatenation of γ_i and of $\overline{xy}$ (maybe with reverse orientation) is non-contractible, and also simple, hence generating; it is strictly shorter than γ; so γ was not a shortest generating cycle. This proves the claim.

[5] In a previous version of this paper, the result by Cabello et al. was not available, and we used a shortest non-separating cycle, computed with an algorithm of Erickson and Har-Peled [15, Lemma 5.4], instead. However, the result by Cabello et al. improves the time complexity of our preprocessing step from $O(n^2 \log n)$ to $O(gn \log n)$.

Thus, γ can be viewed as a circuit in $G(\mathcal{M})$, that is, a closed walk without repeated vertices. So we are actually looking for a shortest generating circuit in $G(\mathcal{M})$. Now, by a result by Reif [25, Propositions 1 and 2], a shortest generating circuit in $G(\mathcal{M})$ is the dual of a minimum cut in $G^*(\mathcal{M})$. We can compute such a minimum cut in $O(n \log n)$ time by using an algorithm of Frederickson [17, Theorem 7]; its dual is the desired cycle, and indeed has multiplicity one.

(f) Colin de Verdière and Lazarus describe an algorithm for this problem [8]. Let δ_1, δ_2, and δ_3 denote the boundaries of $\mathcal{M}$, and suppose we want a cycle homotopic to δ_1. Compute a shortest arc α between δ_2 and δ_3 (by part (b)) and compute the shortest generating cycle γ in the annulus $\mathcal{M}\setminus\alpha$ (by part (e)), of multiplicity one on $\mathcal{M} \setminus \alpha$. This cycle is simple; Proposition 2.4 implies that it is tight in $\mathcal{M}$. Since it has multiplicity one on $\mathcal{M} \setminus \alpha$ and since α has multiplicity one, it has multiplicity at most two on $\mathcal{M}$. ∎

3. Tight Octagonal Decompositions. In this section and the following one, we assume that $\mathcal{M}$ has genus at least two and has no boundary ($g \geq 2$, $b = 0$), and we prove our Main Theorem in this case. We will first describe the preprocessing phase of our path- and cycle-shortening algorithm; the algorithm for tightening curves will be described in the next section.

An *octagonal decomposition* of a surface is an arrangement of simple cycles in which every vertex has degree four and every face has eight sides. See Figure 3.1.

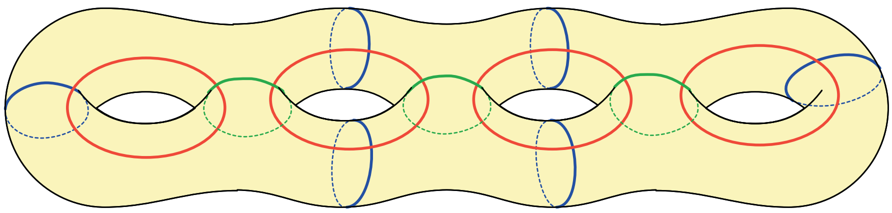

FIG. 3.1. *An octagonal decomposition built by our algorithm.*

If we lift the cycles of an octagonal decomposition to the universal cover of the surface, we obtain a tiling of the plane with octagons, where each vertex has degree four. This tiling is actually isomorphic to the tiling of the hyperbolic plane by regular right-angled octagons; hence, our decomposition imposes a crude regular hyperbolic structure on any cross-metric surface, thereby allowing us to exploit classical results in combinatorial hyperbolic geometry and combinatorial group theory, primarily in §3.3 and §4.1.

3.1. Construction. THEOREM 3.1. *Let $\mathcal{M}$ be a cross-metric surface with complexity n, genus $g \geq 2$, and no boundary. In $O(gn \log n)$ time, we can construct a tight octagonal decomposition of $\mathcal{M}$ in which each cycle has multiplicity $O(1)$.*

Proof. Our construction algorithm has four phases and uses repeatedly Proposition 2.7.

Phase 1: Unzipping. We begin by 'unzipping' the surface into a disk using one cycle and $2g-1$ paths. Let τ_1 be a tight cycle in $\mathcal{M}$ that is tight, simple, non-separating, and has multiplicity at most six (Proposition 2.7(d)). Let β_1 be the shortest arc βetween the two boundary components of $\mathcal{M}\setminus\tau_1$, and let $\mathcal{M}_1 = \mathcal{M}\setminus(\tau_1 \cup \beta_1)$. For each i from 1 to $g-1$, let α_{i+1} be the shortest non-separating αrc in $\mathcal{M}_i$; let β_{i+1} be the shortest arc βetween the two boundaries of $\mathcal{M}_i \setminus \alpha_{i+1}$; and let $\mathcal{M}_{i+1} = \mathcal{M}_i \setminus (\alpha_{i+1} \cup \beta_{i+1})$. For each i, the surface $\mathcal{M}_i$ has genus $g-i$ and one boundary cycle. See Figure 3.2(a). Proposition 2.7 implies that the *union* of the cycle τ_1 and the arcs α_i and β_i has multiplicity at most six on $\mathcal{M}$ (since no arc intersects edges of G^* crossed by the cycle and the arcs created before) and that we can compute these curves in $O(gn \log n)$ total time.

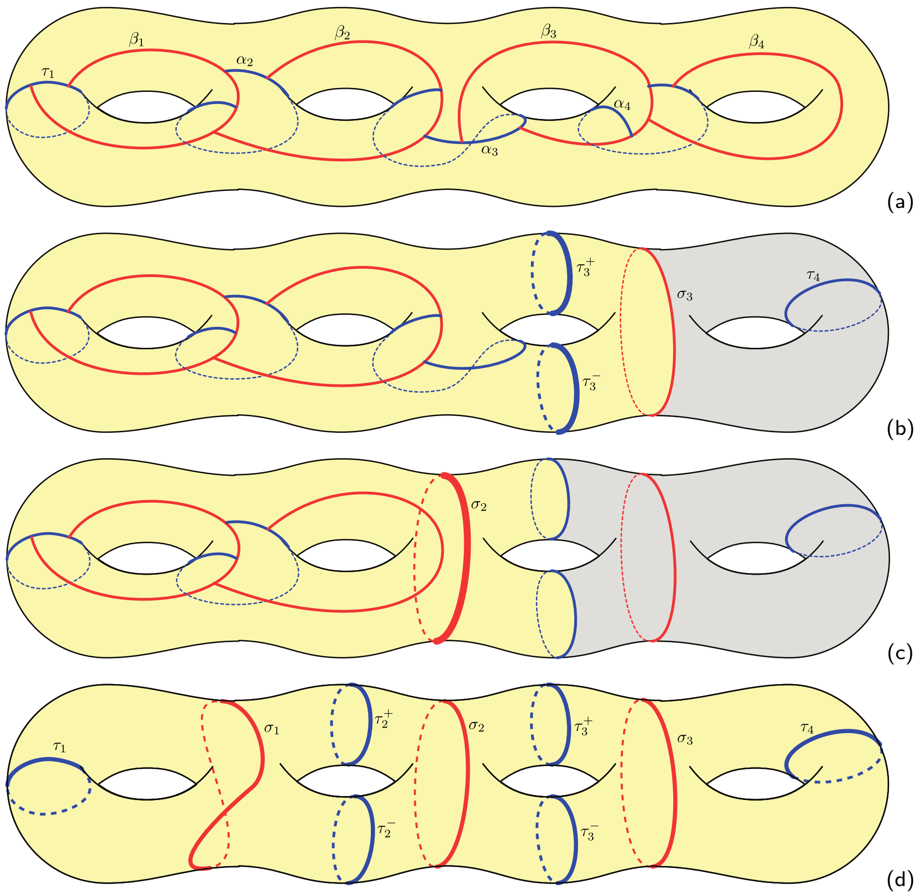

FIG. 3.2. *(a) The surface $\mathcal{M}$ unzipped. (b) Computing τ_3^+ and τ_3^-. (c) Computing σ_2. (d) The final pants decomposition.*

Phase 2: Pants decomposition. Next, we use the curves in the previous phase to help construct a set of $3g-3$ tight simple cycles, each with multiplicity $O(1)$, that decompose $\mathcal{M}$ into $2g-2$ pairs of pants (Figure 3.2). Henceforth, *all the cycles considered are simple and tight*: the tightness in $\mathcal{M}$ follows from Proposition 2.4.

Let τ_g denote the shortest generating cycle in the annulus $\mathcal{M}_{g-1} \setminus \alpha_g$. Let σ_{g-1} be the shortest cycle in $\mathcal{M}_{g-1} \setminus \tau_g$ homotopic to the boundary of $\mathcal{M}_{g-1}$. For each i from $g-1$ down to 2, let τ_i^+ be the shortest cycle in $\mathcal{M}_{i-1} \setminus (\alpha_i \cup \sigma_i)$ homotopic to a boundary of $\mathcal{M}_{i-1} \setminus \alpha_i$; let τ_i^- be the shortest cycle, in the component of $\mathcal{M}_{i-1} \setminus (\alpha_i \cup \sigma_i \cup \tau_i^+)$ that is a pair of pants, homotopic to a boundary of $\mathcal{M}_{i-1} \setminus \alpha_i$; and let σ_{i-1} be the shortest cycle in $\mathcal{M}_{i-1} \setminus (\tau_i^+ \cup \tau_i^-)$ homotopic to the boundary of $\mathcal{M}_{i-1}$. Recall that τ_1 is our original starting cycle. All these cycles are tight in $\mathcal{M}$ by repeated applications of Proposition 2.4.

Every edge of $\mathcal{M}$ is split into at most three sub-edges in $\mathcal{M}_i$ and $\mathcal{M}_{i-1} \setminus \alpha_i$, since the union of the α_i's and β_i's has multiplicity at most 2. Each of the σ_i and $\tau_i^\pm$ is a shortest cycle homotopic to a boundary of $\mathcal{M}_i$ or $\mathcal{M}_{i-1} \setminus \alpha_i$, hence of multiplicity at most 2 on that surface (Proposition 2.6), and thus of multiplicity at most 6 on $\mathcal{M}$. So, in each case, we are computing a shortest cycle homotopic to a boundary in a pair of pants of complexity $O(n)$, which can be done in $O(n \log n)$ time by Proposition 2.7(f).

The $3g-3$ cycles $\tau_1, \sigma_1, \tau_2^+, \tau_2^-, \sigma_2, \ldots, \tau_{g-1}^+, \tau_{g-1}^-, \sigma_{g-1}, \tau_g$ split $\mathcal{M}$ into $2g-2$ pairs of pants. Specifically, the cycles σ_i partition $\mathcal{M}$ into a chain of punctured tori $T_1 \cup T_2 \cup \cdots \cup T_g$, where T_1 and T_g each have one boundary (σ_1 and σ_{g-1}, respectively), and every other T_i has two boundaries (σ_{i-1} and σ_i). The first torus T_1 (resp. the last torus T_g) is cut into a pair of pants by τ_1 (resp. τ_g), and each intermediate torus T_i is cut into two pairs of pants by the cycles τ_i^+ and τ_i^-.

Phase 3: Around the holes. In the next phase, we find tight simple cycles that go 'around the hole' of each punctured torus T_i, crossing the cycle(s) $\tau_i^\pm$ exactly once. Henceforth, we assign infinitesimal weights to the edges of the curves already constructed, allowing to apply Lemma 2.3.

First consider the torus T_1. Let α be the shortest non-contractible arc in $T_1 \setminus \tau_1$ with both endpoints on the boundary σ_1, and let β be the shortest non-contractible arc in $T_1 \setminus \alpha$. Finally, let ϕ_1 be the shortest generating cycle in the annulus $T \setminus \beta$. Because τ_1 is homotopic to the boundary of $T_1 \setminus \alpha$, the arc β crosses τ_1 exactly once, so ϕ_1 also crosses τ_1 exactly once (by applying Lemma 2.3 twice). See Figure 3.3. Since τ_1 and σ_1 each have constant multiplicity, so does ϕ_1.

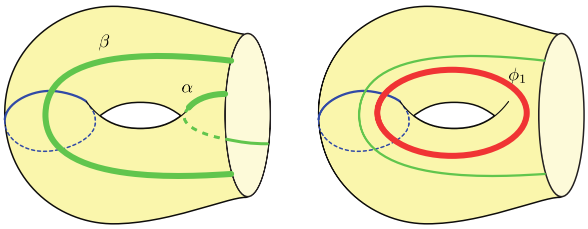

FIG. 3.3. *Left: α and β. Right: ϕ_1.*

A symmetric construction finds a tight cycle ϕ_g in the torus T_g that crosses τ_g exactly once.

Now, for some $2 \le i \le g-1$, consider the torus T_i, whose boundary consists of σ_{i-1} and σ_i. Let α^- be the shortest non-contractible arc with endpoints on σ_{i-1} in $T_i \setminus (\tau_i^+ \cup \tau_i^-)$. Similarly, let α^+ be the shortest non-contractible arc with endpoints on σ_i in $T_i \setminus (\tau_i^+ \cup \tau_i^-)$. These arcs α^- and α^+ split T_i into two annuli, one containing τ_i^+ and the other τ_i^-. Let β^+ and β^- be shortest arcs, one on each of these annuli, joining a point of σ_{i-1} and a point of σ_i. The arc β^+ crosses τ_i^+ once (Lemma 2.3) and does not cross τ_i^-; symmetrically, β^- crosses τ_i^- once and does not cross τ_i^+. Finally, let ϕ_i be the shortest generating cycle in the annulus $T_i \setminus (\beta^+ \cup \beta^-)$; this cycle crosses each of τ_i^+ and τ_i^- exactly once (Lemma 2.3). See Figure 3.4. Since $\tau_i^\pm$ and σ_i each have constant multiplicity, so does ϕ_i.

Proposition 2.7 implies that each curve ϕ_i is computed in time $O(n_i \log n_i)$, where n_i denotes the complexity of T_i. Since $\sum_i n_i = O(n)$, the overall running time of this phase is $O(n \log n)$.

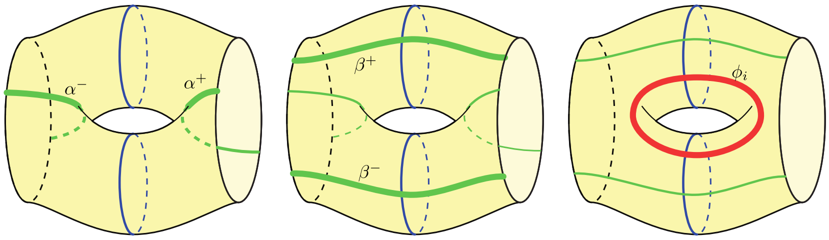

FIG. 3.4. *Left:* α^+ *and* α^-. *Middle:* β^+ *and* β^-. *Right:* ϕ_i.

Phase 4: Around the handles. Finally, for each i between 1 and $g-1$, let M_i be the pair of 'monkey pants' formed by gluing together the two pairs of pants with σ_i as their common boundary. The boundaries of M_i are τ_i^+, τ_i^-, τ_{i+1}^+, and τ_{i+1}^-. (Here τ_1^+ and τ_1^- denote the two copies of τ_1 in M_1, and τ_g^+ and τ_g^- denote the two copies of τ_g in M_{g-1}.) Let γ^+ be the shortest arc in $M_i \setminus (\phi_i \cup \phi_{i+1})$ between τ_i^+ and τ_{i+1}^+, and let γ^- be the shortest arc in $M_i \setminus (\phi_i \cup \phi_{i+1} \cup \gamma^+)$ between τ_i^- and τ_{i+1}^-. The arcs γ^+ and γ^- are disjoint and cross σ_i exactly once (Lemma 2.3 applied twice). Finally, let θ_i be the shortest generating cycle in the annulus $M_i \setminus (\gamma^+ \cup \gamma^-)$. It follows from Lemma 2.3 that θ_i crosses σ_i exactly twice and ϕ_i and ϕ_{i+1} each exactly once. See Figure 3.5. Finally, since all the earlier cycles have constant multiplicity, so does θ_i. As in the previous phase, Proposition 2.7 implies that this phase of the algorithm runs in $O(n \log n)$ time.

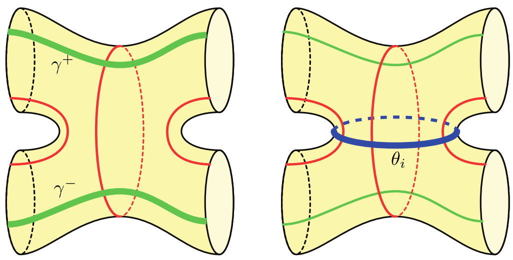

FIG. 3.5. *Left:* γ^+ *and* γ^-. *Right:* θ_i.

To summarize, the tight simple cycles $\tau_i^\pm$, ϕ_i, and θ_i decompose the surface $\mathcal{M}$ into octagons exactly as shown in Figure 3.1. Proposition 2.4 implies that each of these cycles is tight in $\mathcal{M}$. ∎

If we do not discard the cycles σ_i, we obtain a tight *hexagonal* decomposition of $\mathcal{M}$. Our remaining results use an octagonal decomposition, but this hexagonal decomposition could be used as well, applying exactly the same arguments.

3.2. Limiting Crossings. In the actual curve-shortening algorithm, we need to bound the number of times an input curve crosses the cycles in our octagonal decomposition. To that end, we will actually construct our decomposition on a refinement of the input surface $\mathcal{M}$.

Let $G^+ = G^+(\mathcal{M})$ be the graph obtained by overlaying the primal graph $G(\mathcal{M})$ and the dual graph $G^*(\mathcal{M})$ (see Figure 1.1; recall that in the present case, $\mathcal{M}$ has no boundary). The vertices of G^+ are either vertices of G, vertices of G^*, or intersections between an edge e of G and its dual edge e^* in G^*. Each edge of G and dual edge in

G^* is partitioned into two edges in G^+. Finally, each face of G^+ is a quadrilateral.

To treat $\mathcal{M}$ as a cross-metric surface with 'dual' graph G^+, we assign a crossing weight to each edge e^+ of G^+ as follows. If e^+ is contained in an edge of G^*, it has the same crossing weight as that dual edge. If it is contained in an edge of G, its crossing weight is a fixed formal infinitesimal $\varepsilon' << \varepsilon$. Any curves that are tight with respect to G^+ are also tight with respect to G^*. Among all tight curves that cross each other as few times as possible, our algorithms choose curves that cross the edges of G as few times as possible.

We actually apply Theorem 3.1 in this augmented cross-metric surface. In time $O(gn \log n)$, we obtain an octagonal decomposition $\mathcal{O}$ of $\mathcal{M}$ where each cycle is tight, each edge of G^* is crossed $O(1)$ times by each cycle in $\mathcal{O}$, and each edge of G is crossed $O(1)$ times by each cycle in $\mathcal{O}$. In particular, any walk in G of length k crosses the cycles in $\mathcal{O}$ a total of $O(gk)$ times.

3.3. Universal Cover. Let $\mathcal{O}$ be a tight octagonal decomposition of a surface $\mathcal{M}$ without boundary. As we mentioned earlier, the universal cover of this decomposition is a four-valent octagon tiling of the plane, as is the regular tiling of the hyperbolic plane by right-angled octagons (see Figure 3.6), which can also be viewed as an infinite arrangement of hyperbolic lines. Building on this intuition, we call any lift of a cycle in $\mathcal{O}$ to the universal cover $\widetilde{\mathcal{M}}$ a *line*. The set of lines is denoted by $\widetilde{\mathcal{O}}$. This terminology is further motivated by Corollary 3.3 and Lemma 3.5.

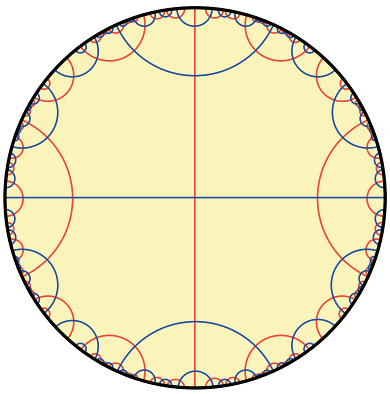

FIG. 3.6. *Universal cover of an octagonal decomposition.*

LEMMA 3.2 (Dehn [9]; see also Stillwell [27, p. 188]). *Let S be the non-empty union of finitely many octagons in the four-valent octagon tiling of the plane. Some octagon in S has five consecutive sides on the boundary of S.*

A trivial but important corollary is:

COROLLARY 3.3. *Let S be the non-empty union of finitely many octagons in the four-valent octagon tiling of the plane. Then at least five distinct lines contain edges on the boundary of S. In particular, two lines in the tiling cross at most once.*

The *perimeter* of a set of octagons is the number of edges on its boundary.

LEMMA 3.4. *Any union of N octagons, $1 \leq N < \infty$, in the four-valent octagon tiling of the plane has perimeter at least $2N + 6$.*

Proof. Removing an octagon with at least five consecutive sides on the boundary of the union (Lemma 3.2) reduces the perimeter by at least two. The base case is a single octagon. ∎

LEMMA 3.5. *Let $\tilde{p}$ be a path in $\widetilde{\mathcal{M}}$, with endpoints x and y; let L be the set of lines in $\widetilde{\mathcal{O}}$ crossed an odd number of times by $\tilde{p}$. Let $\tilde{p}'$ be a shortest path with endpoints x and y, where the lines in $\widetilde{\mathcal{O}}$ are assigned infinitesimal crossing weight ε. Then $\tilde{p}'$ crosses exactly once each line in L and no other line.*

Proof. Since each line ℓ in $\widetilde{\mathcal{O}}$ separates the plane (Lemma 2.1), a path connecting x and y crosses ℓ an odd number of times if and only if ℓ separates x and y. Hence it suffices to prove that $\tilde{p}'$ crosses each line at most once.

If $\tilde{p}'$ crosses some line ℓ at least twice, at points u and v, then $\tilde{p}'$ and ℓ form a bigon. Since ℓ is a lift of a tight cycle, every subpath of ℓ is a shortest path, by Proposition 2.5. Thus, we can remove the bigon from $\tilde{p}'$ by replacing the subpath from u to v with the shortest path in ℓ. Since any pair of lines in $\widetilde{\mathcal{O}}$ intersect at most once, this exchange results in a path with fewer line crossings (and possibly shorter length), which is impossible. $\square$

4. Tightening Curves on Surfaces Without Boundary. In this section, we explain how to compute a tight path or cycle homotopic to a given path or cycle in polynomial time, again in the case $g \geq 2$ and $b = 0$. Our algorithm is much faster than previous results [7, 8, 6], and unlike those results, it does not require the input curve to be simple.

Consider an arbitrary path p on a surface $\mathcal{M}$. Let $\tilde{p}$ be a lift of p to the universal cover $\widetilde{\mathcal{M}}$, and let $\tilde{p}'$ be a shortest path in $\widetilde{\mathcal{M}}$ between the endpoints of $\tilde{p}$. Projecting $\tilde{p}'$ back down to $\mathcal{M}$ gives us a shortest path homotopic to p. Our algorithm exploits this characterization by constructing a subset of $\widetilde{\mathcal{M}}$ of small complexity that contains both $\tilde{p}$ and some shortest path $\tilde{p}'$. Compared to previous approaches [26, 11], the construction of this part of the universal cover is very simple once we have computed an octagonal decomposition of the surface.

4.1. Building the Relevant Region. Let $\mathcal{O}$ be the tight octagonal decomposition of $\mathcal{M}$ as computed in §3. Consider a path p in $\mathcal{M}$, and let $\tilde{p}$ be a lift of p to the universal cover $\widetilde{\mathcal{M}}$. For any line ℓ in $\widetilde{\mathcal{O}}$, let ℓ^+ denote the component of $\widetilde{\mathcal{M}} \setminus \ell$ that contains the starting point $\tilde{p}(0)$.

Let $\ell_1, \ell_2, \ldots, \ell_z$ be the sequence of lines in $\widetilde{\mathcal{O}}$ crossed by $\tilde{p}$, in order of their first crossing. Let $\mathcal{L}_0 = \emptyset$, and for any integer i between 1 and z, let $\mathcal{L}_i = \mathcal{L}_{i-1} \cup \{\ell_i\}$. For each i, let $\widetilde{\mathcal{M}}_i$ be the subset of $\widetilde{\mathcal{M}}$ reachable from $\tilde{p}(0)$ by crossing only (a subset of) lines in $\mathcal{L}_i$, in any order. Combinatorially, the region $\widetilde{\mathcal{M}}_i$ is a 'convex polygon' formed by intersecting the 'half-planes' ℓ^+ for all lines ℓ not in the set $\mathcal{L}_i$. By Lemma 3.5, some shortest path $\tilde{p}'$ between the endpoints of $\tilde{p}$ crosses only a subset of the lines that $\tilde{p}$ crosses; so $\tilde{p}'$ is contained in $\widetilde{\mathcal{M}}_z$. For this reason, $\widetilde{\mathcal{M}}_z$ is called the *relevant region* of $\widetilde{\mathcal{M}}$ (with respect to $\tilde{p}$).

LEMMA 4.1. *For any line ℓ and any $i \geq 0$, $\ell \cap \widetilde{\mathcal{M}}_i$ is either empty or connected.*

Proof. Let $\ell[x, y]$ be the segment of ℓ between two points x and y in $\ell \cap \widetilde{\mathcal{M}}_i$, and suppose some line ℓ' crosses $\ell[x, y]$. Since two lines cross at most once, the points x and y are on different sides of ℓ'. Since ℓ' separates the plane but $\widetilde{\mathcal{M}}_i$ is connected, this line must be in the set $\mathcal{L}_i$. It follows that the entire segment $\ell[x, y]$ belongs to $\widetilde{\mathcal{M}}_i$. $\square$

Since ℓ_i separates $\widetilde{\mathcal{M}}$, Lemma 4.1 implies that ℓ_i intersects $\widetilde{\mathcal{M}}_{i-1}$ on its boundary, along a connected set of octagons $O_1, O_2, \ldots, O_u$. For each j between 1 and u, let O'_j be the reflection of O_j across ℓ_i. See Figure 4.1. The octagons O'_j do not belong to $\widetilde{\mathcal{M}}_{i-1}$.

LEMMA 4.2. $\widetilde{\mathcal{M}}_i = \widetilde{\mathcal{M}}_{i-1} \cup O'_1 \cup \cdots \cup O'_u$.

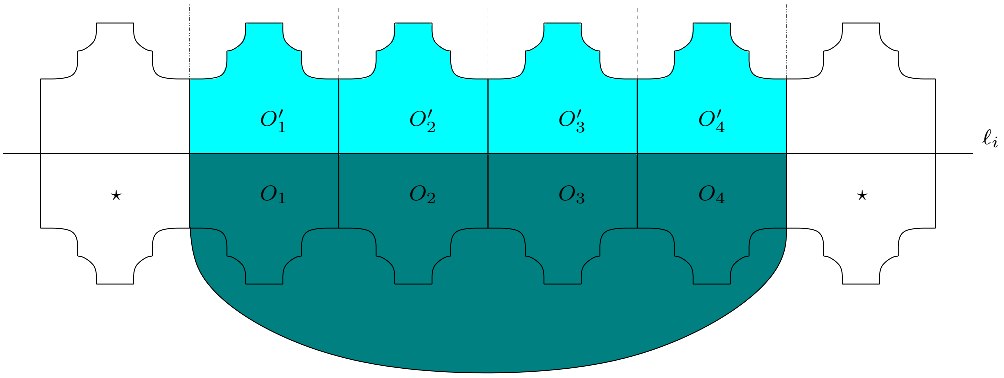

FIG. 4.1. *From $\widetilde{\mathcal{M}}_{i-1}$ (dark shaded) to $\widetilde{\mathcal{M}}_i$ (all shaded).*

Proof. Let $\widetilde{\mathcal{N}} = O'_1 \cup \cdots \cup O'_u$ (the lightly shaded region in Figure 4.1). To prove the lemma, it suffices to show that none of the lines bounding $\widetilde{\mathcal{N}}$ are in the set $\mathcal{L} = \{\ell_1, \ldots, \ell_{i-1}\}$. Obviously ℓ_i is not in this set. Each octagon O'_j is bounded by eight lines: ℓ_i, two *inner* lines that cross ℓ_i at a vertex of O'_j, and five *outer* lines.

If some outer line ℓ of some octagon O'_j intersected ℓ_i, then it would also intersect an inner line $\ell' \neq \ell$ of O'_j. Thus the lines ℓ, ℓ', and ℓ_i would pairwise intersect, and these three lines would bound a disk in the tiling $\widetilde{\mathcal{O}}$, contradicting Lemma 3.2.

Hence, since every line in $\mathcal{L}$ has a point in ℓ_i^+, no outer line can be in $\mathcal{L}$. Only the first and last inner lines contribute a side to the boundary of $\widetilde{\mathcal{N}}$. Neither of these two lines is in $\mathcal{L}$, for otherwise one of the starred octagons in Figure 4.1 would also belong to $\widetilde{\mathcal{M}}_{i-1}$. ◻

LEMMA 4.3. *$\widetilde{\mathcal{M}}_z$ contains at most $7z + 1$ octagons.*

Proof. Let v be a vertex on the boundary of $\widetilde{\mathcal{M}}_z$. Depending on whether zero or one line incident to v belong to $\{\ell_1, \ldots, \ell_z\}$, either one or two octagons incident to v belong to $\widetilde{\mathcal{M}}_z$. In the former case, we say that v is an *extremal boundary vertex*, and in the latter case, we say that v is a *flat boundary vertex*. If there is no flat boundary vertex, then there is exactly one octagon, and the lemma holds.

Every flat boundary vertex is the intersection of some line ℓ_i with the boundary of $\widetilde{\mathcal{M}}_z$. There are at most $2z$ such vertices by Lemma 4.1. Between two consecutive flat boundary vertices, there are trivially at most 6 extremal boundary vertices, all on the boundary of the same octagon. Thus, the perimeter of $\widetilde{\mathcal{M}}_z$ is at most $14z$. The lemma now follows directly from Lemma 3.4. ◻

Constructing the relevant region $\widetilde{\mathcal{M}}_z$ is now straightforward. $\widetilde{\mathcal{M}}_0$ is a copy of the octagon containing $p(0)$, the starting point of p. To compute $\widetilde{\mathcal{M}}_i$, we follow $\tilde{p}$ until it exits the previous region $\widetilde{\mathcal{M}}_{i-1}$. At the exit point, the path is crossing ℓ_i into some octagon O'_j (with the notation of Figure 4.1). To complete $\widetilde{\mathcal{M}}_i$, we append the octagons $O'_1, \ldots, O'_u$.

4.2. Tightening Paths. We now prove the Main Theorem in the case $g \geq 2$, $b = 0$, and if the input curve is a path.

The preprocessing consists of building the tight octagonal decomposition $\mathcal{O}$ on the cross-metric surface defined by $G^+(\mathcal{M})$. Let x be the number of crossings of p with $\mathcal{O}$; as we argued earlier, $x = O(gk)$. Let $\tilde{p}$ be an arbitrary lift of p to the universal cover $\widetilde{\mathcal{M}}$ of $\mathcal{M}$. We first compute the relevant region of $\widetilde{\mathcal{M}}$ with respect to $\tilde{p}$, ignoring the internal structure of the surface within each octagon of $\widetilde{\mathcal{O}}$. In other words, we construct a subset of the abstract regular octagon tiling. The construction is described in Lemma 4.2. We require only constant time for every edge of p, every crossing between p and $\mathcal{O}$, and every relevant octagon; Lemma 4.3 shows that this phase takes time $O(k + x)$.

Now let L be the set of lines crossed an odd number of times by $\tilde{p}$. By Lemma 3.5, there is a shortest path $\tilde{q}$ with the same endpoints as $\tilde{p}$ that crosses each line in L exactly once and no other line. Let x' be the number of lines in L. We now compute the set of octagons accessible from $\tilde{p}(0)$ by crossing only these lines, this time building also the internal surface structure of the octagons. There are $O(x')$ such octagons by Lemma 4.3; since each cycle in $\mathcal{O}$ has constant multiplicity, each octagon has complexity $O(n)$. Thus, the relevant region has total complexity $O(x'n)$ and can be constructed in time $O(x'n)$.

Finally, we compute a shortest path $\tilde{p}'$ between the endpoints of $\tilde{p}$ in this relevant region, in time $O(x'n)$ using the planar shortest path algorithm by Henzinger et al. [20]. By giving the lines L infinitesimal crossing weight, we guarantee that $\tilde{p}'$ crosses each line in L exactly once. The projection $p' = \pi(\tilde{p}')$ onto $G(\mathcal{M})$ is the desired output path. The total time spent is $O(k + x + x'n)$.

The complexity k' of p' (the number of edges of G^* crossed by p') is at most $O(x'n) = O(xn) = O(gnk)$. To complete the time analysis, we observe that $x' = O(g \min\{k, k'\})$.

4.3. Tightening Cycles. We still consider the case $g \geq 2$, $b = 0$, and we prove the Main Theorem for cycles. Our algorithm uses similar ideas as for paths, but is more complicated. As we saw, tightening a path can be done by (1) computing a lift of that path in the universal cover, (2) computing a shortest path between the endpoints of that lift, in a relevant region of the universal cover, and (3) projecting back to the surface. To tighten a cycle, we will (1) compute a lift of that cycle in the annular cover generated by that cycle, (2) compute a shortest generating cycle in a relevant region of that annulus using Proposition 2.7(e), and (3) project back to the surface. The main difficulty is to compute the region of the annular cover that is relevant for our purposes. Before describing the algorithm, we have to explain some structural properties of the annular cover.

Consider a non-contractible cycle γ; let $\widehat{\mathcal{M}}$ be the annular cover generated by γ. Let $\widehat{A} \subseteq \widehat{\mathcal{M}}$ be a union of octagons of $\widehat{\mathcal{M}}$ that contains a shortest generating cycle in $\widehat{\mathcal{M}}$ and is, topologically, an annulus (for example, $\widehat{A} = \widehat{\mathcal{M}}$). Since the cycles of the octagonal decomposition $\mathcal{O}$ are simple, the maximal pieces of their lifts in $\operatorname{Int} \widehat{A}$ are simple curves of three possible types: (1) simple generating cycles on $\widehat{A}$; (2) simple arcs with both endpoints on the same boundary component of $\widehat{A}$; and (3) simple arcs with one endpoint on each of the boundary components of $\widehat{A}$.

Lemma 4.4. *Any shortest generating cycle in $\widehat{A}$ (using infinitesimal crossing weights for the lifts of the cycles in $\mathcal{O}$) crosses no curve of type (1) and (2), and crosses exactly once every arc of type (3).*

Proof. By the assumption on $\widehat{A}$, a shortest generating cycle in $\widehat{A}$ is also a shortest generating cycle in $\widehat{\mathcal{M}}$. Let $\widehat{c}$ be a lift in $\widehat{A}$ of a cycle in $\mathcal{O}$.

If $\widehat{c}$ is of type (1), then a generating cycle crossing $\widehat{c}$ lifts, in $\widetilde{\mathcal{M}}$, to a curve crossing a given line several (actually infinitely many) times; thus such a cycle cannot be a shortest homotopic cycle by Lemma 3.5 and Proposition 2.5.

If $\widehat{c}$ is of type (2), it separates $\widehat{A}$ into an annulus and a disk. If a generating cycle crosses $\widehat{c}$, a maximal part of it inside the disk forms a bigon with $\widehat{c}$. This implies that the cycle lifts in $\widetilde{\mathcal{M}}$ to a curve crossing a lift of $\widehat{c}$ at least twice, hence, again by Lemma 3.5 and Proposition 2.5, this cycle cannot be a shortest homotopic cycle.

If $\widehat{c}$ is of type (3), then a generating cycle in $\widehat{A}$ has to cross it an odd number of times. If some generating cycle crosses it at least twice, then it must cross it twice consecutively in opposite directions. Then, as in the previous case, this generating cycle is not as short as possible in its homotopy class. ∎

Corollary 4.5. *In $\widehat{\mathcal{M}}$, there exists at least one lift of a cycle in $\mathcal{O}$ with one endpoint on each boundary component of $\widehat{\mathcal{M}}$.*

Proof. Otherwise, by Lemma 4.4, the shortest generating cycle in $\widehat{\mathcal{M}}$ would cross no lift of the lines in $\widetilde{\mathcal{O}}$. But, on $\mathcal{M}$, the cycles crossing no curve of the octagonal decomposition are contractible. ∎

Let γ be the input cycle of our algorithm. We choose an arbitrary point u on γ and view temporarily γ as a loop with basepoint u. We consider an arbitrary lift u_0 of u in $\widetilde{\mathcal{M}}$. Let $(u_i)_{i\in\mathbb{Z}}$ be points in $\widetilde{\mathcal{M}}$ such that u_{i+1} is the target of the lift of γ starting at u_i. (Hence the u_i's are uniquely determined once u_0 has been fixed.)

As above, x and x' denote the number of crossings of the input and output cycles with $\mathcal{O}$; let $\bar{x} = \min\{x, x'\}$. Again, in the first phase of the algorithm, we ignore the $O(n)$ internal complexity of the octagons. We first build the relevant part of the universal cover $\widetilde{\mathcal{M}}$ associated to a lift of γ^3 starting at u_0 (that is, the set of octagons one can reach from the source of this lift by crossing only lines crossed by this lift); Lemma 4.2 explains how to do this computation, and Lemma 4.3 shows that this takes $O(k + x)$ time. We also build points u_0, u_1, u_2, and u_3.

We can easily identify in $O(x)$ time the set of lines separating u_1 and u_2 but not u_0 and u_1 nor u_2 and u_3, which is non-empty by Corollary 4.5 (unless the input cycle is contractible, in which case the algorithm is trivial). Let ℓ_2 be such a line, oriented such that u_1 is on its left and u_2 is on its right. As above, for each $i \in \mathbb{Z}$, we define the lines ℓ_i by translation of ℓ_2 (in particular, ℓ_i has u_{i-1} on its left and u_i on its right). We can find ℓ_3 in $O(x)$ time.

Lemma 4.6. *For each i, ℓ_i has ℓ_{i-1} on its left and ℓ_{i+1} on its right.*

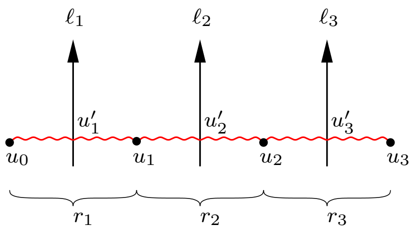

Fig. 4.2. *The notations for the proof of Lemma 4.6.*

Proof. See Figure 4.2. It follows from the definition that ℓ_i is different from ℓ_{i-1} and ℓ_{i+1}. Since the ℓ_j are all lifts of a given simple cycle, two ℓ_j must be disjoint, unless they are identical.

Let r_2 be a shortest path from u_1 to u_2, with infinitesimal crossing weights for the lines in $\widetilde{\mathcal{O}}$. By Lemma 3.5, r_2 crosses every line at most once. Let u'_2 be the intersection point of r_2 with ℓ_2. We also define r_j and u'_j, $j \in \mathbb{Z}$, by translation of r_2 and u'_2.

Since, by definition, ℓ_2 does not separate u_2 and u_3, it does not cross r_3, so u'_3 is on the right of ℓ_2. Hence ℓ_3 is on the right of ℓ_2. Similarly, ℓ_1 is on the left of ℓ_2. The lemma follows by translation. ☐

By Lemma 4.6, we are in the situation depicted in Figure 4.3. Let v_2 be an intersection point of ℓ_2 with the lift of γ between u_1 and u_2; define the points v_j by translation of v_2. In particular, we can determine in $O(k+x)$ time the position of v_2 and v_3 in the abstract octagon tiling. Let p_i be the part of the lift of γ connecting v_i to v_{i+1}. We are looking for a shortest cycle homotopic to the cycle corresponding to the projection of p_2 onto $\mathcal{M}$.

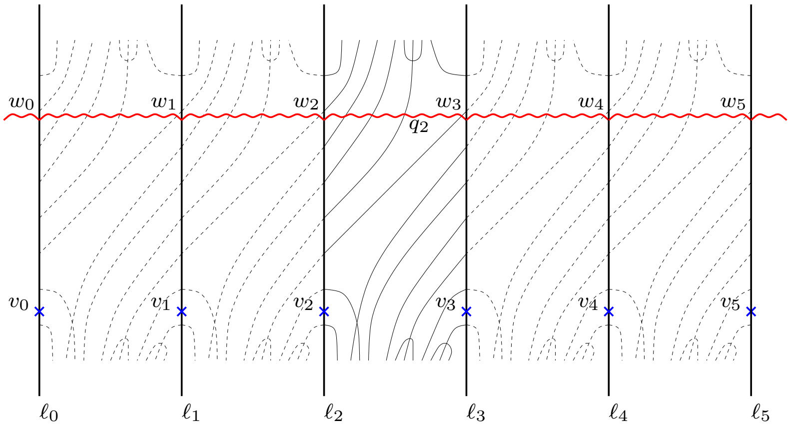

FIG. 4.3. *The part of the universal cover containing the lifts of γ and δ, and the lines ℓ_i. Other lines in this area are also shown.*

We consider the strip S of the universal cover that is the region comprised between lines ℓ_2 and ℓ_3, and we identify these two lines so that the points v_2 and v_3 are identified. We thus obtain the annular cover $\widehat{\mathcal{M}}$ of $\mathcal{M}$. The path p_2 projects, in $\widehat{\mathcal{M}}$, to a cycle $\widehat{\gamma}$, which itself projects to γ in $\mathcal{M}$. Finding the shortest cycle homotopic to γ now boils down to finding the shortest cycle homotopic to $\widehat{\gamma}$ in $\widehat{\mathcal{M}}$. The idea is to determine a subset of small complexity of $\widehat{\mathcal{M}}$ that is relevant for this computation.

We still ignore the $O(n)$ internal structure of the surface, only manipulating abstract octagons. We consider a set R of octagons in $\widetilde{\mathcal{M}}$ that is the intersection of the following two regions of $\widetilde{\mathcal{M}}$:

- the strip S, and
- the relevant space of the concatenation $p_{0,\dots,4}$ of the p_i's for $i = 0, \dots, 4$, as described in §4.1. Recall that this is the set of octagons of $\widetilde{\mathcal{M}}$ reachable from v_0 by crossing only lines crossed by $p_{0,\dots,4}$. This region has complexity $O(x)$ by Lemma 4.3.

Identifying the portions of ℓ_2 and ℓ_3 of R so as to make p_2 a cycle, we obtain an annular region $\widehat{A}$ that is a subset of the annular cover $\widehat{\mathcal{M}}$. Again, this takes $O(k+x)$ time; R has complexity $O(x)$.

LEMMA 4.7. *Some shortest cycle homotopic to γ (using infinitesimal crossing weights for the cycles in $\mathcal{O}$) has a lift that is a generating cycle in $\widehat{A}$.*

Proof. Refer to Figure 4.3. Let δ be a shortest cycle homotopic to γ, assuming infinitesimal crossing weights for the cycles of the octagonal decomposition. The homotopy from γ to δ lifts, in $\widehat{\mathcal{M}}$, to a homotopy from $\widehat{\gamma}$ to a generating cycle $\widehat{\delta}$ that is a lift of δ. This generating cycle crosses exactly once the projection of the lines ℓ_i, by Lemma 4.4, so, in S, it corresponds to a path q_2 from a point $w_2 \in \ell_2$ to a point

$w_3 \in \ell_3$. As for γ, we construct a sequence $(q_i)_{i\in\mathbb{Z}}$ of lifts in $\widetilde{\mathcal{M}}$ of the projection of q_2, and a sequence $(w_i)_{i\in\mathbb{Z}}$ of points, such that q_i has endpoints w_i and w_{i+1}.

No line crosses the concatenation of the q_i's more than once, because otherwise δ would not be tight by Proposition 2.5 and Lemma 3.5. Hence any line disjoint from q_2 and separating q_2 from both v_0 and v_5 crosses four consecutive lines ℓ_i. Also, if a line crosses q_2, then it either crosses $p_{0,\ldots,4}$, or it crosses three consecutive lines ℓ_i. It follows that q_2 belongs entirely to R (which concludes the proof), unless a line ℓ crosses three consecutive lines ℓ_i; so we can now assume that it is the case.

There is at most one line that crosses both ℓ_2 and ℓ_3, for otherwise there would be, in the four-valent octagon tiling, a region bounded by four lines, which is impossible by Corollary 3.3. Thus, since ℓ crosses three consecutive lines ℓ_i, the translation of the universal cover that takes v_i to v_{i+1} leaves ℓ invariant. This implies that there is exactly one such line and that ℓ is a lift of a cycle homotopic to γ. This cycle is tight by Proposition 2.5. Now, if we take q_2 to run along this line, on the adequate side, the same arguments as in the previous paragraph show that q_2 belongs to R. This concludes the proof. □

Since every generating cycle in $\widehat{A}$ projects to a cycle homotopic to γ, it suffices to compute a shortest generating cycle in $\widehat{A}$. To improve the time complexity, we further discard some unnecessary pieces of this annulus.

The maximal parts of the lifts of the cycles in $\mathcal{O}$ in $\operatorname{Int}\widehat{\mathcal{M}}$ are of three possible types, according to their intersection with $\widehat{A}$, as explained above. Using a traversal of the arrangement of the octagons on $\widehat{A}$, we can easily determine in $O(x)$ time the type of each curve. If there is a curve of type (1), then it is tight by Proposition 2.5, so the output of the algorithm is its projection on $\mathcal{M}$, which can be computed in $O(k + x + n\bar{x}) = O(gk + gn\bar{k})$ time.

So we assume there are only curves of type (2) and (3). Let $\widehat{\delta}$ be the shortest generating cycle of the annular part $\widehat{A}$, using infinitesimal crossing weights on the octagonal decomposition. By Lemma 4.4, $\widehat{\delta}$ does not cross any curve of type (2).

Given a curve of type (2), we run a tandem search on the components of the complement of this arc to compute the Euler characteristic of one of them and determine which component is a disk and which component is an annulus. We discard this disk from the surface, since we know $\widehat{\delta}$ does not enter it. The type of the curves is unchanged when cutting off the disk from the annulus. We iterate this process over all curves of type (2). The time necessary for each of these operations is linear in the complexity of the part discarded, so this takes $O(x)$ total time.

Finally, we obtain an annular subset $\widehat{A}'$ of $\widehat{\mathcal{M}}$ containing δ and made of curves of type (3) only; note that each of them has to be crossed by the output cycle. We build the $O(n)$ internal structure of these octagons and compute the shortest generating cycle of $\widehat{A}'$; the output of the algorithm is its projection on $\mathcal{M}$.

Each of the $O(x')$ curves of type (3) in $\widehat{A}'$ crosses the projection of ℓ_2 on $\widehat{A}'$ at most twice, by a reasoning analogous to the end of the proof of Lemma 4.7. Thus every such curve of type (3) in $\widehat{A}'$ corresponds, in the strip S, to at most three different lines. The annulus $\widehat{A}'$ is included in the relevant space built by allowing to cross these $O(x')$ lines, and thus has complexity $O(x'n)$ by Lemma 4.3. So the running time for this computation is $O(nx' \log(nx'))$ by Proposition 2.7(e). The total time complexity is $O(x + nx' \log(nx'))$. Recall that $x = O(gk)$ and $x' = O(gk')$; furthermore, we have $x' = O(x)$; this completes the proof.

5. Tightening Curves on a Torus. We here prove the Main Theorem in the case where $\mathcal{M}$ is a torus ($g = 1$, $b = 0$). The ingredients are essentially similar to the case $g \geq 2$, $b = 0$, so we only briefly indicate the differences.

It is not difficult to prove that a torus cannot have an octagonal decomposition, using Euler's formula and the fact that the Euler characteristic of the torus is zero. Instead of an octagonal decomposition, a *quadrilateral* decomposition can be used. It is comprised of two tight non-separating cycles that cross exactly once, which we can compute in $O(n \log n)$ time as follows. The first cycle γ_1 is a tight non-separating cycle, obtained by Proposition 2.7(d). Cutting along it yields an annulus; each point on a boundary has a "twin" point on the other one. We can compute the $O(n)$ shortest arcs in this annulus between each pair of twin vertices; the shortest such arc gives the second cycle γ_2. The multiplicity is $O(1)$. These shortest paths can be computed in $O(n \log n)$ total time using an algorithm by Klein [22].

The polygonal schema of this quadrilateral decomposition has four sides, and the lift of the quadrilateral decomposition in the universal cover is combinatorially isomorphic to the Euclidean grid. The major difference, compared to the hyperbolic case, is that the relevant region, obtained by allowing to cross x lines, has complexity $O(x^2)$, and not necessarily $O(x)$ as in Lemma 4.3, hence the difference in the time complexity.

For $i = 1, 2$, let m_i be the number of times the input path or cycle crosses γ_i from left to right minus the number of times it crosses it from right to left; this can be computed in $O(k + x)$ time. For clarity of the exposition we assume m_1 and m_2 to be nonnegative; we have $m_1 \cdot m_2 = O(x^2)$. These two numbers determine uniquely the homotopy class of the input curve, since the fundamental group of the torus is Abelian.

If the input curve is a path p, we build an $(m_1 + 1) \times (m_2 + 1)$-grid, also building the $O(n)$ internal structure of each quadrilateral. Then we compute the shortest path between the lift of $p(0)$ in the leftmost lower square and the lift of $p(1)$ in the rightmost upper square. Its projection is the desired shortest homotopic path, by a reasoning analogous to Lemma 3.5. This takes $O(nx'^2)$ time. The total running time is $O(k + x + nx'^2)$, where $x = O(k)$, $x' = O(k)$, and $x' = O(k')$; hence, in total, this is $O(k + n\bar{k}^2)$.

If the input curve is a cycle γ, we can assume $(m_1, m_2) \neq (0, 0)$; for example, $m_1 \neq 0$. We build an $m_1 \times (m_2 + 1)$-grid, and we glue together the vertical side incident to the leftmost lower vertex to the vertical side incident to the rightmost upper vertex; this gives an annulus $\widehat{A}$ that is a part of the annular cover $\widehat{\mathcal{M}}$ with respect to the cycle γ. Some shortest cycle γ' homotopic to γ must have a lift, in the annular cover, that belongs actually to this region $\widehat{A}$. We compute the shortest generating cycle using Proposition 2.7(e). The total time spent is $O(k + n\bar{k}^2 \log(n\bar{k}^2))$.

6. Surfaces with Boundary. In this section, we consider the case where the surface has at least one boundary component: $b \geq 1$. We remark that the algorithm for surfaces without boundary gives, without much effort, an algorithm for surfaces with boundary. Indeed, given the cross-metric surface with boundary $\mathcal{M}$, attach a handle to each boundary cycle, obtaining a surface without boundary $\overline{\mathcal{M}}$, and assign infinite crossing weights to the edges of the handles and of the original boundary cycles. Then two curves in $\mathcal{M} \subset \overline{\mathcal{M}}$ are homotopic in $\mathcal{M}$ if and only if they are homotopic in $\overline{\mathcal{M}}$, and a shortest curve homotopic to a given curve in $\mathcal{M}$ must be in $\mathcal{M}$. So it suffices to tighten the curve in $\overline{\mathcal{M}}$.

However, this approach is a bit artificial. In this section, we propose a simpler approach. We also preprocess the surface with a topological decomposition; but this decomposition is made of *disjoint, simple* curves, in contrast to the case of surfaces without boundary, where crossings were necessary in the octagonal decomposition. Also, no hyperbolic geometry is involved, and the preprocessing step has running time $O(n \log n + (g+b)n)$, which is slightly better than the $O((g+b)n \log n)$ running time of the approach sketched above.

As for the case of surfaces without boundary, a preliminary step builds a decomposition of the surface with tight curves, allowing the shortening algorithm to explore a portion of the universal cover of reasonable complexity.

6.1. Tight Systems of Paths. We first describe a slight extension of a paper by Erickson and Whittlesey [16]. Let $\mathcal{N}$ be a cross-metric surface with complexity n, genus g, and without boundary. Let $P = \{p_1, \dots, p_h\}$ be a set of points on $\mathcal{N}$. A *system of paths with vertex set* P is a set of simple paths or loops with endpoints in P that are pairwise disjoint except possibly at common endpoints, cut the surface into a disk, and such that every point in P is an endpoint of at least one path. By Euler's formula, any system of paths contains exactly $2g + h - 1$ paths, and the boundary of its associated polygonal schema has $4g + 2h - 2$ edges.

LEMMA 6.1. *In $O(n \log n + (g+h)n)$ time, we can compute a tight system of paths with vertex set P on $\mathcal{N}$, in which each path has multiplicity at most two.*

Proof. Our algorithm is a variant of the algorithm by Erickson and Whittlesey [16] to compute the shortest system of loops, based at some given point p, of a cross-metric surface $\mathcal{N}$ without boundary. This algorithm maintains a set of loops L, initially empty, and iteratively adds to L the shortest loop ℓ with basepoint p such that $\mathcal{N} \setminus (L \cup \{\ell\})$ is connected. This greedy process actually provides a shortest system of loops with basepoint p; each loop is tight and has multiplicity at most two. It can be implemented so as to run in $O(n \log n + gn)$ time by a single-source shortest path tree T rooted at p with Dijkstra's algorithm: the candidate loops are made of a shortest path in T with starting point p, a single crossing of an edge in $G \setminus T$, and a shortest path in T with terminal point p.

In our variant, we run Dijkstra's algorithm from all points in P simultaneously, obtaining a shortest path forest F. The candidate loops are made of a shortest path in F with starting point in P, a single crossing of an edge in $G \setminus F$, and a shortest path in F with final endpoint in P. All the arguments showing that this greedy process terminates and that every resulting path is tight and has multiplicity at most two carry through without modification. This takes $O(n \log n + (g+h)n)$ time. ☐

We call the tight system of paths obtained in the previous proposition the *greedy system of paths* with vertex set P. Let S be such a greedy system. Cutting $\mathcal{N}$ along S also cuts the edges of G^* into *sub-edges*. A sub-edge is *external* if at least one of its endpoints is an endpoint of the original edge of G^*; otherwise, this sub-edge is *internal*. The *corners* of a system of paths with vertex set P are the copies of the points in P on the boundary of the associated polygonal schema. The *sides* of a system of paths are the parts of the boundary of the polygonal schema between consecutive corners.

LEMMA 6.2. *Let S be a greedy system of paths on $\mathcal{N}$; let D be the associated polygonal schema. Then the shortest path between any two points in D has multiplicity at most four on $\mathcal{N}$.*

Proof. Let $q_1, \dots, q_{4g+2h-2}$ be the corners of D, in this cyclic order around the boundary of D. Recall that the greedy system of paths is built using a shortest path forest F on $\mathcal{N}$: each path is the concatenation of two shortest paths in F plus a crossing of an edge that is outside F and belongs to no other path of the system.

Internal edges come from the fact that two such shortest paths may share edges in F and are thus running along each other. In particular, an internal sub-edge necessarily connects adjacent sides of D, that is, for some i, the side $q_i q_{i-1}$ of D to the side $q_i q_{i+1}$ (here all indices are taken modulo $4g + 2h - 2$); see Figure 6.1. In that case, the internal sub-edge is said to *belong to corner* q_i.

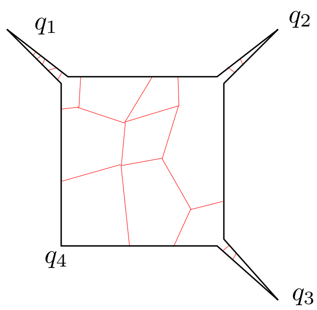

FIG. 6.1. *A polygonal schema and the sub-edges. The internal sub-edges have both endpoints on the boundary of the polygonal schema, while the external sub-edges have at least one endpoint at a vertex of* G^*.

Let p be a shortest path in D; we claim that it crosses internal sub-edges belonging to at most two corners. For let q_i and q_j be the corners to which belong the first and last internal sub-edge crossed by p; let q_k be a corner different from q_i and q_j. Since the internal sub-edges belonging to q_i and q_j are on the same side of any internal sub-edge belonging to q_k, the internal sub-edges belonging to q_k must be crossed an even number of times by p, but also at most once by p since p is a shortest path. Hence p crosses only internal sub-edges belonging q_i and q_j. This proves the claim.

Each edge of $\mathcal{N}$ corresponds to at most two external sub-edges in D, and to at most one internal sub-edge belonging to q_i (for each i). Hence, by the claim in the previous paragraph, a shortest path in D has multiplicity at most four on $\mathcal{N}$. $\square$

6.2. Preprocessing Step. Our preprocessing step consists of the construction of a *triangulated system of arcs* on $\mathcal{M}$: a set of simple, pairwise disjoint arcs that cut the surface into disks, each disk being incident to exactly three (sides of) arcs. By Euler's formula and double-counting of the vertex-edge incidences, any triangulated system of arcs contains exactly $6g + 3b - 3$ arcs.

THEOREM 6.3. *In $O(n \log n + (g+b)n)$ time, we can compute a tight triangulated system of arcs on $\mathcal{M}$ such that each arc has multiplicity at most four.*

Proof. We fill the b boundaries of $\mathcal{M}$ with disks $D_1, \dots, D_b$, and assign weights to the edges of G^* incident to the disks D_i such that all the weights of $\mathcal{M}$ are infinitesimally small compared to them; this gives a new cross-metric $\mathcal{N}$ without boundary. Let $P = \{p_1, \dots, p_b\}$ be a set of points in $\mathcal{N}$, one inside each disk D_i. We now apply Lemma 6.1, obtaining, in $O(n \log n + (g+b)n)$ time, a greedy system of paths S with vertex set P. Let D be the polygonal schema associated with S; it has complexity $O((g+b)n)$.

Let $q_1, \dots, q_{4g+2b-2}$ be the corners of the polygonal schema D. To triangulate D, we first compute a shortest-path tree inside D with root q_1, in $O((g+b)n)$ time using the linear-time algorithm by Henzinger et al. [20]. Let S' be the set of shortest paths in D between q_1 and q_i, for $i = 3, \dots, 4g+2b-3$. Each face of the arrangement of S' on D is a triangle.

We now view the elements of S' as paths or loops in $\mathcal{N}$. Since the curves in $S \cup S'$ have endpoints in $\{p_1, \ldots, p_b\}$, and because the weights on the boundaries of the disks D_i are huge, none of these curves meets a disk D_i, except to leave the disk containing its starting point and to enter the disk containing its terminal point. We now remove the disks D_i and the pieces of the paths in $S \cup S'$ inside them: we re-obtain the cross-metric surface $\mathcal{M}$, with a triangulated set of arcs $\bar{S} \cup \bar{S}'$.

Each arc of $\bar{S}$ is tight, since each element of S is tight. Each arc of $\bar{S}'$ is tight, because it is a shortest arc inside the polygonal schema delimited by the tight arcs in $\bar{S}$ and by Proposition 2.4. Each arc in $\bar{S}$ has multiplicity at most two (Lemma 6.1). Each arc in $\bar{S}'$ has multiplicity at most four (Lemma 6.2). This concludes. $\square$

As for the case of the octagonal decomposition, we will actually apply Theorem 6.3 in the overlay $G^+(\mathcal{M})$ of $G(\mathcal{M})$ and $G^*(\mathcal{M})$. More precisely, this graph G^+ is defined as follows for surfaces with boundary. The vertices of G^+ are either vertices of G, vertices of G^*, or intersections between an edge e of G and its dual edge e^* in G^*. All the non-boundary edges of G and G^* are overlaid in G^+; therefore, the non-boundary edges of G, and some non-boundary edges of G^*, are partitioned into two subedges in G^+. Finally, boundary edges are created between consecutive vertices on the boundary of $\mathcal{M}$. Each face of G^+ is a quadrilateral.

The edges e^+ of G^+ are assigned crossing weights as follows. If e^+ is contained in a non-boundary edge of G^*, it has the same crossing weight as that dual edge. If it is contained in a non-boundary edge of G, its crossing weight is a fixed formal infinitesimal $\varepsilon' << \varepsilon$. If e^+ is a boundary edge, it has crossing weight ∞. Again, any curves that are tight with respect to G^+ are also tight with respect to G^*.

In $O(n \log n + (g + b)n)$ time, we obtain a triangulated system of arcs $\mathcal{S}$ of $\mathcal{M}$ where each of the $O(g + b)$ arcs is tight, each edge of G^* is crossed $O(1)$ times by each arc in $\mathcal{S}$, and each edge of G is crossed $O(1)$ times by each arc in $\mathcal{S}$. In particular, any walk in G of length k crosses the arcs in $\mathcal{S}$ a total of $O((g + b)k)$ times. The faces of the triangulated system of arcs have $O(n)$ complexity. These properties are thus exactly the same as for the octagonal decomposition, with $g + b$ instead of g.

6.3. Tightening Curves. We here prove the Main Theorem in the case where $\mathcal{M}$ is a surface with at least one boundary ($b \geq 1$). Our algorithms for tightening paths in this case essentially follow the technique of Hershberger and Snoeyink [21], only using a tight triangulated system of arcs in place of a triangulation.

We start with the triangulated system of arcs $\mathcal{S}$ described in the previous section. We call the faces of $\mathcal{S}$ *triangles* (although their boundary is made of three sides of arcs and three pieces of boundaries of $\mathcal{M}$). Let $\widetilde{\mathcal{S}}$ be the set of lifts of arcs in $\mathcal{S}$ in the universal cover $\widetilde{\mathcal{M}}$ of $\mathcal{M}$. A crucial property is that each such lift is separating (Lemma 2.1), as the lines of an octagonal decomposition.

We first explain how to tighten a path p; let $\tilde{p}$ be one of its lifts in $\widetilde{\mathcal{M}}$. By Lemma 2.3, and since each lift of an arc is separating, some shortest path $\tilde{p}'$ with the same endpoints as $\tilde{p}$ does not cross the lifts in $\widetilde{\mathcal{S}}$ crossed an even number of times by $\tilde{p}$. We compute the union of the triangles traversed by $\tilde{p}$, without building the internal structure of the triangles. The construction is classical, see Hershberger and Snoeyink [21]: start with a copy of the triangle containing $p(0)$; follow p until it exits the currently built union of triangles, and add to this region a copy of the triangle necessary to prolongate the lift $\tilde{p}$ of p; iterate. Then we consider the lifts in $\widetilde{\mathcal{S}}$ that are crossed an odd number of times by $\tilde{p}$, and compute the *relevant space*, that is, the portion of $\widetilde{\mathcal{M}}$ accessible from $\tilde{p}(0)$ by crossing only (a subset of) these lifts; we build the $O(n)$ internal structure of these triangles. Finally, we compute the shortest path, in the relevant space, between the endpoints of $\tilde{p}$, using the linear-time algorithm by Henzinger et al. [20]. Its projection is the desired output path. The analysis of the time complexity is exactly the same as in the case $g \geq 2$, $b = 0$, so we omit it.[6]

To tighten a cycle γ, we use ideas similar as those in §4.3, proceeding as in the previous paragraph to build the relevant space. More precisely, Lemma 4.4, Corollary 4.5, and Lemma 4.6 extend directly to the case of surfaces with boundary. We obtain a sequence $(a_i)_{i\in\mathbb{Z}}$ of lifts, in $\widetilde{\mathcal{M}}$, of an arc in $\widetilde{\mathcal{S}}$ such that a_{i-1} is on the left of a_i and a_{i+1} is on its right, and a sequence $(v_i)_{i\in\mathbb{Z}}$ of points such that $v_i \in a_i$, and we are looking for the shortest cycle homotopic to the cycle corresponding to the projection of any path from v_i to v_{i+1}.

We consider the relevant space of the part of the lift of γ between v_2 and v_3; it is made of $O(x)$ triangles. The shortest cycle homotopic to γ has a lift in this region that connects a point of a_2 to the corresponding point of a_3, and crosses only arcs that separate a_2 and a_3. We thus compute the space reachable from v_2 by crossing only these lifts, and then build the internal structure of the triangles. Finally, we compute the shortest generating cycle of the annulus $\widehat{A}'$ obtained from this space by identifying a_2 and a_3; its projection is the output of the algorithm. The complexity analysis is, again, identical to the one done in §4.3.

7. Better Analysis of Curve Shortening. The goal of this section is to prove that the algorithms of Colin de Verdière [6] and Colin de Verdière and Lazarus [7, 8] for tightening sets of simple, pairwise disjoint curves on combinatorial surfaces run in time polynomial in the complexity of their input: the number of vertices, edges, and faces of the surface, plus the total number of edges of the input curves. Previously, these algorithms were only known to work in time polynomial in the size of the input and in the ratio α between the largest and smallest length of an edge of the input surface.

These three algorithms all use the same high-level approach, which we rephrase now for convenience:

(1) Given a set s of simple, pairwise disjoint arcs [6] (resp. cycles [8]) on a combinatorial surface $\mathcal{M}$, the algorithm extends s to a so-called *cut system by paths* (resp. *doubled pants decomposition*): a set of simple, pairwise disjoint arcs (resp. cycles) $s' = (s'_1, \dots, s'_p)$ such that each component of $\mathcal{M}$ cut along all curves of s' except one has a simple topology: a disk (resp. a pair of pants).

(2) The second step is a succession of *elementary steps*, each of which tightens a curve s'_i in the surface $\mathcal{M}$ cut along $s' \setminus \{s'_i\}$ (this is easy because this surface is topologically simple). Elementary steps are applied to each curve s'_1, s'_2, $\dots$, s'_p in order, again and again, until stability in length is reached. It is guaranteed that, once this stability is reached, every arc (resp. cycle) is tight in $\mathcal{M}$.

(3) Finally, the curves corresponding to the ones added in Step (1) are removed.

The algorithm for tightening arcs also allows, given a graph embedding on the surface, to compute the shortest graph embedding isotopic, with fixed vertices, to it [6]: Simply puncture the surface $\mathcal{M}$ at each vertex of the graph; each edge of the graph becomes an arc on this new surface $\mathcal{M}'$; apply the algorithm above, thus tightening each arc; filling the punctures of $\mathcal{M}'$ to re-obtain $\mathcal{M}$ gives indeed the desired graph embedding. The (earlier) algorithm of Colin de Verdière and Lazarus [7] is a

[6] We could also build directly the relevant space, without the first step that ignores the internal surface structure of the triangles, by computing the *crossing word* of p with $\mathcal{S}$ and by reducing it [7]. This is simpler in practice but has asymptotically the same complexity.

slight variation of this algorithm, in the special case where the graph embedding is a system of loops.

Anyway, Step (1) is known to run in polynomial time in its input (and trivially also Step (3)). The bottleneck in complexity is Step (2), which was known to run in time polynomial in the complexity of its input and in α. We prove that the running time is polynomial, irrespectively of α:

THEOREM 7.1. *On a combinatorial (g, b)-surface with complexity n,*

(a) the algorithm of Colin de Verdière [6, Theorem 3.13] for shortening a cut system by paths of complexity k has running time $O((g+b)^4 nk^4)$;

(b) the algorithm of Colin de Verdière and Lazarus [8] for shortening a doubled pants decomposition of complexity k has running time $O((g+b)^4 nk^4 \log(nk))$.

As mentioned above, the algorithm to shorten a graph [6, Theorem 3.2] is essentially the same as the algorithm for shortening a cut system by paths [6, Theorem 3.13]. So (a) implies that this algorithm also runs in $O((g+b+h)^4 nk^4)$, where h is the number of vertices of the graph. Similarly, the algorithm to tighten a system of loops [7] runs in $O(g^4 nk^4)$ time.

Proof. We will use the following known facts, valid for both algorithms:

(1) The number of elementary steps is $O(g+b)$ times the maximum, over i, of the number of crossings between the input set of curves s' and any shortest curve t_i homotopic to some input curve s'_i.

(2) An elementary step can be performed in $O(n+k')$ time for case (a) and in $O((n+k')\log(n+k'))$ time for case (b), where k' is the complexity of the current set of curves.

Below, we prove the two following properties:

(3) The number of crossings between s' and t_i is $O((g+b)k^2)$.

(4) At each elementary step, the complexity of the set of curves increases by $O(n)$.

These two properties conclude. Indeed, (1) and (3) imply that the number of elementary steps is $O((g+b)^2 k^2)$. With (4), we obtain that the total complexity of the curves at any given stage of the algorithm is $O(k+(g+b)^2 k^2 n) = O((g+b)^2 k^2 n)$. Now, (2) implies that the time spent for an elementary step in case (a) is $O((g+b)^2 k^2 n)$; therefore the time complexity for case (a) is $O((g+b)^4 k^4 n)$. The same analysis (with an extra log factor) holds for case (b).

To prove (3), let $\mathcal{U}$ be a tight octagonal decomposition of $\mathcal{M}$ (if $b=0$ and $g \geq 2$), a quadrilateral decomposition of $\mathcal{M}$ (if $b=0$ and $g=1$), or a tight triangulated system of arcs of $\mathcal{M}$ (if $b \neq 0$), whose existence follows from Theorem 3.1, §5, and Theorem 6.3, respectively. Below, by a *polygon*, we mean an octagon, a quadrilateral, or a triangle, respectively. The existence of these decompositions where each curve has multiplicity $O(1)$ in the refined graph $G^+(\mathcal{M})$ implies that a curve s'_i has at most $O((g+b)k)$ crossings with $\mathcal{U}$. So the shortest curve t_i homotopic to s'_i has at most $O((g+b)k)$ crossings with $\mathcal{U}$. Each subpath of t_i inside a polygon is a shortest path within this polygon. Since the boundary of a polygon is made of at most 8 different curves, each of which having multiplicity $O(1)$ on $\mathcal{M}$, this implies that each subpath of t_i inside a polygon passes $O(1)$ times through each face of G^*, and is therefore crossed $O(k)$ times by s'. There are $O((g+b)k)$ such subpaths of t_i. Thus s' and t_i cross $O((g+b)k^2)$ times, which was to be proved.

We now prove (4) in case (a). At each elementary step, an arc s'_i is replaced by a shortest arc s''_i in $\mathcal{M}$ cut along $s' \setminus \{s'_i\}$. We consider all the edges of G^*, split into sub-edges by the crossings with the arcs in $s' \setminus \{s'_i\}$. Call the sub-edges that contain one or two endpoints of an edge of G^* *external*, and the other ones *internal*. The arc s''_i crosses each sub-edge at most once. There are $O(n)$ external sub-edges, so the number of crossings of s''_i with the external sub-edges is $O(n)$. Assume that s''_i crosses a given internal sub-edge e; it does so exactly once. The endpoints of e are on the boundary of the disk obtained by cutting $\mathcal{M}$ along $s' \setminus \{s'_i\}$. Since s''_i crosses e exactly once and has the same endpoints as s'_i inside the disk, this implies that s'_i also crosses e. We can thus charge this crossing of s''_i with e by the crossing of s'_i with e (because s'_i is removed). Thus, the complexity of s''_i equals the complexity of s'_i plus $O(n)$.

To prove (4) in case (b), recall that each elementary step consists of replacing, in a pair of pants P, a cycle homotopic to one boundary of P by a shortest homotopic cycle in P. As above, we consider the sub-edges obtained by splitting the edges of G^* at the crossing points with the cycles in $s' \setminus \{s'_i\}$. Since the shortest cycle homotopic to a given boundary of a pair of pants has multiplicity at most two (Proposition 2.7(f)), and since there are $O(n)$ external sub-edges, the number of crossings of s''_i with the external sub-edges is $O(n)$. Finally, note that each lift of each internal sub-edge in P separates the universal cover of P (Lemma 2.1); this implies that the number of crossings of s''_i with an internal sub-edge is no more than the number of crossings of s'_i with this internal sub-edge. $\square$

8. Conclusion. To conclude, we briefly discuss the case of non-orientable surfaces. Our tightening algorithm requires orientability at several places. The results of Hass and Scott [18] (stated in Proposition 2.2) were proven in the orientable case only. Also, the annular cover is defined for two-sided curves only (for one-sided curves, it would be a Möbius band); in particular, Proposition 2.5 holds for orientable surfaces only. Finally, the construction of the octagonal decomposition does not work directly for non-orientable surfaces (in particular, one-sided cycles should be handled separately).

For the problem of path tightening, we can circumvent this problem easily. Let $\mathcal{M}$ be a non-orientable surface, possibly with boundary. Let $\mathcal{M}^2$ be the *orientable double cover* of $\mathcal{M}$: it is an orientable surface that is a covering space of $\mathcal{M}$, such that each point of $\mathcal{M}$ lifts to exactly two points in $\mathcal{M}^2$. Given a path p on $\mathcal{M}$, computing a shortest path homotopic to p amounts to lifting p to $\mathcal{M}^2$, computing in $\mathcal{M}^2$ the shortest path homotopic to that lift, and projecting it back to $\mathcal{M}$. If $\mathcal{M}$ is a non-orientable surface of genus g with b boundaries, then $\mathcal{M}^2$ is an orientable surface of genus $g-1$ (by Euler's formula) and $2b$ boundaries. Hence we obtain an algorithm with the following complexity in the non-orientable case:

	preprocessing step	path tightening
$g \geq 3$, $b = 0$	$O(gn \log n)$	$O(g(k + n\bar{k}))$
$g = 2$, $b = 0$	$O(n \log n)$	$O(k + n\bar{k}^2)$
$g = 1$, $b = 0$	0	$O(k + n)$
$b \geq 1$	$O(n \log n + (g+b)n)$	$O((g+b)(k + n\bar{k}))$

(In the case $g = 1$, $b = 0$, the surface $\mathcal{M}^2$ is a sphere, and we only have to compute a shortest path joining the endpoints of the lift of the path.)

However, the problem of tightening cycles on a non-orientable surface is much more complicated. The difficulty comes from the fact that the lift of a cycle to $\mathcal{M}^2$ is not necessarily a cycle. (One may hope to use the fact that the lift of the square of that cycle is a cycle, but this is not helpful, since that lift may be contractible even if the original cycle is not; consider for example a non-contractible simple cycle in the projective plane.) We leave this as an open problem.

Acknowledgments. The authors would like to thank Kim Whittlesey for helpful comments on topology and combinatorial group theory, Francis Lazarus for several discussions (notably on the use of the annular cover and on a part of the proof of Theorem 7.1) that arose when writing previous papers [7, 8], and the referees for carefully reading the document and suggesting helpful modifications. Jeff would like to thank Michel Pocchiola for his invitation to visit ENS, where this work was initiated.

REFERENCES

[1] S. Bespamyatnikh, *Computing homotopic shortest paths in the plane*, Journal of Algorithms, 49 (2003), pp. 284–303.

[2] S. Cabello and E. W. Chambers, *Multiple source shortest paths in a genus g graph*, in Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2007, pp. 89–97.

[3] S. Cabello, M. DeVos, J. Erickson, and B. Mohar, *Finding one tight cycle*, ACM Transactions on Algorithms, 6 (2010), Article No. 61.

[4] S. Cabello, Y. Liu, A. Mantler, and J. Snoeyink, *Testing homotopy for paths in the plane*, Discrete & Computational Geometry, 31 (2004), pp. 61–81.

[5] S. Cabello and B. Mohar, *Finding shortest non-separating and non-contractible cycles for topologically embedded graphs*, Discrete & Computational Geometry, 37 (2007), pp. 213–235.

[6] É. Colin de Verdière, *Raccourcissement de courbes et décomposition de surfaces [Shortening of Curves and Decomposition of Surfaces]*, PhD thesis, University of Paris 7, Dec. 2003. Available at http://www.di.ens.fr/~colin/textes/these.html.

[7] É. Colin de Verdière and F. Lazarus, *Optimal system of loops on an orientable surface*, Discrete & Computational Geometry, 33 (2005), pp. 507–534.

[8] É. Colin de Verdière and F. Lazarus, *Optimal pants decompositions and shortest homotopic cycles on an orientable surface*, Journal of the ACM, 54 (2007), Article No. 18.

[9] M. Dehn, *Transformation der Kurven auf zweiseitigen Flächen*, Mathematische Annalen, 72 (1912), pp. 413–421.

[10] T. K. Dey and S. Guha, *Transforming curves on surfaces*, Journal of Computer and System Sciences, 58 (1999), pp. 297–325.

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

[12] A. Efrat, S. G. Kobourov, and A. Lubiw, *Computing homotopic shortest paths efficiently*, Computational Geometry: Theory and Applications, 35 (2006), pp. 162–172.

[13] D. Eppstein, *Dynamic generators of topologically embedded graphs*, in Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2003, pp. 599–608.

[14] D. B. A. Epstein, *Curves on 2-manifolds and isotopies*, Acta Mathematica, 115 (1966), pp. 83–107.

[15] J. Erickson and S. Har-Peled, *Optimally cutting a surface into a disk*, Discrete & Computational Geometry, 31 (2004), pp. 37–59.

[16] J. Erickson and K. Whittlesey, *Greedy optimal homotopy and homology generators*, in Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2005, pp. 1038–1046.

[17] G. N. Frederickson, *Fast algorithms for shortest paths in planar graphs, with applications*, SIAM Journal on Computing, 16 (1987), pp. 1004–1022.

[18] J. Hass and P. Scott, *Intersections of curves on surfaces*, Israel Journal of Mathematics, 51 (1985), pp. 90–120.

[19] A. Hatcher, *Algebraic topology*, Cambridge University Press, 2002. Available at http://www.math.cornell.edu/~hatcher/.

[20] M. Henzinger, Ph. Klein, S. Rao, and S. Subramanian, *Faster shortest-path algorithms for planar graphs*, Journal of Computer and System Sciences, 55 (1997), pp. 3–23.

[21] J. Hershberger and J. Snoeyink, *Computing minimum length paths of a given homotopy class*, Computational Geometry: Theory and Applications, 4 (1994), pp. 63–98.

[22] Ph. N. Klein, *Multiple-source shortest paths in planar graphs*, in Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2005, pp. 146–155.

[23] M. Kutz, *Computing shortest non-trivial cycles on orientable surfaces of bounded genus in almost linear time*, in Proceedings of the 22nd Annual ACM Symposium on Computational

Geometry (SOCG), 2006, pp. 430–438.

[24] F. Lazarus, M. Pocchiola, G. Vegter, and A. Verroust, *Computing a canonical polygonal schema of an orientable triangulated surface*, in Proceedings of the 17th Annual ACM Symposium on Computational Geometry (SOCG), 2001, pp. 80–89.

[25] J. H. Reif, *Minimum s-t cut of a planar undirected network in $O(n \log^2(n))$ time*, SIAM Journal on Computing, 12 (1983), pp. 71–81.

[26] H. Schipper, *Determining contractibility of curves*, in Proceedings of the 8th Annual ACM Symposium on Computational Geometry (SOCG), 1992, pp. 358–367.

[27] J. Stillwell, *Classical Topology and Combinatorial Group Theory*, Springer-Verlag, New York, 1993.

[28] G. Vegter and C.-K. Yap, *Computational complexity of combinatorial surfaces*, in Proceedings of the 6th Annual ACM Symposium on Computational Geometry (SOCG), 1990, pp. 102–111.